

Notas de Clase Estadística III 3009137

Capítulo 3: El análisis de series de tiempo

Nelfi González Alvarez
Profesora Asociada

Departamento de Estadística
Universidad Nacional de Colombia, Sede Medellín



UNIVERSIDAD NACIONAL DE COLOMBIA

2025

Índice general

3. El análisis de series de tiempo	58
3.1. Introducción	58
3.2. Clases de pronósticos - pronósticos cuantitativos	59
3.3. Ventajas y desventajas de los métodos cuantitativos	59
3.4. La precisión de un pronóstico vs. el horizonte de tiempo	60
3.5. Las tareas en la realización de un pronóstico	60
3.6. La metodología de los datos de series de tiempo	60
3.6.1. Datos históricos	60
3.6.2. Periodo muestral para análisis	61
3.6.3. Periodo de estimación	62
3.6.4. Periodos de pronóstico	62
3.7. Estudio descriptivo de las series de tiempo	63
3.7.1. Componente de tendencia (T_t)	63
3.7.2. Componente estacional (S_t)	63
3.7.3. Componente cíclica (C_t)	63
3.7.4. Componente horizontal, irregular o error (E_t)	64
3.8. Modelos de descomposición	64
3.9. Guía para el análisis de las componentes de una serie	70
3.9.1. Determinación del tipo de modelo a considerar entre modelos de componentes aditivas vs. multiplicativas	70
3.9.2. Determinación de la presencia de la componente estacional periódica exacta o casi exacta	72
3.9.3. Evaluación de la tendencia	72
3.10. Ejemplo de análisis de las componentes de una serie	75
A. Función decompose: Descomposición estacional clásica mediante medias móviles	79
Bibliografía	83

Índice de figuras

3.1.	Línea de tiempo vs. ajuste y pronóstico de una serie de tiempo. Fuente: Gaynor and kirkpatrick, 1994.	61
3.2.	Ejemplo de la definición de la línea de tiempo: Muestra de ajuste y de validación cruzada con pronósticos Ex-post.	62
3.3.	Ejemplos patrones de tendencia determinísticas	63
3.4.	Ejemplo patrón estacional periodico exacto	63
3.5.	Ejemplo de fluctuaciones cíclicas	64
3.6.	Ejemplo patrón componente irregular	64
3.7.	(a) Ejemplo serie estacional y aditiva: Variabilidad alrededor de su tendencia es constante; (b) Ejemplo serie estacional multiplicativa: La variabilidad alrededor de su tendencia cambia en la dirección de ésta.	65
3.8.	(a) Observaciones entre enero de 1990 a mayo de 2011; (b) Descomposición aditiva con función <code>R decompose</code> . Esta serie se considera aditiva pues su variabilidad alrededor de su tendencia no cambia en el tiempo. Observe además que tiene componente estacional.	66
3.9.	Estimaciones de las componentes serie de la Figura 3.8(a), con filtro de la descomposición aditiva: (a) Componente de tendencia (realmente queda tendencia más ciclos); (b) Componente estacional; (c) Componente de error (realmente es un residuo)	67
3.10.	Análisis de la estimación de tendencia con filtro de descomposición de la serie en la Figura 3.8(a): (a) Tendencia y ciclos; (b) Posible tendencia pura; (c) Ciclos obtenidos como la diferencia entre (a) y (b).	67
3.11.	(a) Observaciones entre enero de 1967 - diciembre de 1994 (miles de dólares), Fuente: Diebold, 2001; (b) Descomposición multiplicativa con función <code>R decompose</code> . Esta serie se considera multiplicativa pues su variabilidad alrededor de su tendencia cambia en el tiempo en la dirección de la tendencia. Observe que tiene componente estacional.	68
3.12.	(a) Observaciones entre enero de 1967 - diciembre de 1994 (log. miles de dólares); (b) Descomposición aditiva con función <code>R decompose</code> . El logaritmo de los datos se considera una serie aditiva, pues su variabilidad alrededor de su tendencia es constante en el tiempo. Observe que el patrón estacional se conserva con la transformación.	68
3.13.	Estimaciones de las componentes de la serie en la Figura 3.12(a), con filtro de la descomposición aditiva: (a) Componente de tendencia (realmente queda tendencia más ciclos); (b) Componente estacional; (c) Componente de error (realmente es un residuo)	69
3.14.	Análisis de la estimación de tendencia con filtro de descomposición de la serie en la Figura 3.12(a): (a) Tendencia y ciclos; (b) Posible tendencia pura; (c) Ciclos obtenidos como la diferencia entre (a) y (b).	69
3.15.	(a) La varianza de la serie alrededor de su tendencia es constante (componentes aditivas); (b) La varianza de la serie alrededor de su tendencia no es constante y cambia en la dirección de la tendencia (componentes multiplicativas)	70
3.16.	Caso aditivo: (a) Serie en escala original; (b) Serie en escala logaritmo natural.	71
3.17.	Caso multiplicativo: (a) Serie en escala original; (b) Serie en escala logaritmo natural.	71
3.18.	Caso multiplicativo: (a) \hat{E}_t en la descomposición aditiva de la serie en la Figura 3.17; (b) \hat{E}_t en la descomposición multiplicativa de la serie en la Figura 3.17. El patrón de varianza de \hat{E}_t en la descomposición multiplicativa es más estable que el de \hat{E}_t en la descomposición aditiva.	72
3.19.	Caso aditivo: (a) \hat{E}_t en la descomposición aditiva de la serie en la Figura 3.16; (b) \hat{E}_t en la descomposición multiplicativa de la serie en la Figura 3.16. El patrón de varianza de \hat{E}_t en la descomposición multiplicativa no es más estable que el de \hat{E}_t en la descomposición aditiva.	72
3.20.	(a) Gráfica para identificar que existe un patrón que se repite cada año, donde las líneas verticales que separan los años revelan un patrón repetitivo con forma de V; (b) Las líneas centrales de cada caja representan la mediana, y vemos que su posición varía de acuerdo al mes del año	73
3.21.	Caso aditivo: (a) Serie aditiva en su escala original; (b) Componente de tendencia según la descomposición aditiva de la serie usando la función <code>R decompose()</code> . Los valores máximo y mínimo del eje vertical en esta última gráfica deben coincidir con los de la serie.	73
3.22.	Caso multiplicativo: (a) Serie multiplicativa en su escala original; (b) Logaritmo natural de la serie; (c) Tendencia del logaritmo natural según su descomposición aditiva con la función <code>R decompose()</code> . Los valores máximo y mínimo del eje vertical en esta última gráfica deben coincidir con los del logaritmo natural de la serie.	74

3.23. (a) La serie mensual de desempleo en Francia, promedio mensual de mujeres (1000 personas), enero-1983 a mayo-2012, fuente: Eurostat, datamarket.com, es de componentes aditivas (la varianza es constante); (b) Componente de tendencia de la descomposición aditiva de la serie obtenida con la función R <code>decompose()</code> . En esta gráfica es claro que la trayectoria de largo plazo es fuertemente afectada por variaciones cíclicas, por tanto un modelo de regresión de la forma dada en (3.1) será difícil de formular con un buen ajuste y con una calidad predictiva aceptable.	75
3.24. Vista de la base de datos: “Mean monthly air temperature (Deg. F) Nottingham Castle.csv”	75
3.25. (a) Serie en “Mean monthly air temperature (Deg. F) Nottingham Castle.csv”; (b) Descomposición aditiva de la serie con la función R <code>decompose()</code>	76
3.26. (a) Gráfica de la tendencia estimada por la función R <code>decompose()</code> , sin ajustar rango del eje vertical; (b) Gráfica de la tendencia estimada por la función R <code>decompose()</code> , ajustando el rango de valores del eje vertical al de la serie completa leída desde el archivo “Mean monthly air temperature (Deg. F) Nottingham Castle.csv”.	76
3.27. (a) Boxplots de la distribución según mes del año, de la serie en el archivo “Mean monthly air temperature (Deg. F) Nottingham Castle”. (b) Gráfica de los primeros cinco años, con líneas de referencia vertical indicando los meses de diciembre.	77
A.1. Descomposición de una serie multiplicativa. Fuente:O.D. Anderson (1976) and O'Donovan (1983). Ventas mensuales compañía X, ene 65 - may 71.	80

Capítulo 3

El análisis de series de tiempo

3.1. Introducción

“El análisis de series de tiempo consiste en el desarrollo de modelos estadísticos para explicar el comportamiento de una variable aleatoria que varía en el tiempo, con el fin de construir pronósticos de los valores futuros de dicha variable con fines de planeación y toma de decisiones” (Makridakis et. al., 1983).

El pronóstico realizado se basa en la estructura o comportamiento pasado de la serie, suponiendo que éste se conserva en el futuro (**supuesto de continuidad**). Los métodos de pronóstico son una rama particular de la estadística aplicada y por lo tanto reúnen sus propios conceptos y métodos específicos. Los pronósticos son además una de las principales herramientas de tipo cuantitativo aplicadas en los negocios en sus procesos de planeación y toma de decisiones, y también han sido utilizados en las ciencias sociales, ambientales y en el área militar.

Pronosticar puede definirse como “el esfuerzo de prever eventos futuros examinando el pasado” (Gaynor and Kirkpatrick, 1994), con el fin de controlar el riesgo y la incertidumbre. Sin embargo, las técnicas de pronóstico pueden estar basadas también en la experiencia o la opinión de “expertos”, o en modelos matemáticos (simulaciones dinámicas) que describen los patrones pasados. La mayoría de los pronósticos se basan en el supuesto de que la variable o evento a pronosticar es afectada o por el tiempo o por un conjunto de variables con las cuales puede tener ya sea una relación de causa efecto o bien, una relación de tendencia.

El estudio de las series de tiempo además de permitir formular modelos y pronósticos, también permite evaluar los efectos de acciones o decisiones pasadas, conocidas como intervenciones hechas sobre la serie. Estas intervenciones pueden tener un efecto momentáneo o bien un efecto permanente que modifica la estructura de la serie de tiempo intervenida.

Es importante reconocer el papel del pronóstico en la organización al igual que los métodos de pronósticos a aplicar. Una organización establece metas y objetivos, trata de predecir los factores ambientales de su negocio, y selecciona las acciones que espera le lleven a sus metas.

Las áreas empresariales en las cuales los pronósticos juegan un papel importante son (Makridakis et. al., 1983):

- **Programación de los recursos existentes (Pronósticos de corto plazo):** Para el uso eficiente de los recursos se requiere programar la producción, el transporte, el efectivo, el personal, etc. Para ello el pronóstico del nivel de la demanda de los productos, materiales, mano de obra, la financiación requerida, etc., son esenciales para la buena programación de los recursos existentes.
- **Adquisición de recursos adicionales (pronósticos de mediano plazo):** El lead time para la adquisición de materia prima, la contratación de personal, la compra de maquinaria y equipo, etc. puede variar desde días a varios años. Se requiere pues el pronóstico para establecer los requerimientos futuros de recursos.
- **Determinación de qué recursos son deseados (pronósticos de largo plazo):** Todas las empresas u organizaciones requieren determinar los recursos que desean o necesitan a largo plazo. Decisiones como éstas dependen de las oportunidades de mercado, de los factores ambientales y de sus recursos internos.

*Las empresas deben desarrollar múltiples aproximaciones para **pronosticar** eventos inciertos y construir un sistema para pronosticar. Esto a su vez exige conocimiento y habilidad en al menos las siguientes áreas:*

- Identificación y definición de problemas a pronosticar
- Aplicación de una variedad de métodos de pronóstico
- Procedimientos para seleccionar los métodos apropiados para una situación específica
- Un soporte organizacional para aplicar y usar métodos de pronóstico formalizados

3.2. Clases de pronósticos - pronósticos cuantitativos

Básicamente, la gran clasificación de los métodos de pronóstico es según si es *cualitativo* (cuando prácticamente no se cuentan con datos históricos) - *tecnológico o cuantitativo* (se disponen de datos históricos). En este curso sólo se abordarán algunos métodos cuantitativos. *Los métodos cuantitativos* pueden aplicarse cuando se dan las siguientes tres condiciones (Makridakis et. al., 1983):

1. Se dispone de información histórica.
2. La información disponible puede cuantificarse en forma de datos numéricos.
3. Puede asumirse que algunos aspectos de los patrones dados en el pasado continuarán en el futuro (**supuesto de continuidad**).

Los métodos cuantitativos varían considerablemente y han sido desarrollados por diversas disciplinas para diferentes propósitos, sus costos y precisiones varían y deben ser consideradas al momento de seleccionarlos.

El pronóstico formal implica objetividad y emplea un método estadístico para obtener pronósticos usando una base regular con fechas específicas de inicio y fin (Campos-Santillan, 2001). Según Makridakis et. al., 1983, involucran también la extrapolación pero de una manera sistemática buscando minimizar el error de pronóstico. Los métodos cuantitativos de pronóstico se dividen en (Makridakis et. al., 1983),

- *Los métodos de series de tiempo* (suavizamiento, descomposición, métodos Box - Jenkins).
- *Los métodos causales o econométricos* (regresión y correlación).

Los métodos de series de tiempo están basados en el análisis de una sucesión cronológica de observaciones de la variable de interés y bajo el **supuesto de continuidad** antes mencionado, realizan la predicción del futuro basados en datos históricos. Por tanto, el objetivo de estos métodos es descubrir el patrón subyacente en la serie de datos históricos y extrapolar ese patrón al futuro. Los datos pueden ser anuales, trimestrales, mensuales, semanales, diarios, horarios, etc., la variable de interés puede ser de naturaleza macroeconómica, tal como una medida de agregación de la economía o la industria, o de naturaleza micro económica como aquellas medidas para una empresa.

