



Blaconá, María Teresa*

Bussi, Javier**

**Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística. Consejo de Investigaciones, Universidad Nacional de Rosario.*

E-mail: mblacona@fcecon.unr.edu.ar

*** Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario.*

COMPORTAMIENTO DE LA SERIE DE TIEMPO TASA DE DESOCUPACIÓN DEL GRAN ROSARIO EN EL PERÍODO 1974-2005¹

1 INTRODUCCIÓN

La tasa de desocupación del aglomerado urbano Gran Rosario, elaborada a partir de la Encuesta Permanente de Hogares (EPH), la cual es relevada por el Instituto de Estadística y Censo (INDEC), se viene publicando desde el año 1974.

Su comportamiento ha sido de un crecimiento sostenido, y es a partir de 1995 donde comienza a ser más inestable, alcanzando en la primer onda del 2001 un valor récord de 24,3 %.

En trabajos anteriores (Blaconá et al 2003, Blaconá et al 2004, Blaconá y Bussi, 2004), se realizaron diversos análisis de la serie de tiempo tasa de desocupación, en base a información bianual en el período 1974-2002. En dichos trabajos se detectó, a través de la utilización de los modelos de espacio de estados (MEE), que la serie presentaba nivel aleatorio, valores atípicos, cambios de nivel y estacionalidad determinística.

En este estudio se incrementa el período de observación hasta el primer trimestre del 2005. Es importante recordar, que a partir del año 2003 ha cambiado la periodicidad de captación de datos y la de su publicación. La información se releva en forma continua y se publica en forma trimestral.

Para analizar series de tiempo con las características antes mencionadas existen diversos métodos (Blaconá et al, 2005), de los cuales se usará en este trabajo el de MEE con distintas periodicidades (Harvey, 1989). Los MEE, también llamados modelos de componentes no observables, resultan una herramienta eficaz para este tipo de circunstancia, permitiendo encontrar el comportamiento estacional en la periodicidad trimestral y realizar pronósticos en esta nueva periodicidad.

En síntesis, en esta presentación se analiza la serie de tiempo tasa de desocupación del aglomerado urbano Gran Rosario (GR), en el período 1974 - 2005, teniendo en cuenta las dos periodicidades existentes en la serie. En especial, se estudia la tendencia, se detec-

¹ Este trabajo se realizó en el marco del proyecto: La información estadística como base para el diagnóstico de la desocupación en Rosario.



tan los puntos de quiebre, se determina el patrón estacional de la nueva periodicidad y se realiza el pronóstico del segundo trimestre del año 2005.

En la sección 2 se presentan brevemente los MEE y su extensión para distintas periodicidades. En la sección 3 se enumeran algunas características y supuestos para el estudio de la tasa de desocupación de GR, los resultados se muestran en la sección 4 y en la sección 5 se realizan las consideraciones finales.

2 MODELO DE ESPACIO DE ESTADOS (MEE)

Los modelos de espacio de estados, también conocidos como modelos de componentes no observables, pueden ser interpretados como regresiones sobre funciones del tiempo en las cuales los parámetros varían. Estos modelos consideran que los componentes no observables tales como la tendencia y la estacionalidad pueden ser estocásticos, los mismos son estimados a través de la utilización del filtro de Kalman, el cual permite la presencia de observaciones faltantes en las series modeladas.

2.1 MODELO BÁSICO DE ESPACIO DE ESTADOS

El modelo básico de espacio de estados, que también se conoce como modelo lineal Gaussiano de espacio de estados, se puede expresar como

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, & \boldsymbol{\varepsilon}_t &\sim \text{NID}(\mathbf{0}, \mathbf{H}_t), \\ \boldsymbol{\alpha}_t &= \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{R}_t \boldsymbol{\eta}_t, & \boldsymbol{\eta}_t &\sim \text{NID}(\mathbf{0}, \mathbf{Q}_t), \end{aligned} \quad (2.1.1)$$

donde \mathbf{y}_t es un vector de orden $p \times 1$ de observaciones, $\boldsymbol{\alpha}_t$ es un vector de orden $m \times 1$ no observable denominado vector de estado, las matrices $\mathbf{Z}_t(p \times m)$, $\mathbf{T}_t(m \times m)$ y $\mathbf{R}_t(m \times g)$ son conocidas y $\boldsymbol{\eta}_t$ es un vector aleatorio de orden $g \times 1$. \mathbf{R}_t se denomina matriz de selección y está formada de ceros y unos de acuerdo a si los $\boldsymbol{\eta}$ son determinísticos o aleatorios.

