2 Conceptos fundamentales de series de tiempo

2.1 Definición de una serie de tiempo

Una serie de tiempo es el conjunto de mediciones que describen la evolución de un fenómeno o variable a lo largo del tiempo (Pepió 2001).

De manera formal, es una la secuencia cronológicamente ordenada de valores de medición sobre el estado de una variable cuantitativa de un fenómeno o proceso. Dichas mediciones están ordenadas respecto al tiempo y son generalmente dependientes entre sí. Esta dependencia entre las observaciones jugará un papel importante en el análisis de la serie.

Así entonces, una serie de tiempo puede representar desde los precios de un artículo, las tasas de desempleo, la temperatura máxima diaria, la velocidad del viento, hasta los esfuerzos y temperaturas en diversos puntos de una obra civil instrumentada.

Las observaciones de una serie de tiempo pueden ser denotadas como:

$$Y_1,\ Y_2,\ Y_3,\ Y_4, \dots, Y_{t-1}, Y_t$$

Donde:

$$Y_t: es\ el\ valor\ tomado\ por\ el\ proceso\ en\ el\ instante\ t$$

$$t = 1,2,3, \dots, etc.$$

Al referirnos a las series tiempo cabe distinguir entre aquellas cuyas observaciones son tomadas en cada punto de tiempo o series simultáneas y las que no tienen un valor instantáneo en cada punto, dado que son acumulación de valores, como las series temporales acumuladas. Esta distinción no tiene efectos sobre el análisis, puesto el método y la teoría de las series de tiempo se aplican por igual, a ambas modalidades (Arnau 2001).

Los valores de una serie de tiempo se clasifican en continuos o discretos. Los valores continuos son aquellos que se miden con cifras decimales. Por ejemplo: $Y_1 = 12.35,\ Y_2 = 45.67, etc.$ Valores discretos son aquellos que sólo se miden en enteros. Por ejemplo: $Y_1 = 24,\ Y_2 = 56, etc.$ (Pankratz 1983).

Por otro lado los valores de una serie de tiempo se pueden medir en periodos de tiempo continuos o discretos. Una serie medida en periodos de tiempo continuos se presenta cuando los valores de estado de la variable se obtienen de forma permanente o fluida en el tiempo. Ofrece la medida instantánea del estado de la variable observada, teóricamente para momentos de tiempo tan pequeños como se quiera. Por ejemplo: $Y_{1.1},\ Y_{1.2},\ Y_{1.3}, etc.$

En una serie medida en periodos de tiempo discretos, los valores se obtienen para intervalos de tiempo usualmente homogéneos y representa una magnitud acumulada del estado de la variable durante cada uno de esos intervalos. Por ejemplo: $Y_1,\ Y_2,\ Y_3,\ etc.$

Conviene resaltar que en la mayoría de la bibliografía disponible se utiliza el término *serie discreta* para cuando los valores se miden en periodos de tiempo discretos y serie continua cuando los valores se miden en periodos de tiempo continuos.

2.2 Tratamiento de series de tiempo

Como se mencionó en el capitulo anterior, el manejo de la información producto de un monitoreo es de suma importancia para poder conocer el estado que guarda una estructura. Si se aprovecha al máximo esta información, se obtendrán resultados útiles y se podrán alcanzar los objetivos planteados al implementar un sistema de monitoreo estructural.

La manera más común de tratar la información es en forma directa. Esto significa que se analiza la información de manera visual, gráfica y con la obtención de los parámetros estadísticos más representativos tales como la media, varianza, desviación estándar, etc. De esta manera se obtienen estimaciones confiables acerca del estado actual de una estructura. No obstante, la interpretación final depende de la experiencia adquirida del analista a través del tiempo. De esta manera, aunado al gran volumen de datos recopilados, en algunas ocasiones puede ser difícil el análisis e interpretación y se puede inducir con ello a decisiones equivocadas.

Por estas razones se hace necesario y se propone algún tipo de tratamiento que ayude a extraer resultados y conclusiones útiles que complementen el análisis actual de la información, especialmente cuando se tiene gran cantidad de datos.

Las características de la información capturada en un monitoreo estructural permite que su análisis sea bajo el marco de ***series de tiempo***, ya que los parámetros recopilados están asociados a un momento del tiempo determinado, es decir, se conoce el valor de la variable en momento de tiempo específico. Asimismo, estas capturas de información normalmente se hacen separadas equidistantemente en el tiempo, cada hora, cada día, cada semana, etc., según se requiera. Característica que como se verá en próximos capítulos, será importante para el análisis de la serie.

Dentro del marco de series de tiempo, resalta el disponer de información histórica para poder intentar predecir, aunque sea de forma aproximada, un comportamiento futuro que facilite la toma de decisiones cuando pretendemos modificar alguna determinada realidad, sobre todo con el fin de planificar, prever o prevenir.

Es claro que ante una toma de decisión existen diversas alternativas posibles. Sin embargo estas se encuentran en un ambiente de fluctuaciones de forma permanente, se hace presente, por tanto, una gran incertidumbre ante el futuro.

