3 Modelación clásica

Es a finales del siglo XVII que se tienen los primeros indicios de estudios a conductas sociales, movimientos demográficos, enfermedades, así como tasas de nacimientos a partir de registros históricos. Se sentaron algunas de las bases de la metodología que está vigente en la actualidad. No es hasta el siglo XIX cuando se realizan los primeros pronósticos para enfermedades ajustando curvas a series de tiempo y extrapolando valores. Ya en el siglo XX se proponen las ideas de la modelación clásica en el tratamiento de series de tiempo. En un principio se desarrollo este método para el análisis de series económicas y al ir perfeccionándose, el método se empezó a aplicar en diversos ámbitos como en la salud, transportes, contaminación, etc., pudiéndose aplicar a la mayoría de series de tiempo de diversos campos de conocimiento (Aguirre 1994).

3.1 Enfoque determinístico

La principal característica de la modelación clásica es el enfoque determinístico con que se estudia y analiza una serie de tiempo.

El enfoque determinístico se refiere a que cualquier acontecimiento, responde a una causa, y así, una vez dada la causa, el acontecimiento ha de seguir sin posible variación y con ello supone la existencia de leyes físicas estrictas que rigen el comportamiento de las cosas (Rodríguez 2000).

Por medio de dichas leyes es posible predecir la totalidad de los acontecimientos futuros, esto sin aceptar cambios o variaciones, ya que no se acepta la existencia de hechos aleatorios o caóticos en la realidad. Estos últimos se consideran como errores (por fallas humanas y/o imperfecciones de los instrumentos de medición), o simplemente como una parte de la serie imposible de estudiar y modelar.

Específicamente en una serie de tiempo, el enfoque determinístico dicta que la variable observada presenta un patrón de comportamiento fijo, con lo cual las irregularidades de la serie se contemplan como una desviación respecto a una pauta de comportamiento sistemática (Rodríguez 2000).

Es importante conocer el enfoque determinístico para comprender cabalmente la metodología expuesta en las siguientes páginas y así identificar fácilmente las ventajas y desventajas que se derivan de ella.

La modelación clásica bajo el enfoque determinista, parte de considerar a la serie como un proceso dependiente del tiempo el cual rige el comportamiento de la variable en la serie. En otros términos, el modelo buscado para representar al proceso que gobierna el comportamiento de la serie es una función del tiempo (Aguirre 1994).

Una función del tiempo es una relación en la cual entra una unidad de tiempo y de la cual sale el valor correspondiente de la serie, con un cierto margen de error.

Esquemáticamente se puede representar la idea de función del tiempo de la siguiente manera:

$$t \longrightarrow \boxed{f} \longrightarrow Y_t \pm \varepsilon_t$$

Donde Y_t representa a los valores observados de la variable Y en el momento t y ε_t representa, según este enfoque determinista, un margen de error producido por la existencia en la serie de una perturbación aleatoria imposible de modelar.

Como se ha mencionado previamente, esta perturbación se atribuye a imperfecciones en la observación, recopilación y transmisión de la información obtenida, todas ellas presentes en cualquier sistema de medición. Como se observa en el esquema, la variable t es la variable independiente y la variable resultante Y_t es a su vez la variable dependiente. De forma abreviada se puede representar el esquema con la notación:

$$Y_t = f(t)$$

Con este concepto en mente, podemos decir que modelación clásica se centra en la búsqueda de la expresión analítica del proceso que gobierna el comportamiento observado en la serie.

Dicha función $Y_t = f(t)$ puede construirse por la composición de cuatro elementos que pueden aparecer como patrones básicos en una serie. A estos patrones o regularidades también se les suele denominar componentes, descriptores o determinantes. El procedimiento de análisis se suele llamar *descomposición* (Rodríguez 2000).

Las cuatro componentes de una serie son (Uriel 1995):

Tendencia (T_t). Movimiento ó dirección general de la variable en periodos prolongados de tiempo. *Figura 3.1*

Estacionalidad (E_t). Fluctuaciones periódicas de la variable, más o menos regulares, en periodos relativamente cortos de tiempo, con una oscilación repetitiva para lapsos de tiempo contiguos. *Figura 3.2*

Ciclo (C_t). Movimientos de la variable similares a la estacionalidad, pero relativos a periodos de tiempo mucho más prolongados. Sólo se detecta en series suficientemente largas. *Figura 3.3*

Aleatoriedad (a_t). Es el movimiento irregular de la variable, determinado por el azar, impredecible de forma determinística. A esta componente se le suele llamar residuo aleatorio o ruido. *Figura 3.4*

