



Casa Abierta al Tiempo

Universidad Autónoma Metropolitana

IZTAPALAPA

Division de Ciencias Basicas e Ingenieria

**PRONOSTICO DE PUNTOS DE
CAMBIO EN SERIES DE TIEMPO**

T E S I S

Que para obtener el Grado de
MAESTRIA EN MATEMATICAS
P R E S E N T A

Eleonora Romero Vadillo

Asesor: VICTOR M. GUERRERO

MEXICO D.F. FEBRERO 1991.

PRONOSTICO DE PUNTOS DE CAMBIO EN SERIES DE TIEMPO

INDICE

	pag.
CAPITULO 1. INTRODUCCION	1
CAPITULO 2. ELEMENTOS PARA EL ANALISIS DE PUNTOS DE CAMBIO	5
2.1. Conceptos básicos	5
2.2. Puntos de cambio	16
2.3. Detección de puntos de cambio	23
CAPITULO 3. METODOS PARA EL PRONOSTICO DE PUNTOS DE CAMBIO	32
3.1. Sistema de predicción frecuencial	33
3.1.1. Calibración del SPF	39
3.1.2. SPF para el pronóstico de puntos aberrantes y cambios de nivel	44
3.2. Pronóstico de cambios de nivel	46
3.3. Enfoque Bayesiano	52
3.4. Ataques especulativos contra el peso mexicano	57
CAPITULO 4. APLICACIONES	62
4.1. Simulación univariada	62
4.2. Simulación bivariada	81
CAPITULO 5. CONCLUSIONES Y COMENTARIOS	91
APENDICE A. DISTRIBUCION PREDICTIVA	95

APENDICE B. GENERACION DE REALIZACIONES DE VARIABLES ALEATORIAS	99
APENDICE C. PROGRAMA PARA DETECTAR PUNTOS ABERRANTES Y CAMBIOS DE NIVEL	107
APENDICE D. PROGRAMA PARA PRONOSTICAR PUNTOS DE CAMBIO CUMBRE Y ABISMO	111
APENDICE E. ALGORITMO PARA EL PRONOSTICO DE CAMBIOS DE NIVEL Y PUNTOS ABERRANTES	115
APENDICE F. PROGRAMA PROBABILIDAD	122
REFERENCIAS.	126

PRONOSTICO DE PUNTOS DE CAMBIO EN SERIES DE TIEMPO

CAPITULO 1. INTRODUCCION

El propósito de este trabajo es exponer algunos métodos que permitan pronosticar *Puntos de Cambio*, para una variable cuyos datos sean observados como una serie de tiempo discreta.

El pronóstico de puntos de cambio es un problema de gran importancia, principalmente para series de tiempo relacionadas con los negocios y la economía, ya que consiste en determinar a futuro, a partir de una serie observada, el momento en el cual la serie puede presentar un cambio en el comportamiento de la variable. Por ejemplo, si no se espera un cambio en la demanda de cierto producto, la empresa que lo produce puede continuar con la política seguida hasta el momento, sin embargo, si se anticipara un punto de cambio en el futuro cercano, entonces pueden requerirse cambios substanciales en las políticas de producción y distribución de dicho producto. Asimismo, el pronóstico de cambios en el Producto Nacional Bruto, en la inflación, o bien el pronóstico de devaluaciones, son problemas de gran importancia para los analistas macroeconómicos en todo el mundo.

El pronóstico de puntos de cambio es un tema que no ha sido estudiado suficientemente y los trabajos que existen son bastante recientes. En el presente estudio se consideran principalmente los métodos y definiciones proporcionados por Wecker (1979), Kling (1987), Leiva (1987) y Zellner (1987), dado que son los únicos que se encontró, que proporcionan métodos para pronosticar puntos de cambio.

En la primera sección del Capítulo 2 se exponen brevemente los

conceptos que sirven de base para el presente estudio, como son el de serie de tiempo, tanto univariada como multivariada, el de operadores y polinomios de retraso, y el de proceso Gaussiano de ruido blanco univariado y multivariado; se exponen también los modelos que se emplean para representar el comportamiento de series de tiempo univariadas y multivariadas, así como el concepto de estacionariedad. En la segunda sección se dan algunas definiciones de punto de cambio, éstas son la de punto de cambio cumbre y abismo, cambio de nivel, punto salto, punto aberrante aditivo, punto aberrante innovado y cambio transitorio, y se estudia la relación que guardan entre ellas. En la tercera sección se estudia brevemente la detección de cambios mediante un enfoque propuesto por Tsay (1987), el cual consiste en realizar pruebas de la significación de los cambios; este método es aplicable en la detección de cambios de nivel, puntos aberrantes aditivos e innovados y para cambios transitorios, pero no puede aplicarse directamente en la detección de puntos de cambio cumbre y abismo, no obstante, éstos pueden transformarse en los anteriores mediante la diferenciación de la serie.

El Capítulo 3 se divide en cuatro secciones, en las primeras tres se exponen los que se considera son los principales métodos para el pronóstico de puntos de cambio, estos son, en primer lugar el método propuesto por Wecker (1979), el cual consiste en construir un sistema de predicción frecuencial a partir de la función de distribución predictiva de la serie, este trabajo fue generalizado por Kling para el caso de series de tiempo multivariadas. Se propone además un método para construir el sistema de predicción frecuencial para cambios de nivel, puntos aberrantes aditivos e innovados y cambios transitorios, se estudia además en esta sección algunas pruebas que pueden aplicarse para probar la buena calibración de un sistema de predicción frecuencial. En la segunda sección se expone el método propuesto por Leiva (1987) para el pronóstico de cambios de nivel, este método consiste en separar las componentes de nivel y de ruido de la serie, para así, después de suponer una distribución

para los tiempos de espera, calcular la probabilidad de un cambio de nivel en determinado momento. En la tercera sección se expone el enfoque Bayesiano proporcionado por Zellner (1987), este método se aplica en el pronóstico de puntos de cambio cumbre y abismo partiendo de los datos históricos y la función de distribución a priori de los parámetros del modelo generado por las mismas, a partir de los cuales se obtiene la distribución predictiva de probabilidad de las observaciones futuras. De la integración de esta función se obtiene la probabilidad de que un punto de cambio cumbre o abismo ocurra determinados periodos adelante, se proponen aquí algunas estructuras de pérdida para decidir si pronosticar que un punto de cambio ocurrirá o no. En la cuarta sección se describe un modelo desarrollado por Blanco y Garber (1986) con el fin de predecir el momento y la magnitud de una devaluación en sistemas de tipo de cambio fijo.

En el Capítulo 4 se realiza la simulación de dos series, la primera es univariada y se emplea para ilustrar el método de detección de cambios y los métodos de pronóstico univariados ya mencionados, con excepción del enfoque Bayesiano dado que la construcción de la función de distribución predictiva y la integración de la misma, requieren de paquetes de computo con los cuales no se cuenta. La segunda serie es bivariada y se emplea para observar las diferencias que se presentan al considerar variables explicativas y al no considerarlas.

Para finalizar el trabajo se exponen las conclusiones y comentarios derivados de los capítulos anteriores, es decir se plantean las ventajas y desventajas que se observan, como son la dificultad para decidir la significación de puntos de cambio cumbre y abismo bajo la definición propuesta por Wecker y lo laborioso que resulta construir sistemas de predicción frecuencial. Asimismo la dificultad para determinar la distribución de los tiempos de espera en el método propuesto por Leiva, cuando han ocurrido pocos cambios de nivel en la serie observada, no obstante la sencillez de

aplicación de este método.

Cabe mencionar que el propósito de este trabajo no es el de estimar modelos que describan los puntos de cambio, sino el de pronosticar en qué momento se puede producir el siguiente punto de cambio. Otros temas de gran interés, relacionados con puntos de cambio, son el de pronosticar cambios en los parámetros del modelo y cambios en la varianza de la serie, pero dichos temas no serán tratados en el presente trabajo por motivos de extensión.

Este trabajo va dirigido entonces, a todas aquellas personas que, con conocimientos básicos de probabilidad y estadística, se interesan en el análisis de series de tiempo como una herramienta más que sirva para el desarrollo de la ciencia, en particular en lo que toca a sus aplicaciones para solucionar problemas de carácter económico. Cabe aclarar que las aplicaciones del presente trabajo están dadas para series de tiempo económicas, pero los resultados aquí expuestos pueden ser aplicados en series de tiempo relacionadas con cualquier materia de estudio, de tal manera que no existen limitantes de aplicación y sí, en cambio, una amplia posibilidad de aplicaciones en la práctica.

CAPITULO 2. ELEMENTOS PARA EL ANALISIS DE PUNTOS DE CAMBIO

Antes de iniciar el análisis de puntos de cambio, es necesario proporcionar la definición de algunos conceptos importantes en los cuales se basa el desarrollo de este trabajo, por lo cual se comienza por definir, en la sección 1, lo que se entiende por *serie de tiempo*, tanto univariada como vectorial; se definen también en esta sección, los modelos que son más usuales para describir el comportamiento de una serie de tiempo y otros conceptos básicos como son la estacionariedad y la no-estacionariedad. Posteriormente, en la segunda sección, se estudian algunas definiciones de punto de cambio, estas son la de *Punto aberrante innovado*, *Punto aberrante aditivo*, la de *Cambio de nivel* y la de *Cambio transitorio*. Para finalizar este capítulo se estudia el problema de detección de cambios, tomando como base el metodo propuesto por Tsay (1987).

El conocimiento de estos elementos es fundamental para el pronóstico de puntos de cambio en series de tiempo, tema que se estudia en el capítulo 3 y que es el objetivo central de este trabajo.

2.1. Conceptos básicos

La información que se requiere para hacer un análisis estadístico de variables que dependen del tiempo (ya sean económicas o de otro tipo), debe presentarse numéricamente como el resultado de mediciones que se realizan a intervalos de tiempo fijo, esta información ordenada cronológicamente es lo que se conoce como una serie de tiempo.

DEFINICION-2.1.1. Una serie de tiempo univariada, denotada por $\{x_t\}$, es el resultado de las mediciones que se realizan a intervalos de tiempo fijo, de cierta variable de interés. En este caso, la variable es x_t , con el subíndice t denotando la dependencia del momento de observación.

DEFINICION-2.1.2. Una serie de tiempo k -variada, denotada por $\{x_t\}$ donde $x_t = (x_{1t}, x_{2t}, \dots, x_{mt})'$, con el apóstrofe denotando transposición, es el resultado de la medición simultánea, a intervalos de tiempo fijo de m variables de interés.

Una serie de tiempo se puede ver entonces, como una familia de variables aleatorias asociadas a un conjunto de índices de números reales, que representan el tiempo, con la característica de que a cada elemento del conjunto de índices le corresponde una y sólo una variable aleatoria.

Si el conjunto de índices de tiempo T es un intervalo de números reales, entonces la serie de tiempo es continua y si T es un conjunto numerable, la serie de tiempo se dice discreta. El hecho de que la serie de tiempo sea continua o discreta no indica que la medición de la variable aleatoria deba ser continua o discreta. En lo sucesivo se hará referencia únicamente a series de tiempo discretas.

Existen modelos de gran utilidad para representar fenómenos reales observados como series de tiempo, estos modelos serán descritos a continuación en forma muy breve, pero pueden ser estudiados con más detalle en Box y Jenkins (1970) o en Guerrero (1983).

Antes de describir los modelos, es necesario definir operadores y polinomios de retraso, que permitan la representación de estos en forma concisa.

DEFINICION-2.1.3 El operador de retraso, denotado por B , es un operador tal que

$$B^k x_t = x_{t-k} \quad \text{para toda } k = 0, 1, 2, \dots \text{ y toda } t$$

Nótese que si únicamente se tienen las observaciones $\{x_1, x_2, \dots, x_n\}$, al aplicar el operador B^k , la serie se transforma en $\{x_{1-k}, x_{2-k}, \dots, x_{n-k}\}$ donde $x_{1-k}, x_{2-k}, \dots, x_0$ son desconocidas y por lo tanto la serie se reduce a únicamente $n-k$ observaciones.

DEFINICION-2.1.4. El operador diferencia, denotado por ∇ , es un operador tal que

$$\nabla x_t = x_t - x_{t-1} \quad \text{para toda } t$$

Obsérvese que

$$\nabla x_t = (1-B)x_t$$

y que de la aplicación sucesiva de ∇ , se obtiene

$$\nabla^k x_t = \sum_{j=1}^k \frac{k!}{j!(k-j)!} (-1)^j x_{t-j} \quad \text{para } k = 1, 2, \dots$$

DEFINICION-2.1.5. Un polinomio de retraso, denotado por $G(B)$, es un operador lineal tal que

$$G(B)x_t = x_t - g_1 x_{t-1} - g_2 x_{t-2} - \dots - g_k x_{t-k}$$

$$= x_t - \sum_{j=1}^k g_j x_{t-j} \quad \text{para } k = 1, 2, \dots$$

donde los coeficientes $g_1, g_2, \dots, g_k$ son constantes que ponderan la importancia de los retrasos con los que están asociados.

Un concepto intimamente ligado con los modelos que se usan para representar fenómenos observados como series de tiempo, es el de proceso de ruido blanco, el cual se define a continuación.

DEFINICION-2.1.6. Un *proceso Gaussiano univariado de ruido blanco*, denotado por $\{a_t\}$, es la realización de variables aleatorias independientes, normalmente distribuidas con media cero y varianza σ_a^2 , ordenadas cronológicamente.

En esta definición se aprecia entonces que la única diferencia entre un proceso de ruido blanco y una muestra aleatoria, es el hecho de que el proceso de ruido blanco presenta un orden cronológico de realización de variables, mientras que en la muestra aleatoria no importa dicho orden.

DEFINICION-2.1.7. Un *proceso Gaussiano multivariado de ruido blanco*, denotado por $\{a_t\}$, es la realización ordenada cronológicamente de vectores aleatorios independientes, con distribución normal multivariada $N(0, \Sigma_m)$, donde Σ es la matriz de varianza-covarianza

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2m} \\ \dots & \dots & \dots & \dots \\ \sigma_{1m} & \sigma_{2m} & \dots & \sigma_m^2 \end{bmatrix}$$

A continuación se describen en primer lugar, los modelos que sirven para representar el comportamiento de la mayor parte de las series de tiempo univariadas y en segundo lugar se describen los modelos empleados para representar la conducta de series de tiempo multivariadas.

DEFINICION-2.1.8. Se dice que la serie de tiempo $\{x_t\}$ está representada por un modelo de *promedios móviles* de orden q (MA(q)), si está dada mediante la expresión

$$x_t - \mu = \Theta(B)a_t$$

donde μ denota a la media o nivel global de la serie, $\{a_t\}$ es un proceso de ruido blanco y $\Theta(B)$ es el polinomio de retraso dado por

$$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

con $\theta_1, \theta_2, \dots, \theta_q$ los correspondientes parámetros de promedios móviles y $\theta_q \neq 0$.

DEFINICION-2.1.9. Se dice que la serie $\{x_t\}$ está representada por un modelo *autorregresivo* de orden p (AR(p)), si está dada por la siguiente expresión

$$\Phi(B)(x_t - \mu) = a_t$$

donde μ es el nivel de la serie, $\{a_t\}$ es un proceso de ruido blanco y $\Phi(B)$ es el polinomio de retraso dado por

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

con $\phi_1, \phi_2, \dots, \phi_p$ los respectivos parámetros autorregresivos y $\phi_p \neq 0$.

DEFINICION-2.1.10. Se dice que $\{x_t\}$ está representada por un modelo *autorregresivo y de promedios móviles* de orden p, q (ARMA(p, q)), si está dada por la combinación de los dos modelos anteriores, esto es, tiene la representación

$$\Phi(B)(x_t - \mu) = \Theta(B)a_t$$

DEFINICION-2.1.11. Se dice que $\{x_t\}$ esta representada por un modelo autorregresivo integrado y de promedios moviles de orden p, d, q (ARIMA(p, d, q)) si está dada por

$$\Phi(B)\nabla^d x_t = \theta_0 + \theta(B)a_t$$

con θ_0 un parámetro que sirve para tomar en cuenta una posible tendencia polinomial determinista.

El problema de identificación y estimación de estos modelos ha sido ampliamente estudiado por Box y Jenkins (1970); para un estudio más simplificado, consúltese Guerrero (1983).

En la práctica, el comportamiento de una variable observada como serie de tiempo, generalmente se encuentra ligado con el comportamiento de otras variables igualmente observables, de ahí la necesidad de crear modelos que expliquen simultáneamente la conducta de k series de tiempo, algunos de estos modelos serán descritos a continuación en forma muy concisa y pueden ser estudiados con más detalle en Guerrero (1987).

DEFINICION-2.1.12. El comportamiento de la serie de tiempo vectorial $\{x_t\}$ está explicado por un vector de promedios móviles de orden q (MA(q) vectorial) si está dado por la expresión

$$x_t = D + \Theta(B)a_t$$

donde $\Theta(B)$ denota a la matriz de retrasos dada por

$$\Theta(B) = \begin{bmatrix} \theta_{11}(B) & \theta_{12}(B) & \dots & \theta_{1m}(B) \\ \theta_{21}(B) & \theta_{22}(B) & \dots & \theta_{2m}(B) \\ \dots & \dots & \dots & \dots \\ \theta_{m1}(B) & \theta_{m2}(B) & \dots & \theta_{mm}(B) \end{bmatrix}$$

con $\theta_{ij}(B) = 1 - \theta_{ij,1}B - \theta_{ij,2}B^2 - \dots - \theta_{ij,q}B^q$, para $i, j = 1, 2, \dots, m$, $q \geq 1$. D representa a un vector de factores deterministas y $\{a_t\}$ denota un proceso multivariado de ruido blanco.

DEFINICION-2.1.13. El comportamiento de la serie de tiempo vectorial $\{x_t\}$ está explicado por un vector autorregresivo de orden p (AR(p) vectorial), si está dado por la expresión

$$\Phi(B)x_t = D + a_t$$

donde $\Phi(B)$ denota a la matriz de retrasos dada por

$$\Phi(B) = \begin{bmatrix} \phi_{11}(B) & \phi_{12}(B) & \dots & \phi_{1m}(B) \\ \phi_{21}(B) & \phi_{22}(B) & \dots & \phi_{2m}(B) \\ \dots & \dots & \dots & \dots \\ \phi_{m1}(B) & \phi_{m2}(B) & \dots & \phi_{mm}(B) \end{bmatrix}$$

con $\phi_{ij}(B) = 1 - \phi_{ij,1}B - \phi_{ij,2}B^2 - \dots - \phi_{ij,p}B^p$ para $i, j = 1, 2, \dots, m$, $p \geq 1$. D representa a un vector de factores deterministas (que incluye el nivel de las series) y $\{a_t\}$ denota un proceso multivariado de ruido blanco.

DEFINICION-2.1.14. La conducta de la serie $\{x_t\}$ está representada por un modelo autorregresivo y de promedios móviles vectorial de orden p, q (ARMA(p, q) vectorial), si se expresa por una combinación de los dos modelos anteriores, esto es si

$$\Phi(B)x_t = D + \Theta(B)a_t$$

En la práctica, la identificación y estimación de los parámetros de un modelo ARMA vectorial presenta aún serias dificultades, por eso en general es preferible utilizar un modelo AR

vectorial, el cual puede verse como una aproximación del modelo ARMA vectorial, pero es más fácil su identificación y estimación. La identificación y la estimación de los modelos AR vectorial ha sido estudiada por Quinn (1980) según lo indica Guerrero (1987).

Otro concepto de gran importancia en el estudio de series de tiempo es el de *estacionariedad*, el cual se estudia a continuación de manera muy breve pero puede ser estudiado más extensamente en Box y Jenkins (1970).

Supóngase que la serie $\{x_t\}$ está expresada como función de $\{a_t\}$, mediante la relación lineal

$$x_t = \mu + \Psi(B)a_t$$

donde μ representa el nivel de la serie y $\Psi(B)$ es el polinomio de retraso dado por

$$\Psi(B) = 1 - \psi_1 B - \psi_2 B^2 - \dots$$

Entonces la esperanza de x_t está dada por

$$E(x_t) = E(\mu + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)$$

ahora bien, si

$$1 + \sum_{i=1}^{\infty} |\psi_i| < \infty \quad (2.1.1)$$

se tiene que

$$E(x_t) = \mu$$

Es decir, si (2.1.1) se cumple, la media es igual que el nivel de la serie, ya que dicho nivel no depende del tiempo.

La varianza de $\{x_t\}$ está dada por

$$\begin{aligned}\gamma_0 &= E[(x_t - \mu)^2] \\ &= E[(a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)^2] \\ &= E(a_t^2 - \psi_1^2 a_{t-1}^2 - \psi_2^2 a_{t-2}^2 + \dots) - E\left[\sum_{i,j} \psi_i \psi_j a_{t-i} a_{t-j}\right]\end{aligned}$$

dado que $\{a_t\}$ es un proceso de variables aleatorias independientes con media cero, la esperanza de los productos cruzados es cero y por lo tanto

$$\gamma_0 = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2$$

lo cual converge siempre y cuando (2.1.1) se cumpla, y en tal caso γ_0 tampoco depende del tiempo.

Es necesario estudiar también la covarianza entre x_t y x_{t+k} , la cual se denota por γ_k y está dada por

$$\begin{aligned}\gamma_k &= E[(x_t - \mu)(x_{t+k} - \mu)] \\ &= E[(a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)(a_{t+k} - \psi_1 a_{t+k-1} - \psi_2 a_{t+k-2} - \dots)] \\ &= E(-\psi_k a_t^2 + \psi_1 \psi_{k+1} a_{t-1}^2 + \psi_2 \psi_{k+2} a_{t-2}^2 + \dots) \\ &\quad - E(\text{productos cruzados de } \psi_i a_{t-i} \text{ y } \psi_j a_{t-j}) \\ &= \sigma_a^2 (-\psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \dots) \\ &= \sigma_a^2 \left[\sum_{i=1}^{\infty} \psi_i \psi_{k+i} - \psi_k \right]\end{aligned}$$

lo cual tiene sentido únicamente si $\sum_{i=1}^{\infty} \psi_i \psi_{k+i}$ converge, y esto ocurre cuando (2.1.1) se cumple, en este caso γ_k sólo depende de la separación k entre las variables y no del tiempo.

DEFINICION-2.1.15. Una serie de tiempo se dice estacionaria de segundo orden, si se cumple que sus momentos de primero y segundo orden ($E(x_t)$, γ_k , $k \geq 0$) no dependen del tiempo. Esto es, si (2.1.1) se cumple.

En la figura (2.1.1) se muestra la gráfica típica de una serie de tiempo estacionaria. En la práctica, la mayor parte de las series de tiempo no son estacionarias, aunque muchas de éstas presentan una *no-estacionariedad homogénea*, esto es, únicamente el nivel de la serie se ve afectado por la no-estacionariedad, en tal caso las series presentan una tendencia polinomial adaptiva, del tipo mostrado en la figura (2.1.2). Este tipo de series puede ser estacionarizada mediante la aplicación sucesiva del operador diferencia ∇ , dado en la definición (2.1.4).

En la figura (2.1.3) se presenta la gráfica de la primera diferencia de la serie mostrada en la figura (2.1.2). Se puede observar que esta serie aún no es estacionaria, pero sí es localmente estacionaria, esto es, hasta el tiempo $t = 27$ la serie fluctúa alrededor de cierto nivel, pero a partir de este momento la serie cambia de nivel, comenzando a fluctuar alrededor de un nivel diferente y ocurre lo mismo al tiempo $t = 64$. Este tipo de comportamiento da origen a la definición siguiente.

DEFINICION-2.1.14. Un nivel local subyacente se define como aquel punto alrededor del cual la serie de tiempo fluctúa durante cierto intervalo de tiempo $[t_0, t_1]$.

Se puede observar que la serie de la figura (2.1.3) presenta tres niveles locales subyacentes, asociados con los intervalos $[1, 28)$, $[28, 63)$ y $[64, 100]$, estos niveles son respectivamente 0.35, 0.117, 0.74.

FIGURA 2.1.1

Gráfica de una serie de tiempo estacionaria

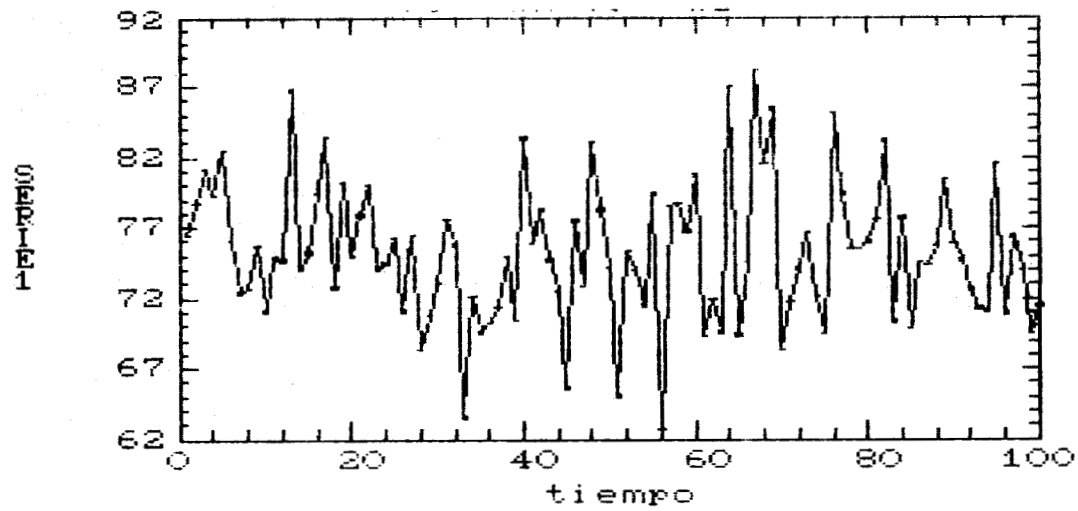


FIGURA 2.1.2

Gráfica de una serie de tiempo que presenta tendencia polinomial adaptiva

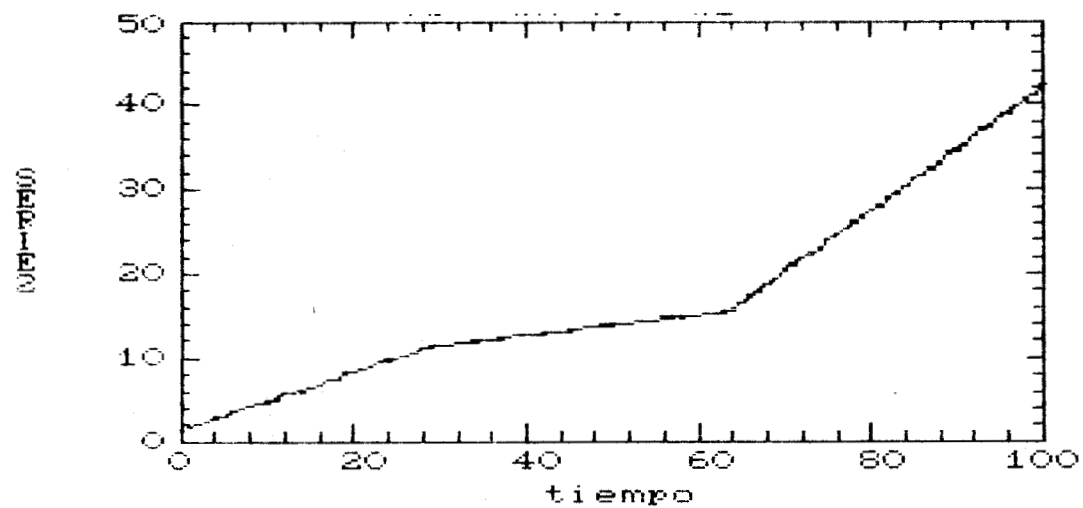
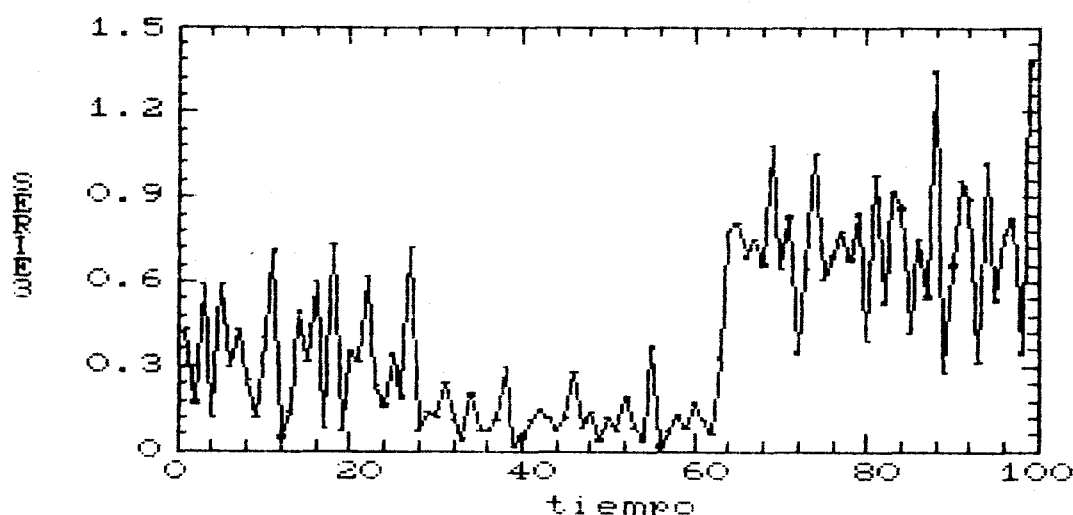


FIGURA 2.1.3
Gráfica de la primera diferencia de la serie no-estacionaria de la figura 2.1.2



2.2 Puntos de cambio.

El problema del pronóstico de puntos de cambio, comienza con la dificultad para dar una definición que sea "adecuada" en todo contexto, ya que un punto de cambio puede tener diferente significado dependiendo del fenómeno que se estudie.

