Análisis estadístico de la tendencia por polinomios ortogonales

Por

F. J. URBELZ IBARROLA

Catedrático de Estadística
Actuario de Seguros

INTRODUCCION

1. En el estudio analítico de las series económicas si se eliminan las variaciones estacionales y cíclicas tenemos una serie residual compuesta de la tendencia y la perturbación aleatoria.

Es costumbre la descomposición de una serie estadística si se emplea un modelo aditivo en una suma de componentes:

tendencia
variaciones estacionales y
variaciones cíclicas.

Se añade a las anteriores componentes las variaciones de azar del proceso.

La tendencia se considera como una función determinista, y es preciso señalar que eliminadas las otras componentes obtendremos la tendencia más una perturbación aleatoria. Este valor es el que consideramos para analizar la tendencia.

2. Así como hemos expuesto que la serie puede considerarse como un modelo aditivo pudiera plantearse un modelo multiplicativo: este caso se reduce al anterior si tomamos logaritmos.

Pero en la práctica, a veces, cuando los elementos multiplicativos son negativos es conveniente combinar ambos métodos.

3. La tendencia, variaciones estacionales y cíclicas se estudian más racionalmente por análisis espectral; es decir: en el dominio de la frecuencia. Pero también se utilizan modelos de tendencia de tipo parabólico cuando previamente se han eliminado las influencias de las componentes periódicas.

4. Elegido el tipo de tendencia interesa analizar la conveniencia de expresar la función temporal por una parábola de grado k o por medio de polinomios ortogonales centrados.

Para mejor comprensión del tema, el Capítulo I lo dedico al estudio y propiedades de los Polinomios Ortogonales, teoremas que se deducen y fórmulas generales de los polinomios centrales.

5. En el Capítulo II justifico el modelo de la tendencia y establezco unas pocas fórmulas utilizadas como filtros para eliminar los efectos de las variaciones estacionales y cíclicas e igualmente estudio relaciones entre modelos parabólicos y los mismos expresados en forma de polinomios ortogonales.

Termino este capítulo con una axiomática de la tendencia más la perturbación aleatoria.

6. El Capítulo III lo dedico a las distribuciones de probabilidad de la función conjunta del vector ϵ y su función característica —n— dimensional e, igualmente, descomposiciones del vector ϵ para conocer sus distribuciones marginales y condicionadas y funciones características.

Igualmente, por una transformación obtengo la función de densidad —n— dimensional del modelo estocástico (tendencia más vector aleatorio) y funciones características, etc,

7. El Capítulo IV lo dedico a la estimación por punto obteniendo las estimaciones de los coeficientes regresores poblacionales, así como algunas de sus propiedades, como son de ser centrados y la matriz de covarianzas, deduciendo que estos estimadores son independientes.

Esta demostración es tan importante que no podemos afirmar lo mismo de las estimaciones de los coeficientes de una parábola, por lo que este método de representar por polinomios ortogonales supera a aquél.

Las desviaciones y sus cuadrados se analizan, así como la estimación de la varianza del modelo.

Terminamos este capítulo estudiando si puede obtenerse un estimador centrado, óptimo mejor que el obtenido, y comprobamos que es precisamente un estimador ELIO.

8. El Capítulo V lo dedico a la estimación por intervalo. Y, previamente, estudio las distribuciones de los estimadores, funciones características —n— dimensionales y unidimensionales, así como las funciones de densidad.

La independencia entre las desviaciones y los coeficientes regresores nos permite aplicar el Teorema de Descomposición de la X^2 de Pearson para conocer distintas distribuciones de los estadísticos y establecer intervalos de confianza según se conozca o no la varianza.

Además de estudiar intervalos de confianza se estudian esquemas para el análisis del grado de la tendencia.

9. El Capítulo VI se dedica a la Teoría de la Predicción de la Tendencia.

Y estableciendo un predictor lineal se estudian las condiciones para que sea centrado y óptimo. Esta optimización de la varianza nos permite obtener, conocida la distribución del predictor, intervalos de confianza que contenga la tendencia teórica —no aleatoria— o también el proceso estocástico de la tendencia.

10. El Capítulo VII y último está dedicado a complementar las ideas expuestas: Aplicar a un caso el método de Polinomios Ortogonales, y se acompaña una página de nuestras Tablas Estadísticas para comprender mejor el ajuste y la estimación por Polinomios Orotogonales.

Se forman diversos cuadros para la mejor comprensión de los valores ajustados (polinomios ortogonales hasta segundo y cuarto grado), intervalos de confianza de los estimadores, de la varianza, hipótesis sobre el grado de tendencia y la aplicación a la predicción.

Como observación importante hemos adoptado el punto para expresar la coma decimal, notación que el lector que utilice computadores le será fácil de comprender.

11. Terminamos el trabajo con una pequeña bibliografía y un índice sobre las materias tratadas.

EL AUTOR

CAPITULO I

Polinomios Ortogonales

SECCIÓN 1.ª DEFINICIONES

I. *Funciones ortogonales de variables estadísticas*

1. Dada la variable estadística x definida en D (para $\forall x_i \in D$) [1.1], las funciones $\{\varphi_j(.)\}$ [1.2] ($j = o, 1, 2 \dots k$) son ortogonales si cumplen las condiciones siguientes:

$$\sum_{i=1}^{n} \varphi_j(x_i)\, \varphi_r(x_i)\, h_i = 0 \qquad \text{Para } j \neq r \qquad [1.3]$$

$$\sum_{i=1}^{n} \varphi_j(x_i)^2\, h_i \neq 0 \qquad \text{Para } j = r \qquad [1.4]$$

donde h_i es la frecuencia relativa de x_i:

$$\sum_{i=1}^{n} h_i = 1 \qquad [1.5]$$

Si la [1.4] fuere para todo j la unidad, las funciones [1.2] se denominan normalizas. La variable estadística x tiene dos campos: campo de variabilidad $x \in D$ y campo de frecuencias h_i ($i = 1, \dots n$) · [1.6]

2. Casos particulares: si las funciones [1.2] son polinomios $\{P_o(x), \dots P_s(x), \dots P_k(x)\}$ {1.6} y el de grado $P_s(x)$ es:

$$P_s(x) = b_{so} + b_{s1}\, x + \dots + b_{s,\, s-1}\, x^{s-1} + x^s \qquad [1.7]$$

asociados a sus frecuencias h_i para $x_i \in D$ son otorgonales si cumplen las condiciones [1.3] y [1.4].

El polinomio $P_o(x) = 1$ para $\forall x_i \in D$. [1.8]

II. *Polinomios ortogonales definidos en un conjunto de números naturales* $E = \{1, 2, \dots n\}$ [1.9]

1. Dada la variable $x \in E$ y, los polinomios $\{P_o(x), \dots P_s(x), \dots P_k(x)\}$ [1.10] (el de grado s es la forma [1.7]) son ortogonales si cumplen las condiciones siguientes:

$$\sum_{x=1}^{n} P_j(x)\ P_r(x)=0 \qquad \text{para } j \neq r \qquad [1.11]$$

$$\sum_{x=1}^{n} P_j(x)^2 \neq 0 \qquad \text{para } j=r \qquad [1.12]$$

donde $P_o(x)=1$ para $\forall x \in E$ [1.13]

2. Observemos una notable diferencia respecto a la definición anterior: allí, la variable está definida en $x \in D$, mientras que E es el conjunto parcial de números naturales y no se asocia a frecuencias.

III. *Polinomios ortogonales centrados definidos en un conjunto de números naturales* E [1.9]

Dados los polinomios $\{P_o(x), \dots P_s(x), \dots P_k(x)\}$ [1.14] definidos en E, donde $P_o(x)=1$ y el de grado de $P_s(x)$ — es:

$$P_s(x)=b_{so}+b_{s1}(x-\bar{x})+\dots+b_{sj}(x-\bar{x})^j+\dots+b_{s,s-1}(x-\bar{x})^{s-1}+(x-\bar{x})^s \text{ para } s=1,2,3\dots k \qquad [1.16]$$

son ortogonales si cumplen las condiciones [1.11] y [1.12].

En esta monografía utilizamos estos polinomios centrados y los llamamos simplemente polinomios. La notación será distinta para la variable que utilizamos la t representativa del parámetro tiempo.

SECCIÓN 2.ª

Teorema Fundamental

1. *Teorema*

Hipótesis

Dado un sistema de polinomios ortogonales centrados (Def. III) y dos números enteros $o \leq h < s$

Tesis

Para $\forall h < s$ se cumple:

$$\sum_{t=1}^{n} (t-\bar{t})^h\ P_s(t)=o \qquad t \in E \qquad [1.17]$$

siendo $\bar{t}=\dfrac{n+1}{2}$

Demostración

Si la [1.17] se verifica para $h-1$

$$\sum_{t=1}^{n} (t-\bar{t})^{h-1} P_s(t) = 0 \qquad h-1<s \qquad [1.17']$$

(el exponente puede ser desde o hasta $h-1$), demostremos se cumple para h si $h<s$.

Por hipótesis, $P_h(t)$ y $P_s(t)$ son ortogonales. Si el polinomio $P_h(t)$ [1.16] tenemos:

$$\sum_{t=1}^{n} P_h(t)\, P_s(t) = o = \sum_{t=1}^{n} [b_{ho} + b_{h1}(t-\bar{t}) + \ldots + b_{h,h-1}(t-\bar{t})^{h-1} + (t-\bar{t})^{h}]\, P_s(t) =$$

$$= b_{ho} \sum_{t=1}^{n} P_s(t) + b_{h1} \sum_{t=1}^{n} (t-\bar{t})\, P_s(t) + \ldots + b_{h,h-1} \sum_{t=1}^{n} (t-\bar{t})^{h-1}\, P_s(t) +$$

$$+ \sum_{t=1}^{n} (t-\bar{t})^{h}\, P_s(t)$$

Todos los sumandos que preceden al último son nulos por hipótesis [1.17'] y el último por ortogonalidad. El primer sumando no es sino $\Sigma\, P_o(t)\, P_s(t) = \Sigma\, P_s(t) = o$ según [1.12].

Luego

$$\sum_{t=1}^{n} (t-\bar{t})^{h}\, P_s(t) = o \qquad [1.17'']$$

La [1.17'] implica se cumple para el siguiente si $h<s$. Luego por ser:

$$\sum_{t=1}^{n} P_o(t)\, P_s(t) = o \qquad \sum_{t=1}^{n} (t-\bar{t})\, P_s(t) = o \qquad o<s$$

$$\sum_{t=1}^{n} P_1(t)\, P_s(t) = o = \Sigma [b_{1o} + (t-\bar{t})] \quad P_s(t) =$$

$$= b_{1o} \sum_{t=1}^{n} P_s(t) + \sum_{t=1}^{n} (t-\bar{t})\, P_s(t) \Rightarrow$$

$$\sum_{t=1}^{n} (t-\bar{t})\, P_s(t) = o \qquad 1<s$$

El teorema queda totalmente demostrado, porque [1.17] se cumple para $h=o$, 1 y se cumple para $h=2$ por la [1.17'] y así sucesivamente siempre que $h<s$.

2. *Corolarios*

1. Si $h=s$

$$\sum_{t=1}^{n} (t-\bar{t})^{s}\, P_s(t) = \sum_{t=1}^{n} P_s(t)^2 = p_s^2 > 0 \qquad [1.18]$$

porque

$$\sum_{t=1}^{n} P_s(t)\,P_s(t)=\sum_{t=1}^{n}[b_{so}+\ldots+b_{sh}(t-\bar{t})^h+\ldots+b_{s,s-1}(t-\bar{t})^{s-1}+(t-\bar{t})^s]\,P_s(t)$$

Todos los sumandos son nulos por [1.17] excepto el último, que es [1.18].

2. Para todo sistema de polinomios centrados según la Def. III y dos números $o\leq h<s$ tenemos:

$$\sum_{t=1}^{n} t^h\,P_s(t)=o \qquad [1.19]$$

así:

$$\sum_{t=1}^{n}(\,(t-\bar{t})+\bar{t})^h\,P_s(t)=\sum_{t=1}^{n}\sum_{j=o}^{h}\binom{h}{j}(t-\bar{t})^j\,\bar{t}^{\,h-j}\,P_s(t)=$$

$$=\sum_{h=o}^{h}\binom{h}{j}\bar{t}^{\,h-j}\sum_{t=1}^{n}(t-\bar{t})^j\,P_s(t)=o$$

Por ser $j\leq h<s$, luego queda probado la [1.19].

3. Para todo sistema de polinomios ortogonales centrados también se cumple:

$$\sum_{t=1}^{n} t^s\,P_s(t)=\sum_{t=1}^{n} P_s(t)^2 \qquad [1.20]$$

Esta demostración es inmediata:

$$\sum_{t=1}^{n} t^s\,P_s(t)=\sum_{j=1}^{s}\binom{s}{j}\,\bar{t}^{\,s-j}\sum_{t=1}^{n}(t-\bar{t})^j\,P_s(t)$$

De esta expresión son nulos todos los sumandos por [1.17] excepto para $j=s$ y no es sino la [1.18].

SECCIÓN 3.ª

1. *Transformación de polinomios*

1. Dado un conjunto de constantes λ_h $(h=o,1,\ldots k)$ [1.21] no nulas si multiplicamos los polinomios [1.15] por estas constantes, obtenemos otros polinomios relacionados por la expresión:

$$P_h^1(t)=\lambda_h\,P_h(t) \qquad h=o,1,2,\ldots k \qquad [1.22]$$

y también son ortogonales.

En efecto:

$$\sum_{t=1}^{n} P_s^1(t)\,P_h^1(t)=\lambda_s\lambda_h\sum_{t=1}^{n} P_s(t)\,P_h(t)\;=0 \qquad s\neq h$$

porque $\lambda_s,\lambda_h\neq 0$ y por la Def. III.

2. Polinomios normalizados. Pueden obtenerse si elegimos las constantes [1.21], según [1.18].

$$\lambda_h = \frac{1}{p_h} \qquad [1.23]$$

y sustituidas en [1.21] tenemos

$$P_h^{|}(t) = \frac{P_h(t)}{p_h} \qquad [1.24]$$

luego

$$\sum_{t=1}^{n} P_h^{|}(t)^2 = \frac{1}{p_h^2} \sum_{t=1}^{n} P_h(t)^2 = 1 \qquad [1.25]$$

por [1.18]. Esta es la condición de normalización.

Los polinomios $\{P_s^{|}(t)\}$ son ortonormalizados.

3. *Polinomios enteros.* Aunque $t \in E$ los polinomios [1.15] no tienen por qué ser enteros, pero pueden elegirse las constantes [1.21] para que lo sean.

En nuestras Tablas Estadísticas (1) hemos tabulado polinomios ortogonales eligiendo las constantes [1.21] para que sus transformados [1.22] sean números enteros. Seguimos criterio semejante al de Fisher y Yates.

En las tablas aparece λ_h como un quebrado

$$\lambda_h = \frac{m_h}{d_h} \qquad [1.26]$$

donde m (multiplicador) y d (divisor) depende de h y de n.

Estos números se consignan al pie de cada polinomio, como también los cuadrados de los valores de éstos.

SECCIÓN 4.ª

Fórmula general de un polinomio

1. *Momentos totales centrados*

Definimos momento central total de orden h a la expresión:

$$M_h = \sum_{t=1}^{n} (t-\bar{t})^h \qquad t \in E \qquad [1.27]$$

donde E es el conjunto indicado en [1.9].

(1) URBELZ IBARROLA, F. J., y PÉREZ MAURA, Alvaro: *Tablas Estadísticas.* Departamento de Estadística, Empresariales, Santander.

El valor de M_h depende de K:

$$M_o = n \qquad M_{2k} \neq o \qquad M_{2k-1} = o \qquad [1.28]$$

2. *Sistema de ecuaciones*

Si el valor de $P_s(t)$ [1.16] lo sustituimos en [1.17], tenemos la siguiente expresión, y con la notación de [1.27] es:

$$\sum_{t=1}^{n} (t-\bar{t})^h P_s(t) = \sum_{t=1}^{n} (t-\bar{t})^h \left(\sum_{j=o}^{s-1} b_{sj} (t-\bar{t})^j + (t-\bar{t})^s \right) =$$

$$= \sum_{j=o}^{s-1} b_{sj} \sum_{t=1}^{n} (t-\bar{t})^{j+h} + \sum_{t=1}^{n} (t-\bar{t})^{h+s} =$$

$$= b_{so} M_h + b_{s1} M_{h+1} + b_{s2} M_{h+2} + \ldots + b_{s,s-1} M_{h+s-1} + M_{h+s} = 0 \qquad [1.29]$$

Válida para todo $h < s$ según [1.17].

La [1.29] es una ecuación para un valor concreto de h si damos valores a h $(h = o, 1, 2 \ldots s-1)$ tenemos un sistema de s ecuaciones con otras tantas incógnitas que nos permite determinar los coeficientes b_{sj} $(j = o, \ldots, s-1)$.

3. *Fórmula general de los polinomios*

La [1.16] puede escribirse pasando $P_s(t)$ al otro miembro:

$$b_{s0} + b_{s1}(t-t) + b_{s2}(t-\bar{t})^2 + \ldots + b_{s,s-1}(t-\bar{t})^{s-1} + (t-\bar{t})^s - P_s(t) = o \qquad [1.30]$$

Las s ecuaciones [1.29] y la [1.30] forma un sistema de $s+1$ ecuaciones con s incógnitas (b_{sj} $j = o, 1, \ldots s-1$) y para su compatibilidad es necesario y suficiente que el determinante formado por los coeficientes de las incógnitas y los términos independientes sea nulo. (Rouché-Frobenius.)

Así, de [1.29] dando valores a $h = o, 1, \ldots s-1$ y con la [1.30], formamos el determinante siguiente para que sea compatible:

$$\begin{vmatrix} M_o & M_1 & M_2 & \ldots M_{s-1} & M_s \\ M_1 & M_2 & M_3 & \ldots M_s & M_{s+1} \\ \ldots\ldots\ldots & & & & \\ M_{s-1} & M_s & M_{s+1}\ldots & M_{2s-2} & M_{2s-1} \\ 1 & (t-\bar{t}) & (t-\bar{t})^2\ldots & (t-\bar{t})^{s-1} & (t-\bar{t})^s - P_s(t) \end{vmatrix} = 0$$

$(t-\bar{t})^s - P_s(t)$ es un elemento de la última columna del determinante.

Si añadimos a los elementos de la última columna de [1.31] un cero podemos descomponer en suma de dos determinantes que tienen las restantes columnas iguales:

$$\begin{vmatrix} M_o & M_1 & \dots & M_{s-1} & M_s \\ M_1 & M_2 & \dots & M_s & M_{s+1} \\ \dots & \dots & \dots & \dots & \dots \\ M_{s-1} & M_s & \dots & M_{2s-2} & M_{2s-1} \\ 1 & (t-\bar{t}) & \dots & (t-\bar{t})^{s-1} & (t-\bar{t})^s \end{vmatrix} + \begin{vmatrix} M_o & M_1 & \dots & M_{s-1} & O \\ M_1 & M_2 & \dots & M_s & O \\ \dots & \dots & \dots & \dots & \dots \\ M_{s-1} & M_s & \dots & M_{2s-2} & O \\ 1 & (t-\bar{t}) & \dots & (t-\bar{t})^{s-1} & -P_s(t) \end{vmatrix} =$$

$$= N_s(t) - P_s(t) \begin{vmatrix} M_o & M_1 & \dots & M_{s-1} \\ M_1 & M_2 & \dots & M_s \\ \dots & \dots & \dots & \dots \\ M_{s-1} & M_s & \dots & M_{2s-2} \end{vmatrix} = o \Rightarrow \qquad [1.31']$$

$$P_s(t) = \frac{N_s(t)}{D_s} \qquad [1.31'']$$

utilizando la notación $N_s(t)$ para el primer sumando de [1.31'] y D_s para el coeficiente de momentos del denominador.

La fórmula general del polinomio s es:

$$P_s(t) = \frac{\begin{vmatrix} M_o & M_1 & M_2 & \dots & M_s \\ M_1 & M_2 & M_3 & \dots & M_{s+1} \\ M_2 & M_3 & M_4 & \dots & M_{s+2} \\ \dots & \dots & \dots & \dots & \dots \\ M_{s-1} & M_s & M_{s+1} & \dots & M_{2s-1} \\ 1 & (t-\bar{t}) & (t-\bar{t})^2 & \dots & (t-\bar{t})^s \end{vmatrix}}{\begin{vmatrix} M_o & M_1 & M_2 & \dots & M_{s-1} \\ M_1 & M_2 & M_3 & \dots & M_s \\ \dots & \dots & \dots & \dots & \dots \\ M_{s-1} & M_s & M_{s+1} & \dots & M_{2s-2} \end{vmatrix}} \qquad [1.31''']$$

Observaciones:

1.ª La diagonal principal y sus paralelas pares no son nulas. Las paralelas impares son nulas [1.28] excepto los elementos de la última línea del determinante del numerador.

2.ª Los coeficientes b_{sj} $(j=o, 1, 2, \dots s-1)$ se obtienen por simple desarrollo y son los adjuntos de $(t-\bar{t})^j$ divididos por D_s.

3.ª Los numeradores de [1.31'''], $\{s=0, 1, \dots k\}$ son ortogonales si hacemos una transformación [1.22] y elegimos las constantes $\lambda_h = D_h$ [1.32].

SECCIÓN 5.ª

Fórmulas para los primeros polinomios

1. *Obtención del primer polinomio*

Según la Def. III $P_o(t)=1$. $\quad s=o$.

En la fórmula [1.31'''] si hacemos $s=1$, el numerador es un determinante de segundo orden y el denominador es M_o.

Luego:

$$P_1(t) = \frac{\begin{vmatrix} M_o & O \\ 1 & (t-\bar{t}) \end{vmatrix}}{M_o} = t-\bar{t} = t - \frac{n+1}{2} \qquad [1.33]$$

porque $M_o = n$. y $\bar{t} = \dfrac{n+1}{2}$ y $M_1 = o$ según [1.28].

2. *Formación del segundo polinomio*

Recordemos la observación 1.ª de la sección anterior y hagamos en [1.31''] $s=2$:

$$P_2(t) = \frac{\begin{vmatrix} M_o & O & M_2 \\ O & M_2 & O \\ 1 & (t-\bar{t}) & (t-\bar{t})^2 \end{vmatrix}}{\begin{vmatrix} M_o & O \\ O & M_2 \end{vmatrix}} = (t-\bar{t})^2 - \frac{M_2}{M_o} \qquad [1.34]$$

Pero

$$\frac{M_2}{M_o} = \frac{\sum_{t=1}^{n} (t-\bar{t})^2}{n} = \frac{2}{n} \sum_{t \geq t}^{n} (t-\bar{t})^2 \qquad [1.35]$$

por cuanto $(t-\bar{t})^2$ es una función par.

Dos casos pueden presentarse según *n:*

a) *n* impar:

La $\bar{t} = \frac{n+1}{2}$ es un número entero =

$$\sum_{t \geq t}^{n} (t-\bar{t})^2 = o^2 + 1^2 + 2^2 + \ldots + (n - \frac{n+1}{2})^2 =$$

$$= 1^2 + 2^2 + \ldots + \left(\frac{n-1}{2}\right)^2 = \frac{\left(\frac{n-1}{2}\right)\left(\frac{n-1}{2}+1\right)\left(\frac{2\,(n-1)}{2}+1\right)}{6} =$$

$$= \frac{(n^2-1)\,n}{2.12} \qquad [1.36]$$

Hemos aplicado la célebre fórmula de la suma de los cuadrados de los *n* primeros números naturales:

$$S_n = \frac{n\,(n+1)\;(2n+1)}{6} \qquad [1.37]$$

siendo aquí *n* el número entero $\frac{n-1}{2}$ porque *n* es impar.

b) *n* par:

$$\sum_{t > t}^{n} (t-\bar{t})^2 = \left(\frac{1}{2}\right)^2 = \left(\frac{3}{2}\right)^2 + \ldots + \left(\frac{n-1}{2}\right)^2 = \frac{1^2 + 3^2 + \ldots + (\mathrm{n}-1)^2}{4}$$

Pero de la [1.37] deducimos:

$$s_n = 1^2 + 2^2 + 3^2 + 4^2 + \ldots + (n-1)^2 + n^2 = [1^2 + 3^2 + \ldots + (n-1)^2] +$$

$$+ 2^2 [\, 1^2 + 2^2 + \ldots + \left(\frac{n}{2}\right)^2 \,] \qquad \Rightarrow$$

$$\sum_{t > t}^{n} (t-\bar{t})^2 = \frac{1^2 + 3^2 + \ldots + (n-1)^2}{4} = \frac{s_n - 2^2\,(1^2 + 2^2 + \ldots + (\frac{n}{2})^2}{4} =$$

$$= \frac{n\,(n^2-1)}{2.12} \qquad [1.37']$$

Las fórmulas [1.36] y [1.37'] son idénticas.

Llegando a la [1.35] y sustituyendo en [1.34] el polinomio ortogonal de segundo orden en función de n (independiente de ser par o impar) es:

$$P_2(t) = (t - \frac{n+1}{2})^2 - \frac{n^2-1}{12} \qquad [1.38]$$

donde $\bar{t}$ se ha sustituido por la media.