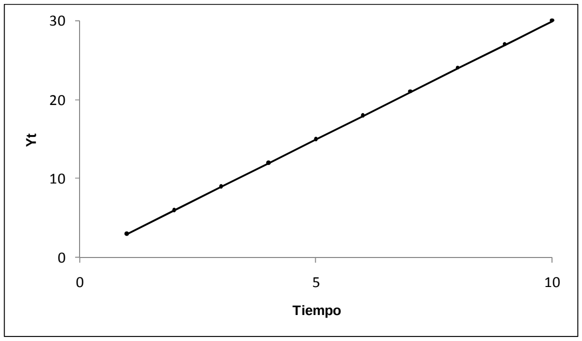


Figura 3.1 Tendencia

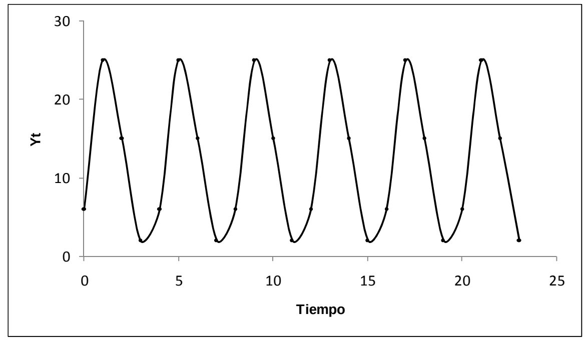


Figura 3.2 Estacionalidad

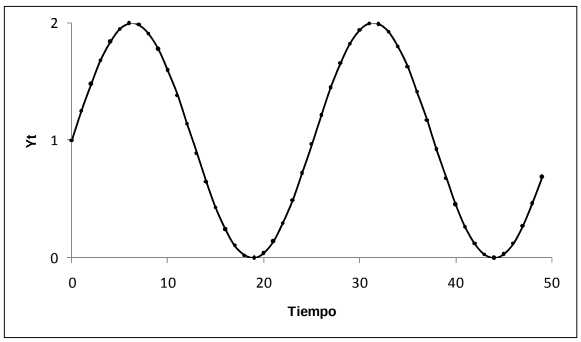


Figura 3.3 Ciclo

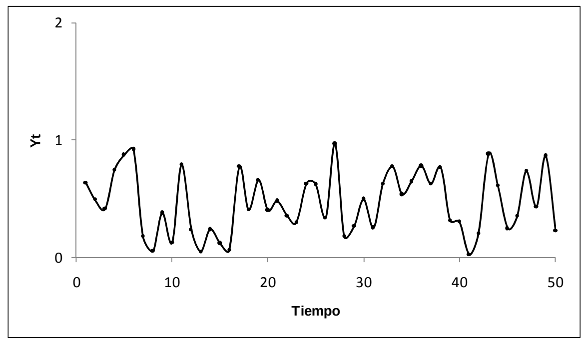


Figura 3.4 Aleatoriedad

3.2 Esquemas de integración

Cada componente detectada en la serie es modelada por separado y una vez estimadas las expresiones analíticas más apropiadas para cada una, el modelo que engloba a cada componente se construye como una integración o composición de los modelos individuales.

Los esquemas de integración tienen dos objetivos principales:

- Ayudar a comprender el comportamiento del fenómeno mostrando de manera simple, es decir, el crecimiento o declinación de la serie
- Producir pronósticos de corto y largo plazo de valores futuros de una serie de tiempo

La integración de las componentes de la serie puede ser mediante (Rodríguez 2000):

1. Esquema aditivo. Los modelos de cada componente se suman entre sí, para formar la función f como $T_t + E_t + C_t + a_t$

2. Esquema multiplicativo. Los modelos de cada componente se multiplican entre sí, para formar la función f como $T_t \times E_t \times C_t \times a_t$

3. Esquema mixto. Surge a partir de la combinación de las dos anteriores. Se forma la función f como $(T_t \times E_t \times C_t) + a_t$

El esquema multiplicativo es el más utilizado en la práctica y significa que el valor de la serie en cualquier momento de tiempo, viene dado por el producto de los valores correspondientes a las cuatro componentes. Este esquema suele aplicarse cuando los valores de la serie muestra oscilaciones cada vez mayores o menores, es decir, la variabilidad de la serie no se mantiene constante.

Por el contrario, cuando la variabilidad de la serie permanece constante, un esquema aditivo da mejores resultados.

El caso mixto se utiliza cuando las oscilaciones estacionales tienden a crecer con el tiempo, y la variabilidad de estas oscilaciones se espera sean constantes.

En la mayoría de las series de tiempo, es difícil manejar los componentes cíclicos de las series de tiempo, por eso dentro de la modelación clásica se considera de poca relevancia esta componente. Además no se cuenta con las herramientas necesarias para la modelación de la componente aleatoria (Aguirre 1994). Por estas razones se estudian con mayor detalle la tendencia y estacionalidad.

3.3 Componentes de una serie de tiempo

3.3.1 Tendencia

La tendencia es la componente de la serie que describe la dirección general hacia la cual se extreman sus valores, como movimiento de dirección global (Aguirre 1994).

En una serie de tiempo no siempre se encuentran las cuatro componentes presente, por esta razón el primer paso para el tratamiento de la tendencia es identificar su presencia. Una vez detectada, el segundo paso en su tratamiento es la modelación mediante el ajuste de una ecuación que relacione los valores de la serie y al tiempo.

La identificación de la tendencia principalmente se hace realizando una inspección visual al gráfico de la serie con el fin de detectar inclinaciones o elevaciones de los valores de la serie que revelarán la presencia de esta componente en la serie. En la mayoría de los casos, la simple inspección visual confirma la existencia de tendencia en la serie.

Si en una serie la identificación no fuera posible mediante la exploración del gráfico o simplemente exista duda de la presencia de tendencia, existen métodos de suavizamiento que ayudan a ver de forma clara la existencia de algún patrón de tendencia en los datos. Dichos métodos son llamados: *suavizamiento por media móvil, exponencial de Brown, exponencial de Holt y estacional de Winter*. Los últimos tres diseñados con fines de pronóstico, pero pueden utilizarse para detectar tendencia (Aguirre 2004). Sólo se estudia el suavizamiento por medias móviles, por ser el de más difusión y cuyos resultados despejan la duda de la presencia de la tendencia en una serie de tiempo.

Suavizamiento por medias móviles: La media móvil se aplica a una serie marcadamente irregular para detectar su tendencia, determinando la forma de la tendencia. La media móvil es un promedio de los valores de la serie de datos desplazándose de un extremo a otro de la misma y sustituyendo esos valores por el resultado promedio (Hanke y Wichern 2006).

La cantidad de valores de la variable a considerar en la media móvil y los coeficientes utilizados para ponderar esos valores, definen la clase de media móvil empleada para el suavizamiento de la serie.