A la primera ecuación de (2.1.1) por lo general se la llama ecuación de medida y a la segunda, ecuación de transición.

2.2 MEE PARA SERIES DE TIEMPO CON DIFERENTES PERIODICIDADES

En el caso en que se deba analizar una serie con distintas periodicidades, se debe describir más detalladamente el MEE que es necesario formular.

Se considera el caso correspondiente a una serie con dos periodicidades, siendo la periodicidad inicial la que corresponde al período más largo entre dos observaciones registradas. En consecuencia, la periodicidad final corresponde al registro de observaciones en forma más frecuente.



El modelo general resulta:

$$y_t^v = \mathbf{z}'\boldsymbol{\alpha}_t + \varepsilon_t, \quad \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2 \quad (2.2.1)$$

$$\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \mathbf{R}\boldsymbol{\eta}_t, \quad \text{Var}(\boldsymbol{\eta}_t) = \mathbf{Q}$$

Se incluye en la notación el supraíndice v con el fin de destacar que la serie no es observada en todos los períodos de tiempo $t = 1, 2, 3, \dots, T$. Las observaciones que se indican de la forma y_τ , $\tau = 1, \dots, T_1$, están registradas cada δ períodos

$$y_t = y_{\delta\tau}^v \quad \tau = 1, \dots, T_1.$$

En estos casos para los $t \neq \delta\tau$, y_t^v no es observada. Las restantes observaciones de T_1+1 a T se presentan en la periodicidad más frecuente sin datos faltantes.

El MEE cuando existen observaciones faltantes debido al cambio de periodicidad, puede ser formulado de dos maneras. Una forma es plantear el modelo de acuerdo a los períodos más cortos (denominada de ahora en más FEE(t,1)) con $t = 1, \dots, T$. Otra forma de plantear el modelo es de acuerdo a los períodos más extensos, denominada FEE(τ,δ).

Ambas formas de plantear el MEE permiten que sean estimados los hiperparámetros, pero solo la FEE(t,1) admite la interpolación, es decir la estimación de valores en los momentos en los que no fueron registrados debido a la periodicidad en intervalos más largos. Otra ventaja de plantear el modelo de esta forma es que es posible hacer predicciones para ambas periodicidades como así también estimaciones de los componentes no observables en los períodos más cortos.

Esta formulación del MEE, la cual permite utilizar las dos partes de la serie de tiempo, resulta en una forma del filtro de Kalman que puede ser utilizada para el pronóstico de observaciones futuras y faltantes. El filtro para la forma del modelo FEE(t,1) se ejecuta de forma directa, el mismo saltea la actualización de las ecuaciones en los períodos intermedios donde no se cuenta con información, pero los considera en la estimación de la variancias.

Por ejemplo, sin perdida de generalidad, para un modelo sólo con nivel resulta:

$$y_t = \mu_t + \varepsilon_t \quad (2.2.2)$$

$$\mu_t = \mu_{t-1} + \eta_t$$

con $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$ y $\text{Var}(\eta_t) = \sigma_\eta^2$.



Suponiendo que no se cuenta con la observación y_k , la serie $y = y_t^v$, $t = 1, \dots, T$

donde $t_\tau = \begin{cases} \tau & \tau = 1, \dots, k-1 \\ \tau+1 & \tau = k, \dots, \tau \end{cases}$, se tiene el modelo para el momento $k+1$

$$\begin{aligned} y_{k+1} &= \mu_{k+1} + \varepsilon_{k+1} \\ \mu_{k+1} &= \mu_k + \eta_{k+1} \end{aligned} \tag{2.2.3}$$

como no se tiene información sobre μ_k se reemplaza por $\mu_{k-1} + \eta_k$, en consecuencia