Los modelos causales por su parte asumen que el factor o variable a ser pronosticada (variable dependiente) mantiene una relación de causa -efecto con una o más variables independientes o predictoras, por tanto, el comportamiento de estas últimas permite pronosticar el comportamiento de la variable dependiente. Por ejemplo, las ventas pueden ser explicadas por los gastos en publicidad, el precio de ventas, el ingreso disponible de los consumidores, el nivel de competencia en el mercado, etc. Por tanto, *el modelo causal tiene como propósito descubrir la forma (función matemática) de esa relación y usarla para pronosticar los valores futuros de la variable dependiente*.

Puede verse entonces que la diferencia fundamental entre los métodos de series de tiempo y los causales consiste en que *los primeros tratan la información como una caja negra sin intentar descubrir los factores que están afectando la variable que se analiza de modo que el sistema es un proceso no identificado*.

3.3. Ventajas y desventajas de los métodos cuantitativos de pronóstico

En general los métodos cuantitativos exhiben las siguientes ventajas:

1. Los pronósticos están basados en valores predeterminados y pueden ser por tanto objetivos.
2. Es posible medir la exactitud de los pronósticos.
3. Después de construir los modelos, se puede ahorrar tiempo en la generación de los pronósticos y utilizar repetitivamente dichos modelos.
4. Se pueden obtener pronósticos puntuales (un valor específico para cada periodo) o por intervalos (rango de valores basados en un intervalo con un nivel de confianza).

Pero los métodos cuantitativos exhiben también desventajas, tales como:

1. Sólo son apropiados para pronósticos de corto y mediano plazo con una precisión razonable.
2. No hay un método para contabilizar las influencias externas o los cambios que pudieran afectar los resultados. Un pronóstico cuantitativo puede ser muy bueno si las condiciones que han existido hasta el momento se mantienen; sin embargo, un cambio inducido por fuerzas externas en cierto punto del horizonte de pronóstico puede hacer bastante malo al pronóstico.

3.4. La precisión de un pronóstico vs. el horizonte de tiempo

Un aspecto fundamental a considerar en los pronósticos es la precisión, la cual es una función inversa del horizonte de tiempo (Campos-Santillan, 2001): mientras más cercano en el tiempo está el futuro a pronosticar, mayor exactitud del pronóstico. Los pronósticos de largo plazo suelen ser más imprecisos, debido a variaciones inflacionarias y factores macroeconómicos imprevistos que introducen incertidumbre e inexactitud. Los métodos cuantitativos deben ser acompañados de los cualitativos para un mejor pronóstico del futuro lejano.

En los pronósticos de mediano plazo los factores macroeconómicos mayores tienden a ser menos influyentes pero su impredecibilidad hace que deban ser tenidos en cuenta. A mediano plazo los factores que más causan inexactitud son las variaciones comerciales específicas del producto (la moda) y la estacionalidad, que obligan a revisar periódicamente los pronósticos dentro del ciclo de mediano plazo.

El corto plazo es afectado positivamente por la inercia de muchas de las actividades económicas en las cuales los modelos de series de tiempo pueden generar pronósticos más o menos exactos.

3.5. Las tareas en la realización de un pronóstico

De acuerdo a Gaynor and Kirkpatrick, 1994, se deben establecer las siguientes prioridades en la construcción de un modelo de pronóstico:

- Tarea 1: Determinar el objetivo primario del proyecto de pronóstico y especificar claramente la variable o variables a ser pronosticadas, la periodicidad del pronóstico (semanales, mensuales, trimestrales, anuales, etc.), la longitud de los pronósticos (inmediatos, corto, mediano o largo plazo) y si relaciones causales son importantes en el pronóstico.
- Tarea 2: Determinar la disponibilidad de datos chequeando todas las fuentes primarias apropiadas de datos, tanto públicas como privadas.
- Tarea 3: Recoger los datos, y asegurarse que la tabulación de estos es comparable y consistente sobre el tiempo.
- Tarea 4: Establecer la línea de tiempo mediante consideración cuidadosa del tipo de modelo de pronóstico, incluyendo el número de observaciones requeridas para la estimación del modelo y el número de períodos de tiempo en el periodo de pronóstico ex ante (periodo futuro del cual no se tienen observaciones)
- Tarea 5: Graficar los datos y examinar sus patrones, para hacer una determinación final de cuál tipo de modelo es apropiado considerar.
- Tarea 6: Estimar los modelos seleccionados.
- Tarea 7: Evaluar el modelo sobre el periodo de estimación y seleccionar el mejor modelo comparando tanto las medidas resumen de pronóstico y los gráficos de los valores actuales versus los pronosticados y los errores. Examinar los gráficos para la habilidad del modelo para capturar los puntos de cambio.
- Tarea 8: Evaluar el modelo usando el pronóstico ex post (Periodo de tiempo de observaciones conocidas después del fin del periodo muestral usado en el ajuste). Si es posible, usar tanto medidas de resumen de pronóstico y gráficos.
- Tarea 9: Generar el pronóstico ex ante con límites de confianza apropiados y preparar la presentación del pronóstico.
- Tarea 10: Después de establecer el modelo, hacer seguimiento periódico al modelo actualizándolo y revaluándolo.

3.6. La metodología de los datos de series de tiempo y pronósticos

Siguiendo a Gaynor and Kirkpatrick, 1994, en la construcción de un modelo de pronóstico, por un método cuantitativo, se toma una muestra de observaciones del conjunto de datos disponibles. El método de recolección de los datos y los pronósticos dependen de los factores representados en la Figura 3.1.

3.6.1. Datos históricos

Corresponden desde el valor del periodo de tiempo de inicio Y_{inicio} hasta el valor del periodo de tiempo final Y_{final} . El tiempo de inicio es el periodo en el cual la recolección de datos sobre la variable comienza y el tiempo final, corresponde al tiempo de la observación más reciente de esa variable. Este último periodo debe ser considerado como el “hoy”.

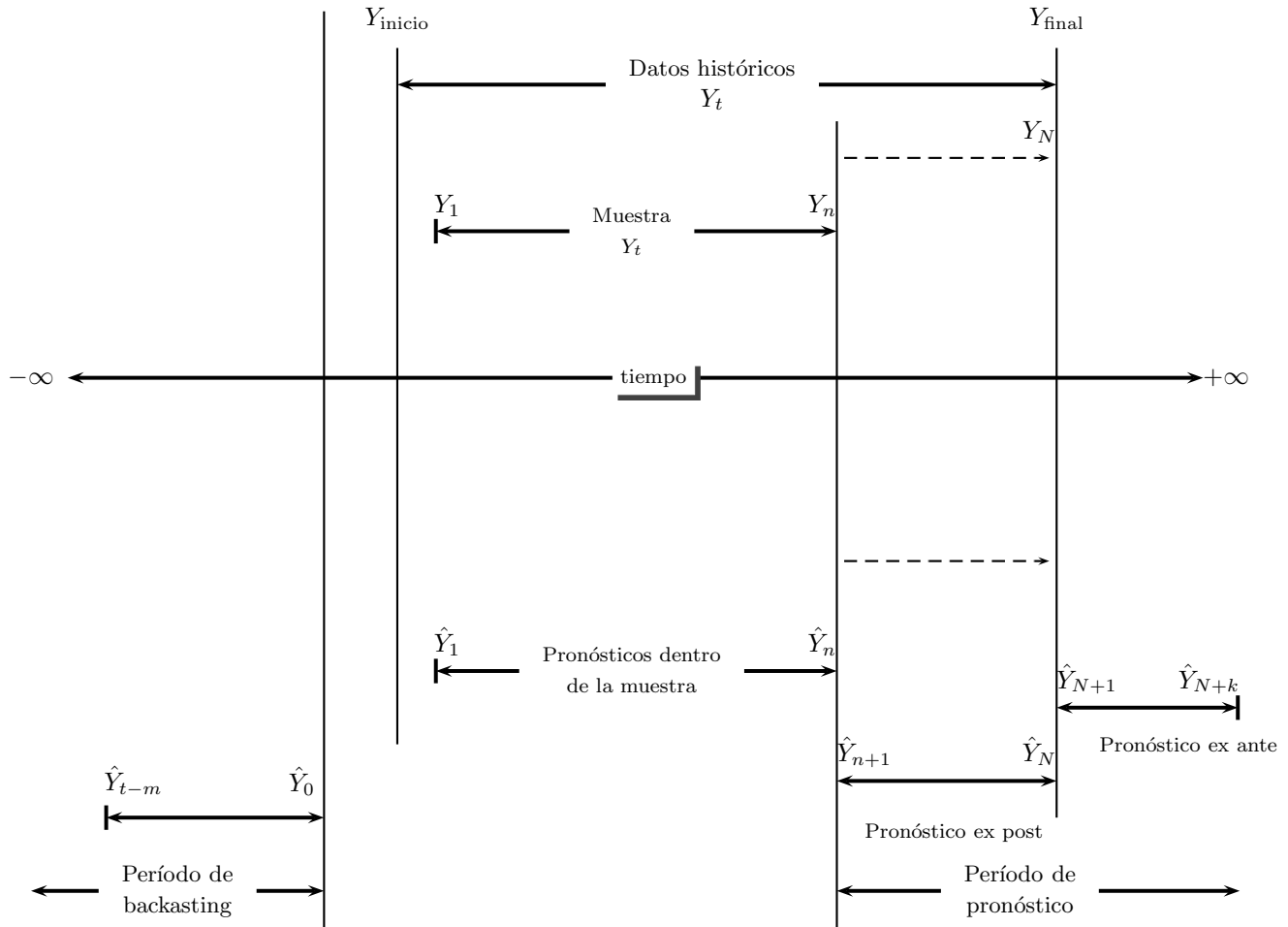


Figura 3.1: Línea de tiempo vs. ajuste y pronóstico de una serie de tiempo. Fuente: Gaynor and kirkpatrick, 1994.

3.6.2. Período muestral para análisis

Cubre los valores Y_1, \dots, Y_n , es el periodo sobre el cual será construido y estimado el modelo de pronóstico. Y_{inicio} y Y_1 no necesariamente corresponden. En la Figura 3.2 se ilustra un caso en el cual la muestra es un subconjunto de los datos históricos. La razón por la cual Y_{inicio} y Y_1 pueden diferir es que en algún punto del tiempo pueden haber ocurrido cambios estructurales o en las relaciones subyacentes que modifican el patrón de los datos. Los modelos de series son construidos bajo el supuesto de continuidad, en tanto que los modelos causales son generados con base en presunciones sobre relaciones de estructura y comportamiento permanentes sobre un periodo de tiempo, entonces puede ser que el número de periodos de tiempo desde un cambio identificado sobre la estructura o comportamiento, no es suficiente para la estimación de la serie y/o del modelo causal. En ese caso debe tenerse cuidado, y mantenerse en mente los cambios ocurridos. El periodo muestral concluye con la observación Y_n , pero este periodo podría extenderse hasta la observación Y_{final} , es decir $Y_n = Y_{\text{final}}$. Aunque podría procederse de tal manera cuando hay muchas observaciones entre Y_n y Y_{final} , en la práctica no se procede así, primero porque muchas de las variables económicas son entregadas como estimaciones sujetas a una posterior revisión, y estos datos revisados pueden ser muy diferentes de los estimados preliminarmente, por esto no se incluyen las estimaciones en la muestra para el ajuste del modelo de pronóstico. En segundo lugar, el proceso de reestimar y probar continuamente un modelo de pronóstico, consume tiempo y costos. Por tanto la exactitud de los modelos es monitoreada constantemente y sólo después que ciertos estándares de calidad

subjetivos dejen de ser satisfechos, entonces el modelo es reestimado y probado.

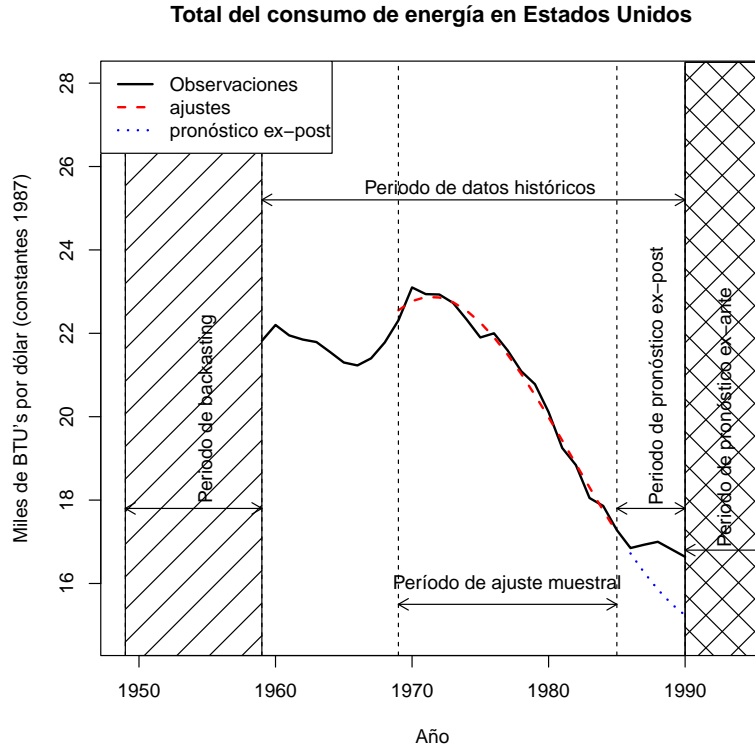


Figura 3.2: Ejemplo de la definición de la línea de tiempo: Muestra de ajuste y de validación cruzada con pronósticos Ex-post.