Una media móvil simple de tres puntos se obtiene mediante:

$$Y_t = \frac{Y_{t-1} + Y_t + Y_{t+1}}{3}$$

Cuando la cantidad de puntos en la media móvil es par, se toma la mitad de los valores extremos. Esto se hace para centrar el valor Y_t que será sustituido por la media móvil en cada momento. Así, la media móvil de cuatro puntos es:

$$Y_t = \frac{1/2(Y_{t-2}) + Y_{t-1} + Y_t + Y_{t+1} + 1/2(Y_{t+2})}{4}$$

Como ejemplo de la aplicación de este método, se hace referencia a la serie de datos que se muestra en la *Tabla 3.1* de la cual se obtendrá la media móvil de 4 puntos. Su gráfico correspondiente se muestra en la *Figura 3.5*

t	Y_t	t	Y_t	t	Y_t	t	Y_t
1	500	8	300	15	250	22	500
2	350	9	350	16	550	23	400
3	250	10	200	17	550	24	650
4	400	11	150	18	400	25	850
5	450	12	400	19	350	26	600
6	350	13	550	20	600	27	450
7	200	14	350	21	750	28	700

Tabla 3.1 Serie de tiempo

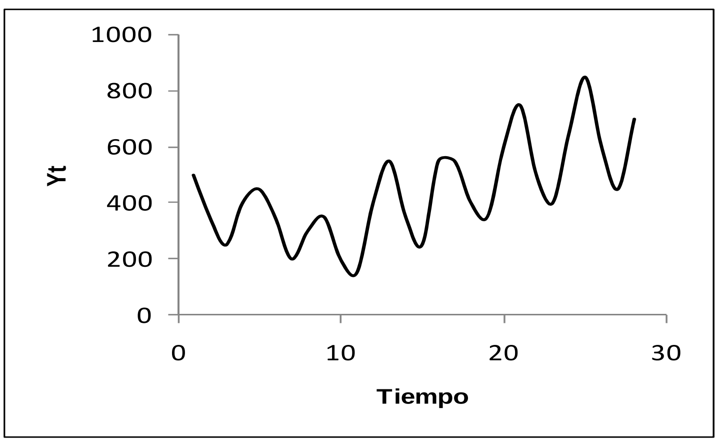


Figura 3.5 Gráfico de la serie

La media móvil de 4 puntos se puede calcular a partir de Y_3:

$$Y_3 = \frac{1/2(Y_{3-2}) + Y_{3-1} + Y_3 + Y_{3+1} + 1/2(Y_{3+2})}{4}$$

$$Y_3 = \frac{1/2(Y_1) + Y_2 + Y_3 + Y_4 + 1/2(Y_5)}{4}$$

$$Y_3 = \frac{1/2(500\ ) + 350 + 250 + 400 + 1/2(450)}{4}$$

$$Y_3 = 368.75$$

Aplicando el mismo procedimiento, se obtienen las medias móviles restantes. Los resultados se muestran en la *Tabla 3.2* Gráficamente el resultado se presenta en la *Figura 3.6*.

t	Y_t	Y'_t	t	Y_t	Y'_t	t	Y_t	Y'_t	t	Y_t	Y'_t
1	500		8	300	281.3	15	250	425.0	22	500	568.8
2	350		9	350	256.3	16	550	431.3	23	400	587.5
3	250	368.8	10	200	262.5	17	550	450.0	24	650	612.5
4	400	362.5	11	150	300.0	18	400	468.8	25	850	631.3
5	450	356.3	12	400	343.8	19	350	500.0	26	600	643.8
6	350	337.5	13	550	375.0	20	600	537.5	27	450	
7	200	312.5	14	350	406.3	21	750	556.3	28	700	

Tabla 3.2 Obtención media móvil de 4 puntos

Como se puede observar en la gráfica, la media o promedio móvil de 4 puntos muestra un incremento gradual en los valores de la serie. Mediante este procedimiento comprobamos la presencia de la componente tendencial en la serie.

Una vez que se ha detectado la existencia de tendencia en la serie y nos hemos hecho una idea de su forma, es necesario obtener un modelo matemático para representar este componente. El método más usual para la modelación de la tendencia consiste en ajustar a la serie a una determinada curva geométrica cuya expresión algebraica sea conocida. Los modelos algebraicos más conocidos y utilizados para modelar la tendencia son los modelos lineales, cuadráticos, exponenciales, potenciales y logarítmicos.

Se ahondará sólo en los modelos lineales y cuadráticos por ser los más sencillos de obtener y que con ellos se obtienen buenos resultados.

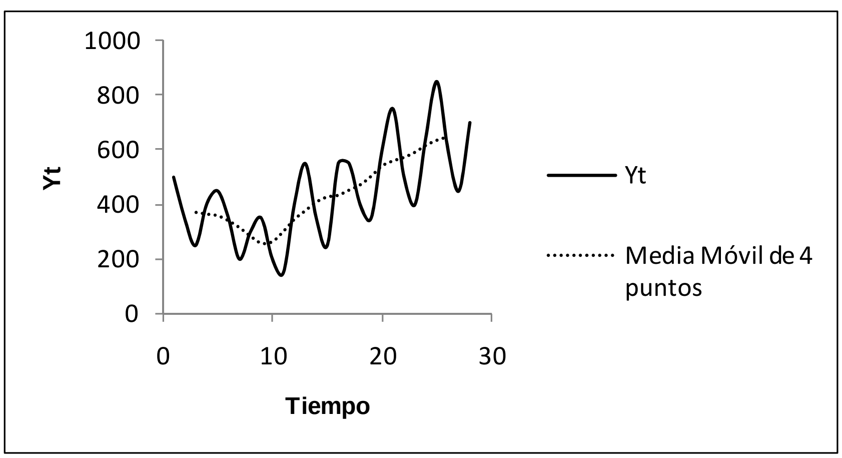


Figura 3.6 Suavizado con media móvil de 4 puntos

Modelo lineal: El modelo lineal de la tendencia representa a una línea recta como se muestra en la *Figura 3.7*. Esta línea puede ser ascendente o descendente. Es muy buena aproximación cuando la serie muestra una curvatura pequeña.