Para reducir precisamente este grado de incertidumbre se suele recurrir a la elaboración de previsiones que tratan de anticipar la evolución de algún fenómeno. Justamente bajo el marco de series temporales, el objetivo principal al analizar las series de tiempo es reducir la incertidumbre sobre el futuro e inclusive en el mismo presente del fenómeno que se desee estudiar. Es decir, el análisis de series de tiempo se puede utilizar también con el único propósito de describir la evolución de cierto fenómeno o conocer el estado actual de cierta variable, que a simple vista resultaría difícil de detectar.

La estadística ha desarrollado teoría y métodos que apuntan a resolver el problema de predicción mediante el uso de series de tiempo. Sin embargo, este no puede ser resuelto por argumentos puramente matemáticos, debe ser el resultado de una combinación matemático-especialista.

Existen distintos tipos de análisis que han surgido a lo largo del tiempo. Dos de ellos resaltan por su eficacia y su amplia divulgación: ***La modelación clásica*** y ***La modelación ARIMA***. Cada uno de ellos trabajan bajo enfoques distintos: el enfoque determinista y el enfoque estocástico, respectivamente. Ambos se detallarán en los capítulos 3 y 4 respectivamente.

2.3 Herramientas básicas para el análisis de series de tiempo

Para entrar a detalle al estudio de series de tiempo es necesario aprender el uso de algunas herramientas que proporcionan información importante acerca de ciertas particularidades presentes en series de datos. Podemos mencionar, entre algunas otras particularidades, la detección patrones ascendentes o descendentes, oscilatorios, periódicos, etc.

Las herramientas que se presentan a continuación sirven tanto para la modelación clásica como para la modelación ARIMA.

2.3.1 Gráfico de la serie

Por muy simple y obvio que pudiera parecer, la exploración gráfica es (o debe ser) el punto de partida para cualquier análisis que se pretenda realizar a una serie de datos. Mediante el gráfico de la serie se puede detectar si los valores siguen alguna regularidad o algún movimiento armónico, también cambios sistemáticos en la media e inclusive cambios en la varianza.

Aunque la exploración del gráfico de una serie no es prueba definitoria para comprobar la existencia de algún tipo de patrón (tendencia, estacionalidad, ciclo, etc.), sí puede proporcionar una idea primaria sobre alguna de estas características.

2.3.2 Funciones de autocorrelación

Una parte fundamental en el estudio de series de tiempo, bajo la modelación clásica y especialmente en la modelación ARIMA (que se estudiará posteriormente) gira en torno a la relación que pueda existir entre los valores de una serie de tiempo. A esta relación o correlación entre los valores de una misma variable se le llama autocorrelación.

La autocorrelación se define como la relación mutua existente entre valores de una serie de tiempo en diferentes periodos y describe lo que tiende a sucederle a un valor si se da un cambio en el otro

Se distinguen dos tipos de autocorrelación que sirven para conocer tanto la existencia de correlación entre los datos de la serie como el grado o intensidad y el tipo de la misma (Bowerman, O'Conell, Koehler 2007).

Autocorrelación simple: Mide la relación lineal entre las observaciones de una serie de datos Y_t, distanciados en un lapso de tiempo k. A este lapso de tiempo k se le conoce como retardo o retraso. Este retardo denota el periodo de tiempo entre los valores de la serie, para el cual se mide el tipo y grado de correlación de la variable considerada.

Cabe aclarar que la autocorrelación simple ofrece una medida de la relación para todos los valores de la serie respecto a los valores de la serie observados k unidades de tiempo antes, comparando la variación conjunta de todos los valores de Y_t en intervalos de k unidades de tiempo, con la variación de la secuencia respecto a esos valores.

Autocorrelación parcial: Es una medida asociada a la autocorrelación simple. De forma sencilla se puede describir la autocorrelación parcial como la estimación de la autocorrelación simple, para el mismo retardo k, con la eliminación del efecto producido por las autocorrelaciones para retardos menores a k, las cuales están presentes en la estimación de la autocorrelación simple. Dicho de otra forma, la autocorrelación parcial no considera las autocorrelaciones acumuladas para el retardo k para el que se estima (Aguirre 1994).

La diferencia específica entre estos dos tipos de autocorrelación se halla en que la autocorrelación simple brinda para un retardo k tanto la relación entre las observaciones con una diferencia de k retardos de tiempo, como la relación para retardos menores, mientras que la autocorrelación parcial brinda sólo la relación para la diferencia estricta en k retrasos de tiempo (Aguirre 1994).

El grado de relación, tanto para la autocorrelación simple y parcial, se mide mediante el coeficiente de autocorrelación, que también se distingue entre simple y parcial. Ambos coeficientes de autocorrelación varían entre +1 y -1.