3.6.3. Periodo de estimación

En este período se generan los valores de pronóstico $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_n$, los cuales conforman el pronóstico del modelo dentro de la muestra. A partir de los valores actuales y pronosticados para Y_t , es posible calcular los errores de ajuste e_1, \dots, e_n del modelo, y por tanto la exactitud de éste (dentro del periodo de estimación). Todos los valores más allá de la observación Y_n , deben ser pronosticados, lo que en la línea de tiempo se denomina periodo de pronóstico. Todos los pronósticos generados dentro de este periodo se conocen como pronósticos fuera de la muestra.

3.6.4. Periodos de pronóstico

El periodo de pronóstico es dividido en los siguientes periodos:

- **Periodo de pronóstico ex post:** Periodo que va desde la primera observación después del final del periodo muestral hasta la observación más reciente. Para las predicciones después del ajuste, definimos la función de predicción puntual para L periodos después de $t = n$, denotada por $\hat{Y}_n(L)$, la cual es simplemente obtenida evaluando el modelo ajustado en el tiempo $t = n + L$, de modo que las predicciones ex-post son: $\hat{Y}_n(1) = \hat{Y}_{n+1}$, $\hat{Y}_n(2) = \hat{Y}_{n+2}$, \dots , $\hat{Y}_n(m) = \hat{Y}_N$, con $m = N - n$. Este periodo se caracteriza porque se dispone de los valores observados de la variable de la serie de tiempo para estos periodos, Y_{n+1}, \dots, Y_N . Por tanto, es posible determinar la exactitud del modelo fuera de la muestra de ajuste usando los valores observados y pronosticados de Y_t dentro del periodo de pronósticos ex post.

Si la exactitud del modelo fuera de la muestra de ajuste es cuestionable, se tienen dos caminos: buscar un modelo alternativo con mayor exactitud o reestimar el modelo actual extendiendo el periodo de estimación hasta el periodo ex post. En este último caso se generarían las estimaciones de la serie $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_N$ y no se tendría un pronóstico ex post.

- **Periodo de pronóstico ex ante:** Periodo de tiempo en el cual no se tienen observaciones de la serie de tiempo de la variable bajo estudio o de cualquier otra variable importante. En ese caso se pronostica dentro del futuro. Los pronósticos de este periodo son denotados por $\hat{Y}_N(L)$, siendo L el número de periodos después de $t = N$ hasta la fecha de predicción, luego, corresponden a $\hat{Y}_N(1) = \hat{Y}_{N+1}$, $\hat{Y}_N(2) = \hat{Y}_{N+2}$, \dots , $\hat{Y}_N(k) = \hat{Y}_{N+k}$, obtenidos con el modelo ajustado usando todos los N datos, y para las cuales no es posible determinar a priori la exactitud de los pronósticos ex ante.

- **Periodo anterior a Y_{inicio} :** Caracterizado por la ausencia de datos sobre la variable. Es posible realizar back-casting (estimación de valores pasados y desconocidos) de valores de la serie, lo cual es hecho en ocasiones (según el modelo o metodología aplicada) cuando el periodo muestral no contiene suficientes observaciones para el análisis o ajuste.

3.7. Estudio descriptivo de las series de tiempo - componentes en una serie de tiempo

Para la selección apropiada de los modelos para una serie de tiempo es preciso identificar los patrones básicos o componentes en ésta. Estas componentes son:

3.7.1. Componente de tendencia (T_t)

Patrón de largo plazo caracterizado por la persistencia a un crecimiento o a un decrecimiento de los valores de la serie, por tanto refleja el crecimiento o declinación de la serie (Gaynor and Kirkpatrick, 1994). Cuando su comportamiento es predecible, puede ser representada con una función suave del tiempo¹, por ejemplo como las ilustradas en la Figura 3.3.

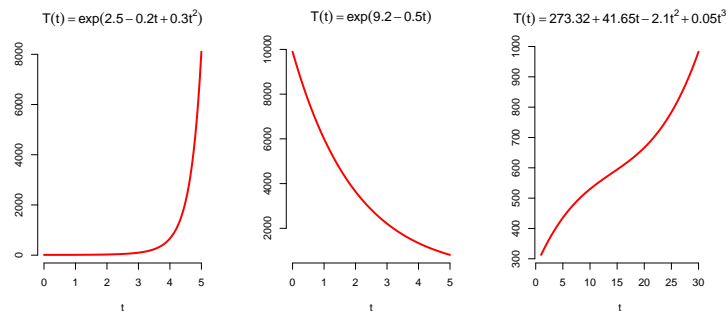


Figura 3.3: Ejemplos patrones de tendencia determinísticas

3.7.2. Componente estacional (S_t)

En el caso determinístico es definida como el patrón de cambio regular que se completa dentro de un año calendario y que se repite sobre una base anual. Este comportamiento es debido a factores tales como los consumos o las estaciones climáticas, es decir factores que ocurren con una periodicidad bien sea semanales, mensuales, trimestrales, o semestrales. Las ventas de productos como bebidas, helados, prendas de vestir, juguetería, etc. están sujetos a este tipo de patrones (Gaynor and Kirkpatrick, 1994). La Figura 3.4 ejemplifica un modelo determinístico para la componente estacional periódica exacta.

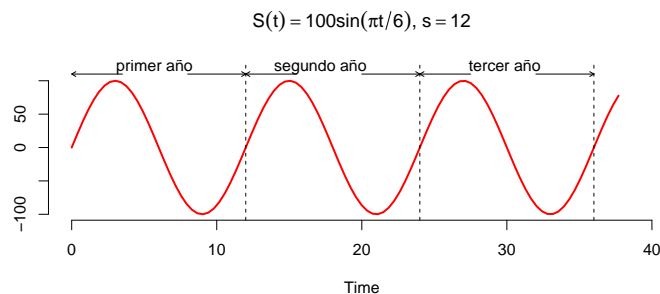


Figura 3.4: Ejemplo patrón estacional periodico exacto

3.7.3. Componente cíclica (C_t)

Corresponden a cambios o movimientos hacia arriba y hacia abajo que ocurren sobre una duración de 2 o más años, debido a la influencia de fluctuaciones económicas de largo plazo como las asociadas con los ciclos de negocios. Por ejemplo las ventas de propiedad raíz, el ciclo de negocios medido en términos del PIB (Gaynor and Kirkpatrick,

¹Una función continua, derivable y con pocos cambios de concavidad

1994): Un ciclo es medido desde un pico a otro o desde una depresión a otra, de modo que un periodo de prosperidad es seguido por un periodo de recesión, lo cual hace que cambie la tendencia de los datos hacia arriba (expansión) a una tendencia hacia abajo (contracción); el movimiento hacia abajo alcanza su punto más mínimo luego de lo cual comienza el movimiento hacia arriba hasta un punto máximo. Los ciclos son las componentes más difíciles de pronosticar debido a sus intervalos de tiempo más largos. La diferencia entre ciclos y patrones estacionales es que los segundos tienen una longitud y amplitud constante que se repite sobre una base regular periódica, mientras que los ciclos varían en longitud y magnitud. El propósito de varios de los métodos de series de tiempo es eliminar las irregularidades e influencias estacionales y proyectar la serie de datos más bien con base en sus patrones de tendencia/ciclo (Campos-Santillan, 2001). La Figura 3.5 ilustra este tipo de patrones.

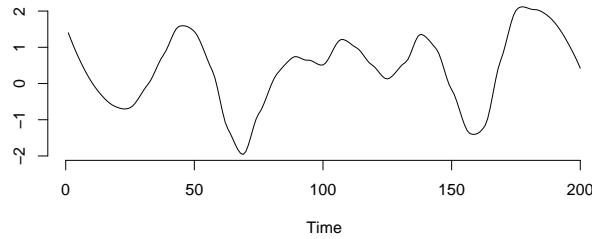


Figura 3.5: Ejemplo de fluctuaciones cíclicas

3.7.4. Componente horizontal, irregular o error (E_t)

También es llamado ruido blanco, caracterizado por fluctuaciones erráticas sin un patrón definido alrededor de una media constante. Muy a menudo estas fluctuaciones son debidas a eventos externos que sólo ocurren en un tiempo y de forma impredecible (Gaynor and Kirkpatrick, 1994). Se dice de una serie que sólo exhibe tal comportamiento que es estacionaria en su media. Tal serie por ejemplo, aparece o es asumida en las situaciones del control estadístico de procesos. La Figura 3.6 ilustra una realización de este tipo de componente.

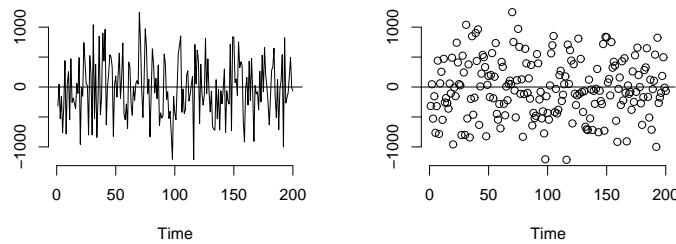


Figura 3.6: Ejemplo patrón componente irregular

3.8. Modelos de descomposición

Una serie de tiempo puede ajustarse a uno de los cuatro patrones descritos o ser una combinación de todos ellos. En principio solo considere las componentes de tendencia, estacionalidad y error. Estas tres componente pueden combinarse así:

- En forma aditiva (serie de componentes aditivas),

$$Y_t = T_t + S_t + E_t, \quad (3.1)$$

- en forma multiplicativa (serie de componentes completamente multiplicativas),

$$Y_t = T_t \times S_t \times E_t. \quad (3.2)$$

Sin embargo, adoptaremos la siguiente ecuación, que no requiere que el término de error sea una variable no nula y positiva,

$$Y_t = T_t \times S_t \times \exp(E_t). \quad (3.3)$$

- Otra posible forma de descomposición es la parcialmente multiplicativa,

$$Y_t = T_t \times S_t + E_t. \quad (3.4)$$

Nota 3.1. De acuerdo a los modelos generales plantados en las ecuaciones (3.1), (3.2), (3.3) y (3.4), los ciclos y cualquier otro fenómeno temporal no explicado y la componente de error quedarán mezclados en E_t . Sin embargo, se tomará preliminarmente como supuesto que los errores satisfacen $E_t \stackrel{iid}{\sim} N(0, \sigma^2)$.

Para escoger entre una descomposición aditiva o multiplicativa, *es necesario evaluar la variabilidad de la serie alrededor de su media de largo plazo: es decir alrededor de la tendencia*, de modo que la serie será

- Aditiva, si tal variabilidad es constante. Ver Figura 3.7(a)
- Multiplicativa, si tal variabilidad no es constante y cambia en la dirección de la tendencia (crece cuando la tendencia crece y decrece cuando la tendencia decrece). Ver Figura 3.7(b). Sin embargo, no podemos gráficamente establecer cual modelo multiplicativo sería más apropiado, entre las versiones completa y parcialmente multiplicativas.

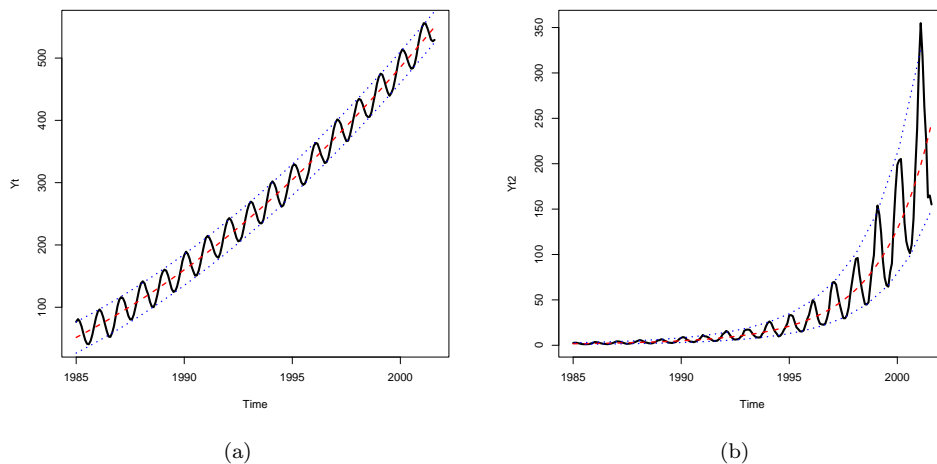


Figura 3.7: (a) Ejemplo serie estacional y aditiva: Variabilidad alrededor de su tendencia es constante; (b) Ejemplo serie estacional multiplicativa: La variabilidad alrededor de su tendencia cambia en la dirección de ésta.