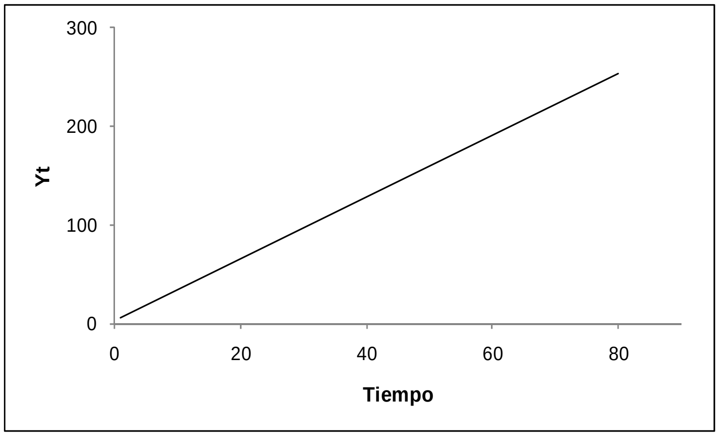


Figura 3.7 Modelo lineal

La forma general de este modelo es:

$$T_t = a + bt$$

Donde

$$T_t\text{: } es\ el\ valor\ predicho\ para\ la\ tendencia\ en\ el\ tiempo\ t$$

$$t\text{: } es\ la\ variable\ independiente\ que\ representa\ el\ tiempo\ y\ supone\ valores\ enteros\ (1, 2, 3, etc.)$$

$$a\text{: } representa\ la\ intersección\ con\ el\ eje\ \ T_t$$

$$b\text{: } representa\ el\ incremento\ promedio\ o\ decremento\ en\ T_t\ para\ cada\ incremento\ de\ tiempo.$$

Por medio del método de mínimos cuadrados puede deducirse la ecuación del modelo lineal que represente a la tendencia. Este método selecciona los valores de los coeficientes en la ecuación de la tendencia (a y b), de forma que los valores de la tendencia estimada T_t son los más cercanos a los valores reales de Y_t.

La proximidad se mide mediante el criterio de la suma del cuadrado de los errores residuales SSE (Bowerman, O'Conell, Koehler 2007):

$$SSE = \sum (Y_t - T_t)^2$$

La línea que mejor se ajusta a un conjunto de datos T_t, t es aquella que minimiza la suma del cuadrado de las distancias desde los puntos hasta la línea, midiéndola en dirección vertical ó T_t. Se conoce como recta de mínimos cuadrados o recta de regresión ajustada.

Así, el método de mínimos cuadrados elige los valores para a y b para reducir el error de la suma de los cuadrados:

$$SSE = \sum (Y_t - T_t)^2 = \sum (Y_t - a - bt)^2$$

Usando un poco de cálculo diferencial, se pueden hacer derivaciones sobre expresiones algebraicas específicas para obtener los valores de mínimos cuadrados.

Específicamente:

$$b = \frac{\sum\left(t - \bar{t}\right)(Y_t - \overline{Y_t})}{\sum(t - \bar{t})^2}$$

$$a = \overline{Y_t} - b\bar{t}$$

El ajuste lineal de la tendencia es el método más usado debido a lo práctico que resulta su manejo.

No obstante, existen modelos no lineales para el ajuste de la tendencia que permiten un mejor ajuste, pero que a cambio su uso resulta en ocasiones no muy práctico o difícil de realizar. Para este tipo de modelos es aconsejable el uso de programas de cómputo que realicen el ajuste correspondiente.

Modelo cuadrático de la tendencia: El modelo cuadrático se utiliza cuando la tendencia se asemeja a un movimiento curvilíneo. Este modelo se representa en la *Figura 3.8* El modelo cuadrático general es:

$$T_t = a + bt + ct^2$$

Modelo exponencial de la tendencia: El modelo exponencial caracteriza a la tendencia con curvas similares a la cuadrática, pero de aumento o disminución más fuerte (*Figura 3.9*). La expresión general es:

$$T_t = ab^t$$

Modelo logarítmico de la tendencia: Este modelo se caracteriza por iniciar con curvatura más pronunciada que la exponencial, pero más suave en su crecimiento posterior (*Figura 3.10*). Su forma general es:

$$T_t = a + b \log t$$

Para los modelos no lineales, al igual que para el modelo lineal, la manera de obtener una expresión matemática es mediante el método de mínimos cuadrados.

El juicio y la experiencia podrían desempeñar una función significativa en la selección y uso de una curva de tendencia. Para utilizar una curva de tendencia para pronosticar, se debe tener la capacidad de argumentar que se ha seleccionado la tendencia adecuada y que, con toda probabilidad, el futuro será como el pasado.