3. *Polinomios de tercero, cuarto y quinto grado.*

De acuerdo con [1.31'''] y [1.28] para $s=3$ este polinomio es:

$$P_3(t) = \frac{\begin{vmatrix} M_o & O & M_2 & O \\ O & M_2 & O & M_4 \\ M_2 & O & M_4 & O \\ 1 & (t-\bar{t}) & (t-\bar{t})^2 & (t-\bar{t})^3 \end{vmatrix}}{\begin{vmatrix} M_o & O & M_2 \\ O & M_2 & O \\ M_2 & O & M_4 \end{vmatrix}}$$

y si hacemos operaciones tenemos la siguiente fórmula:

$$P_3(t) = (t - \frac{n+1}{2})^3 - \frac{3n^2-7}{20}(t - \frac{n+1}{2}) \qquad [1.38a]$$

Para $s=4$ es:

$$P_4(t) = (t - \frac{n+1}{2})^4 - \frac{3n^2-13}{14}(t - \frac{n+1}{2})^2 + \frac{3(n^2-1)(n^2-9)}{560} \qquad [1.38b]$$

y para $s=5$, igualmente:

$$P_5(t) = (t - \frac{n+1}{2})^5 - \frac{5n^2-7}{18}(t - \frac{n+1}{2})^3 + \frac{15n^4-230n^2+407}{1.008}$$

[1.38c]

Hemos omitido estas demostraciones sencillas (algo pesadas) y hacemos las aclaraciones siguientes:

1.ª Los polinomios centrados $(t-\bar{t})$ impares carecen de término independiente y los coeficientes de sus términos de lugar par son nulos.

2.ª Los polinomios centrados $(t-\bar{t})$ pares carecen de los términos impares y tienen términos independientes.

CAPITULO II

Modelo de la tendencia

SECCIÓN 1.ª

Problemas preliminares

1. *Componentes de un proceso estocástico*

1. Los datos experimentales de una serie temporal deben analizarse para conocer las componentes dinámicas que contiene y la estructura interna del proceso: si es de naturaleza estacionaria o evolutiva; tendencia, variaciones estacionales o cíclicas. Si la serie es estacionaria interesa conocer el tipo de proceso a que pertenece: perturbación aleatoria, medias móviles, autorregresivo, mixto, etc.

En anteriores artículos (1) he tratado extensamente estos temas.

En las ciencias experimentales de fenómenos físicos o semejantes pueden estudiarse mejor los procesos que en las series económicas.

Los economistas se han preocupado y se preocupan en estudiar estas series. Y admiten como componentes esenciales de los procesos las características siguientes:

— La tendencia.
— Las variaciones estacionales.
— Las variaciones cíclicas.
— Y, finalmente, las oscilaciones erráticas.

Alguna de ellas puede faltar.

2. En muchas ocasiones interesa aislar estas componentes para analizar la influencia de cada una y su posible composición ulterior.

3. Un estudio profundo y serio del proceso es prácticamente imposible cuando la información de la serie es escasa y también cuando las series económicas son subjetivas, porque no son muy aptas para un tratamiento estadístico.

Pero otras series económicas reflejan perfectamente la evolución de un fenómeno.

(1) Véase bibliografía.

Las series actuariales se prestan mejor a un tratamiento estadístico e, igualmente, las series económicas con numerosos datos objetivamente medibles.

4. Los errores de las series brutas deben eliminarse, y para atenuar y subsanar sus deficiencias conviene emplear «filtros» adecuados, que permitan mayor rigor en el estudio de las series derivadas.

5. La serie temporal observada es una muestra, y pretender conocer el proceso si la muestra no es numerosa resulta imposible.

2. *Proceso estocástico de la tendencia*

1. Examinemos con carácter general la descomposición de una serie temporal y sus correcciones necesarias para obtener la *componente de la tendencia* y eliminar otras componentes si las tuviera,

2. En primer lugar, aconsejamos hacer *una estimación del espectro del proceso* (2), porque por simple análisis nos indica sus componentes y la importancia de cada una de ellas.

3. Existen series actuariales que, basadas en ecuaciones diferenciales, se conocen las formas generales de la tendencia. Tales son, por ejemplo, la logística, las leyes de mortalidad, etc., sobre las cuales no tratamos ahora.

4. Si de una serie temporal estimamos las funciones de covarianza y la trasladamos al campo de la frecuencia obteniendo la transformada de Fourier, tenemos la función de densidad espectral, y si existieran acumulaciones de masas espectrales («con picos») en torno a la frecuencia $\lambda = o$ es muy significativo que la serie tenga tendencia: La tendencia se caracteriza por el movimiento profundo que tiene período infinito.

5. Pero si además de revelarnos la tendencia, la función de densidad espectral tuviere otros «picos» para diferentes valores de λ ($\lambda \neq o$), precisamente en estos valores *existirán variaciones estacionales o cíclicas.*

6. Como tratamos de la tendencia, nos interesa eliminar de la serie dadas las otras componentes. Los métodos son: o en el dominio de la frecuencia (utilizando el análisis espectral) o el dominio del tiempo.

La *serie temporal observada se transforma por ciertas operaciones en otra nueva serie temporal liberada* (o atenuada) *de las otras componentes,* y estas operaciones de cálculo con *la serie antigua para obtener la nueva* se denomina «*filtrado*» (3).

(2) URBELZ, F. Javier: «Estimación del espectro de las series temporales». *Anales I. Actuarios,* 1982.

(3) URBELZ, F. J.: «Procesos básicos econométricos y de sus transformaciones lineales». *I. Actuarios,* 1983.

Las funciones de densidad espectrales de la nueva y la antigua serie están relacionadas por una función que se denomina función de transferencia.

7. La serie transformada $\{ y_t, t \in E \}$ *liberada* de las otras componentes es la que estudiamos en esta monografía con algún detenimiento. Y *suponemos que* el filtro *no incida* sobre la perturbación aleatoria.

Por eso cuando escribimos sobre el *modelo de tendencia* debe entenderse que nos referimos al proceso estocástico $\{ y_t, t \in E \}$ que es una de las componentes de los procesos estocásticos según expusimos anteriormente.

La tendencia suele estudiarse como función exclusiva del tiempo y sin componentes aleatorias. Es más: puede ser función potencial combinada con términos trigonométricos, etc. Indicamos el concepto, pero prescindimos de ellos, aunque *le añadimos la perturbación aleatoria para* estudiar la serie observada como muestra procedente de un proceso estocástico.

3. *Filtros para determinar la tendencia*

1. Entre los filtros que indicamos a título orientativo (4) están los de medias móviles y el más sencillo de éstos es de Witsein:

$$a_j = \frac{1}{2m+1} \qquad j = 0, \pm 1, \pm 2 \ldots \pm m \qquad [2.1]$$

La aplicación práctica es que los datos observados y_t^* por medias consecutivas de $2m+1$ términos variando t se obtiene la nueva serie:

$$y_t = \sum_{j=-m}^{m} \frac{y_{t+j}^*}{2m+1} \qquad [2.2]$$

Y podemos asegurar:

1.º Contiene exactamente la tendencia, si m es pequeño (relacionado con el año).

2.º Atenúa las componentes estacionales y cíclicas.

La propiedad del filtro [2.1] que aplicado sobre la serie antigua nos da la nueva con la misma tendencia lo conocemos por el valor de la función de transferencia para $\lambda = o$: el límite de esta función para $\lambda \to o$ es precisamente la unidad (4).

---

(4) URBELZ, F. Javier: «Los procesos estocásticos econométricos y sus transformaciones lineales». *Anales I. Actuarios*, 1983.

La serie obtenida $\{ y_t, t \in E \}$ *contiene la tendencia*, pero *desconocemos si se han eliminado* las *otras componentes de tipo cíclico*. La función de densidad espectral de salida para $\lambda = o$ es idéntica a la de entrada.

2. Otras fórmulas dadas en combinaciones lineales (igualmente, de medias móviles) más específicas son:

$$y_t = \frac{1}{24} (y^*_{t-6} + 2y^*_{t-5} + ... + 2y^*_{t+5} + y^*_{t+6}) \qquad [2.3]$$

y se utiliza *cuando los datos observados son mensuales* y que eliminan o atenúan componentes de tipo estacional.

3. Existen series actuariales interesantes, y por citar algunas (basadas también en combinaciones lineales), citaremos la de Spencer (de 15 puntos):

$$\gamma_t = \frac{1}{320} [74 y^*_t + 67 (y^*_{t-1} + y^*_{t+1}) + 46 (y^*_{t-2} + y^*_{t+2}) +$$

$$+ 21 (y^*_{t-3} + y^*_{t+3}) + 3 (y^*_{t-4} + y^*_{t+4}) - 5 (y^*_{t-5} + y^*_{t+5}) -$$

$$- 6 (y^*_{t-6} + y^*_{t+6}) - 3 (y^*_{t-7} + y^*_{t+7}) \qquad [2.4]$$

e, igualmente, la de 21 puntos del mismo autor; las parábolas de Karup, Wolhouse, Higam, etc.

4. Las medidas móviles son particularmente interesantes, porque, a veces, expresiones tan sencillas son consecuencia de parábolas de segundo o tercer grado, como son, por ejemplo, las de Wolhouse, Higam, etc.

Un método bastante general de construcción de estimadores de medias móviles que tengan por base una parábola es utilizar un polinomio de segundo, tercero, cuarto grado e inclusive tomar $2m + 1$ términos de la serie y por mínimos cuadrados obtener expresiones en función de los datos observados. Kendall (5) utiliza mucho estas expresiones y obtiene numerosas fórmulas, de forma parecida a la siguiente:

$$y_t = \sum_{s=-m}^{m} c_s y^*_{t+s} \qquad c_s = c_{-s} \qquad [2.5]$$

donde las c_s se han obtenido previamente por ajuste y no son sino medias móviles. Determinado m y haciendo el ajuste, las c_s son idénticas, y puede variar t y filtrar la serie.

Métodos más complicados pueden verse en Hannan (6) y en Murray Rosenblatt (7).

(5) KENDALL, M. G., and STUART, A.: *The advanced theory of statistics*. Vol. III, pág. 366 y siguientes.

(6) HANNAN, E. J.: «Regresion for time series». Australian National University. *Rosemblatt*. Wiley, 1962.

(7) *Time Series Analysis* (Symposium). MURRAY ROSEMBLATT. *John Wiley*, 1962, New York.

5. Existe un método utilizado con éxito por Parzen, consistente en ajustar a los datos observados una regresión con retardos mensuales, bimensuales..., hasta treinta y seis meses, y después analizar aquellos términos que significativamente difieran de cero. Y con estos términos, efectuar operaciones para precisar mejor la tendencia, efectuando nuevos ajustes.

La serie se trata hasta que el espectro prácticamente es constante (el de la perturbación aleatoria). De esta forma obtiene la tendencia. Por ser interesante tenemos en estudio una aplicación a una serie económica por este método.

6. Si utilizamos un «filtro» para analizar la tendencia, es condición esencial *que la nueva serie contenga la tendencia y no se haya alterado.*

Un buen «filtro» será aquél que elimine componentes estacionales y cíclicas y deje pasar la tendencia intacta si es esta la finalidad perseguida.

La función de transferencia (que relaciona las funciones de densidades espectrales del proceso entrante y saliente) nos indica si elimina las componentes estacionales y cíclicas.

En mi artículo citado trato del *filtro ideal,* y de forma concluyente cómo contiene la tendencia en un intervalo en torno al origen y que contenga las masas espectrales fundamentales de la tendencia.

Exponemos su fórmula, aunque no es práctica:

$$y_t = \frac{\lambda_o y_t^*}{\pi} + \sum_{j=1}^{\infty} \frac{sen\,\lambda_o j}{\pi j} (y_{t-j}^* + y_{t-j}^*) \qquad |\lambda_o| < \epsilon \qquad [2.5]$$

De este filtro pueden tomarse aproximaciones limitando la sumatoria λ_o es muy pequeño.

7. Las series económicas caracterizadas por una f. de densidad espectral de naturaleza continua no tienen ciclos perfectos, como sucede en las series físicas, y por eso las concentraciones espectrales en torno a determinados valores de la frecuencia son características del «pseudo-ciclo» o la «pseudo-tendencia».

Si el espectro fuese de barras (es decir, función de cuantía para $\lambda = o$) tendríamos una perfecta tendencia. Y, análogamente, para los ciclos. En estos casos, el proceso estocástico puede descomponerse perfectamente en funciones *no estocásticas* que reflejarán la tendencia y las variaciones estacionales y cíclicas. En este caso, sí podríamos determinar una función matemática que recogiese las componentes citadas.

La realidad económica es muy distinta y el espectro de estas series es de tipo continuo, no discreto.

SECCIÓN 2.ª

Estudio del modelo

1. *Distintos modelos de tendencia*

Hemos comentado las dificultades para encontrar el modelo adecuado de la tendencia. Si por la naturaleza del mismo planteamos una ecuación matemática y resolvemos la tendencia en función del tiempo, determinaremos, por ajuste, sus coeficientes.

La tendencia puede ser de los tipos siguientes:

1. Parabólica.
2. Exponencial.
3. Hiperbólica.
4. Logarítmica.
5. Armónica.
6. Combinadas.

Pero aunque sean de tipo distinto al parabólico y tengamos el fundamento teórico que la tendencia (por ejemplo, la logística) sean de un tipo conocido, generalmente podemos desarrollar en serie y con más o menos grado de precisión tomaremos los términos necesarios para acotar el error.

2. *Modelo parabólico de la tendencia*

1. Si elegimos un modelo de tipo parabólico para expresar la tendencia, la componente y_t del vector **Y** puede escribirse:

$$y_t = (1, t, t^2, \dots t^k) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_k \end{pmatrix} + \epsilon_t = \sum_{h=0}^{k} \beta_h t^h + \epsilon_t \qquad [2.6]$$

Y dando a t valores $t = 1, 2, \dots n \quad t \in E$ tenemos todas las componentes del vector **Y** de n dimensiones.

Matricialmente, la [2.6] para $t \in E$ puede escribrirse

$$\mathbf{Y} = \mathbf{T}\beta + \epsilon \qquad k + 1 < n \qquad [2.6']$$

donde la matriz **T** es de orden $n \times (k + 1)$; β es el vector de orden $(k + 1) \times 1$ y **Y** y ϵ son ambos vectores de $n \times 1$.

2. El modelo parabólico de grado k puede representarse en función de los polinómios ortogonales estudiados en el capítulo anterior.

Así, la componente [2.5] puede escribirse:

$$\mathbf{Y}_t = \{ P_o(t), P_1(t), P_2(t), \ldots P_k(t) \} \begin{pmatrix} \gamma_o \\ \gamma_1 \\ . \\ . \\ . \\ \gamma_k \end{pmatrix} + \epsilon_t$$

$$= \sum_{h=o}^{k} \gamma_h P_h(t) + \epsilon_t \qquad [2.6"]$$

todas las observaciones pueden representarse:

$$\mathbf{Y} = \mathbf{P}\gamma + \epsilon \qquad [2.7]$$

Remitimos al lector al capítulo primero para conocer las propiedades de los polinomios *ortogonales centrados* que son los que utilizaremos.

3. Los modelos [2.6] y [2.7] son polinomios idénticos si desarrollamos los polinomios ortogonales en función de las potencias enteras de t.

4. Las estimaciones y las funciones de distribución y de densidad conjunta de los estimadores de β son más complejas que la función de densidad conjunta de las estimaciones de las γ.

3. *Relaciones entre los coeficientes de las distintas representaciones de la tendencia parabólica*

1. Aunque no son únicas las representaciones parabólicas del mismo grado [2.6] y [2.7] estudiamos esta identidad.

2. Si igualamos la [2.6] y [2.6"] y simplificamos, tenemos la siguiente relación:

$$\sum_{h=o}^{k} \beta_h t^h = \sum_{j=o}^{k} P_j(t)\gamma_j \qquad [2.8]$$

La [2.8] expresa la tendencia como una función del tiempo sin la componente aleatoria.

Si multiplicamos la [2.8] por $P_j(t)$ y sumamos $\forall t \in E$, tenemos:

$$\sum_{h=o}^{k} \beta_h \sum_{t=1}^{n} t^h P_j(t) = \sum_{j=1}^{n} P_j^2(t)\gamma_j = \gamma_j p_j^2$$

porque $\sum_{t=1}^{n} P_j(t) P_h(t) = o \quad j \neq h$ según la def. III, [1.12] y [1.18].

La anterior se simplifica por [1.19], luego:

$$\gamma_j p_j^2 = \sum_{h=j}^{k} \beta_h \sum_{t=1}^{n} t^h P_j(t) \quad \Rightarrow$$

$$\gamma_j = \frac{\sum_{h=j}^{k} \beta_h \sum_{t=1}^{n} t^h P_j(t)}{p_j^2} \qquad [2.9]$$

por anularse $\sum_{t=1}^{n} t^h P_j(t)$ para $h<j$.

3. Las expresiones matriciales [2.6] y [2.7] son idénticas, luego las funciones no aleatorias $\mathbf{T}\boldsymbol{\beta}$ y $\mathbf{P}\boldsymbol{\gamma}$ también:

$$\mathbf{T}\boldsymbol{\beta} = \mathbf{P}\boldsymbol{\gamma} \quad \Rightarrow \qquad [2.10]$$

$$\boldsymbol{\beta}'\mathbf{T}' = \boldsymbol{\gamma}'\mathbf{P}' \quad \Rightarrow \qquad [2.11]$$

Premultiplicando [2.10] por la matriz $\mathbf{T}'$ el vector columna $\boldsymbol{\beta}$ es:

$$\boldsymbol{\beta} = (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{P}\boldsymbol{\gamma} \qquad [2.12]$$

Igualmente, premultiplicando la misma ecuación matricial por $\mathbf{P}'$ y despejando $\boldsymbol{\gamma}$, tenemos:

$$\boldsymbol{\gamma} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\mathbf{T}\boldsymbol{\beta} \qquad [2.13]$$

Las [2.12] y [2.13] nos dan las relaciones entre los parámetros $\boldsymbol{\beta}$ en función de los parámetros γ o viceversa, aunque la componente γ_j $(j=o, 1, 2 \ldots k)$ la hemos obtenido en [2.9].

Por las propiedades de los polinomios ortogonales estudiadas en el capítulo I, la [2.13] en forma desarrollada (previo cálculo de las operaciones $(\mathbf{P}'\mathbf{P})^{-1}$ y $\mathbf{T}\boldsymbol{\beta}$) es:

$$\boldsymbol{\gamma} = \begin{pmatrix} \frac{1}{p_o^2}, o, o, \ldots o \\ o, \frac{1}{p_1^2}, o, \ldots o \\ o \quad o \quad \frac{1}{p_j^2}, \ldots o \\ \cdots\cdots \\ o \quad o \quad o \ldots \frac{1}{p_k^2} \end{pmatrix} \begin{pmatrix} P_o(1) \ldots P_o(t) \ldots P_o(n) \\ P_1(1) \ldots P_1(t) \ldots P_1(n) \\ \cdots\cdots \\ P_j(1) \ldots P_j(t) \ldots P_j(n) \\ \cdots\cdots \\ P_k(1) \ldots P_k(t) \ldots P_k(n) \end{pmatrix} \begin{pmatrix} \Sigma \beta_h \\ \Sigma 2^h \beta_h \\ \cdots\cdots \\ \Sigma t^h \beta_h \\ \cdots\cdots \\ \Sigma n^h \beta_h \end{pmatrix}$$

En el último vector, las sumatorias se extienden de $h = o$ a k. La primera matriz es la matriz diagonal de [2.13']:

$$(\mathbf{P'P})^{-1} = \begin{pmatrix} \frac{1}{p_o^2} & & \\ & \frac{1}{p_h^2} & \\ & & \frac{1}{p_k^2} \end{pmatrix} \Rightarrow \qquad [2.13']$$

$$\gamma = \begin{pmatrix} \frac{1}{p_o^2} \cdots\cdots & & & \\ \cdots \frac{1}{p_1^2} \cdots & & & \\ \cdots\cdots \frac{1}{p_j^2} \cdots & & & \\ \cdots\cdots\cdots & & & \\ \cdots\cdots\cdots \frac{1}{p_k^2} & & & \end{pmatrix} \begin{pmatrix} \sum_{t=1}^{n} P_o(t) \sum_{h=o}^{k} t^h \beta_h \\ \cdots\cdots\cdots \\ \sum_{t=1}^{n} P_j(t) \sum_{h=o}^{k} t^h \beta_h \\ \cdots\cdots\cdots \\ \sum_{t=1}^{n} P_k(t) \sum_{h=o}^{k} t^h \beta_h \end{pmatrix} = \begin{pmatrix} \frac{\Sigma P_o(t) \Sigma t^h \beta_h}{p_o^2} \\ \cdots\cdots\cdots \\ \frac{\Sigma P_j(t) \Sigma t^h \beta_h}{p_j^2} \\ \\ \frac{\Sigma P_k(t) \Sigma t^h \beta_h}{p_k^2} \end{pmatrix} \qquad [2.14]$$

La componente $j + 1$ del vector γ es:

$$\gamma_j = \frac{\sum_{t=1}^{n} P_j(t) \sum_{h=o}^{k} t^h \beta_h}{p_j^2} = \frac{\sum_{h=o}^{n} \beta_h \sum_{t=1}^{n} t^h P_j(t)}{p_j^2}$$

como obtuvimos en [2.9]

Y de acuerdo con las condiciones de ortogonalidad de la def. II y de la [1.19] tenemos:

$$\gamma_j = \frac{\sum_{h=j}^{k} \beta_h \sum_{t=1}^{n} t^h P_j(t)}{p_j^2} \qquad (h \geq j) \qquad [2.14']$$

que es, repetimos la [2.9].

Un caso particularmente interesante de la [2.9] es cuando $j = k$:

$$\gamma_k = \frac{\beta_k \sum_{t=1}^{n} t^k P_k(t)}{p_k^2} \qquad [2.15]$$

Observaciones:

1.ª La relación entre los coeficientes β_k y γ_k nos indica una importante consecuencia: si el modelo fuere de grado $k-1$, por ser $\beta_k = o \Rightarrow \gamma_k = o$.

2.ª Al hacer $j = k-1$ de la [2.9] deducimos:

$$\gamma_{k-1} = \frac{\beta_{k-1} \sum_{t=1}^{n} t^{k-1} P_{k-1}(t) + \beta_k \sum_{t=1}^{n} t^k P_k(t)}{p_{k-1}^2} \qquad [2.16]$$

que nos indica otra consecuencia: si la tendencia del modelo fuere del grado k, el coeficiente polinómico ortogonal de grado $k-1$, es decir, γ_{k-1} depende de los coeficientes de grado $k-1$ y del grado k, pero no de los de grado inferior.

3.ª Y la [2.9] nos amplía lo expuesto:

El coeficiente del modelo en forma poligonal γ_j de $\beta_j, \beta_{j+1}, \dots \beta_k$.

4. De la relación matricial [2.10] analicemos las componentes del vector β en función de las del vector γ [2.12].

Aplicamos las propiedades de ortogonalidad def. III de la Secc. 1.ª del Cap. I, y denominamos a los elementos de la matriz simétrica $(\mathbf{T}'\mathbf{T})^{-1} = \{ t_{jh} \}$ donde t_{jh} es el elemento de la fila $i-1$ y columna $h-1$. La matriz es de orden $(k+1) \times (k+1)$.

La fila $h+1$ de la matriz traspuesta de $\mathbf{T}$ es:

$$\{ 1, 2^h, \dots t^h \dots n^h \}$$

El vector columna $j+1$ de matriz P es:

$$P'_j = \{ P_j(1), \dots P_j(t) \dots P_j(n) \}$$

El elemento $a_{h+1,j+1}$ de la matriz producto

$$a_{h+1,j+1} = \sum_{t=1}^{n} t^h P_j(t) \qquad [2.17]$$

Y tiene las siguientes propiedades por el corolario 3 del teorema fundamental [1.19] del capítulo anterior:

$$\begin{array}{lll} a_{h+1,h+1} \neq o & h > 1 & \\ a_{h+1,j+1} = o & h < j & \\ a_{h+1,j+1} \neq o & h > j & j, h = o, 1, 2, \dots k. \end{array} \qquad [2.18]$$

La matriz $(k+1) \times (k+1)$ formada por los elementos [2.18] es una matriz triangular, y los elementos no nulos son la diagonal principal y los elementos superiores a ésta, y no es sino $\mathbf{T}'\mathbf{P}$ de la [2.12].

Si llamamos a la matriz simétrica $(\mathbf{T'}\,\mathbf{T})^{-1} = \{ t_{jh} \}$, el producto $(\mathbf{T'}\,\mathbf{T})^{-1}$ $\mathbf{T'}\,\mathbf{P}$ nos da la matriz cuadrada $k+1, k+1$;

$$(\mathbf{T'}\,\mathbf{T})^{-1}\,\mathbf{T'}\,\mathbf{P} = \{ q_{j+1,h+1} \} \tag{2.19}$$

$$q_{j+1,h+1} = \sum_{r=0}^{k} t_{j+1,r+1} \cdot a_{r+1,h+1} \tag{2.20}$$

Esta matriz $\mathbf{Q} = \{ q_{j+1,h+1} \}$ nos permite obtener la relación del elemento β_j en función de los de γ_h: de acuerdo con [2.12], [2.19] y γ:

$$\beta_j = \sum_{h=0}^{k} q_{j+1,h+1}\,\gamma_h \tag{2.21}$$

Una breve consideración al analizar la [2.21] y recordando [2.20].