Estos coeficientes además de detectar la existencia de relación ó correlación, miden el tipo y grado de esta. Por ejemplo, un valor aproximado de +1 implica una fuerte relación entre dos observaciones separadas k unidades de tiempo. El signo positivo (+), denota una relación directa entre las dos observaciones, es decir, que cuando el valor de alguna de ellas aumenta, el otro también tenderá a aumentar; o viceversa, si disminuye un valor, el otro tenderá a disminuir también. Un coeficiente de autocorrelación cercano a -1 expresa que la relación entre las observaciones es inversa, el aumento del valor de la variable se asocia a la disminución de la otra. Un coeficiente de 0 indica que los dos valores de una misma serie de tiempo no están relacionados (Makridakis y Wheelwright 2000).

Los patrones de datos, que incluyen componentes como la tendencia y estacionalidad, se pueden estudiar si se usa esta herramienta. Las autocorrelaciones proporcionan información importante acerca de la estructura de un conjunto de datos y de sus patrones. Por ejemplo, en un conjunto de datos completamente aleatorios la autocorrelación entre valores sucesivos estará cercana a 0, o será igual a 0, pero los valores de datos de fuerte naturaleza estacional o cíclica estarán sumamente autocorrelacionados (Hanke 2006).

El cálculo del coeficiente de autocorrelación simple (r_k) entre observaciones que se encuentran a k periodos de distancia se realiza por medio de la siguiente expresión (Bowerman, O'Conell, Koehler 2007):

$$r_k = \frac{\sum_{t=1}^{n-k}\left(Y_t - \overline{Y}\right)\left(Y_{t+k} - \overline{Y}\right)}{\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2}$$

Donde:

$$r_k: coeficiente\ de\ autocorrelación\ simple\ para\ un\ retraso\ de\ k\ periodos$$

$$\overline{Y}: media\ de\ los\ valores\ de\ la\ serie$$

$$Y_t: observación\ en\ el\ periodo\ t$$

$$Y_{t+k}: observación\ en\ el\ periodo\ con\ k\ retrasos$$

$$n: número\ total\ de\ periodos$$

Una vez conocida la expresión para el cálculo de este coeficiente, es necesario aclarar que en la práctica cuando se tienen datos sin correlación, r_k puede ser diferente de cero, con lo que surge el problema de establecer un intervalo alrededor de $r_k = 0$ que ayude a juzgar si la diferencia a cero del coeficiente r_k es significativa o no.

El intervalo de significación para r_k se construye usualmente con el error típico:

$$S_{r_k} = \begin{cases} \dfrac{1}{\sqrt{n}} & si\ k = 1 \\ \\ \sqrt{\dfrac{1 + 2\sum_{i=1}^{k-1} {r_i}^2}{n}} & si\ k = 2{,}3 \ldots \end{cases}$$

En conjunto con este valor, se puede estimar la estadística t_{r_k}del error típico que esta asociada a la estimación del coeficiente de autocorrelación simple, para juzgar sobre la significación del mismo:

$$t_{r_k} = \frac{r_k}{S_{r_k}}$$

Por el momento no se profundizará acerca del uso de S_{r_k} y t_{r_k}, tan sólo se presentan para referencia futura.

De manera similar, el cálculo del coeficiente de autocorrelación parcial (r_{kk}) entre observaciones que se encuentran a k periodos de distancia se realiza por medio de la siguiente expresión:

$$r_{kk} = \begin{cases} r_1 & si\ \ k = 1 \\ \\ \dfrac{r_k - \sum_{j=1}^{k-1}\ r_{k-1\ j}\ r_{k-j}}{1 - \sum_{j=1}^{k-1}\ r_{k-1\ j}\ r_j} & si\ k = 2{,}3 \ldots \end{cases}$$

Donde:

$$r_{kj} = r_{k-1\ j} - r_{kk} r_{k-1\ k-j} \qquad para\ j = 1{,}2, \ldots, k-1$$

El error estándar de r_{kk} es:

$$S_{r_{kk}} = \frac{1}{\sqrt{n}}$$

La estadística $t_{r_{kk}}$es:

$$t_{r_{kk}} = \frac{r_{kk}}{S_{r_{kk}}}$$

Los coeficientes de autocorrelación, ya sea simple o parcial, que se obtienen para los diferentes periodos $k = 1,2,3 \ldots etc$ se grafican de la manera en que se muestra en la *Figura 2*.1. A esta gráfica se le conoce como correlograma o función de autocorrelación.

El correlograma o función de autocorrelación es una gráfica de las autocorrelaciones para varios retrasos de tiempo de una serie de tiempo.

