ESTADÍSTICA APLICADA

MET. Alejandra Cerda
alejandra.cerdarz@uanl.edu.mx

CLASIFICACIÓN

SERIES DE TIEMPO

MET. Alejandra Cerda

SERIES ESTACIONARIAS

Se dice que una serie de tiempo es estacionaria cuando su distribución y sus parámetros no varían con el tiempo; es decir, la media y la varianza de una **serie estacionaria** no cambian con el tiempo, y tampoco siguen una tendencia.

La serie de tiempo puede expresarse como una función de una tendencia, una componente estacional y un error. En esencia, existen dos métodos para aislar la tendencia y la componente estacional de una serie de tiempo:

- Método paramétrico: se basa en proponer modelos paramétricos para expresar la relación que guardan la tendencia y la componente estacional con el tiempo, ajustar dichos modelos a la serie de tiempo y aislar la tendencia y la componente estacional por medio de los modelos ajustados.
- Método no paramétrico: se basa en asumir “suavidad” en la relación que guardan la tendencia y la componente estacional con el tiempo.

¿PROCESOS ESTOCÁSTICOS?

Desde un punto de vista intuitivo, un proceso estocástico se describe como una secuencia de datos que evoluciona con el tiempo. Las series de tiempo se definen como un caso particular de los procesos estocásticos.

Una serie estacionaria es mucho más fácil de predecir. Si se comportaba de una manera en el pasado podremos suponer que tiene una gran probabilidad de continuar comportándose de la misma forma en el futuro. **Procesos estocásticos lineales estacionarios**: Si una serie de tiempo es estacionaria, su media, su varianza y su autocovarianza (en diferentes rezagos o periodos de tiempo) permanecen iguales sin importar el momento en el cual se midan; es decir, son invariantes respecto al tiempo.

MODELOS

SERIES ESTACIONARIAS

MODELO AR(p)

Los **modelos autoregresivos** se basan en la idea de que el valor actual de la serie x_t , puede explicarse en función de los p valores pasados, donde p determina el número de rezagos necesarios para pronosticar un valor actual. De tal modo que un modelo AR(1) implica un proceso donde x_t esta determinado únicamente por el valor pasado x_{t-1}

Expresión matemática: $x_t = \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \emptyset_3 x_{t-3} + \cdots \emptyset_p x_{t-p} + e_t$

donde se supone cada observación como combinación lineal de observaciones pasadas.

MODELO MA(q)

En el método de **promedio móvil s**e puede especificar un número constante de puntos o datos al inicio y se puede calcular una media (o promedio) para las observaciones más recientes. Conforme se tienen nuevas observaciones, se calcula una nueva media al sumar el valor más reciente y al eliminar el más antiguo. Este promedio móvil se usa para pronosticar el siguiente periodo.

MODELO MA(q)

Modelos determinados por una “fuente externa”. Éstos modelos suponen linealidad, el valor actual de la serie x_t , esta influenciado por los valores de la fuente externa. Un promedio móvil se construye sustituyendo cada valor de una serie por la media obtenida con esa observación y algunos de los valores inmediatamente anteriores y posteriores. Los promedios móviles también se pueden construir tomando en cuenta valores adyacentes de las observaciones.

Expresión matemática:

$$x_t = e_t + \beta_1 e_{t-1} + \cdots + \beta_q e_{t-q}$$

MODELO MA(q)

Una estimación mediante promedio móvil de orden K se puede calcular mediante la ecuación:

$$\hat{y}_{t+1} = \frac{y_t + y_{t-1} + y_{t-2} + \cdots + y_{t-k+1}}{k}$$

donde $\hat{y}_{t+1}$ = valor pronosticado para el siguiente periodo

y_t = valor real en el periodo t

k = número de términos en el promedio móvil

El promedio móvil para el periodo t es la media aritmética de las observaciones más recientes k. Sin embargo, no es el único método de estimación ya que se puede hacer promedio de promedios. El método anterior sería un caso simple con l propósito de entender el proceder del modelo MA.

MODELO ARMA(p,q)

Es muy probable que una serie de tiempo x_t tenga características AR y MA a la vez, por consiguiente, sea **ARMA(p,q)**, en este proceso habrá p términos autoregresivos y q términos de media móvil.

Cada observación es combinación lineal de observaciones pasadas y de errores pasados y presentes. Las series temporales presentan variaciones que pueden ser debidas al propio modelo o al medio en el que se tomaron las observaciones o a las dos causas. Para tener en cuenta los factores externos se introducen los modelos estacionales.

Expresión matemática:

$$x_t = \underbrace{\emptyset_1 x_{t-1} + \cdots \emptyset_p x_{t-p}}_{\text{p términos bajo el modelo AR}} + \underbrace{\beta_1 e_{t-1} + \cdots + \beta_q e_{t-q}}_{\text{q términos bajo el modelo MA}} + c$$

CORRELOGRAMAS

SERIES DE TIEMPO

MET. Alejandra Cerda

AUTOCORRELACIÓN

En ocasiones en una serie de tiempo acontece, que los valores que toma una variable en el tiempo no son independientes entre sí, sino que un valor predeterminado depende de los valores anteriores; existen dos formas de medir esta dependencia:

Función de Autocorrelacion ACF

La función de autocorrelación mide la correlación entre dos variables separadas por k periodos, se define como:

$$\rho_j = corr\left(x_j, x_{j-k}\right) = \frac{cov(x_j, x_{j-k})}{\sqrt{V(x_j)}\sqrt{V(x_{j-k})}}$$

Con propiedades:

- $\rho_0 = 1$
- $-1 \leq \rho_j \leq 1$
- $simetria\ \rho_j = \rho_{-j}$

AUTOCORRELACIÓN PARCIAL

Función de Autocorrelación Parcial PACF

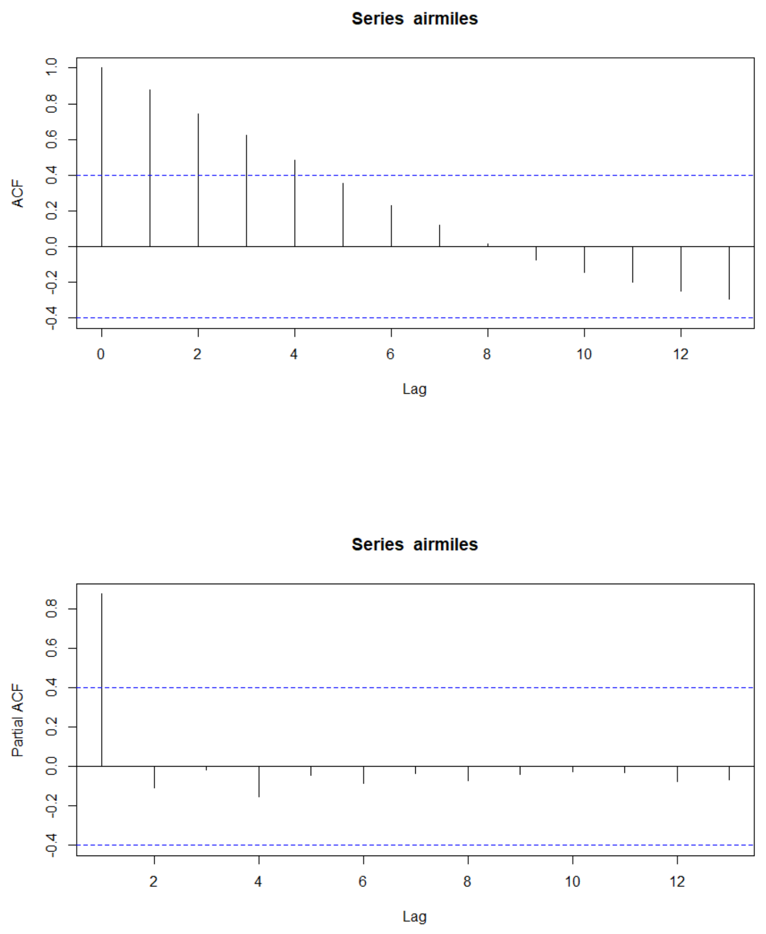
La función de autocorrelación parcial mide la correlación entre dos variables separadas por k periodos cuando no se considera la dependencia creada por los retardos existentes entre ambas, se define como:

$$\pi_j = corr(x_j, x_{j-k}|x_{j-1}x_{j-2} \ldots x_{j-k+1})$$

Los gráficos de autocorrelación son utilizados comúnmente para checar la aleatoriedad de los datos, que se comprueba analizando la relación entre pares de datos de la muestra, a diferentes rezagos. Si existe una correlación significativa en algún nivel de rezago, se puede decir que los datos no son completamente aleatorios.

CORRELOGRAMAS

Las gráficas de auto correlación destacan en azul un intervalo de confianza predeterminado del 95%. Se busca el punto a partir del cual las autocorrelaciones parciales comienzan a no ser significativas respecto a la hipótesis nula: auto correlación cero.



ASOCIACIÓN CON MODELOS

MODELO		FUNCIÓN DE AUTOCORRELACIÓN	FUNCIÓN DE AUTOCORRELACIÓN PARCIAL
Autorregresivo AR(p)		Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente	Se anula a partir del retardo p
Promedios móviles MA(q)		Se anula a partir del retardo q	Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente
ARMA(p,q)	p<q	Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente partir del retardo q-p+1, es decir que los primeros q-p+1 retardos no tienen pauta fija	Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente
	p>q	Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente	Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente pero a partir del retardo p-q+1, es decir que los primeros p-q+1 retardos no tienen pauta fija
	p=q	Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente pero a partir del retardo 1, es decir que el primer retardo no sigue este patrón	Decrece exponencialmente hacia o mediante ondas seno-coseno con amplitudes decreciendo exponencialmente pero a partir del retardo 1, es decir el primer retardo no varía según este patrón.

CRITERIO DE DESEMPEÑO

MODELOS DE SERIES DE TIEMPO

CRITERIO AIC

El AIC proporciona un medio para la selección del modelo, ya que considera la bondad de ajuste del modelo y la complejidad del modelo; sin llegar a ser una prueba de hipótesis estadística. El AIC calcula el **'criterio de información ' de Akaike** para uno o varios objetos de modelo ajustados para los que se puede obtener un valor de log-verosimilitud, según la fórmula

-2*log-verosimilitud + k*npar

donde npar representa el número de parámetros en el modelo ajustado, y k = 2 para el AIC usual, o k = log (n) (n siendo el número de observaciones) para el llamado BIC.

Lo que se busca es tener modelos que generen valores AIC lo más pequeño posible.

A modo de resumen se tiene:

- Buscamos series estacionarias para modelar.
- Los modelos pueden ser AR(p), MA(q) o ARMA (p,q).
- Los autocorrelogramas podrían servir como sugerencia para el modelo a bucar
- El criterio de selección del mejor modelo o el más adecuado es el valor AIC.
- La conclusión será la selección del modelo adecuado para la serie.