Los coeficientes paramétricos β_j de la tendencia expresada en forma de una función entera de grado k, dependen de todos los coeficientes paramétricos de la tendencia cuando se expresa en forma de polinomios ortogonales.

La [2.15] nos indica mejor la realización cuando $j = k$;

$$\beta_k = \frac{p_k^2 \cdot \gamma_k}{\sum_{t=1}^{n} t^k P_k(t)} \tag{2.21'}$$

y también a partir de la [2.9] podemos sustituir los valores en la [2.21] para comprobar la identidad de los elementos y relaciones entre los valores de la matriz $\mathbf{Q}$ e interesantes propiedades, sobre todo cuando se refiere a las estimaciones de γ y β.

SECCIÓN 3.ª

Axiomática del modelo

El módelo de tendencia lo fundamentamos en este breve esquema de axiomas.

A-1. La tendencia puede expresarse por una parábola de grado k y puede escribirse de cualquiera de las formas que se indican en [2.8], *con un proceso estocástico puramente aleatorio.*

A-2. La tendencia es función del tiempo y de sus parámetros estructurales, siendo independiente de la perturbación aleatoria.

A-3. Las características de la perturbación aleatoria son las siguientes:

$$E\epsilon = \mathbf{O}$$

$$E\epsilon\,\epsilon' = \sigma^2\,\mathbf{I} = \Sigma \tag{2.22}$$

siendo $E\epsilon_i \epsilon_s = \sigma^2 \cdot \delta_{t,s}$; $\mathbf{I}$ indica la matriz unidad, y $\delta_{t,s}$, el símbolo de Kronecker.

La [2.22] representa una matriz diagonal, lo que permite indicar que la covarianza es homocedástica.

A-4. El vector aleatorio ϵ sigue una distribución nomal $N(\mathbf{O}, \sigma^2 \mathbf{I})$. [2.23]

A-5. El vector aleatorio $\mathbf{Y}$ es la suma de la tendencia más la perturbación aleatoria de ϵ.

En el capítulo próximo nos dedicamos a un estudio general de las distribuciones de probabilidad de ϵ y de $\mathbf{Y}$, que son fundamentales para conocer las distribuciones de las estimaciones de los parámetros poblacionales cuando las efectuemos.

CAPITULO III

Distribuciones de probabilidad

SECCIÓN 1.ª

Perturbación aleatoria

1. *Función de densidad conjunta del vector ϵ y función característica*

1. De las hipótesis efectuadas para ϵ_t tenemos que por ser normal $N(o, \sigma)$:

$$f(\epsilon_t) = \frac{e^{-\frac{\epsilon_t^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \qquad [3.1]$$

y la función característica es:

$$\varphi_t(r_t) = E e^{ir_t\epsilon_t} = e^{-\frac{r_t^2 \sigma^2}{2}} \qquad [3.2]$$

según es conocida en los cursos de Estadística, donde r_t es el parámetro que corresponde a ϵ_t y cuya notación usual es t, y aquí no la empleamos para evitar el confusionismo con el parámetro tiempo.

2. La función de densidad —n— dimensional $\epsilon = (\epsilon, \epsilon_2 \ldots \epsilon_n)$ en caso de homocedasticidad (A-3):

$$f(\epsilon) = \frac{e^{-\frac{\epsilon'\epsilon}{2\sigma^2}}}{\sigma^n (2\pi)^{\frac{n}{2}}} = \frac{e^{-\frac{\epsilon'\Sigma\epsilon^{-1}}{2}}}{/\,\Sigma\,/^{\frac{n}{2}} (2\pi)^{\frac{n}{2}}} \qquad [3.3]$$

donde la matriz de covarianzas [2.22] y axioma 4 es:

$$\Sigma = \sigma^2 \mathbf{I} \Rightarrow \Sigma^{-1} = \frac{1}{\sigma^2} \mathbf{I} \qquad [3.4]$$

porque los elementos covariantes de Σ son:

$$E\,\epsilon_t\,\epsilon_s = \sigma^2\,\delta_{t,s} \qquad [3.5]$$

Si la matriz de covarianzas es una matriz diagonal, la inversa también lo es, y la expresión [3.4] nos aclara que todos los elementos diagonales son $1/\sigma^2$.

La f. característica del proceso vector ϵ es:

$$\varphi\ (\tau) = E\ e^{i\tau'\epsilon} = E\ e^{i\Sigma r_t \epsilon_t} = \prod_{t=1}^{n} E\ e^{ir_t\epsilon_t} =$$

$$= e^{-\frac{\tau' \mathbf{I} \sigma^2 \tau}{2}} \qquad [3.6]$$

El exponente de la [3.6] está formado por el vector paramétrico arbitrario $\tau' = \{ r_1, r_2, \ldots r_n \}$; la matriz unidad $\mathbf{I}$ y la varianza σ^2, porque, según hemos indicado, equivale a la matriz de covarianzas [3.4].

2. *Funciones de densidad marginales de* ϵ

1. Si el vector ϵ lo descomponemos en dos vectores $\epsilon^{(1)}$ y $\epsilon^{(2)}$ $\epsilon' = (\epsilon^{(1)'}, \epsilon^{(2)})$. El primero de p elementos y el segundo de $n-p$, denominando el dominio de integración R_{n-p}:

$$f(\epsilon^{(1)}) = \int_{R_{n-p}} \frac{e^{-\frac{\epsilon'\Sigma^{-1}\epsilon}{2}}}{\left|\Sigma\right|^{\frac{1}{2}} (2\pi)^{\frac{n}{2}}}\, d\epsilon^{(2)} =$$

$$= \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \int_{R_{n-p}} e^{-\frac{(\epsilon^{(1)'}, \epsilon^{(2)'}) \frac{1}{\sigma^2} \binom{\epsilon^{(1)}}{\epsilon^{(2)}}}{2}}\, d\epsilon^{(2)} \qquad [3.7]$$

La integral es una integral múltiple de $n-p$ dimensiones y $d\epsilon^{(2)} = d\epsilon_{p+1} \cdot d\epsilon_{p+2} \ldots d\epsilon_n$.

Hagamos una participación de la matriz unidad:

$$\mathbf{I} = \begin{pmatrix} \mathbf{I}_p & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{n-p} \end{pmatrix} \Rightarrow [\epsilon^{(1)'}, \epsilon^{(2)'}] \begin{pmatrix} \mathbf{I}_p & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{n-p} \end{pmatrix} \begin{pmatrix} \epsilon^{(1)} \\ \epsilon^{(2)} \end{pmatrix} \qquad [3.8]$$

Las matrices $\mathbf{O}$ son de órdenes $p \times (n-p)$ y viceversa, según esté en la participación de la primera fila o la segunda, respectivamente.

$$\mathbf{I} = \epsilon^{(1)'} \mathbf{I}_p \epsilon^{(1)} + \epsilon^{(2)'} \mathbf{I}_{n-p} \epsilon^{(2)} \qquad [3.8']$$

Sustituida en [3.7] tenemos:

$$f(\epsilon^{(1)}) = \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \int_{R_{n-p}} e^{-\frac{\epsilon^{(1)'} \mathbf{I}_p \epsilon^{(1)}}{2\sigma^2}} \, e^{-\frac{\epsilon^{(2)'} \mathbf{I}_{n-p} \epsilon^{(2)}}{2\sigma^2}} \, d\epsilon^{(2)} =$$

$$= \frac{e^{-\frac{\epsilon^{(2)} \mathbf{I}_{n-p} \epsilon^{(2)}}{2\sigma^2}}}{\sigma^n (2\pi)^{\frac{n}{2}}} \int_{R_{n-p}} e^{-\frac{\epsilon^{(2)'} \mathbf{I}_{n-p} \epsilon^{(2)}}{2\sigma^2}} \, d\epsilon^{(2} =$$

$$= \frac{e^{-\frac{\epsilon^{(1)} \mathbf{I}_p \epsilon^{(1)}}{2\sigma^2}}}{\sigma^p (2\pi)^{\frac{p}{2}}} \qquad [3.9]$$

porque

$$\int_{R_{n-p}} e^{-\frac{\epsilon^{(2)'} \mathbf{I}_{n-p} \epsilon^{(2)}}{2\sigma^2}} \, d\epsilon^{(2)} = (\sigma \sqrt{2\pi})^{n-p} \qquad [3.10]$$

ya que el dominio de integración es en el espacio R_{n-p}, y la forma de ponerse cada integral es:

$$\int_{-\infty}^{\infty} e^{-\frac{\epsilon_i^2}{2\sigma^2}} \, d\epsilon_i = \sigma \sqrt{2\pi} \qquad [3.11]$$

por ser R_{n-p} un dominio de $n-p$ dimensiones.

2. Un caso particular es cuando a partir de [3.9] deseemos obtener la función de densidad marginal de ϵ_i. En este caso, el vector ϵ lo descompondríamos en dos (ϵ_i, $\epsilon^{(2)'}$), donde el primero sería unidimensional y el segundo, de n—1 dimensiones.

En tal caso tendríamos (de la fórmula [3.10]):

$$f\,(\epsilon^{(1)} = \epsilon_i) = \frac{e^{-\frac{\epsilon_i \epsilon_i}{2\sigma^2}}}{\sigma\sqrt{2\pi}} = \frac{e^{-\frac{\epsilon_i^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

porque la traspuesta de ϵ_i es el mismo elemento; la matriz unidad se reduce al elemento unidad y coincide con la [3.1].

3. *Función característica de la distribución marginal*

La sencillez de esta expresión si descomponemos el vector arbitrario paramétrico $r' = r^{(1)'} + \mathbf{O}'$, siendo $\mathbf{O}$ un vector columna de n—p dimensiones, y la matriz unidad [3.8] sustituyendo en [3.6] tenemos:

$$\varphi(\tau^{(1)}) = e^{-\frac{(\tau^{(1)'}\ \mathrm{O}')\begin{pmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{O} & \mathrm{I} \end{pmatrix}\begin{pmatrix} \tau \\ \mathrm{O} \end{pmatrix}\sigma^2}{2}} = e^{-\frac{\tau'\sigma \mathrm{I}\, \tau}{2}} \qquad [3.12]$$

no es sino la función característica de una variable normal de p dimensiones. Si el vector aleatorio de —p— dimensiones es $\epsilon^{(1)}$, su función de densidad marginal es la [3.9], y por la independencia (obsérvese la matriz de covarianzas), es:

$$f\,(\epsilon^{(1)}) = \frac{e^{-\frac{\Sigma \epsilon^2}{2\sigma^2}}}{\sigma^p (2\pi)^{\frac{p}{2}}} = \frac{e^{-\frac{\epsilon'\epsilon}{2\sigma^2}}}{\sigma^p (2\pi)^{\frac{p}{2}}} \qquad [3.13]$$

La sumatoria de la expresión [3.13] indica que puede sumarse del conjunto $\{\epsilon_1, \epsilon_2, \ldots \epsilon_n\}$ p variables cualesquiera no necesariamente consecutivas.

4. *Funciones condicionadas de* $\epsilon^{(1)}$

De las [3.3] y [3.9] deducimos funciones de densidad condicionadas de las $\epsilon^{(1)} / \epsilon^{(2)}$.

$$f(\epsilon^{(1)} / \epsilon^{(2)}) = \frac{e^{\frac{\epsilon'\epsilon}{2\sigma^2}}}{\sigma^n (2\pi)^{\frac{n}{2}}} : \frac{e^{\frac{\epsilon^{(1)'}\epsilon^{(1)}}{2\sigma^2}}}{\sigma^{n-p}(2\pi)^{\frac{n-p}{2}}} =$$

$$= \frac{e^{\frac{\epsilon^{(1)'}\epsilon^{(1)}}{2\sigma}}}{\sigma^p (2\pi)^{\frac{p}{2}}} = f(\epsilon^{(1)}) \qquad [3.14]$$

por ser $\epsilon' \mathbf{I}_n \epsilon = \epsilon' \epsilon = \epsilon^{(1)'} \mathbf{I}_p \epsilon^{(1)} + \epsilon^{(2)'} \mathbf{I}_{n-p}{}^{(2)} \epsilon^{(2)}$ [3.15]

La [3.14] nos indica que $\epsilon^{(1)} / \epsilon^{(2)}$ no depende para nada del vector aleatorio de componentes $\epsilon^{(2)}$ y, en consecuencia, el grupo de variables $\epsilon^{(1)}$ y $\epsilon^{(2)}$ son independientes.

Huelga decir que la función característica es precisamente la [3.12].

El análisis de la [3.14] nos indica que, dado un conjunto aleatorio de la perturbación aleatoria $\epsilon^{(2)}$, éstas *no influyen en la perturbación de* $\epsilon^{(1)}$: es decir, son totalmente independientes.

SECCIÓN 2.ª

Modelo estocástico

1. *Función de densidad conjunta y función característica*

Dado un sistema de polinomios $P(t) = \{ P_o(t), P_1(t) \dots P_k(t) \}$ [3.16]

definido en el conjunto $E = \{ 1, 2, \dots n \}$, $t \in E$, el vector aleatorio ϵ, el modelo estocástico $\{ y_t, t \in E \}$ o el vector $\mathbf{Y}$ con un conjunto paramétrico $\gamma' = \{ \gamma_o, \gamma_1, \dots \gamma_k \}$ [3.17] puede expresarse según [2.16]:

$$\mathbf{Y} = \mathbf{P}\gamma + \epsilon \qquad [3.18]$$

La matriz $\mathbf{P}'$ es la traspuesta de $\mathbf{P}$ y la línea t es la expresada en [3.16], donde $t \in E$. El vector columna aleatorio $\mathbf{Y}$ está formado por sus componentes estocásticos $\{ Y_t, t \in E \}$.

$\mathbf{P}\gamma$ no es aleatorio, según expusimos en el modelo.

Tomando esperanzas matemáticas tenemos:

$$E\mathbf{Y} = \mathbf{P}\gamma \qquad [3.19]$$

según el axioma A-3.

El jacobiano de la transformación [3.18] es la unidad $\Rightarrow$

$$\mathrm{f}(\mathbf{Y}) = \mathrm{f}(\epsilon) = \mathrm{f}(\mathbf{Y} - \mathbf{P}\gamma) =$$

$$+ \frac{e^{-\frac{(\mathbf{Y} - \mathbf{P}\gamma)' \mathbf{I} (\mathbf{Y} - \mathbf{P}\gamma)}{2\sigma^2}}}{\sigma^n (2\pi)^{\frac{n}{2}}} = \frac{e^{-\frac{\sum_{t=1}^{n} [Y_t - \mathbf{P}(t)\gamma]^2}{2\sigma^2}}}{\sigma^n (2\pi)^{\frac{n}{2}}} \qquad [3.20]$$

La componente t del vector $\mathbf{Y}$ es:

$$\mathbf{Y}_t = \mathbf{P}(t)\gamma + \epsilon_t = \sum_{h=0}^{k} P_h(t)\gamma_h + \epsilon_t \qquad [3.21]$$

donde $\mathbf{P}(t)$ es la fila t de la matriz $\mathbf{P}$.

2. *Función característica de Y*

La función característica del modelo estocástico del vector $\mathbf{Y}$ es:

$$\varphi_{\mathbf{Y}}(\tau) = E\, e^{i\tau'\mathbf{Y}} = E\, e^{i\tau'(P\gamma + \epsilon)} =$$

$$= e^{i\tau' P\gamma}\, E\, e^{i\tau'\epsilon} = e^{i\tau' P\gamma - \frac{\tau' \mathbf{I} \sigma^2 \tau}{2}} \qquad [3.22]$$

recordando la función característica [3.16].

Esta función característica del vector —n— dimensional es normal con los siguientes parámetros:

Medias $\mathbf{P}\gamma$ [3.23]

Matriz de covarianzas $\sigma^2 \mathbf{I}$

Por ser esta matriz diagonal, *los elementos componentes aleatorios del vector* **Y** *son independientes.*

3. *Consecuencias*

1. Función de densidad de Y_t:

$$\mathrm{f}(Y_t) = \frac{e^{-\frac{(Y_t - \mathbf{P}(t)\gamma)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} = \frac{e^{-\frac{(Y_t - \sum P_h(t)\gamma_h)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \qquad [3.24]$$

recordando [3.6].

2. Función característica de Y_t:

$$\varphi_t(r_t) = e^{i r_t \mathbf{P}(t)\gamma - \frac{r_t' \sigma^2 r_t}{2}} =$$

$$= e^{i r_t \sum_{h=0}^{k} P_h(t)\gamma_h - \frac{\sigma^2 r_t^2}{2}} \qquad [3.25]$$

3. La distribución de Y_t es normal, de parámetros:

Media	$\sum_{h=o}^{k} P_h(t)\gamma_h$	(tendencia)	[3.26]
Varianza	σ^2		[3.27]

4. La función de densidad marginal de una partición cualquiera de Y recordando [3.9], *y si* $Y^{(1)}$ tiene ρ componentes:

$$f(\mathbf{Y}^{(1)}) = \frac{e^{-\sum_{t(p)} \frac{[Y_t - \sum_{h=0}^{k} P_h(t)\gamma_h]^2}{2\sigma^2}}}{\sigma^p (2\pi)^{\frac{p}{2}}} \qquad [3.28]$$

que coincide con la condicionada y son $\mathbf{Y}^{(1)}$ e $\mathbf{Y}^{(2)}$ independientes donde la $\Sigma_{t(p}$ se extiende a un subconjunto $t \in E$ de variables componentes del vector $\mathbf{Y}$.

5. De forma análoga a como expusimos en [3.22] la f. característica de $\mathbf{Y}^{(1)}$ recordando [3.6] y [3.8]:

$$\varphi_{\mathbf{Y}^{(1)}}(\gamma^{(1)}) = e^{i\tau^{(1)'}\mathbf{P}^{(1)}\gamma - \frac{\tau^{(1)'} I_p \sigma^2 \tau^{(1)}}{2}} \qquad [3.29]$$

donde $\mathbf{P}^{(1)}$ es una partición de la matriz $\mathbf{P}$ asociada a las variables $Y^{(1)}$ según los valores del subconjunto $t \in E$ para los que hemos elegido el vector $\epsilon^{(1)}$ o su equivalente del modelo $Y^{(1)}$.

CAPITULO IV

Estimación por punto

SECCIÓN 1.ª

Estimación de los parámetros del modelo

1. *Coeficientes regresores*

Del modelo [2.7] deducimos

$$\epsilon'\epsilon = (\mathbf{Y} - \mathbf{P}\gamma)'(\mathbf{Y} - \mathbf{P}\gamma) = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{P}\gamma - \gamma'\mathbf{P}'\mathbf{Y} + \gamma'\mathbf{P}'\mathbf{P}\gamma =$$
$$= \mathbf{Y}'\mathbf{Y} - 2\,\gamma'\mathbf{P}'\mathbf{Y} + \gamma'\mathbf{P}'\mathbf{P}\gamma \qquad [4.1]$$

Derivemos vectorialmente la [4.1] respecto a γ. Por tener $k+1$ componentes, la derivada nos da el siguiente vector expresado matricialmente:

$$\frac{\partial\,\epsilon'\epsilon}{\partial\,\boldsymbol{\gamma}} = -\,2\,\mathbf{P}'\mathbf{Y} + 2\,\mathbf{P}'\mathbf{P}\gamma \qquad [4.2]$$

La matriz $\mathbf{P}'\mathbf{P}$ es simétrica y diagonal; el vector $\mathbf{Y}$ es de dimensionalidad n y $\boldsymbol{\gamma}$ es de $k+1$ dimensiones. Variando los parámetros $\boldsymbol{\gamma}$ varía el vector [4.2] y si en todo el campo de variabilidad de $\boldsymbol{\gamma}$ existe un vector $\mathbf{c}$ que, sustituido en lugar de γ en [4.2] nos da el vector nulo, a referido vector $\mathbf{c}$ denominamos estimación por punto de $\boldsymbol{\gamma}$:

$$\mathbf{O} = -\,2\,\mathbf{P}'\mathbf{Y} + 2\,\mathbf{P}'\mathbf{P}\mathbf{c} =$$
$$\mathbf{P}'\mathbf{P}\mathbf{c} = \mathbf{P}'\mathbf{Y} =$$
$$\mathbf{c} = (\mathbf{P}'\mathbf{P})^{-1}\,\mathbf{P}'\mathbf{Y} \qquad [4.3]$$

solución matricial de la estimación de los parámetros del modelo.

2. *Forma de los coeficientes regresores*

1. Examinemos la inversa de $\mathbf{P}'\mathbf{P}$ y después el vector $\mathbf{P}'\mathbf{Y}$ para, finalmente, determinar la componente $h+1$ de $\mathbf{c}$.

Por ortogonalidad tenemos:

$$\mathbf{P}'\mathbf{P} = \{\sum_{t=1}^{n} P_h(t)\}\,P_j(t)\,] \;=\; \{p_h^{2}\} \qquad h = o, 1, \ldots k. \;\; j = h \qquad [4.4]$$

matriz diagonal cuyos elementos son [1.18] y los no diagonales de la matriz son nulos según la Def. III del Cap. I.

Pero la matriz inversa de la diagonal es otra matriz diagonal. Luego

$$(\mathbf{P}'\ \mathbf{P})^{-1} = \left\{ \frac{1}{p_h^2} \right\} \qquad h = o, 1, \dots k \qquad [4.5]$$

según expusimos en [2.13]. (I).

2. El producto $\mathbf{P}'\gamma_t$ que aparece en [4.3] es inmediato:

$$\mathbf{P}'\mathbf{Y} = \begin{bmatrix} \Sigma\, P_o(t)\, Y_t \\ \dots\dots \\ \Sigma\, P_h(t)\, Y_t \\ \dots\dots \end{bmatrix} \qquad [4.6]$$

[4.6] Prescindimos de las soluciones posibles singulares si $|\mathbf{P}'\mathbf{P}| = o$ es un vector de $k+1$ elementos.

3. Según las [4.3], [4.5] y [4.6] la componente $h+1$ de $\mathbf{c}$ o sea, el estimador de γ_h es:

$$c_h = \frac{\sum\limits_{t=1}^{n} P_h(t)\, Y_t}{p_h^2} = \frac{\sum\limits_{t=1}^{n} P_h(t)\, Y_t}{\sum\limits_{t=1}^{n} P_h(t)^2} = \qquad [4.7]$$

También si sustituimos Y_t por [3.21], el estimador c_h puede escribirse:

$$c_h = \gamma_h + \sum_{t=1}^{n} \frac{P_h(t)\,\epsilon_t}{p_h^2} = \gamma_h + \frac{\epsilon' P_h}{p_h^2} \qquad [4.7']$$

recordando las [1.9] y [1.22'].

4. La [4.7] si los datos de la serie temporal son dados en una *estimación por punto del parámetro γ_h. Proporciona un solo valor que aún siendo el que más garantías ofrece, su probabilidad es prácticamente nula.*

Si por el contrario: $\{ Y_t, t \in E \}$ es un proceso Y_t es una variable aleatoria, c_h[4.7] o [4.7'] también es una variable aleatoria —n— dimensional. Y las expresiones precedentes *son estimadores o funciones aleatorias, no estimaciones.*

3. *Estimadores centrados*

El vector [4.3] si el vector Y es un proceso (no una realización) es una variable aleatoria o estimador que tiene una esperanza matemática y coincide con los parámetros poblacionales.

Pero la matriz inversa de la diagonal es otra matriz diagonal. Luego

$$(\mathbf{P}'\ \mathbf{P})^{-1} = \left\{ \frac{1}{p_h^2} \right\} \qquad h = o, 1, \dots k \qquad [4.5]$$

según expusimos en [2.13]. (I).

2. El producto $\mathbf{P}'\gamma_t$ que aparece en [4.3] es inmediato:

$$\mathbf{P}'\mathbf{Y} = \begin{bmatrix} \Sigma\, P_o(t)\, Y_t \\ \dots\dots \\ \Sigma\, P_h(t)\, Y_t \\ \dots\dots \end{bmatrix} \qquad [4.6]$$

[4.6] Prescindimos de las soluciones posibles singulares si $|\mathbf{P}'\mathbf{P}| = o$ es un vector de $k+1$ elementos.

3. Según las [4.3], [4.5] y [4.6] la componente $h+1$ de $\mathbf{c}$ o sea, el estimador de γ_h es:

$$c_h = \frac{\sum\limits_{t=1}^{n} P_h(t)\, Y_t}{p_h^2} = \frac{\sum\limits_{t=1}^{n} P_h(t)\, Y_t}{\sum\limits_{t=1}^{n} P_h(t)^2} = \qquad [4.7]$$

También si sustituimos Y_t por [3.21], el estimador c_h puede escribirse:

$$c_h = \gamma_h + \sum_{t=1}^{n} \frac{P_h(t)\,\epsilon_t}{p_h^2} = \gamma_h + \frac{\epsilon'\, P_h}{p_h^2} \qquad [4.7']$$

recordando las [1.9] y [1.22'].