En estudios de series de tiempo económicas, la noción de punto de cambio, denotado de aquí en adelante como PC, es la de un punto en el tiempo en el cual, una serie que ha venido creciendo (o decreciendo) cambia y por un tiempo decrece (o crece). Esto es, por un PC se entiende cualquier momento, en el cual se observe un cambio de signo en la pendiente de la tendencia de la serie. El problema que se presenta luego, es determinar de qué magnitud debe ser la duración de ese cambio, para que no sea considerado únicamente como una perturbación meramente aleatoria y esto dependerá básicamente del contexto del problema. Una definición de PC que puede usarse en forma general es la de un punto en el tiempo, en el cual la pendiente de la tendencia de la serie cambia de signo,

conservándolo al menos durante los τ periodos siguientes. Con τ elegido por el analista de acuerdo con el contexto del problema.

Ahora bien, el comportamiento de una serie de tiempo generalmente está ligado al comportamiento de otras series, las cuales se pueden introducir en el problema como variables explicativas; esto es, un PC en una serie de tiempo, depende comúnmente de los cambios de comportamiento de otras series. Debido a esto, se proporcionará la definición de PC de tal manera que sea aplicable también a series de tiempo vectoriales.

Es necesario además, distinguir entre un PC que se presenta como máximo local, al cual se llamará PC *Cumbre* y se denotará por PCC, y aquél PC que aparece como mínimo local, al cual se denominará PC *Abismo* y se denotará por PCA.

DEFINICION-2.2.1. Sean $x_t = (x_{1t}, x_{2t}, \dots, x_{mt})$ para $t = 1, 2, \dots, n$, los valores observados de una serie de tiempo vectorial $\{x_t\}$. Se dice que un PCC ocurre al tiempo t en la serie i , si

$$x_{it} > x_{i,t+k}, \text{ para } k = -\tau, -\tau+1, \dots, -1, 1, \dots, \tau-1, \tau.$$

Y se dice que un PCA ocurre al tiempo t en la serie i , si

$$x_{it} < x_{i,t+k}, \text{ para } k = -\tau, -\tau+1, \dots, -1, 1, \dots, \tau-1, \tau.$$

En la figura 2.2.1, por ejemplo, si se toma $\tau = 1$ se tienen 13 PCC, estos ocurren al tiempo $t = 6, 9, 15, 25, 32, 38, 42, 56, 69, 67, 69, 78$ y 92 . Si $\tau = 2$ se tienen 7 PCC, al tiempo $t = 6, 25, 42, 56, 69, 67$ y 78 . Para $\tau = 3$ se tienen únicamente 4, al tiempo $t = 6, 25, 56$ y 67 . Para PCA, si $\tau = 1$ se tienen 11, al tiempo $t = 2, 8, 14, 22, 35, 39, 58, 61, 69, 74$ y 88 . Pero si $\tau = 2$, se tienen PCA en los tiempos $t = 8, 22, 35, 58, 61, 69$ y 76 . Con $\tau = 3$, los PCA se dan al tiempo $t = 22, 35$ y 61 .

Nótese que las series de tiempo que presentan PCC y/o PCA son al menos a corto plazo (dependiendo esto del contexto del problema) series no estacionarias, las cuales presumiblemente siguen una tendencia polinomial adaptiva a corto plazo, esto es, pueden ser series que presentan no-estacionariedad, la cual se supondrá que es homogénea. En la gráfica de la figura 2.2.1 se presenta una serie que vista en su conjunto es presumiblemente estacionaria, sin embargo, considerando porciones de ésta, se observa cierta tendencia polinomial en la que se detectan puntos de cambio; de aquí la importancia de especificar perfectamente el problema antes de proceder a su estudio.

Desde luego, podrían emplearse otras definiciones de PC más específicas, que podrían ser de mayor utilidad, dependiendo del contexto del problema que se desea analizar. Así, podría pensarse que un PC es un punto en el tiempo en el cual, una serie que se encuentra fluctuando en cierto nivel, cambia y comienza a fluctuar en otro nivel distinto, estos puntos se denotarán da aquí en adelante por CN. Un ejemplo de este caso se representa en la figura 2.2.2, en la cual se pueden apreciar dos niveles subyacentes de la serie. La serie cambia de nivel en los tiempos $t = 7, 13$ y 21 .

FIGURA 2.2.1
Serie de tiempo que presenta PCC y PCA

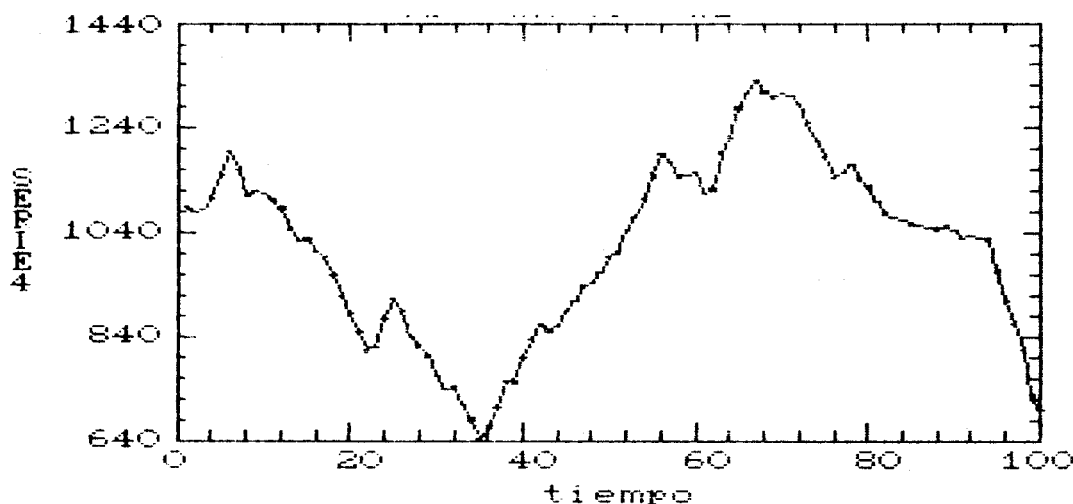
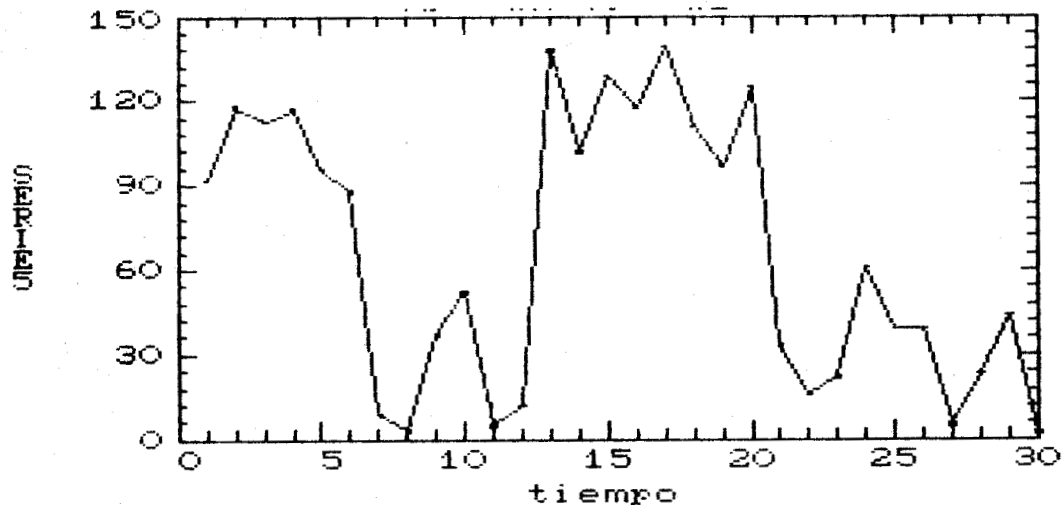


FIGURA 2.2.2
Serie de tiempo que presenta cambios de nivel



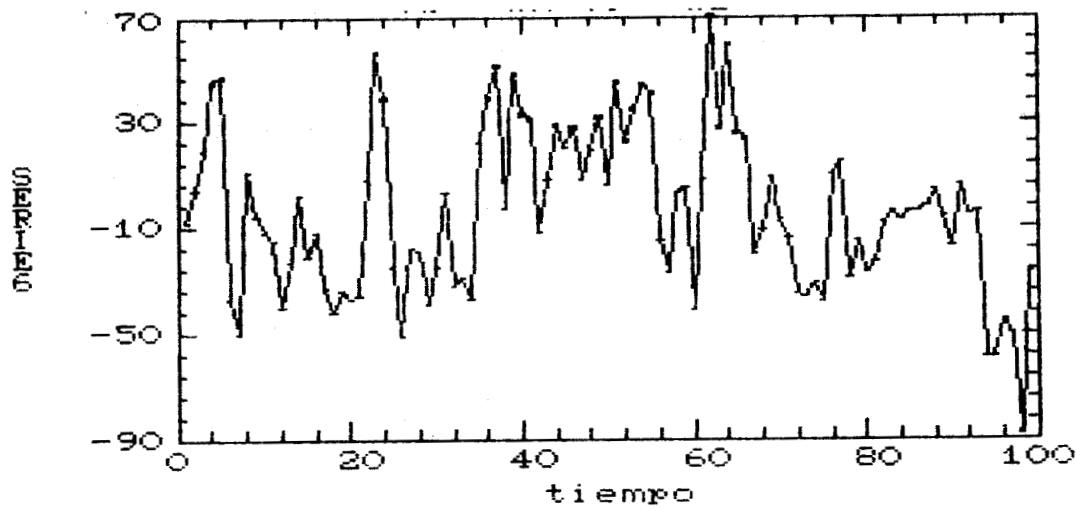
Nótese entonces que una serie que presenta cambios de nivel es una serie localmente estacionaria.

Obsérvese además que un PCC o un PCA puede traducirse, mediante la diferenciación, en un cambio de nivel. A manera de ilustración se presenta, en la figura 2.2.3, la gráfica de la primera diferencia de la serie dada en la figura 2.2.1.

Más aún, cualquier cambio en la pendiente de una serie de tiempo se traduce, mediante la diferenciación, en un CN, como puede observarse en las figuras 2.1.2 y 2.1.3, donde se presentan respectivamente, una serie de tiempo con cambios de pendiente y su primera diferencia, la cual presenta cambios de nivel.

FIGURA 2.2.3

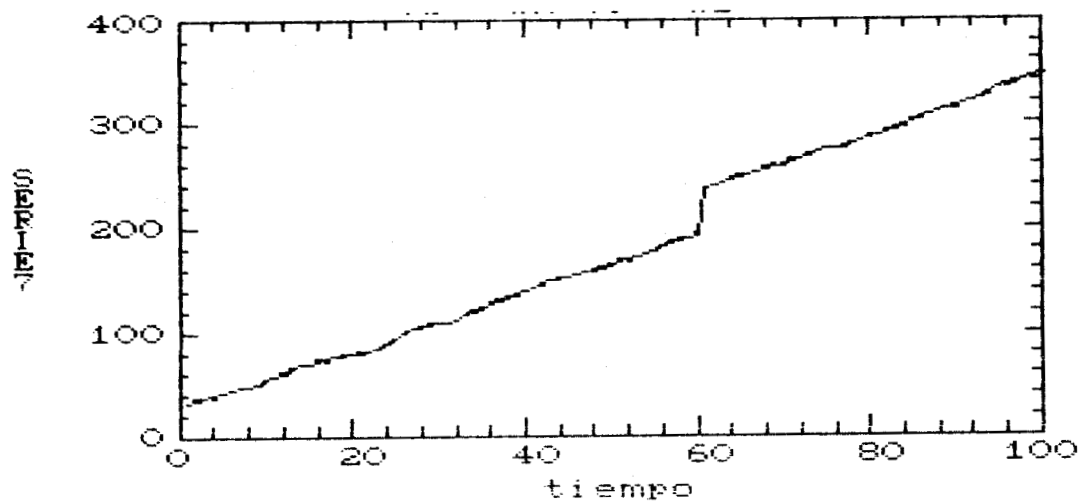
Gráfica de la primer diferencia de la serie dada en la figura 2.2.1



Un PC puede verse también como un punto en el tiempo en el cual, una serie que ha venido creciendo (o decreciendo) con cierta pendiente, presenta un salto pero conserva su pendiente, a este tipo de cambio se le denominará *Punto Salto* (PS) y se ilustra en la figura 2.2.4.

FIGURA 2.2.4

Serie de tiempo que presenta un punto salto



Puede encontrarse también, que una serie que fluctúa alrededor de cierto nivel, en determinado momento presenta un punto alejado de éste, a este punto se le llama *Punto Aberrante Aditivo (PA)* y se ilustra en la figura 2.2.5.

Nótese que un PS puede transformarse en un PA, mediante la diferenciación de la serie. De hecho, la figura 2.2.5 representa la primera diferencia de la serie ilustrada en la figura 2.2.4.

Otra forma de *Punto Aberrante* es el llamado *Punto Aberrante Innovado (PI)*, el cual afecta a varias observaciones sucesivas, un ejemplo de este caso se muestra en la serie de la figura 2.2.6.

Por último, otro tipo de cambio con el que es posible enfrentarse, es el *Cambio Transitorio (CT)*, en el cual el nivel de la serie se ve afectado momentáneamente y el efecto disminuye gradualmente al paso del tiempo, esto se ilustra en la gráfica de la figura 2.2.7.

FIGURA 2.2.5
Serie de tiempo que presenta un punto aberrante aditivo

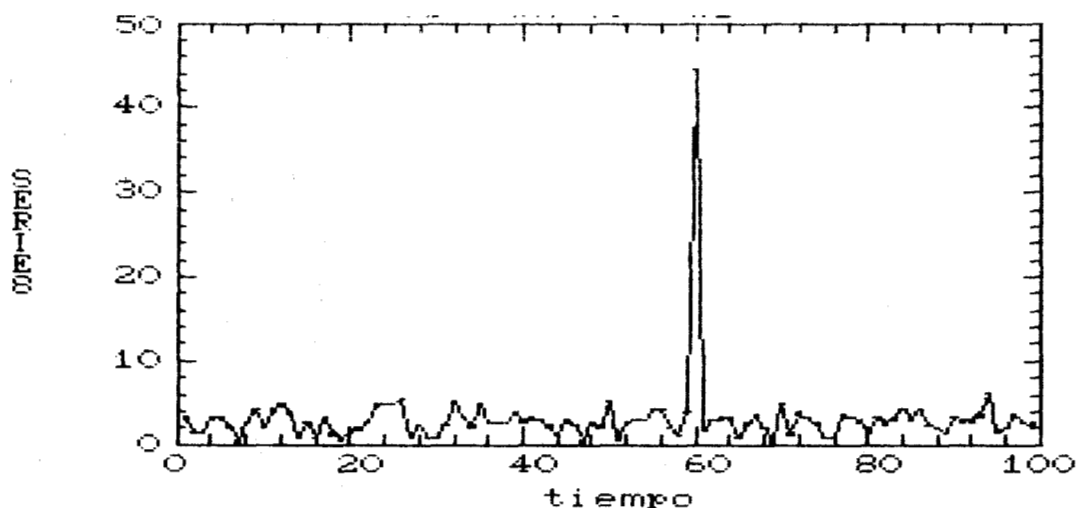


FIGURA 2.2.6

Serie de tiempo que presenta un punto aberrante innovado

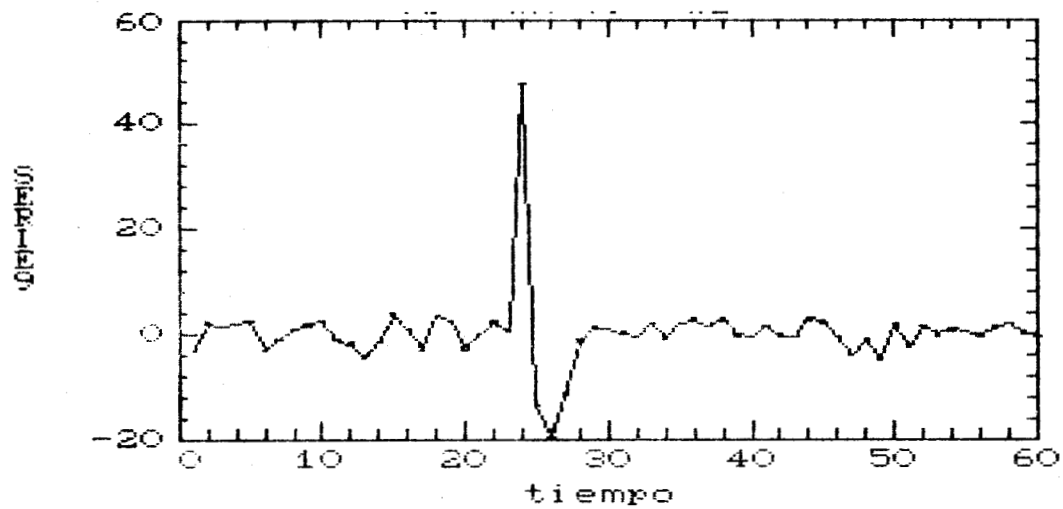
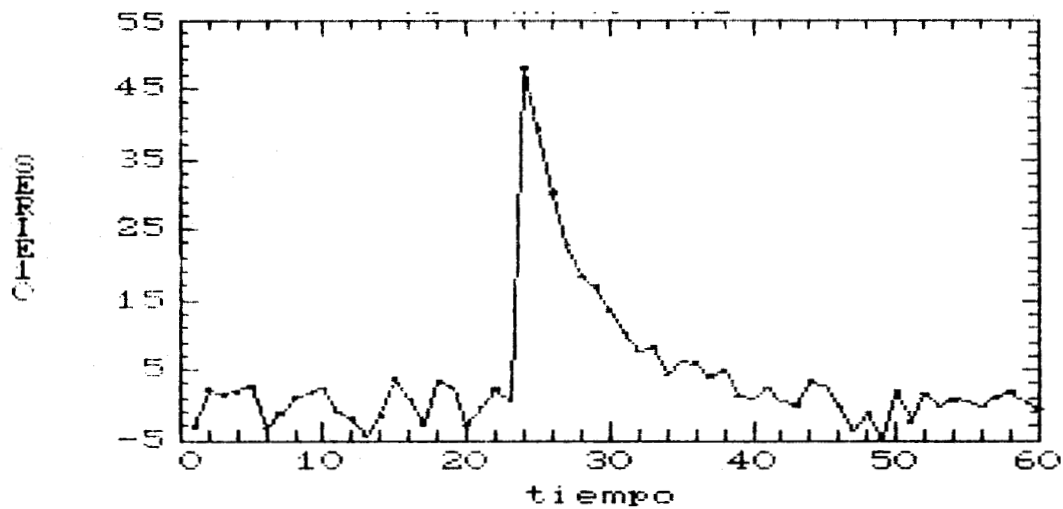


FIGURA 2.2.7

Serie de tiempo que presenta un cambio transitorio



2.3 Detección de puntos de cambio

En esta sección se expone un método de detección de puntos de cambio en series de tiempo, propuesto originalmente por Tsay (1987) dentro de otro contexto.

Varios autores han estudiado el problema de detectar puntos de cambio en series de tiempo, por ejemplo Box y Tiao (1965) desarrollan un método para detectar cambios de nivel en series de tiempo, con el inconveniente de que dicho método fue desarrollado únicamente para los casos particulares donde la serie de tiempo $\{x_t\}$ puede ser representada por los modelos IMA(1,1) o AR(1) y la generalización del método resulta ser algebraicamente muy complicada.

Posteriormente, Box y Tiao (1976) desarrollan un estadístico de prueba, para detectar cambios estructurales en series de tiempo, este estadístico puede estudiarse también en Guerrero (1983, Cap. 7). Existen otros trabajos relacionados con la detección de cambios en series de tiempo, como son los de Chernof y Zacks (1964), Gardner (1969), Jun (1989) y Tiao y Chen (1990).

Se eligió el método propuesto por Tsay (1987), por considerarse el más adecuado al propósito de este trabajo, debido a que en él se basan algunas ideas empleadas en el siguiente capítulo. La exposición de otros trabajos relacionados con la detección de cambios podría extenderse demasiado y opacar el propósito real del presente trabajo.

Considérese una serie de tiempo univariada $\{x_t\}$, la cual sigue un modelo ARMA(p,q)

$$\Phi(B)x_t = \Theta(B)a_t \quad (2.3.1)$$

Supóngase que todos los ceros de $\Phi(B)$ y de $\Theta(B)$ se encuentran

sobre o fuera del círculo unitario y que tales polinomios no tienen factores comunes; además, si alguno de los ceros de $\Phi(B)$ está sobre el círculo unitario, se supone que la serie comienza en un punto del tiempo fijo, con los primeros valores conocidos.

Ahora bien, en la serie pueden ocurrir cambios exógenos esperados o inesperados, por lo cual un modelo más realista para una serie de tiempo observada $\{Y_t\}$ sería

$$Y_t = U f(t) + x_t \quad (2.3.2)$$

donde $f(t)$ es una función que representa los cambios de x_t y U denota la magnitud de los mismos. Por simplicidad se considera el caso de un solo cambio, aunque el caso de varios cambios puede ser estudiado en la misma forma. Ahora supóngase que $f(t)$ es de la forma

$$f(t) = \frac{\omega(B)}{\delta(B)} \xi_t^{(v)} \quad (2.3.3)$$

donde $\xi_t^{(v)} = 1$ si y solamente si $t = v$, significando la ocurrencia de un cambio al tiempo v , $\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s$ y $\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r$ son polinomios de retraso de grados s y r respectivamente, siendo las ω 's y las δ 's parámetros fijos.

Algunos casos especiales de (2.3.3) son los siguientes:

Caso 1:

$$\frac{\omega(B)}{\delta(B)} = \frac{\Theta(B)}{\Phi(B)} \quad (2.3.4)$$

Este caso representa un PI en la serie $\{a_t\}$ y tiene un efecto dinámico en x_t , por propagación a través de las ponderaciones definidas por

$$\Pi(B) = \frac{\Phi(B)}{\Theta(B)}. \quad (2.3.5)$$

Entonces el punto aberrante innovado afecta a x_t para $t \geq v$ (ver figura 2.2.6).

Caso 2:

$$\frac{\omega(B)}{\delta(B)} = 1 \quad (2.3.6)$$

Este es el caso de un PA, el cual afecta únicamente a x_v (ver figura 2.2.5).

Caso 3:

$$\frac{\omega(B)}{\delta(B)} = \frac{1}{(1-B)} \quad (2.3.7)$$

Este es un modelo de cambio de nivel, el cambio ocurre al tiempo $t = v$ con magnitud U y es permanente (ver figura 2.2.2).

Caso 4:

$$\frac{\omega(B)}{\delta(B)} = \frac{1}{(1-\delta B)} \quad \text{con } 0 < \delta < 1 \quad (2.3.8)$$

Este modelo describe un efecto que afecta a muchos valores sucesivos de x_t , partiendo del tiempo v , pero el efecto disminuye gradualmente, de acuerdo con δ que denota la tasa de decaimiento. El cambio en este caso es transitorio, esto es, desaparece eventualmente (ver figura 2.2.7).

Aunque sólo se estudiarán estos casos, puede tomarse una combinación de ellos para modelar otros patrones de respuesta.

Se desea ahora hacer inferencia acerca del parámetro de cambio

U, lo cual puede hacerse usando la técnica de regresión lineal simple, como se verá a continuación.

Defínase

$$y_t = \frac{\Phi(B)}{\Theta(B)} Y_t$$

y

$$z_t = \frac{\Phi(B)}{\Theta(B)} f(t)$$

$$= \frac{\Phi(B)}{\Theta(B)} \frac{\omega(B)}{\delta(B)} \xi_t^{(v)}$$

entonces

$$\frac{\Theta(B)}{\Phi(B)} y_t = U \frac{\omega(B)}{\delta(B)} \xi_t^{(v)} + \frac{\Theta(B)}{\Phi(B)} a_t$$

luego

$$y_t = U \frac{\Phi(B)}{\Theta(B)} \frac{\omega(B)}{\delta(B)} \xi_t^{(v)} + a_t$$

por lo tanto se obtiene

$$y_t = U z_t + a_t \quad (2.3.9)$$

Utilizando el método de Mínimos Cuadrados Ordinarios para estimar U, se tiene que

$$U = \frac{\sum_{t=1}^n y_t z_t}{\sum_{t=1}^n z_t^2} \quad (2.3.10)$$

y

$$\text{var}[U] = \hat{\sigma}_a^2 \left[\sum_{t=1}^n z_t^2 \right]^{-1} \quad (2.3.11)$$

donde n es el tamaño de la muestra y $\hat{\sigma}_a^2$ es el error cuadrático medio.

Usando esta técnica se obtiene la estimación de U para los cuatro casos mencionados anteriormente, esto es, para el caso 1 se tiene

$$\begin{aligned} \hat{U}_{I,v} &= \frac{\sum_{t=1}^n y_t \frac{\Phi(B) \omega(B)}{\Theta(B) \delta(B)} \xi_t^{(v)}}{\sum_{t=1}^n \left[\frac{\Phi(B) \omega(B)}{\Theta(B) \delta(B)} \xi_t^{(v)} \right]^2} \\ &= \frac{\sum_{t=1}^n y_t \xi_t^{(v)}}{\sum_{t=1}^n \left[\xi_t^{(v)} \right]^2} \\ &= y_v \end{aligned} \quad (2.3.12)$$

con varianza

$$\begin{aligned} \text{var} \left[\hat{U}_{I,v} \right] &= \hat{\sigma}_a^2 \left[\sum_{t=1}^n \left[\xi_t^{(v)} \right]^2 \right]^{-1} \\ &= \hat{\sigma}_a^2 \end{aligned} \quad (2.3.13)$$

Para el caso 2

$$\hat{U}_{A,v} = \frac{\sum_{t=1}^n y_t \frac{\Phi(B)}{\Theta(B)} \xi_t^{(v)}}{\sum_{t=1}^n \left[\frac{\Phi(B)}{\Theta(B)} \xi_t^{(v)} \right]^2}$$

$$= \frac{\sum_{l=0}^{n-v} \pi_l y_{v+l}}{\pi_0^2 + \pi_1^2 + \dots + \pi_{n-v}^2}$$

donde π_l es el coeficiente de $\xi_l^{(v)}$, esto es,

$$\Pi(B) = \frac{\Phi(B)}{\Theta(B)} = \pi_0 - \pi_1 B - \pi_2 B^2 - \dots$$

de donde $\Pi(B)\xi_0 = 0$, $\Pi(B)\xi_1 = 0$, ..., $\Pi(B)\xi_v = \pi_0$, $\Pi(B)\xi_{v+1} = \pi_1$,
..., $\Pi(B)\xi_n = \pi_{n-v}$

y

$$\sum_{t=1}^n \left[\frac{\Phi(B)}{\Theta(B)} \xi_t^{(v)} \right]^2 = \pi_0^2 + \pi_1^2 + \dots + \pi_{n-v}^2$$

Sea

$$\rho_{A,v}^2 = \left(\pi_0^2 + \pi_1^2 + \dots + \pi_{n-v}^2 \right)^{-1} \quad (2.3.14)$$

entonces

$$\hat{U}_{A,v} = \rho_{A,v}^2 \sum_{l=0}^{n-v} \pi_l y_{v+l} \quad (2.3.15)$$

y

$$\text{var} \left[\hat{U}_{A,v} \right] = \rho_{A,v}^2 \hat{\sigma}_a^2 \quad (2.3.16)$$

Para el caso 3 se tiene que

$$\hat{U}_{N,v} = \rho_{N,v}^2 \sum_{i=0}^{n-v} \eta_i y_{v+1} \quad (2.3.17)$$

y

$$\text{var} \left(\hat{U}_{N,v} \right) = \rho_{N,v}^2 \hat{\sigma}_a^2 \quad (2.3.18)$$

donde η_i es el coeficiente de B^i en el polinomio

$$\eta(B) = \Pi(B) (1-B)^{-1} \quad (2.3.19)$$

y

$$\rho_{N,v}^2 = \left[\eta_0^2 + \eta_1^2 + \dots + \eta_{n-v}^2 \right]^{-1} \quad (2.3.20)$$

En el caso 4 se tiene que

$$\hat{U}_{T,v} = \rho_{T,v}^2 \sum_{i=0}^{n-v} \beta_i y_{v+1} \quad (2.3.21)$$

y

$$\text{var} \left(\hat{U}_{T,v} \right) = \rho_{T,v}^2 \hat{\sigma}_a^2 \quad (2.3.22)$$

donde β_i es el coeficiente de B^i en el polinomio

$$\beta(B) = \Pi(B) (1-\delta B)^{-1}, \quad (2.3.23)$$

$$\rho_{T,v}^2 = \left[\beta_0^2 + \beta_1^2 + \beta_2^2 + \dots + \beta_{n-v}^2 \right]^{-1} \quad (2.3.24)$$

y δ es una constante preespecificada que satisface $0 < \delta < 1$.

Basándose en los resultados anteriores, se pueden construir estadísticos para probar la existencia de un cambio al tiempo v . Las hipótesis que se plantean son las siguientes:

- H_0 : no existe cambio en x_v ,
- H_I : existe un punto aberrante innovado en x_v ,
- H_A : existe un punto aberrante aditivo en x_v ,
- H_N : existe un cambio de nivel en x_v ,
- H_T : existe un cambio de nivel transitorio en x_v .