$$\mu_{k+1} = \mu_{k-1} + \eta_k + \eta_{k+1}, \tag{2.2.4}$$

llamando $\bar{\eta}_{k+1} = \eta_k + \eta_{k+1}$, se deduce que

$$\text{Var}(\bar{\eta}_{k+1}) = 2\sigma_\eta^2. \tag{2.2.5}$$

Si se considera que la primera parte de la serie se registra en forma bianual y la segunda en forma trimestral, el filtro se aplica a FEE(t,1) y usa la primera observación como valor inicial para un modelo que considera sólo nivel. Como los componentes estacionales trimestrales no son observados en los datos bianuales, en esta primera etapa, los elementos asociados a ésta no aparecen en el vector de estado. Una vez que se arriba a los datos trimestrales, se expande el vector de estado para incluir los tres elementos estacionales y se necesitarían 4 observaciones como valores iniciales, pero en este caso sólo se requieren 3 porque ya se utilizó la primera observación de la serie bianual como valor inicial del nivel.

La función de verosimilitud exacta se obtiene a través de la descomposición del error usando los errores de predicción.

2.3 PUNTOS ABERRANTES Y CAMBIOS DE NIVEL EN MEE

La aparición de un punto aberrante (*outlier*) o de un cambio de nivel en la serie representan casos particulares del análisis de intervención. Las intervenciones están relacionadas con el hecho de estimar el efecto de eventos conocidos. Esto puede realizarse a través de variables *dummy*, las cuales son incluidas en el MEE.

$$y_t = \mathbf{z}_t' \boldsymbol{\alpha}_t + \lambda \omega_t + \varepsilon_t \quad t = 1, \dots, T \tag{2.3.1}$$

La variable ω_t representa la variable *dummy* siendo λ_t el coeficiente correspondiente. La



forma que toma la variable depende de si es un punto aberrante o un cambio de nivel.

En el caso de un punto aberrante, se supone que la intervención ocurre en un cierto tiempo t^* y su efecto se limita a ese momento. La variable ω_t toma la forma:

$$\omega_t = \begin{cases} 0, & t \neq t^* \\ 1, & t = t^* \end{cases}$$

Si se está en presencia de un cambio de nivel, la variable ω_t puede definirse:

$$\omega_t = \begin{cases} 0, & t < t^* \\ 1, & t \geq t^* \end{cases}$$

Este tipo de efecto podría ser también modelado como un cambio momentáneo en la ecuación del nivel.

3 CARACTERÍSTICAS Y SUPUESTOS SOBRE LA SERIE DE TIEMPO TASA DE DESOCUPACIÓN DEL GR

La tasa de desocupación de GR se viene relevando y publicando, en forma bianual, desde el año 1974. A partir del año 2003, dadas las nuevas características de la inserción laboral y el nuevo escenario emergente en relación a la condición social en general y al mercado de trabajo en particular, se hizo necesario adecuar globalmente los instrumentos de medición para dar cuenta de estos procesos y de su dinámica.

Es por ello que, sin alterar los propósitos iniciales del relevamiento, se realiza un intensivo proceso de reformulación integral de la EPH, con el objetivo de reelaborar la metodología de medición y las formas de operación atendiendo a las características socioeconómicas actuales, a las nuevas modalidades de inserción en el mercado de trabajo y a su dinámica de cambio.

La reformulación de la EPH abarca aspectos temáticos, en función de la adecuación de los instrumentos de captación y las dimensiones de análisis y presentación de resultados. Uno de los aspectos temáticos reformulados fueron los muestrales, tendientes a adecuar el diseño a cambios más frecuentes en los fenómenos medidos; y también fueron reconsiderados los aspectos organizativos, ajustando la modalidad de trabajo y los procesos informáticos a los diseños temáticos y muestrales establecidos (INDEC, 2003^(a), INDEC, 2003^(b)).

En función del desarrollo técnico de esos aspectos de la reformulación, la nueva en-



cuesta cuenta con cuestionarios rediseñados, aplicados en una muestra distribuida en el tiempo, bajo una modalidad de relevamiento continuo y con mayor frecuencia de presentación de resultados, que a partir del año 2003 se realiza en forma trimestral.