4. La [4.7] si los datos de la serie temporal son dados en una *estimación por punto del parámetro γ_h. Proporciona un solo valor que aún siendo el que más garantías ofrece, su probabilidad es prácticamente nula.*

Si por el contrario: $\{ Y_t, t \in E \}$ es un proceso Y_t es una variable aleatoria, c_h[4.7] o [4.7'] también es una variable aleatoria —n— dimensional. Y las expresiones precedentes *son estimadores o funciones aleatorias, no estimaciones.*

3. *Estimadores centrados*

El vector [4.3] si el vector Y es un proceso (no una realización) es una variable aleatoria o estimador que tiene una esperanza matemática y coincide con los parámetros poblacionales.

En efecto: si en la [4.3] sustituimos Y por el modelo [3.18] tenemos

$$\mathbf{C} = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\,(\mathbf{P}\gamma + \epsilon) \quad = \qquad [4.8]$$
$$\mathbf{C} = \gamma + (\mathbf{P}'\,\mathbf{P})^{-1}\,\mathbf{P}'\epsilon$$

El único elemento aleatorio es el vector ϵ y la esperanza de [4.8] es:

$$E\ \mathbf{C} = \gamma \qquad [4.9]$$

por cuanto $E\epsilon = \mathbf{O}$ según el axioma A-3.

Luego las componentes c_h:

$$E\,c_h = \gamma_h \qquad [4.10]$$

que también de [4.7] y [4.7′] se obtiene el mismo resultado.

4. *Matriz de covarianzas de* **C**

De [4.9] deducimos:

$$\mathrm{Cov}\,(\mathbf{C}) = E(c - \gamma)(c - \gamma)' = E(\mathbf{P}'\mathbf{P})^{-1}\,\mathbf{P}'\epsilon\epsilon'\,\mathbf{P}\,(\mathbf{P}'\mathbf{P})^{-1} =$$
$$= (\mathbf{P}'\mathbf{P})^{-1}\,\mathbf{P}'(E\epsilon\epsilon')\,\mathbf{P}\,(\mathbf{P}'\,\mathbf{P})^{-1} = \sigma^2(\mathbf{P}'\mathbf{P})^{-1}\,\mathbf{P}'\mathbf{P}\,(\mathbf{P}'\mathbf{P})^{-1} = \sigma^2(\mathbf{P}'\,\mathbf{P})^{-1} =$$
$$\mathrm{Cov}\,(\mathbf{C}) = \sigma^2 \left\{\frac{1}{p_h^{\,2}}\right\} \qquad h = o, 1, \dots k \qquad [4.11]$$

de acuerdo con [4.5], porque $(\mathbf{P}'\mathbf{P})^{-1}$ es una matriz diagonal.

De [4.11] se deduce:

$$E(c_h - \gamma_h)(c_j - \gamma_j) = o \qquad \text{para } h \neq j \qquad [4.12]$$
$$E\,(c_h - \gamma_h)^2 = \mathrm{Var}\,(c_h) = \sigma_h^{\,2} = \frac{\sigma^2}{p_h^{\,2}} \qquad h = j \qquad [4.13]$$

La [4.12] nos indica la incorrelación de estimadores c_h y c_j y la [4.13] la varianza del estimador regresor γ_h.

SECCIÓN 2.ª

Estimación del modelo y de su varianza

1. *Estimación del modelo*

Postmultiplicando la matriz **P** por el vector [4.3] tenemos los valores ajustados:

$$\hat{\mathbf{Y}} = \mathbf{P}\,\mathbf{C} \qquad [4.14]$$

y la componente $\hat{Y}_t$ de [4.14] no es sino:

$$\hat{Y}_t = \sum_{h=o}^{k} c_h P_h(t) \qquad t = 1, 2, \dots n. \qquad [4.15]$$

donde c_h es el valor obtenido en la [4.7].

2. *Desviaciones*

Las desviaciones de los valores del modelo [3.18] y los ajustados [4.14] son:

$$\mathbf{D} = Y - \mathbf{P}\,\mathbf{C} \qquad [4.16]$$

sustituyendo en [4.16], Y por [3.18] y $\mathbf{C}$ por [4.8] el vector aleatorio de desviaciones [4.16] puede escribirse:

$$\mathbf{D} = \mathbf{P}\gamma + \epsilon - \mathbf{P}[\gamma + (\mathbf{P}'\,\mathbf{P})^{-1}\mathbf{P}'\epsilon] =$$

$$\mathbf{D} = [\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}']\bullet \qquad [4.16']$$

El vector fila de [4.16] es:

$$\mathbf{D}' = \epsilon'(\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}') \qquad [4.16'']$$

al ser simétrica la matriz de [4.16′] que premultiplica ϵ.

A esta matriz la llamaremos M y estudiamos algunas propiedades en el siguiente apartado.

3. *Propiedades de la matriz M*

Según hemos indicado:

$$M = \mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' \qquad [4.17]$$

Tenemos las siguientes propiedades:

1.ª La matriz **M** es simétrica.

Se comprueba inmediatamente.

2.ª La matriz **M** es idempotente.

En efecto:

$$\mathbf{M}^2 = (\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}')(\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}') =$$

$$= \mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' - \mathbf{P}(\mathbf{P}'\mathbf{P})'^{-1}\mathbf{P}' + \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' = +\mathbf{M}$$

3.ª Los números característicos de la matriz **M** son unos y ceros.

4.ª La traza de esta matriz es la suma de elementos de la diagonal:

$$\text{Traza } \mathbf{M} = \text{Traza } \mathbf{I} - \text{Traza } \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' = n\text{-traza } (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\mathbf{P} =$$

$$= n - (k+1) \qquad [4.18]$$

4. *Estimación de los cuadrados de las desviaciones*

De acuerdo con las [4.16′] y [4.16″] deducimos que los cuadrados de las desviaciones pueden escribirse:

$$\mathbf{D}'\mathbf{D} = \epsilon'[\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'][\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}']\epsilon$$
$$= \epsilon'[\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}']\epsilon = \sum_{t=1}^{n}[Y_t - \sum_{h=o}^{k} c_h P_h(t)]^2 \qquad [4.19]$$

La primera expresión nos servirá para estimar $\mathbf{D}'\mathbf{D}$ y la segunda, como simple operatoria.

5. *Esperanza de los cuadrados y estimador centrado de la varianza*

Tomando esperanzas de la [4.19] tenemos:

$$E\,\mathbf{D}'\mathbf{D} = \mathbf{E}\,\epsilon'[\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}']\epsilon \qquad [4.20]$$

La matriz que no son ceros son los elementos no diagonales, porque $E\,\epsilon_s\,\epsilon_t = \sigma^2\,\delta_{s,t}$, siendo $\delta_{s,k}$ el símbolo de Kronecker y de acuerdo con Axioma 3.

La esperanza de la [4.20] recordando la propiedad de la matriz M [4.18] es:

$$E\,\mathbf{D}'\mathbf{D} = \sigma^2\,\mathrm{traz}\,\mathbf{M} = \sigma^2[n - (k+1)] \qquad [4.21]$$

La varianza poblacional del modelo es, pues, también:

$$E\frac{\mathbf{D}'\mathbf{D}}{n-(k+1)} = \sigma^2 \Rightarrow \qquad [4.22]$$

$$\hat{\sigma}^2 = \frac{\mathbf{D}'\mathbf{D}}{n-(k+1)} = \frac{\sum_{t=1}^{n}[Y_t - \sum_{t=1}^{n} c_h P_h(t)]^2}{n-(k+1)} \qquad [4.23]$$

y este estimador es centrado, porque por la [4.22] comprobamos esta propiedad.

6. *Cuadrados de las desviaciones D′ D en función de los coeficientes regresores*

Según [4.16] los cuadrados de las desviaciones [4.19] pueden también escribirse:

$$\mathbf{D}'\mathbf{D} = (Y - \mathbf{PC})'(Y - \mathbf{PC}) = Y'Y - \mathbf{C}'\mathbf{P}'Y \qquad [4.24]$$

Al sustituir en [4.24] el vector fila el valor estimado C [4.3] tenemos

$$\mathbf{D}'\mathbf{D} = Y'Y - Y'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'Y \qquad [4.25]$$

Examinemos el valor del término sustractivo:

Por una parte $P'\,Y$ es un vector de $k+1$ componentes, y el elemento $h+1$ es $\sum_{t=1}^{n} P_h(t)\,Y_t$. Su traspuesta es el vector fila con los mismos elementos. La matriz $(P'\,P)^{-1}$ es la matriz diagonal [4.5].

Por otra parte, de la [4.7] deducimos:

$c_h\,p_h^2 = \Sigma\,P_h(t)\,Y_t$ y sustituido en $P'\,Y$ y en su traspuesto y la matriz [4.5] el sustraendo de [4.24] es:

$$Y'\,\mathbf{P}\,(\mathbf{P}'\,\mathbf{P})^{-1}\,\mathbf{P}'\,Y = [c_o p_o^2 \,....\, c_h p_h^2 \,...\, c_k p_k^2] \begin{pmatrix} \frac{1}{p_o^2} & & & \\ & \ddots & & \\ & & \frac{1}{p_h^2} & \\ & & & \ddots \; \frac{1}{p_k^2} \end{pmatrix} \begin{bmatrix} c_o p_o^2 \\ \vdots \\ c_h p_h^2 \\ \cdots \\ c_k p_k^2 \end{bmatrix} =$$

$$= c_o^2\,p_o^2 + c_1^2\,p_1^2 + \ldots + c_k^2\,p_k^2. \qquad [4.25']$$

La [4.25′] sustituida en [4.24] nos da los cuadrados de las desviaciones en función de los estimadores de los coeficientes regresores:

$$\mathbf{D}'\,\mathbf{D} = Y'\,Y - (c_o^2\,p_o^2 + \ldots + c_h^2\,p_h^2 + \ldots + c_k^2\,p_k^2) \qquad [4.26]$$

A veces, la [4.26] se representa de otra forma. Si sabemos $P_o(t) = 1$ (para $\forall t \in E$) def. III Cap. I.

$$p_o^2 = \sum_{t=1}^{n} P_o(t)^2 = n \quad \text{y de la [4.7] para } h = o \Rightarrow C_o = \frac{\Sigma\,Y_t}{n} = \bar{y}$$

Luego, la [4.26] puede escribirse:

$$\mathbf{D}'\,\mathbf{D} = (n - [k+1])\,\hat{\sigma}^2 = \sum_{t=1}^{n} Y_t^2 - n\bar{y}^2 - \sum_{h=1}^{n} c_h^2\,p_h^2 = ns_{ry}^2 \qquad [4.27]$$

según se utilice el estimador $\hat{\sigma}^2$ insesgado o el estimador sesgado s_{ry}^2 de la varianza σ^2.

SECCIÓN 3.ª

Estimadores Lineales Optimos regresores

1. *Fundamentos de este tipo de estimadores*

En la Sección 1.ª demostramos que todo estimador [4.7] c_h del parámetro γ_h es combinación lineal de los datos observados. Así:

$$\frac{P_h(t)}{\Sigma\,P_h(t)^2}$$

es el coeficiente del valor Y_t. Y también probamos que estos estimadores son centrados. (Secc. 1.ª 3.)

En la Sección 1.ª 4 demostramos la matriz de covarianzas [4.11], y la varianza para un estimador concreto dado [4.13].

El problema es que los estimadores ELIO tienen que cumplir un requisito importante: además, su varianza sea mínima.

Se trata de encontrar (no la varianza mínima en el sentido de Cramer-Rao), sino de si otro estimador que fuere lineal tiene menor varianza que la del estimador [4.7] que es la [4.13].

Intentemos formar un estimador de γ_h que sea *combinación lineal de los valores* Y_t y tenga las propiedades siguientes:

1.ª Sea insesgado.
2.ª Tenga varianza mínima.

2. *Planteamiento del problema*

Llamemos $\hat{c}_p$ a este estimador lineal con coeficientes c_{ht} desconocidos.

$$\hat{c}_p = \sum_{t=1}^{n} c_{ht} Y_t \qquad [4.34]$$

¿Encontraremos otro estimador de γ_p lineal [4.34] distinto del [4.7] con las propiedades de ser centrado y tener varianza inferior a [4.13]?

Este problema es el que resolvemos ahora.

En el estimador lineal [4.34] sustituiremos Y_t por su expresión equivalente [2.6′]:

$$c_h = \sum_{t=1}^{n} c_{ht} \left(\sum_{j=o}^{k} P_j(t)\gamma_j + \epsilon_t \right) =$$

$$= \sum_{j=o}^{k} \gamma_j \sum_{t=1}^{n} c_{ht} P_j(t) + \sum_{t=1}^{n} c_{ht} \epsilon_t \qquad [4.35]$$

La [4.35] es una variable aleatoria y para que sea centrado:

$$E\,\hat{c}_h = \gamma_h$$

debe cumplir las siguientes condiciones:

$$\sum_{t=1}^{n} c_{ht} P_j(t) = o \qquad \forall j \neq h \qquad [4.36]$$

$$\sum_{t=1}^{n} c_{ht} P_h(t) = 1 \qquad [4.37]$$

puesto que la esperanza del tercer sumando es nula según el Axioma 3 de la Sección 3.ª del Capítulo II.

Las [4.36] y [4.37] nos permite reescribir el estimador [4.34] o su equivalente [4.35] en forma más reducida:

$$\hat{c}_h = \gamma_h + \sum_{t=1}^{n} c_{ht} \epsilon_t \qquad [4.38]$$

La varianza de este estimador lineal insesgado es:

$$\sigma^2 \hat{c}_h = E(\hat{c}_h - \gamma_h)^2 = \sigma^2 \sum_{t=1}^{n} c_{ht}^2 \qquad [4.39]$$

Esta varianza depende de $c_{ht} (t = 1, 2, \ldots n)$ y hemos de encontrar estos valores que minimicen la [4.39] y al propio tiempo cumplan las c_{ht} las condiciones impuestas en [4.36] y [4.37]. Estas condiciones son en total $k + 1$, ya que en la primera $j = o$, 1, 2 ... k (excepto $j = h$) son k condiciones.

Para resolver el mínimo de [4.39] prescindiremos del término constante σ^2 y aplicaremos el método de multiplicadores de Lagrange.

A las ecuaciones [4.36] multiplicaremos por $-2\,\lambda_j, j \neq h$ y a la [4.37] por $-2\,\lambda_h$: habiendo previamente pasado al primer término, y formaremos la función:

$$\Phi = \sum_{t=1}^{n} c_{ht}^2 - 2 \sum_{j \neq h} \lambda_j \sum_{t=1}^{n} c_{ht} P_j(t) - 2\,\lambda_h \Big(\sum_{t=1}^{n} c_{ht} P_h(t) - 1 \Big) \qquad [4.40]$$

El primer sumando es el término variable de [4.39]; el segundo sumando son las k condiciones de [4.36] y, finalmente, el último sumando proviene de multiplicar la [4.37] por $-2\,\lambda_h$.

La [4.40], el único término no nulo es el primero y multiplicado por la constante σ^2 tenemos [4.39]. El problema está planteado y los valores c_{ht} $(t = 1, 2, \ldots n)$ cumplirán las condiciones [4.36] y [4.37] si determinamos el mínimo de [4.40].

2. *Solución*

Derivemos la [4.40] con respecto a una de las variables c_{ht} e igualemos a cero por ser condición necesaria para la existencia de mínimo:

$$\frac{\partial \Phi}{\partial c_{ht}} = 2\,c_{ht} - 2 \sum_{j \neq h} \lambda_j P_j(t) - 2\lambda_h P_h(t) = o \Rightarrow$$

$$c_{ht} = \sum_{j \neq h} \lambda_j P_j(t) + \lambda_h P_h(t) \qquad [4.41]$$

Si la [4.41] la sustituimos en [4.37] tenemos:

$$1 = \sum_{t=1}^{n} c_{ht} P_h(t) = \sum_{t=1}^{n} \left[\sum_{j \neq h} \lambda_j P_j(t) + \lambda_h P_h(t) \right] P_h(t) =$$

$$= \lambda_h \sum_{t=1}^{n} P_h(t)^2 = \lambda_h p_h^2$$

Luego:

$$\lambda_h = \frac{1}{p_h^2} \qquad h = o,\ 1, 2 \ldots k \tag{4.42}$$

recordando [1.18].

Igualmente, si la [4.41] la sustituimos en [4.36]:

$$\sum_{t=1}^{n} c_{ht} P_j(t) = o = \sum_{t=1}^{n} \left[\sum_{j \neq h} \lambda_j P_j(t) + \lambda_h P_h(t) \right] P_j(t) =$$

$$= \lambda_j \sum_{t=1}^{n} P_j(t)^2 = \lambda_j p_j^2 = o \Rightarrow$$

$$\lambda_j = o \qquad \text{porque } p_j^2 \neq o \qquad j \neq h. \tag{4.43}$$

$$j = o,\ 1, \ldots k.$$

Los multiplicadores [4.43] y [4.42] nos permiten determinar las variables c_{ht} [4.41] que minimizan la varianza:

$$c_{ht} = \frac{P_h(t)}{p_h^2} \tag{4.44}$$

Los valores [4.44] llevados a la [4.34] precisamente nos dan el estimador [4.7] e igualmente en la expresión [4.39] nos da la misma varianza de la fórmula [4.13].

Luego el estimador ELIO $\hat{c}_h$ es el mismo que el c_h determinado en [4.7].

1. *Notas importantes*

1.ª La estimación de los parámetros deducida en la Sección 1.ª se basa en la teoría mínimo-cuadrática, y es válida, aunque suprimamos el axioma 4 de la distribución normal de la perturbación aleatoria (Sec. 3.ª Cap. II).

2.ª Admitido el axioma 4, la función de verosimilitud es:

$$L = \frac{1}{(\sigma 2 \pi)^{\frac{n}{2}}} \; e^{-\frac{(Y - P\gamma)'(Y - P\gamma)}{2\sigma}} \tag{4.45}$$

y el máximo de esta función es equivalente al mínimo del exponente coincidiendo con el método de mínimos cuadrados.

CAPITULO V

Teoría de la estimación por intervalo

SECCIÓN 1.ª

Distribuciones de los estimadores

1. *Función de densidad conjunta de* **C**

De la [4.8] deducimos la función característica del vector regresor **C**:

$$E\, e^{i\tau'\mathbf{C}} = E\, e^{i\tau'(\gamma + [\mathbf{P}'\mathbf{P}]^{-1}\mathbf{P}'\epsilon)} =$$

$$= e^{i\tau'\gamma}\, E\, e^{i\tau'(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\epsilon} \qquad [5.1]$$

Si hacemos la transformación del parámetro τ

$$\mathbf{S} = \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\tau \quad \Rightarrow$$

$$\mathbf{S}' = \tau'(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'$$

Si $\mathbf{S}'$ lo sustituimos en [5.1] la esperanza matemática es la función característica de ϵ, pero de parámetro **S** y según [3.6] tenemos:

$$\varphi_c(\tau) = e^{i\tau'\gamma}\, E\, e^{i\mathbf{S}'\epsilon} = e^{i\tau'\gamma} \quad \frac{\mathbf{S}'\mathbf{I}\sigma^2\mathbf{S}}{2} \qquad [5.2]$$

pero

$$\mathbf{S}'\mathbf{I}\sigma^2\mathbf{S} = \sigma^2\mathbf{S}'\mathbf{S} = \sigma^2\tau'(\mathbf{P}'\mathbf{P}')^{-1}\mathbf{P}'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\tau =$$

$$= \frac{\tau'(\mathbf{P}'\mathbf{P})^{-1}\sigma^2\mathbf{I}\,\tau}{2}$$

La función característica del vector variable C de —$k+1$— dimensiones es:

$$\varphi_{\mathbf{C}}(\tau) = e^{i\tau'\gamma} \cdot \frac{\tau'(\mathbf{P}'\mathbf{P})^{-1}\sigma^2\tau}{2} \qquad [5.3]$$

Esta función característica es de una variable normal de los siguientes parámetros:

Medias: γ [5.3a]

Matriz de covarianzas: $(\mathbf{P}'\mathbf{P})^{-1}\sigma^2$ [5.3b]

2. *Función característica y función de densidad del coeficiente regresor c_h*

El vector fila paramétrico r' de $-k+1-$ dimensiones todos sus componentes podemos hacerlos nulos excepto el del lugar $h+1$:

Hagamos $r^1 = (o, o, \ldots r_h \ldots o)$ y la [5.3] queda:

$$\varphi_{c_h}(r_h) = e^{ir_h\gamma_h - \frac{r_h\sigma^2 r_h}{2p_h^2}} \qquad [5.4]$$

ya que

$$[o, o, .. r_h \ldots o] = \begin{bmatrix} \frac{1}{p_o^2} & & \\ & \frac{1}{p_h^2} & \\ & & \frac{1}{p_k^2} \end{bmatrix} \begin{bmatrix} o \\ \vdots \\ r_h \\ \vdots \\ o \end{bmatrix} = \frac{r_h^2}{p_h^2} \qquad [5.5]$$

Según [1.18] y [2.13'].

Recordemos que r_h es un parámetro semejante a t empleado en la notación de la función característica.

O sea, la función característica de la variable aleatoria c_h unidimensional normal tiene los siguientes parámetros

Media	γ_h	[5.6a]
Varianza	$\frac{\sigma^2}{p_h^2}$	[5.6b]

Demostramos en la Sección 3.ª del Capítulo anterior que las estimaciones c_h del parámetro poblacional γ_h de una muestra de tamaño n, eran estimadores ELIO y ahora recordando la Sección 3.ª A-4 del Capítulo II las distribuciones de estos estimadores son normales con los parámetros indicados en las [5.6]:

$$c_h \sim N(\gamma_h, \frac{\sigma}{p_h})$$

2. *Distribución de los cuadrados de las desviaciones*

2.1. Desviaciones

En la Sección 2.ª 2. del capítulo anterior [4.16] vimos que el vector desviaciones puede ponerse así:

$$\mathbf{D} = [\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}']\epsilon = \mathbf{M}\epsilon$$

siendo **M** la matriz idempotente con las propiedades indicadas en la misma sección.

Por el Teorema de Adición, toda combinación lineal de variables normales, es normal. Luego **D** es normal.

Su función característica es:

$$\varphi_{\mathbf{D}}(\tau) = E e^{i\tau'\mathbf{D}} = E e^{i(\tau'\mathbf{M})\epsilon} = e^{-\frac{\tau'\mathbf{M}\mathbf{I}(\tau'\mathbf{M})'\sigma^2}{2}} =$$

$$= e^{-\frac{\tau'(\mathbf{I}-\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}')\tau\sigma^2}{2}} \qquad [5.7]$$

según vimos en [3.6] y por las propiedades de **M**.

La [5.7] es la función característica de una distribución normal n dimensional, pero la matriz **M** es de $n-(k+1)$ dimensiones.

2.2. Desviaciones al cuadrado. Distribución de $D'D/\sigma^2$

1. De las [4.15] y [4.16] obtenemos la componente t del vector desviaciones:

$$D_t = Y_t - \sum_{h=o}^{k} P_h(t)\, c_h \qquad [5.8]$$

La variable D_t es normal y esperanza matemática nula. Es normal, porque es combinación lineal de variables normales según [3.20] y [5.3].

Luego, la [4.19] es la suma de los cuadrados de las desviaciones normales y la variable

$$\chi^2 = \frac{D'D}{\sigma^2} \qquad [5.9]$$

es una χ^2 de Person con $n-(k+1)$ grados de libertad, pues es suma de cuadrados de variables normales y cuya esperanza son los grados de libertad [4.21].

2. La distribución de $D'D/\sigma^2$ puede obtenerse de

$$\chi_n^2 = \frac{\epsilon'\epsilon}{\sigma^2} = \frac{\sum_{t=1}^{n} (Y_t - \sum_{h=o}^{k} P_h(t)\,\gamma_h)^2}{\sigma^2}$$

porque por hipótesis $\frac{\epsilon_t^2}{\sigma^2}$ es una variable aleatoria normal tipificada al cuadrado. Si se introduce en el segundo miembro

$$\hat{Y}_t = \sum_{h=o}^{k} P_h(t)\, c_h$$

sumando y restando para que aparezcan las desviaciones D_t, elevando al cuadrado y simplificando, queda la expresión

$$\chi_n^2 = \frac{\mathbf{D'D}}{\sigma^2} + \sum_{h=o}^{k} \left(\frac{c_h - \gamma_h}{\frac{\sigma}{p_h}} \right)^2 = \frac{\mathbf{D'D}}{\sigma^2} + \chi_{k+1}^2 \qquad [5.10]$$

porque cada término de la sumatoria:

$$\chi_1^2 = \left(\frac{c_h - \gamma_h}{\frac{\sigma}{p_h}} \right)^2 \qquad [5.11]$$

es por [5.4′] una normal tipificada al cuadrado que sigue una χ^2 con un grado de libertad y por el teorema de adición de la χ^2 la suma del segundo término de [5.10] es otra χ^2 de Pearson con $k+1$ grados de libertad. Y por el Teorema de Partición $D'D/\sigma^2$ sigue también una χ^2 de Pearson con $n-(k+1)$ grados de libertad, y las variables $\mathbf{D'D}$ y $c_h - \gamma_h$ *son independientes.*

2.3. Cuasivarianza

De la suma de los cuadrados $\mathbf{D'D}$ dividida por los grados de libertad $n-(k+1)$ tenemos un estimador para la varianza del modelo σ^2:

$$\hat{\sigma}^2 = \frac{D'D}{n-(k+1)} \qquad [5.12]$$

y según [4.22] este estimador es insesgado de σ^2.