Los estadísticos son

$$\lambda_{I,v} = \frac{\hat{U}_{I,v}}{\hat{\sigma}_a}, \quad (2.3.25)$$

$$\lambda_{A,v} = \frac{\hat{U}_{A,v}}{\left[\rho_{A,v} \hat{\sigma}_a \right]}, \quad (2.3.26)$$

$$\lambda_{N,v} = \frac{\hat{U}_{N,v}}{\left[\rho_{N,v} \hat{\sigma}_a \right]}, \quad (2.3.27)$$

y

$$\lambda_{T,v} = \frac{\hat{U}_{T,v}}{\left[\rho_{T,v} \hat{\sigma}_a \right]}, \quad (2.3.28)$$

Bajo H_0 , todas los estadísticos se distribuyen como $N(0,1)$.

Aplicando el modelo ARMA a x_t , el tipo de cambio y el punto en

el tiempo v son desconocidos, por lo cual Tsay (1987) propone el siguiente proceso iterativo para detectar y modelar disturbios.

Paso 1. Suponer tentativamente que no existen perturbaciones y especificar un modelo ARMA para la serie $\{Y_t\}$, estimar los parámetros y obtener los residuales.

Paso 2. Usando el modelo del paso 1, calcúlese los estadísticos 2.3.4 a 2.3.7 para todo punto en el tiempo t . Entonces, para cada estadístico, localícese su máximo en valor absoluto e identifíquese el correspondiente punto en el tiempo. Por ejemplo, considérese el caso 1 y defínase

$$\lambda_{I,\max} = \max_{1 \leq t \leq n} \left\{ \left| \lambda_{I,t} \right| \right\},$$

además, sea v_I el punto en el tiempo donde $\lambda_{I,\max}$ ocurre. De la misma forma se definen $\lambda_{A,\max}$ y v_A , $\lambda_{N,\max}$ y v_N , $\lambda_{T,\max}$ y v_T .

Paso 3. Compárese cada uno de estos máximos con el valor crítico dado en las tablas de la curva normal estándar. Si todos los máximos son menores que este valor, el proceso iterativo termina y no se rechaza la hipótesis H_0 . En otro caso, una perturbación se detecta. En este caso, el tipo y tiempo de cambio se determina tomando aquél que proporcione el máximo de

$$\left\{ \lambda_{I,\max}, \lambda_{A,\max}, \lambda_{N,\max}, \lambda_{T,\max} \right\}.$$

Paso 4. Remuévase el efecto de la perturbación identificada, substrayendo de Y_t el impacto dinámico, calculando $\hat{U}_f(t)$ y regrésese al paso 1.

CAPITULO 3. METODOS PARA EL PRONOSTICO DE PUNTOS DE CAMBIO

En este capítulo se expondrán tres métodos para pronosticar puntos de cambio. El primer método que se estudiará se presenta en la primer sección de este capítulo y es el desarrollado por Wecker (1979) para series de tiempo univariadas, el cual fue generalizado por Kling (1987) al caso de series de tiempo múltiples; este método consiste en pronosticar puntos de cambio cumbre y abismo a partir de un sistema de predicción frecuencial, generado a partir de la función de distribución predictiva de la serie. Se propone además en esta sección, una generalización del caso univariado para el pronóstico de puntos aberrantes innovados y aditivos, cambios de nivel y cambios transitorios; esta generalización se basa en el método de detección propuesto por Tsay (1987), que fué mostrado en la tercera sección del capítulo 2.

El segundo método será estudiado en la sección 2 y se desprende del artículo de Leiva (1987), mediante este método puede determinarse la probabilidad de que en cierto intervalo de tiempo ocurran determinado número de cambios de nivel.

En la tercera sección se desarrolla un método que surge del trabajo de Zellner (1987) y que proporciona el enfoque Bayesiano al problema de pronosticar puntos de cambio. Con este método puede obtenerse la probabilidad de un cambio basándose, como es de esperarse en un enfoque Bayesiano, en la distribución a priori del vector de parámetros del modelo generado por la serie.

Para concluir este capítulo, se presenta el trabajo realizado por Blanco y Garber (1986), quienes desarrollan un modelo con el fin de predecir el momento y magnitud de devaluaciones forzadas por ataques especulativos sobre sistemas de tipo de cambio fijo; analizando en particular el tipo de cambio mexicano en el periodo 1973-1982. Aunque en este trabajo no se plantea un método que pueda ser empleado en general para el pronóstico de puntos de cambio, es

importante dedicarle una sección ya que éste puede ser tomado como ejemplo para el desarrollo de modelos y pronóstico de cambios en serie que puedan ser estudiadas en forma semejante.

Existen otros trabajos relacionados con el tema, como son el de Cho y Byung (1981) acerca del análisis de series de tiempo estacionales no estacionarias, el de McLaughlin (1982) sobre un modelo de recesión y recuperación, el de Makridakis (1978) sobre el análisis y pronóstico en series de tiempo y el de Tiao (1990) con aproximaciones ARIMA al problema de pronóstico; sin embargo estos trabajos no plantean ningún método para el pronóstico de puntos de cambio. Lo que se plantea en dichos trabajos es encontrar un modelo que logre detectar con gran precisión los momentos en que un cambio ocurre.

3.1. Sistema de predicción frecuencial

En primer lugar se estudia el caso más simple, que fue propuesto por Wecker (1979) y en el cual $\{x_t\}$ es una serie de tiempo univariada. La definición de punto de cambio que se emplea es un caso particular de la definición 2.2.1.

Sean $\{x_1, x_2, \dots, x_n\}$ los valores observados de una serie de tiempo y sea $\{z_t\}$ una sucesión de variables indicadoras tales que

$$z_t = \begin{cases} 1 & \text{si un PCC ocurre al tiempo } t \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.1)$$

Usando el caso de la definición 2.2.1, en el cual $\{x_t\}$ es una serie de tiempo univariada y $\tau = 1$, se tiene que

$$z_t = \begin{cases} 1 & \text{si } x_t > x_{t+1} \quad \text{y} \quad x_t > x_{t-1} \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.2)$$

En este caso, si se observaron n valores de la serie $\{x_t\}$, sólo se pueden obtener $n-2$ valores de z_t . Los valores futuros x_{n+1} , x_{n+2} , ..., x_{n+k} son inciertos y esta incertidumbre se puede expresar mediante la distribución predictiva (ver apéndice A)

$$g_{n+1, n+2, \dots, n+k}(x_{n+1}, x_{n+2}, \dots, x_{n+k} | x_n, x_{n-1}, \dots, x_1) \quad (3.1.3)$$

Box y Jenkins (1970) prueban que, bajo el supuesto de que $(x_1, x_2, \dots, x_n)$ y $(x_{n+1}, \dots, x_{n+k})$ son variables aleatorias dependientes normalmente distribuidas, la distribución predictiva (3.1.3) es normal multivariada con vector de medias

$$\mu_{21} = \mu_2 + \Sigma'_{12} \Sigma_{11}^{-1} (x_1 - \mu_1)$$

y matriz de varianzas y covarianzas

$$\Sigma_{2211} = \Sigma_{22} - \Sigma'_{12} \Sigma_{11}^{-1} \Sigma_{12}$$

donde

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{y} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{bmatrix}$$

son, respectivamente el vector de medias y la matriz de varianzas y covarianzas del vector

$$x' = [x'_1, x'_2], \text{ con } x'_1 = [x_1, x_2, \dots, x_n] \text{ y } x'_2 = [x_{n+1}, \dots, x_{n+k}].$$

Los valores futuros de la sucesión indicadora son función de

los valores aleatorios $x_{n+1}, x_{n+2}, \dots, x_{n+k}$, por lo cual los valores futuros de z_t son asimismo aleatorios. Esta aleatoriedad se expresa con una distribución predictiva de variables indicadoras

$$f_{n, n+1, \dots, n+k-1}(z_n, z_{n+1}, \dots, z_{n+k-1} | x_n, x_{n-1}, \dots, x_1) \quad (3.1.4)$$

Para estimar esta distribución predictiva se utiliza el siguiente método numérico:

a) Generar k números pseudoaleatorios con la distribución predictiva estimada para $x_{n+1}, x_{n+2}, \dots, x_{n+k}$, denótese el resultado obtenido por $x_{n+1}^{(1)}, x_{n+2}^{(1)}, \dots, x_{n+k}^{(1)}$. En el apéndice B se proporcionan algunos métodos, que se utilizan para generar números pseudoaleatorios con distintas distribuciones.

b) A partir de la expresión 3.1.2, de los valores observados x_{n-1} y x_n y de los valores calculados $x_{n+1}^{(1)}, x_{n+2}^{(2)}, \dots, x_{n+k}^{(1)}$, calcular los valores de las variables indicadoras $z_n, z_{n+1}, \dots, z_{n+k-1}$.

c) Repetir los pasos a) y b) N veces, para encontrar la sucesión indicadora $z_n^{(N)}, z_{n+1}^{(N)}, \dots, z_{n+k-1}^{(N)}$. El número N de repeticiones, se elige de acuerdo al criterio del analista; Wecker por ejemplo, utilizó $N=1000$.

Una generalización de este método fué propuesta por Kling (1987) y es la que proporciona predicciones de puntos de cambio en series de tiempo multivariadas. Este método se expone a continuación.

Sean $x_t = (x_{1t}, x_{2t}, \dots, x_{mt})$, $t=1, 2, \dots, n$ los valores observados de la serie de tiempo vectorial $\{x_t\}$, y sean $z_t^C = (z_{1t}^C, z_{2t}^C, \dots, z_{mt}^C)$ y $z_t^A = (z_1^A, z_2^A, \dots, z_m^A)$ las sucesiones de variables indicadoras tales que:

$$z_{it}^C = \begin{cases} 1 & \text{si un PCC ocurre al tiempo } t, \text{ en la serie } i \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.5a)$$

y

$$z_{it}^A = \begin{cases} 1 & \text{si un PCA ocurre al tiempo } t, \text{ en la serie } i \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.5b)$$

Entonces, al usar la definición 2.2.1 se tiene que

$$z_{it}^C = \begin{cases} 1 & \text{si } x_{it} > x_{i,t+k}, k = -\tau, -\tau+1, \dots, -1, 1, \dots, \tau-1, \tau \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.6a)$$

y

$$z_{it}^A = \begin{cases} 1 & \text{si } x_{it} < x_{i,t+k}, k = -\tau, -\tau+1, \dots, -1, 1, \dots, \tau-1, \tau \\ 0 & \text{en otro caso.} \end{cases} \quad (3.1.6b)$$

Claramente, estas sucesiones de variables indicadoras son aleatorias, ya que están en función de los valores futuros (aleatorios) de la serie de tiempo $\{x_t\}$. La incertidumbre en $\{x_t\}$ y en $\{Z_t\}$, puede ser caracterizada en términos de sus funciones de distribución predictivas, esto es

$$g_{t+1, \dots, t+k}(x_{t+1}, \dots, x_{t+k} | x_t, x_{t-1}, \dots, x_1) \quad (3.1.7)$$

$$f_{t-\tau+1, \dots, t+k-\tau}^C(z_{t-\tau+1}^C, \dots, z_{t+k-\tau}^C | x_t, x_{t-1}, \dots, x_1) \quad (3.1.8a)$$

$$f_{t-\tau+1, \dots, t+k-\tau}^A(z_{t-\tau+1}^A, \dots, z_{t+k-\tau}^A | x_t, x_{t-1}, \dots, x_1) \quad (3.1.8b)$$

A continuación se estudiarán las distribuciones de probabilidad de los eventos $W_t^C = (W_{1t}^C, \dots, W_{mt}^C)'$, que representan el tiempo hasta que el siguiente PCC ocurra en cada serie, y $W_t^A = (W_{1t}^A, \dots, W_{mt}^A)'$ el tiempo hasta que el siguiente PCA ocurra. Estos vectores son aleatorios y se definen por

$$W_{it}^C = w(z_{1,t-\tau+1}^C, z_{1,t-\tau+2}^C, \dots) = k, \quad \text{para } i = 1, \dots, m \quad (3.1.9a)$$

donde k es un entero tal que $z_{1,t+k}^C = 1$ y $z_{1,t+j}^C = 0$ para toda $j < k$; asimismo

$$W_{it}^A = w(z_{1,t-\tau+1}^A, z_{1,t-\tau+2}^A, \dots) = k, \quad \text{para } i = 1, \dots, m \quad (3.1.9b)$$

donde k es tal que $z_{1,t+k}^A = 1$ y $z_{1,t+j}^A = 0$ para toda $j < k$.

W_{it}^C y W_{it}^A tienen entonces una distribución de probabilidad discreta, dentro del intervalo de números enteros

$$\{-\tau+1, -\tau+2, \dots, -1, 0, 1, 2, \dots\},$$

luego la predicción al tiempo n con los valores conocidos $\{x_1, x_2, \dots, x_n\}$, proporciona distribuciones de probabilidad P_{in}^C y P_{in}^A respectivamente, para los eventos aleatorios W_{in}^C y W_{in}^A , con $i = 1, 2, \dots, m$.

DEFINICION-3.1.1. A la regla que asocia una realización de $P_n^C = \{P_{1n}^C, \dots, P_{mn}^C\}$ y $P_n^A = \{P_{1n}^A, \dots, P_{mn}^A\}$ con cada valor de n y con todo posible conjunto de observaciones $\{x_1, x_2, \dots, x_n\}$, se le denomina Sistema de Predicción Frecuencial (SPF).

El SPF que se emplea para generar las distribuciones de probabilidad de las W_t^C y W_t^A , está dado por un método numérico que se describe más adelante. Mientras que la estructura de serie de tiempo múltiple que se supondrá para el vector aleatorio x_t es, sin mucha

pérdida de generalidad, un VAR.

Un VAR para $\{x_t\}$, combinado con las suposiciones distribucionales de los errores $\{a_t\}$, proporciona entonces la información necesaria para estimar la función de distribución predictiva de $x_{t+1}, \dots, x_{t+k}$. Para estimar las distribuciones predictivas de z_t^C y z_t^A , (3.1.8a) y (3.1.8b), así como las distribuciones de probabilidad P_t^C y P_t^A sobre los eventos W_t^C y W_t^A , se utiliza el siguiente método numérico:

- a) Para un conjunto dado de n observaciones, se generan k vectores pseudoaleatorios (ver apéndice B) con la función de distribución predictiva estimada para $x_{n+1}, \dots, x_{n+k}$. Denótese el resultado obtenido por $x_{n+1}^{(1)}, \dots, x_{n+k}^{(1)}$.
- b) Se calculan $z_{n-\tau+1}^C(1), \dots, z_{n-\tau+k}^C(1)$ y $z_{n-\tau+1}^A(1), \dots, z_{n-\tau+k}^A(1)$ con las ecuaciones (3.1.6a) y (3.1.6b) respectivamente, así como con los valores observados $x_{n-2\tau+1}, \dots, x_n$ y con los valores generados $x_{n+1}^{(1)}, \dots, x_{n-\tau+k}^{(1)}$.
- c) Se calculan $w_n^C(1)$ y $w_n^A(1)$ mediante las ecuaciones (3.1.9a) y (3.1.9b) respectivamente, empleando los valores $z_{n-\tau+1}^C(1), \dots, z_{n-\tau+k}^C(1)$ y $z_{n-\tau+1}^A(1), \dots, z_{n-\tau+k}^A(1)$.
- d) Se repiten los pasos a) a c) N veces, hasta crear las sucesiones $w_n^C(1), \dots, w_n^C(N)$ y $w_n^A(1), \dots, w_n^A(N)$, donde N es elegido según el criterio del analista.

Las distribuciones de probabilidad P_n^C y P_n^A se estiman entonces mediante la distribución frecuencial empírica de $w_n^C(j)$ y $w_n^A(j)$, $j = 1, 2, \dots, N$.

3.1.1. Calibración del SPF

Es necesario ahora, probar la suficiencia completa del SPF estimado, esto es, se debe probar que el SPF cumple con el principio frecuencial propuesto por Dawid (1984). Este principio se expone a continuación.

PRINCIPIO FRECUENCIAL: Sea P un SPF para X , $x = \{x_n\}$ la realización de X y $P = \{P_n\}$ la correspondiente distribución predictiva frecuencial producida por P . El principio frecuencial establece que toda estimación probabilística de x , depende de P únicamente a través de la sucesión P .

Dawid (1984) propone el criterio de calibración completa para probar que el SPF estimado, cumple con el principio frecuencial. El criterio de calibración completa para el caso binario (o sea, cuando $P_n = P_n(z_n=1) = P(z_n=1 \mid z_1, z_2, \dots, z_{n-1})$) requiere que un promedio de las P' s, sobre un subconjunto elegido, coincida al menos asintóticamente, con el correspondiente promedio de las z' s. Este criterio se justifica debido a que esta propiedad ocurre con probabilidad 1, lo cual fue probado por Dawid (1982).

Para variables aleatorias continuas, considérese la variable Y_t con función de distribución acumulativa F_t , es fácil ver que los cuantiles aleatorios

$$U_t = F_t(Y_t), \quad t = 1, \dots, n$$

son variables aleatorias independientes, uniformemente distribuidas en el intervalo $[0,1]$.

Antes de probar esto, es necesario recordar la definición de variable aleatoria uniformemente distribuida.

DEFINICION-3.1.2. Una variable aleatoria γ se encuentra

uniformemente distribuida en el intervalo $[0,1]$ si $\gamma \in [0,1]$ y $P(a \leq \gamma \leq b) = b-a$, esto es, la probabilidad de que γ tome un valor en el intervalo $[a,b] \subseteq [0,1]$ es igual a la longitud del intervalo.

Ahora bien, supóngase que $F_t(r) = a$. Entonces, dado que F_t es una función monótona creciente y $U_t = F_t(Y_t)$ es una variable aleatoria con valores en el intervalo $[0,1]$, se tiene que

$$F_t(Y_t) \leq a \quad \Leftrightarrow \quad Y_t \leq r$$

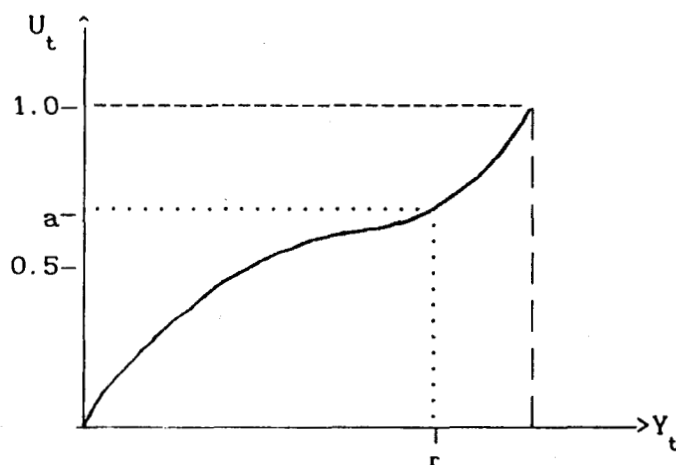
entonces,

$$P(F_t(Y_t) \leq a) = P(Y_t \leq r) = F_t(r) = a$$

La gráfica de la figura 3.1.1 ilustra este resultado.

FIGURA 3.1.1

Gráfica de una función de distribución acumulativa



En este caso, probar la suficiencia completa de un SPF se reduce a probar la hipótesis de que la sucesión observada esta

formada por variables $u_t = F_t(y_t)$, que se distribuyen $U[0,1]$. Si esta hipótesis no puede rechazarse, el SPF se considera bien calibrado.

Adviértase que los vectores aleatorios W_t^C y W_t^A no son continuos ni binarios. Entonces, si el investigador se interesa en la probabilidad de un punto de cambio dentro de k periodos adelante, W_{it}^C y W_{it}^A , $i = 1, \dots, m$ se distribuyen sobre el conjunto de enteros

$$\{-\tau+1, -\tau+2, \dots, -1, 0, -1, 2, \dots, k\}$$

pero su distribución es desconocida.

El criterio de calibración completa para eventos binarios puede aplicarse a eventos discretos de suceso múltiple. Esto es, las probabilidades asignadas a cada evento pueden compararse con las frecuencias de ocurrencia.

Sea $U_{it}^C = F_{it}^C(W_{it}^C)$ tal que F_{it}^C es la función de distribución acumulativa de P_{it}^C , por ejemplo, la probabilidad al tiempo t de que un PCC haya ocurrido en la serie i al momento q , está dada por

$$F_{it}^C(q) = \sum_{j=-\tau+1}^q P_{it}^C(j) \quad q < k$$

donde $P_{it}^C(j)$ es la probabilidad de un PCC en el momento j , esto es $\Pr(W_{it}^C=j)$; entonces U_{it}^C es una variable aleatoria con valores en el intervalo $[0,1]$. Una estimación de la función de distribución acumulativa $F_U(U_{it}^C)$ para U_{it}^C , se obtiene al considerar la sucesión observada

$$u_{it}^C = F_{it}^C(w_{it}^C), \quad t = 1, 2, \dots, n$$

y al calcular

104441

$$F_U \left[u_{it}^c(j) \right] = \frac{j}{n}, \quad j = 1, 2, \dots, n \quad (3.1.9)$$

si el SPF está bien calibrado, entonces

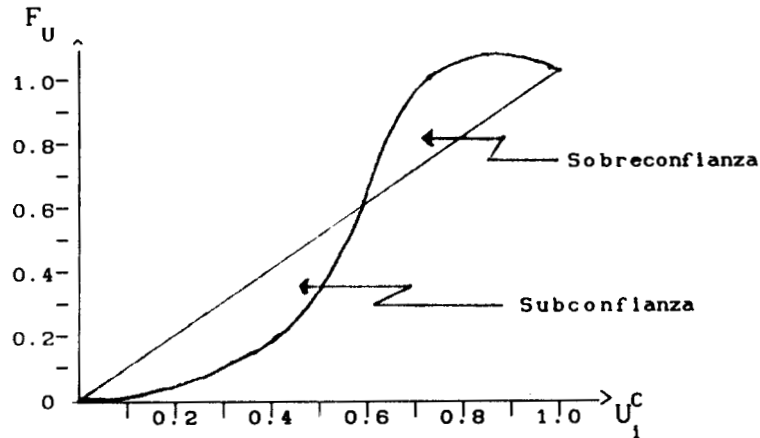
$$\Pr \left[u_{it}^c \leq \frac{j}{n} \right] = \frac{j}{n}, \quad j = 1, 2, \dots, n \quad (3.1.10)$$

De hecho, para todo SPF bien calibrado, las u_{it}^c deben comportarse, al menos asintóticamente, como variables aleatorias independientes $U[0,1]$. Nótese que la probabilidad dada en (3.1.7) es incondicionada y supone que no hay conocimiento de los valores actuales de u_{it}^c a priori, al tiempo t .

Entonces, para un SPF bien calibrado, la función de calibración debe verse como una línea de 45° que pasa por el origen. Sin embargo, en la figura 3.1.2 se aprecia que existe una región de sobreconfianza y una región de subconfianza.

FIGURA 3.1.2

Gráfica de una función de calibración



Existen dos métodos para determinar cuándo la sucesión u_{it}^c , $t = 1, 2, \dots, n$ proviene de un SPF bien calibrado. En primer lugar, se estiman dos medidas de sesgo para determinar la cercanía de la

curva de calibración a la línea de 45°. El sesgo absoluto promedio (SAP) se define como

$$SAP_i^c = \frac{1}{n} \sum_{j=1}^n \left| u_i^c(j) - \left[\frac{j}{n} \right] \right|$$

mientras que el sesgo promedio (SP) está dado por

$$SP_i^c = \frac{1}{n} \sum_{j=1}^n \left[u_i^c(j) - \left[\frac{j}{n} \right] \right]$$

Si el SPF está bien calibrado, las estadísticas SAP y SP son muy cercanas a cero. Sin embargo, si el SAP es grande, el SP indicará si en promedio el SPF está sobreconfiado o subconfiado.

Como un segundo método, la estadística de prueba Ji-cuadrada se usa para probar la hipótesis de que la sucesión $u_i^c(j)$, $j = 1, 2, \dots, n$ proviene de un SPF bien calibrado, es decir, se prueba que las u_i^c se distribuyen $U[0,1]$. Esto es, si el estadístico

$$\chi^2 = n \sum_{j=1}^n \frac{\left[u_i^c(j) - \left[\frac{j}{n} \right] \right]^2}{\left[\frac{j}{n} \right]}$$

al ser comparado con los puntos porcentuales de $\chi^2_{1-\alpha, n-1}$ no conduce al rechazo de la hipótesis de que las u_i^c siguen una distribución $U[0,1]$, entonces se concluye que el SPF está bien calibrado.

Nótese que $\{u_{it}^c\}$, $t = 1, 2, \dots, n$ no forma una muestra aleatoria en el sentido usual, debido al proceso secuencial mediante el cual ha sido determinada. No obstante, para una muestra grande, el número de u_{it}^c , cuyos valores caen en un subintervalo arbitrario de $[0,1]$, es igualmente probable que el número de valores que caen

en cualquier otro subintervalo arbitrario de igual longitud.

3.1.2. SPF para el pronóstico de puntos aberrantes y cambios de nivel.

El siguiente método se propone en este trabajo como una extensión del caso univariado del método anterior, para el pronóstico de puntos aberrantes aditivos e innovados, cambios de nivel y cambios transitorios. Considérese la serie de tiempo univariada $\{Y_t\}$ dada en la sección 2.3, esto es

$$Y_t = Uf(t) + x_t$$

donde

$$f(t) = \frac{\omega(B)}{\delta(B)} \xi_t^{(v)}$$

determina el tipo de cambio que se observa.

Recordando que $\xi_t^{(v)} = 1$ si y solamente si ocurre un cambio al tiempo v , se observa que se trata de una variable aleatoria indicadora, cuya aleatoriedad se expresa mediante su distribución predictiva.

$$f_{n,n+1,\dots,n+k} \left(\xi_{n+1}^{(v)}, \xi_{n+2}^{(v)}, \dots, \xi_{n+k}^{(v)} \mid Y_n, Y_{n-1}, \dots, Y_1 \right)$$

A continuación se proporciona un método numérico para estimar esta distribución predictiva en cada uno de los cuatro casos estudiados en la sección 2.3, para lo cual en lo sucesivo se usará la notación siguiente

$\xi_{I,t}^{(v)}$ es la variable que indica la ocurrencia de un punto aberrante innovado,

$\xi_{A,t}^{(v)}$ es la variable que indica la ocurrencia de un punto aberrante aditivo,

$\xi_{N,t}^{(v)}$ es la variable que indica un cambio de nivel

y

$\xi_{T,t}^{(v)}$ es la variable que indica un cambio de nivel transitorio.

A manera de ilustración se describe el método numérico para estimar la función de distribución predictiva para la variable indicadora de un punto aberrante innovado. Todos los otros casos son análogos. Los pasos a seguir son los siguientes:

- 1) Estimar un modelo ARMA para $\{Y_t\}$.
- 2) Generar k números pseudoaleatorios con la distribución predictiva estimada para $Y_{n+1}, Y_{n+2}, \dots, Y_{n+k}$.
- 3) Usando el modelo del paso 1, los valores generados en el paso 2 y los valores observados $\{Y_1, Y_2, \dots, Y_n\}$, calcular la estadística $\lambda_{I,t}$, dada en la sección 2.3, para toda $t=n+1, n+2, \dots, n+k$ y localizar el máximo de estos valores, o sea

$$\lambda_{I,\max} = \max \left\{ \lambda_{I,n+1}, \lambda_{I,n+2}, \dots, \lambda_{I,n+k} \right\}$$

sea éste,

$$\lambda_{I,v}, \quad \text{con } n+1 \leq v \leq n+k,$$

si $\lambda_{I,v}$ es mayor que el correspondiente valor crítico, se toma

$$\xi_{I,v}^{(1)} = 1 \quad \text{y} \quad \xi_{I,t}^{(1)} = 0 \quad \text{para toda } t \neq v, \quad n+1 \leq t \leq n+k$$

Si $\lambda_{I,v}^{(1)}$ es menor que el correspondiente valor crítico, sea

$$\xi_{I,t}^{(1)} = 0 \quad \text{para toda } n+1 \leq t \leq n+k$$

4) Repetir los pasos 2) y 3) N veces, para encontrar la sucesión indicadora

$$\xi_{I,n+1}^{(N)}, \xi_{I,n+2}^{(N)}, \dots, \xi_{I,n+k}^{(N)}$$

3.2 Pronóstico de cambios de nivel.

En esta sección se expone un método para pronosticar *Cambios de Nivel* en series de tiempo, el cual fue sugerido por Leiva (1987). Con este fin se separan las series en dos componentes, que son el nivel y el ruido. A estas series se les llama Series de Tiempo de Nivel más Ruido (STN+R), las cuales pueden considerarse como series de tiempo localmente estacionarias.

Existen tres componentes que rigen en forma separada la aleatoriedad presente en las STN+R:

1.- Un proceso de niveles subyacentes que se denota por $\{N_k^*\}_{k=0}^{\infty}$, cuya realización establece para cada tiempo, el nivel en el que puede estar localizada la realización de la STN+R.

2.- Un proceso de ruidos que se denota por $\{R_k\}_{k=1}^{\infty}$, cuya realización establece las perturbaciones con que se observa en cada tiempo, el nivel en que se encuentra la realización de la serie.

3.-Un proceso de tiempos de espera que se denota por $\{W_k\}_{k=1}^{\infty}$, cuya realización establece los tiempos de permanencia de la realización de la STN+R en los distintos niveles alcanzados.

Asociado con el último proceso, se encuentra el proceso de tiempos de cambio que se denota por $\{E_k\}_{k=0}^{\infty}$, y que se define por

$$E_0 = 0 \quad \text{y} \quad E_k = \sum_{i=1}^k W_i \quad \text{para } k = 1, 2, \dots$$

la realización de este proceso determina los momentos en los cuales se producen los cambios de nivel de la STN+R.

En la figura 3.2.1, se ilustra la realización de una STN+R, en la cual se pueden apreciar la componente de niveles subyacentes, la componente de tiempos de espera y la de tiempos de cambio. En la figura 3.2.2 se puede observar la componente de ruido.