Para este trabajo se va a considerar que los cambios de definición y metodología en la construcción de la tasa de desocupación son absorbidos por la serie de tiempo, pero no afectan considerablemente la estructura de la misma. Por otro lado, el cambio en la periodicidad con que se capta la información y se publican dichas series, hace que se deban buscar metodologías de análisis de series de tiempo que permitan analizar, por ejemplo, el nuevo patrón estacional, aprovechando la información completa, es decir, la serie bimestral hasta el año 2003 y trimestral a partir de dicho año. También se pretende realizar pronósticos en la nueva periodicidad.

4 APLICACIÓN DE MEE CON DISTINTAS PERIODICIDADES PARA EL ANÁLISIS DE LA TASA DE DESOCUPACIÓN DE GR

En esta sección se presenta el análisis de la tasa de desocupación del GR considerando ambas periodicidades con las que fue registrada, es decir, bianual y trimestral, en el período 1974-2005, a través de los MEE. Por convención se adopta que la primera y segunda onda de la serie bianual corresponden respectivamente al segundo y cuarto trimestre en la serie trimestral. Para todos los análisis se utiliza el procedimiento UCM (Unobserved Components Model) del paquete estadístico SAS versión 9.1.3.

En un análisis previo de la serie de tiempo tasa de desocupación (Blaconá et al., 2004), se encontró un MEE con nivel aleatorio, estacionalidad determinística, un *outlier* y cambios de nivel. El ajuste se realizó utilizando la información bianual publicada por la EPH, desde el año 1974 hasta el 2002.

En base al modelo postulado previamente, se ajusta un modelo similar, pero la estacionalidad fue considerada en forma trimestral.

En consecuencia, el modelo presenta nivel aleatorio, estacionalidad trimestral determinística, un *outlier* el 2do. trimestre de 1989 y cuatro cambios de nivel (2do. trimestre de 1995 y 1997, 4to. trimestre de 2001 y 3er. trimestre 2003). El mismo se presenta a continuación:

$$y_t = \mu_t + \gamma_t + OU_{89.2} + CN_{95.2} + CN_{97.2} + CN_{01.4} + CN_{03.3} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \quad t = 74.1 \text{ al } 04.4,$$
$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \quad (4.1)$$
$$\gamma_t = -\sum_{j=1}^3 \gamma_{t-j},$$



donde μ_t , γ_t y ε_t representan la tendencia, estacionalidad e irregular, respectivamente, OU el outlier y CN representa los cambios de nivel. En la tabla 4.1 se presentan las estimaciones del modelo.

Tabla 4.1: Coeficientes estimados para el nivel al final del estado, el OU y los CN

Coeficiente	Estimación	Desvío st.	Valor t	Prob. asoc.
Nivel	11.97	3.22	3.72	
OU _{89.2}	5.55	1.31	4.24	<0.0001
CN _{95.2}	7.13	1.56	4.56	<0.0001
CN _{97.2}	-3.83	1.56	-2.45	0.0143
CN _{01.4}	4.06	1.56	2.60	0.0094
CN _{03.3}	-4.69	1.37	-3.41	0.0006
Estacionalidad	γ_1	-0.05		
	γ_2	0.01		
	γ_3	1.16		
	γ_4	-1.12		

Tabla 4. 2: Significación de los coeficientes al final del estado

Coeficiente	Grados de libertad	Chi-cuadrado	Prob. asoc.
Irregular	1	0.65	0.4188
Nivel	1	14.71	0.001
Estacionalidad	3	21.50	<.0001

$$\hat{\sigma}_\varepsilon^2 = 0.72, \hat{\sigma}_\eta^2 = 0.69.^2$$

El nivel estocástico al final del periodo llega a 11.97, y combinado con los efectos de los cambios de nivel resulta igual a 15.98. La estacionalidad es significativa y los coeficientes estacionales determinísticos son -0.05 para el primer trimestre, 0.01 para el segundo, 1.16 para el tercero y -1.11 para el cuarto. Los cambios de nivel se presentan en el segundo trimestre del año 1995, el que resulta igual a 7.13, en el segundo trimestre del año 1997 es de -3.83, 4.06 en el cuarto trimestre del año 2001 y -4.69 en el tercer trimestre del 2003.

En el gráfico 4.1 se muestra la tendencia de la serie, reflejada en las estimaciones suavizadas del nivel más los efectos de los cambios de nivel.