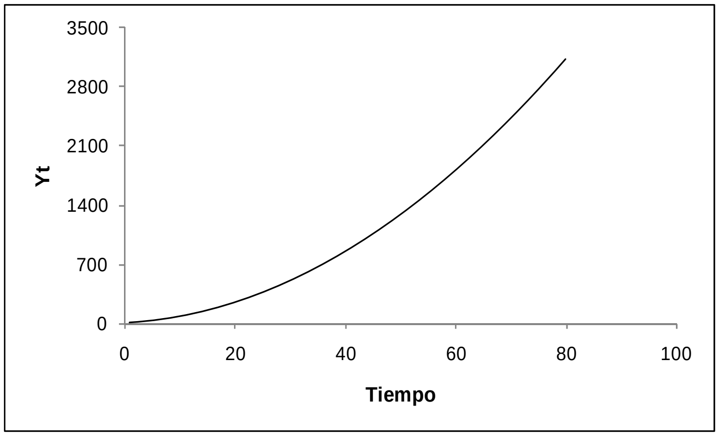


Figura 3.8 Modelo cuadrático

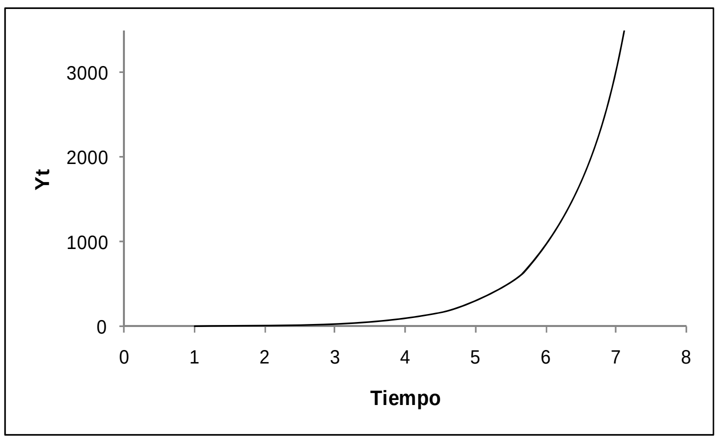


Figura 3.9 Modelo exponencial

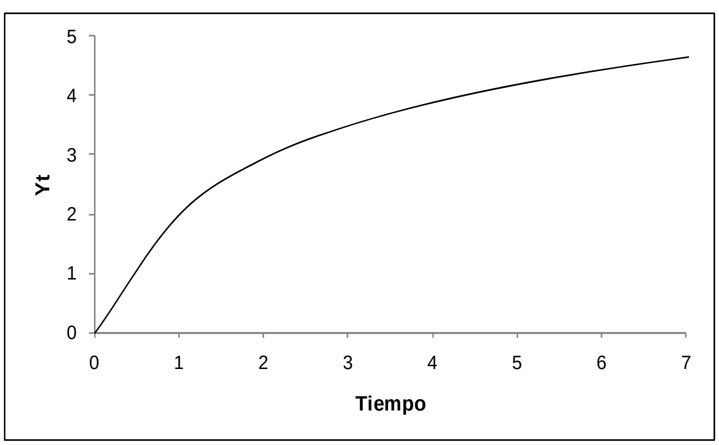


Figura 3.10 Modelo logarítmico

3.3.2 Estacionalidad

Es aquella componente que se caracteriza por una oscilación o fluctuación de tipo local en las observaciones de la serie de tiempo en periodos relativamente cortos de tiempo, manifestada con cierta regularidad. Estos efectos en una serie de tiempo se deben a factores recurrentes cada determinado tiempo (Aguirre 1994).

Al igual que en el tratamiento de la tendencia, el primer paso a seguir es identificar la presencia de esta componente, para luego encontrar un modelo matemático que la represente adecuadamente.

La determinación de existencia de estacionalidad en una serie va ligada a definir el lapso de tiempo en el cual se produce este movimiento regular y reincidente de la variable. A este lapso de tiempo se le conoce con el nombre de *periodo estacional*.

Al igual que en la detección de la tendencia, identificar la estacionalidad en una serie comienza con la inspección visual de su gráfico. Este primer reconocimiento consiste en identificar los grandes movimientos de tipo más general y repetitivos con cierta regularidad.

Por ejemplo, en la *Figura 3.11* se muestran los datos de una serie con marcada estacionalidad. Claramente se aprecian fluctuaciones en los valores de la serie que se repiten constantemente e inclusive aumentan gradualmente al avanzar en el tiempo. Este patrón observado indica la presencia de estacionalidad. Cabe destacar que junto al reconocimiento visual, es importante conocer la naturaleza del proceso estudiado y los factores que de alguna manera pudieran afectar a la variable de manera periódica, como pudiera ser las estaciones del año.

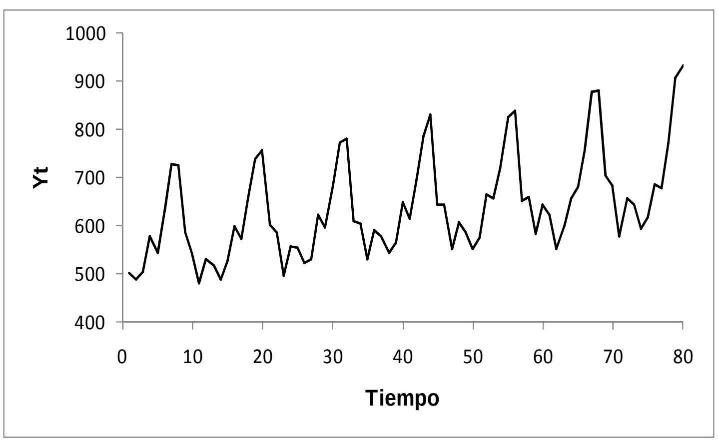


Figura 3.11 Estacionalidad

Ahora bien, aunque en diversas ocasiones es posible distinguir la presencia de estacionalidad en una serie e inclusive determinar su periodo de ocurrencia simplemente observando el gráfico de la serie, en diversas series de tiempo la inspección visual no es prueba definitoria para comprobar la existencia de estacionalidad.