FIGURA 3.2.1
Gráfica de una STN+R

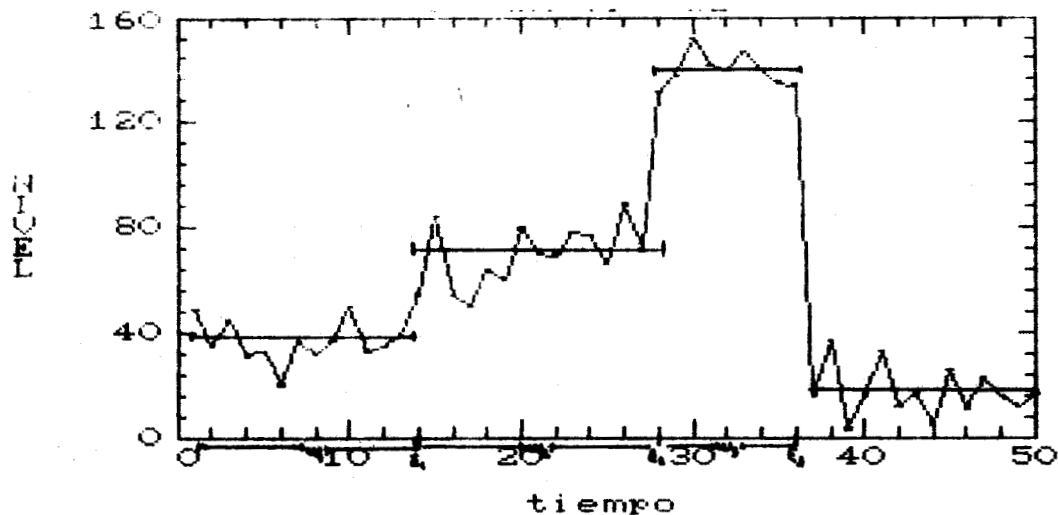
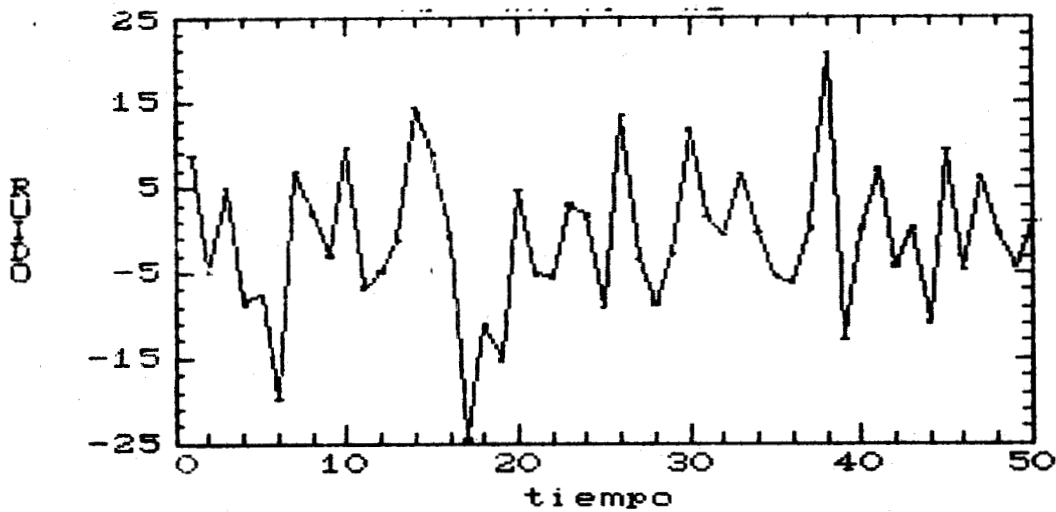


FIGURA 3.2.2

Grafica de la componente de ruido de la serie de la figura 3.2.1



DEFINICION-3.2.1: Sean $\{N_k^*\}_{k=1}^{\infty}$, $\{R_k\}_{k=1}^{\infty}$ y $\{W_k\}_{k=1}^{\infty}$ sucesiones de variables aleatorias tales que

1.- $\{W_k\}_{k=1}^{\infty}$ es una sucesión de renovación discreta, es decir $\{W_k\}$ es una sucesión de variables aleatorias no negativas, independientes e idénticamente distribuidas con valores enteros, tales que $W_k \sim F_W$ con $E[W] < \infty$ para $k = 1, 2, \dots$

2.- $E[N_k^*] = E[R_k] = 0$ para toda $k \in \mathbb{N}$.

Se dice entonces que $\{x_n\}_{n=1}^{\infty}$ es una STN+R discreta, si se verifica que x_n está definida para toda $n \in \mathbb{N}$ por

$$x_n = N_{k-1}^* + R_n \quad \text{si} \quad E_{k-1} \leq n \leq E_k \quad \text{donde } E_0 \equiv 0$$

y

$$E_k = \sum_{j=1}^k W_j \quad \text{para toda } k = 1, 2, \dots$$

DEFINICION-3.2.2. Dado que el nivel correspondiente a todos los x_n , cuyo valor de n es tal que $E_{k-1} \leq n < E_k$, es el mismo N_{k-1}^* , entonces se define un proceso de niveles reales $\{N_i\}_{i=0}^{\infty}$ como

$$N_n = N_{k-1}^*, \quad \text{si } E_{k-1} \leq n < E_k$$

por lo tanto

$$x_n = N_n + R_n$$

La siguiente definición tiene como fin simplificar el desarrollo teórico de esta sección.

DEFINICION-3.2.3. Dados $h, k \in \mathbb{N}$ con $h \leq k$ y dado v tal que $0 \leq v \leq k-h$, se dice que h y k son v -próximos, respecto al proceso de tiempo de cambio $\{E_k\}$, si y sólo si existe $j \in \mathbb{N}$ tal que $E_{j-1} \leq h < E_j$ y $E_{j+v-1} \leq k < E_{j+v}$.

Nótese que el hecho de que h y k sean v -próximos es equivalente a decir que en el intervalo de tiempo $[h, k]$ ocurren v cambios de nivel. Entonces tiene sentido hablar de la probabilidad de ocurrencia de que h y k sean v -próximos, ya que esto es un evento aleatorio.

Sea $R(i) = \sup \{m: E_m \leq i\}$, esto es, $R(i)$ cuenta el número de cambios que ocurren hasta el momento i , en consecuencia

$$R(i) = n \quad \Leftrightarrow \quad E_n \leq i < E_{n+1}.$$

Leiva (1987) estableció el teorema que se expone a continuación, en el cual se proporciona un método para pronosticar el número de cambios de nivel que pueden ocurrir en cierto intervalo de tiempo. Es decir, mediante el resultado que se obtiene en este teorema, se puede calcular la probabilidad de que ocurra un cierto

número de cambios en el intervalo de tiempo dado.

TEOREMA-3.2.1: Para $h, k, v \geq 0$,

$$P[h \text{ y } k \text{ sean } v\text{-próximos}] = P[R(k-h)=v]$$

entonces,

$$P[h, k \text{ sean } v\text{-próximos}] = \begin{cases} 1-F_w(k-h) & \text{si } v=0 \\ F_v(k-h)-F_{v+1}(k-h) & \text{si } v=1, 2, \dots \end{cases}$$

donde $F_v(t) = P[E_v \leq t]$

DEMOSTRACION: Para $h \leq k$ y $0 \leq v \leq k-h$

$$\begin{aligned} P[h, k \text{ sean } v\text{-próximos}] &= \sum_{i=1}^{h+1} P[E_{i-1} \leq h < E_i, E_{i+v-1} \leq k < E_{i+v}] \\ &= \sum_{i=1}^{h+1} P[R(h)=i-1, R(k)=i+v-1] \\ &= \sum_{i=1}^{h+1} P[R(h)=i-1, R(k)-R(h)=v] \\ &= \sum_{i=1}^{h+1} P[R(h)=i-1, R(k-h)=v] \\ &= P[R(k-h)=v] \sum_{i=1}^{h+1} P[R(h)=i-1] \end{aligned}$$

y dado que

$$\sum_{i=1}^{h+1} P[R(h)=i-1] = 1$$

entonces,

$$P[h, k \text{ sean } v\text{-próximos}] = P[R(k-h)=v].$$

Ahora bien,

$$\begin{aligned} P[R(k-h)=0] &= P[W_k > k-h] \\ &= 1 - P[W_k \leq k-h] \\ &= 1 - F_W(k-h) \end{aligned}$$

por lo tanto, si $v = 0$,

$$P[h, k \text{ sean } v\text{-próximos}] = 1 - F_W(k-h)$$

por otra parte, si $v \neq 0$ se quiere probar que

$$P[R(k-h)=v] = P[E_v < k-h] - P[E_{v+1} < k-h]$$

para esto obsérvese que

$$\begin{aligned} P[R(k-h)=v] &= P[E_v < k-h] P[E_{v+1} > k-h] \\ &= P[E_v < k-h][1 - P[E_{v+1} < k-h]] \\ &= P[E_v < k-h] - P[E_v < k-h] P[E_{v+1} < k-h] \end{aligned}$$

entonces, dado que E_v y E_{v+1} son variables aleatorias

independientes, se tiene que

$$\begin{aligned} P[E_v < k-h]P[E_{v+1} < k-h] &= P[E_v < k-h \text{ y } E_{v+1} < k-h] \\ &= P[E_{v+1} < k-h]. \end{aligned}$$

por lo tanto

$$\begin{aligned} P[R(k-h)=v] &= P[E_v < k-h] - P[E_{v+1} < k-h] \\ &= F_v(k-h) - F_{v+1}(k-h) \end{aligned}$$

Se ha probado que la probabilidad de que ocurran v cambios entre el momento h y el momento k , está dada por $1-F_w(k-h)$ si $v = 0$ y por $F_v(k-h) - F_{v+1}(k-h)$ si $v \neq 0$.

Este método podría emplearse también, para el pronóstico de puntos de cambio cumbre o abismo y puntos salto, ya que como se vió en la sección 2.2, un punto de cambio de este tipo, puede traducirse en un cambio de nivel, mediante la diferenciación de la serie. Entonces, pronosticar un cambio de nivel en la serie diferenciada, es equivalente a pronosticar un punto de cambio cumbre o abismo o un punto salto en la serie original. El método puede generalizarse además para el pronóstico de puntos aberrantes aditivos e innovados y para cambios transitorios.

3.3. Enfoque Bayesiano.

En esta sección se proporciona un enfoque Bayesiano, para lo cual se hará uso del análisis Bayesiano en general, el cual puede ser estudiado con detalle, por ejemplo en Broemeling (1985).

El siguiente análisis, propuesto por Zellner (1987), puede emplearse bajo cualquier definición que se adopte de PC, sin embargo, a manera de ilustración se empleará un caso particular de la definición 2.2.1 en el desarrollo de esta sección. De hecho, se usará la definición 2.2.1 con $\tau = 1$ y sin variables explicativas. Esto es, si $x_{t-1} < x_t$ y $x_t > x_{n+1}$, entonces un PCC ocurre al tiempo t . Y si $x_{t-1} < x_t$ y $x_t \leq x_{t+1}$ entonces no se presenta un PCC al tiempo t (NPCC). Si $x_{t-1} > x_t$ y $x_t < x_{t+1}$ entonces un PCA se observa al tiempo t . Y si $x_t \geq x_{t+1}$ entonces no se da un PCA al tiempo t (NPCA).

Dado un modelo generado por las observaciones $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$, sea $P(\mathbf{x} | \theta)$ la función de verosimilitud para los datos históricos, donde θ es el vector de parámetros que especifican el modelo, $\Pi(\theta | I)$ es la distribución a priori para θ y $P(\theta | D) \propto \Pi(\theta | I)P(\mathbf{x} | \theta)$ es la distribución a posteriori para θ , donde $D=(\mathbf{x}, I)$ representa el conjunto formado por la muestra y por la información a priori.

Entonces, si $P(x_{n+1} | \theta, D)$ representa la función de densidad de probabilidad (fdp) para x_{n+1} dado θ y las observaciones, la función de distribución predictiva de probabilidad (ver apéndice B) para x_{n+1} , está dada por

$$P(x_{n+1} | D) = \int_{\Theta} P(x_{n+1} | \theta, D) P(\theta | D) d\theta \quad -\infty < x_{n+1} < \infty \quad (3.3.1)$$

donde Θ es el espacio parametral tal que $\theta \in \Theta$.

La función de distribución predictiva de probabilidad (3.3.1) se emplea para obtener probabilidades asociadas con los eventos dados en la definición, por ejemplo, si $x_{n-1} < x_n$ la probabilidad de un PCC al tiempo n está dada por

$$P_C = P(x_{n+1} \leq x_n | D) = \int_{-\infty}^{x_n} P(x_{n+1} | D) dx_{n+1} \quad (3.3.2)$$

y la probabilidad de que no ocurra un PCC al tiempo n es $P_{NC} = 1 - P_c$.

Si $x_{n-1} < x_n$, entonces de la definición se sigue que la probabilidad de un PCA es $P_A = 0$. Una vez calculada la probabilidad de ocurrencia de un PCC, se debe decidir si se pronostica un PCC o si no se pronostica un PCC al tiempo n . Para esto, considérese la estructura de pérdida dada en la tabla (3.3.1).

T A B L A 3.3.1
Estructura de pérdida lineal para el pronóstico de PCC

Acción	Posible ocurrencia		Pérdida esperada
	C	NC	
Pronosticar C	0	K_1	$K_1 (1 - P_c)$
Pronosticar NC	K_2	0	$K_2 P_c$

Obsérvese que si

$$K_1 (1 - P_c) < K_2 P_c,$$

esto es, si

$$1 < \frac{K_2}{K_1} \left[\frac{P_c}{1 - P_c} \right] \quad (3.3.3)$$

o equivalentemente, si

$$P_c > \frac{K_1}{K_1 + K_2} \quad (3.3.4)$$

pronosticar un PCC produce pérdidas menores que pronosticar un NPCC. Nótese que si $K_1 = K_2$, la ecuación (3.3.4) conduce a predecir un PCC

si $P_c > 1/2$. Entonces, la decisión de predecir un Punto de Cambio es muy sensible al valor de K_2/K_1 .

Ahora se considera el problema conjunto de pronosticar PC's y obtener la predicción de los puntos. La estructura de pérdida que se utilizará a manera de ilustración, se muestra en la tabla 3.3.2, en donde $u = x_{n+1}$ y $\hat{u}$ representa la estimación de x_{n+1} .

T A B L A 3.3.2
Estructura de pérdida cuadrática para el pronóstico de PCC

Acción	Posible ocurrencia	
	C	NC
Pronosticar C	$K_1(\hat{u}-u)^2$	$K_2(\hat{u}-u)^2 + b_1$
Pronosticar NC	$K_1(\hat{u}-u)^2 + b_2$	$K_2(\hat{u}-u)^2$

Aquí se utilizan pérdidas cuadráticas, $K_1(\hat{u}-u)^2$ y $K_2(\hat{u}-u)^2$ para C y NC, además de los elementos de pérdida aditivos, b_1 y b_2 para reflejar errores en el pronóstico de PC's.

La pérdida esperada al pronosticar un punto cumbre al tiempo n está dada por

$$E(L_c) = P_c K_1 E[(\hat{u}-u)^2|C] + (1-P_c) \{K_2 E[(\hat{u}-u)^2|NC] + b_1\} \quad (3.3.5)$$

El valor de $\hat{u}$ que minimiza (3.3.5), sujeto a la condición $\hat{u} < x_n$, se denota por $\hat{u}_c$. Entonces (3.3.5) se evalúa en $\hat{u} = \hat{u}_c$. La pérdida esperada al pronosticar NPCC está dada por

$$E(L_{NC}) = P_c \{K_1 E[(\hat{u}-u)^2|C] + b_2\} + (1-P_c) K_2 E[(\hat{u}-u)^2|NC] \quad (3.3.6)$$

El valor de $\hat{u}$ que minimiza (3.3.6), sujeto a la condición $\hat{u} \geq x_n$, se denota por $\hat{u}_{NC}$. Se evalúa (3.3.6) en $\hat{u} = \hat{u}_{NC}$.

Entonces si

$$E(L_C)(\hat{u}=\hat{u}_C) < E(L_{NC})(\hat{u}=\hat{u}_{NC})$$

se pronostica un PCC y se usa $\hat{u}_C$ como valor predicho para x_{n+1} .
Alternativamente, si

$$E(L_C)(\hat{u}=\hat{u}_C) > E(L_{NC})(\hat{u}=\hat{u}_{NC})$$

se pronostica que no habrá un PCC y se utiliza $\hat{u}_{NC}$ como valor predicho para x_{n+1} .

El desarrollo de este método se ha llevado a cabo usando el caso particular de la definición 2.2.1 en el cual $\tau = 1$, sin embargo el método puede extenderse al caso en que $\tau = q$. Por ejemplo, para el caso $\tau = 2$, si $x_{n-2} < x_{n-1} < x_n$, la probabilidad de un punto cumbre al tiempo n está dado por

$$\begin{aligned} P_C &= P\left[x_n > x_{n+1} > x_{n+2} | D\right] \\ &= \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n+1}} P\left[x_{n+1}, x_{n+2} | D\right] dx_{n+2} dx_{n+1} \end{aligned} \quad (3.3.7)$$

suponiendo que

$$-\infty < x_{n+1} < \infty \quad \text{y} \quad -\infty < x_{n+2} < \infty.$$

Este procedimiento de pronóstico puede generalizarse para el caso de series de tiempo vectoriales. Dada una definición de punto de cambio y un modelo para series de tiempo vectoriales, la probabilidad de un punto de cambio puede calcularse al obtener la función de distribución de probabilidad predictiva conjunta.

Entonces, dada una estructura de pérdida se pueden obtener predicciones óptimas de puntos de cambio.

Este método se presenta entonces como alternativa al método expuesto en la sección 3.1, pero a diferencia de aquel, aquí se supone el conocimiento de la distribución a priori de los parámetros del modelo generado por las observaciones de la serie.

3.4. Ataques especulativos contra el peso mexicano.

En esta subsección se describe un modelo desarrollado por Blanco y Garber (1986) con el fin de predecir el momento y magnitud de las devaluaciones forzadas por ataques especulativos sobre sistemas de tipo de cambio fijo; en particular se analiza el tipo de cambio mexicano durante el periodo 1973-1982.

El modelo que aquí se presenta se basa en el siguiente hecho: Cuando los requerimientos para financiar tanto el déficit del gobierno como para apoyar el tipo de cambio exceden algún límite de los recursos del Banco Central, el tipo de cambio en vigencia será abandonado, esto es, cuando las reservas netas alcanzan un nivel crítico $\bar{R}$, medido en unidades de divisas internacionales, el Banco Central establece un nuevo tipo de cambio $\hat{e}_t$ usando una regla permanente de política económica; $\hat{e}_t$ debe ser siempre un tipo de cambio viable, esto requiere que $\hat{e}_t$ exceda algún valor mínimo, el cual será establecido mas adelante. El mercado de dinero es la componente principal del modelo que se establece en la siguiente forma:

$$m_t - p_t = \beta + \Omega y_t - \alpha i_t + w_t \quad (3.4.1)$$

donde m_t , p_t y y_t representan el logaritmo del acervo de dinero, del nivel de precios y del nivel de producto agrícola respectivamente,

i_t es la tasa de interés interna y w_t es el término aleatorio de demanda de dinero; asimismo se supone que el nivel de precios y la tasa de interés están dadas por

$$i_t = i_t^* + Ee_{t+1} + I^{-e}_t \quad (3.4.2)$$

$$p_t = p_t^* + e_t + u_t \quad (3.4.3)$$

donde el asterisco indica que se trata de una variable exógena. Por su lado, e_t y u_t son los logaritmos del tipo de cambio nominal y real, respectivamente. El operador E representa expectativas condicionales a través del tiempo. Sustituyendo (3.4.2) y (3.4.3) en (3.4.1) se obtiene

$$\tilde{b}_t = -\alpha E\tilde{e}_{t+1} + [1 + \alpha]\tilde{e}_t \quad (3.4.4)$$

donde

$$b_t = \log[D_t + \bar{R} \exp(\bar{e})] - \beta - \Omega y_t + \alpha i_t^* - p_t^* - u_t - w_t$$

D_t es el componente de crédito interno de la base monetaria en el tiempo t . $\bar{R}$ se convierte en moneda nacional usando el tipo de cambio fijo $\bar{e}$ vigente. El proceso estocástico que obedece la variable b_t se supone que es un proceso autorregresivo de primer orden, exógeno al tipo de cambio, esto es

$$b_t = \theta_1 + \theta_2 b_{t-1} + v_t \quad (3.4.5)$$

donde v_t es un proceso de ruido blanco con función de densidad normal $g(v)$, con media cero y desviación estándar σ . Resolviendo las ecuaciones en diferencia (3.3.4) y (3.4.5) se obtiene el tipo de cambio flexible

$$\tilde{e}_t = \mu\alpha\theta_1 + \mu b_t \quad (3.4.6)$$

donde

$$\mu = \frac{1}{[(1+\alpha)-\alpha\theta_2]} ,$$

lo cual se sigue de suponer que $\alpha\theta_2/(1 + \alpha) < 1$. El nuevo tipo de cambio se obtiene mediante

$$\hat{e}_t = \tilde{e}_t + \delta v_t \quad (3.4.7)$$

siendo δ un parámetro no negativo, entonces $\hat{e}_t$ es el tipo de cambio fijo nuevo, que se establecerá si el nivel de reservas alcanza el valor $\bar{R}$ en el tiempo t . Ahora bien, dado que es equivalente que e_t exceda la tasa fija actual, a que ocurra una devaluación en el tiempo t , la probabilidad de devaluación en el tiempo $t + 1$ basada en la información disponible al tiempo t es

$$P(\mu\alpha\theta_1 + \mu b_{t+1} + \delta v_{t+1} > \bar{e}) \quad (3.4.6)$$

donde $\bar{e}$ es el valor de la tasa fija en el tiempo t . Equivalentemente, la probabilidad de devaluación es

$$1 - F(k_t) = P(v_{t+1} > k_t) \quad (3.4.7)$$

donde

$$k_t = [1/(\mu + \delta)][\bar{e} - \mu\alpha\theta_1 - \mu(\theta_1 + \theta_2 b_t)]$$

y $F(k_t)$ es la función distribución acumulativa asociada con $g(v)$

Conociendo la función de densidad pueden formarse expectativas de los tipos de cambio futuros a partir del promedio del tipo de cambio actual y el tipo esperado a materializarse si una devaluación ocurre, ambos ponderados por sus respectivas probabilidades de ocurrencia, esto es

$$Ee_{t+1} = F(k_t)\bar{e} = [1 - F(k_t)] E [e_{t+1} | v_{t+1} > k_t]$$

de (3.4.7) se tiene que

$$E[e_{t+1} | v_{t+1} > k_t] = \mu\theta_1(1 + \alpha) + \mu\theta_2 b_t + [\mu + \delta]E[v_{t+1} | v_{t+1} > k_t]$$

donde

$$E[v_{t+1} | v_{t+1} > k_t] = \int_{k_t}^{\infty} \frac{v g(v)}{1-F(k_t)} dv$$

dado que $g(v)$ es una función de densidad normal, el pronóstico no condicional del tipo de cambio para $t + 1$ es:

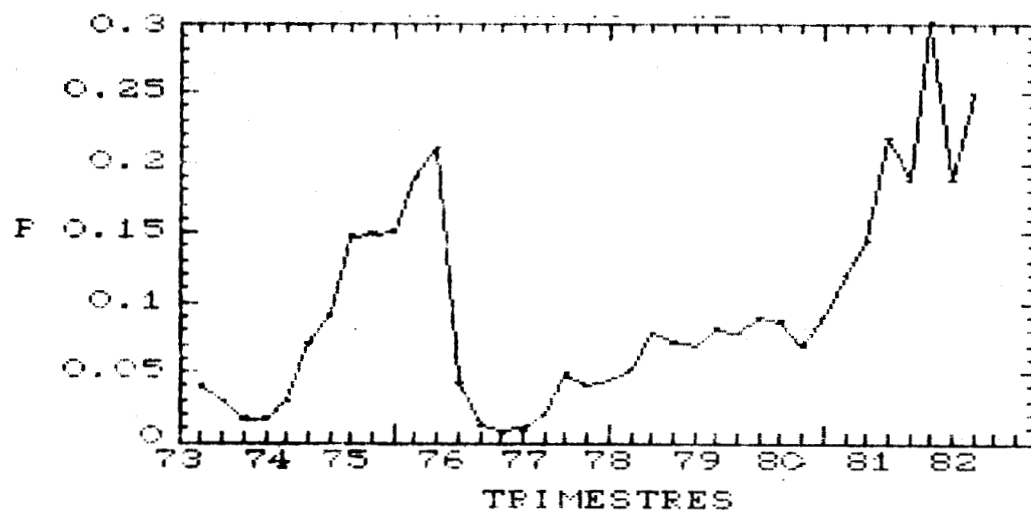
$$\begin{aligned} Ee_{t+1} &= F(k_t)\bar{e} + [1 - F(k_t)][\mu\theta_1(1 + \alpha) + \mu\theta_2 b_t] \\ &+ \frac{\sigma(\mu+\delta)}{(2\pi)^{1/2}} \exp[-.5(k_t/\sigma)^2] \end{aligned}$$

Blanco y Garber analizan con buenos resultados el tipo de cambio mexicano en el periodo 1973 a 1982; la figura 3.4.1 muestra las probabilidades obtenidas por ellos de que una devaluación ocurra en el siguiente periodo, las líneas verticales indican los periodos donde realmente ocurrió una devaluación; como puede observarse las probabilidades más altas se obtuvieron justamente en los periodos en que las devaluaciones ocurrieron.

Desafortunadamente, estos resultados no pueden ser reproducidos a partir de la información del artículo en donde se expone el método, debido a que no se proporcionan todos los datos necesarios.

FIGURA 3.4.1

Probabilidad de una devaluación en el siguiente periodo



104441

CAPITULO 4. APLICACIONES

Este capítulo tiene como objetivo mostrar la utilidad práctica de los conceptos descritos en los capítulos 2 y 3, para lo cual se realizó la simulación de dos series; la primera serie es univariada y se expone en la sección 4.1, esta serie presenta puntos de cambio cumbre y abismo, además en la serie diferenciada se observan cambios de nivel. En primer lugar se obtendrán los SPF para PCC y PCA empleando la definición 2.2.1 con $\tau = 1$ y $\tau = 2$, posteriormente se procederá a diferenciar la serie para detectar los cambios de nivel y obtener el correspondiente SPF para éstos, además se realizará el pronóstico de cambios de nivel empleando los métodos expuestos en las secciones 3.2 y 3.3.

En la sección 4.2 se simula una serie bivariada. en esta serie se aplica el método para pronóstico de puntos de cambio con series de tiempo multivariadas propuesto por Kling como una extensión del trabajo de Wecker.

En el desarrollo de las aplicaciones se emplean los paquetes de cómputo SCA y STATGRAPHICS, así como los programas que se muestran en los apéndices C, D, E y F.

4.1. Simulación univariada.

En esta sección se ilustran, a partir de una serie simulada, el método de detección visto en la sección 2.3 y los métodos de pronóstico desarrollados en el capítulo 3. La simulación de la serie se realizó como se indica a continuación.

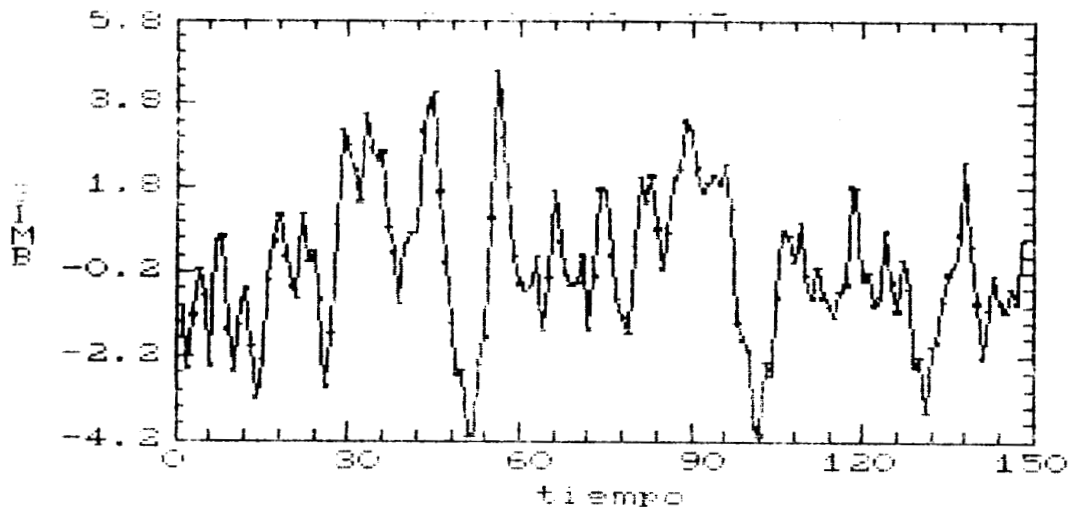
Mediante el paquete de cómputo SCA se generaron 148 valores aleatorios con distribución $N(0,1)$, los cuales se emplean como ruido blanco en el modelo siguiente

$$(1 - 0.6B)SIMB_t = (1 + 0.4B)a_t$$

A partir de este modelo se genera la serie $SIMB_t$, tomando como valor inicial $SIMB_0 = 0$. La realización de esta serie se muestra en la figura 4.1.1.

FIGURA 4.1.1

Serie simulada a partir del modelo $(1-0.6B)SIMB_t=(1+0.4B)a_t$



En esta serie se introducen ahorapuntos aberrantes aditivos, los cuales se obtienen al generar 10 tiempos de espera con distribución de Poisson y parámetro $\lambda = 15$. La magnitud U de los PA se obtiene eligiendo 10 números aleatorios uniformemente distribuidos dentro del intervalo $[-10,10]$ y denotados por M , los cuales se multiplican por el factor de escala $\sigma_{SIMB} = 1.78$. La tabla 4.1.1 muestra los tiempos de espera W , así como los tiempos de cambio E , definidos por

$$E_j = \sum_{i=1}^j W_i,$$

y la magnitud U de los puntos aberrantes generados.