El análisis gráfico de las series de tiempo es una buena manera de identificar estas componentes. Para ello se puede recurrir al denominado **filtro de la descomposición clásica** el cual opera bajo las ecuaciones (3.1) y (3.2) y que está implementado en R en la función `decomposise()` (ver Apéndice al Final de este capítulo). La Figura 3.8 presenta una serie estacional aditiva (Índice mensual de salario real sector industrial, enero de 1990 - mayo de 2011, Fuente: MMM-DANE, año base 1990) y sus componentes obtenidas por el filtro de la descomposición aditiva:

- Tendencia
- Estacionalidad
- Error

En R podemos obtener la descomposición de una serie a través `decompose()`. A continuación se ilustra el código R usado sobre la serie en la Figura 3.8, el código incluye los datos leídos a través de la función `scan()` y su conversión a un objeto R tipo serie de tiempo con la función `ts()`, en la cual se informa con su argumento `freq=12` que la serie es mensual y con su argumento `start=c(1990,1)` se indica que la serie tiene como fecha inicial enero de 1990, luego la graficación de la serie, la descomposición de ésta de forma aditiva con la función `decompose()` cuyos resultados son presentados en la Figura 3.8(b), y la extracción de cada una de las tres componentes, las cuales son presentadas en las gráficas de la Figura 3.9.

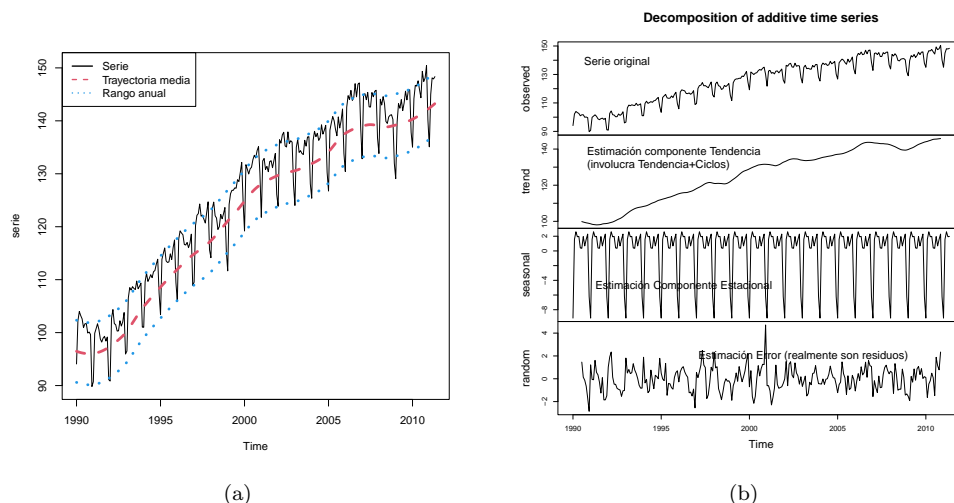


Figura 3.8: (a) Observaciones entre enero de 1990 a mayo de 2011; (b) Descomposición aditiva con función R `decompose`. Esta serie se considera aditiva pues su variabilidad alrededor de su tendencia no cambia en el tiempo. Observe además que tiene componente estacional.

Código R 3.1. Lectura de datos, creación serie de tiempo y gráficas descriptivas para la serie en la Figura 3.8, disponible en el Script R: *ProgramaREjemplodescripcionSerieSalario.R*:

```
salario=ts(scan(,dec="."),freq=12,start=c(1990,1))
94.08 102.23 103.99 103.21 102.59 100.95 101.67 101.71 99.93 100.11
99.65 89.81 90.71 99.88 100.75 101.67 100.58 98.82 98.21 98.96
99.3 98.4 99.17 91.12 90.85 102.67 104.36 101.71 101.48 100.29
100.33 102.24 101.24 102.02 103.79 95.95 96.52 106.55 108.5 108.04
108.68 108.34 108.14 109.75 108.28 108.76 109.36 101.03 101.01 109.63
111.2 109.67 110.24 110.94 110.38 111.5 111.73 112.94 113.87 107.39
103.38 113.84 114.96 116 113.54 113.25 113.35 114.88 113.81 116.03
117.5 109.5 105.87 117.69 119.22 118.8 118.48 116.65 116.54 117.03
116.04 116.87 118.6 108.91 109.53 120.21 122.15 122.48 124.3 121.99
122.2 121.41 120.69 122.72 124.7 115.65 114.15 123.45 124.66 122.08
121.64 119.48 120.02 122.33 121.5 122.45 123.63 115.43 111.66 124.21
126.22 126.88 126.88 127.33 127.27 128.88 128.54 130.28 132.15 124.65
119.19 131.72 130.87 129.96 130.92 131.24 132.67 133.5 132.99 134.38
135.13 130.08 121.78 132.43 132.89 130.82 131.39 130.39 132.31 132.47
131.14 132.5 133.83 126.09 123.87 136.44 135.78 137.33 137.84 135.26
135.06 137.91 135.72 136.25 135.97 127.91 124.02 135.03 136.45 134.04
134.37 133.84 133.64 136.33 135.22 136.95 137 128.09 125.38 136.31
138.02 137.01 137.65 136.31 136.8 139.25 137.49 138.23 139.16 130.28
126.74 138.23 139.6 140.81 140.03 138.36 137.18 140.74 138.65 140.04
142.24 133.84 130.43 142.58 144.76 144.01 145.75 143.87 144.32 147.02
144.53 146.91 147.16 138.52 132.83 144.78 145.13 143.98 145.36 142.64
142.6 145.35 144.54 145.58 145.77 136.78 133.83 144.09 142.98 144.57
143.24 139.88 139.52 140.74 141.06 141.16 140.26 131.98 129.08 139.77
143.25 142.38 144.01 141.93 143.37 145.61 143.95 145.68 145.87 137.84
135.03 144.19 146.7 146.56 147.56 145.87 144.36 149.35 147.43 148.08
150.45 140.15 135.09 144.99 148.11 147.91 148.35

win.graph()
plot(salario)

#Grafica de la descomposicion aditiva de la serie mediante la funcion decompose
win.graph()
plot(decompose(salario,type="additive"))

#Grafica individual de la tendencia estimada por el filtro
win.graph(width=4.0, height=3.5)
plot(decompose(salario)$trend,ylim=c(min(salario),max(salario)),ylab="Tendencia (Tendencia+ciclos)",lwd=3)

#Grafica individual de la estacionalidad estimada por el filtro
win.graph(width=4.0, height=3.5)
plot(decompose(salario)$seasonal,ylab="Estacionalidad")

#Grafica individual de los residuos (lo que el filtro llama error "random")
win.graph(width=4.0, height=3.5)
plot(decompose(salario)$random,ylab="Error (residuos)");abline(h=0,col=2)
```

La función `decompose` (ver Apéndice A) filtra la tendencia y estima los efectos estacionales mediante promedios móviles. Podemos valernos de esta función como una herramienta descriptiva útil en la definición de modelos estadísti-

cos basados en regresión.

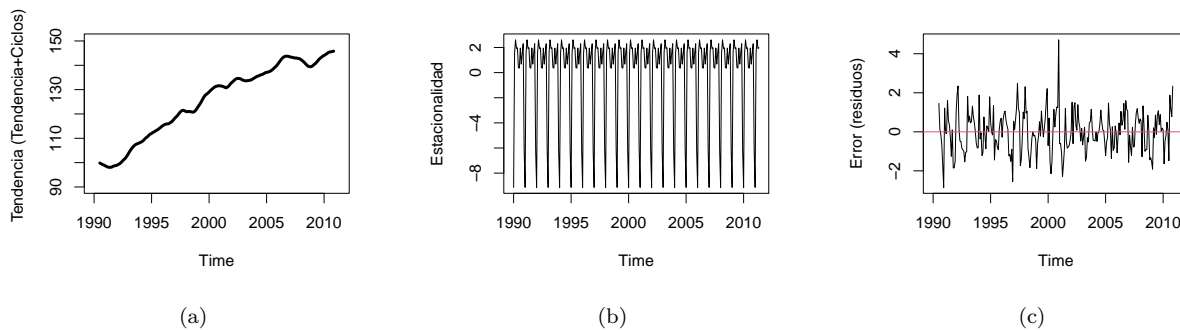


Figura 3.9: Estimaciones de las componentes serie de la Figura 3.8(a), con filtro de la descomposición aditiva: (a) Componente de tendencia (realmente queda tendencia más ciclos); (b) Componente estacional; (c) Componente de error (realmente es un residuo)

Note además que lo que el filtro arroja como estimación de tendencia (ver Figura 3.9(a)) no es una curva lisa sino que presenta deformaciones como cimas y que corresponden a los movimientos generados por la componente cíclica, es decir, tal estimación de tendencia realmente involucra tendencia y ciclos, de modo que al momento de tratar de identificar alguna función del tiempo para la tendencia debemos hacerlo pensando en la trayectoria más suave posible sin pretender seguir los ciclos, por ejemplo, como ilustra la siguiente Figura.

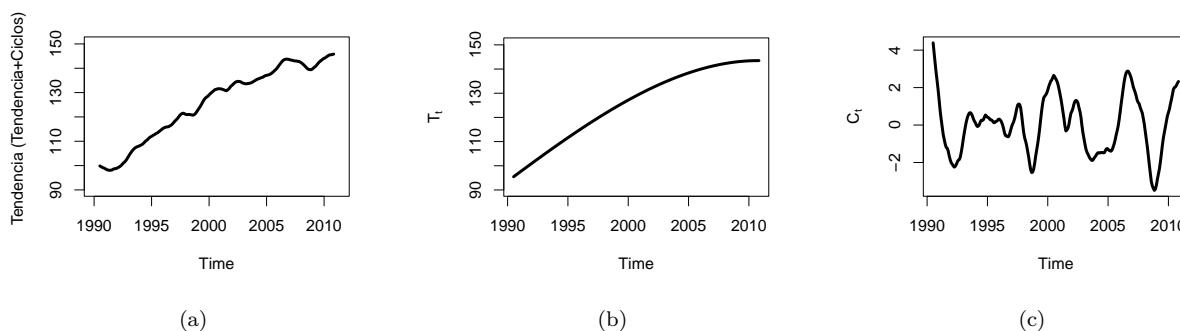


Figura 3.10: Análisis de la estimación de tendencia con filtro de descomposición de la serie en la Figura 3.8(a): (a) Tendencia y ciclos; (b) Posible tendencia pura; (c) Ciclos obtenidos como la diferencia entre (a) y (b).

Una serie de componentes multiplicativas es presentada en la Figura 3.11(a), esta serie corresponde al valor de las ventas de licor en Estados Unidos (miles de dólares nominales), enero de 1967 - diciembre de 1994, Fuente: Diebold, 2001. Al igual que en la serie anterior, en ésta última es visible el patrón estacional de periodicidad anual y una tendencia a crecer a la larga, pero la variabilidad alrededor de su trayectoria de largo plazo está creciendo conforme crece la tendencia, por eso el modelo de descomposición a considerar debe ser de tipo multiplicativo y de ahí que en la Figura 3.11(b) se presente la gráfica obtenida con el filtro de la descomposición clásica multiplicativa.

También podemos obtener por separado cada una de las estimaciones de las componentes de tendencia, estacionalidad y error (residuos realmente) como se mostró en el ejemplo previo, sin embargo, vamos a considerar a continuación lo siguiente: Note que bajo el modelo de descomposición en la ecuación (3.2), si todas las componentes en el lado derecho son positivas y no nulas, este modelo se puede transformar a uno aditivo a través de la transformación logaritmo natural,

$$\log(Y_t) = \log(T_t) + \log(S_t) + \log(E_t) \quad (3.5a)$$

$$Y_t^* = T_t^* + S_t^* + E_t^*. \quad (3.5b)$$

En (3.5) aparece la transformación del modelo multiplicativo y observe que redefiniendo los logaritmos en (3.5a) por las correspondientes variables en (3.5b) la transformación nos lleva a un modelo aditivo sobre los logaritmos de la serie original.

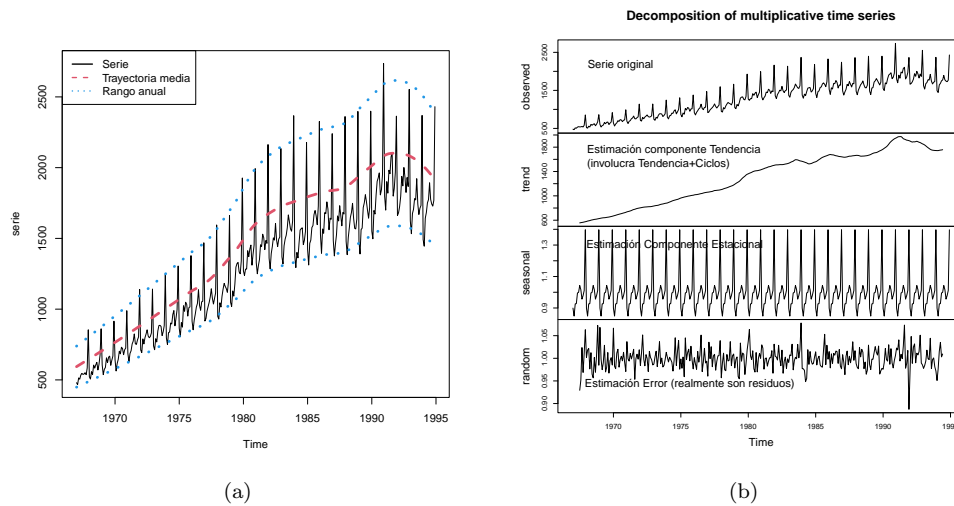


Figura 3.11: (a) Observaciones entre enero de 1967 - diciembre de 1994 (miles de dólares), Fuente: Diebold, 2001; (b) Descomposición multiplicativa con función R `decompose`. Esta serie se considera multiplicativa pues su variabilidad alrededor de su tendencia cambia en el tiempo en la dirección de la tendencia. Observe que tiene componente estacional.