La [5.12] desarrollada puede verse en [4.23].

2.4. Teorema de independencia (1)

Hipótesis: Dadas las variables aleatorias $\mathbf{D'D}$ y las $c_h - \gamma_h$ donde $\mathbf{D'D}$ es la suma de cuadrados de las desviaciones [4.19] y $c_h - \gamma_h$ es la desviación del coeficiente regresor respecto a su parámetro poblacional [4.1′].

Tesis: Entre las variables indicadas existe una incorrelación y por ser $\mathbf{D}$ y $\mathbf{C}$ normales son independientes.

Demostración.

(1) Este Teorema ha sido demostrado en 2.2 por los Teoremas de la χ^2. Ahora, lo que hacemos es demostrar la incorrección.

Para demostrar la incorrelación calculemos la esperanza de las variables indicadas:

$$E\,\mathbf{D}'\mathbf{D}\,(c_p - \gamma_p)$$

Si sustituimos $\mathbf{D}'\mathbf{D}$ y $c_p - \gamma_p$ por sus valores [4.19] y [4.7′] y recordamos [4.17] la anterior (que es un escalar), puede escribirse

$$E\,\mathbf{D}'\mathbf{D}\,(c_h - \gamma_h) = E\,\epsilon'\mathbf{M}\,\mathbf{M}'\epsilon\,\epsilon'\mathbf{P}_h p_h^2 =$$

$$E\epsilon'\mathbf{M}\,\epsilon\,\epsilon'\,\frac{\mathbf{P}_h}{p_h^2} = \frac{1}{p_h^2}\,E\,\mathbf{P}_h'\mathbf{M}'\epsilon\,\epsilon'\epsilon = \frac{\mathbf{P}_h'\mathbf{M}}{p_h^2}\,E\,\epsilon\,\epsilon'\epsilon' \qquad [5.13]$$

Por las propiedades del cálculo matricial y de $\mathbf{M}$ según 3. de la Sección 2.ª del Capítulo anterior $= \mathbf{P}_h'\mathbf{M}$ es un vector fila de n dimensiones, y $E\epsilon\,\epsilon'\epsilon$ es un vector columna de n dimensiones. Demostremos que:

$$\mathbf{P}_h'\mathbf{M} = \mathbf{O}$$

es un vector fila nulo.

En efecto:

$$\mathbf{P}_h'\mathbf{M} = \mathbf{P}_h'[\mathbf{I} - \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'] = \mathbf{P}_h' - \mathbf{P}_h'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' \qquad [5.14]$$

pero:

$$\mathbf{P}_h'\mathbf{P} = [P_h(1), \dots P_h(\mathrm{t}) \dots P_h(n)] \begin{pmatrix} P_o(1) \dots P_h(1) \dots P_k(1) \\ \dots\dots\dots\dots \\ P_o(n) \dots P_h(n) \dots P_k(n) \end{pmatrix} =$$

$$= [o, \dots p_h^2 \dots o] \qquad [5.15]$$

por las condiciones de ortogonalidad def. III del Cap. I y [1.18].

Sustituyendo $(\mathbf{P}'\mathbf{P})^{-1}$ por su valor [4.5] y premultiplicando por [5.15] tenemos que

$$\mathbf{P}_h'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1} = [0, \dots p_h^2 \dots o] \begin{bmatrix} \frac{1}{p_o^2} & & \\ & \frac{1}{p_h^2} & \\ & & \frac{1}{p_k^2} \end{bmatrix} = \qquad [5.16]$$

$$= [o, o, \dots 1, \dots o]$$

Finalmente, el sustraendo de la [5.14] es:

$$\mathbf{P}_h'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' = [o, o, \dots 1 \dots o] \begin{pmatrix} P_o(1) \dots P_o(t) \dots P_o(n) \\ \dots\dots\dots\dots \\ P_h(1) \dots P_h(t) \dots P_h(n) \\ \dots\dots\dots\dots \end{pmatrix} =$$

$$= [P_h(1), \dots P_h(t), \dots P_h(n)] = \mathbf{P}_h' \quad \Rightarrow$$

$$\mathbf{P}'\mathbf{M} = \mathbf{P}_h' - \mathbf{P}_h'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' = \mathbf{O} \qquad [5.17]$$

Llevando este valor a: [5.13] tenemos:

$$E\,\mathbf{D}'\mathbf{D}\;(c_h - \gamma_h) = E\epsilon'\mathbf{M}\,\epsilon\epsilon'\frac{\mathbf{P}_h}{p_h^2} = \frac{P_h'\,M}{p_h^2}\;E\epsilon\epsilon'\epsilon = \mathbf{O} \qquad [5.18]$$

3. *Distribución del estimador c_h cuando se desconoce σ*

Por [5.4′] c_h es normal y tipificando, tenemos

$$\frac{c_h - \gamma_h}{\dfrac{\sigma}{p_h}} \sim N(o, 1) \qquad - \qquad \tau$$

Y por [5.9] $\mathbf{D}'\mathbf{D}/\sigma^2$ es una χ^2 con $n-(k+1)$ grados de libertad. Según el teorema 2.3 estas variables son independientes. Luego, podemos formar una t de Student:

$$t_{n-(k+1)} = \frac{c_h - \gamma_h}{\sigma}\cdot p_h : \sqrt{\frac{\mathbf{D}'\mathbf{D}}{\sigma^2(n-[k+1])}} =$$

$$= \frac{(c_h - \gamma_h)\,p_h}{\sqrt{\mathbf{D}'\mathbf{D}}}\sqrt{n-(k+1)} = \frac{(c_h - \gamma_h)\,p_h}{\hat{\sigma}} \qquad [5.19]$$

representando por $\hat{\sigma}^2$ la cuasivarianza centrada [5.12].

4. *Distribución de la regresión estimada.* $\hat{Y}$

Examinemos la distribución $-n-$ dimensional para después estudiar la de la componente Y_i:

La regresión ajustada es:

$$\hat{\mathbf{y}} = \mathbf{P}\,\mathbf{C} \qquad [5.20]$$

donde C son variables normales y, en consecuencia, y es normal, ya que la matriz $\mathbf{P}$ no es aleatoria.

Hallemos su *f.* característica:

$$\varphi_{\hat{Y}}(\tau) = E\,e^{i\tau'\hat{Y}} = E\,e^{i\tau'\mathbf{PC}} \qquad [5.21]$$

Si ponemos $\tau'\mathbf{P} = \mathbf{S}' \Rightarrow \quad \mathbf{S} = \mathbf{P}'\tau$ [5.22]

Sustituyendo en [5.21] y por ser r un vector fila $-n-$ dimensional, también $\mathbf{S}'$ es vector paramétrico. Recordando [5.1] y [5.2] tenemos:

$$\varphi_{\hat{Y}}(\tau) = E\;e^{i\mathbf{S}'\mathbf{C}} = e^{i\mathbf{S}'\gamma - \frac{\mathbf{S}'(\mathbf{P}'\mathbf{P})^{-1}\sigma^2\mathbf{S}}{2}}$$

y reestableciendo el parámetro τ [5.22] tenemos:

$$\varphi_{\hat{Y}}(\tau) = E\, e^{i\tau' \mathbf{P}\mathbf{C}} = e^{i\tau' \mathbf{P}\gamma - \frac{\tau' \mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\sigma^2 \mathbf{P}'\tau}{2}} \qquad [5.23]$$

La ecuación característica [5.23] nos informa que la función de densidad conjunta de y es una normal —n— dimensional con los siguientes parámetros estructurales:

Medias: $\mathbf{P}\,\boldsymbol{\gamma}$ [5.24a]

Matriz de covarianzas: $\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\sigma^2$ [5.24b]

5. *Distribución del valor ajustado* $\hat{Y}_t$

1. *Caso de conocer* σ

Podemos hacerlo directamente o por medio de la función característica anterior haciendo las componentes distintas de τ_t nulas.

En este caso, examinemos el exponente de [5.23] para el valor particular del parámetro fila:

$$\boldsymbol{\tau}^1 = (o, o, \ldots r_t \ldots o)\,. \qquad [5.25]$$

La [5.23] se reduce

$$\varphi_{\hat{Y}_t(r_t)} = e^{i r_t \sum_{h=0}^{k} P_h(t)\gamma_h - \frac{r_t^2 \sigma^2}{2} \sum_{h=0}^{k} \frac{P_h(t)^2}{\rho_h}} \qquad [5.26]$$

porque de [5.25] sustituido en la [5.2], tenemos:

$$r^1 P\gamma = (o, \ldots r_t \ldots o) \begin{pmatrix} P_o(1) \ldots P_h(1) \ldots P_k(1) \\ \ldots\ldots\ldots\ldots\ldots \\ P_o(t) \ldots P_h(t) \ldots P_k(t) \\ \ldots\ldots\ldots\ldots\ldots \\ P_o(n) \ldots P_h(n) \ldots P_k(n) \end{pmatrix} \begin{pmatrix} \gamma_o \\ \vdots \\ \gamma_h \\ \vdots \\ \gamma_k \end{pmatrix} =$$

$$= r_t (P_o(t), \ldots P_h(t) \ldots P_k(t)) \begin{pmatrix} \gamma_o \\ \vdots \\ \gamma_h \\ \vdots \\ \gamma_k \end{pmatrix} = r_t \sum_{h=o}^{n} P_h(t)\,\gamma_h \qquad [5.27a]$$

De forma análoga: De [5.25]

$$\boldsymbol{\tau}'\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\boldsymbol{\sigma}^2\mathbf{P}'\boldsymbol{\tau} = r_t^2\sigma^2\{(P_o(t)\dots P_h(t)\dots P_k(t)\}\ \{\frac{1}{p_h^2}\}\begin{pmatrix} P_o(t) \\ \dots\dots \\ P_h(t) \\ \dots\dots \\ P_k(t)\end{pmatrix} =$$

$$= r_t^2\sigma^2 \sum_{h=o}^{k} \frac{P_h(t)^2}{p_h^2} \qquad [5.27b]$$

y queda justificada la ecuación característica [5.26].

La distribución de $\hat{Y}_t$ es según [5.26]:

$$\hat{Y}_t \sim N\left(\sum_{h=o}^{k} P_h(t)\gamma_h,\ \sigma\sqrt{\sum_{h=o}^{k}\frac{P_h(t)^2}{p_h^2}}\right), \qquad [5.28]$$

NOTA: Esta distribución puede obtenerse aplicando el Teorema de adición de variables normales.

2.º *Caso de desconocerse* σ^2

Formemos una t de Student tipificando la 5.28 y diviendo por el mismo divisor cuando estudiamos la distribución del estimador c_h (es decir: la raíz cuadrada de la media de la χ^2 con n-k-1 g. e.):

$$t_{n-(k+1)} = \frac{\hat{Y}_t - \sum_{h=o}^{k} P_h(t)\gamma_h}{\sigma\sqrt{\sum_{h=o}^{k}\frac{P_h(t)^2}{p_h^2}}} : \sqrt{\frac{\mathbf{D}'\mathbf{D}}{\sigma^2[n-(k+1)]}} =$$

$$= \frac{\hat{Y}_t - \sum_{h=o}^{k} P_h(t)\gamma_h}{\hat{\sigma}\sqrt{\sum_{h=o}^{k}\frac{P_h(t)^2}{p_h^2}}} \qquad [5.29]$$

donde $\hat{\sigma}^2$ representa la cuasivarianza centrada [5.12].

SECCION 2.ª

Estimación por intervalo

1. *Generalidades*

La estimación por punto nos permite determinar un solo valor del parámetro poblacional cuando se nos da un conjunto de datos observados y aunque sea verosímil, la probabilidad que el valor estimado coincida con el verdadero parámetro poblacional prácticamente es nula.

Para evitar este riesgo y a su vez tener cierta seguridad en las estimaciones se establecen intervalos aleatorios dentro de los cuales existe una probabilidad determinada que al repetir el experimento, el parámetro poblacional se encuentre dentro del intervalo.

Se elige un nivel de significación cuyas notaciones ϵ o α (no hay que confundir ni con la perturbación ni con la media).

Son las probabilidades de error, o sea, que el parámetro poblacional caiga fuera del intervalo.

Con estas aclaraciones previas y conociendo las:

Distribuciones de los parámetros regresores
Distribución de los cuadrados
Distribución de la regresión estimada

podemos estudiar los correspondientes intervalos de confianza cuando se conozca o no la varianza poblacional.

En esta sección estudiamos los intervalos de las estimaciones efectuadas.

2. *Intervalo de confianza para* $\hat{\sigma}^2$

Sabemos por [4.12] que

$$\hat{\sigma}^2 = \frac{\mathbf{D}'\mathbf{D}}{n-(k+1)}$$

es un estimador centrado de σ^2 y que

$$\frac{\mathbf{D}'\mathbf{D}}{\sigma^2}$$

sigue una χ^2 de Pearson con $n-(k+1)$ grados de libertad [5.9].

Si elegimos unos niveles de significación para ambas colas ϵ_1 y ϵ_2 $\epsilon_1 + \epsilon_2 = \epsilon$, las Tablas de la χ^2, está tabulada el área de la derecha para un valor de la variable χ^2 y con los grados de libertad correspondiente. El área

ϵ_1 de la izquierda es equivalente al $1-\epsilon_1$ del área de la derecha buscando en las Tablas, y de esta forma podemos utilizar las mismas y el estimador [5.9] que tenga probabilidad $1-\epsilon$ de estar entre los valores:

$$\chi^2_{n-(k+1)}(1-\epsilon_1) < \frac{\mathbf{D}'\mathbf{D}}{\sigma^2} < \chi^2_{n-(k+1)}(\epsilon_2) \qquad \Rightarrow$$

$$\frac{\mathbf{D}'\mathbf{D}}{\chi^2_{n-(k+1)}(\epsilon_2)} < \sigma^2 < \frac{\mathbf{D}'\mathbf{D}}{\chi^2_{n-(k+1)}(1-\epsilon_1)} \qquad [5.30]$$

nos permite determinar un intervalo para σ^2.

Este intervalo *será aleatorio o no* según se considere **D′D** como estimador o se efectúe una estimación.

En el primer caso, de ajustarse el experimento reiteradamente, tendremos que un $(1-\epsilon)$ 100 de las veces σ^2 caerá en el interior del intervalo.

Si por el contrario, **D′D** es *una estimación*, el intervalo no es aleatorio y *nos da una confianza que contenga al parámetro poblacional*, con un coeficiente fiducial $1-\epsilon$ (no probabilidad).

Importa indicar que no es necesariamente $\epsilon_1 = \epsilon_2 = \frac{\epsilon}{2}$ y lo decimos, porque existe también el problema que, dado el intervalo aleatorio [5.30] sea el menor posible, y esto depende de las elecciones apropiadas de los niveles de significación. Es decir, con un mismo nivel de significación, podemos obtener intervalos más selectivos.

3. *Intervalo de confianza para c_h*

Dos casos pueden presentarse: que se conozca o no la varianza poblacional.

a) *Caso en que conozcamos la varianza σ^2*

Si tipificamos c_h según la [5.4] tenemos la variable normal $u \sim N(o, 1)$

$$u = \frac{c_h - \gamma_h}{\frac{\sigma}{p_h}} = \frac{(c_h - \gamma_h)\, p_h}{\sigma} \qquad [5.31]$$

Elegido un nivel de significación a dos bandas, la variable aleatoria u estará comprendida con probabilidad $1-\epsilon$ entre los límites:

$$-u_{\frac{\epsilon}{2}} < \frac{c_h - \gamma_h}{\sigma} \cdot p_h < u_{\frac{\epsilon}{2}} \qquad [5.32]$$

donde $u_{\frac{\epsilon}{2}}$ lo tomamos de la normal de forma que la probabilidad a la derecha de $u_{\frac{\epsilon}{2}}$ sea precisamente $\frac{\epsilon}{2}$.

De la [5.32] deducimos el siguiente *intervalo aleatorio si consideramos c_h como una función estimadora:*

$$c_h - \frac{\sigma}{p_h} \cdot u_{\frac{\epsilon}{2}} < \gamma_h < c_h + \frac{\sigma}{p_h} u_{\frac{\epsilon}{2}} . \qquad [5.33]$$

Y para una estimación de c_h tomada de una muestra con el mismo nivel de significación, el intervalo *no aleatorio* es:

$$c_h - \frac{\sigma}{p_h} u_{\frac{\epsilon}{2}} < \gamma_h < c_h + \frac{\sigma}{p_h} u_{\frac{\epsilon}{2}} . \qquad [5.33']$$

NOTA: Las dos expresiones dependen de considerar c_h como estimador o como estimación. En este último caso, el intervalo no es aleatorio y la probabilidad se dice probabilidad fiducial. c_h es una estimación concreta.

b) Caso en que desconozcamos σ^2.

De la ley de distribución obten para c_h [5.19] cuando no conocemos el valor de σ^2, pero sí su estimación $\hat{\sigma}^2$ centrada [4.12] es una ley de Student con $n-(k+1)$ grados de libertad:

$$t_{n\ (k+1)} = \frac{(c_h - \gamma_h)\, p_h}{\hat{\sigma}} \qquad [5.34]$$

Y eligiendo un nivel de significación ϵ a dos bandas ($\frac{\epsilon}{2}$ para cada cola) con una probabilidad $1-\epsilon$, el cociente [5.34] caerá dentro del intervalo:

$$-t_{n\ (k+1)}\left(\frac{\epsilon}{2}\right) < \frac{c_h - \gamma_h}{\hat{\sigma}}\, p_h < t_{n\ (k+1)}\left(\frac{\epsilon}{2}\right)$$

donde $t_{n\ (k+1)}\left(\frac{\epsilon}{2}\right)$ lo tomamos de las tablas de la t de Student.

El intervalo de confianza para el parámetro regresor γ_h, en este caso, es:

$$c_h - \frac{\hat{\sigma}}{p_h} t_{n\ (k+1)} \left(\frac{\epsilon}{2}\right) < \gamma_h < c_h + \frac{\hat{\sigma}}{p_h} t_{n\ (k+1)} \left(\frac{\epsilon}{2}\right) \qquad [5.35]$$

Al comparar este intervalo con la [5.33] observamos unas fórmulas parecidas: El valor σ *hay que sustituirlo por la estimación* $\hat{\sigma}$ y, el valor $u\left(\frac{\epsilon}{2}\right)$ *de la* *normal hay que sustituirlo por el de las Tablas de la t de Student con $n-(k+1)$ grados de libertad.*

Si $n-(k+1)$ fuere grande (prácticamente, superior a 60), podemos emplear la normal.

4. *Intervalo de confianza para la tendencia ajustada que contenga a la tendencia real*

Existen dos casos según conozcamos o no σ.

a) *Caso en que se conozca σ.*

Sabemos que estos estimadores son ELIO y deseamos determinar para un valor de t el intervalo que comprenda la tendencia teórica que según [3.26] sabemos es $E\,Y_t = E\,Y_t = P(t)\gamma$.

Y por [5.28] conocemos la distribución $\hat{Y}_t$ que es normal y tipificando:

$$u = \frac{\hat{Y}_t - \sum_{h=o}^{k} P_h(t)\gamma_h}{\sigma\sqrt{\sum \frac{P_h(t)^2}{p_h^2}}} \Rightarrow$$

deducimos para el nivel de significación ϵ (a dos bandas) el intervalo de confianza para la tendencia, y tomando $u_{\frac{\epsilon}{2}}$ de la Tabla de la Normal:

$$\hat{Y}_t - \sigma_{\hat{Y}_t}\cdot u_{\frac{\epsilon}{2}} < \sum_{t=o}^{k} P_h(t)\gamma_h < \hat{Y}_t + \sigma_{\hat{Y}_t} u_{\frac{\epsilon}{2}} \qquad [5.36]$$

donde hemos puesto:

$$\sigma_{\hat{Y}_t} = \sigma\sqrt{\sum_{h=o}^{k} \frac{P_h(t)^2}{p_h^2}} \qquad [5.37]$$

El intervalo [5.36] nos indica con probabilidad $1-\epsilon$ que la tendencia *por los puntos observados* se encontrara entre los límites extremos indicados.

b) *Caso en que se desconozca σ*

Nos encontramos que la distribución de $\hat{Y}_t$ estudiada en [5.29] sigue la t de Student.

$$t_{n-(k+1)} = \frac{\hat{Y}_t - \sum_{h=o}^{k} P_h(t)\gamma_h}{\hat{\sigma}\sqrt{\sum_{h=o}^{k} \frac{P_h(t)^2}{p_h^2}}}$$

que nos permite elegido el nivel de significación ϵ tomando valores de la t en las Tablas con n— $(k+1)$ grados de libertad deducir el intervalo de confianza que contenga la tendencia:

$$\hat{Y}_t - \hat{\sigma}\sqrt{\Sigma \frac{P_h(t)^2 \gamma_h}{p_h^2}}\; t_{n-(k+1)}\left(\frac{\epsilon}{2}\right) < \Sigma P_h(t)\gamma_h < \hat{Y}_t + \hat{\sigma}\sqrt{\Sigma \frac{P_h(t)^2}{p_h^2}} \cdot t_{n-(k+1)}\left(\frac{\epsilon}{2}\right) \quad [5.38]$$

Este intervalo hay que considerarlo dentro del recinto muestral sin el concepto (que estudiaremos) de extrapolación. Recuérdese que $\hat{\sigma}^2$ es la cuasivarianza [5.12].

5. *Intervalos de confianza de las variables del modelo*

1. *Intervalo de confianza de ϵ_t*

De las hipótesis efectuadas para el modelo en la Sección 2.ª Capítulo II tenemos que:

$$\epsilon_t \quad N(o,\sigma) \quad \Rightarrow \quad \left|\frac{\epsilon_t}{\sigma}\right| < u_{\frac{\epsilon}{2}} \quad [5.38']$$

nos permitirá conocer si en un conjunto de observaciones experimentadas el $(1-\epsilon)\,100$ es inferior en valor absoluto a $u\left(\frac{\epsilon}{2}\right)$.

Pero las ϵ_t no se conocen directamente sino a través del modelo y en el supuesto que los parámetros sean conocidos:

$$\epsilon_t = Y_t - \sum_{h=o}^{k} P_h(t)\gamma_h$$

que nos permite ciertas consideraciones.

2. *Intervalo de confianza de Y_t*

Conocemos su distribución por [3.21], luego es fácil deducir el intervalo para Y_t:

$$\sum_{h=o}^{k} P_h(t)\gamma_h - \sigma\, t_{\frac{\epsilon}{2}} < Y_t < \sum_{h=o}^{k} P_h(t)\gamma_h + \sigma\, t_{\frac{\epsilon}{2}} \quad [5.38'']$$

que nos permite comprobar si están comprendidos los valores observados entre los valores extremos si se conocen los parámetros del modelo.

SECCIÓN 3.ª

Análisis sobre el grado de la tendencia

1. *Relación entre la parábola de grado k y los polinomios*

En el Capítulo II, Sección 2.ª.2 estudiamos el planteamiento del problema para establecer la relación de enlace entre la tendencia expresada por una parábola de grado k y esta misma tendecia expresada en forma de polinomios ortogonales [2.8].

También estudiamos la fórmula [2.9] que nos permite conocer los coeficientes γ_j en función de los β_h $(h=j, \dots m)$. Y recordemos también la [2.15], que nos indica que si la parábola es de grado $k - 1$ $(\beta_k = o)$ *necesariamente* es $\gamma_k = o$ (pero no viceversa).

Estas fórmulas nos permiten el planteamiento de hipótesis nulas $\beta_k = o$ (parábola de grado inferior) frente a las alternativas $\beta_k \neq o$.

2. *Descomposición de los cuadrados en distintos órdenes*

En [4.19] vimos que

$$\mathbf{D'D} = \sum_{t=1}^{n} [Y_t - \sum_{h=o}^{k} c_h P_h(t)]^2 = \sum_{t=1}^{n} Y_t^2 - n\bar{y}^2 - \sum_{h=1}^{k} c_h^2 p_h^2 \qquad [5.39]$$

si el modelo tiene una tendencia parabólica de grado k.