TABLA 4.1.1

Tiempos de espera, tiempos de cambio y magnitud de los PAA.

W	5	19	12	16	14	12	8	27	17	12
E	5	24	36	52	66	78	86	113	130	142
M	7	-9	-1	3	8	-3	-3	8	-9	5
U	14.4	-16.0	-1.7	5.3	14.2	-5.3	-5.3	14.2	-16.0	8.9

La serie que se obtiene al introducir estos PAA se denota por $SIMPA_t$ y se muestra en la figura 4.1.2. En la figura 4.1.3 se ilustra la serie $SIMN_t$, la cual se obtiene de la integración de la serie $SIMPA_t$, tomando como valor inicial $SIMN_0 = 0$, de la integración de esta última se obtiene a su vez la serie SIM_t , donde $SIM_0 = 500$, esta serie se ilustra en la figura 4.1.4. Obsérvese que la serie $SIMN_t$ presenta cambios de nivel, los tiempos en los cuales ocurren estos cambios se obtendrán más adelante ya que es necesario primero identificar y estimar un modelo que describa el comportamiento de la serie, para así poder aplicar el método de detección de cambios visto en la sección 2.3. Al usar la definición 2.2.1 con $m = 1$ y $\tau = 1$ se encuentra que la serie SIM_t presenta PCC en los tiempos 13, 52, 135 y 142 y PCA en los tiempos 6, 45, 58, 138, 143. Considerando $\tau = 2$ se tienen los mismos PCC y PCA excepto para $t = 142$ y 143.

FIGURA 4.1.2

Serie simulada que presenta PA

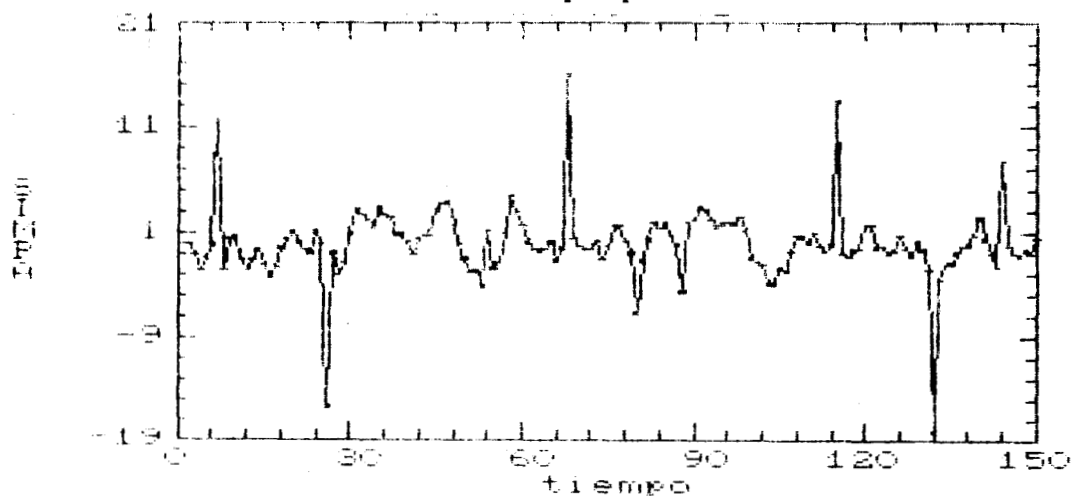


FIGURA 4.1.3
Serie que presenta cambios de nivel

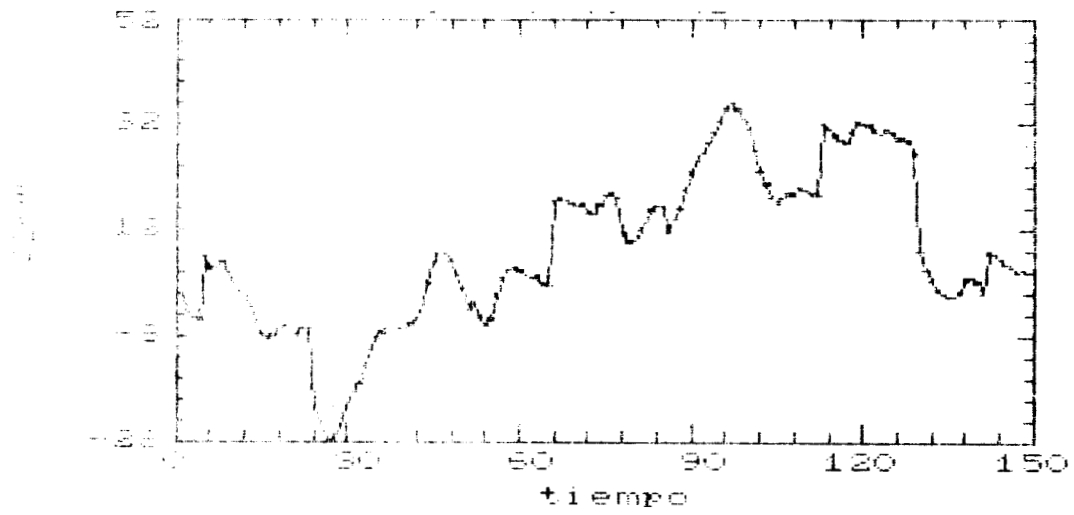
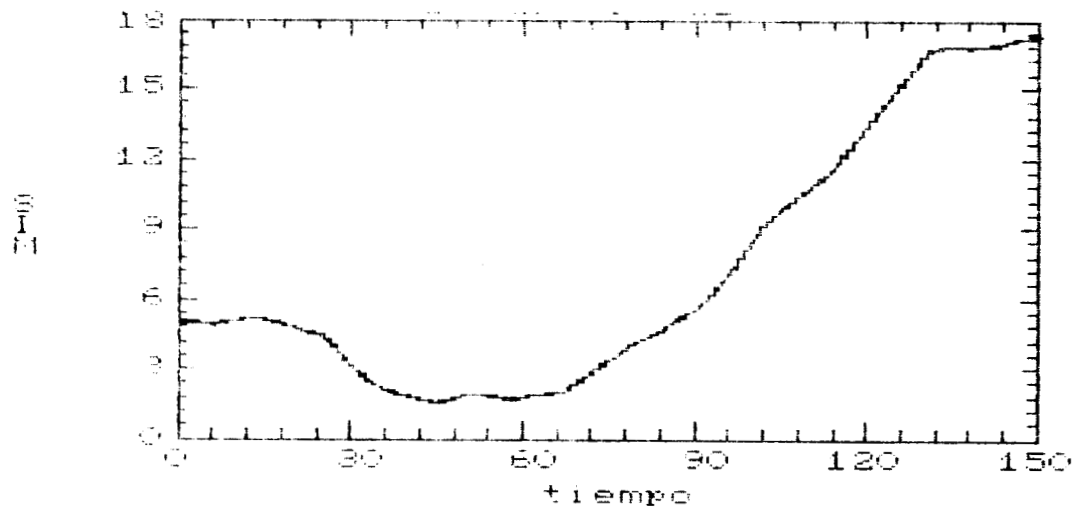


FIGURA 4.1.4
Serie que presenta PCC y PCA



Mediante la estrategia de construcción de modelos de Box y Jenkins y el método de la FAME (vease Islas (1989)) se identifica un modelo ARIMA(1,2,1) para la serie SIM_t , y el modelo estimado resulta ser

$$(1 - 0.4B)\nabla^2 SIM_t = (1 - 0.2B)a_t \quad 4.1.1$$

Nótese que este modelo estimado para $\nabla^2 SIM_t = SIMPA_t$ difiere del modelo a partir del cual fue simulada la serie $SIMB_t$, esto es, al introducir los PA se produce un cambio en los parámetros, que afecta principalmente al parámetro de promedios móviles, pero no afecta el orden del modelo.

Ahora pueden detectarse los cambios de nivel en la serie $SIMN_t$ empleando el método presentado en la sección 2.3 y el programa que se proporciona en el apéndice C. La tabla 4.1.2 muestra los tiempos en que ocurren cambios de nivel y la magnitud U de los mismos, con un nivel de significación $\alpha = 0.05$

T A B L A 4 . 1 . 2
*Tiempos de ocurrencia, tiempos de espera y magnitud
de cambios de nivel en la serie $SIMN_t$, con $\alpha = 0.05$*

t	6	25	67	79	87	114	131	143
W	6	19	42	12	8	27	17	12
U	11.71	-15.72	16.01	-6.61	-4.58	13.45	-18.25	7.83

Obsérvese que los tiempos de espera obtenidos son los mismos que los simulados, excepto el primero que, al igual que los tiempos de cambio, se encuentra desfasado una unidad debido a la integración, y el tercero, que está dado por la suma del tercero, cuarto y quinto de los simulados. Esto se debe a que los cambios que fueron simulados en los tiempos 36 y 52 con magnitudes $-1\sigma_x$ y $3\sigma_x$,

no resultaron significativos durante el proceso de detección.

Se desea ahora ilustrar los métodos de pronóstico vistos en la sección 3.1, para esto se consideran únicamente los últimos 50 valores de las series $\{SIM_t\}$ y $\{SIMN_t\}$, los cuales, para tener una visión más clara de lo que ocurre en este periodo, se muestran en las figuras 4.1.5 y 4.1.6 respectivamente. El hecho por el cual se estudian únicamente 50 observaciones es que en la aplicación de los métodos vistos en la sección 3.1 es necesario invertir la matriz de varianzas y covarianzas de la serie, cuya dimensión está dada por el número de observaciones y es difícil conseguir paquetes de cómputo que inviertan matrices demasiado grandes, además de que se le suma a esto la pérdida de precisión en los cálculos. Con el fin de mostrar las variaciones que sufre el SPF debido al conocimiento de un mayor o menor número de observaciones, el pronóstico se realiza eligiendo tres diferentes puntos de partida: $n = 136, 138$ y 140 para la serie $\{SIM_t\}$ y $n = 135, 137$ y 139 para la serie $\{SIMN_t\}$.

FIGURA 4.1.5
Serie $\{SIM_t\}$, para el periodo 100-150

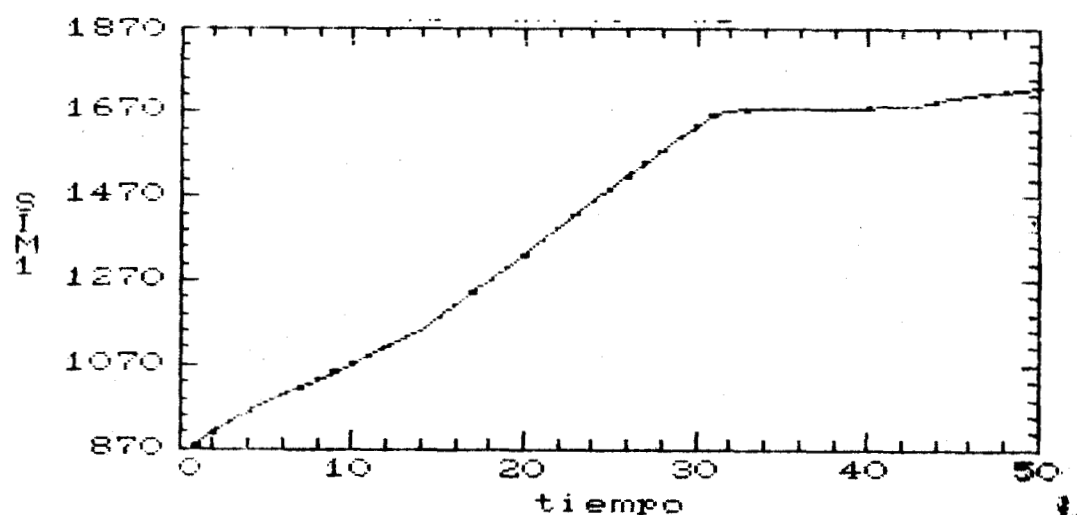
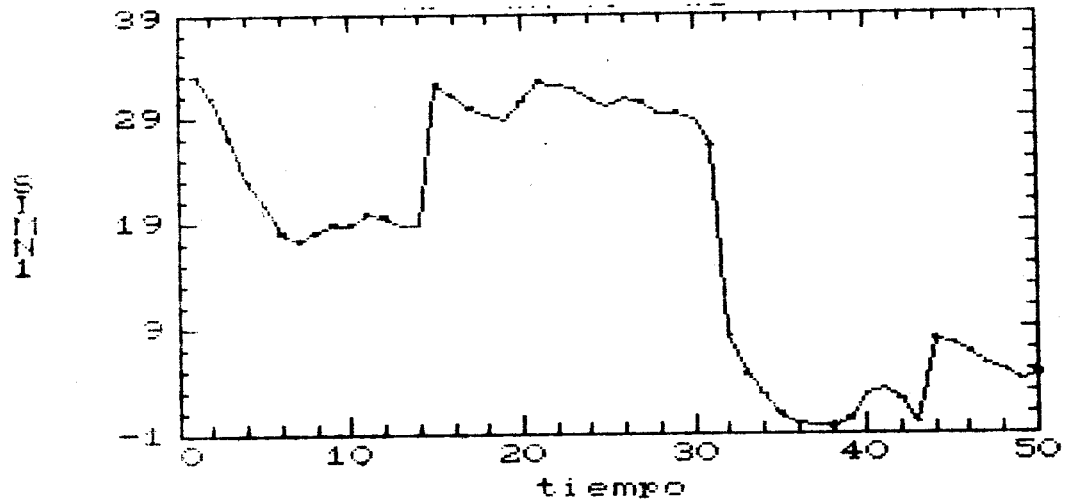


FIGURA 4.1.6
Serie (SIM_t) , para el periodo 99-149



Para obtener los SPF de la serie $\{SIM_t\}$, deben determinarse primero las sucesiones $SIM_{n+1}^{(N)}$, $SIM_{n+2}^{(N)}$, ..., $SIM_{n+k}^{(N)}$, generadas a partir de la distribución predictiva 3.1.3, para lo cual es necesario determinar

$$\mu_{21} = \mu_2 - \Sigma_{12}' \Sigma_{11}^{-1} (SIM_1 - \mu_1)$$

y

$$\Sigma_{2211} = \Sigma_{22} + \Sigma_{12}' \Sigma_{11}^{-1} \Sigma_{12}$$

donde $SIM_1 = (SIM_{100}, SIM_{101}, \dots, SIM_n)'$. Por su lado, Σ_{11} puede construirse a partir de las autocorrelaciones, teniendo en cuenta que

$$\gamma_k = \text{cov}(SIM_t, SIM_{t+k}) = \gamma_0 \rho_k$$

donde γ_0 denota la varianza de la serie y ρ_k la autocorrelación de

SIM_t con SIM_{t+k} . Para obtener μ_2 , Σ_{12} y Σ_{22} se requiere conocer SIM_{n+1} , SIM_{n+2} , ..., SIM_{150} , valores que se suponen desconocidos ya que el pronóstico se desea hacer sobre este periodo, entonces a partir del modelo 4.1.1 se realizan 100 simulaciones para este periodo y se obtiene el valor medio, el cual se emplea como μ_2 estimado,

$$\hat{\mu}_2 = (\overline{SIM}_{n+1}, \overline{SIM}_{n+2}, \dots, \overline{SIM}_{150})'$$

Al calcular la función de autocorrelación y la varianza, puede construirse Σ_{22} , mientras que Σ_{12} se obtiene calculando la función de autocorrelación de la serie

$$SIM_{12} = (SIM_{100}, SIM_{101}, \dots, SIM_n, \overline{SIM}_{n+1}, \overline{SIM}_{n+2}, \dots, \overline{SIM}_{150})'$$

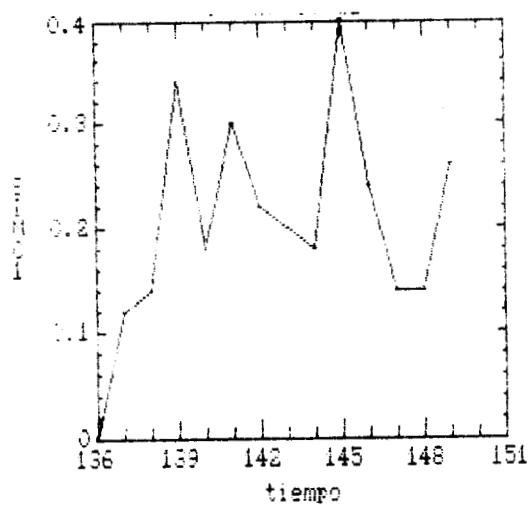
y multiplicando la matriz de autocorrelaciones por la desviación estándar de SIM_1 y la desviación estándar de $\hat{\mu}_2$.

Una vez obtenida la sucesión $SIM_{n+1}^{(N)}$, $SIM_{n+2}^{(N)}$, ..., $SIM_{150}^{(N)}$, se obtiene el SPF utilizando el programa PROBABILIDAD que se desarrolló con este fin y se expone en el apéndice D (se eligió en este caso $N = 100$). Las figuras 4.1.7 a), b) y c) muestran los SPF de un PCC con $\tau = 1$, para $n = 136, 138, 140$ respectivamente. Asimismo, las figuras 4.1.8 a), b) y c) muestran los SPF de un PCA con $\tau = 1$, para $n = 136, 138$ y 140 respectivamente, para $\tau = 2$ se obtuvieron los SPF que se muestran en las figuras 4.1.9 y 4.1.10

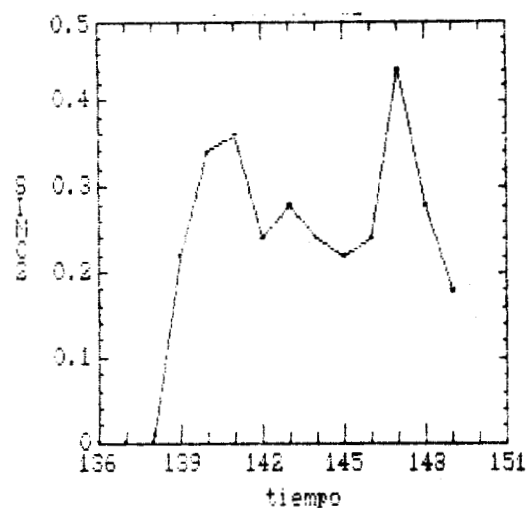
FIGURA 4.1.7

SPF de un PCC, con $\tau = 1$, para $\{SIM_t\}$,

a)
 $n = 136$



b)
 $n = 138$



c)
 $n = 140$

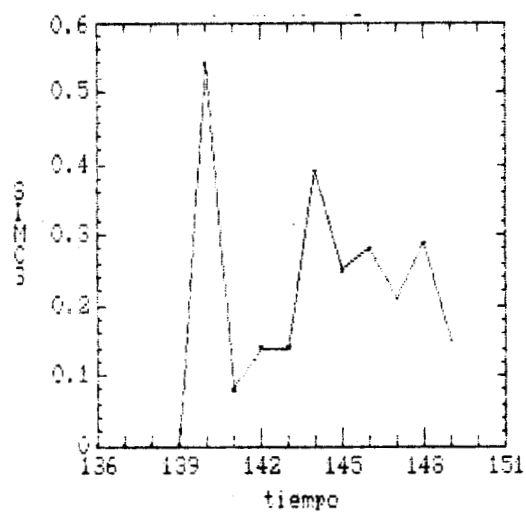
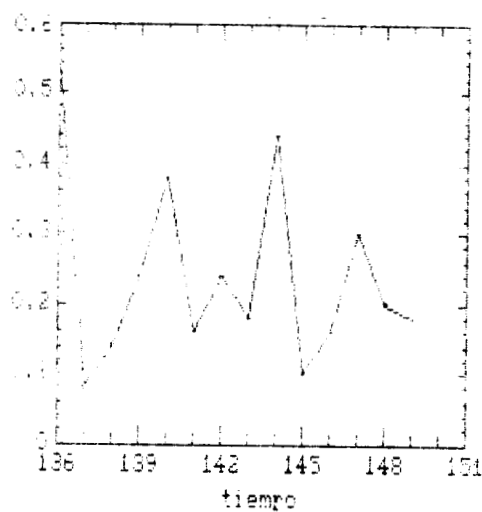


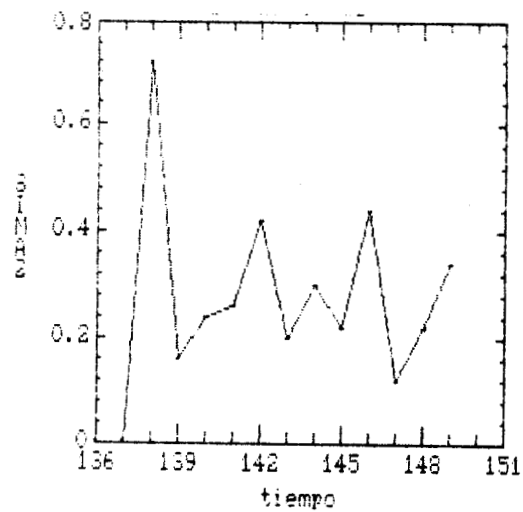
FIGURA 4.1.8

SPF de un PCA con $\tau = 1$, para (SIM_t)

a)
 $n = 136$



b)
 $n = 138$



c)
 $n = 140$

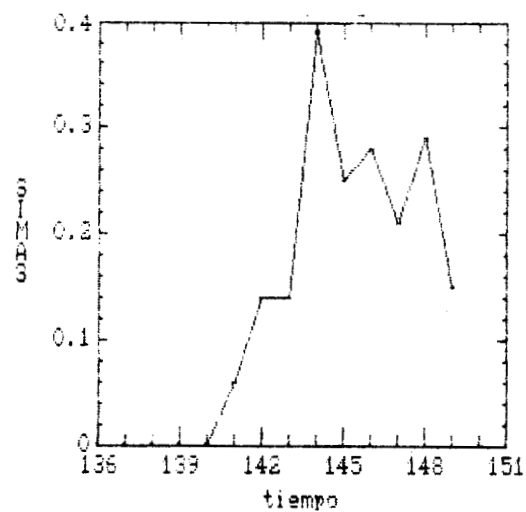
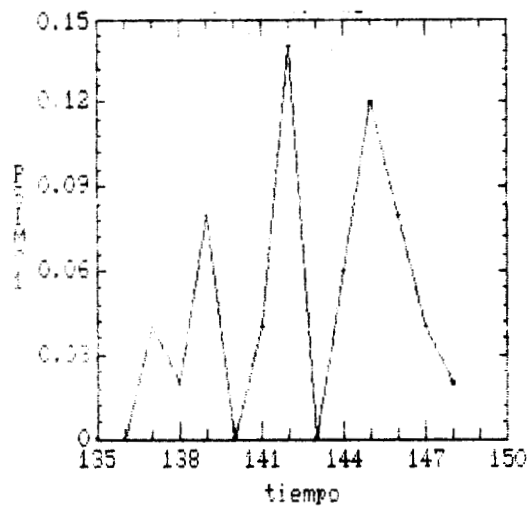


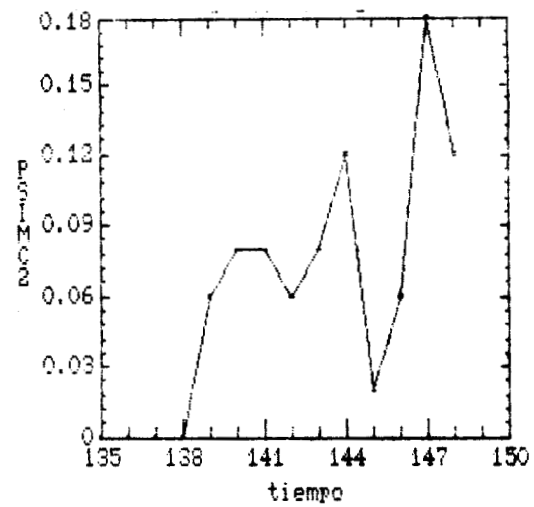
FIGURA 4.1.9

SPF para un PCC con $\tau = 2$ para (SIM_t)

a)
 $n = 136$



b)
 $n = 138$



c)
 $n = 140$

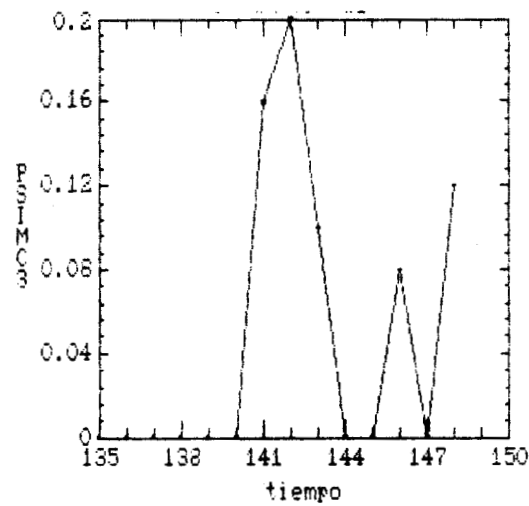
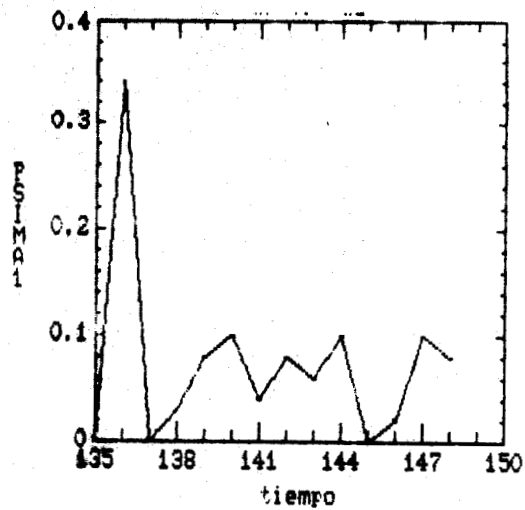


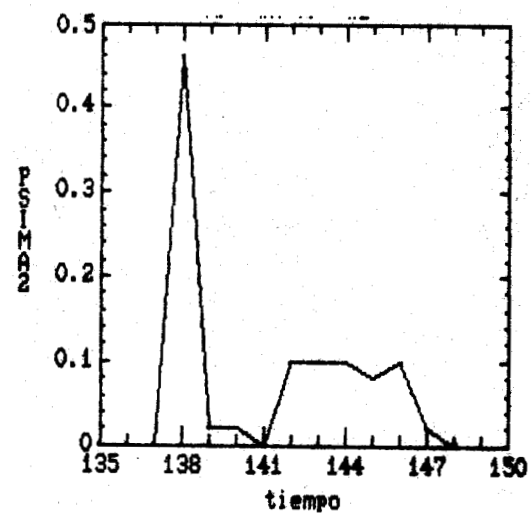
FIGURA 4.1.10

SPF para un PCA con $\tau = 2$ para (SIN_t)

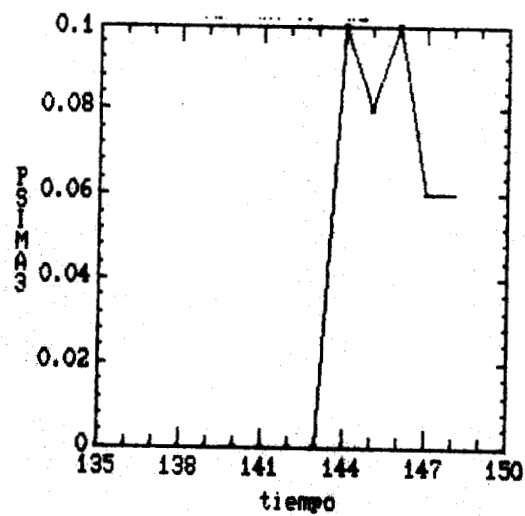
a)
 $n = 136$



b)
 $n = 138$



c)
 $n = 140$



Como puede observarse, cuando se emplea la definición 2.2.1 con $\tau = 1$, el área bajo la curva de los SPF es mayor que uno, esto se debe a que, el hecho de que ocurra un cambio en determinado momento no excluye la posibilidad de que ocurra un cambio en otro periodo más adelante, dentro del intervalo de tiempo que se está considerando. Al usar $\tau = 2$ ocurre lo mismo, sólo que en el presente ejemplo, las probabilidades de PCC y PCA que se obtuvieron para este caso fueron bajas; esta dificultad no puede superarse ya que bajo esta definición es imposible determinar cuándo un cambio es más significativo que otro. Por este motivo, no tiene sentido probar la buena calibración del SPF como lo propone Kling, ya que un SPF bien calibrado presupone que los eventos involucrados son mutuamente excluyentes.

Para $\tau = 1$ y $n = 136$ los SPF se obtienen para el intervalo de tiempo $[136, 149]$, ya que para determinar si un PCC o un PCA ocurre en el momento $t = 136$ se requiere conocer la observación 137, la cual se supone desconocida, por este motivo la ocurrencia de un PCC o un PCA en el momento 136 resulta incierta y puede calcularse su probabilidad. Por otra parte, para determinar la ocurrencia de un PCC o un PCA al momento $t = 150$, se requiere conocer la observación en $t = 151$, la cual no entra en el periodo en estudio, es por esto que no se determina la probabilidad de ocurrencia de un PCC o un PCA al tiempo 150. Análogamente, para $n = 138$ y 140 el SPF se determina dentro de los intervalos de tiempo $[138, 149]$ y $[140, 149]$ respectivamente.