También bajo la ecuación (3.3), este modelo se puede transformar a uno aditivo a través de la transformación logaritmo natural,

$$\log(Y_t) = \log(T_t) + \log(S_t) + E_t \quad (3.6a)$$

$$Y_t^* = T_t^* + S_t^* + E_t \quad (3.6b)$$

pero en este último caso, no es necesario atribuir a E_t una distribución de probabilidad que garantice con probabilidad 1 que los errores sean positivos y no nulos y podemos trabajar bajo el supuesto de errores $E_t \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$, en cambio en (3.5a) no podemos asumir errores normales centrados en cero, ya que bajo tal distribución hay 0.5 de probabilidad de valores negativos para los cuales no estaría definido su logaritmo natural. En conclusión, para series multiplicativas de valores positivos y no nulos, se cumple que su logaritmo natural es de componentes aditivas, ver por ejemplo la Figura 3.12(a) del logaritmo natural de la serie de ventas de licor en Estados Unidos.

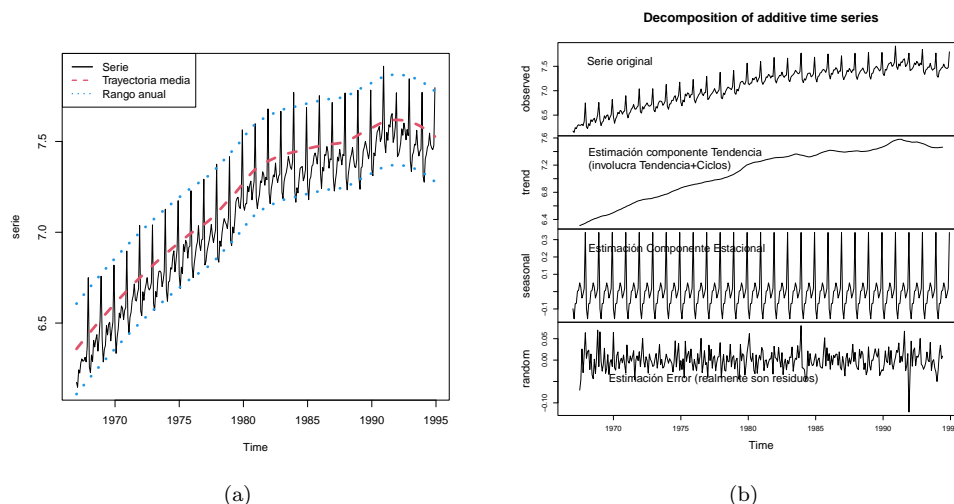


Figura 3.12: (a) Observaciones entre enero de 1967 - diciembre de 1994 (log. miles de dólares); (b) Descomposición aditiva con función R `decompose`. El logaritmo de los datos se considera una serie aditiva, pues su variabilidad alrededor de su tendencia es constante en el tiempo. Observe que el patrón estacional se conserva con la transformación.

De otro lado, en los modelos de regresión que veremos, la modelación de las componentes para series multiplicativas es realizada primero en la escala logaritmo natural, por ello, en lugar de analizar las componentes de la descomposición multiplicativa de la serie original, procedemos con las componentes de la descomposición aditiva del logaritmo natural de la serie, como puede observarse en la Figura 3.12(b), que presenta la descomposición aditiva de la serie de los

logaritmos de las ventas de licor. Adicionalmente, podemos ver por separado cada componente estimada por el filtro, como ilustra la Figura 3.13. De nuevo, note que la estimación de tendencia por el filtro de la descomposición aditiva sobre el logaritmo natural de la serie de las ventas nominales de licor en la Figura 3.13(a), no es solo tendencia sino tendencia más ciclos, de modo que para la formulación de un modelo de regresión de esta componente es necesario pensar en la curva más suave posible para la trayectoria de largo plazo (la media de largo plazo) sin incluir los ciclos, como se sugiere en la Figura 3.14.

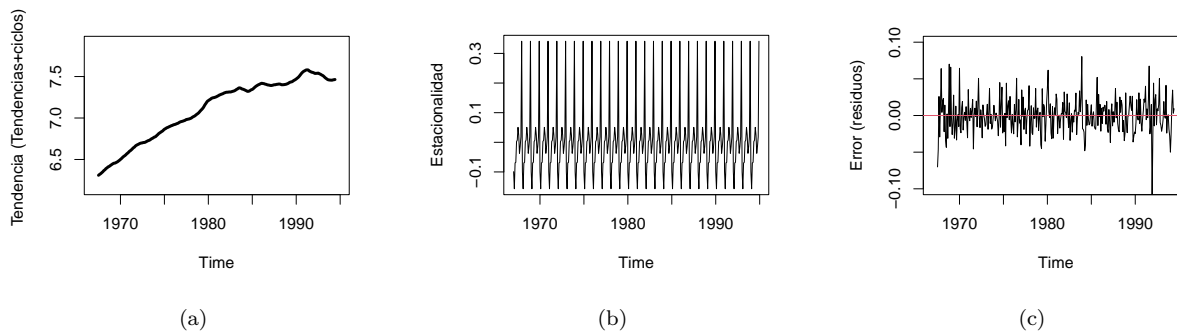


Figura 3.13: Estimaciones de las componentes de la serie en la Figura 3.12(a), con filtro de la descomposición aditiva: (a) Componente de tendencia (realmente queda tendencia más ciclos); (b) Componente estacional; (c) Componente de error (realmente es un residuo)

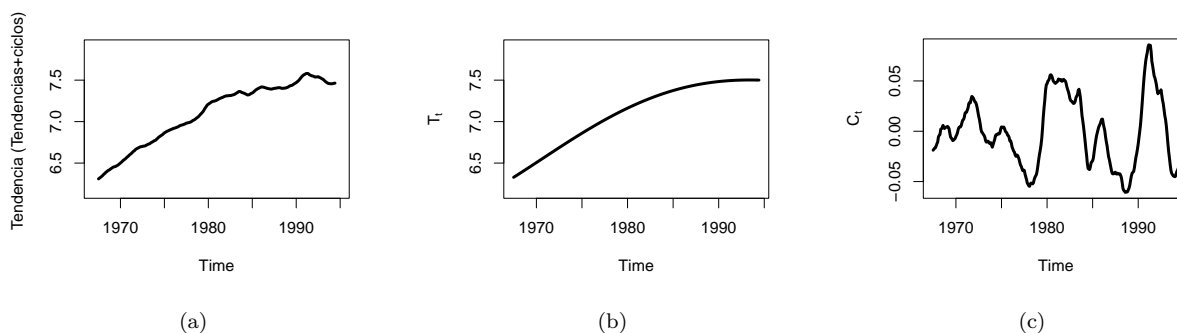


Figura 3.14: Análisis de la estimación de tendencia con filtro de descomposición de la serie en la Figura 3.12(a): (a) Tendencia y ciclos; (b) Posible tendencia pura; (c) Ciclos obtenidos como la diferencia entre (a) y (b).

A continuación se muestra el código R utilizado sobre los datos de las ventas de licor en Estados Unidos.

Código R 3.2. *Lectura de datos, creación serie de tiempo y gráficas descriptivas para la serie en la Figura 3.11 y para su logaritmo natural, disponible en el Script R:*

Programa REjemplodescripcionSerieLicorUSA.R:

```
licorUSA=ts(scan(),frequency=12,start=c(1967,1))
480 467 514 505 534 546 539 541 551 537 584 854
522 506 558 538 605 583 607 624 570 609 675 861
605 537 575 588 656 623 661 668 603 639 669 915
643 563 616 645 703 684 731 722 678 713 725 989
687 629 687 706 754 774 825 755 751 783 804 1139
711 693 790 754 799 824 854 810 798 807 832 1142
740 713 791 768 846 884 886 878 813 840 884 1245
796 750 834 838 902 895 962 990 882 936 997 1305
866 805 905 873 1024 985 1049 1034 951 1010 1016
1378 915 854 922 965 1014 1040 1137 1026 992 1052
1056 1469 916 934 987 1018 1048 1086 1144 1077 1036
1076 1114 1595 949 930 1045 1015 1091 1142 1182 1161
1145 1119 1189 1662 1048 1019 1129 1092 1176 1297
1322 1330 1263 1250 1341 1927 1271 1238 1283 1283
1413 1371 1425 1453 1311 1387 1454 1993 1328 1250
1308 1350 1455 1442 1530 1505 1421 1485 1465 2163
1361 1284 1392 1442 1504 1488 1606 1488 1442 1495
1509 2135 1369 1320 1448 1495 1522 1575 1666 1617
1567 1551 1624 2367 1377 1294 1401 1362 1466 1559
1569 1575 1456 1487 1549 2178 1423 1312 1465 1488
1577 1591 1669 1697 1659 1597 1728 2326 1529 1395
```

```

1567 1536 1682 1675 1758 1708 1561 1643 1635 2240
1485 1376 1459 1526 1659 1623 1731 1662 1589 1683
1672 2361 1480 1385 1505 1576 1649 1684 1748 1642
1571 1567 1637 2397 1483 1390 1562 1573 1718 1752
1809 1759 1698 1643 1718 2399 1551 1497 1697 1672
1805 1903 1928 1963 1807 1843 1950 2736 1798 1700
1901 1820 1982 1957 2076 2107 1799 1854 1968 2364
1662 1681 1725 1796 1938 1871 2001 1934 1825 1930
1867 2553 1624 1533 1676 1706 1781 1772 1922 1743
1669 1713 1733 2369 1491 1445 1643 1683 1751 1774
1893 1776 1743 1728 1769 2431

#Grafica serie en su escala original
plot(licorUSA,ylab="serie")

#Grafica descomposicion multiplicativa serie en su escala original
win.graph()
plot(decompose(licorUSA,type="multiplicative"))

#Grafica serie en escala log natural
win.graph()
plot(log(licorUSA),ylab="serie")

#Grafica descomposicion aditiva de logaritmo natural de la serie
win.graph()
plot(decompose(log(licorUSA),type="additive"))

#Graficas de las componentes separadas, descomposicion aditiva de log natural de la serie
win.graph(width=4.0, height=3.5)
plot(decompose(log(licorUSA),type="additive")$trend,ylim=c(min(log(licorUSA)),max(log(licorUSA))),
      ylab="Tendencia (Tendencias+ciclos)",lwd=3)

win.graph(width=4.0, height=3.5)
plot(decompose(log(licorUSA),type="additive")$seasonal,ylab="Estacionalidad")

win.graph(width=4.0, height=3.5)
plot(decompose(log(licorUSA),type="additive")$random,ylab="Error (residuos)",ylim=c(-0.1,0.1))
abline(h=0,col=2)

```

3.9. Guía para el análisis de las componentes de una serie

3.9.1. Determinación del tipo de modelo a considerar entre modelos de componentes aditivas vs. multiplicativas

Como previamente se definió, una serie es de componentes aditivas si su varianza es constante alrededor de su trayectoria de largo plazo (la tendencia), en caso contrario, se considera que la serie es de componentes multiplicativas. Entonces podemos evaluar esto mediante,

La gráfica de Y_t vs. t :

Por ejemplo, como en las gráficas de la Figura 3.15.

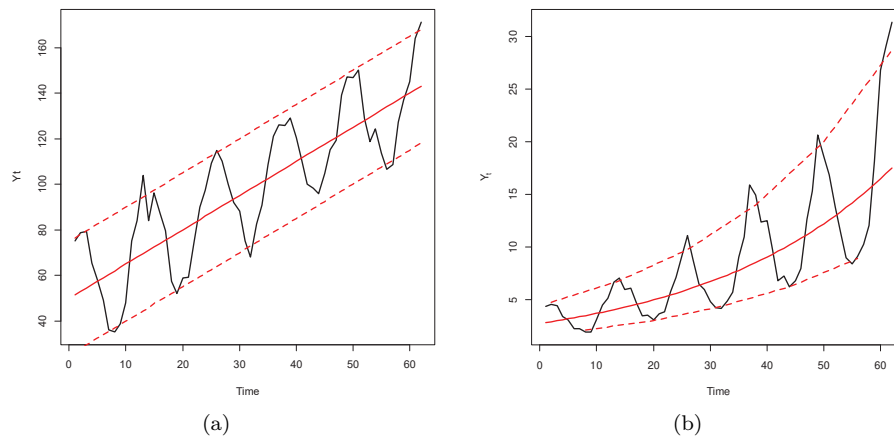


Figura 3.15: (a) La varianza de la serie alrededor de su tendencia es constante (componentes aditivas); (b) La varianza de la serie alrededor de su tendencia no es constante y cambia en la dirección de la tendencia (componentes multiplicativas)

Comparación del patrón de la varianza de Y_t vs. t contra el patrón de la varianza de $\log(Y_t)$ vs. t

Si ambas gráficas muestran el mismo patrón de varianza y sólo hay diferencias en la escala de los datos, o si el patrón de la varianza para $\log(Y_t)$ es peor que la exhibida por Y_t , entonces es mejor considerar que la serie es de componentes aditivas. Pero si en la gráfica de $\log(Y_t)$ vs. t se observa un patrón de varianza más estable que en la gráfica de Y_t vs. t , entonces es mejor considerar que la serie es de componentes multiplicativas. Por ejemplo, para la serie en la Figura 3.16 el patrón de varianza para $\log(Y_t)$ no es mejor (empeoró) que el patrón de la varianza de la serie sin transformar.