² Los hiperparámetros se estimaron por MV en base a las 66 observaciones disponibles.



Se presenta en el Gráfico 4.2 la serie ajustada a través de la combinación de las estimaciones del nivel y sus cambios más la estacionalidad.

En el Gráfico 4.3 se puede observar una representación de los coeficientes estacionales para los 4 trimestres. La estimación de los coeficientes estacionales debe ser analizada con cautela debido al escaso número de observaciones para su estimación, como así también por parecer que el último tramo de la serie de tiempo presentaría cambios de comportamiento.

Gráfico 4.1: Tendencia de la serie tasa de desocupación de GR

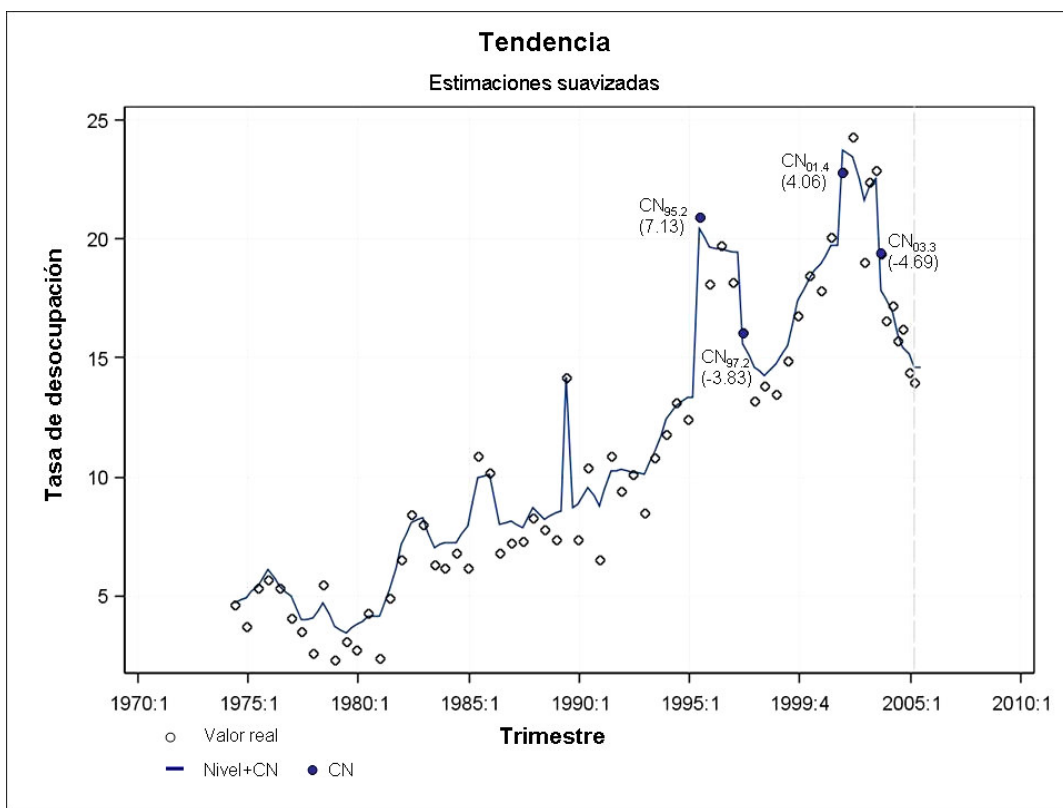