La [5.39] puede escribirse también en la siguiente forma:

$$\mathbf{D'D} = \sum_{t=1}^{n} \left(Y_t^2 - \overline{n\bar{y}^2} - \sum_{h=1}^{i} c_h^2 p_h^2 \right) - \sum_{h=i+1}^{k} c_h^2 p_h^2 \Rightarrow \qquad [5.40]$$

$$\mathbf{D'D}(i) = \mathbf{D'D} + \sum_{h=i+1}^{k} c_h^2 p_h^2 \qquad [5.41]$$

Llamamos $\mathbf{D'D}(i)$ a los cuadrados de las desviaciones observadas cuando el ajuste es hasta el polinomio $P_i(t)$. En consecuencia, para $i = k$:

$$\mathbf{D'D} = \mathbf{D'D}(k) \qquad [5.42]$$

La $\mathbf{D'D}$ son sumas de cuadrados de $n-(k+1)$ g. de libertad. Si la [5.41] dividimos por σ^2 tenemos:

$$\frac{\mathbf{D'D}}{\sigma^2}(i) = \frac{\mathbf{D'D}}{\sigma^2} + \sum_{h=i+1}^{k} \frac{c_h^2 p_h^2}{\sigma^2} \qquad [5.43]$$

En la hipótesis de considerar $\gamma_{i+1} = \dots = \gamma_k = o$, las variables aleatorias $\frac{c_h p_h}{\sigma}$ son normales N (o, 1) y según [5.11] la segunda sumatoria de [5.43] no es sino la suma de $k - i$ variables normales al cuadrado e independientes: o sea, una χ^2 de Pearson con $k - i$ grados de libertad. Y según hemos dicho, $\mathbf{D}'\mathbf{D}/\sigma^2$ sigue otra χ^2 de Pearson con $n - (k+1)$ g. de libertad, luego recordando el teorema de Adición $\mathbf{D}'\mathbf{D}(i)/\sigma^2$ sigue también la misma distribución con los grados de libertad la suma de ambos $= n - (k+1) + k - i =$ $= n - (i+1)$ [5.44].

Los cuadrados (y varianzas) pueden descomponerse como indicamos en [5.41] o también:

$$\mathbf{D}'\mathbf{D}(i) = \mathbf{D}'\mathbf{D} + p_{i+1}^2\, c_{i+1}^2 + \dots + p_k^2\, c_k^2 \qquad [5.45]$$

La expresión del primer miembro es la suma de los cuadrados de las desviaciones *cuando se considera la hipótesis que la tendencia sea de grado i.*

De la [5.40] para $i = o$ tenemos:

$$\mathbf{D}'\mathbf{D}(o) = \sum_{t=1}^{n} Y_t^2 - n\,y^2 = n\,s_y^2 \quad \Rightarrow$$

$$\frac{\mathbf{D}'\mathbf{D}(o)}{\sigma^2} = \chi^2_{n-1} = \frac{ns_y^2}{\sigma^2} \Rightarrow s_y^2 = \frac{\Sigma y_t^2}{n} - \bar{y}^2 \qquad [5.46]$$

recordando [5.44]. La [5.46] es cuando no tiene tendencia.

La relación [5.45] nos permite formar esquemas para analizar la varianza y hacer hipótesis sobre las estimaciones de los coeficientes regresores y también del grado de la tendencia.

3. *Esquemas prácticos*

1. A continuación exponemos el esquema siguiente para el análisis de la varianza y coeficientes regresores.

i	$c_i^2 p_i^2$	Suma de residuos	Media residual $D'D(i)/n-(i+1)$	Función de Snedecor $F_{1,n-(i+1)}$	Tablas de F Snedecor $F_{1,n-i-1}(\epsilon)$
o	—	$D'D + \sum_{h=1}^{k} c_h^2 p_h^2$	$(D'D + \sum_{t=1}^{k} c_h^2 p_h^2) : (n-1)$		
1	$c_1^2 p_1^2$	$D'D + \sum_{h=2}^{k} c_h^2 p_h^2$	$(D'D + \sum_{i=2}^{k} c_h^2 p_h^2) : (n-2)$	$F_{1,n-2}$	$F_{1,n-2}(\epsilon)$
2	$c_2^2 p_2^2$	$D'D + \sum_{h=3} c_h^2 p_h^2$	$(D'D + \sum_{i=3}^{k} c_h^2 p_h^2) : (n-3)$	$F_{1,n-3}$	$F_{1,n-3}(\epsilon)$
3	$c_3^2 p_1^2$	$D'D + \sum_{h=4} c_h^2 p_h^2$	$(D'D + \sum_{i=4}^{k} c_h^2 p_h^2) : (n-4)$	$F_{1,n-4}$	$F_{1,n-4}(\epsilon)$
4	$c_4^2 p_3^2$	$D'D + \sum_{i=5} c_h^2 p_h^2$	$(D'D + \sum_{i=5}^{k} c_h^2 p_h^2) : (n-5)$	$F_{1,n-5}$	$F_{1,n-5}(\epsilon)$
. . .	. . .	. . . .	. . . .	. . . . .	. . . . . .
j	$c_j^2 p_j^2$	$D'D + \sum_{h=j+1}^{k} c_h^2 p_h^2$	$(D'D + \sum_{i=j+1}^{k} c_h^2 p_h^2) : (n-(j+1))$	$F_{1,n-(j+1)}$	$F_{1,n-(j+1)}(\epsilon)$

La suma de residuos de la fila $i = o, 1, 2...$ (tercera columna), no es sino $D' D(i)$ según [5.45]. Estos valores divididos por σ^2 siguen una χ^2 de Pearson con $n - (i + 1)$ grados de libertad.

La cuarta columna son estimadores centrados de la varianza del modelo.

Por ser $c_h \sim N(\gamma_h, \frac{\sigma}{p_h})$ en la hipótesis $\gamma_h = o$, las variables de la segunda columna son variables normales al cuadrado. [5.31], y si la dividimos por σ^2 es:

$$\left(\frac{c_i}{\frac{\sigma}{p_i}}\right)^2 = \frac{c_i^2 p_i^2}{\sigma^2} = \chi_1^2 \qquad [5.47]$$

en la hipótesis de que $\gamma_i = o$.

La quinta columna es el cociente de la segunda y la cuarta columnas, y es una F. de Snedecor: si, repetimos $\gamma_i = o$:

$$F_{1, n-(i+1)} = \frac{\frac{c_i^2 p_i^2}{\sigma^2}}{\frac{D' D(i)}{\sigma^2 (n - c - 1)}} = \frac{c_i^2 p_i^2 (n - i - 1)}{D' D + c_{i+1}^2 p_{i+1}^2 + ... + c_k^2 p_k^2}$$

con un grado de libertad en el numerador y $n - (i + 1)$ en el denominador.

2. Otro esquema puede plantearse en sentido inverso. De la [5.45] deducimos:

$$\mathbf{D' D}(i) = \mathbf{D' D} + \sum_{h=i+2}^{k} c_h^2 p_h^2 + c_{i+1}^2 p_{i+1}^2$$

Los dos primeros sumandos no es sino $D' D(i + 1) \quad \Rightarrow$

$$\mathbf{D' D}(i + 1) = \mathbf{D' D}(i) - C_{i+1}^2 p_{i+1}^2 \qquad [5.49]$$

Recuérdese por [5.46] que $\mathbf{D' D}(o) = \Sigma Y_i^2 - n\bar{y}^2$.

La [5.49] es importante y, por recurrencia, pueden obtenerse cuadrados de las desviaciones.

El esquema anterior comienza por la fila $i = o$ en la columna tercera con los residuos [5.46] y por la ley de recurrencia [5.49] obtenemos la columna tercera. Previamente precisamos la segunda columna. Las columnas cuarta y quinta se calculan como explicamos anteriormente.

4. *Análisis del grado de la tendencia*

1. El primer planteamiento sería por el análisis de la varianza.

Si el estimador $\frac{D' D(i)}{n - (i + 1)}$ [5.50] de la varianza tiende a estabilizarse (dentro de los límites de las fluctuaciones muestrales) para un valor de $i = k - 1$, entonces puede tomarse el estadístico [5.50] que para la hipótesis de rechazar $H_o : \gamma_k = o$ (o sea, que el grado de la tendencia sea inferior a k), elegido un nivel de significación ϵ rechazaremos la hipótesis nula aceptando la alternativa $H_1 : \gamma_k \neq o$ si se verifica:

$$F_{1, n-(k+1)}(\epsilon) < \frac{c_k^2 p_k^2 (n-k-1)}{\mathbf{D'D}} \qquad [5.51]$$

La anterior puede escribirse:

$$\frac{c_k^2 p_k^2}{D' D} > \frac{F_{1, n-k-1}(\epsilon)}{n-k-1} \qquad [5.52]$$

y nos indica si el primer miembro supera al segundo que el grado de la tendencia no es inferior a k, y si las estimaciones centradas de la varianza prácticamente están estabilizadas, podemos afirmar que el grado de la tendencia es k.

Luego si la desigualdad [5.52] es contraria, aceptamos la hipótesis H_o: $\gamma_k = o$ y el grado de la tendencia será inferior a k, porque esta hipótesis es que el polinomio k no contenga el modelo.

2. En nuestra hipótesis de conocer mejor el modelo si al examinar la columna de las medias de las estimaciones de las varianzas éstas se estabilizan a partir de un valor i en adelante, podremos formar el estadístico con una significación ϵ en la hipótesis $H_o : \gamma_{i+1} = \ldots = \gamma_k = o$ [5.53] y la decisión a rechazar γ_i.

$$F_{1, n-(i+1)}(\epsilon) < \frac{c_i^2 p_i^2 (n-i-1)}{D' D + c_{i+1}^2 p_{i+1}^2 + \ldots + c_k^2 p_k^2} \qquad [5.54]$$

Si rechazamos $\gamma_i = o$ tenemos que aceptar la alternativa de que la tendencia es de grado no inferior a i.

Este es un esquema por *pasos sucesivos* para determinar el grado de la tendencia.

3. Pero examinemos el siguiente esquema que nos permite con un nivel de significación ϵ, determinar el grado de la tendencia con una probabilidad fiducial $1 - \epsilon$.

Pretendemos utilizar la [5.43] y que caso de ser de grado i la tendencia, la hipótesis nula es:

$$H_o = \gamma_{i+1} = \gamma_{i+2} = \ldots = \gamma_k = o \qquad [5.53']$$

En este caso de [5.43] tenemos:

$$\chi^2_{k-i} = \sum_{h=i+1}^{k} \frac{c_h^2 p_h^2}{\sigma^2} \qquad [5.55]$$

y que intervienen *todos los coeficientes estimados* desde $i + 1$ hasta k.

Y como $D' D/\sigma^2$ es otra χ^2 de Pearson con $n-(k+1)$ grados de libertad el cociente de las medias (dividido entre sus grados de libertad) es una F. de Snedecor:

$$F_{k-i,\, n-(k+1)} = \frac{\sum_{h=i+1}^{k} c_h^2 p_h^2}{(k-i)\,\sigma^2} : \frac{D' D}{(n-(k+1))\,\sigma^2} =$$

$$= \frac{\sum_{h=i+1}^{k} c_h^2 p_h^2}{(k-i)\,\hat{\sigma}^2} \qquad [5.56]$$

donde $\hat{\sigma}^2$ es la cuasivarianza [5.12].

Si elegido un nivel de significación ϵ y tomamos de las Tablas de la F. de Snedecor con $k-i$ grados de libertad del numerador y $n-(k+1)$ del denominador si

$$F_{k-i,\, n-(k+1)}(\epsilon) > \frac{\sum_{h=i+1}^{k} c_h^2 p_h^2}{(k-i)\,\hat{\sigma}^2} \qquad [5.57]$$

admitiremos la hipótesis [5.49] y el grado de la tendencia será igual o inferior a i con un nivel de significación ϵ.

Con la ayuda de la [5.52] formamos el esquema para el

ANÁLISIS DEL GRADO DE LA TENDENCIA

i	$\sum_{h=i+1}^{k} c_h^2 p_h^2$	Número sumandos	Medias	$F_{k-i,\, n-(k+1)}$	Tablas $F_{k-i,\, n-(k+1)}(\epsilon)$
o	$\sum_{1}^{k} c_h^2 p_h^2$	k	$\left(\sum_{2}^{k} c_h^2 p_h^2\right)\frac{1}{k}$	$F_{k,\, n-(k+1)}$	$F_{k,\, n-(k+1)}(\epsilon)$
1	$\sum_{2}^{k} c_h^2 p_h^2$	$k-1$	$\left(\sum_{2}^{k} c_h^2 p_h^2\right)\frac{1}{k-1}$	$F_{k-1,\, n-(k+1)}$	$F_{k-1,\, n-(k+1)}(\epsilon)$
2	$\sum_{3}^{k} c_h^2 p_h^2$	$k-2$	$\left(\sum_{3}^{k} c_h^2 p_h^2\right)\frac{1}{k-2}$	$F_{k-2,\, n-(k+1)}$	$F_{k-2,\, n-(k+1)}(\epsilon)$
. . .	. . . . . .	. . . .	. . . .	. . . .	. . . .
j	$\sum_{j+1}^{k} c_h^2 p_h^2$	$k-j$	$\left(\sum_{j+1}^{k} c_h^2 p_h^2\right)\frac{1}{k-j}$	$F_{k-j,\, n-(k+1)}$	$F_{k-j,\, n-(k+1)}(\epsilon)$

La quinta columna se forma de la cuarta dividida por la estimación de la cuasivarianza [5.12]. Los valores así obtenidos los comparamos con los de las Tablas, y si éstos son superiores, se admite la hipótesis del grado expresado por la primera columna i [5.49], porque puede admitirse la hipótesis nula $H_o: \gamma_{i+1} = \ldots = \gamma_k = o$.

Si los valores de la columna quinta son superiores a los de la Tabla, la desigualdad [5.57] es contraria y rechazaríamos la hipótesis H_0 y admitiríamos la alternativa, o sea, que el grado de la tendencia es superior a i al nivel determinado.

CAPITULO VI

Predicción de la tendencia

SECCIÓN 1.ª

Posibilidad de modificaciones del modelo

1. *Generalidades*

El ajuste por polinomios ortogonales de una serie temporal nos permite estimar los parámetros poblacionales ($\gamma_0, \ldots \gamma_k$) que mejor expliquen los resultados observados.

Pero nunca olvidemos que de un proceso de naturaleza estocástica pudieran extraerse (al menos teóricamente) infinitas realizaciones del tipo

$$\{Y_{ti,} \, t \in E\} \qquad i = 1, 2, 3 \ldots \qquad [6.1]$$

E i representa una muestra o serie temporal de tamaño n ($t \in E$).

Para $i = 1$ tendremos una muestra del proceso $\{ Y_t, t \in E\}$ para $i = 2$ tendremos otra serie temporal $\{Y_{t2,} \, t \in E\}$ del mismo proceso. Si para cada serie ajustamos el mismo modelo, las funciones estimadoras de los parámetros poblacionales explican el comportamiento de la tendencia basada en esta nueva serie temporal dentro del dominio observado.

Ahora bien: El planteamiento de un modelo para explicar la trayectoria de una serie observada ($t \in E$) es distinto de la predicción. Podemos hacer un ajuste casi perfecto a los datos observados y, sin embargo, no reflejar la verdadera naturaleza de la tendencia. Si decimos esto es para orientar a los lectores y evitarles que cometan errores. En principio haríamos unas sencillas preguntas: ¿El tamaño de la muestra era representativa? ¿La estimación de la varianza estaba estabilizada? ¿La ley de probabilidad de la perturbación aleatoria futura es la misma o diferente? ¿El modelo representativo del fenómeno es el mismo o evoluciona? ¿Sigue tendencia parabólica u otra expresión matemática?

Estas preguntas tienen una respuesta: nuestro desconocimiento de la estructura interna del proceso. Si conociéramos las representaciones espectrales sabríamos más sobre la naturaleza del proceso estocástico y sería posible encontrar fórmulas óptimas (en sentido lineal) de extrapolación. En mi trabajo (1) basado en el conocimiento de la forma de la función de densidad espectral trato este problema de la extrapolación de *procesos estocásticos de naturaleza estacionaria.*

El tamaño de la muestra es importante, ya que, a veces, la escasez de datos impide el conocimiento del modelo. En las series de tipo económico este tamaño generalmente es insuficiente y es aconsejable con muestras parciales, la estimación de la varianza sí cambia con el transcurso del tiempo para el mismo modelo. Esto nos hace pensar que el proceso estocástico, el modelo considerado y estudiado por ajuste de una muestra de la serie temporal se altera la tendencia y su ley de probabilidad.

2. *Covarianza estacionaria*

Si sostenemos la hipótesis de que la tendencia es parabólica, pero la estacionariedad en covarianza se ha modificado, si llamamos $\alpha(t)$ a la esperanza matemática en el momento t, tenemos:

$$Cov(t, s) = E\ [Y_t - \alpha(t)][Y_s - \alpha(s)] \qquad [6.2]$$

donde el símbolo E es el de esperanza matemática.

Esta covarianza, si depende de la diferencia de tiempos exclusivamente:

$$Cov(t, s) = B(t-s) = E[Y_t - \alpha(t)]\ [Y_s - \alpha(s)] \qquad [6.3]$$

nos permite si conociésemos la forma de $B(u)$ [6.3'].

($u = t - s$) completar el proceso en dos:

1. Un proceso de tipo funcional parabólico

$$\alpha(t) = \sum_{h=o}^{k} \gamma_h P_h(t)$$

2. Otro proceso de naturaleza estacionaria y que tiene por representación espectral (2)

$$\zeta_t = \int_{-\pi}^{\pi} e^{i\lambda t} d\zeta(\lambda) \qquad [6.4]$$

con función de densidad espectral de tipo conocido.

En estos casos, la predicción sería compuesta por dos sumandos:

La tendencia parabólica del proceso en el punto t más la correspondiente extrapolación del proceso de tipo estacionario (3) y según sea la función de

(1) URBELZ IBARROLA, F. J.: «Interpolación, extrapolación y filtraje». *Anales Ins. Actuarios E.*, 1977.

(2) Véase mi artículo citado.

(3) JAVIER URBELZ, F.: «Interpolación, extrapolación y filtraje». *Anuario Ins. Actuario.*

densidad espectral de la función de covarianza [6.3] la extrapolación sería óptima. En este caso, nos encontraríamos con procesos autorregresivos, de medias móviles, mixtos, etc., combinados con la tendencia.

3. *Condiciones para la predicción*

En principio hemos adoptado un modelo y hemos ajustado sus parámetros.

Nunca olvidemos nuestras hipótesis: hemos eliminado las variaciones estacionales y cíclicas. También estas variaciones (que pueden representarse por funciones de Fourier) pueden añadirse a la forma parabólica. Otro día estudiaremos este problema y que complementarían la tendencia.

Y puesto que partimos de la hipótesis al analizar la serie temnporal en $t \epsilon E$ si ahora es diferente $t = s \notin E$ resulta para $s > n$ (n máximo valor de $t \epsilon E$) se nos plantean muchos problemas de los anteriormente expuestos y que ahora restringimos con el fin de simplificar y que el predictor recoja la esencia de la experimentación:

1.ª Se base sobre datos del pasado.

2.ª Los axiomas fundamentales del modelo y leyes de probabilidad sean invariantes.

3.ª Sea un estimador ELIO: lineal, insesgado y óptimo.

Distintos planteamientos pueden hacerse:

1.º Si el valor $\hat{Y}_s$ puede ser un valor extrapolado, o sea, el predictor óptimo, y

2.º Si el predictor es el valor desconocido de la trayectoria de los datos $Y_t \ldots Y_s$. En este caso, este elemento de la trayectoria (o su predictor) sería el óptimo.

Los planteamientos son distintos y las varianzas también. Por eso es conveniente distinguir entre estas matizaciones.

La conveniencia de comparar los resultados del predictor con los valores futuros es importante y nos permite admitir el predictor o modificar nuestras hipótesis sobre el modelo.

SECCIÓN 2.ª

Estimador ELIO del predictor

1. *Naturaleza lineal*

Recordando las hipótesis teníamos una muestra de una trayectoria $\{y_1, y_2, \ldots y_n\}$ [6.5] y donde

$$y_t \sim N\left(\sum_{h=o}^{k} \gamma_h P_h(t), \sigma\right) \qquad [6.6]$$

es válido para todo $t \in E \quad \{1, 2, \ldots n\}$. [6.7]

Si ahora hacemos $t = s \qquad s \notin E$ [6.8]

$$E \quad Y_s = \sum_{h=o}^{k} \gamma_h P_h(s) \qquad [6.9]$$

de acuerdo con los axiomas.

Formemos, a partir de la muestra [6.5], un predictor para $E\,Y_s$ que denominaremos a ese predictor

$$\hat{Y}_s = d_1 y_1 + d_2 y_2 + \ldots + d_t y_t + \ldots + d_n y_n \qquad [6.10]$$

siendo $s > n$.

La primera pregunta que se nos presenta es la siguiente:

$$\text{¿será } \hat{Y}_s = \sum_{h=o}^{k} c_h P_h(s) \quad ? \qquad [6.11]$$

donde c_h son los valores estimados a partir de la serie temporal [6.5].

Demostremos que la respuesta es afirmativa.

aunque $\qquad s \notin$ E

2. Insesgado

El planteamiento lineal para el predictor $\hat{Y}_s$ lo completaremos con la condicion de que $\hat{Y}_s$ sea centrado.

Para ello

$$\mathrm{E}\,(\Sigma\, d_t y_t - y_s) = o \qquad [6.12]$$

Pero el predictor $\hat{Y}_s$ puede escribirse sustituyendo su valor por [6.10] y los de Y_t por los del modelo [2.6′]

$$\hat{Y}_s = \sum_{t=1}^{n} d_t y_t = \sum_{t=1}^{n} d_t \left(\sum_{h=o}^{k} \gamma_h P_h(t) + \epsilon_t\right) =$$

$$= \sum_{h=o}^{k} \gamma_h \sum_{t=1}^{n} d_t P_h(t) + \sum_{t=1}^{n} d_t \epsilon_t \qquad [6.13]$$

$$E\,\hat{Y}_s = \sum_{h=o}^{k} \gamma_h \sum_{t=1}^{n} d_t\, P_h(t) = E\, Y_s = \sum_{h=o}^{k} \gamma_h\, P_h(s) \Rightarrow \qquad [6.14]$$

$$\sum_{t=1}^{n} d_t\, P_h(t) = P_h(s) \qquad h = o, 1, \dots k. \qquad [6.14]$$

porque los coeficientes de γ_h de ambos miembros tienen que ser idénticos.

Las [6.14] nos dan $k+1$ ecuaciones condicionadas para que $\hat{Y}_s$ sea centrado. Si estas condiciones las sustituimos en [6.13] el estimador Y_s puede escribirse:

$$\hat{Y}_s = \sum_{h=o}^{k} \gamma_h\, P_h(s) + \sum_{t=1}^{n} d_t\, \epsilon_t \qquad [6.15]$$

La expresión aleatoria del predictor depende de las funciones no aleatorias $P_h(s)$ en el punto s; $(s \notin E)$ de los parámetros d_t (que determinaremos) y de las perturbaciones aleatorias en $t \in E$.

3. *Optimización de la varianza*

En la Sección 3.ª del Capítulo IV estudiamos un problema semejante, aunque reducido al campo $t \in E$ aplicaremos el teorema de Markof para que la varianza del predictor [6.15] sea mínima recordando las condiciones impuestas para que sea centrado.

Primero, por su condición de ser $\hat{Y}_s$ insesgado

$$E\,\hat{Y}_s = \sum_{h=o}^{k} \gamma_h\, P_h(s) + \sum_{t=1}^{n} d_t\, E\,\epsilon_t = \sum_{h=o}^{k} \gamma_h\, P_h(s) \qquad [6.16]$$

varianza de $\hat{Y}_s$ es

$$\operatorname{Var} \hat{Y}_s = E(\hat{Y}_s - \sum_{h=o}^{k} \gamma_p\, P_h(s)\,)^2 = E\left[\sum_{t-1}^{n} d_t^2\,\epsilon_t^2 + \sum_{t \neq t'} d_{t'}\, d_t\, \epsilon_t\, \epsilon_{t'}\right]$$

$$\sigma_{\hat{Y}_s}^{\,2} = \sigma^2 \sum_{t=1}^{n} d_t^2 \qquad [6.17]$$

según la [6.15] y por las propiedades de $E\,\epsilon_t\, \epsilon_{t'} = \sigma^2\, \delta_{t,t'}$.

La varianza de cualquier predictor lineal [6.10] depende de d_t, $t \in E$. Y la [6.17] se hará mínima si determinamos la variabilidad de $d_t\,(t \in E)$ de forma que cumplan las condiciones [6.14].

El problema se reduce a determinar el mínimo de la [6.17] condicionado a la $k+1$ ecuaciones [6.14]. Aplicamos el sencillo método de los multiplicadores de Lagrange.

Aquí, σ^2 es el parámetro del modelo, y hemos de encontrar los valores de d_t $(t = 1, 2, \dots n)$ que, sustituidos en [6.17] nos den la mínima varianza del estimador lineal [6.10].

Formemos la función:

$$\Phi = \sum_{t=1}^{n} d_t^2 - 2 \sum_{h=o}^{k} \lambda_h \left(\sum_{t=1}^{n} d_t P_h(t) - P_h(s) \right) \qquad [6.18]$$

La última expresión son las condiciones impuestas al predictor en [6.14] multiplicada cada una por $-2\lambda_h$.

Calcular el mínimo de la varianza [6.17] condicionado a las [6.14] equivale a calcular el mínimo de [6.18], porque el primer sumando multiplicado por σ^2 nos da [6.17] y los otros son nulos.