Nótese en la figura 4.1.7 que al incluir más información, el SPF sufre cambios; tomando como punto de partida $n = 136$ un PCC se espera con probabilidad 0.34 al momento 139, con probabilidad 0.30 al momento 141 y con probabilidad 0.40 al momento 145, incluyendo en la información las observaciones 137 y 138, la probabilidad de que un PCC ocurra al momento 139 disminuye a 0.22, mientras que para $t = 141$ ésta se incrementa a 0.36 y para $t = 145$ sufre un decremento, tomando el valor 0.22. Al parecer los puntos más probables sufrieron

104441

un desfaseamiento de dos unidades de tiempo a la derecha, ya que en este caso las probabilidades más altas se observan en los tiempos 141, 143 y 147 con valores 0.36, 0.28 y 0.44 respectivamente, tomando como punto de partida $t = 140$ este comportamiento no se repite, ya que las probabilidades más altas están dadas en los tiempos 140, 144 y 148 con valores de 0.54, 0.39 y 0.3 respectivamente, sin embargo puede observarse un comportamiento muy interesante al tiempo 140, en el cual se introduce más información y la probabilidad de un PCC se incrementa. Esto es, al tomar como punto de partida $t = 136$, la probabilidad de que un PCC ocurra al tiempo 140 es de 0.18 con relativamente alta probabilidad de que ocurra antes; se introducen ahora dos observaciones más, la 137 y la 138, y puede observarse que realmente no ocurrió un cambio en el momento 137, el cual se esperaba con probabilidad de 0.12 y la probabilidad de que un PCC ocurra al tiempo 138 se anula por completo ya que $SIM_{137} = 1679.09$ es mayor que $SIM_{138} = 1678.62$, para $t = 139$ la probabilidad disminuye, pero la probabilidad de que un PCC ocurra al tiempo 140 aumenta a 0.34, introduciendo las observaciones 139 y 140 se tiene que realmente no ocurrió un PCC en $t = 139$ pero la probabilidad de que ocurra al tiempo 140 toma un valor de 0.54, se esperaría entonces que un PCC ocurriera en $t = 140$, no obstante el cambio ocurrió realmente en $t = 142$ donde era poco probable.

Obsérvense ahora los SPF para PCA, tomando como punto de partida $t = 136$ se espera que un PCA ocurra en $t = 136$ con probabilidad 0.54 y en $t = 144$ con probabilidad 0.44, sin embargo en $t = 136$ y 137 realmente no ocurre un PCA y dados los valores de estas observaciones la probabilidad de que un PCA ocurra en $t = 138$ se incrementa tomando el valor 0.66; es importante observar que en este punto realmente ocurre un PCA. Tomando ahora como punto de partida $t = 140$ se espera, con probabilidad de 0.4, que un PCA ocurra al tiempo 144, no obstante el cambio ocurre realmente al tiempo 143. El hecho de que los puntos de cambio no siempre ocurran en el momento más esperado no debe ser un factor determinante para

descartar el método, ya que éste calcula las probabilidades de ocurrencia y de ninguna forma puede pensarse como algo determinista. Cabe subrayar que una observación que se incluya puede hacer cambiar por completo el SPF.

Se desea estudiar también la probabilidad de ocurrencia de cambios de nivel en la serie $\{SIMN_t\}$, para esto y siguiendo el mismo procedimiento que se empleó para generar la sucesión $SIM_{n+1}^{(N)}$, $SIM_{n+2}^{(N)}$, ..., $SIM_{150}^{(N)}$, se genera la sucesión $SIMN_{n+1}^{(N)}$, $SIMN_{n+2}^{(N)}$, ..., $SIMN_{149}^{(N)}$, pero considerando la serie

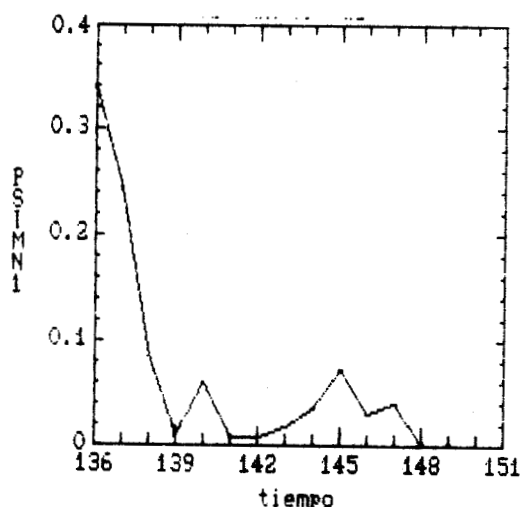
$$SIMN_1 = (SIMN_{99}, SIMN_{100}, \dots, SIMN_n)$$

en lugar de la serie SIM_1 , y $n = 135, 137$ y 139 . Se aplica ahora el método propuesto en la subsección 3.1.2, para lo cual se utiliza el programa que se muestra en el apéndice E, con $N = 1000$. Las figuras 4.1.12 a), b) y c) muestran los SPF para CN tomando $n = 135, 137$, y 139 respectivamente.

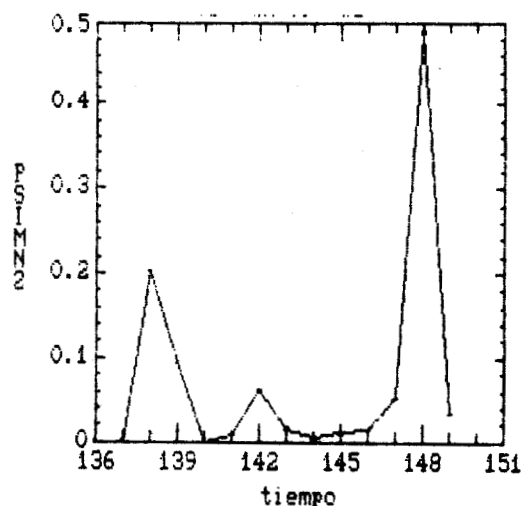
FIGURA 4.1.12

SPF para un CN en la serie $\{SIMN_t\}$

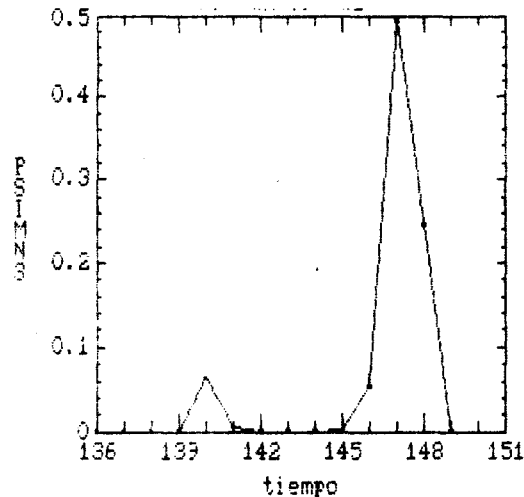
a)
 $n = 135$



b)
 $n = 137$



c)
 $n = 139$



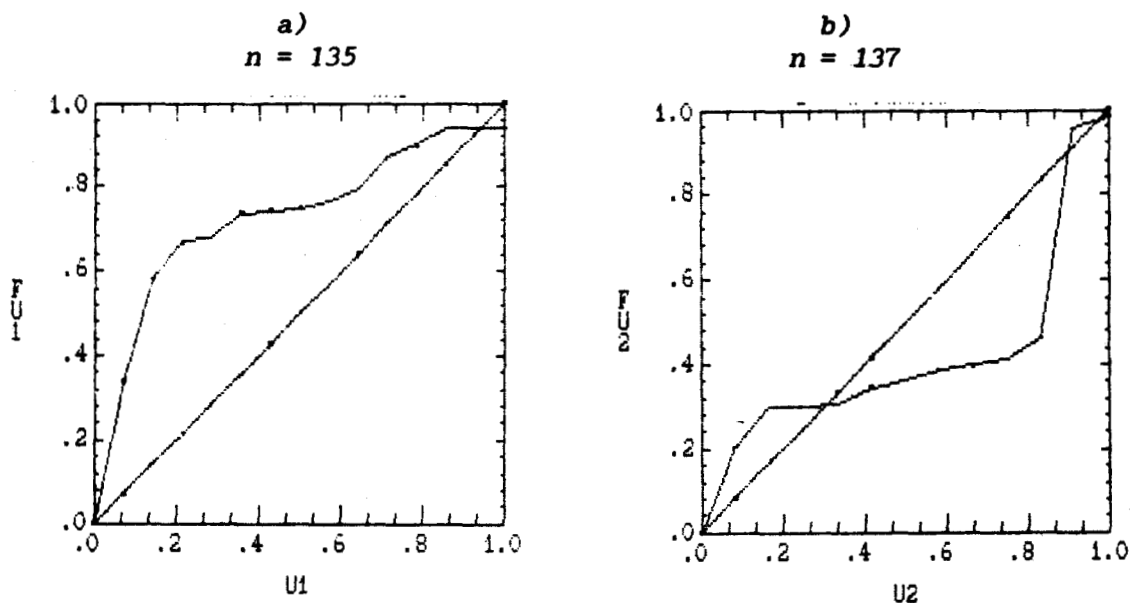
Como puede observarse en la figura 4.1.12, al introducir mayor información el SPF generado mediante este método también sufre cambios. Al tomar como punto de partida $t = 135$, se obtiene que la probabilidad de que un CN ocurra al tiempo $t = 136$ es de 0.34 siendo muy cercana a cero para todos los puntos restantes, ahora bien, dado que realmente no se observó CN en $t = 136$ y 137 , la probabilidad de que ocurra al tiempo $t = 138$ se ve incrementada de 0.08 a 0.2, al igual que para $t = 139$ que pasa de 0 a 0.09, pero la probabilidad de que un CN ocurra al tiempo $t = 148$ también se incrementa de 0 a 0.48. Dado que un CN tampoco ocurre en $t = 138$ y 139 , la probabilidad de que ocurra en $t = 140$ pasa de 0 a 0.06, sin embargo el punto más alto del SPF se traslada una unidad de tiempo a la izquierda, esto es, la probabilidad de que un CN ocurra en $t = 147$ pasa de 0.06 a 0.5 y la probabilidad de que ocurra al momento 148 decae a 0.25.

En este caso sí puede mostrarse la buena calibración del SPF, ya que éste se construyó de tal forma que para cada sucesión

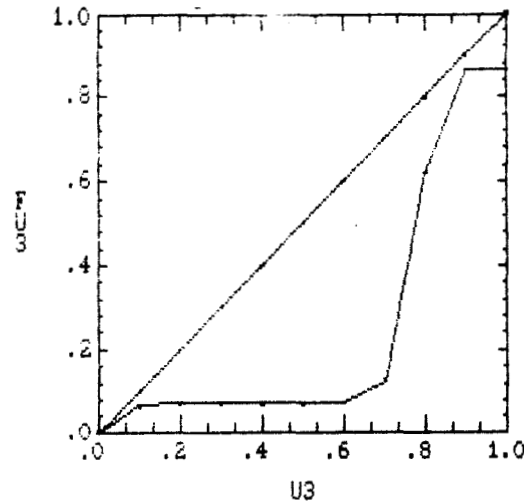
$SIMN_{n+1}^{(r)}$, $SIMN_{n+2}^{(r)}$, ..., $SIMN_{149}^{(r)}$, con $r = 1, 2, \dots, 1000$, se consideró únicamente el cambio más significativo. Para probar la buena calibración, se calculan las estadísticas de prueba SAP, SP, y χ^2 definidas en la subsección 3.1.1, para los SPF obtenidos tomando como puntos de partida $n = 135, 137$ y 139 , los cuales se denotarán como SPF1, SPF2 y SPF3 respectivamente. Las figuras 4.1.13 a), b), y c) muestran las funciones de calibración de los SPF1, SPF2 y SPF3. Para el SPF1 se tiene $SAP = 0.23$ y $SP = 0.22$, lo cual indica que se encuentra sobreconfiado, por otra parte $\chi^2 = 67.77$, lo cual conduce a concluir que el SPF1 no está bien calibrado, con un nivel de significación de cero. Para SP2 se obtiene que $SAP = 0.142$, $SP = 0.08$ y $\chi^2 = 11.56$, la hipótesis de que el SPF se encuentra bien calibrado se rechazaría con un nivel de significación de 0.25, por lo cual puede concluirse que SP2 sí está bien calibrado. Por último, para SPF3, $SAP = 0.26$ y $SP = 0.26$ lo que indica que el SPF3 se encuentra subconfiado; por su parte $\chi^2 = 19.1$, conduciendo al rechazo de la hipótesis de buena calibración con un nivel de significación de 0.1.

FIGURA 4.1.13

Gráfica de la función de calibración del SPF para CN



c)
n = 139



Es muy importante resaltar que si se espera un CN en determinado momento, un PCC o PCA se espera con probabilidad mayor o igual en el mismo instante. Para ilustrar esto, considérense las figuras 4.1.7 a), 4.1.8 a) y 4.1.9 a) y obsérvense los momentos más esperados para que ocurran CN; en $t = 136$ se espera un CN con probabilidad 0.34 y un PCA se espera en ese momento con probabilidad 0.54, en $t = 140$ la probabilidad de un CN es de 0.06, siendo la probabilidad de un PCA de 0.38, por último, la probabilidad de un CN en $t = 145$ es de 0.08, con probabilidad de un PCC para ese momento de 0.4. Adviértase que

$$P(\text{CN al tiempo } t) \leq P(\text{PCC o PCA al tiempo } t)$$

$$= P(\text{PCC al tiempo } t) + P(\text{PCA al tiempo } t).$$

Sin embargo este hecho puede tener excepciones cuando en la serie no diferenciada se espere un cambio de pendiente, sin cambiar el signo de la misma; como se inició en la sección 2.2, cualquier cambio de pendiente en una serie, se transforma bajo la diferenciación en un cambio de nivel.

Se estudiará ahora la probabilidad de cambios de nivel

empleando el método propuesto por Leiva, el cual fue presentado en la sección 3.2. Para esto se establece primero una distribución para los tiempos de cambio W ; se supondrá una distribución de Poisson con parámetro $\lambda = 17.875$, determinado por la media de los tiempos de espera obtenidos anteriormente, esto es

$$\lambda = \frac{6 + 19 + 42 + 12 + 8 + 27 + 17 + 12}{8} = 17.875$$

A partir del teorema 3.2.1 se determina la probabilidad de que un cierto número de cambios ocurran hasta determinado momento. El último cambio se observó al tiempo $t = 131$, por lo cual se calcula la probabilidad de que ningún cambio ocurra entre este momento y los tiempos $k = 136, 137, \dots, 149$. Es decir, se calculan las probabilidades de que 131 y k sean 0-próximos. También se calculan las probabilidades de que exactamente un cambio ocurra, o lo que es lo mismo que 131 y k sean 1-próximos, estas probabilidades se calculan empleando el programa que se expone en el apéndice F. La tabla 4.1.3 muestra las probabilidades de que ningún cambio ocurra y las probabilidades de que exactamente un cambio ocurra hasta el momento k .

T A B L A 4 . 1 . 3
Probabilidad de que 131 y k sean v -próximos, $v = 0, 1$

	k						
	136	137	138	139	140	141	142
0-prox	0.9996	0.9988	0.9968	0.9924	0.9835	0.9677	0.9419
1-prox	0.0003	0.0011	0.0031	0.0076	0.0159	0.0320	0.0580

	k						
k	143	144	145	146	147	148	149
0-prox	0.9036	0.8509	0.7836	0.7034	0.6138	0.5196	0.4260
1-prox	0.0960	0.1491	0.2200	0.3010	0.3910	0.4820	0.5700

Se espera entonces, con probabilidad 0.57, que al menos un CN ocurra entre el momento 136 y 149, aunque la probabilidad de que ningun cambio de nivel ocurra es bastante alta, lo cual indica que la probabilidad de que dos CN ocurran es muy pequeña, esto es

$$\begin{aligned}
 P(131 \text{ y } 149 \text{ sean 2-próximos}) &< 1 - [P(131 \text{ y } 149 \text{ sean 0-próximos}) \\
 &+ P(131 \text{ y } 149 \text{ sean 1-próximos})] \\
 &= 1 - [0.426 + 0.57] = 0.004.
 \end{aligned}$$

Cabe señalar que en realidad únicamente un CN ocurrió en este periodo.

4.2. Simulación Bivariada.

Esta sección tiene como objetivo comparar en forma práctica los SPF que se obtienen de una serie, al considerar por un lado, que es generada por un proceso bivariado y por otro, si se considera como una serie univariada. Con este fin se hizo la simulación de una serie bivariada a partir del modelo siguiente:

$$\begin{aligned}
 x_{1,t} &= 0.8 x_{1,t-1} + a_{1,t} \\
 x_{2,t} &= 0.4 x_{1,t-1} + 0.6 x_{2,t-1} + a_{2,t}
 \end{aligned}
 \tag{4.2.1}$$

donde $a_t = \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix}$ es un proceso gaussiano de ruido blanco con vector

de medias $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ y matriz de varianzas y covarianzas $\sigma = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$

Este model fue elegido por su simplicidad y porque a la vez satisface el que la variable $x_{2,t}$ dependa no solo de su pasado, sino del comportamiento de $x_{1,t}$, mientras que $x_{1,t}$ únicamente es función de su propia historia. La realización de estas series se muestra en las figuras 4.2.1 y 4.2.2.

FIGURA 4.2.1

Gráfica de la serie $x_{1,t}$ simulada a partir del modelo (4.2.1)

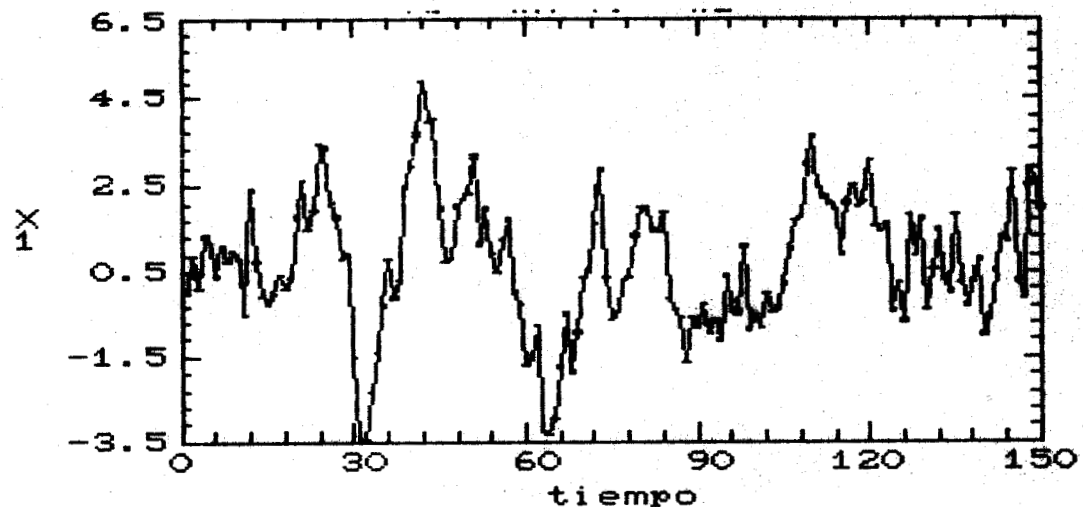
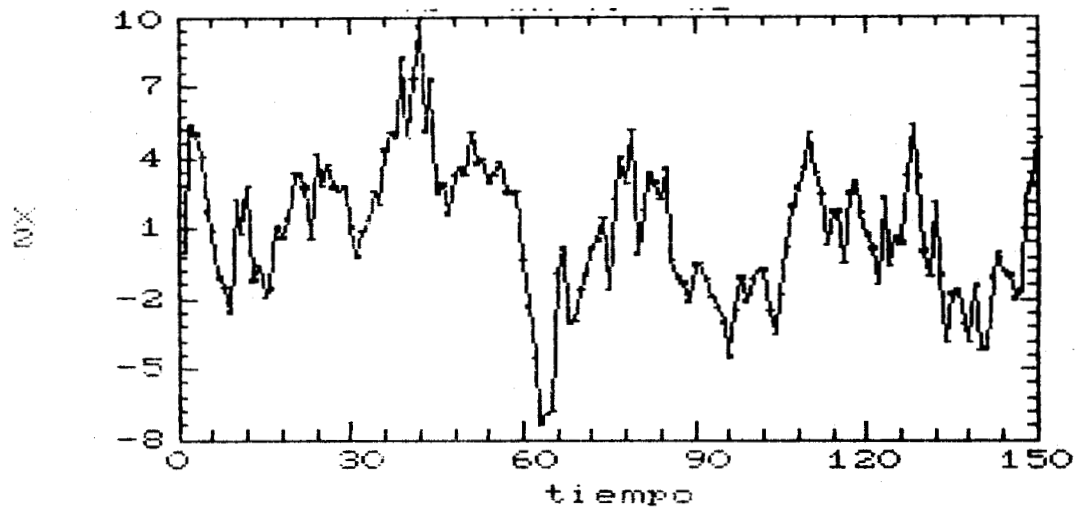


FIGURA 4.2.2

Gráfica de la serie x_2 simulada a partir del modelo (4.2.1).



De la primera integración de la serie $x_{2,t}$ se obtiene la serie $SIM2_t$ la cual se muestra en la figura 2.2.3, donde puede observarse que presenta PCC y PCA, la tabla 4.2.1 muestra los momentos en que éstos ocurren según la definición 2.2.1 con $\tau = 1$ y $\tau = 2$.

FIGURA 4.2.3

Gráfica de la serie $SIM2_t$

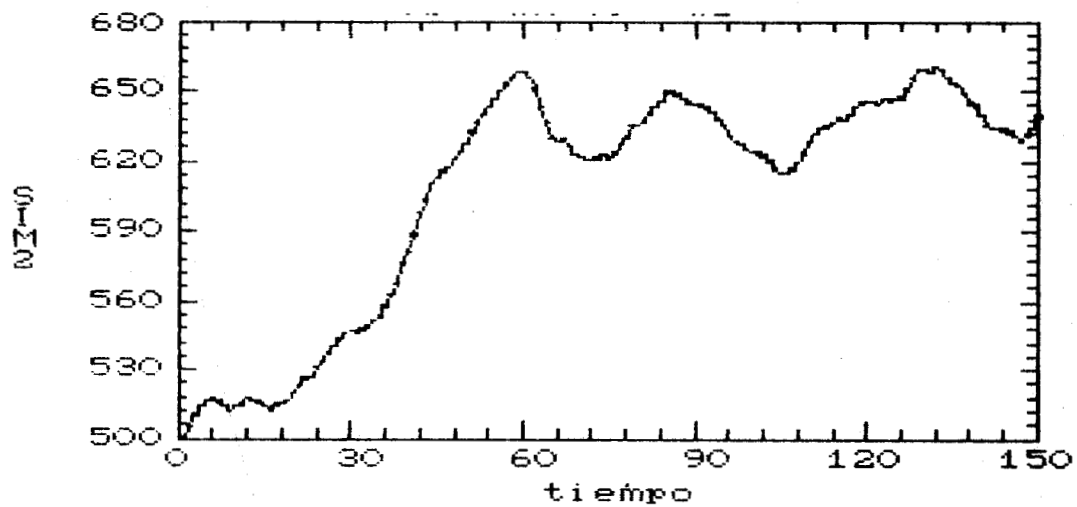


TABLA 4.2.1

*Tiempos de ocurrencia para PCC y PCA en la serie $SIM2_t$
para $\tau = 1$ y $\tau = 2$*

		TIEMPO DE OCURRENCIA												
$\tau = 1$	PCC	6	12	59	67	74	79	85	115	121	123	130	132	143
	PCA	9	15	66	71	75	80	105	122	124	131	142	147	
$\tau = 2$	PCC	6	12	59	85									
	PCA	9	15	71	105	147								

Con el fin de ilustrar el método de pronóstico estudiado en la sección 3.1.1 para series multivariadas, se identifica y estima un vector autorregresivo empleando el método de mínimos cuadrados ordinarios y el estadístico de Hannan y Quin (ver Guerrero (1987)). El modelo estimado resulta ser

$$\begin{aligned} x_{1,t} &= 0.74 x_{1,t-1} + a_{1,t} \\ x_{2,t} &= 0.15 x_{1,t-1} + 0.74 x_{2,t-1} + a_{2,t} \end{aligned} \quad (4.2.2)$$

con matriz de varianzas y covarianzas estimada

$$\hat{\sigma} = \begin{bmatrix} 1.04 & 0.97 \\ 0.97 & 3.52 \end{bmatrix}$$

y

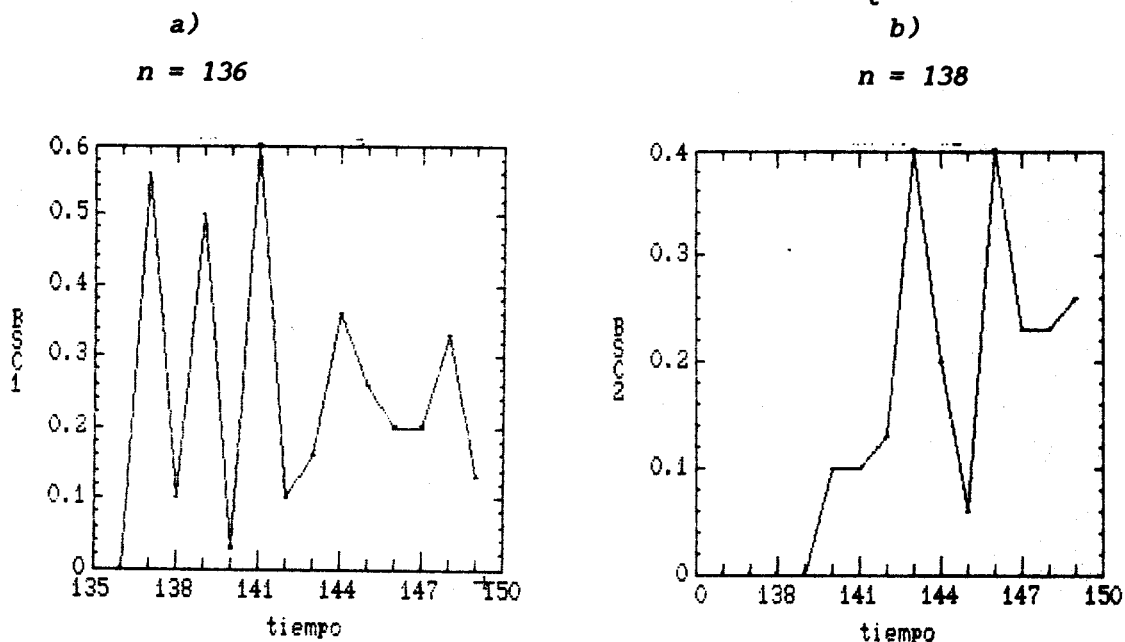
$$x_{2,t} = \nabla SIM2_t$$

Para la construcción de los SPF se emplean los últimos 50 valores de la serie $SIM2_t$, en este caso también se eligen tres puntos de partida para el pronóstico: $n = 136$, $n = 138$ y $n = 140$.

Para obtener los SPF se requiere determinar las sucesiones $SIM2_{n+1}^{(N)}$, $SIM2_{n+2}^{(N)}$, ..., $SIM2_{n+k}^{(N)}$, para lo cual es necesario estimar μ_2 en cada punto de partida. La estimación se obtiene al promediar 100 simulaciones de la serie $\{SIM2_{n+1}, SIM2_{n+2}, \dots, SIM2_{n+k}\}$ generadas a partir del modelo (4.2.2). Para la construcción de los SPF se sigue el procedimiento descrito en la sección anterior, las gráficas 4.2.4 a), b) y c) muestran los SPF para PCC con $\tau = 1$, obtenidos tomando como punto de partida $n = 136$, $n = 138$ y $n = 140$ respectivamente. Las gráficas 4.2.5 a), b) y c) muestran los SPF para PCA respectivamente para $n = 136$, $n = 138$ y $n = 140$.

FIGURA 4.2.4

SPF de un PCC, con $\tau = 1$, $m = 2$, para $\{SIM2_t\}$



c)
 $n = 140$

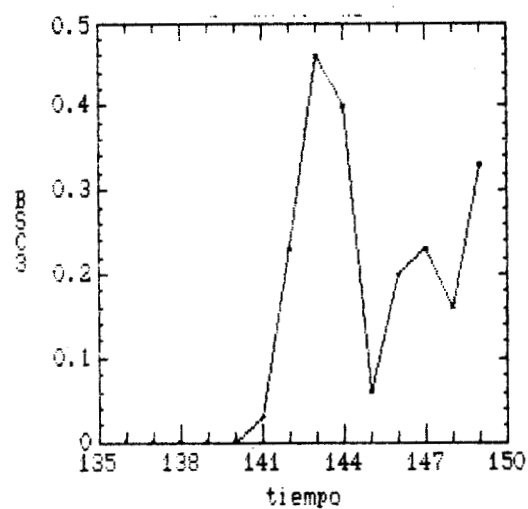


FIGURA 4.2.5

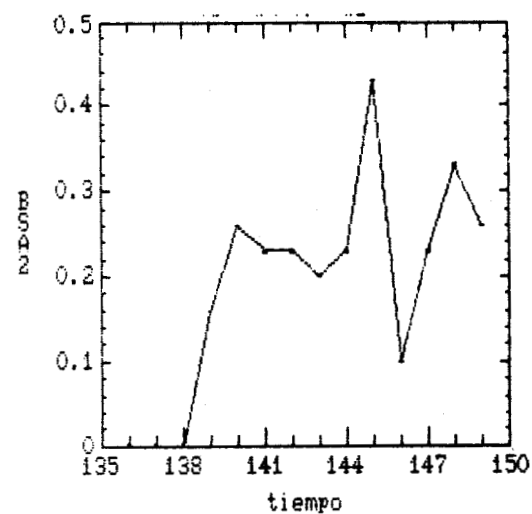
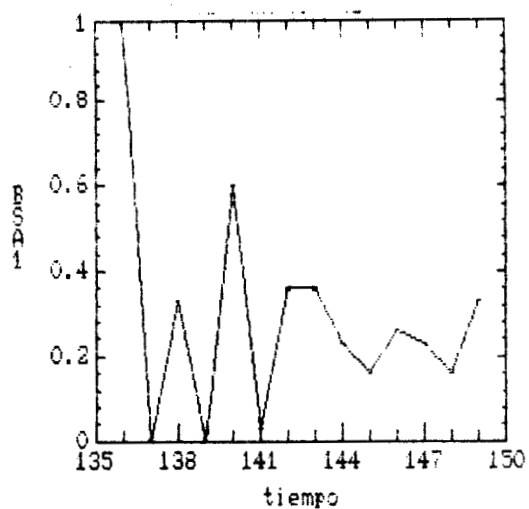
SPF de un PCA con $\tau = 1$, $m = 2$, para (SIM_2)

a)

$n = 136$

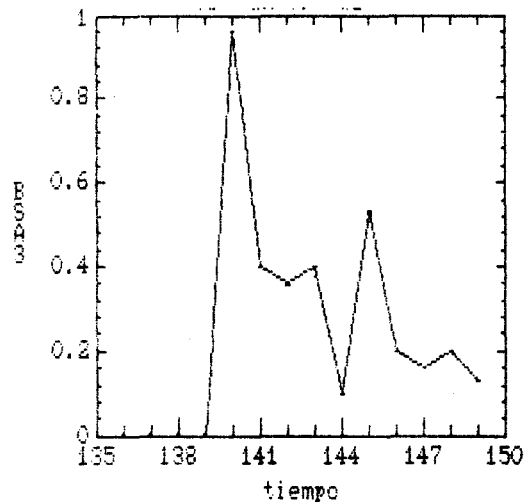
b)

$n = 138$



c)

$n = 140$



Puede observarse que la probabilidad de que un PCC ocurra en el momento 143, donde realmente ocurrió, se va incrementando al introducir más información, esto es, tomando como punto de partida $n = 136$, la probabilidad de que un PCC ocurra en $t = 143$ es de 0.16, la cual se incrementa a 0.4 al introducir las observaciones 137 y 138, llegando a 0.46 cuando se dispone de las observaciones 139 y 140.