Por ese motivo usualmente se recurre a algunos instrumentos que sirven para comprobar de manera más exacta y confiable la existencia de este patrón. Las más comunes son las funciones de autocorrelación y el periodograma. Las funciones de autocorrelación miden para cada lapso de tiempo, el tipo y grado de correlación existente entre los valores de la serie separados entre sí por un periodo de tiempo igual a este retardo. De acuerdo con esto, si existe estacionalidad en una serie, los valores separados entre sí por lapsos de tiempo iguales al periodo estacional, estos deben estar correlacionados de alguna forma, pues tienden a repetir un similar movimiento en instantes iguales correspondientes a cada periodo estacional (Hanke y Wichern 2006).

Para mostrar cómo la función de autocorrelación simple y la función de autocorrelación parcial sirven como instrumento en la detección confiable de la estacionalidad y determinación de su periodo, supóngase la serie presentada anteriormente en el gráfico de la *Figura 3.11*.

En primera instancia, una rápida inspección visual advirtió la presencia de estacionalidad. Se observan fluctuaciones que se repiten continuamente en el tiempo en un periodo aproximado de cada 12 o 13 datos.

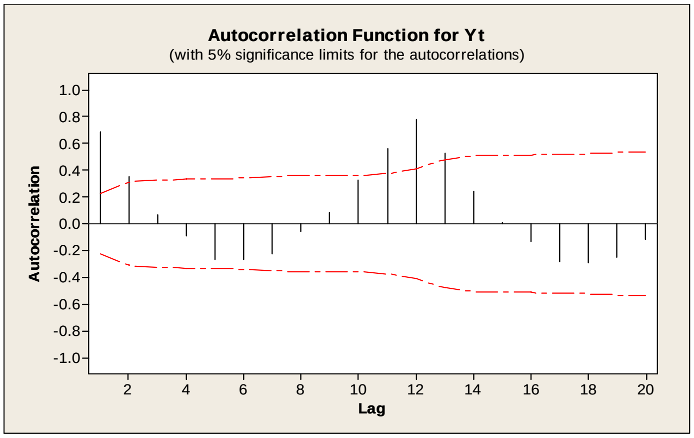
Ahora bien, se obtienen las funciones de autocorrelación simple y parcial. Al prestar especial atención a la *Figura 3.12* que muestra la FAS de la serie, se observa un valor inicial de 0.68 para el primer retardo y posteriormente decaen los valores subsecuentes a negativo hasta el sexto retardo donde vuelven a aumentar hasta el retardo 12 y así sucesivamente. En este análisis se busca detectar una fuerte autocorrelación en el retardo (Lag) 12, lo cual indicaría que los valores separados por 12 retardos están fuertemente correlacionados. Se distingue el coeficiente del retardo 12 respecto a otros debido a su magnitud (0.78) y a que rebasa la línea punteada. Esta espiga es una fuerte evidencia de la existencia de un patrón que se repite constantemente en el tiempo con periodo igual a 12.

Esta conclusión se puede confirmar al analizar la FAP (*Figura 3.13*) donde se aprecia que en el retardo 12, se presenta un valor alto positivo (también llamada espiga). Lo cual confirma una alta relación entre los valores separados por 12 unidades de tiempo. En esa misma figura en el retardo 13 existe un valor importante que resalta de los demás, pero que a diferencia del valor del retardo 12, el retardo 13 es un valor negativo y este no figura de manera significativa en la FAS.

A modo de resumen, se puede concluir que la detección de la estacionalidad en una serie mediante la FAS y la FAP se realiza observando los valores significativamente altos (espigas) en comparación con otros en la misma gráfica (que en la mayoría de las veces resaltan a primera vista) y que se presentan en ambas funciones de autocorrelación simple y parcial. En un amplio número de series de tiempo la estacionalidad puede ser reconocida con gran precisión utilizando esta herramienta y con frecuencia suelen presentar periodos trimestrales, semestrales y anuales.

El uso de la FAS y la FAP en el reconocimiento de la estacionalidad y la determinación de su periodo para el analista puede presentar un poco de dificultad en un principio, pero al transcurso del tiempo, se desarrolla la habilidad necesaria para detectar patrones de este tipo.

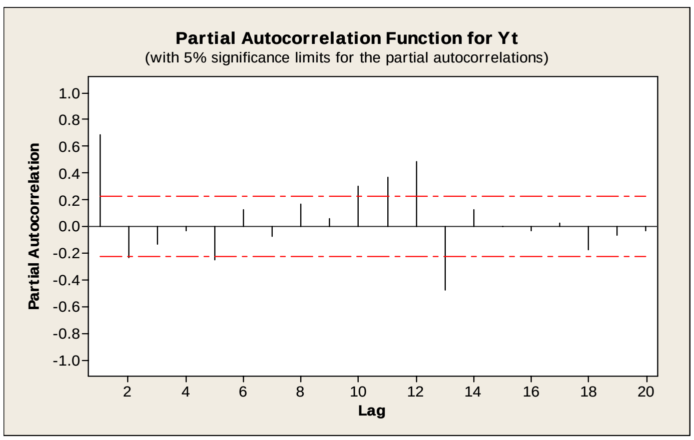
Sin embargo, aún cuando se adquiera cierta habilidad para detectar patrones estacionales, existen series en las cuales aún con las funciones de autocorrelación se hace difícil determinar con precisión la existencia de esta componente y determinar el periodo con que se repiten. A veces también sólo existe incertidumbre acerca de la presencia de estacionalidad y se requiere una herramienta que confirme los resultados que arrojan la FAS y FAP. Frecuentemente se hace uso del periodograma para realizar dicha función. El periodograma es una herramienta que sirve para detectar la presencia de estacionalidad y determinar su periodo con mayor precisión.