Derivando parcialmente respecto a d_t tenemos:

$$\frac{\partial \Phi}{\partial d_t} = 2 d_t - 2 \sum_{h=o}^{k} \gamma_h P_h(t) \qquad [6.18']$$

La condición de extremo es que la [6.18′] se anule.

Luego

$$d_t = \sum_{j=o}^{k} \gamma_j P_j(t) \qquad [6.19]$$

En la expresión anterior hemos sustituido el subíndice h por j para mayor claridad en la demostración siguiente.

Llevado el valor [6.19] a la [6.14] tenemos:

$$P_h(s) = \sum_{t=1}^{n} d_t P_h(t) = \sum_{t=1}^{n} \left(\sum_{j=o}^{k} \gamma_j d_t P_j(t)) P_h(t) \right) = \gamma_h \sum_{t=1}^{n} P_h(t)^2 \Rightarrow$$

$$\gamma_h = \frac{P_h(s)}{\sum_{t=1}^{n} P_h(t)^2} = \frac{P_h(s)}{p_h^2} \qquad [6.20]$$

por ortogonalidad según la Def. III del Cap. I y recordando la notación introducida en [1.18].

Sustituyendo este valor en la expresión [6.19] tenemos los valores de los coeficientes lineales:

$$d_t = \sum_{j=o}^{k} \frac{P_j(s)}{p_j^2} . P_j(t) = \sum_{j=o}^{k} \frac{P_j(t)}{p_j^2} . P_j(s) \qquad [6.21]$$

La varianza mínima del estimador lineal [6.10] es sustituir [6.21] en [6.17]:

$$\sigma_{\hat{Y}s}^2 = \sigma^2 \sum_{t=1}^{n} \left(\sum_{j=o}^{k} \frac{P_j(s) P_j(t)}{p_j^2} \right)^2 =$$

$$= \sigma^2 \sum_{t=1}^{n} \left[\sum_{j=o}^{k} \frac{P_j(t)^2 P_j(s)^2}{p_j^4} + \sum_{j \neq h} \frac{P_j(t) P_h(t) P_j(s) P_h(s)}{p_j^2 p_h^2} \right] =$$

$$= \sigma^2 \left[\sum_{j=o}^{k} \frac{P_j(s)^2}{p_j^4} \sum_{t=1}^{n} P_j(t)^2 + \sum_{j \neq h} \frac{P_j(s) P_h(s)}{p_j^2 p_h^2} \sum_{j=1}^{n} P_j(t) P_h(t) \right]$$

Y por las condiciones de ortogonalidad indicadas tenemos:

$$\sigma^2_{\hat{Y}_s} = \sigma^2 \sum_{j=o}^{k} \frac{P_j(s)^2}{p_j^2} \qquad [6.22]$$

Comparemos esta expresión con la varianza de $\hat{Y}_t$ [5.27b] para $t \in E$ y observamos que en el caso del predictor $\hat{Y}_s$, la varianza es de forma semejante, pero recordando que $s \notin E$.

SECCIÓN 3.ª

Predictor de la tendencia

1. *Predictor ELIO* $\hat{Y}_t$

En 1. de la Sección 2.ª establecimos el predictor $\hat{Y}_s$ *como una combinación lineal de* $\{y_1\, y_2 \ldots y_n\}$ para $s > n$ y nos preguntamos si sería de la forma [6.11].

En [6.21] determinamos los valores de d_t que, sustituidos en [6.10], nos da el predictor que cumple las condiciones de linealidad, insesgadez y de mínima varianza en sentido lineal.

Así, pues:

$$\hat{Y}_s = \sum_{t=1}^{n} d_t\, y_t = \sum_{t=1}^{n} \left(\sum_{j=o}^{k} \frac{P_j(t)}{p_j^2} P_j(s) \right) . \; Y_t =$$

$$= \sum_{j=o}^{k} P_j(s) \sum_{t=1}^{n} \frac{P_j(t)}{p_j^2} . \; Y_t = \sum_{j=o}^{k} c_j\, P_j(s) \qquad [6.23]$$

La [6.23] nos indica que para $t = s \quad s \notin E$, el predictor $\hat{Y}_s$ coincide con el valor de la tendencia regresora estimada $\hat{Y}_t$ [6.11] dada, pero en el punto s.

El predictor $\hat{Y}_t$ de la tendencia regresora es un estimador ELIO.

SECCIÓN 4.ª

Varianzas del predictor

1. *Varianza* $\hat{Y}_s$ *con respecto a su media teórica*

La media teórica de $\hat{Y}_s$ es [6.9] y en consecuencia su varianza es la indicada en [6.22].

Esta varianza puede escribirse

$$\sigma^2_{\hat{Y}s} = \sigma^2 \left[\frac{P_o(s)^2}{p_o^2} + \frac{P_1(s)^2}{p_1^2} + \dots + \frac{P_k(s)^2}{p_k^2} \right] \qquad [6.23]$$

Pero como $P_o(s) = 1$ y $P_o^2 = \sum_{t=1}^{n} P_o(t) = n$ la [6.23] *es:*

$$\hat{\sigma}^2_{\hat{Y}s} = \sigma^2 \left[\frac{1}{n} + \frac{P_1(s)^2}{p_1^2} + \dots + \frac{P_k(s)^2}{p_k^2} \right] \qquad [6.23']$$

Repetimos: *esta varianza es respecto a la línea*

$$E\,Y_s = \sum_{h=o}^{k} \gamma_h P_h(s)$$

de tendencia considerada ésta como función no aleatoria.

2. *Varianza de $\hat{Y}_s$ respecto a Y_s*

La serie temporal muestral de n valores, en el momento s, tomará otro valor Y_s, en el supuesto de eliminadas componentes estacionales y cíclicas.

En este caso, la esperanza de Y_s coincide también con la tendencia (no aleatoria y función exclusiva del tiempo). Luego

$$E(\hat{Y}_s - Y_s) = o \Rightarrow E\hat{Y}_s = EY_s$$

$$\sigma^2_{Y_s}(\hat{Y}_s) = E[(\hat{Y}_s - E\hat{Y}_s) - (Y_s - EY_s)]^2 =$$

$$= \text{Var}\ \hat{Y}_s + \text{Var}\ Y_s - 2\,E(\hat{Y}_s - E\hat{Y}_s)\,(Y_s - EY_s)$$

$$= \sigma^2 \left(1 + \sum_{h=o}^{k} \frac{P_h(s)^2}{p_h^2}\right) = \sigma^2 \left[1 + \frac{1}{n} + \sum_{h=1}^{k} \frac{P_h(s)^2}{p_h^2} \right] \qquad [6.24]$$

porque $Y_s - E\,Y_s = \epsilon_s \;\Rightarrow\; E\,\hat{Y}_s\,\epsilon_s = o = E\{(E\,\hat{Y}_s)\,\epsilon_s\} = o$, ya que la variable aleatoria $\hat{Y}_s$ depende de $\{\epsilon_t,\ t \in E\}$ según [6.15] y $E\,\hat{Y}_s = \sum_{h=o}^{k} \gamma_h P_h(s)$ es no aleatoria $\Rightarrow E(E\,\hat{Y}_t)\,\epsilon_s\} = \sum_{h=1}^{k} \gamma_h P_h(t)\,E\,\epsilon_s = o$.

3. *Representaciones matriciales de las varianzas*

1. La de la tendencia estimada sobre la teórica.

$$\sigma^2_{\hat{Y}s} = \sigma^2\,\mathbf{P}(s)'\,\mathbf{A}^{-1}\,\mathbf{P}(s) \qquad [6.25]$$

donde $\mathbf{A}$ es la matriz diagonal

$$\mathbf{A} = \mathbf{P}'\mathbf{P} = \{p_h^2\} \Rightarrow$$

$$\mathbf{A}^{-1} = \left\{\frac{1}{p_h^2}\right\}$$

2. La del proceso Y_s respecto a la tendencia estimada:

$$\sigma^2_{Y_s}(\hat{Y}_s) = \sigma^2(\mathbf{I} + \mathbf{P}(\mathbf{S})'\mathbf{A}^{-1}\mathbf{P}(\mathbf{S})) \qquad [6.26]$$

SECCIÓN 5.ª

Estimación por intervalo para la tendencia teórica

1. *Distribución del estimador* $\hat{Y}_s$

De la [5.4′] deducimos que por ser $c_j \sim N(\gamma_i, \frac{\sigma}{p_i}) \Rightarrow \hat{Y}_s$ combinación lineal [6.23] es normal según el Teorema de Adición, y sus parámetros son:

$$\text{Media} \quad \sum_{i=0}^{k} \gamma_i P_i(s) \qquad [6.27a]$$

$$\text{Varianza } \sigma^2 \sum_{i=0}^{k} \frac{P_i(s)^2}{p_i^2} \qquad [6.27b]$$

y nos permite formar intervalos de confianza para el *predictor de la regresión* EY_s.

Examinemos dos casos: según se conozca o no la varianza σ^2.

1.º *Caso en que se conozca* σ^2

Tipificando el estimador [6.23] tenemos:

$$u = \frac{\hat{Y}_s - \sum_{i=0}^{k} \gamma_i P_i(s)}{\sigma\sqrt{\sum_{i=0}^{k} \frac{P_i(s)^2}{p_i^2}}} \qquad [6.28]$$

donde $u \sim N(o, 1)$. Para un nivel de significación ϵ el cociente de aceptación a dos bandas será:

$$\hat{Y}_s - u\left(\frac{\epsilon}{2}\right)\sigma\sqrt{\sum_{i=0}^{k} \frac{P_i(s)^2}{p_i^2}} < \sum_{i=0}^{k} \gamma_i P_i(s) < \hat{Y}_s + u\left(\frac{\epsilon}{2}\right)\sigma\sqrt{\sum_{i=0}^{k} \frac{P_i(s)^2}{p_i^2}} \qquad [6.29]$$

La [6.29] nos indica con una confianza de $1-\epsilon$ que la *tendencia teórica se encuentra dentro del intervalo.* $u(\frac{\epsilon}{2})$ indica el valor de las tablas de la normal cuya área desde ese valor hasta el infinito es $\frac{\epsilon}{2}$.

Decimos la tendencia teórica o la esperanza matemática del proceso en el punto s en las condiciones de no alterarse el modelo.

Observemos que la desviación típica $\hat{\sigma}_{\hat{Y}_s}$ aumenta con $s\ (s > n)$ luego el intervalo [6.29] tiene mayor amplitud, por lo que la tendencia teórica será más difícil de precisar por punto y los intervalos son menos selectivos cuanto más distantes se encuentren los valores extrapolados.

1.º *Caso en que se desconozca* σ^2

En este caso empleamos la t de Student con $n - (k+1)$ grados de libertad:

$$t_{n-(k+1)} = \frac{u}{\sqrt{\dfrac{D'D}{\sigma^2(n-(k+1))}}} = \frac{\hat{Y}_s - \sum_{j=o}^{k} \gamma_j P_j(s)}{\hat{\sigma}\sqrt{\sum_{j=o}^{k} \dfrac{P_j(s)}{p_j^2}}} \qquad [6.30]$$

Para llegar a la [6.30] hemos sustituido u por [6.28] y $D'D/\ (n-(k+1)$ por la varianza estimada [5.12].

Elegido un nivel de significación ϵ entraremos en las tablas de la t de Student y el intervalo semejante a [6.29] excepto en los valores de $t_{n-(k+1)}$ y $\hat{\sigma}$ es:

$$\hat{y}_s \pm t_{n-(k+1)}\left(\frac{\epsilon}{2}\right)\hat{\sigma}\sqrt{\sum_{j=o}^{k} \frac{P_j(s)}{p_j^2}} \qquad [6.31]$$

con probabilidad $1 - \epsilon$ que contenga la tendencia teórica [6.27a].

SECCIÓN 6.ª

Estimación por intervalo de $\{Y_s \quad s \notin E)$

1. A veces, con la línea de regresión de la tendencia estimada deseamos establecer un intervalo de confianza que comprenda los valores del proceso estocástico en el punto $s\{\ Y_s, s \epsilon E\}$.

Es preciso aclarar bien conceptos para no utilizar fórmulas erróneas en las aplicaciones.

La varianza [6.24] es muy distinta a la [6.23') que es la del predictor ajustado respecto de la tendencia teórica.

Ahora, el problema es distinto: La varianza de Y_s se mide respecto a la regresión estimada.

a) *Caso en que se conozca* σ^2

Por hipótesis, el modelo es normal y también

$$Y_s - \hat{Y}_s \sim N\left(o, \sigma\sqrt{1 + \sum_{h=o}^{k} \frac{P_h(s)^2}{p_h^2}}\right) \qquad [6.32]$$

aplicando el Teorema de Adición de variables normales.

Tipificando y tomando el nivel de significación ϵ a dos bandas, tenemos que el valor

$$\hat{Y}_s - \sigma \sqrt{1 + \sum_{h=o}^{k} \frac{P_h(s)^2}{p_h^2}} u \left(\frac{\epsilon}{2}\right) < y_s < \hat{y}_s + \sigma \sqrt{1 + \sum_{h=o}^{k} \frac{P_h(t)^2}{p_h^2}} u \left(\frac{\epsilon}{2}\right) \quad [6.33]$$

donde $\hat{Y}_s$ es la [6.11].

b) *Caso en que se desconozca σ^2*

Aplicaremos la t de Student con $n-(k+1)$ grados de libertad utilizando la cuasivarianza centrada y, sencillamente y para no repetir, llegamos al intervalo de confianza:

$$\hat{Y}_s - \hat{\sigma} \sqrt{1 + \sum_{h=o}^{k} \frac{P_h(s)^2}{p_h^2}} \cdot t_{n-(k+1)}\left(\frac{\epsilon}{2}\right) < y_s < \hat{y}_s + \hat{\sigma}\sqrt{1 + \sum \frac{P_s(s)^2}{p_h^2}} t_{n-(k+1)}\left(\frac{\epsilon}{2}\right) \quad [6.34]$$

La fórmula se utiliza para comprobar si los valores de la «serie temporal futura» se encuentran en el intervalo [6.34] y lo harán con una confianza de un $(1-\epsilon)$ 100 siempre que no se modifique la estructura del modelo.

CAPITULO VII

Aplicaciones

SECCIÓN 1.ª

1. *Bases fundamentales*

1. Se nos da una muestra de una serie temporal que consta de 23 términos consecutivos. Con el fin de no repetirla aparece en la columna cuarta de la Tabla siguiente. Esta serie ha sido originada de una muestra del proceso estocástico:

$$Y_t = 24 - 2\, P_1(t) + o,4\, \mathrm{P}_2(t) + \epsilon_t \quad [7.1]$$

donde $\gamma_o = 24$ $\gamma_1 = -2$ $\gamma_2 = o,4$ (parámetros teóricos) y $\epsilon_t\, N(o, 1.5)$ $\sigma = +1.5$ [7.2]

2. Los polinomios son las expresiones

$$P_1(t) = t - \bar{t} = t - 12 \quad [7.3]$$

$$P_2(t) = (t-12)^2 - \frac{23^2 - 1}{12} \quad [7.4]$$

por ser $n = 23$, según las fórmulas [1.33] y [1.38]. Los polinomios $P_3(t)$ y $P_4(t)$ son [1.38a] y [1.38b].

3. La segunda columna (tendencia teórica) es precisamente

$$E\,Y_t = 24 - 2\,P_1(t) + o,4\,P_2(t) \qquad [7.5]$$

porque $E\epsilon_t = o$.

4. La realización de la perturbación aleatoria se ha tomado de nuestras Tablas Estadísticas (4), de la Tabla de Números Aleatorios de la Distribución normal. $N(o, 1)$ multiplicada por la desviación típica $\sigma = 1.5$ eligiendo dos decimales solamente por fines didácticos.

5. La suma de las columnas 2.ª y 3.ª nos da la *simulación de la realización muestral* de la serie que es la columna 4.ª de Y_t.

2. *Preparación*

Se trata de:

1.º Determinar la tendencia aplicando los Polinomios ortogonales, estimando los parámetros poblacionales.

2.º Análisis de la varianza y su aplicación para conocer el grado de la tendencia. (Niveles 1% y 5%).

3.º Intervalos de confianza del parámetro poblacional σ^2.

4.º Valores representativos de la tendencia observada por ajuste e intervalos de la regresión de la tendencia al 1%.

5.º Predicción para $t = 24 \leq t \leq 35$, desviaciones típicas e intervalos de confianza.

6.º Conclusiones.

3. *Solución. 1.º Ajuste*

1.º En el CUADRO I, columna 4.ª, vienen los valores de la serie temporal. Para $n = 23$ tomamos de nuestras Tablas Estadísticas los valores que *para mejor comprensión aparecen en el recuadro,* y que son los valores de los polinomios $P_1(t)$, $P_2(t)$, $P_3(t)$ y $P_4(t)$. Los Polinomios impares son de signo contrario y los pares son simétricos respecto al valor central $t = 12$. Estas indicaciones nos permiten completar las columnas 5.ª a la 8.ª.

Las columnas 9.ª, 10.ª, 11.ª y 12.ª son, respectivamente, los productos de la columna 4.ª por la 5.ª, por la 6.ª, por la 7.ª y por la 8.ª. La columna 13.ª son los cuadrados de la columna 4.ª.

2.ª Las sumas de las columnas 4.ª, 9.ª, 10.ª, 11.ª, y 12.ª y 13.ª son necesarias.

(4) URBELZ IBARROLA, F. J., y PÉREZ MAURA, A.: *Tablas Estadísticas*. Dep. Estadística. Empresariales. Santander.

Método de ajuste de los coeficientes de la tendencia por polinomios ortogonales

	$Y = EY_t$	ε_t	Y_t	P_1	P_2	P_3	P_4	YP_1	YP_2	YP_3	YP_4	Y^2
1	76.8	— .54	76.26	—11	77	—77	1643	— 838.86	5872.02	—5872.02	125295.18	5815.58
2	66.4	2.34	68.74	—10	56	—35	133	— 687.40	3849.44	—2405.90	9142.42	4725.18
3	56.8	— .52	56.28	— 9	37	— 3	—627	— 506.52	2082.36	— 168.84	—35287.56	3167.44
4	48	—1.90	46.10	— 8	20	20	—950	— 368.80	922.00	922.00	—43795.	2125.21
5	40	2.79	42.79	— 7	5	35	—955	— 299.53	213.95	1497.65	—40864.45	1830.98
6	32.8	—1.43	31.37	— 6	— 8	43	—747	— 188.22	—250.96	1348.91	—23433.39	984.08
7	26.4	—1.48	24.72	— 5	—19	45	—417	— 124.60	—473.48	1121.40	—10391.64	621.01
8	20.8	1.90	22.70	— 4	—28	42	— 42	— 90.80	—635.60	953.40	— 953.40	515.29
9	16	— .31	15.69	— 3	—35	35	315	— 47.07	—549.18	549.18	4942.35	246.18
10	12	—2.4	9.6	— 2	—40	25	605	— 19.2	—384.	240	5808.	92.16
11	8.8	— .07	8.73	— 1	—43	13	793	— 8.73	—375.39	113.49	6922.89	76.21
12	6.4	1.78	8.18	0	—44	0	858	0	—359.92	0	7018.44	66.91
13	4.8	— .98	3.82	1	—43	—13	793	3.82	—164.26	— 49.66	3029.26	14.59
14	4	—1.53	2.47	2	—40	—25	605	4.94	— 98.80	— 61.75	1494.35	6.10
15	4	1.85	5.85	3	—35	—35	315	17.55	—204.75	— 204.75	1842.75	34.22
16	4.8	—1.22	3.58	4	—28	—42	— 42	14.32	—100.24	— 150.36	— 150.36	12.82
17	6.4	.45	6.85	5	—19	—45	—417	34.25	—130.15	— 308.25	— 2857.45	46.92
18	8.8	— .96	7.84	6	— 8	—43	—747	47.04	— 62.72	— 337.12	— 5856.48	61.47
19	12	— .07	11.93	7	5	—35	—955	83.51	59.65	— 417.55	—11393.15	142.32
20	16	1.80	17.80	8	20	—20	—950	142.40	356.00	— 356.00	—16910.00	316.84
21	20.8	2.91	23.71	9	37	3	—627	213.39	877.27	71.13	—14866.17	562.16
22	26.4	.98	27.38	10	56	35	133	273.80	1533.28	958.30	4641.54	749.66
23	32.8	— .60	32.20	11	77	77	1643	354.20	2479.40	2479.40	52904.60	1036.84
			554.79	1012		32890		—1990.51		— 77.37		
					35420		13123110		14455.95		15283.73	
				1	1	1	7	— 1.9667		— .00039		
			24.103	1	1	6	12		.40813		.00068	23.250.20

Anexo página de nuestras Tablas Estadísticas donde hemos recogido los valores recuadrados para $n = 23$.

TABLAS ESTADISTICAS

Polinomios ortogonales centrados

	n = 21 P′1	P′2	P′3	P′4	n = 22 P′1	P′2	P′3	P′4	n = 23 P′1	P′2	P′3	P′4
	0	—110	0	594	1	—20	—12	702	0	—44	0	858
	1	—107	—54	540	3	—19	—35	585	1	—43	—13	793
	2	—98	—103	385	5	—17	—55	365	2	—40	—25	605
	3	—83	—142	150	7	—14	—70	70	3	—35	—35	315
	4	—62	—166	—130	9	—10	—78	—258	4	—28	—42	—42
	5	—35	—170	—406	11	—5	—77	—563	5	—19	—45	—417
	6	—2	—149	—615	13	1	—65	—775	6	—8	—43	—747
	7	37	—98	—680	15	8	—40	—810	7	5	—35	—955
	8	82	—12	—510	17	16	0	—570	8	20	—20	—950
	9	133	114	0	19	25	57	57	9	37	3	—627
	10	190	285	969	21	35	133	1197	10	56	35	133
									11	77	77	1463
PP′	770	201894	432630	5720330	3542	7084	96140	8748740	1012	35420	32890	13123110
m	1	3	5	7	2	1	1	7	1	1	1	7
d	1	1	6	12	1	2	3	12	1	1	6	12

	n = 24				n = 25				n = 26			
	P′ 1	P′ 2	P′ 3	P′ 4	P′ 1	P′ 2	P′ 3	P′ 4	P′ 1	P′ 2	P′ 3	P′ 4
	1	−143	−143	143	0	−52	0	858	1	−28	−84	1386
	3	−137	−419	123	1	−51	−77	803	3	−27	−247	1221
	5	−125	−665	85	2	−48	−149	643	5	−25	−395	905
	7	−107	−861	33	3	−43	−211	393	7	−22	−518	466
	9	−83	−987	27	4	−36	−258	78	9	−18	−606	−54
	11	−53	−1023	−87	5	−27	−285	−267	11	−13	−649	−599
	13	−17	−949	−137	6	−16	−287	−597	13	−7	−637	−1099
	15	25	−745	−165	7	−3	−259	−857	15	0	−560	−1470
	17	73	−391	−157	8	12	−196	−982	17	8	−408	−1614
	19	127	133	−97	9	29	−93	−897	19	17	−171	−1419
	21	187	847	33	10	48	55	−517	21	27	161	−759
	23	253	1771	253	11	69	253	253	23	38	598	506
					12	92	506	1518	25	50	1150	2530
PP′	4600		17760600		1300		1480050		5850		7803900	
		394680		394680		53820		14307150		16380		40060020
m	2	3	10	1	1	1	5	5	2	1	5	7
d	1	1	3	12	1	1	6	12	1	2	3	12

TABLAS ESTADISTICAS

Polinomios ortogonales centrados

	$n = 27$				$n = 28$				$n = 29$			
	$P'\,1$	$P'\,2$	$P'\,3$	$P'\,4$	$P'\,1$	$P'\,2$	$P'\,3$	$P'\,4$	$P'\,1$	$P'\,2$	$P'\,3$	$P'\,4$
	0	182	0	1638	1	—65	— 39	936	0	—70	0	2184
	1	—179	— 18	1548	3	—63	—115	840	1	—69	—104	2080
	2	—170	— 35	1285	5	—59	—185	655	2	—66	—203	1775
	3	—155	— 50	870	7	—53	—245	395	3	—61	—292	1290
	4	—134	— 62	338	9	—45	—291	81	4	—54	—366	660
	5	—107	— 70	— 262	11	—35	—319	— 259	5	—45	—420	— 66
	6	— 74	— 73	— 867	13	—23	—325	— 590	6	—34	—449	— 825
	7	— 35	— 70	—1400	15	— 9	—305	— 870	7	—21	—448	—1540
	8	10	— 60	—1770	17	7	—255	—1050	8	— 6	—412	—2120
	9	61	— 42	—1872	19	25	—171	—1074	9	11	—336	—2460
	10	118	— 15	—1587	21	45	— 49	— 879	10	30	—215	—2441
	11	181	22	— 782	23	67	115	— 395	11	51	— 44	—1930
	12	250	70	690	25	91	325	455	12	74	182	— 780
	13	325	130	2990	27	117	585	1755	13	99	468	1170
									14	126	819	4095
PP'	166[illegible]	712530	101790	56448210	7308	95004	2103660	19634160	2030	113274	4207320	107987880
m	1	3	1	7	2	1	2	7	1	1	5	7
d	1	1	6	12	1	1	3	24	1	1	6	12

Dentro del recuadro vienen por cada columna del polinomio tres valores:

— En tresbolillo aparecen la suma de los cuadrados de los polinomios.