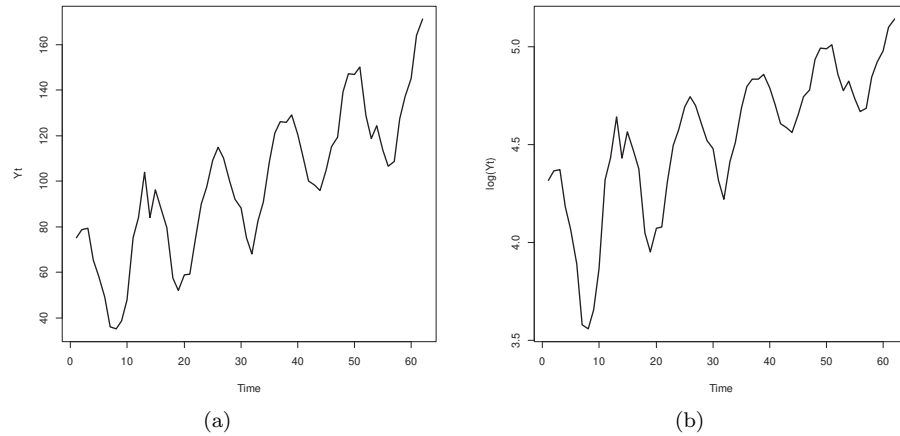


Figura 3.16: Caso aditivo: (a) Serie en escala original; (b) Serie en escala logaritmo natural.

En cambio, para la serie en la Figura 3.17 el patrón de varianza para $\log(Y_t)$ es más estable que el patrón de la varianza de la serie sin transformar.

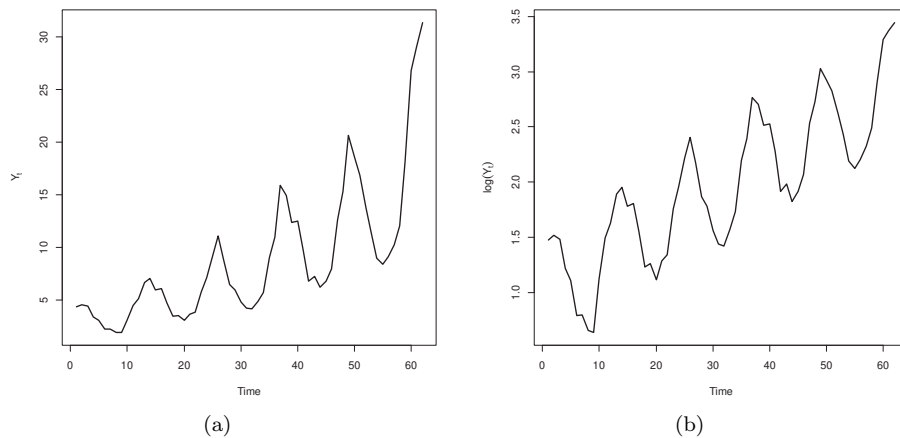


Figura 3.17: Caso multiplicativo: (a) Serie en escala original; (b) Serie en escala logaritmo natural.

Comparación del patrón de la varianza de los residuos en la descomposición aditiva de la serie vs. el patrón de la varianza de los residuos en la descomposición multiplicativa

Las gráficas de los residuos en ambas descomposiciones son obtenidas combinando las funciones R `decompose()` y `plot()` como se muestra a continuación:

```
plot(decompose(serie,type="additive")$random); abline(h=0)
plot(decompose(serie,type="multiplicative")$random); abline(h=1)
```

El objeto `serie` son los datos pero definidos como un objeto serie de tiempo usando la función `ts()`. Observe que los residuos \hat{E}_t en la descomposición aditiva es una serie centrada en cero mientras que en la descomposición multiplicativa es una serie centrada en 1. Si los \hat{E}_t de la descomposición multiplicativa muestra un patrón de varianza más estable que \hat{E}_t de la descomposición aditiva, entonces es mejor considerar a la serie Y_t como de componentes multiplicativas (Ver Figura 3.18). Pero si los \hat{E}_t de la descomposición multiplicativa muestra patrón de varianza igual o menos estable que el de los \hat{E}_t de la descomposición aditiva, entonces la serie es de componentes aditivas (Ver Figura 3.19).

Nota 3.2. Si se identifica que las componentes de la serie deben considerarse como multiplicativas, para la proposición de modelos de regresión debe analizarse la descomposición aditiva del logaritmo natural de la serie para determinar qué funciones del tiempo pueden formularse para las componentes de tendencia y estacionalidad en la escala logarítmica.

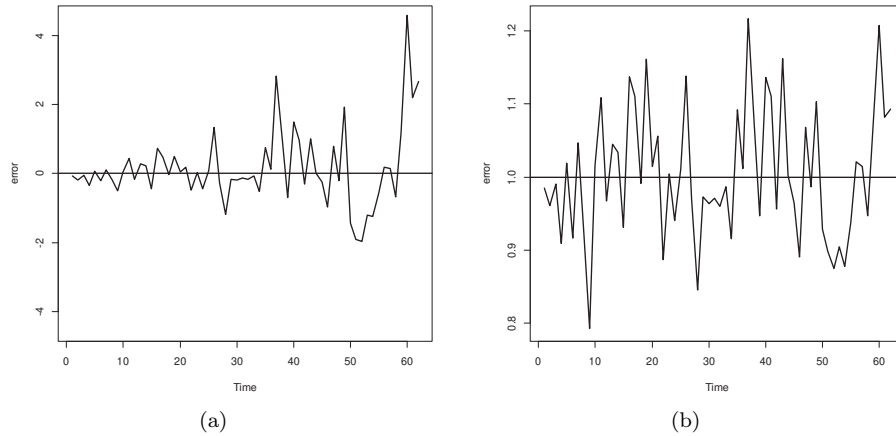


Figura 3.18: Caso multiplicativo: (a) \hat{E}_t en la descomposición aditiva de la serie en la Figura 3.17; (b) \hat{E}_t en la descomposición multiplicativa de la serie en la Figura 3.17. El patrón de varianza de \hat{E}_t en la descomposición multiplicativa es más estable que el de \hat{E}_t en la descomposición aditiva.

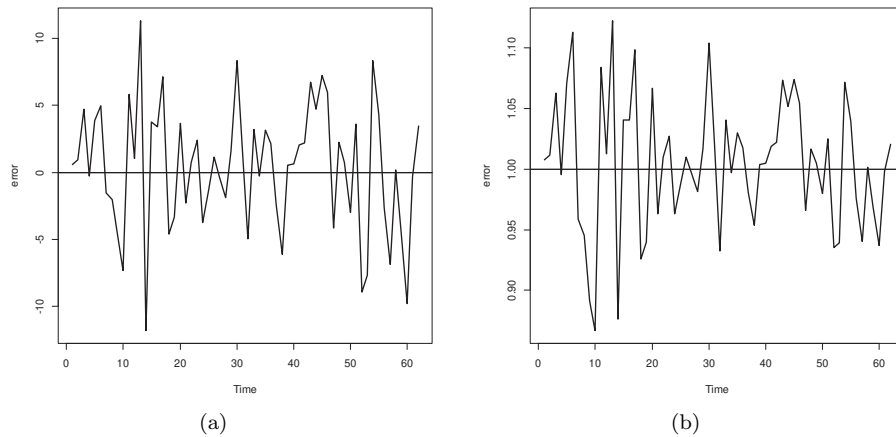


Figura 3.19: Caso aditivo: (a) \hat{E}_t en la descomposición aditiva de la serie en la Figura 3.16; (b) \hat{E}_t en la descomposición multiplicativa de la serie en la Figura 3.16. El patrón de varianza de \hat{E}_t en la descomposición multiplicativa no es más estable que el de \hat{E}_t en la descomposición aditiva.

3.9.2. Determinación de la presencia de la componente estacional periódica exacta o casi exacta

Para determinar si en una serie existe patrón estacional periódico exacto o casi exacto, podemos proceder mediante:

- Identificación de la existencia en la serie de un patrón que se repite año tras año, y cuya forma es aproximadamente la misma a lo largo de los años (ver Figura 3.20(a)).
- Determinar si la localización (media o mediana) de la distribución de la serie cambia según el período del año (meses en series mensuales o trimestres en series trimestrales). Si esto ocurre entonces se concluye que cada período en que se divide el año calendario determina una estación, la cual tiene incidencia sobre el valor medio de la serie (ver Figura 3.20(b)).
- El periodograma (será estudiado más adelante): Evalúa la existencia de asociaciones entre una serie de tiempo y ondas sinusoidales armónicas.

3.9.3. Evaluación de la tendencia

Para un modelo de regresión sobre una serie de componentes aditivas se debe identificar cuál función suave del tiempo pudiera representar a la componente T_t que arroja el filtro de la descomposición aditiva, como se ilustró previamente en la Figura 3.10, en tanto que para una serie de componentes multiplicativas debemos identificar primero

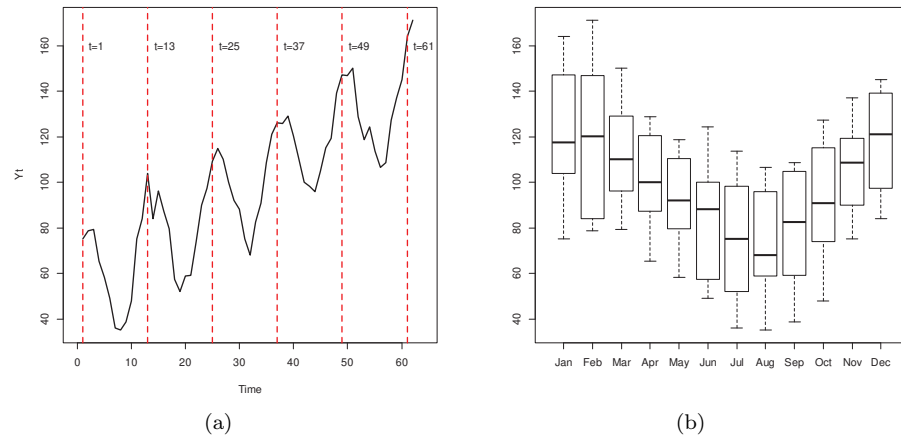


Figura 3.20: (a) Gráfica para identificar que existe un patrón que se repite cada año, donde las líneas verticales que separan los años revelan un patrón repetitivo con forma de V; (b) Las líneas centrales de cada caja representan la mediana, y vemos que su posición varía de acuerdo al mes del año

tal función sobre la tendencia filtrada de forma aditiva sobre la serie con transformación logaritmo natural como se ilustró en la Figura 3.14. La extracción de esta componente es como sigue,

```
#Caso serie aditiva
decompose(serie,type="additive")$trend

#Caso serie multiplicativa: aplicamos filtro sobre su logaritmo natural
decompose(log(serie),type="additive")$trend
```

Esta componente debe graficarse en el rango de variación de los datos (o del logaritmo de los datos si se analiza en tal escala). En el caso aditivo procedemos así (ver Figura 3.21):

```
Tt=decompose(serie,type="additive")$trend
plot(Tt,ylim=c(min(serie),max(serie)))
```

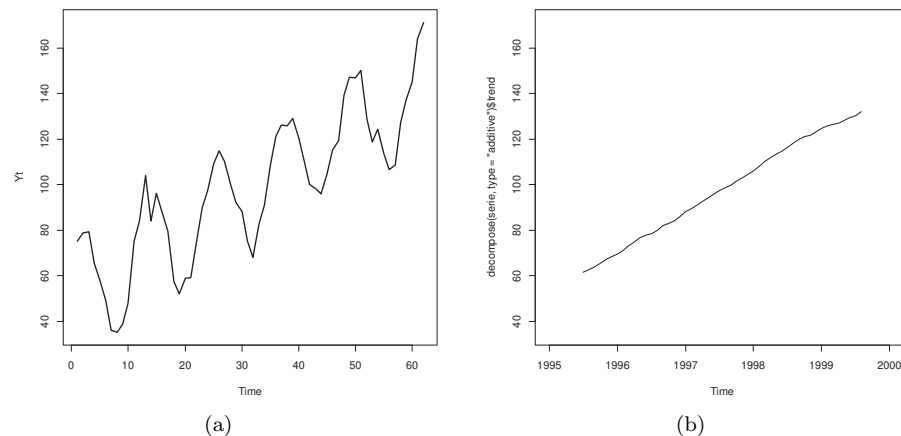


Figura 3.21: Caso aditivo: (a) Serie aditiva en su escala original; (b) Componente de tendencia según la descomposición aditiva de la serie usando la función R `decompose()`. Los valores máximo y mínimo del eje vertical en esta última gráfica deben coincidir con los de la serie.

En el caso multiplicativo, con la serie bajo transformación logaritmo natural (ver Figura 3.22),

```
Tt.log=decompose(log(serie),type="additive")$trend
plot(Tt.log,ylim=c(min(log(serie)),max(log(serie))))
```

Nota 3.3. Si para la tendencia del logaritmo natural se identifica por ejemplo que su modelo puede ser $T_t^* = \beta_0 + \beta_1 t + \beta_2 t^2$, entonces estamos diciendo que para la serie en su escala original su tendencia corresponde a $T_t = \exp(T_t^*) = \exp(\beta_0 + \beta_1 t + \beta_2 t^2)$.

