Capítulo 7

Predicción econométrica

El objetivo de este capítulo es realizar un estudio de índole econométrica relativamente sencillo, de manera que podamos comparar el enfoque tradicional con el nuevo método predictivo que proponemos. A continuación se describe tal método de series de tiempo tradicional[1].

7.1. El Modelo

Se escogió para hacer el estudio un modelo comúnmente usado para analizar series de tiempo, un proceso autoregresivo de orden 1 o proceso $AR(1)$.[2] Específicamente este es un proceso de la forma:

$$y_t = \alpha + \rho_1 y_{t-1} + e_t, \quad t = 1, 2, \cdots \tag{7.1}$$

El punto de comienzo de la secuencia es y_0 (en $t = 0$), y la secuencia $\{e_t : t = 1, 2, \cdots\}$ se asume *independiente e idénticamente distribuida*, o i.d.d. con media cero y varianza σ_e^2. También asumimos que los e_t son independientes de y_0 y que la esperanza matemática del valor inicial cumple $E(y_0) = 0$.

Ya sabemos la dudosa credibilidad práctica de todos estos supuestos, pero por un momento los asumimos, con el objeto de usar el modelo. Para una discusión extensa sobre estos y otros problemas de los modelos que trabajan series de tiempo, consultar [25].

[1] Agradezco profundamente al Dr. David L. Wetzell quien en su estancia en la UDLA, Puebla en el Verano de 2004, proporcionó su tiempo y ayuda para perfeccionar el contenido de este capítulo.

[2] Del Inglés autoregressive process of order one.

La idea del modelo es simple y práctica, se infiere que cada observación y_t depende de la observación anterior, y_{t-1} (como ya vimos, se trata de un modelo con memoria a corto plazo) mediante el factor lineal ρ_1, la ordenada al origen α y un término de error e_t, mientras que conjunto de los $\{e_t\}$ se espera cumplan con las propiedades ya señaladas. Para saber más restricciones sobre el modelo, consultar [7] y [5].

Dicho lo anterior, en nuestro caso tomaríamos las 312 Observaciones de nuestro período de estudio como nuestra muestra (ver capítulo anterior) y aplicaríamos el siguiente modelo $AR(1)$ imitando (7.1):

$$IPC_t = \alpha + \rho_1 IPC_{t-1} + e_t \tag{7.2}$$

El siguiente paso es calcular el valor de ρ_1 y α, esto se puede hacer fácilmente con el método de los mínimos cuadrados ordinarios, o el de mínimos cuadrados ponderados, y un paquete de cómputo. En nuestro caso, Usamos el paquete econométrico STATA para ejecutar mínimos cuadrados ordinarios. Este método es el más común en la econometría y no es el objetivo de esta tesis explicarlo ya que puede encontrarse en cualquier libro de econometría básico. Para una excelente discución al respecto, remitimos al lector a [5]. Simplemente mencionaremos que el método, aplicando propiedades algebraicas y estadísticas de una regresión como nuestro modelo, cálcula el valor de ρ_1 tal que la línea resultante sea la de *mejor ajuste*, a todos los puntos de la muestra y que, por lo tanto, según este enfoque, describa mejor su comportamiento. Este método se basa en el cumplimiento del teorema de Gauss-Markov (ver [5] ,[7] y [25]), y sus distintas propiedades, que son del tipo que ya criticamos en el marco teórico de esta tesis (normalidad, independencia etc.), y que como ya mencionamos, ignoraremos por esta vez. Este método se agiliza significativamente, aplicando operaciones matriciales mediante la computadora, STATA hizo este trabajo por nosotros.

Al aplicar el método ya descrito sobre la muestra, surgió una complicación técnica importante, a saber la detección de raices unitarias al aplicar en la computadora el test de Dickey-Fuller[3]. Por esto, el modelo exigía ciertas modificaciones, que logramos superar y explicamos en la siguiente sección. Para más detalles sobre este problema, ver el apéndice ya señalado en la nota al pie de página.