Autocorrelation Function: Yt

Lag	ACF	T	LBQ
1	0,687530	6,15	39,25
2	0,350975	2,25	49,61
3	0,070558	0,43	50,04
4	-0,091251	-0,55	50,75
5	-0,265137	-1,59	56,90
6	-0,265132	-1,54	63,13
7	-0,222982	-1,26	67,60
8	-0,061563	-0,34	67,95
9	0,080913	0,45	68,55
10	0,321961	1,78	78,27
11	0,556622	2,96	107,72
12	0,778087	3,75	166,13
13	0,526712	2,18	193,29
14	0,244818	0,96	199,25
15	0,007646	0,03	199,25
16	-0,130096	-0,50	200,99
17	-0,284871	-1,10	209,44
18	-0,289036	-1,10	218,28
19	-0,254200	-0,95	225,23
20	-0,115358	-0,43	226,68

Figura 3.12 Función de autocorrelación simple



Partial Autocorrelation Function: Yt

Lag	PACF	T
1	0,687530	6,15
2	-0,230840	-2,06
3	-0,135700	-1,21
4	-0,030075	-0,27
5	-0,251269	-2,25
6	0,122443	1,10
7	-0,072793	-0,65
8	0,164090	1,47
9	0,057992	0,52
10	0,304054	2,72
11	0,364778	3,26
12	0,482634	4,32
13	-0,477800	-4,27
14	0,127944	1,14
15	0,001305	0,01
16	-0,030195	-0,27
17	0,020971	0,19
18	-0,177264	-1,59
19	-0,066854	-0,60
20	-0,030026	-0,27

Figura 3.13 Función de autocorrelación parcial

Por medio del programa *Degtra* se obtiene el periodograma perteneciente a la serie de ejemplo (*Figura 3.14*).

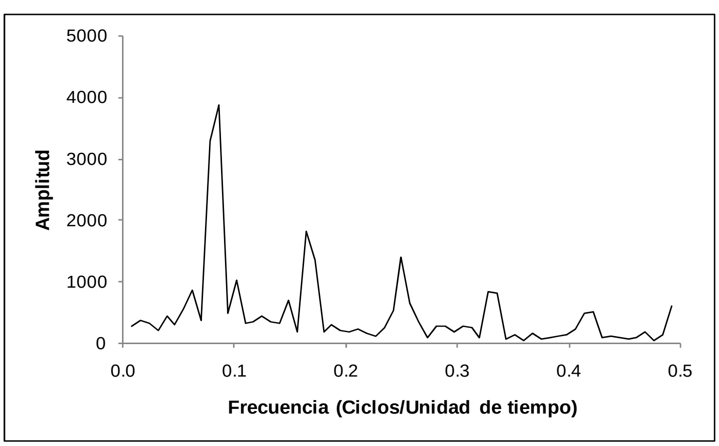


Figura 3.14 Periodograma

De acuerdo al capítulo anterior, la interpretación del periodograma consiste en la localización de amplitudes destacables preferentemente en las primeras frecuencias del gráfico. En este caso se presentan 5 amplitudes que resaltan de las demás. Son de interés las primeras tres para determinar la periodicidad de la serie. Para esta labor normalmente se utiliza la tabla adjunta de datos de salida del programa que elabora el periodograma. Para nuestro periodograma de ejemplo, la *Figura 3.15* muestra la lista de las amplitudes destacables con su respectiva frecuencia.

Frecuencia	Amplitud
0.0833	3864
0.1641	1811
0.2500	1401

Figura 3.15 Amplitudes destacables

Ahora bien, con la relación:

$$p = \frac{1}{f}$$

Se obtienen los respectivos periodos:

$$p_1 = \frac{1}{0.0833} = 12$$

$$p_2 = \frac{1}{0.1641} = 6$$

$$p_3 = \frac{1}{0.2500} = 4$$

Del análisis del periodograma finalmente resalta el periodo igual a 12, el cual confirma que este el correspondiente al patrón estacional de la serie, tal y como se concluyó en la inspección de la serie y con las funciones de autocorrelación. Los periodos 6 y 4 corresponden a submúltiplos del periodo estacional.

Una vez detectada la presencia de estacionalidad en una serie y definido su periodo, el siguiente paso es la obtención del modelo E_t para esta componente. Se han desarrollado diversos métodos para modelar la variación estacional. La idea básica de todos ellos es en primer lugar estimar y eliminar la tendencia de la serie original, dejando únicamente las componentes estacionales y aleatorias. Después, los valores estacionales se recopilan y se resumen para producir un número para cada intervalo observado del año.

La eliminación de la tendencia en una serie se realiza por medio de los esquemas de integración. Si el esquema elegido es aditivo, entonces la tendencia se remueve mediante la ecuación:

$$Y_t - T_t = E_t + a_t$$

Si el esquema elegido es multiplicativo, la remoción de la tendencia es:

$$\frac{Y_t}{T_t} = E_t \times a_t$$

En ambos casos la tendencia se elimina para dejar exclusivamente la componente estacional (si es que la hay) y la componente aleatoria. Así se facilita la detección y modelado de la estacionalidad.

A diferencia de la tendencia que se representa por una curva mejor ajustada, la estacionalidad tiene que ser calculada mediante un conjunto de valores para cada uno de los intervalos observado del año (semana, mes o trimestre) y con frecuencia se presenta en la forma de un número índice. Estos son porcentajes que muestran los cambios a través del tiempo, es decir, muestran cuales periodos del año son relativamente bajos y cuáles son relativamente altos. Los índices estacionales en conjunto describen el patrón estacional.