Los puntos abismo ocurrieron en los momentos 142 y 147, las probabilidades de un PCA para el momento 142 fueron variando tomando los valores 0.36, 0.23 y 0.36 respectivamente para los puntos de partida 136, 138 y 140. Por su parte, para el momento 147 las probabilidades de un PCA toman los valores 0.23, 0.23 y 0.16, observándose así una disminución al tomar como punto de partida $n = 140$.

Por otra parte se considera la serie $\{SIM_2\}$ en forma univariada. Con el fin de determinar los SPF para los mismos puntos

de partida, se identifica y estima un modelo ARMA(p,q), el cual resulta ser

$$(1 + 0.2B)V^2SIM2_t = a_t$$

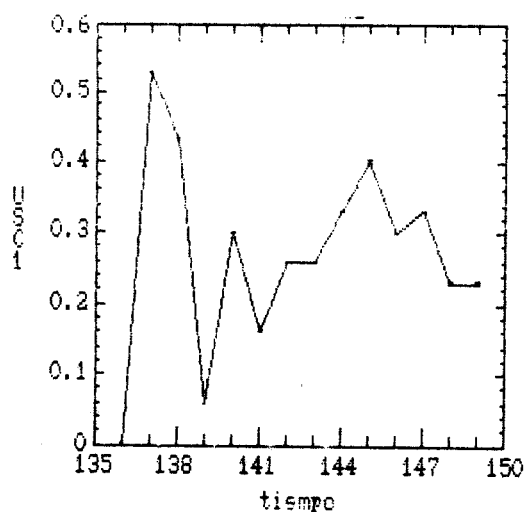
Los SPF se construyen en la forma descrita en la sección anterior. Las graficas 4.2.6 a), b) y c) muestran los SPF con para PCC con $\tau = 1$ y $m = 1$, obtenidos para $n = 136$, $n = 138$ y $n = 140$ respectivamente. Por su parte, la figura 4.2.7, muestra los SPF para PCA con $\tau = 1$ y $m = 1$.

FIGURA 4.2.6

SPF de un PCC con $\tau = 1$ y $m = 1$, para $(SIM2_t)$

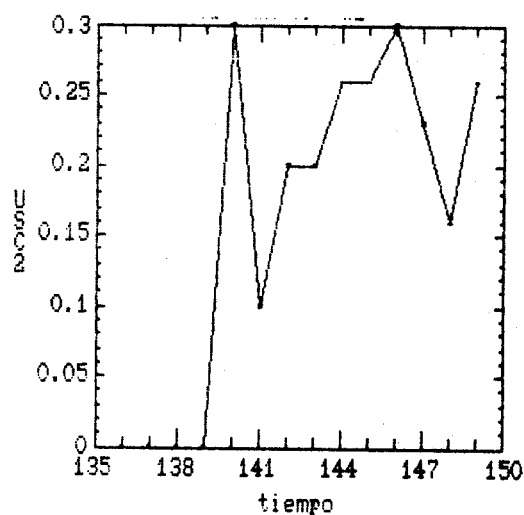
a)

$n = 136$



b)

$n = 138$



c)
 $n = 140$

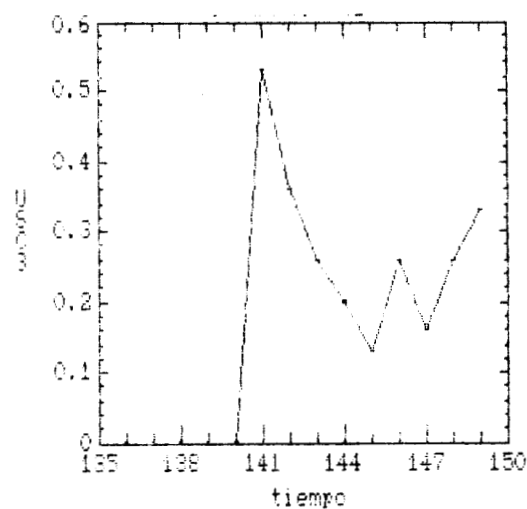
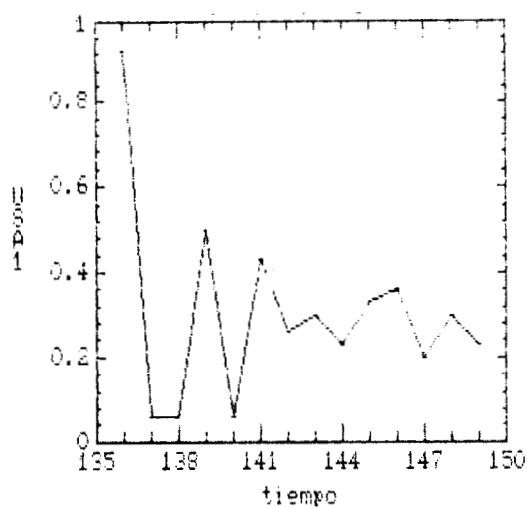


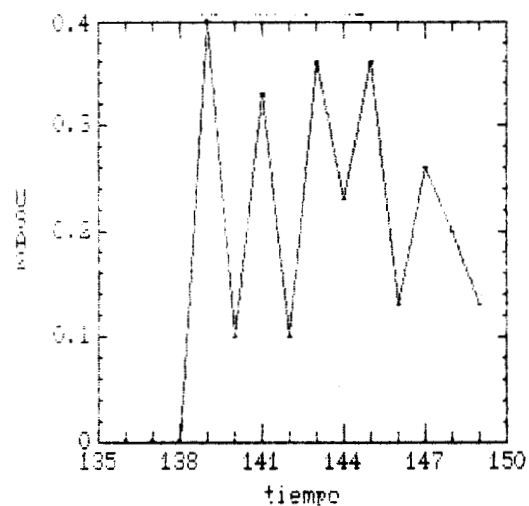
FIGURA 4.2.7

SPF de un PCA con $\tau = 1$ y $m = 1$, para (SIM_2)

a)
 $n = 136$

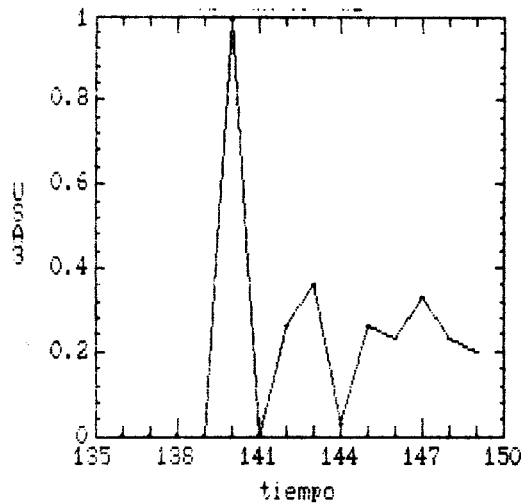


b)
 $n = 138$



c)

$n = 140$



Para este caso, las probabilidades de que un PCC ocurra al momento 143 toman los valores 0.26, 0.2 y 0.26 para los puntos de partida $n = 136$, 138 y 140 respectivamente, siendo éstas menores a las obtenidas en el caso bivariado. Las probabilidades de PCA al momento 142 toman los valores 0.26, 0.1 y 0.26, siendo éstas también menores a las del caso bivariado. Para el momento 147 se observa que las probabilidades de PCA toman los valores 0.36, 0.26 y 0.33, resultando ser en este caso mayores a las alcanzadas en el caso bivariado.

Obsérvese también que la probabilidad de un PCA al tiempo 136 es muy cercana a uno, tanto en el caso bivariado como en el univariado; lo mismo ocurre para el momento 140, siendo éste el punto de partida. Cuando el punto de partida es $n = 138$ la probabilidad de que un PCC se produzca en este mismo instante es de cero para ambos casos. Resulta entonces que los SPF presentan cierta similitud para el caso bivariado y univariado, sin embargo, por los resultados antes analizados, puede pensarse que el pronóstico bivariado es más preciso que el univariado.

CAPITULO 5. CONCLUSIONES Y COMENTARIOS.

El problema de pronóstico de puntos de cambio tiene como primer punto que afrontar, el proporcionar una definición adecuada al problema en estudio; la definición empleada por Wecker para PCC y PCA tiene la desventaja de que no se puede establecer cuándo un PCC o un PCA es significativo, así bajo la definición 2.2.1 con $\tau = 1$, los cambios que se presentan en la serie dada en la figura 5.1, son igualmente importantes, esto no sucede cuando se estudia en su lugar la serie diferenciada, ya que como puede observarse en la figura 5.2, la cual proporciona la primer diferencia de la serie dada en la figura 5.1, bajo la diferenciación dichos puntos de cambio se transforman en PA, para los cuales sí existen criterios para determinar su significación y magnitud. Si los PCC y PCA de una serie satisfacen la definición 2.2.1 con $\tau > 1$, éstos se transforman bajo la diferenciación en CN. Adviertase que los puntos de cambio observados en la figura 5.1, también pueden denominarse PS. Cabe recordar que cambios en la pendiente de una serie también pueden observarse como cambios de nivel en la serie diferenciada. Cabe observar también que la definición 2.2.1 tiene sentido cuando se trabaja con series que presentan tendencia polinomial, ya que bajo esta definición, una serie estacionaria presenta un sinnúmero de PCC y PCA, esto puede notarse en la figura 2.2.1. Recuérdese que una serie que a corto plazo presenta tendencia polinomial, a largo plazo puede resultar estacionaria.

Con respecto al pronóstico de puntos de cambio, el método propuesto por Wecker expuesto en la sección 3.1, presenta dificultades debido a las desventajas derivadas de no poder determinar si un cambio es realmente significativo o es producto de una perturbación meramente aleatoria. Esto causa problemas en la construcción del SPF, ya que para cada realización $x_{n+1}^{(j)}$, $x_{n+2}^{(j)}$, ..., $x_{n+k}^{(j)}$, con $j = 1, 2, \dots, N$ pueden encontrarse varios PCC o PCA, sin tener un criterio para descartar los que sean de poca importancia. Se tiene entonces que la probabilidad de que un PCC (o un PCA) ocurra al tiempo $n + 1$, no excluye la posibilidad de que un

PCC (o un PCA) ocurra al tiempo $n + j$, con $j > 1$ e $i, j = 1, 2, \dots, k$; en consecuencia se obtiene un SPF en el que no tiene sentido hablar de buena calibración.

FIGURA 5.1

Gráfica de una serie que presenta PCC y PCA

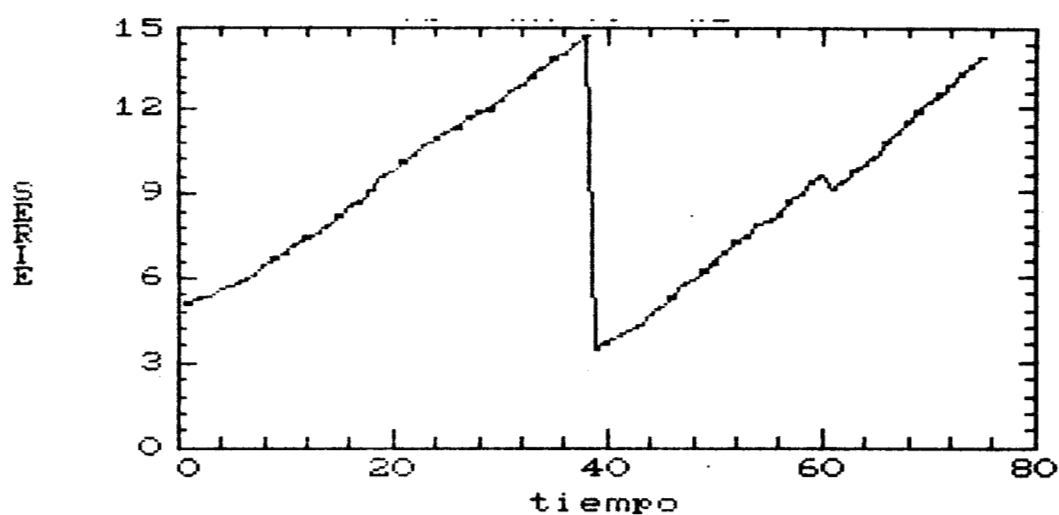
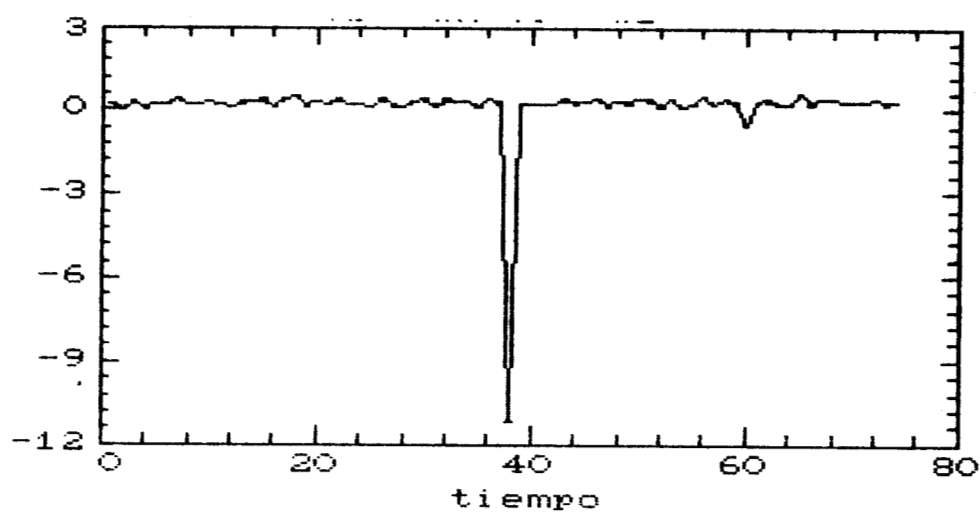


FIGURA 5.2

Gráfica de la primera diferencia de la serie dada en la figura 5.1



En la subsección 3.1.1 se propone un método alternativo, que consiste en pronosticar CN, PAI, PAA o CT de una manera excluyente, a partir de este método se deriva un SPF en el cual puede probarse la buena calibración, ya que para cada sucesión $x_{n+1}^{(j)}$, $x_{n+2}^{(j)}$, ..., $x_{n+k}^{(j)}$, se puede determinar en que momento ocurre un cambio realmente significativo y elegir únicamente aquél que presente mayor significación estadística.

No obstante, en ambos métodos el SPF es muy susceptible a cambios derivados del conocimiento de nuevas observaciones, esto es, el SPF obtenido a partir de las observaciones $x_1, x_2, \dots, x_n$ puede ser muy distinto del SPF obtenido a partir de las observaciones $x_1, x_2, \dots, x_{n+1}$. Esto se debe en parte a que si bajo el SPF obtenido a partir de las primeras n observaciones, la probabilidad de que un cambio ocurra al tiempo $n+2, n+3, \dots, n+k$ es $p_2^{(1)}, p_3^{(1)}, \dots, p_k^{(1)}$, el SPF obtenido a partir de $n+1$ observaciones ya tiene conocimiento de lo que realmente ocurrió en el momento $n+1$, entonces la probabilidad de que un cambio ocurra en el momento $n+2, n+3, \dots, n+k$ está condicionada a lo ocurrido en $n+1$. Por otra parte, llevar a la práctica estos métodos resulta un tanto laborioso, ya que obtener la función de distribución predictiva (3.1.3) implica invertir la matriz de varianzas y covarianzas de la serie observada, la cual es de orden $n \times n$, luego, si n es grande (mayor que 80) es difícil encontrar paquetes de cómputo que realicen tal operación, por lo cual no es posible emplear siempre toda la información de que se dispone.

Por su parte, el método desarrollado en la sección 3.2 para pronóstico de cambios de nivel es fácil de aplicar, sin embargo requiere de un basto conocimiento de la serie ya que, si en la serie observada únicamente han ocurrido uno o dos cambios, es difícil establecer una distribución adecuada para los tiempos de cambio. Por otra parte, este método únicamente tiene en cuenta la información relacionada con los momentos en que ocurren los cambios, sin considerar las autocorrelaciones de los datos y las posibles estacionalidades de la serie, es por esto que resulta conveniente

relacionada con los momentos en que ocurren los cambios, sin considerar las autocorrelaciones de los datos y las posibles estacionalidades de la serie, es por esto que resulta conveniente analizar el comportamiento de la serie antes de proceder a aplicar dicho método.

Para el caso multivariado resulta un tanto más complicado construir el SPF, ya que se requiere hacer simulación de la parte de la serie que se quiere pronosticar para así obtener una estimación de μ_2 , dicha simulación se realiza a partir del vector autorregresivo estimado para las series involucradas, lo cual presenta mayores dificultades para un número mayor de variables. No obstante, el pronóstico obtenido teniendo en cuenta variables explicativas es más preciso que el obtenido sin tener en cuenta estas variables.

Con respecto al método Bayessiano, éste se basa también en la definición 2.2.1, por lo cual presenta la misma dificultad que el método propuesto por Wecker, de no poderse determinar cuándo un cambio es estadísticamente significativo. Por otra parte, la aplicación resulta un tanto complicada, pues requiere de programas de cómputo más complicados, con los cuales no se cuenta.

APENDICE A

DISTRIBUCION PREDICTIVA

El propósito de este apéndice es definir y explicar en forma concisa el concepto de función de densidad predictiva, empleado en la tercera sección del capítulo 3. El material aquí expuesto fue tomado de Aitchison y Dunsmore (1975), donde puede abundarse sobre el tema de Pronóstico Estadístico.

Para entender la definición de función de densidad predictiva deben conocerse otros conceptos íntimamente ligados con éste, como son la función de densidad marginal, la distribución a priori y la distribución a posteriori. Se comenzará entonces por recordar la definición de función de densidad marginal, cuya utilidad se verá claramente más adelante. En el desarrollo de este apéndice se usará el término función de densidad para denotar tanto la función de probabilidad en el caso discreto, como la función de densidad de probabilidad en el caso continuo. Además se usará el símbolo $\int$ para denotar integración o suma, según sea aplicable.

DEFINICION-A.1. Supóngase que se especifica una función de densidad $p(\theta)$ en un espacio Θ y que, para cada $\theta \in \Theta$, se especifica una función de densidad $p(x|\theta)$ en un espacio X . La función de densidad marginal $p(x)$ en X , está dada por

$$p(x) = \int_{\Theta} p(x|\theta) p(\theta) d\theta$$

Nótese que $p(\theta)$ y $p(x)$ son funciones completamente diferentes. Supóngase que se está interesado en estudiar los posibles resultados de un experimento futuro f , y que para la descripción probabilística de éste, se tiene un espacio muestral Y y una clase de posibles funciones de densidad

$$\{ p(y|\theta): \theta \in \Theta \}$$

en Y .

El espacio Θ se supone conocido, pero el verdadero parámetro θ^* es desconocido. La incertidumbre acerca del verdadero valor parametral es lo que da origen al problema de predicción. Si se conoce θ^* , se puede conocer la función de densidad $p(y|\theta^*)$, describiendo así la aleatoriedad de experimento futuro f .

La incertidumbre acerca del verdadero parámetro, puede ser mitigada por la información que proporcionan la distribución a priori acerca de θ y la información experimental.

i) La información a priori acerca de θ .

Supóngase que se desconoce el verdadero parámetro θ^* , esta aleatoriedad se puede expresar mediante una función de densidad conocida $p(\theta)$ en Θ .

ii) Información experimental.

Supóngase que es posible obtener información mediante un experimento e , en el espacio muestral X . Este experimento proporciona información acerca del verdadero parámetro, bajo el supuesto de que el espacio parametral de la familia de funciones de probabilidad que describen a e es el mismo Θ , y que el verdadero parámetro es θ^* .

La familia de funciones de densidad para e puede denotarse entonces por

$$\{ p(x|\theta): \theta \in \Theta \}$$

en X .

La información obtenida observando x , después de realizar el experimento e , claramente influye sobre las posibles distribuciones a priori $p(\theta)$, que se han atribuido a los posibles parámetros. La actualización de las posibilidades a la luz de las observaciones x , se efectúa a través del teorema de Bayes, el cual conduce a la función a posteriori:

$$p(\theta|x) = \frac{p(\theta) p(x|\theta)}{p(x)}$$

donde

$$p(x) = \int_{\theta} p(\theta) p(x|\theta) d\theta.$$

Hasta ahora se ha usado la información muestral y la información a priori para estimar el parámetro desconocido θ^* , pero éste no es el objetivo final. El verdadero interés es describir la aleatoriedad del suceso y del experimento futuro f , y se requiere para su descripción utilizar la estimación de los posibles parámetros.

La aleatoriedad de y , dado $p(\theta)$ y x , se puede describir mediante la función de densidad predictiva $p(y|x)$, la cual se define a continuación.

DEFINICION-A.2. Para un experimento futuro f , con familia de funciones de densidad

$$\{ p(y|\theta): \theta \in \Theta \}$$

en Y , un experimento informativo e , con familia de funciones de densidad

$$\{ p(x|\theta): \theta \in \Theta \}$$

en X y una función de densidad a priori $p(\theta)$ en Θ , la función de densidad predictiva $p(y|x)$, se define como

$$p(y|x) = \int_{\Theta} p(y|\theta) p(\theta|x) d\theta$$

$$= \frac{\int_{\Theta} p(y|\theta) p(x|\theta) p(\theta) d\theta}{\int_{\Theta} p(x|\theta) p(\theta) d\theta}$$

Algunas variaciones pueden hacerse a lo aquí expuesto según el problema en estudio. En la sección 3.3, por ejemplo, se obtiene la función de distribución predictiva para el valor futuro x_{n+1} , partiendo del conocimiento a priori de la distribución de los parámetros del modelo generado por las observaciones de una serie de tiempo.

APENDICE B

GENERACION DE REALIZACIONES DE VARIABLES ALEATORIAS

Este apéndice tiene el objetivo de mostrar algunos métodos para generar variables aleatorias, el material aquí expuesto fue desarrollado en Sobol (1976) y en Kennedy y Gentle (1980).

Existen tres procedimientos básicos que se aplican para la obtención de realizaciones de variables aleatorias con una distribución dada, estos son: los generadores de números aleatorios, las tablas de números aleatorios y los métodos para generar números pseudoaleatorios.

Uno de los generadores de números aleatorios más primitivos es la ruleta. Para generar variables aleatorias discretas uniformemente distribuidas en el conjunto $\{1, 2, \dots, n\}$, se puede construir un disco giratorio, dividido en n pedazos de igual tamaño, con una flecha indicadora, como el que se ilustra en la figura A.1 para el caso $n = 10$. Entonces, para generar una cifra aleatoria, se hace girar el disco y posteriormente se le para en seco, la cifra señalada por la flecha es una cifra aleatoria con la distribución antes mencionada.

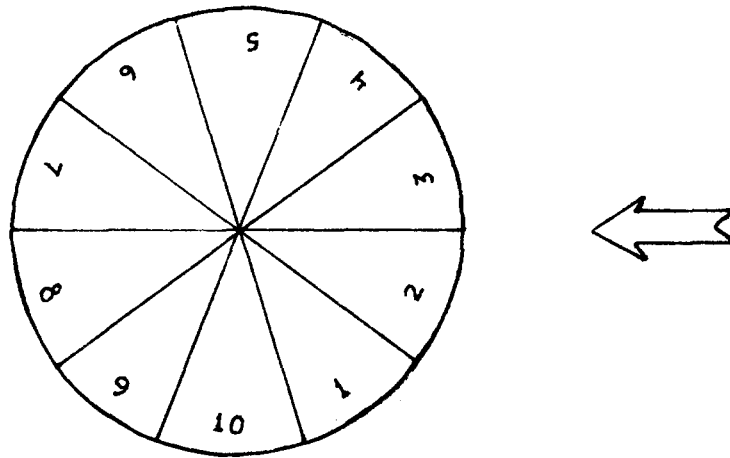
A partir de una ruleta, se pueden generar también, variables aleatorias discretas con distintas distribuciones, por ejemplo, para generar variables aleatorias con la distribución

$$\left[\begin{array}{c|cccc} \text{Valor de la realización} & x_1 & x_2 & \dots & x_n \\ \hline \text{Probabilidad} & p_1 & p_2 & \dots & p_n \end{array} \right]$$

se puede emplear un disco giratorio, dividido en partes proporcionales a p_i , $i = 1, 2, \dots, n$.

FIGURA B.1

Ruleta empleada en la obtención de números aleatorios



En la actualidad se emplean generadores de números aleatorios más sofisticados, como son los generadores electrónicos.

Una tabla de números aleatorios es una lista de cifras, obtenidas a partir de algún generador de números aleatorios, en las cuales, para generar una muestra aleatoria de tamaño n , se elige un número como punto de partida y se toman los n siguientes números como observaciones de la muestra.

Las tablas de números aleatorios, al igual que la mayor parte de los generadores, se usan generalmente cuando los cálculos en los cuales se utiliza la muestra aleatoria, se realizan a mano. Por lo cual se enfocará la atención principalmente hacia algunos métodos para generar números pseudoaleatorios, ya que éstos son utilizados en cálculos que se realizan mediante computadoras.

Cuando se generan números aleatorios, no importa el procedimiento empleado, siempre y cuando se cumplan ciertos requerimientos establecidos mediante pruebas estadísticas, como son las de bondad de ajuste. Por esta razón, estos números pueden ser

calculados mediante una fórmula dada. A los números que se obtienen a partir de una fórmula y que imitan el comportamiento de una variable aleatoria, se les denomina números pseudoaleatorios.

El primer algoritmo destinado a la construcción de números pseudoaleatorios fue propuesto por J. Von Neumann y se conoce como el *método de cuadrados centrales*. Para explicar este método se recurre aquí a un ejemplo.

Considérese un número de cuatro cifras decimales, sea este $\gamma_0 = 0.9876$, al elevarlo al cuadrado se obtiene $\gamma_0^2 = 0.97535376$ de ocho cifras decimales, tómense las cuatro cifras centrales y sea $\gamma_1 = .5353$, elévese este al cuadrado ($\gamma_1^2 = 0.28654609$) para obtener $\gamma_2 = 0.6546$, así sucesivamente se pueden obtener $\gamma_3, \gamma_4, \dots$

Este método no es del todo satisfactorio ya que proporciona más números pequeños de lo necesario, por lo cual han sido creados otros algoritmos más eficientes. La mayor parte de los algoritmos que se emplean tienen la forma recursiva

$$\gamma_{k+1} = F(\gamma_k) \quad (B.1)$$

Dado el número inicial γ_0 , todos los números siguientes se obtienen mediante la fórmula (B.1). El empleo de la función $y = F(x)$ resulta exitoso únicamente si la gráfica de ésta recubre con suficiente densidad el cuadrado unitario.

El método más usado para la obtención de números pseudoaleatorios fué propuesto por D. Lehmer y se conoce como el *método de congruencias*. Con este método se determina una sucesión de números enteros $\{m_k\}$ tomando $m_0 = 1$ y calculando todos los números sucesivos $m_1, m_2, \dots$ a partir de una misma fórmula

$$m_{k+1} \equiv 5^{17} m_k \pmod{2^{40}} \quad k = 1, 2, \dots \quad (B.2)$$

Posteriormente, a partir de estos números m_k , se calculan los números pseudoaleatorios γ_k , mediante la formula

$$\gamma_k = 2^{-40} m_k \quad (B.3)$$

La fórmula (B.2) significa que el número m_k es igual al residuo que se obtiene al dividir $5^{17} m_k$ entre 2^{40} . Este residuo lleva el nombre de residuo positivo módulo 2^{40} mínimo. Para obtener mayor información acerca de congruencias, se puede consultar cualquier texto de Teoría de los Números, como por ejemplo Niven y Zuckerman (1976).