[3]Ver apéndice, sección 10.2

7.2. El Modelo Modificado.

Como señalamos anteriormente, el modelo necesitaba modificarse para evadir el problema de las raices unitarias, esto se logró mediante las siguientes transformaciones:

Definimos una *diferencia porcentual* del IPC, dIPC de la siguiente manera[4]:

$$dIPC_t = \frac{|IPC_{t-1} - IPC_t|}{IPC_{t-1}} \tag{7.3}$$

Después, reformulamos nuestro modelo con la nueva variable descrita en (7.3) para obtener:

$$dIPC_t = \rho_1 dIPC_{t-1} + e_t \tag{7.4}$$

al correr nuevamente el modelo (modificado) en la computadora, y aplicar el test de Dickey-Fuller vimos felizmente que el problema había desaparecido. Cabe señalar que corrimos dos tipos de regresiones: la media (idéntica a la ya descrita) y la mediana. La única diferencia entre las dos estriva en que en la regressión mediana, el paquete hace una operación que minimiza la suma de los valores absolutos de las diferencias entre los valores observados y los predecidos en una regresión, a diferencia de mínimos cuadrados ordinarios, que minimiza el cuadrado de estas diferencias (ver [25]). STATA arrojó los siguientes resultados:

Para la regresión media:

$$\rho_{1media} = 0{,}07272 \qquad \alpha_{media} = 0{,}001625 \tag{7.5}$$

$$t = 1{,}32 \qquad t = 3{,}15 \tag{7.6}$$

Para la regresión mediana:

$$\rho_{1mediana} = 0{,}0983064 \qquad \alpha_{mediana} = 0{,}0129003 \tag{7.7}$$

$$t = 1{,}73 \qquad t = 2{,}42 \tag{7.8}$$

Las ecuaciones (7.6) y (7.8) dan los valores del estadístico t que resulta de dividir el valor de ρ_1 y e_t respectivamente, entre sus desviaciones estándar.

[4]Debo dar crédito al Dr. Wetzell por señalarme esta solución.

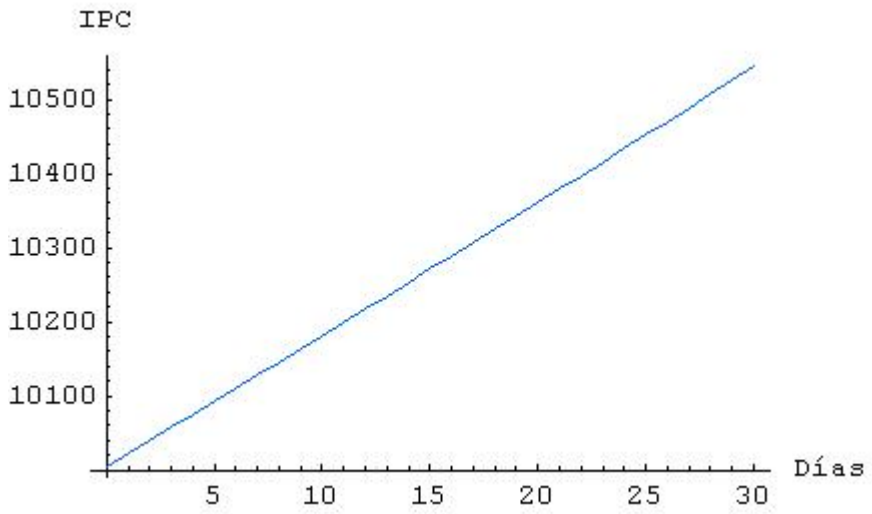

Figura 7.1: Regresión media.

Un valor de t cercano a 2 o mayor [5] nos dice que los valores encontrados son estadísticamente significativos (distintos de cero).

A pesar de que, debido a la transformación realizada en (7.3) los valores obtenidos son muy pequeños, resultaron significativos con un porcentaje de error $\leq 5\,\%$ excepto para ρ_{1media}. Sin embargo podemos ser flexibles ya que sabemos que este valor realmente es cercano a cero, y el t estadístico, simplemente nos comunica esto.