La variación estacional es expresado mediante *variables ficticias* (Bowerman, O'Conell, Koehler 2007):

$$E_t = \beta_{s1}\,\mathrm{X}_{s1,t} + \beta_{s2}\mathrm{X}_{s1,t} + \cdots + \beta_{s(L-1)}\mathrm{x}_{s(L-1),t}$$

Donde:

L: *es el periodo estacional*

$\beta_{s1}, \beta_{s2}, \beta_{s(L-1)}$: *son los números representativos de cada uno de los intervalos observados*

$$X_{s1,t} = \begin{cases} 1 & \text{si el periodo t es la estación 1} \\ 0 & \text{si no es así} \end{cases}$$

$$X_{s2,t} = \begin{cases} 1 & \text{si el periodo t es la estación 2} \\ 0 & \text{si no es así} \end{cases}$$

$$X_{s(L-1),t} = \begin{cases} 1 & \text{si el periodo t es la estación L} - 1 \\ 0 & \text{si no es así} \end{cases}$$

El objetivo de la variable ficticia es asegurar que se incluye en el modelo, un parámetro estacional apropiado para cada periodo. Usualmente los parámetros estacionales o números índice correspondientes a cada intervalo observado se obtienen por medio de programas de computadora tales como SAS y MINITAB. Este trabajo de tesis se limita a la identificación de la estacionalidad y obtención del periodo estacional, elementos importantes en la modelación ARIMA que se estudiará en el siguiente capítulo. Si se desea conocer más acerca de la modelación estacional, puede recurrirse al libro de Bowerman, O'Conell y Koehler (2007) y al de Uriel Jiménez Ezequiel (1995).

3.3.3 Ciclo

El ciclo es una componente similar a la estacionalidad, pero manifestada en periodos considerablemente más extensos de tiempo. El grado en que pueden determinarse los ciclos a partir de datos históricos, tanto sus longitudes, como sus magnitudes (diferencias entre altos y bajos), están lejos de ser constantes. Esta carencia de un patrón de oleaje constante, hace difícil que los ciclos se distingan de las tendencias que evolucionan con suavidad. En general los ciclos se presentan en periodos que van desde año y medio hasta los diez años.

Igualmente que en la estacionalidad, el ciclo se detecta por medio de las funciones de autocorrelación y el periodograma. Por una parte se busca que en retardos altos existan coeficientes de correlación significativos (espigas) tanto en la FAS como en la FAP. Por otro lado en el periodograma se analiza si existen amplitudes destacables en frecuencias bajas.

Sin embargo, cuando existe ciclo en una serie de tiempo pero no contiene una cantidad suficiente de valores observados, su detección puede llegar a confundirse con la tendencia y ser modelada junto a ella. Además, la modelación se torna también difícil o imposible. En consecuencia, para mantener las cosas relativamente simples, se supone que cualquier ciclo en los datos, es parte de la tendencia.

En series suficientemente largas donde se ha detectado el patrón cíclico, la modelación de esta componente se realiza comúnmente con la serie de Fourier.

3.3.4 Aleatoriedad

Son oscilaciones no sistemáticas que en general afectan a la serie en el momento en que ocurren y normalmente tienen una estructura puramente aleatoria. En la modelación clásica esta componente se considera de poca relevancia ya que la considera como un mero factor errático, de igual importancia a los errores de observación, recopilación y transmisión de la información obtenida, presentes en cualquier sistema de medición. Es decir, considera la aleatoriedad como hechos aislados que no repercutirán en el comportamiento futuro de una variable.

No es posible dar un tratamiento a esta componente debido a que no se dispone de las herramientas necesarias para su modelación, además de que la aleatoriedad no contiene información de cara a la predicción. Es por ello que en la modelación clásica una serie de tiempo se concibe en términos globales como la integración de una parte sistémica que pretendemos modelar y otra aleatoria.

3.4 Ventajas y desventajas modelación clásica

A pesar de que el enfoque clásico puede resolver satisfactoriamente un grupo de problemas bastante amplio relativo al tratamiento de series de tiempo, tanto para finalidades descriptivas como de pronóstico, ésta ha caído en desuso, debido al menor grado de desarrollo de las técnicas de modelación en comparación con otros enfoques disponibles en la actualidad.

La modelación clásica se ve seriamente limitada en series cuyas componentes son difíciles de detectar y por tanto modelar adecuadamente con las herramientas desarrolladas para tales propósitos. Específicamente, en ocasiones se torna complicado detectar el patrón estacional, principalmente cuando su periodo es irregular e inconstante. Caso similar sucede con la componente cíclica, donde en primera lugar se depende de la suficiencia de datos contenida en la serie, seguido de la común irregularidad de su periodicidad. Finalmente, la modelación clásica se ve sumamente limitada en series donde existe una fuerte cantidad de ruido (aleatoriedad).

Otra desventaja es la forma en que se ajustan los datos. En la modelación clásica se ajusta la serie a una forma predeterminada y a un modelo matemático preestablecido (modelo lineal, cuadrático, logarítmico, etc.). Contrario a ello, métodos actuales ajustan un modelo, proveniente de una clase general, a la serie de tiempo, es decir, se crea un modelo específico para la serie de tiempo.

Es por esta razón que las nuevas formas de tratamiento de información (analizada bajo el concepto de series de tiempo), proponen modelos más flexibles y con mejor ajuste a las series de tiempo.

No obstante, en la modelación clásica se establecen las herramientas necesarias para comprender las nuevas metodologías para el manejo de información, que se difunden día a día con mayor auge. Desde la inspección visual, la detección de la tendencia, estacionalidad, ciclo, etc., hasta la determinación de los periodos respectivos, facilitan de gran manera la aplicación de otros enfoques a las series de tiempo.

Particularmente la *modelación ARIMA* que se verá en el capitulo siguiente, se vale de conceptos pertenecientes al enfoque clásico que son fundamentales para el desarrollo de la misma. La experiencia en el reconocimiento de componentes presentes en la serie de tiempo (especialmente en lo que se refiere a la tendencia y la estacionalidad) juega un papel importante en el correcto ajuste y modelación de la serie.