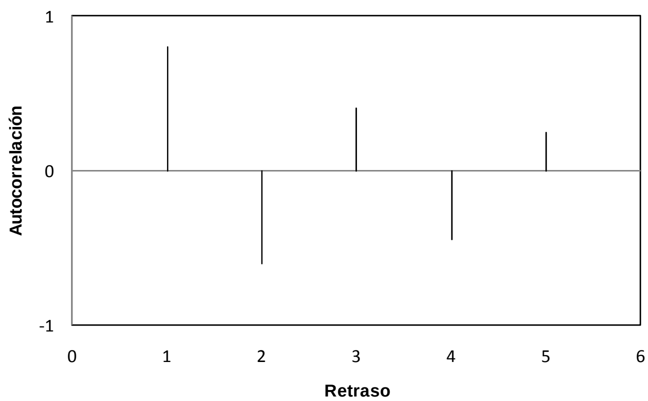

Figura 2.1 Correlograma o función de autocorrelación

En el correlograma de la *Figura 2.1*, la escala horizontal en la parte de debajo de la gráfica muestra cada uno de los retrasos de tiempo que son de interés ($k = 1,2,3 \ldots et$), mientras que la escala vertical del lado izquierdo muestra el posible rango de un coeficiente de autocorrelación de -1 a +1. La línea horizontal a la mitad de la gráfica representa autocorrelaciones de cero.

Para simplificar el uso de estos términos, la *función de autocorrelación simple* se abreviará con las siglas ***FAS*** y a la *función de autocorrelación parcial* como ***FAP***.

Algunos autores como Box y Jenkins sugieren que el número máximo de autocorrelaciones calculadas (tanto simple y parcial) sea de $n/_4$, donde n es el número total de observaciones.

Para ejemplificar el cálculo de los coeficientes de autocorrelación simple y parcial, la obtención de los errores estándar, la estadística t, así como la obtención de la función de autocorrelación (correlograma) se hace referencia a los datos de la *Tabla 2.1* y su correspondiente gráfico que se muestra en la *Figura 2.2*.

Tiempo t	Y_t
1	123
2	130
3	125
4	138
5	145
6	142
7	141
8	146
9	147
10	157
11	150
12	160

Tabla 2.1 Datos serie de tiempo

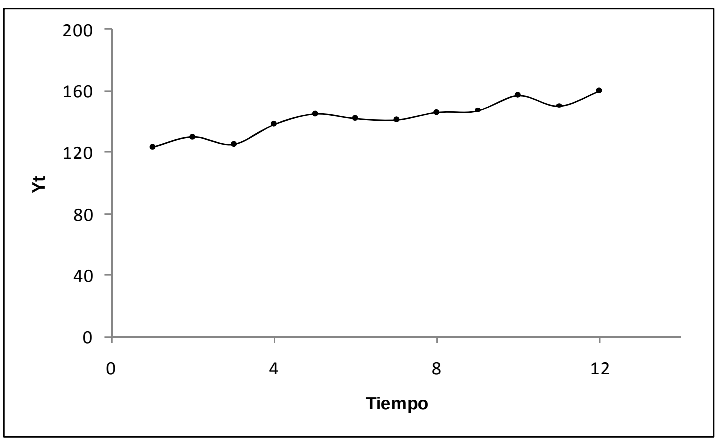

Figura 2.2 Gráfico de la serie de tiempo

Se comenzará por el cálculo de los coeficientes de autocorrelación simple, recordando el número máximo de coeficientes $n/4$. Para la serie de este ejemplo $n = 12\ datos$, así se calcularán tan sólo 3 coeficientes de autocorrelación (r_1, r_2, r_3).

Tomando en cuenta que:

$$r_k = \frac{\sum_{t=1}^{n-k}(Y_t - \overline{Y})(Y_{t+k} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2}$$

Se tiene entonces que para el coeficiente r_1, $n = 12\ y\ k = 1$ y a partir de los datos de la *Tabla 2.1*:

$$r_1 = \frac{\sum_{t=1}^{11}\left(Y_t - \overline{Y}\right)\left(Y_{t+1} - \overline{Y}\right)}{\sum_{t=1}^{12}\left(Y_t - \overline{Y}\right)^2}$$

También:

$$\overline{Y} = \left(\frac{1}{n}\right)\sum_{t=1}^{n} Y_t = \left(\frac{1}{12}\right)1740 = 142$$

Desarrollando tenemos que:

$$r_1 = \frac{\left(Y_1 - \overline{Y}\right)\left(Y_2 - \overline{Y}\right) + \left(Y_2 - \overline{Y}\right)\left(Y_3 - \overline{Y}\right) + \left(Y_3 - \overline{Y}\right)\left(Y_4 - \overline{Y}\right) + \cdots + \left(Y_{11} - \overline{Y}\right)\left(Y_{12} - \overline{Y}\right)}{\left(Y_1 - \overline{Y}\right)^2 + \left(Y_2 - \overline{Y}\right)^2 + \left(Y_3 - \overline{Y}\right)^2 + \cdots + \left(Y_{12} - \overline{Y}\right)^2}$$