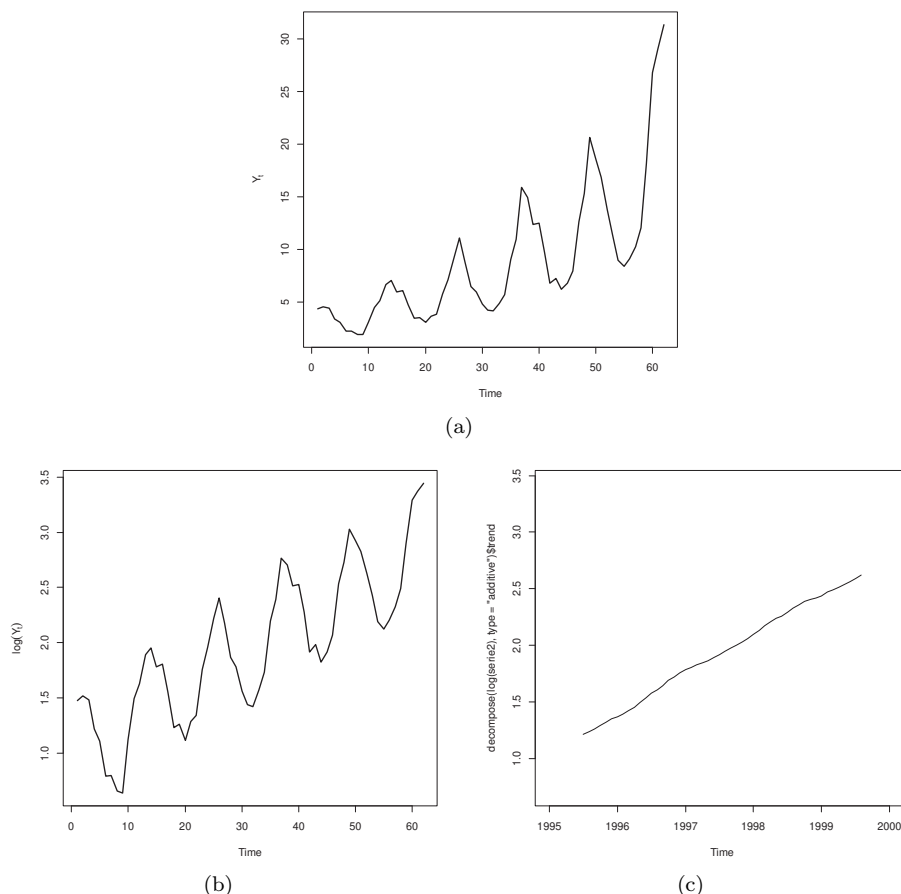


Figura 3.22: Caso multiplicativo: (a) Serie multiplicativa en su escala original; (b) Logaritmo natural de la serie; (c) Tendencia del logaritmo natural según su descomposición aditiva con la función R `decompose()`. Los valores máximo y mínimo del eje vertical en esta última gráfica deben coincidir con los del logaritmo natural de la serie.

Nota 3.4. Muchas series presentan una tendencia filtrada con el filtro de la descomposición clásica sobre la cual es difícil identificar una curva suave que pueda representar apropiadamente una función de tendencia separada de la componente cíclica. En esos casos la modelación por regresión descomponiendo solo en tendencia, estacionalidad y error no conduce a buenos ajustes y predicciones, además de no validar supuestos como el de la independencia entre los errores del modelo. Ver por ejemplo la serie en la Figura 3.23.

Código R 3.3. El código R usado para el ejemplo de la Figura 3.23 es el siguiente, el cual se halla disponible en el Script R “ProgramaREjemplosdeguiaparaanálisiscomponentesdeunaserie.R”:

```
rm(list=ls(all=TRUE))
#Para leer "total unemployment-monthly-average-france-female editado.csv"
datos=read.table(file.choose(),header=T,skip=11,sep=';',dec=',',colClasses=c("NULL","numeric"))
datos=ts(datos,freq=12,start=c(1983,1))
plot(datos)

#Guardando la estimacion de tendencia de la descomposicion aditiva
Tt=decompose(datos,type="additive")$trend

#Grafica de tendencia estimada por descomposicion aditiva,
#controlando rango del eje vertical con argumento ylim en funcion plot
win.graph()
plot(Tt,ylim=c(min(datos),max(datos)))
```

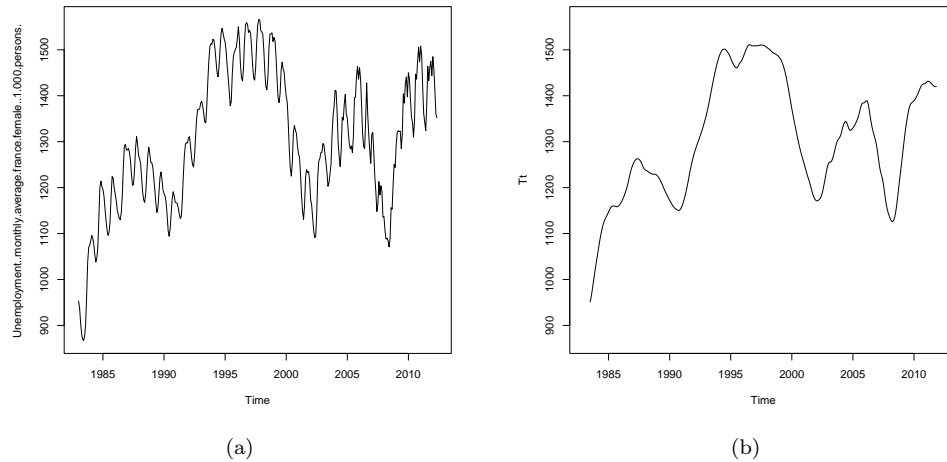


Figura 3.23: (a) La serie mensual de desempleo en Francia, promedio mensual de mujeres (1000 personas), enero-1983 a mayo-2012, fuente: Eurostat, datamarket.com, es de componentes aditivas (la varianza es constante); (b) Componente de tendencia de la descomposición aditiva de la serie obtenida con la función R `decompose()`. En esta gráfica es claro que la trayectoria de largo plazo es fuertemente afectada por variaciones cíclicas, por tanto un modelo de regresión de la forma dada en (3.1) será difícil de formular con un buen ajuste y con una calidad predictiva aceptable.

3.10. Ejemplo de análisis de las componentes de una serie

En la Figura 3.24 se muestra la imagen del archivo de datos sobre la serie de la temperatura ($^{\circ}\text{F}$) promedio/mes, registrada en la localidad del Castillo de Nottingham, Inglaterra, Fuente: Time Series Data Library, datamarket.com; en la Figura 3.25 se muestran las gráficas de esta serie y la de su descomposición aditiva. Note en la Figura 3.25(a) que se trata de una serie estacional y que su varianza alrededor de su trayectoria es constante. En la Figura 3.26 se presentan dos gráficas de la componente de tendencia obtenida en la descomposición aditiva de la serie. Vea en estas últimas cómo afecta el manejo de la escala vertical de la gráfica en nuestra apreciación del tipo o modelo de tendencia a considerar a la hora de formular su ecuación dentro de un modelo de regresión: Con 3.26(a) ¿Una curva muy compleja? o con 3.26(b): a pesar de los ciclos, ¿un modelo tan simple como una recta de pendiente cero?

Month	Mean monthly air temperature (Deg. F) Nottingham Castle 1920-1939
1920-01	40.6
1920-02	40.8
1920-03	44.4
1920-04	46.7
1920-05	54.1
1920-06	58.5
1920-07	57.7
1920-08	56.4
1920-09	54.3
1920-10	50.5
1920-11	42.9
1920-12	39.8
1921-01	44.2
1921-02	39.8
1921-03	45.1
1921-04	47
1921-05	54.1
1921-06	58.7
1921-07	66.3
1921-08	59.9
1921-09	57
1921-10	54.2
1921-11	39.7
1921-12	42.8
1922-01	37.5
1922-02	38.7
1922-03	39.5
1922-04	42.1
1922-05	55.7

Figura 3.24: Vista de la base de datos: “Mean monthly air temperature (Deg. F) Nottingham Castle.csv”

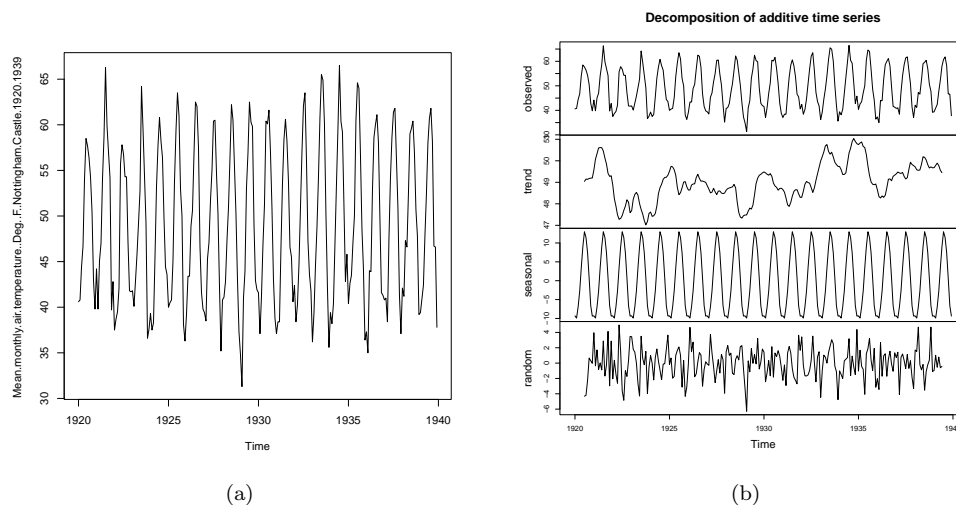


Figura 3.25: (a) Serie en “Mean monthly air temperature (Deg. F) Nottingham Castle.csv”; (b) Descomposición aditiva de la serie con la función R `decompose()`

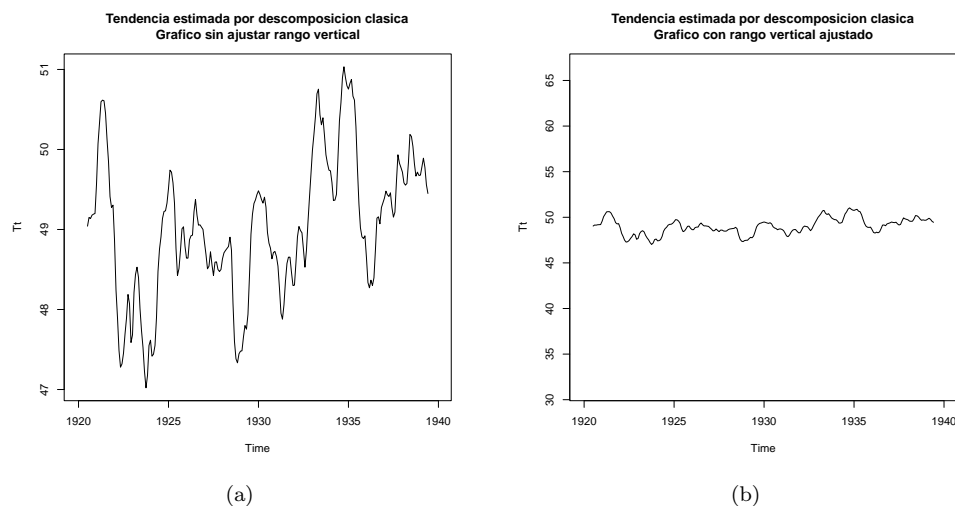


Figura 3.26: (a) Gráfica de la tendencia estimada por la función R `decompose()`, sin ajustar rango del eje vertical; (b) Gráfica de la tendencia estimada por la función R `decompose()`, ajustando el rango de valores del eje vertical al de la serie completa leída desde el archivo “Mean monthly air temperature (Deg. F) Nottingham Castle.csv”.

De la gráfica de la serie vs. tiempo se concluye que aunque hay fenómeno estacional (pues en cada año se observa que en los meses del inicio y final las temperaturas son bajas, mientras que en los meses de mediados del año las temperaturas son altas), a la larga la serie no parece tener una tendencia a crecer o a decrecer, es decir, su componente de tendencia pudiera modelarse como una constante $T_t = \beta_0$; vemos también que la varianza alrededor de esta tendencia es constante por lo que se asumen componentes aditivas.

La conclusión sobre el tipo de tendencia que se obtuvo al evaluar la gráfica de la serie, también debería ser obtenida al evaluar la gráfica de la estimación de esta componente que nos da el filtro de la descomposición aditiva. Sin embargo, observe la Figura 3.26. La gráfica en 3.26(a) se ha realizado sin ajustar el rango de valores del eje vertical, mientras que en la gráfica en 3.26(b) se ha ajustado el rango de valores verticales de modo que coincida con el rango en que se mueve la serie completa. De las dos gráficas de tendencia, se debe considerar la Figura 3.26(b), donde vemos realmente cómo es el patrón de esta componente respecto al rango en que la serie completa varía, y de esta gráfica concluimos también que deberíamos modelar la tendencia como una función constante del tiempo, a pesar de las perturbaciones cíclicas que se observan sobre esa estimación. Si se basa en la Figura 3.26(a), muy probablemente formulará erróneamente tendencias más complejas, como por ejemplo una función cuadrática, lo cual en este caso es absurdo, teniendo en cuenta además la naturaleza de los datos (por lo menos en el rango de tiempo considerado, todavía no se daba como un hecho la posibilidad de tendencias en la temperatura ambiental).