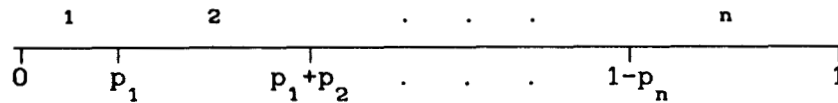
Mediante este método se obtienen realizaciones de variables aleatorias uniformemente distribuidas en el intervalo $[0,1]$. Sin embargo, en la solución de problemas reales, es necesario simular variables aleatorias con diferentes distribuciones, pero este problema no es tan serio, debido a que los valores de prácticamente cualquier variable, aleatoria se pueden obtener efectuando transformaciones de una variable aleatoria uniformemente distribuida en el intervalo $[0,1]$.

Supóngase que es preciso obtener valores de una variable aleatoria discreta ξ con distribución

$$\xi = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

para esto considérese el intervalo $[0,1]$ y divídase en n intervalos de longitud $p_1, p_2, \dots, p_n$. Las coordenadas de los puntos de división serán $p_1, p_1+p_2, \dots, 1$. Indíquese con $1, 2, \dots, n$ los intervalos, como se muestra en la figura A.2, entonces se toma γ uniformemente distribuida en el intervalo $[0,1]$, si este punto aparece en el i -ésimo intervalo, se acepta que $\xi = x_i$.

FIGURA B.2



Dado que γ se encuentra uniformemente distribuida en $[0,1]$, la probabilidad de que γ pertenezca a un intervalo contenido en $[0,1]$, es igual a la longitud del intervalo, por lo tanto

$$P \{0 < \gamma < p_1\} = p_1$$

$$P \{p_1 < \gamma < p_1 + p_2\} = p_2$$

.

$$P \{p_1 + p_2 + \dots + p_{n-1} < \gamma < 1\} = p_n$$

Supóngase ahora que se quieren obtener valores de una variable aleatoria continua ζ , distribuida con densidad $P(x)$ en el intervalo (a,b) . Los valores de ζ se pueden determinar a partir de la ecuación

$$\int_a^{\zeta} P(X) dx = \gamma \quad (B.3)$$

esto es, una vez que se conoce el valor γ , es preciso resolver la ecuación (B.3) para obtener el valor de ζ . Esta afirmación se deriva de las propiedades de la función de densidad y del resultado visto en la sección 3.1, donde se prueba que $U = F(x)$ es una variable aleatoria $U[0,1]$.

Este método resulta poco efectivo para el caso en que ζ es una variable aleatoria normalmente distribuida, ya que la ecuación

$$\frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\zeta} e^{-x^2/2} dx = \gamma$$

no admite solución explícita.

Para generar muestras de una $N[0,1]$ es conveniente recordar el Teorema Central del Limite, según el cual

$$\lim_{n \rightarrow \infty} P \left[\frac{\bar{y}_n - m}{s/n^{1/2}} \leq x \right] = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^x e^{-y^2/2} dy$$

donde m es el valor esperado de la variable aleatoria y , y s^2 la varianza común a las variables $y_1, y_2, \dots, y_n$ que se suponen independientes e idénticamente distribuidas. Entonces, un método conveniente para generar una muestra aleatoria de una $N[0,1]$ es el siguiente:

a) Generar un número suficientemente grande de valores de una variable aleatoria que sea fácil de obtener, con media m y varianza s^2 , sea $y_1, y_2, \dots$,

b) Calcúlese

$$\zeta = \frac{\bar{y}_n - m}{s/n^{1/2}}$$

Entonces ζ es aproximadamente, una realización de una variable $N[0,1]$. Es aconsejable tomar las muestras en a) de una distribución uniforme en $[0,1]$.

Para generar una variable aleatoria $\zeta' \sim N[a, \sigma^2]$ es suficiente con generar $\zeta \sim N[0,1]$ y transformarla mediante

$$\zeta' = a + \sigma\zeta$$

Para probar que ζ' así obtenida, tiene una distribución $N[a, \sigma^2]$, tómense dos números arbitrarios x_1, x_2 , tales que $x_1 < x_2$ y calcúlese

$$\begin{aligned} P \{x_1 < \zeta' < x_2\} &= P \{x_1 < a + \sigma\zeta < x_2\} \\ &= P \left\{ \frac{x_1 - a}{\sigma} < \zeta < \frac{x_2 - a}{\sigma} \right\} \end{aligned}$$

Por consiguiente

$$P \{x_1 < \zeta' < x_2\} = \frac{1}{(2\pi)^{1/2}} \int_{(x_1 - a)/\sigma}^{(x_2 - a)/\sigma} e^{-x^2/2} dx \quad (B.4)$$

Tómese ahora

$$x' = a + \sigma x$$

sustituyendo en (B.4) se obtiene

$$P \{x_1 < \zeta' < x_2\} = \frac{1}{(2\pi)^{1/2} \sigma} \int_{x_1}^{x_2} e^{-(x' - a)^2 / 2\sigma^2} dx'$$

por lo tanto

$$\zeta' \sim N[a, \sigma^2]$$

Supóngase ahora que se desea generar una muestra aleatoria x con distribución normal P -variada, denótese por μ al vector de medias y por V a la matriz de varianzas y covarianzas, entonces

$$f(\mathbf{x}) = 2\pi^{-p/2} |\mathbf{V}|^{-1/2} \exp \left\{ -1/2 (\mathbf{x}-\boldsymbol{\mu})' \mathbf{V}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right\}$$

un método para generar $\mathbf{x}$ consiste en generar $z_1, z_2, \dots, z_p$ variables aleatorias $N(0,1)$ y calcular

$$\mathbf{x} = \mathbf{A} \mathbf{z} + \boldsymbol{\mu}$$

donde $\mathbf{A}$ es tal que $\mathbf{A}\mathbf{A}' = \mathbf{V}$ y puede obtenerse mediante la descomposición de Cholesky, que puede ser estudiada con detalle en Kennedy y Gentle (1980).

APENDICE C

PROGRAMA PARA DETECTAR PUNTOS ABERRANTES Y CAMBIOS DE NIVEL

En éste apéndice se proporciona el algoritmo que fue empleado en la detección de cambios en el capítulo 4 y se basa en el método de Tsay(1987) expuesto en 2.3. El algoritmo fue realizado en Turbo Pascal.

```
PROGRAM DETECTA (INPUT, OUPUT);
USES
  CRT, PRINTER;
TYPE
  VECTOR = ARRAY [0..100] OF REAL;
VAR
  OPT, V, N, K : INTEGER;
  Y, PPI,      : VECTOR;
  SGMA         : REAL;

FUNCTION TRANSFORMA(Y: VECTOR): VECTOR;
VAR
  I, K : INTEGER
BEGIN
  K := 0;
  BEGIN
    K := K+1;
    FOR I := 2 TO N DO
       $Y[K] := Y[I] - \pi_1 Y[I-1] - \pi_2 Y[I-2] - \dots;$ 
    END;
  END;

FUNCTION RHO2(N, V: INTEGER; PPI : VECTOR): REAL;
```

```

VAR
  I : INTEGER;
  S : REAL;
BEGIN
  S := 0;
  FOR I := 0 TO N - V DO
    S := S + SQR(PPI[I]);
    RHO2 := 1/S;
  END;

FUNCTION W (N, V : INTEGER; PPI, Y, :VECTOR): REAL;
VAR
  I : INTEGER;
  S : REAL;
BEGIN
  S := 0;
  FOR I := 0 TO N - V DO
    BEGIN
      S := S + PPI[I] * Y[V+I];
    END;
  W := RHO2(N, V, PPI) * S;
END;

FUNCTION LAM (V : INTEGER; SGMA : REAL) : REAL;
BEGIN
  LAM := W(N, V, PPI, Y) / (SQRT (RHO2(N, V, PPI)) * SGMA;
END;
PROCEDURE DATA_Y_PPI;

BEGIN
  Y[1] :=      ; Y[2] :=      ; . . . Y[N] :=      ;
  PPI[0] :=      ; PPI[1] :=      ; . . . PPI[N-1] :=      ;
END;

FPROCEDURE DATA;

```

```

VAR
  I : INTEGER
BEGIN
  WRITE('OPCION = '); READLN 9OPT);
  IF OPT = 0
  THEN DATA_Y_PPI
  ELSE
    FOR I:= 0 TO N DO
      BEGIN
        WRITE (LST,'Y(',I,'); READ (Y[I]);
        WRITE (LST,'PPI(',I,'); READ (PPI[I]);
      END;
    END;
  END;

  BEGIN
    WRITE ( 'SGMA = '); READ (SGMA);
    WRITE ( 'N = '); READ (N);
    DATA;

    WRITELN ('');
    WRITELN (LST,'V      W      LAM');
    WRITELN ('');

    FOR V := 1 TO N DO
      IF V < 10
      THEN WRITELN(LST,V,'      'W(N,V,PPI,Y),'      ',LAM(V,SGMA))
      ELSE WRITELN(LST,V,'      'W(N,V,PPI,Y),'      ',LAM(V,SGMA));
    END.

```

Mediante este algoritmo pueden detectarse puntos aberrantes innovados, puntos aberrantes aditivos, cambios de nivel y cambios transitorios únicamente estableciendo los valores de PPI[I], que corresponden a los valores de los coeficientes Π_1 en el polinomio

$$\Pi(B) = \frac{\Phi(B)}{\Theta(B)}$$

para la detección de un punto aberrante aditivo; a los valores de los coeficientes η_i en el polinomio

$$\eta(B) = \frac{\Phi(B)}{\Theta(B)} (1-B)^{-1}$$

para la detección de un cambio de nivel; a los valores de los coeficientes β_i en el polinomio

$$\beta(B) = \Pi(B)(1-\delta B)^{-1}$$

para la detección de un cambio transitorio; y por último, $PPI[0] = 1$ y $PPI[I] = 0$ para $I = 1, 2, \dots, N - 1$, en el caso de un punto aberrante innovado.

La variable SGMA representa el valor de la desviación estándar de la serie y VAL_T el valor de $t_{\alpha/2, N-1}$ en la tabla de la distribución t con el nivel de significación deseado.

Las constantes Π_i en FUNCTION TRANSFORMA corresponden a los coeficientes del polinomio $\Pi(B)$. La constante N representa el número de valores observados de la serie en estudio.

APENDICE D

PROGRAMA PARA PRONOSTICAR PUNTOS DE CAMBIO CUMBRE Y ABISMO

El presente programa se emplea en el pronóstico de puntos de cambio cumbre y abismo a partir de la sucesión de variables $x_{n+1}^{(N)}$, $x_{n+2}^{(N)}$, ..., $x_{n+k}^{(N)}$, generadas con la distribución predictiva estimada para x_{n+1} , x_{n+2} , ..., x_{n+k} , como se indica en la sección 3.1

```
PROGRAM PRONOSTICO;
```

```
USES
```

```
  CRT, PRINTER;
```

```
TYPE
```

```
  MAT      = ARRAY[1..50,0..14] OF REAL;
```

```
  MATI     = ARRAY[1..50,0..13 OF INTEGER;
```

```
  VECTOR   = ARRAY[1..13];
```

```
VAR
```

```
  X, Y      : MAT;
```

```
  P         : VECTOR;
```

```
PROCEDURE LEE_ARCH (VAR Y : MAT);
```

```
VAR
```

```
  ARCH      : TEXT;
```

```
  I, J      : INTEGER;
```

```
  VALOR     : REAL;
```

```
  CAR       : CHAR;
```

```
BEGIN
```

```
  ASSIGN (ARCH, 'A: *.DAT');
```

```
  RESET (ARCH);
```

```
  I := 1;
```

```
  WHILE NOT EOF (ARCH) DO
```

```
  BEGIN
```

```
    FOR J := 1 TO 14 DO
```

```

    BEGIN
        READ (ARCH, VALOR);
    END;
    I := I + 1;
END;

END;

PROCEDURE INIX (VAR X : MATI);
VAR
    I, J      : INTEGER;
BEGIN
    FOR I := 1 TO 50 DO
        FOR J := 1 TO 13 DO
            X[I,J] := 0;
        END;
    END;

PROCEDURE INIP (VAR P : VECTOR);
VAR
    I : INTEGER;
BEGIN
    FOR I := 1 TO 13 DO
        P[I] := 0;
    END;

PROCEDURE COMPARA(VAR Y : MAT);
VAR
    I, J : INTEGER;
BEGIN
    Y[I,0] :=      ;
    FOR I := 1 TO 50 DO
        BEGIN
            FOR J := 1 TO 13 DO
                BEGIN
                    IF Y[I,J-1] >= Y[I,J]
                        THEN

```

```

        X[I,J] := 0
    ELSE
        IF Y[I,J] > Y[I,J+1]
            THEN
                X[I,J] := 1
            ELSE X[I,J] := 0
        END;
    END;
END;
END;

```

```

PROCEDURE SUMA (VAR X : MAT);

```

```

VAR

```

```

    I, J, S : INTEGER;

```

```

BEGIN

```

```

    FOR J := 1 TO 13 DO

```

```

        BEGIN

```

```

            S := 0;

```

```

            FOR I := 1 TO 50 DO

```

```

                BEGIN

```

```

                    S := S + X[I,J];

```

```

                    P[J] := S;

```

```

                END;

```

```

            END;

```

```

        END;

```

```

PROCEDURE IMPRIME (VAR P : VECTOR);

```

```

VAR

```

```

    K : INTEGER;

```

```

BEGIN

```

```

    FOR K := 1 TO 13 DO

```

```

        WRITE (P[K]:2 , ' ');

```

```

    WRITELN;

```

```

END;

```

```

PROCEDURE PASADA (VAR Y : MAT);

```

```

VAR
  X : MATI;
  P : VECTOR;
BEGIN
  INIX (X);
  INIP (P);
  COMPARA (Y);
  SUMA (X);
  IMPRIME (P);
END;

BEGIN
  CLRSCR;
  LEE_ARCH (Y);
  PASADA (Y);
END.

```

En PROCEDURE LEE_ARCH, *.dat representa el nombre del archivo donde se guarda la sucesión $x_{n+1}^{(N)}, x_{n+2}^{(N)}, \dots, x_{n+k}^{(N)}$, con $N = 50$ ya que la capacidad de la computadora empleada no permite el uso de archivos de datos demasiado grandes, es por esto que, para $N = 1000$ se requiere almacenar los datos en 20 archivos y repetir la aplicación del presente programa para cada uno de ellos. En PROCEDURE COMPARA, $Y[I,0]$ representa al último valor observado, es decir a x_n .

Este programa fue desarrollado para el pronóstico de puntos de cambio cumbre y abismo empleando la definición 2.2.1 con $m = 1$ y $\tau = 1$, sin embargo PROCEDURE COMPARA puede transformarse para emplear dicha definición con $\tau = k$, incrementando el número de comparaciones.

APENDICE E

ALGORITMO PARA EL PRONOSTICO DE CAMBIOS DE NIVEL Y PUNTOS ABERRANTES

En este apéndice se proporciona el algoritmo empleado en el capítulo 4 para el pronóstico de cambios de nivel. Una vez obtenidas las simulaciones de Y_{n+1} , Y_{n+2} , ..., Y_{n+k} , a partir de la función de distribución predictiva (3.1.3), mediante este algoritmo puede obtenerse la sucesión indicadora

$$\xi_{n+1}^{(N)}, \xi_{n+2}^{(N)}, \dots, \xi_{n+k}^{(N)} \quad \text{E.1}$$

vista en 3.1. El algoritmo fue desarrollado en Turbo Pascal.

```
PROGRAM PRONOSTICO
USES
  CRT, PRINTER;
TYPE
  MAT      = ARRAY [1..100,0..12] OF REAL;
  VECTOR   = ARRAY[0..12] OF REAL;
VAR
  SGMA, VAL_T : REAL;
  PPI          : VECTOR;
  Y            : MAT;

PROCEDURE LEE_ARCH (VAR Y : MAT);
VAR
  ARCH : TEXT;
  I, J : INTEGER;
  VALOR : REAL;
  CAR : CHAR;
BEGIN
  ASSIGN (ARCH, 'A: PRON.DAT');
  RESET (ARCH);
```

```

I := 1;
WHILE NOT EOF (ARCH) DO
BEGIN
  FOR J := 1 TO 12 DO
  BEGIN
    READ (ARCH, VALOR);
    Y [I,J] := VALOR;
    READ (ARCH, CAR);
  END;
  I := I + 1;
END;
END;

```

```

PROCEDURE INICIO (VAR X : MAT);
VAR
  I, J : INTEGER;
BEGIN
  FOR I := 1 TO 100 DO
    FOR J := 1 TO 12 DO
      X[I,J] := 0.0;
    END;
  END;

```

```

PROCEDURE TRANSFORMA (Y : MAT; VAR X; MAT);
VAR
  I, J : INTEGER;
BEGIN
  FOR J := 1 TO 100 DO
    BEGIN
      Y[J,0] := ;
      BEGIN
        FOR I := 1 TO 12 DO
          X[J,I] := Y[J,I] -  $\pi$ *Y[J,I-1];
        END;
      END;
    END;
  END;

```

```

PROCEDURE CALC_R2 (PPI : VECTOR; VAR R2 : VECTOR);
VAR
  I, V, N, COL : INTEGER;
  S              : REAL;
BEGIN
  COL := 0;
  N   := 12;
  FOR V := 1 TO 12 DO
    BEGIN
      S := 0;
      COL := COL + 1;
      FOR I := 0 TO N - V DO
        R2[COL] := S;
      END;
    END;
  END;

PROCEDURE CALC_SUMA (VAR X : MAT; PPI : VECTOR);
VAR
  I, J, V, N, COL : INTEGER;
  S              : REAL;
BEGIN
  N := 12;
  FOR J := 1 TO 100 DO
    BEGIN
      COL := 0;
      FOR V := 1 TO 12 DO
        BEGIN
          S := 0;
          COL := COL + 1;
          S := S + PPI[I] * X[J,V+I];
          X[J,COL] := S;
        END;
      END;
    END;
  END;
END;

```

```

PROCEDURE LEE (VAR SGMA, VAL_T : REAL);
BEGIN
    CLRSCR;
    GOTOXY (4, 4);
    WRITE ( 'SGMA = ');
    READLN (SGMA);
    GOTOXY (4, 14);
    WRITE (DAME EL VALOR DE "T" DE LA TABLA ');
    GOTOXY (4, 16);
    WRITE ( 'VAL_T = ');
    READLN (VAL_T);
END;

PROCEDURE CALC_LAMDA (VAR X : MAT; SGMA : REAL; R2 : VECTOR);
VAR
    I, J : INTEGER;
BEGIN
    FOR I := 1 TO 12 DO
        FOR J := 1 TO 100 DO
            X[J,I] := X[J,I] / (SQRT(R2[I]) * SGMA);
        END;
    END;

PROCEDURE CALC_MAX (VAR X : MAT);
VAR
    MAXIMO : ARRAY[1..100] OF REAL;
    MAX : REAL;
    I, J : INTEGER;
BEGIN
    FOR I := 1 TO 100 DO
        BEGIN
            MAX := ABS(X[I,1]);
            FOR J := 1 TO 11 DO
                IF MAX >= ABS(X[I,J+1])
                THEN
                    MAX := MAX
                ELSE

```

```

        MAX := ABS(X[I,J+1])
        MAXIMO[I] := MAX;
    END;
FOR I := 1 TO 100 DO
    FOR J := 1 TO 12 DO
        IF ABS(X[I,J]) <> MAXIMO
        THEN
            X[I,J] := 0.0
        ELSE
            X[I,J] := MAXIMO[I];
    END;
END;

PROCEDURE COMPARA (VAR X: MAT; VAL_T : REAL);
VAR
    I, J : INTEGER
BEGIN
    FOR I := 1 TO 100 DO
        FOR J := 1 TO 12 DO
            BEGIN
                IF X[I,J]>= VAL_T
                THEN
                    X[I,J] := 1.0
                ELSE
                    X[I,J] := 0.0
            END;
        END;
    END;
END;

PROCEDURE IMPRIME (VAR X : MAT);
VAR
    I, J : BYTE;
BEGIN
    FOR I := 1 TO 100 DO
        BEGIN
            FOR J:= 1 TO 12 DO
                WRITE (LST, X[I,J] : 2 : 1 , ' ');
            WRITELN (LST);
        END;
    END;
END;

```

```

        END;
END;

PROCEDURE DATA_PPI;
BEGIN
    PPI[0] :=      ; PPI[1] :=      ; . . . PPI[11] :=      ;
END;

PROCEDURE PASADA (VAR X : MAT);
VAR
    X      : MAT;
    R2     : VECTOR;
BEGIN
    INICIO (X);
    TRNSFORMA (Y, X);
    DATA PPI;
    CALC_R2; (PPI, R2);
    CALC_SUMA (X, PPI);
    CALC_LAMDA (X, SGMA, R2);
    CALC_MAX (X);
    COMPARA (X, VAL_T);
    IMPRIME (X);
END;

BEGIN
    CLRSCR;
    INICIO (Y);
    LEE (SGMA, VAL_T);
    PASADA (Y);
END.

```

A partir de este algoritmo puede encontrarse también la sucesión indicadora E.1 para el pronóstico de puntos aberrantes innovados, puntos aberrantes aditivos y cambios transitorios, determinando los valores de PPI en la misma forma que en el apéndice C.

El valor de la constante SGMA corresponde a la desviación estándar del pronóstico, VAL_T corresponde al valor de $t_{\alpha/2, k-1}$ en las tablas de la distribución t. PRON.DAT es el nombre del archivo donde se guardan 100 simulaciones de $Y_{n+1}, Y_{n+2}, \dots, Y_{n+k}$, el nombre del archivo puede transformarse según convenga. Para el valor N en la sucesión indicadora D.1 se tomó 100 debido a que la memoria de la máquina no permitió trabajar archivos más grandes, corriendo 10 veces el programa se puede obtener D.1 con $N = 1000$. El valor 12 se estableció para pronosticar 12 pasos adelante, pero este valor corresponde al valor de k y puede transformarse según convenga. El valor de $Y[J, 0]$ en PROCEDURE TRANSFORMA es el valor de Y al tiempo n; si en el modelo el polinomio $\Pi(B)$ fuera de orden r, deben proporcionarse los valores de $Y_n, Y_{n-1}, \dots, Y_{n-r+1}$.

APENDICE F

PROGRAMA PROBABILIDAD

En este apéndice se desarrolla un algoritmo para calcular la probabilidad de que un cambio único de nivel ocurra en determinado intervalo de tiempo. Este algoritmo se basa en el teorema 3.2.1 con $v = 1$, esto es

$$P[h, k \text{ sean 1-próximos}] = P[E_1 \leq k-h] - P[E_2 \leq k-h].$$

Sea $k-h = r$, entonces

$$P[E_1 \leq r] = P[W_1 \leq r]$$

y

$$\begin{aligned} P[E_2 \leq r] &= P[W_1 = 1] P[W_2 \leq r-1] + P[W_1 = 2] P[W_2 \leq r-2] \\ &+ \dots + P[W_1 \leq r-2] P[W_2 \leq 2] \\ &+ P[W_1 = r-1] P[W_2 = 1] \end{aligned}$$

Dado lo laborioso que puede resultar el cálculo de esta suma para r grande ($r > 3$) se presenta el siguiente algoritmo para calcular $P[h, k \text{ sean 1-prox}]$.

```
PROGRAM PROBABILIDAD;
USES
  CRT, PRINTER;
TYPE
  VECTOR = ARRAY[1..18] OF REAL;
VAR
  X, EV, EVI, P : VECTOR;
```

```
PROCEDURE INI_EV (VAR EV : VECTOR);
```

```
VAR
```

```
  I : INTEGER;
```

```
BEGIN
```

```
  FOR I := 1 TO 18 DO
```

```
    EV[I] := 0.0;
```

```
END;
```

```
PROCEDURE INI_EVI (VAR EVI : VECTOR);
```

```
VAR
```

```
  I : INTEGER;
```

```
BEGIN
```

```
  FOR I := 1 TO 18 DO
```

```
    EVI[I] := 0.0;
```

```
END;
```

```
PROCEDURE INI_P (VAR P : VECTOR);
```

```
VAR
```

```
  I : INTEGER;
```

```
BEGIN
```

```
  FOR I := 1 TO 18 DO
```

```
    P[I] := 0.0;
```

```
END;
```

```
PROCEDURE DATA;
```

```
BEGIN
```

```
X[1]:= ; X[2]:= ; ...; X[18]:= ;
```

```
END;
```

```
PROCEDURE CAL_EV ( X : VECTOR; VAR EV : VECTOR);
```

```
VAR
```

```
  I, N : INTEGER;
```

```
  S : REAL;
```

```
BEGIN
```

```
  FOR N := 1 TO 18 DO
```

```
    BEGIN
```

```

        S := 0.0;
        FOR I := 1 TO N DO
        BEGIN
            S := S + X[I];
            EV[N] := S;
        END;
    END;
END;

```

```

PROCEDURE CAL_EVI ( EV : VECTOR; VAR EVI : VECTOR);
VAR
    I, N      : INTEGER;
    S          : REAL;
BEGIN
    FOR N := 1 TO 18 DO
    BEGIN
        S := 0.0;
        FOR I := 1 TO N-1 DO
        BEGIN
            S := S + X[I] * EV[N-I];
            EVI[N] := S;
        END;
    END;
END;

```

```

PROCEDURE CAL_PROB (EV, EVI : VECTOR; VAR P : VECTOR);
VAR
    N : INTEGER;
BEGIN
    FOR N := 5 TO 18 DO
        P[N] := EV[N] - EVI[N];
    END;

```

```

PROCEDURE IMPRIME (P : VECTOR);
VAR

```

```

I : INTEGER;
BEGIN
  FOR I := 5 TO 18 DO
    WRITE (P[I]:5, ' ');
    WRITELN;
  END;

```

```

BEGIN
  CLRSCR;
  DATA;
  INI_EV(EV);
  INI_EVI(EVI);
  INI_P(P);
  CAL_EV(X, EV);
  CAL_EVI(EV, EVI);
  CAL_PROB(EV, EVI, P);
  IMPRIME(P);
END.

```

El algoritmo fue diseñado para calcular la probabilidad de que h y k sean 1-próximos con $r = k-h$ variando entre 5 y 18, ya que así se requería en el ejemplo dado en 4.1, no obstante estos valores pueden cambiarse según convenga. $X[1]$, $X[2]$, ..., $X[18]$, representan aquí $P[W = 1]$, $P[W = 2]$, ..., $P[W = 18]$ respectivamente; estos valores se obtienen directamente de la distribución que se suponga para W .

REFERENCIAS

- Aitchison, J. y Dunsmore, I. R. (1975), *Statistical Prediction Analysis*. Cambridge University Press, New York.
- Blanco, H. y Garber, P. M. (1986), "Un modelo de los ataques especulativos contra el peso mexicano", *Estudios Económicos*, 1, 127-146.
- Box, G. E. P. y Jenkins, G. (1970), *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day.
- Box, G. E. P. y Tiao G. C. (1965), "A Change in Level of a non Stationary Time Series". *Biometrika*, 52, 181-192.
- Box, G. E. P. y Tiao, G. C. (1976). "Comparison of forecast and actuality". *Applied Statistics*, 25, 195.
- Chen, C. y Tiao G. C. (1990), "Random Level-Shift Time Series Models, ARIMA Aproximations, and Level-Shift Detection". *Journal of Business and Economic Statistics*, Vol 8, n^o 1, 83-97.
- Chernof, H. y Zacks, S. (1964), "Estimating the current mean of a Normal Distribution which is subjected to Changes in Time". *The Annals of Mathematical Statistics*, 35, 999-1018.
- Dawid, A. P. (1982), "The Well Calibrated Bayesian". *Journal of the American Statistical Association*, Vol. 77, 605-613.
- Dawid, A. P. (1984), "Statistical Theory the Precuential Approach". *Journal of the Royal Statistical Society: series A*, 147, 278-292.
- Gardner, L. A. (1969), "On detecting changes in the mean of normal variates". *The Annals of Mathematical Statistics*, 40, 116-126.

Guerrero, V. M. (1983), *Análisis Estadístico de Series de Tiempo Económicas*. Libro no publicado, disponible en fotocopias, México.

Guerrero, V. M. (1987), *Los Vectores Autorregresivos como Herramienta de Análisis Económico*. Documento de Investigación del Banco de México. N° 64.

Islas, A. (1989), "Métodos para determinar el orden de modelos autorregresivos y de promedios móviles ARMA(p,q) para series de tiempo". Tesis presentada para obtener el grado de Maestro en Matemáticas. UAM.

Jun, D. B. (1989), "On Detecting and Estimating a Major Level or Slope Change in General Exponential Smoothing". *Journal of Forecasting*, 8, 55-64.

Kennedy, W. J. y Gentle, J. E. (1980), *Statistical computing*. Marcel Dekker Inc.

Kling, J. L. (1987), "Predicting the Turning Points of Business and Economic Time Series" *Journal of Business*, 60, 201-238.

Leiva, R. A. (1987), "Un nuevo Proceso Estocástico para la modelación de fenómenos económicos. Definición y estimación". Trabajo presentado en Mar del Plata, Argentina. XV Coloquio de la Sociedad Argentina de Estadística.

Niven, I. y Zuckerman, H. S. (1976), *Introducción a la Teoría de los Números*. Editorial Limusa, México.

Quinn, B. C. (1980), "Order Determination for a Multivariate Autoregression". *Journal of the Royal Statistical Society, series B*, 42, 182-185.

Sóbol, I. M. (1976), "Método de Montecarlo", *Lecciones Populares de*

~~Matemáticas~~. Editorial Mir, Moscú.

Tsay, R. S. (1987). "Detecting and Modeling Changes in a Time Series". Comentario al artículo de Kitagawa, G.. *Journal of the American Statistical Association*, 82.

Wecker, W. (1979), "Predicting the Turning points of a Time Series", *Journal of Business*, 52, 35-50.

Zellner, A. (1988), "Turning points in Economic Time Series, loss structure and Bayesian forecasting". Presentado en la VIII Reunión Econométrica Latinoamericana, celebrada en agosto de 1988, en San José, Costa Rica.