

Selección de Modelos

1 Criterios de Información

¿Cómo seleccionar un modelo de entre varios candidatos? Vía Criterios de Información. Aquí se describen tres criterios: AIC (Akaike Information Criterium), BIC (Bayesian Criterium) y HQ (Hannan-Quin).

Supongamos que tenemos \bar{k} modelos alternativos $M_1, M_2, \dots, M_k, \dots, M_{\bar{k}}$, donde k representa el número de parámetros que tiene el modelo M_k . Según los criterios de información se elige " k " tal que minimice

$$IC(k) = \ln \hat{\sigma}_k^2 + k \frac{C(T)}{T}$$

donde $\hat{\sigma}_k^2$ es la varianza residual del modelo M_k , T es el tamaño muestral y $C(T)$ es un término de penalización (o "penalty") que varía en función del criterio de información seleccionado. El AIC especifica $C(T) = 2$, el BIC $C(T) = \ln T$ y el HQ $C(T) = 2 * \ln(\ln T)$.

2 Modelos ARMA

Si se elige entre modelos de tipo AR(p), $k = p$; y si se escoge entre modelos de tipo ARMA(p,q) $k = p + q$.

Supongamos que estamos eligiendo entre la familia de modelos AR(p). $\hat{p}(IC)$ se elige tal que

$$IC(\hat{p}(IC)) = \min\{IC(p) | p = 0, 1, \dots, \bar{k}\}.$$

Un estimador \hat{p} del orden del AR(p) se dice que es consistente si

$$p \lim \hat{p}_T = p \quad \text{o} \quad \lim_{T \rightarrow \infty} \Pr(\hat{p} = p) = 1.$$

Proposición: Sea z_t un proceso AR(p) estacionario. Supongamos que el máximo orden considerado es $\bar{k} \geq p_0$ (el verdadero p) y que elegimos \hat{p} tal que minimice el criterio

$$IC(k) = \ln \hat{\sigma}_k^2 + k \frac{C(T)}{T}$$

para $k = 0, 1, \dots, \bar{k}$.

Entonces \hat{p} es consistente si y solo si $C(T) \rightarrow \infty$ y $C(T)/T \rightarrow 0$ cuando $T \rightarrow \infty$.

Podemos concluir que AIC no es consistente, pero sí lo son BIC y HQ (*¿por qué?*). En la prueba de la proposición se demuestra que aunque un criterio como

el AIC no es consistente, asintóticamente con probabilidad uno no eligirá un $\hat{p} < p_0$. Por lo tanto su inconsistencia viene producida por el hecho de que con una probabilidad positiva tiende a seleccionar un $\hat{p} > p_0$. *¿Cuándo esto no representa un problema serio?*

3 Filosofías en la Selección de Modelos

Diferenciamos dos filosofías en la selección de modelos en la práctica:

* De lo particular a lo general: Metodología Box-Jenkins. Identificación, Estimación, Diagnóstico (pasa o no pasa).

* De lo general a lo particular: Vía testing

Consejo: Forma tu propia filosofía teniendo en cuenta para que función objetivo el modelo va a ser usado.