Donde:

$$Y_1 = 123$$

$$Y_2 = 130$$

$$Y_3 = 125$$

$$\downarrow$$

$$Y_{12} = 160$$

Sustituyendo:

$$r_1 = \frac{(123 - 142)(130 - 142) + (130 - 142)(125 - 142) + \cdots + (150 - 142)(160 - 142)}{(123 - 142)^2 + (130 - 142)^2 + (125 - 142)^2 + \cdots + (160 - 142)^2}$$

$$r_1 = \frac{843}{1474}$$

$$r_1 = 0.5719$$

El coeficiente de autocorrelación r_2 para el segundo retardo de tiempo se calcula entonces con $n = 12\ y\ k = 2$:

$$r_2 = \frac{\sum_{t=1}^{10}\left(Y_t - \overline{Y}\right)\left(Y_{t+k} - \overline{Y}\right)}{\sum_{t=1}^{12}\left(Y_t - \overline{Y}\right)^2}$$

Así:

$$r_2 = \frac{(Y_1 - \overline{Y})(Y_3 - \overline{Y}) + (Y_2 - \overline{Y})(Y_4 - \overline{Y}) + (Y_3 - \overline{Y})(Y_5 - \overline{Y}) + \cdots + (Y_{10} - \overline{Y})(Y_{12} - \overline{Y})}{(Y_1 - \overline{Y})^2 + (Y_2 - \overline{Y})^2 + (Y_3 - \overline{Y})^2 + \cdots + (Y_{12} - \overline{Y})^2}$$

$$r_2 = \frac{(123 - 142)(125 - 142) + (130 - 142)(138 - 142) + \cdots + (157 - 142)(160 - 142)}{(123 - 142)^2 + (130 - 142)^2 + (125 - 142)^2 + \cdots + (160 - 142)^2}$$

$$r_2 = \frac{682}{1474}$$

$$r_2 = 0.4627$$

De forma similar se procede para el cálculo de r_3, así entonces:

$$r_3 = \frac{(Y_1 - \overline{Y})(Y_4 - \overline{Y}) + (Y_2 - \overline{Y})(Y_5 - \overline{Y}) + (Y_3 - \overline{Y})(Y_6 - \overline{Y}) + \cdots + (Y_9 - \overline{Y})(Y_{12} - \overline{Y})}{(Y_1 - \overline{Y})^2 + (Y_2 - \overline{Y})^2 + (Y_3 - \overline{Y})^2 + \cdots + (Y_{12} - \overline{Y})^2}$$

$$r_3 = \frac{(123 - 142)(138 - 142) + (130 - 142)(145 - 142) + \cdots + (147 - 142)(160 - 142)}{(123 - 142)^2 + (130 - 142)^2 + (125 - 142)^2 + \cdots + (160 - 142)^2}$$

$$r_3 = \frac{163}{1474}$$

$$r_3 = 0.1106$$

Una vez calculados los coeficientes, creamos la función de autocorrelación simple (FAS) o correlograma, como se muestra en la *Figura 2.3*.

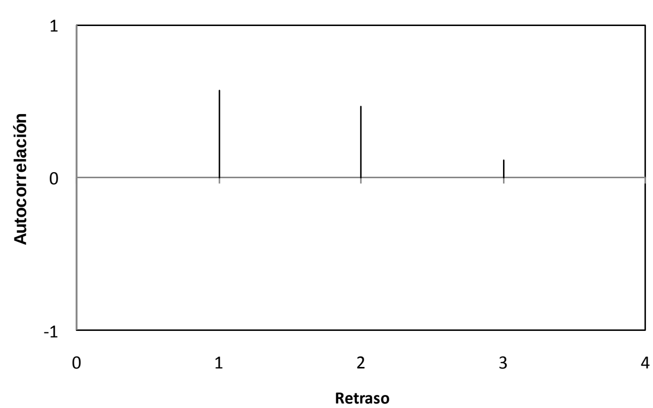

Figura 2.3 Función de autocorrelación simple (FAS)

Ahora bien, a partir de r_1, r_2, r_3, obtenemos sus correspondientes errores estándares y su estadístico t asociado.

Recordando las fórmulas:

$$S_{r_k} = \begin{cases} \dfrac{1}{\sqrt{n}} & si\ k = 1 \\ \\ \sqrt{\dfrac{1 + 2\sum_{i=1}^{k-1} {r_i}^2}{n}} & si\ k = 2{,}3 \ldots \end{cases}$$

$$t_{r_k} = \frac{r_k}{S_{r_k}}$$

Para r_1, $k = 1$ y $n = 12$; entonces:

$$S_{r_1} = \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{12}}$$

$$S_{r_1} = 0.2887$$

$$t_{r_1} = \frac{r_1}{S_{r_1}} = \frac{0.5719}{0.2887}$$

$$t_{r_1} = 1.98$$

Para r_2, $k = 2$ y $n = 12$; entonces:

$$S_{r_2} = \sqrt{\frac{1 + 2\sum_{i=1}^{2-1} {r_i}^2}{12}} = \sqrt{\frac{1 + 2\,{r_1}^2}{12}} = \sqrt{\frac{1 + 2\,(0.5719^2)}{12}}$$