Ahora debemos regresar a nuestra variable original, mediante la siguiente transformación:

$$IPC_t = IPC_{t-1}(1 + dIPC_t) \tag{7.9}$$

7.3. Resultados Gráficos.

A continuación veamos los resultados predictivos de nuestras regresiones: para la regresión media (azul):

$$IPC_t = \alpha_{media} + \rho_{1media} IPC_{t-1} + e_{tmedia} \tag{7.10}$$

para la regresión mediana (morado):

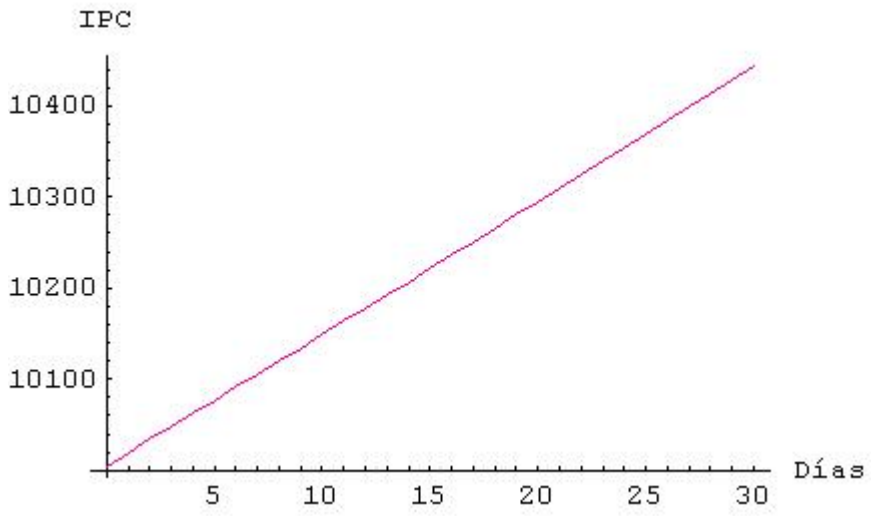

Figura 7.2: Regresión mediana.

$$IPC_t = \alpha_{mediana} + \rho_{1mediana} IPC_{t-1} + e_{tmediana} \tag{7.11}$$

Comparemos ambas con el IPC real:

Como vemos, el análisis econométrico nos indica una tendencia de la serie de tiempo a muy largo plazo, sin embargo en el corto plazo, deja de captar y predecir todo el detalle de sus movimientos, información vital para el inversionista.

En la figura anterior, podemos apreciar que la regresión mediana, lo hace significativamente mejor, pegándose más al gráfico real del IPC. Por lo tanto en el siguiente capítulo sólo hablaremos de la regresión mediana.

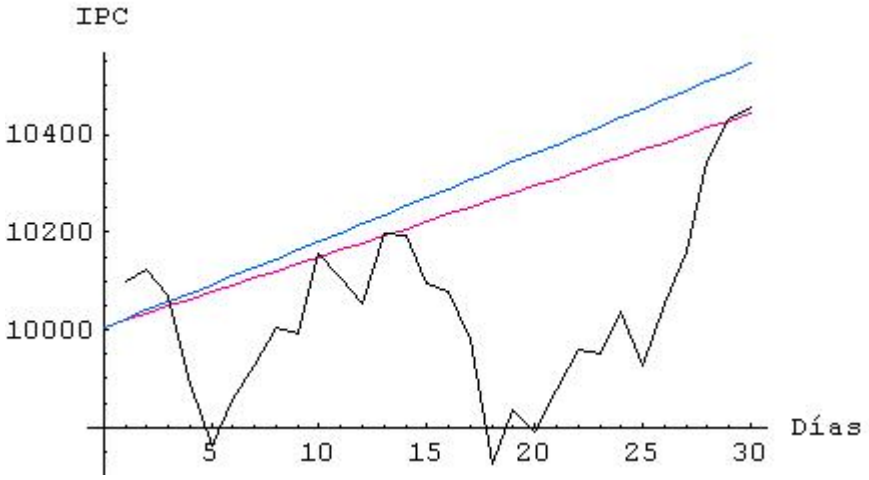

Figura 7.3: Ambas regresiones y el IPC real.