Gráfico 4.2: Ajuste de la serie tasa de desocupación de GR

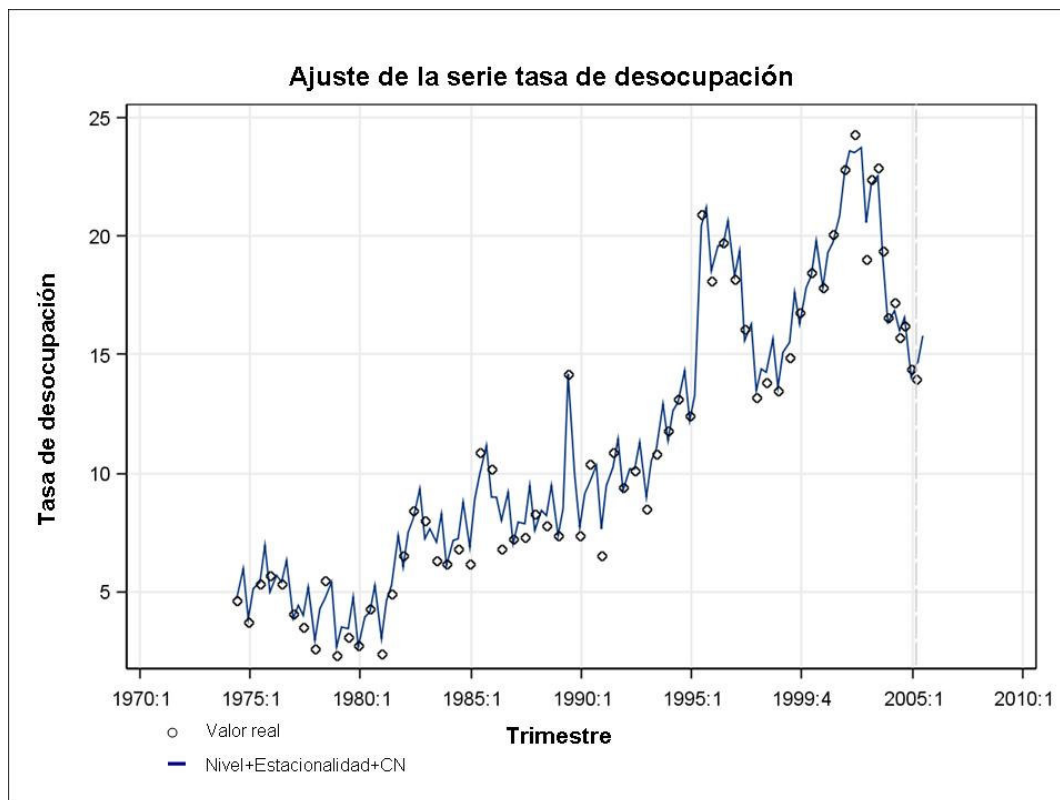
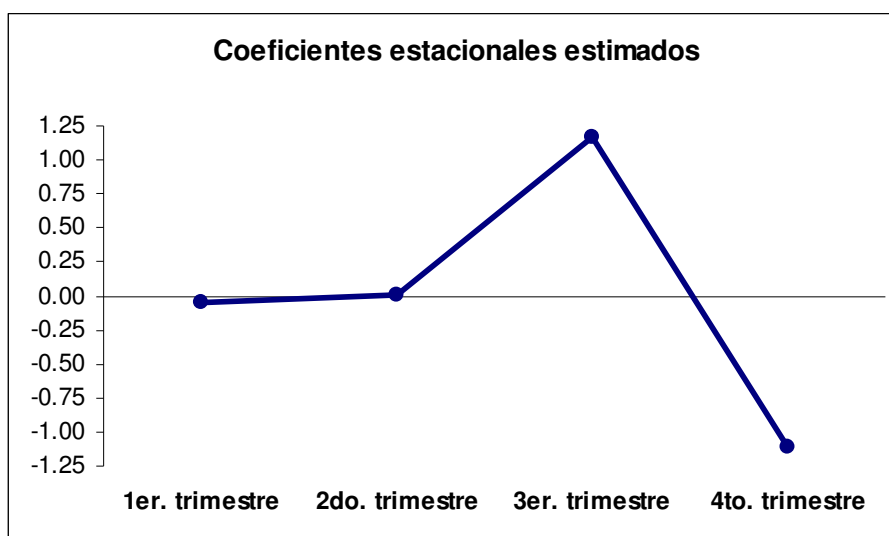


Gráfico 4.3: Coeficientes estacionales estimados de la serie tasa de desocupación de GR



Se realiza el pronóstico del segundo trimestre de 2005, que se presenta en la Tabla 4.1.



Tabla 4.3: Pronósticos para 1er. y 2do. trimestre de 2005

Trimestre	Pronóstico	Error Estándar	Intervalo de confianza 95%	Valor Real	% Error
2005:2	14.6	1.79	11.79 - 17.49	14.0	4.3

El error de pronóstico absoluto porcentual, para el segundo trimestre de 2005, es del 4.3%, ya que el valor publicado por INDEC de la tasa de desocupación para GR es 14.