$$S_{r_2} = 0.3713$$

$$t_{r_2} = \frac{r_2}{S_{r_2}} = \frac{0.4627}{0.3713}$$

$$t_{r_1} = 1.25$$

Para r_3, $k = 3$ y $n = 12$; entonces:

$$S_{r_2} = \sqrt{\frac{1 + 2\sum_{i=1}^{3-1} r_i{}^2}{12}} = \sqrt{\frac{1 + 2\,(r_1{}^2 + r_2{}^2)}{12}} = \sqrt{\frac{1 + 2\,(0.5719^2 + 0.4627^2)}{12}}$$

$$S_{r_2} = 0.4166$$

$$t_{r_3} = \frac{r_3}{S_{r_3}} = \frac{0.1106}{0.4166}$$

$$t_{r_1} = 0.27$$

Siguiendo en el ejemplo, se calcularán ahora los coeficientes de autocorrelación parcial $r_{11},\ r_{22},\ r_{33}$, sus respectivos errores y estadísticos t. Se procederá de forma similar a la mostrada en el cálculo de autocorrelaciones simples.

Teniendo en cuenta que:

$$r_{kk} = \begin{cases} r_1 & si\ \ k = 1 \\ \\ \dfrac{r_k - \sum_{j=1}^{k-1}\ r_{k-1\ j}\, r_{k-j}}{1 - \sum_{j=1}^{k-1}\ r_{k-1\ j}\, r_j} & si\ k = 2{,}3\ ... \end{cases}$$

Entonces para $k = 1$:

$$r_{11} = r_1$$

$$r_{11} = 0.5719$$

Para $k = 2$:

$$r_{22} = \frac{r_2 - \sum_{j=1}^{2-1}\ r_{2-1\ j}\, r_{2-j}}{1 - \sum_{j=1}^{2-1}\ r_{2-1\ j}\, r_j}$$

$$r_{22} = \frac{r_2 - r_{1\,1}\, r_1}{1 - r_{11}\, r_1} = \frac{0.4627 - (0.5719)(0.5719)}{1 - (0.5719)(0.5719)}$$

$$r_{22} = 0.2015$$

Para $k = 3$:

$$r_{33} = \frac{r_3 - \sum_{j=1}^{3-1} \; r_{3-1\,j}\, r_{3-j}}{1 - \sum_{j=1}^{3-1} \; r_{3-1\,j}\, r_j}$$

$$r_{33} = \frac{r_3 - (r_{21}\, r_2 + r_{22}\, r_{1)}}{1 - (r_{21}\, r_1 + r_{22}\, r_2)}$$

Donde:

$$r_{kj} = r_{k-1\,j} - r_{kk} r_{k-1\,k-j}$$

$$r_{21} = r_{11} - \; r_{22}\, r_{11}$$

$$r_{21} = 0.5719 - (0.2015)(0.5719)$$

$$r_{21} = 0.4567$$

Entonces:

$$r_{33} = \frac{0.1106 - [(0.4567)(0.4627) + (0.2015)(0.5719)]}{1 - [(0.4567)(0.5719) + (0.2015)(0.4627)]}$$

$$r_{33} = -0.3345$$

La respectiva función de autocorrelación parcial (FAP) se observa en la *Figura 2.4*.

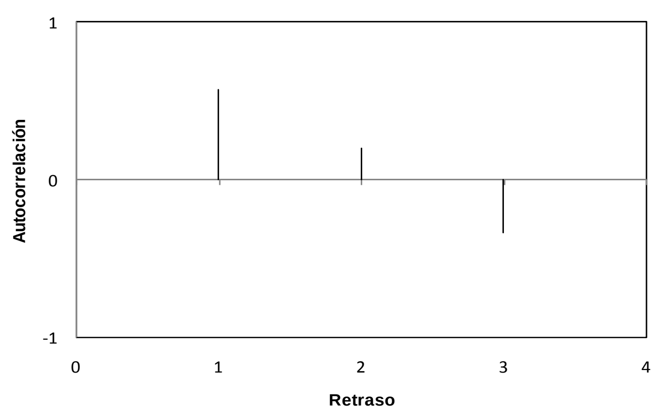

Figura 2.4 Función de autocorrelación parcial (FAP)

Los errores estándar para cada uno de los coeficientes calculados son:

$$S_{r_{kk}} = \frac{1}{\sqrt{n}}$$

$$S_{r_{11}} = S_{r_{22}} = S_{r_{33}} = \frac{1}{\sqrt{12}} = 0.2887$$

La estadística $t_{r_{kk}}$para cada uno de los coeficientes es:

$$t_{r_{kk}} = \frac{r_{kk}}{S_{r_{kk}}}$$

$$t_{r_{11}} = \frac{r_{11}}{S_{r_{11}}} = \frac{0.5719}{0.2887}$$

$$t_{r_{11}} = 1.98$$

$$t_{r_{22}} = \frac{r_{22}}{S_{r_{22}}} = \frac{0.2015}{0.2887}$$

$$t_{r_{22}} = 0.70$$

$$t_{r_{33}} = \frac{r_{33}}{S_{r_{33}}} = \frac{-0.3345}{0.2887}$$

$$t_{r_{33}} = -1.16$$

Para finalizar este ejemplo de aplicación, en la *Figura 2.5* se resumen los cálculos realizados. De acuerdo con los coeficientes de autocorrelación simple se puede inferir (para la serie de ejemplo) que existe una mediana relación de los valores respecto a aquellos distanciados un retardo de tiempo y la relación disminuye para los retardos 2 y 3. Por otra parte los coeficientes de autocorrelación parcial indican una mediana relación directa de los valores con aquellos separados 1 y 2 retrasos de tiempo y una relación inversa con el valor separado 3 retrasos de tiempo.