Para el análisis de la componente estacional se evalúa además de la gráfica de la serie, los boxplots comparativos de la distribución según el mes, como se muestra en la Figura 3.27(a), de la cual se concluye que en promedio, la temperatura no es la misma a lo largo del año calendario y que año tras año varía iniciando con un nivel 40 °F (aprox.) en los dos primeros meses del año, luego aumenta progresivamente hasta alcanzar una temperatura media máxima de aproximadamente 62°F en los meses de julio agosto, y comienza descender progresivamente hasta alcanzar nuevamente una temperatura promedio de 40°F para el mes de diciembre. Este es un comportamiento asociado a las estaciones climáticas. Esta variación del promedio de la serie según el período del año calendario nos indica la presencia de la componente estacional, además, la Figura 3.27(b) nos permite ver que este comportamiento en la temperatura se da año a año de forma similar.

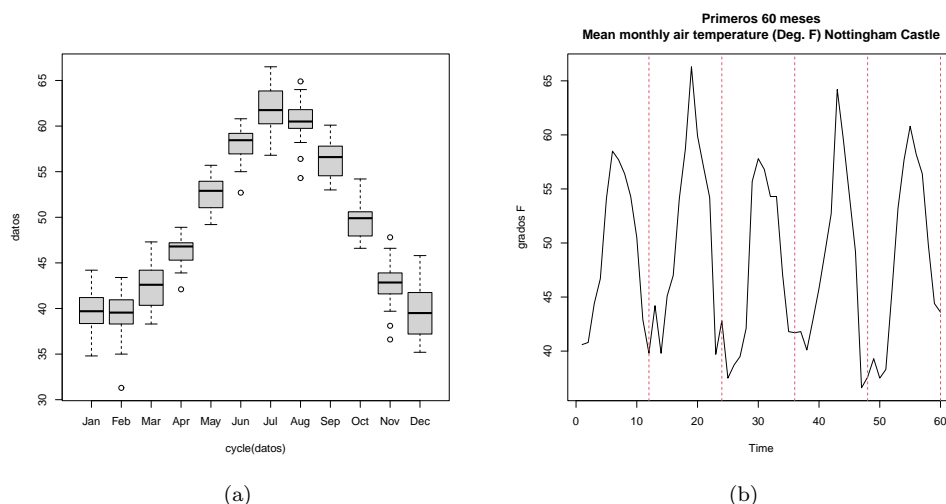


Figura 3.27: (a) Boxplots de la distribución según mes del año, de la serie en el archivo “Mean monthly air temperature (Deg. F) Nottingham Castle”. (b) Gráfica de los primeros cinco años, con líneas de referencia vertical indicando los meses de diciembre.

Código R 3.4. A continuación, el código R utilizado en el ejemplo anterior, el cual se halla disponible en el Script R “ProgramaREjemplosdeguia para analisis componentes de una serie.R”.

```
rm(list=ls(all=TRUE))
#Leer archivo "Mean monthly air temperature (Deg. F) Nottingham Castle.csv"
datos=read.table(file.choose(),header=T,skip=11,sep=';',dec=".",colClasses=c("NULL","numeric"))
#crear serie de tiempo
datos=ts(datos,freq=12,start=c(1920,1))

#Graficando la serie
plot(datos)

#Grafica de la descomposicion aditiva de la serie
win.graph()
plot(decompose(datos,type="additive"))

#Guardando la estimacion de tendencia de la descomposicion
#aditiva
Tt=decompose(datos,type="additive")$trend

#Grafica de tendencia estimada por descomposicion aditiva,
#sin controlar rango del eje vertical
win.graph()
plot(Tt,main="Tendencia estimada por descomposicion clasica\nGrafico sin ajustar rango vertical")

#Grafica de tendencia estimada por descomposicion aditiva,
#controlando rango del eje vertical con argumento ylim en
#funcion plot
win.graph()
plot(Tt,ylim=c(min(datos),max(datos)),
      main="Tendencia estimada por descomposicion clasica\nGrafico con rango vertical ajustado")

#Boxplots para comparar distribucion de la serie segun
#periodos calendario
win.graph()
boxplot(datos~cycle(datos),names=month.abb)

#Grafica de los primeros 60 meses de la serie
```

```
win.graph()
plot.ts(datos[1:60],ylab="grados F",
        main="Primeros 60 meses\nMean monthly air temperature (Deg. F) Nottingham Castle")
abline(v=seq(12,60,by=12),col=2,lty=2) #lineas de referencia en los meses de diciembre
```

Nota 3.5. Ver Capítulo 1, Sección 1.7: Explicación de algunas funciones de interés, para la explicación de funciones de lectura de datos, creación de series de tiempo y construcción de gráficas, entre otras. Ver además el Apéndice A de este Capítulo 3, sobre la función `decompose()`.

Apéndice A

Función `decompose`: Descomposición estacional clásica mediante medias móviles

Esta función R descompone una serie de tiempo en sus componentes de tendencia, estacionalidad y componente de error utilizando medias móviles, para series de componentes aditivas o multiplicativas. Su sintaxis general es la siguiente:

```
decompose(x, type = c("additive", "multiplicative"), filter = NULL)
```

donde,

- **x**: Es una serie de tiempo (objeto del tipo `ts`).
- **type**: El tipo de componente estacional, aditiva ("`additive`") según modelo en (3.1) o multiplicativa ("`multiplicative`") según modelo en (3.2).
- **filter**: Un vector de coeficientes de filtramiento dados en orden inverso del tiempo (como para los coeficientes AR o MA), usado para filtrar la componente estacional. Si su valor es `NULL`, entonces la función aplica una media móvil con ventana simétrica.

La función procede de la siguiente manera: Primero halla la componente de tendencia usando una media móvil (si **filter** es igual a `NULL`, `decompose` usa una ventana simétrica con pesos iguales, es decir, calcula las medias móviles bilaterales) y la remueve de la serie restándola de la serie original (caso aditivo) o dividiendo la serie original por esta componente. Luego, la componente estacional es calculada mediante promedios para cada estación sobre todos los periodos. Esta componente es luego centrada para que los factores estacionales sumen cero (caso aditivo) o normalizados dividiendo por el promedio de los factores estacionales crudos (caso multiplicativo). Finalmente, la componente de error es obtenida eliminando de la serie original la tendencia y la estacionalidad, para esto, si la serie es aditiva, a los datos originales le resta las componentes de tendencia y estacionalidad, y en el caso multiplicativo, divide la serie original por el producto de las componentes de tendencia y estacionalidad.

Valor: La función `decompose` produce un objeto R de la clase "`decomposed.ts`" que tiene los siguientes componentes:

- **seasonal**: La componente estacional para $t = 1, \dots, n$;
- **figure**: Sólo los s factores estacionales estimados;
- **trend**: La componente de tendencia;
- **random**: El error o resto de la serie;
- **type**: El valor especificado para el tipo de serie.

Ejemplo: Considere la serie presentada en la Figura A.1 junto con su descomposición. Esta serie es estacional y es multiplicativa. Sobre esta serie es aplicado el código R A.1 (los datos son leídos mediante la función `scan()`).

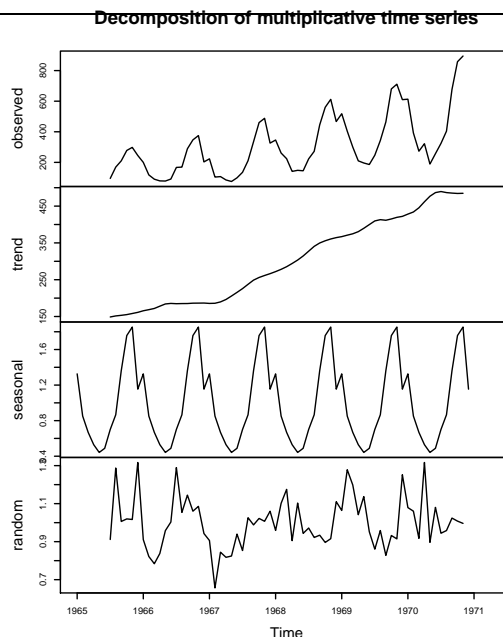


Figura A.1: Descomposición de una serie multiplicativa. Fuente:O.D. Anderson (1976) and O'Donovan (1983). Ventas mensuales compañía X, ene 65 - may 71.

Código R A.1.

```
yt=scan()
154 96 73 49 36 59 95 169 210 278 298 245
200 118 90 79 78 91 167 169 289 347 375 203
223 104 107 85 75 99 135 211 335 460 488 326
346 261 224 141 148 145 223 272 445 560 612 467
518 404 300 210 196 186 247 343 464 680 711 610
613 392 273 322 189 257 324 404 677 858 895 664
628 308 324 248 272

yt=ts(yt,frequency=12,start=c(1965,1))

descom=decompose(yt,type="multiplicative")
descom$seasonal #extrae la serie de tiempo de la estacionalidad estimada
      Jan      Feb      Mar      Apr      May      Jun      Jul
1965 1.3262761 0.8516414 0.6680514 0.5296497 0.4422268 0.4887194 0.7007826
1966 1.3262761 0.8516414 0.6680514 0.5296497 0.4422268 0.4887194 0.7007826
1967 1.3262761 0.8516414 0.6680514 0.5296497 0.4422268 0.4887194 0.7007826
1968 1.3262761 0.8516414 0.6680514 0.5296497 0.4422268 0.4887194 0.7007826
1969 1.3262761 0.8516414 0.6680514 0.5296497 0.4422268 0.4887194 0.7007826
1970 1.3262761 0.8516414 0.6680514 0.5296497 0.4422268 0.4887194 0.7007826
      Aug      Sep      Oct      Nov      Dec
1965 0.8660279 1.3621465 1.7570249 1.8533501 1.1541034
1966 0.8660279 1.3621465 1.7570249 1.8533501 1.1541034
1967 0.8660279 1.3621465 1.7570249 1.8533501 1.1541034
1968 0.8660279 1.3621465 1.7570249 1.8533501 1.1541034
1969 0.8660279 1.3621465 1.7570249 1.8533501 1.1541034
1970 0.8660279 1.3621465 1.7570249 1.8533501 1.1541034

descom$figure #extrae los valores la estimacion de los efectos estacionales
[1] 1.3262761 0.8516414 0.6680514 0.5296497 0.4422268 0.4887194 0.7007826
[8] 0.8660279 1.3621465 1.7570249 1.8533501 1.1541034

descom$trend #extrae la serie de tiempo de la estimacion de T_t
      Jan      Feb      Mar      Apr      May      Jun      Jul      Aug
1965    NA      NA      NA      NA      NA      NA    148.7500 151.5833
1966 165.5833 168.5833 171.8750 178.0417 184.1250 185.5833 184.7917 185.1667
1967 185.5833 186.0000 189.6667 196.2917 205.7083 215.5417 225.7917 237.4583
1968 272.0000 278.2083 285.3333 294.0833 303.4167 314.4583 327.5000 340.6250
1969 367.0833 371.0417 374.7917 380.5833 389.7083 399.7917 409.7083 413.1667
1970 428.2917 434.0417 445.4583 461.7500 476.8333 486.7500 489.6250 486.7500
1971    NA      NA      NA      NA      NA
      Sep      Oct      Nov      Dec
1965 153.2083 155.1667 158.1667 161.2500
1966 185.2917 186.2500 186.3750 186.5833
1967 248.8750 256.0833 261.4583 266.4167
1968 349.7500 355.7917 360.6667 364.3750
1969 411.5417 415.0833 419.4583 422.1250
1970 485.3750 484.4167 484.7917    NA
1971
```

```

descom$random #extrae la serie de tiempo de lo que la funcion llama error
      Jan      Feb      Mar      Apr      May      Jun      Jul
1965    NA      NA      NA      NA      NA      NA      NA 0.9113461
1966 0.9107086 0.8218842 0.7838265 0.8377545 0.9579366 1.0033278 1.2895874
1967 0.9060078 0.6565437 0.8444675 0.8175764 0.8244499 0.9398193 0.8531837
1968 0.9591207 1.1015739 1.1751293 0.9052322 1.1030043 0.9435074 0.9716509
1969 1.0639744 1.2785036 1.1981784 1.0417916 1.1372902 0.9519620 0.8602781
1970 1.0791627 1.0604688 0.9173725 1.3166195 0.8962932 1.0803577 0.9442742
      Aug      Sep      Oct      Nov      Dec
1965 1.2873700 1.0062667 1.0196907 1.0165853 1.3165024
1966 1.0538821 1.1450333 1.0603647 1.0856408 0.9427108
1967 1.0260374 0.9881883 1.0223477 1.0070705 1.0602579
1968 0.9220628 0.9340679 0.8958068 0.9155624 1.1105129
1969 0.9585990 0.8277141 0.9323859 0.9145834 1.2521145
1970 0.9583928 1.0239705 1.0080691 0.9961172      NA

descom$type #arroja la siguiente información
[1] "multiplicative"

#Graficando la serie original y su descomposición, en un panel grafico de cuatro filas
win.graph(width=4,height=5,pointsize=8)
plot(descom)

```

Bibliografía

- Campos Santillán, T. (2001). *Problemario de Pronósticos para la Toma de Decisiones*. Thomson Learning, México.
- Diebold, F. (2001). *Elementos de Pronósticos*. Thomson Learning, México.
- Gaynor, P. E and Kirkpatrick, R. C. (1994). *Introduction to Time-Series Modeling and Forecasting in Business and Economics*. McGrawh-Hill, Inc. New York.
- Makridakis, Wheelwright and McGee, (1983). *Forecasting: Methods and Applications*. Wiley.