5. CONSIDERACIONES FINALES

En este trabajo se realiza un análisis de la serie de tiempo tasa de desocupación del aglomerado urbano Gran Rosario en el período 1974 – 2005. Se debe tener en cuenta que esta serie se presenta con dos periodicidades, de 1974 al 2002 se publica en forma bi-anual, mientras que en los años 2003 al 2005 se lo hace en forma trimestral.

Para estudiar el comportamiento de la tendencia, la estacionalidad y los puntos de quiebre en la serie completa, se emplean los MEE combinando distintas periodicidades. Mediante este último modelo también se consiguen pronósticos en forma trimestral.

Entre las características de la serie se pueden destacar:

- i) el nivel de la serie es aleatorio y tiene un valor al final de la serie, considerando los cambios de nivel, de 16.3;
- ii) la estacionalidad es determinística y significativa, el comportamiento de estos coeficientes debe ser tomado con cautela debido al escaso número de datos con la nueva periodicidad y al hecho que corresponde además a un período que pareciera ser de cambio;
- iii) existe un *outlier* en el segundo trimestre del año 1989 que hace elevar la tasa de desocupación en aproximadamente 5 puntos, pero en los períodos posteriores la tasa vuelve a sus niveles anteriores;
- iv) se presentan varios cambios de nivel, el primero se produce en el segundo trimestre de 1995, con un incremento en la tasa de aproximadamente 7 puntos (el mayor en toda la historia de la serie), las tasas se mantienen en un nivel alto hasta el primer trimestre de 1997 y en el segundo trimestre del mismo año se produce un nuevo cambio de nivel, esta vez de signo negativo, descendiendo la tasa aproximadamente en 4 puntos, a partir del año 2000 comienza a crecer nuevamente, produciéndose un nuevo cambio de nivel positivo en el cuarto trimestre del 2001 con un valor de aproximadamente 4 puntos, se mantienen las tasas altas hasta el primer trimestre del 2003, a partir del cual comienzan a descender, observándose en el tercer trimestre de 2003 un nuevo cambio de nivel negativo de aproximadamente 5 puntos.

Lo expresado en el párrafo anterior lo podemos sintetizar diciendo que los períodos con



altas tasas de desocupación son del segundo trimestre de 1995 al primer trimestre de 1997 y del primer trimestre del 2001 al cuarto del 2002. Existe una relativa estabilidad del cuarto trimestre del 1997 hasta el cuarto de 1999. El comportamiento del año 2004 es similar al del año 1999, con la diferencia que en el año 2004 la tendencia es negativa, mientras en el año 1999 era positiva.

Los comportamientos detectados en este trabajo serán utilizados para relacionarlos con estudios que utilizan microdatos.

6 REFERENCIAS BIBLIOGRÁFICAS

_____ Blaconá, María Teresa, Bussi, Javier. (2004) "Uso del concepto de cointegración para la modelación de la tasa de desocupación", *Actas de la XXXIX Reunión Anual de la Asociación Argentina de Economía Política*, CD. Buenos Aires, 10 -12 noviembre de 2004. Publicado página web: www.aaep.org.ar/anales/publicaciones.

_____ Blaconá, María Teresa, Bussi, Javier, Ventroni, Nora. (2004) "Comparación de bondad de ajuste y pronósticos de modelos para tasa de desocupación. CLATSE IV, Universidad de Concepción, Chile.

_____ (2003) "Análisis de la tasa de desocupación de Gran Rosario y Gran Buenos Aires a través de modelos univariados de series de tiempo- (1º 1974 – 2º 2002)", Octavas Jornadas de Investigación, Facultad de Ciencias Económicas y Estadística, UNR. Página Web: www.fcecon.unr.edu.ar/scyt/jor/jor2003.

_____ Harvey, A.C. (1989). "Forecasting Structural Time Series and the Kalman Filter". Cambridge, U.K. Cambridge, University Press.

_____ INDEC (2003.)^(a). *La nueva Encuesta Permanente de Hogares de Argentina. 2003*.

_____ INDEC (2003)^(b). *Encuesta Permanente de Hogares (EPH). Cambios metodológicos*.