Coeficiente	Valor	Error estándar	Valor	Estadístico t	Valor
r_1	0.572	S_{r1}	0.289	t_{r1}	1.981
r_2	0.463	S_{r2}	0.371	t_{r2}	1.246
r_3	0.111	S_{r3}	0.417	t_{r3}	0.265
r_{11}	0.572	S_{r11}	0.289	t_{r11}	1.981
r_{22}	0.202	S_{r22}	0.289	t_{r22}	0.698
r_{33}	-0.335	S_{r33}	0.289	t_{r33}	-1.159

Figura 2.5 Cuadro resumen

Un comentario que resulta interesante e importante destacar, producto del ejemplo anterior, es que para una serie con $n = 12\, datos$, la obtención de los coeficientes de autocorrelación simples y parciales puede resultar un trabajo laborioso. Es intuitivo que para una serie de tiempo con más observaciones, la dificultad de esta labor se complique aún más. El problema que ello conduce es que la tarea de analizar una serie de tiempo se vea mermada y se retrase de manera significativa. No es práctico detenerse demasiado tiempo en la obtención de las funciones de autocorrelación simple y parcial (FAS y FAP), debido al hecho de que aunque son una herramienta fundamental para el análisis de series de tiempo, de ella no se obtienen directamente los resultados que se buscan al estudiar una serie de datos. Por ese motivo es recomendable y más conveniente usar programas de cómputo estadístico que ayude en la obtención de dichos coeficientes. Algunos paquetes estadísticos pueden ser MINITAB, SAS, algún complemento para MATLAB o EXCEL, entre otros. Estos paquetes además ofrecen diversas opciones para el análisis y tratamiento de series de tiempo (algunas de estas opciones se mencionarán en otros capítulos del presente texto).

Para introducir el uso de los programas de cómputo al tratamiento de las series de tiempo, se compara la salida del programa MINITAB con los resultados del ejemplo expuesto anteriormente. La salida del programa se muestra en la *Figura 2.6* donde se observa el valor de los coeficientes de autocorrelación simple. En la *Figura 2.7* se muestran los coeficientes de autocorrelación parcial.

Autocorrelation Function: Datos Serie

Lag	ACF	T	LBQ
1	0,571913	1,98	5,00
2	0,462687	1,25	8,59
3	0,110583	0,27	8,82

Figura 2.6 Coeficientes de autocorrelación simple

Partial Autocorrelation Function: Datos Serie

Lag	PACF	T
1	0,571913	1,98
2	0,201514	0,70
3	-0,334512	-1,16

Figura 2.7 Coeficientes de autocorrelación parcial

Podemos observar que la salida de MINITAB contiene, además del cálculo de los coeficientes, el estadístico t, mismo que se obtuvo en el ejemplo de los datos de la *Tabla 2.1*. Se obtienen los mismos resultados que en el ejemplo.

MINITAB también ofrece la posibilidad de obtener las gráficas de la FAS y la FAP, tal como se muestra en las *Figuras 2.8* y *2.9* respectivamente. Las dos bandas de puntos que se observan en la gráfica representan el valor de dos desviaciones estándar para cada retraso. Es decir $2(S_{r_k})$ por arriba y $2(S_{r_k})$ por debajo del eje central. Más adelante analizaremos la ayuda que ofrecen estás bandas.

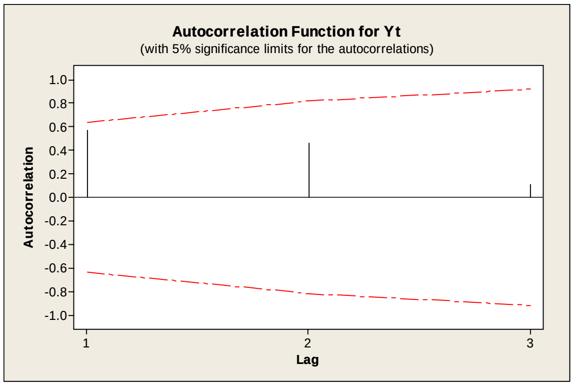

Figura 2.8 Función de autocorrelación simple

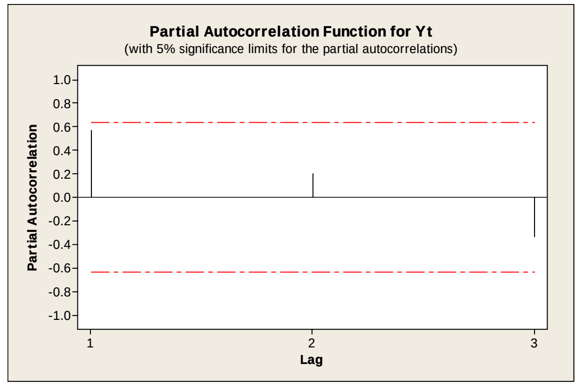

Figura 2.9 Función de autocorrelación parcial

2.3.3 Periodograma

El periodograma es un instrumento que sirve para detectar estacionalidad en una serie y determinar su periodo (Aguirre 1994). Estas cuestiones se estudiaran en los próximos capítulos con más de detalle, por el momento sólo se mencionarán las características básicas de este instrumento.

Para comprender el funcionamiento de esta herramienta y su utilidad, supongamos que una señal en el tiempo presenta movimientos oscilatorios. Así, si establecemos una unidad arbitraria de tiempo, se puede medir la cantidad de ondulaciones o impulsos producidos por la señal durante esa unidad de tiempo. A esta cantidad de ondulaciones o fluctuaciones en la unidad de tiempo se le denomina frecuencia de la señal. Esta frecuencia se mide en ciclos en una unidad de tiempo. Ahora bien, si en vez de medir la cantidad de ciclos o impulsos producidos por la señal en una unidad de tiempo, medimos el tiempo invertido por esa señal en completar un ciclo, estamos obteniendo el periodo de la señal. Como se verá un poco más adelante, es ahí donde hallamos la utilidad de esta herramienta en una serie de tiempo, donde se requiere conocer el periodo estacional; es decir, se desea conocer el tiempo invertido por la variable en realizar un ciclo regular o repetitivo de manera completa. El periodo se mide en unidades de tiempo.

De la definición de frecuencia y periodo se deducen las siguientes relaciones:

$$Periodo = \frac{1}{Frecuencia}$$

y

$$Frecuencia = \frac{1}{Periodo}$$

Asociado a estos términos se incluye otro que resulta de interés: la amplitud. La amplitud en una señal es la distancia desde la base hasta la cima o cresta de un ciclo de esa señal.

El periodograma se construye a partir de transformar la serie de tiempo, llevando la variable del dominio del tiempo (su forma natural) al dominio de frecuencias. Esto se realiza aplicando una transformada a los valores de la serie. Existen diferentes transformadas para la construcción de un periodograma, la más usual y aplicada es la transformada rápida de Fourier. En la *Figura 2.10* se muestra el periodograma para una serie de tiempo. En el presente texto no se detalla el desarrollo de la transformada rápida de Fourier, sólo se mencionan los programas de cómputo *Degtra* (Ordaz y Montoya 2002) y *Scope* que proveen esta herramienta de transformación.

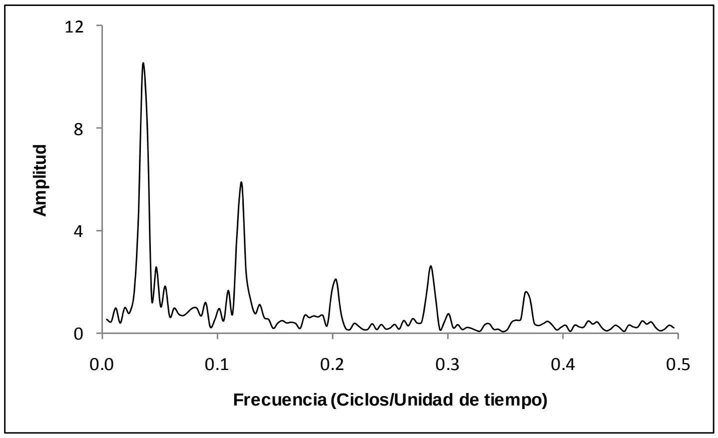

Figura 2.10 Periodograma

Ahora bien, la interpretación de un periodograma se realiza mediante la búsqueda de amplitudes destacables en frecuencias bajas, es decir, picos al inicio del gráfico. Una vez detectadas dichas amplitudes se busca la frecuencia correspondiente y se calcula el periodo correspondiente mediante:

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

De esta manera se obtiene un *periodo estacional posible* de la serie de tiempo. Decir que es un periodo *posible* hace referencia al hecho que el periodograma también puede mostrar amplitudes destacables no sólo en las frecuencias correspondientes a los verdaderos periodos estacionales, sino también en los múltiplos y submúltiplos de dichas frecuencias. Para realizar la discriminación de los verdaderos periodos se debe comparar los resultados con los obtenidos de las funciones de autocorrelación simple y parcial.

Otra consideración es que para una serie extensa, un periodograma puede mostrar varios patrones estacionales y cíclicos. Este último se presenta para frecuencias mucho más bajas que para un patrón estacional (Aguirre 1994).