Así: en la columna P_1^* aparece al pie 1012, para P_2^*, 35420; para P_3^* 32860, y para P_4^*, 1312310.

— También aparecen otros números (multiplicador o divisor):

m	1	1	1	7
d	1	1	6	12

correspondiente al primer polinomio, segundo, tercero o cuarto, respectivamente. Todas estas indicaciones son para nuestro ejemplo, que es $n = 23$.

3.ª Las estimaciones aparecen al pie del Cuadro I y son:

$$c_0 = \frac{565.79}{23} = 24.103$$

$$c_1 = -\frac{1990.51}{1012} \times \frac{1}{1} = -1.9669$$

$$c_2 = \frac{14495.95}{35420} \times \frac{1}{1} = .40813$$

$$c_3 = \frac{-77.37}{32890} \times \frac{1}{6} = -.00039$$

$$c_4 = \frac{15283.73}{13123110} \times \frac{7}{12} = .00068$$

obsérvese que en todos se dividen por los cuadrados. ($\Sigma\ P_0\ (t)^2 = 23$ por ser $P_0\ (t) = 1$). Además, aparecen los valores $\frac{m}{d}$ para multiplicar y obtener la estimación de los valores teóricos.

Existe un ajuste de la tendencia de segundo grado expresado por la regresión:

$$\hat{Y}_t\,(2) = 24.103 - 1.9669\ P_1\,(t) + .40813\ P_2\,(t)$$

o también añadiendo otros dos términos:

$$\hat{Y}_t\,(4) = 24.103 - 1.9669\ P_1\,(t) + .40813\ P_2\,(t) - .00039\ P_3\,(t) + .00068\ P_4\,(t)$$

donde los polinomios $P_1\,(t)$ y $P_2\,(t)$ ya se han consignado, y el $P_3\,(t)$ y $P_4\,(t)$ son las [1.38a] y [1.38b].

2.º *Análisis de la varianza y sus aplicaciones*

1. De la [5.46] y del Cuadro I deducimos:

$$D'\,D\,(o) = 23250.20 - \frac{554.79^2}{23} = 9867.94 \qquad [7.6]$$

En la 1.ª columna del Cuadro II aparece el valor de h ($h = o$, 1, 2, 3, 4) según el término del polinomio; en la 2.ª columna, los productos de los cuadrados de los valores de los polinomios ortogonales por los cuadrados de los coeficientes regresores.

Para determinar la 2.ª columna partimos de los cuadrados de los polinomios Tabulados y que aparecen al pie de la Tabla:

$$p_1^{*2} = 1012 \quad p_2^{*2} = 35420 \quad p_3^{*2} = 32890 \quad p_4^{*2} = 13123110$$

Los polinomios de las Tablas están relacionados por la fórmula

$$P_h^*(\mathrm{t}) = P_h\,(t) \cdot \frac{m}{d} \Rightarrow P_h^*(\mathrm{t})^2 = P_h\,(t)^2\ \frac{m^2}{d^2} \Rightarrow$$

$$\Sigma\ P_h^*(\mathrm{t})^2 = \frac{m^2}{d^2}\Sigma\ P_h\,(t)^2 = \frac{m^2}{d^2} \cdot p_h^{\ 2} = p_h^{*2} \Rightarrow$$

$$p_h^{\ 2} = p_h^{*2}\ \frac{d^2}{m^2} \qquad [7.7]$$

Los valores p_h^{*2} ($h = 1, 2, 3, 4$) para este ejemplo que aparece al pie de cada polinomio del Cuadro I hay que corregirlos con los valores m/d que aparecen también al pie de cada polinomio aplicando la [7.7]. Así, los valores de los cuadrados de los polinomios corregidos son:

$$p_1^{\ 2} = \quad 1012 \times \frac{1}{1} = \quad 1012$$

$$p_2^{\ 2} = \quad 35420 \quad = \quad 35420$$

$$p_3^{\ 2} = \quad 32890 \times \frac{6^2}{1^2} = \quad 1184040$$

$$p_4^{\ 2} = 13123110 \times \frac{12^2}{7^2} = 38565874$$

por ser d/m los valores inversos que aparecen al pie de cada polinomio.

2. Luego:

$p_1^{\ 2}\,c_1^{\ 2} = \quad 1012 \times (-1.9669)^2$	$= 3915.15$	3.915.15
$p_2^{\ 2}\,c_2^{\ 2} = \quad 35420 \times .40813^2$	$= 5899.91$	9.815.06
$p_3^{\ 2}\,c_3^{\ 2} = \quad 1184040 \times (-.00039)^2$	$= 1.82$	9.816.88
$p_4^{\ 2}\,c_4^{\ 2} = 38565874 \times (.000679)^2$	$= 17.80$	9.834.68

Los valores $p_h^2 c_h^2$ aparecen en la 2.ª columna del Cuadro II.

La última columna anterior (no del cuadro) es la acumulación de los anteriores resultados $D' D$.

CUADRO II

Análisis de la Varianza

h	$p_h^2 c_h^2$	$D' D (h)$	G. lib.	Medias $\hat{\sigma}^2$	F. de Snedecor			
					$F_{1, n-h-1}$	Tablas: niveles		
						1%	5%	10%
0		9867.94	22	448.54				
1	3915.11	5952.83	21	283.47	13.81	8.016	4.324	2.960
2	5899.91	52.92	20	2.64	2334.81	8.096	4.351	2.974
3	1.82	51.10	19	2.69	.65	8.185	4.380	2.989
4	17.80	33.30	18	2.36	7.54	8.285	4.413	3.007

3. La columna 3.ª se forma por la ley de recurrencia

$$D' D (h) = D' D (h-1) - p_h^2 c_h^2 \qquad [7.8]$$

demostrada en [5.49], y que nos expresa los cuadrados de las desviaciones si ajustamos hasta el polinomio h inclusive.

El valor inicial [7.6] son los cuadrados de las desviaciones sin considerar la existencia de tendencia:

4. La columna 4.ª indica los grados de libertad. Como las observaciones de la serie temporal son $n = 23$ comienzan en 22 y por polinomio introducido se pierde un grado de libertad [5.54].

5. La columna 5.ª son las medias o la estimación de $\hat{\sigma}^2$ empleando la cuasivarianza (columna 3.ª dividida por la 4.ª).

6. Finalmente, las últimas columnas son para la F. de Snedecor:

La 6.ª columna (cociente de la 2.ª y 5.ª columnas) es una F. de 1 g. de libertad del numerador, y por grados de libertad del denominador los de la columna 4.ª. Y es para admitir o rechazar la hipótesis $H_o : \gamma_h = o$; o sea, si los valores de la columna 6.ª son inferiores a los de las otras columnas, se admite; caso contrario se rechaza. Para esto se han considerado tres niveles de significación: al 1%, 5% y 10%. Están en la región crítica para $h = 1$ y $h = 2$ eligiendo cualquier nivel, o sea, que no pueden admitirse las hipótesis: $\gamma_1 = \gamma_2 = o$. Respecto a $\gamma_3 = o$ puede admitirse tal hipótesis y también $\gamma_4 = o$ con el nivel de significación 1%. Estas conclusiones se deducen comparando los valores de la 6.ª columna con los valores tabulados y obtenidos en las siguientes columnas de la misma línea: si los valores de la columna 6.ª son superiores a los tabulados se rechaza la hipótesis de ser $\gamma_h = o$ y, caso contrario, se admite la hipótesis $\gamma_h \neq o$ (por ejemplo: γ_3).

3.º *Determinación del grado de tendencia*

Otro análisis importante relacionado con el anterior es la determinación del grado de la tendencia que nos permite utilizar los polinomios precisos.

De la [5.45]

$$D'D(i) = D'D + \sum_{h=i+1}^{k} p_h^2 \, c_h^2 \qquad [7.8']$$

donde la $D'D/\sigma^2$ es una χ^2 con $n-(k+1)$ g. de libertad [5.12] y en la hipótesis: $H_o : \gamma_{i+1} = \ldots = \gamma_k = o$ [7.9]

$$\chi^2_{k-i} = \frac{\sum_{h=i+1}^{k} p_h^2 \, c_h^2}{\sigma^2}$$

según [5.43] formamos una F. de Snedecor:

$$F_{k-i,\, n-(k+1)} = \frac{\sum_{h=i+1}^{k} p_h^2 \, c_h^2}{\sigma^2 (k-i)} : \frac{D'D}{\sigma^2 (n-(k+1))} = \frac{\sum_{h=i+1}^{k} p_h^2 \, c_h^2}{(k-i) : \hat{\sigma}^2} \qquad [7.10]$$

con $k-i$ grados de libertad del numerador y $n-(k+1)$ en el denominador. Para distintos valores de i tenemos diversos valores de la F. de Snedecor, y elegidos niveles significativos, si el cociente [7.10] es superior a los de las Tablas, rechazaremos la hipótesis que sea de grado igual o inferior a i; y, caso contrario, aceptaremos la hipótesis [7.9]. La comparación se hace entre valores de una misma fila.

El siguiente esquema facilita el cálculo.

CUADRO III

Análisis del grado de tendencia

h	Sumas $p_h^2 \, c_h^2$	N.º Sumandos	Medias	$F_{k-i,\, n-(k+1)}$ $F_{k-i,\, 18}$	F. Snedecor Tablas al 1%	5%
o						
1	9.834.64	4	2458.66	1041.80	4.579	2.927
2	5.919.53	3	1973.17	836.08	5.091	3.159
3	19.62	2	9.81	4.15	6.012	3.554
4	17.80	1	17.8	7.54	8.285	4.413

El Cuadro III (relacionado con el Cuadro II) requiere poca explicación. La 2.ª columna del Cuadro III se forma de la columna 2.ª del Cuadro II por sumas hasta el final, y la 3.ª columna la hemos puesto para indicar el número de sumandos y conocer los grados de libertad del numerador.

La columna 4.ª (la hemos formado de la 2.ª dividida por la 3.ª), dividida por σ^2 es una media de una χ^2 de Pearson, y con la estimación de σ^2 es 2,36 con 18 grados de libertad (Cuadro II) la columna 5.ª es la 4.ª dividida por 2,36. Este cociente es una F. de Snedecor y tomamos los valores de las Tablas al 1% y 5% (columnas 6.ª y 7.ª) variando los grados de libertad del numerador según la columna 3.ª y tomando los grados de libertad del denominador como $n-(k+1)=23-(k+1)=18$.

Conclusiones:

1.ª Por ser $1041.8>4.579$ (y 2.927) rechazamos la hipótesis de que la tendencia sea de grado *o* (no tenga tendencia) al 1% (para $h=1$).

2.ª Igualmente, por ser $836.08>5.091$ rechazamos la hipótesis $\gamma_2=\gamma_3=$ $=\ldots=o$ *que la tendencia sea de primer grado: es superior.*

3.ª Por ser $4.15<6.012$ *admitimos que la tendencia es expresable por una parábola de segundo grado* con un nivel de significación del 1%, o sea, $H_o:\gamma_3=\gamma_4=o$ es decir: rechazamos la hipótesis alternativa $H_1:\gamma_3\neq 0$, $\gamma_4\neq o$.

4.ª Conclusión análoga obtenemos al rechazar la hipótesis de expresar la tendencia por una parábola de 4.º grado al 1% de nivel de significación.

4.º *Intervalos de confianza de los parámetros poblacionales*

En el Cuadro IV se determinan los valores de los intervalos de confianza para los parámetros poblacionales.

La columna 2.ª son los grados de libertad de la *t* de Student, y la columna 3.ª es la desviación típica estimada tomada de la raíz cuadrada de la columna 5.ª de la cuasivarianza del Cuadro II.

La columna 4.ª es la raíz cuadrada de los cuadrados corregidos y explicados en el punto 2. Análisis de la Varianza.

Las columnas 5.ª y 6.ª son los niveles de significación elegidos y la 7.ª, las estimaciones de c_h halladas en el Cuadro I y puestas al pie de las columnas YP_1, YP_2, YP_3 e YP_4.

Finalmente, tenemos los intervalos de confianza al 1% y al 5%. Cada intervalo consta de los extremos (superior e inferior) y que puede comprobarse al nivel 1% que para c_3 (—.0004, .0039) y para c_4 (—.00003, .0013) que comprenden al cero y estos intervalos con una probabilidad fiducial del 99%.

CUADRO IV

Intervalos de confianza para los estimadores de los parámetros γ_h de la tendencia y su grado

h	g. l.	Estimación $\hat{\sigma}$	P_h	t Student (Tablas) $\epsilon = 1\%$	t Student (Tablas) $\epsilon = 5\%$	Estimaciones c_h	Intervalos de γ_h al 1%	Intervalos de γ_h al 5%
0	22	21.178						
1	21	16.836	31.811	2.831	2.079	—1.9669	—3.4652, —.46859	—3.067, —.8665
2	20	1.624	188.202	2.845	2.086	.40813	.3382, . 4780	.3568, .4593
3	19	1.667	1088.136	2.860	2.093	—.00039	— .0004, .0039	.0035, .0028
4	18	1.536	6210.14	2.878	2.101	.00068	—.00005, .0013	.0001, .0011

5.º *Intervalos de confianza de la varianza*

La [5.30] nos permite elegido niveles de significación obtener intervalos de confianza para valores poblacionales con probabilidad fiducial 1 — ϵ.

Hemos elegido $\epsilon_1 = \epsilon_2 = \frac{\epsilon}{2}$ es decir: tomamos el área de la χ^2 de la izquierda igual que el área de la derecha y para los niveles $\epsilon = 1\%$ y 5%. Esto significa que por tener tabulada el área de la derecha tomamos el valor de la χ^2 $(1 - \epsilon_1)$ y el de χ^2 (ϵ_2). Con esta aclaración quedan explicadas las columnas 4.ª: χ^2 $(1 - 0{,}01/2)$ y 6.ª χ^2 $(1 - 0{,}05/2)$.

Los extremos de los intervalos se obtienen sin dificultad aplicando la fórmula [5.30] recordando los grados de libertad de la 2.ª columna.

CUADRO V

Intervalos de confianza de σ^2

h	g. l.	$D(h)$	Tablas X^2 1% $X_i^2(.995)$	Tablas X^2 1% $X_s^2(.005)$	Tablas X^2 5% $X_i^2(.975)$	Tablas X^2 5% $X_s^2(.025)$	Intervalos de confianza para σ^2 al 1%		Intervalos de confianza para σ^2 al 5%	
0	22	9867.94	8.6427	42.7957	10.9823	36.7807	230.58	1141.76	268.29	898.53
1	21	5952.83	8.0337	41.4011	10.2829	35.4789	143.78	740.98	167.78	578.90
2	20	52.92	7.4338	39.9968	9.5908	34.1696	1.32	7.11	1.54	5.51
3	19	51.20	6.8440	38.5823	8.9065	32.8523	.75	7.48	1.58	5.74
4	18	33.30	6.2648	37.1565	8.2307	31.5264	.89	5.31	1.05	4.04

6.º *Intervalos ajustados de la tendencia* (Cuadro VI)

La elección del grado de polinomio nos da una línea de regresión estimada para la tendencia. Hemos elegido los valores $\hat{Y}_t(2)$ de segundo grado y $\hat{Y}_t$ (4) de cuarto grado, porque según el Cuadro II al nivel del 1% se rechazaba $\gamma_4 = o$, aunque el 5% se admite. (Ver Cuadro III.)

Las desviaciones típicas en la hipótesis de ser de diferente grado se basan en las expresiones [5.29] y [5.45] y en nuestro caso:

Para la de segundo grado, la varianza del valor de regresión estimado sería:

$$\hat{\sigma}_{\hat{Y}_t}(2)^2 = \hat{\sigma}_2^2 \left(\frac{1}{23} + \frac{P_1(t)^2}{1012} + \frac{P_2(t)^2}{35420} \right)$$

y la raíz cuadrada es la desviación expresada en la 4.ª columna.

$P_1(t)$ es el polinomio de primer grado y
$P_2(t)$ es el de segundo grado
$P_1^2 = 1012$
$P_2^2 = 35420$

Para la de cuarto grado que admitimos a un nivel de significación del 5% (véase Cuadro III), la varianza de cada valor estimado sería:

$$\hat{\sigma}_{\hat{Y}_t}^2(4) = \hat{\sigma}_4^2 \left(\frac{1}{23} + \frac{P_1(t)^2}{1012} + \frac{P_2(t)^2}{35420} + \frac{P_3(t)^2}{1184040} + \frac{P_4(t)^2}{38545874} \right)$$

Las desviaciones típicas aparecen en la columna 5.ª del Cuadro V y los estimadores $\hat{\sigma}^2 = 2.64$ y $\hat{\sigma}^2 = 2.36$ aparecen en el Cuadro II.

Tanto las estimaciones de $\hat{Y}_2$ como de $\hat{Y}_4$, sus intervalos de confianza comprenden *la tendencia* teórica poblacional no aleatoria con la probabilidad fiducial que esté dentro de los intervalos $1 - \epsilon$.

Unos breves comentarios sobre estas dos líneas de tendencia ajustadas:

1. Los intervalos de la regresión de segundo grado son más pequeños y con la misma confianza, así para $t = 1$ la amplitud de la de segundo grado es 79.84-74.52 = 5.32, mientras que para la de cuarto grado es 82.69-75.43 = 7.26. La mínima amplitud es para $t = 11$:

Para la de segundo grado es	9.97-7.10 = 2.87
Y para la de cuarto grado	11.14-7.72 = 3.42

2. La tendencia teórica expuesta en la 2.ª columna del Cuadro I está contenida entre los intervalos de ambas tendencias ajustadas, y el Cuadro VI completa la estimación de la tendencia ajustada.

7.º *Predicción* (Cuadro VII)

1. Por el Cuadro VI comprobamos que el ajuste de la tendencia por polinomios ortogonales hasta segundo grado es suficiente.

CUADRO VI

Estimaciones de la tendencia por polinomios hasta 2.º grado $\hat{Y}(2)$ y hasta 4.º grado $\hat{Y}$ (4.º), así como desviaciones e intervalos al 5%

t	$\hat{Y}_t(2)$	$\hat{Y}_t(4)$	$\hat{\sigma}_{\hat{Y}_t}(2)$	$\hat{\sigma}_{\hat{Y}_t}(4)$	Intervalos de la regresión			
					Ajuste	2.º grado	Ajuste	4.º grado
1	77.18	79.08	.933	1.261	74.52	79.84	75.43	82.69
2	66.64	66.88	.780	.797	64.42	68.86	64.58	69.17
3	56.92	56.20	.654	.673	55.06	58.78	54.26	58.13
4	48.01	46.86	.558	.685	46.43	49.60	44.89	48.83
5	39.93	38.73	.494	.685	38.52	41.33	36.76	40.70
6	32.65	31.68	.462	.651	31.34	33.97	29.81	33.56
7	26.20	25.60	.454	.601	24.90	27.49	23.87	27.34
8	20.56	20.41	.463	.564	19.24	21.88	18.78	22.03
9	15.73	16.02	.479	.557	14.37	17.10	14.41	17.62
10	11.72	12.37	.494	.573	10.32	13.13	10.72	14.02
11	8.53	9.43	.505	.594	7.10	9.97	7.72	11.14
12	6.16	7.16	.508	.603	4.71	7.61	5.42	8.90
13	4.60	5.55	.505	.594	3.16	6.04	3.84	7.27
14	3.86	4.62	.494	.573	2.45	5.26	2.97	6.27
15	3.93	4.38	.479	.557	2.57	5.29	2.78	5.99
16	4.82	4.87	.463	.564	3.50	6.14	3.24	6.50
17	6.53	6.15	.454	.601	5.23	7.82	4.42	7.88
18	9.05	8.28	.462	.651	7.74	10.36	6.41	10.15
19	12.39	11.36	.494	.685	10.98	13.80	9.38	13.33
20	16.54	15.48	.558	.685	14.96	18.13	13.51	17.46
21	21.52	20.78	.654	.673	19.65	23.38	18.84	22.72
22	27.30	27.38	.780	.797	25.08	29.52	25.08	29.67
23	33.91	35.43	.933	1.261	31.25	36.56	31.80	39.06

Elegimos como predictor la regresión ajustada $= \hat{Y}_t(2)$ y en la 2.ª columna tomamos la $E\,Y_t$, donde Y_t es el proceso estocástico [7.1], donde ϵ_t es [7.2]. (Cuadro VII.)

2. Predecimos los valores cuando $t = 24$ hasta 35.

El valor de la predicción de la tendencia $\hat{Y}_t(2)$ es la columna 3.ª, y la estimación de su varianza es en la columna 4.ª. La varianza de la 5.ª es cuando el predictor contenga valores aleatorios $\{Y_t,\ t \notin E\}$ del proceso.

3. Los intervalos de confianza complementan ese trabajo. Obsérvese que $E\,Y_t$ está dentro del intervalo, por ejemplo, para $t = 32$ $E\,Y_t = 126.40$ y el intervalo para la tendencia teórica es de 120.79 a 139.32.

4. Respecto del otro intervalo (para Y_t) puede formarla

$$Y_t = E\,Y_t + 1.5\,\epsilon_t$$

Y tomando ϵ_t de las Tablas de Números Aleatorios de la Distribución Normal $N(o, 1)$ formaremos los valores del proceso $\{Y_t, t \notin E\}$ y compararemos si caen dentro del intervalo el 95% de los datos.

5. Las varianzas (o las raíces cuadradas que aparecen en el Cuadro VII) son la [6.26] y la [6.23'], sustituyendo σ por su estimación.

CUADRO VII

Predicción de la tendencia estimada por polinomios ortogonales de 2.º grado, desviaciones estimadas sobre la regresión y sobre el proceso e intervalos de confianza al 5%

t	$E\,Y_t$	$\hat{Y}_t(2)$	$\hat{\sigma}_{\hat{Y}_t}$	$\hat{\sigma}_{Y_t}(\hat{Y}_t)$	Intervalos al 5% Para $E\,Y_t$		Para Y_t	
24	40	41,31	1.111	1.968	38.15	44.47	35.64	46.97
25	48	49.54	1.311	2.088	45.81	53.28	43.53	55.55
26	56.8	58.60	1.532	2.233	54.24	62.96	52.17	65.03
27	66.4	68.47	1.773	2.404	63.42	73.51	61.54	75.39
28	76.8	79.15	2.032	2.602	73.37	84.94	71.66	86.64
29	88	90.65	2.311	2.825	84.08	97.23	82.52	98.79
30	100	102.97	2.608	3.072	95.55	110.39	94.12	111.82
31	112.8	116.10	2.923	3.344	107.79	124.42	106.48	125.73
32	126.4	130.05	3.256	3.639	120.79	139.32	119.58	140.53
33	140.8	144.82	3.607	3.956	134.56	155.08	133.43	156.21
34	156	160.40	3.975	4.295	149.09	171.72	148.04	172.77
35	172	176.80	4.362	4.654	164.39	189.21	163.40	190.20

BIBLIOGRAFIA.

ANDERSON, T. W.: "An introduction to multivariate statistical analysis". *John Wiley*, N. Y., 1958.

BARTLETT, M. S.: *An introduction to stochastic processes*. Cambridge University Press. London, 1966, 2.ª ed.

BOX, G. E. P. and JENKINS, G. M.: "Time series analysis for casting and control". *Holden Day*, 1970.

CRAMER, H.: "Métodos matemáticos de estadística". Aguilar. *Stationary and related stochastic processes*. John Wiley and Sons, 1968, 1.ª ed.

DOOB, J. L.: "Stochastic processes". *John Wiley and Sons*, 6.ª ed. 1967.

DHRYMES, PHOEBUS, J.: *Stastistical foundations and applications*. Harper International Edition, 1970, N. Y.

DURBIN, J.: "Efficient estimations of parameters in moving average models". *Biometrika*, Vol. 46 (1959), págs. 309-316.

"Trend elimination by moving average and variate diference filters". *Bul International de Statistic*, Vol. 39, 1961, págs. 130-141.

GOODMAN, N. R.: "Some comments of spectral analysis of time series". *Technometrics*, May, 1961.

HANNAN, S. J.: "Multiple time series". *John Wiley*, 1970.

LUKACS, E.: "Characteristic functions". *Griffin*, 1970, 2.ª ed.

MALINVAUD, Edmond: *Métodos estadísticos de la econometría*. Ariel, 1967.

URBELZ IBARROLA, F. J.: "Aplicaciones del espacio de Hilbert a la estadística". *Anales I. Actuarios Españoles*, 1975, pág. 85 a 114.

— "Interpolación, extrapolación y filtraje". *Anales I. Actuarios Españoles*, 1977.

— "Introducción a la teoría de los procesos estocásticos". *Anales I. Actuarios Españoles*, 1978.

— "Estimación del espectro de las series temporales". *Anales I. Actuarios Españoles*, 1981-82.

— "Análisis espectral de los procesos econométricos y de sus transformaciones lineales". *Anales I. Actuarios Españoles*, 1983.

YAGLOM, A. M.: "An introduction to the theory of stationary random functions". *Prentice Hall*. New Jersey, 1962.

INDICE

Introducción

CAPITULO I

Polinomios ortogonales.
Sección 1.ª *Definiciones*
I. Funciones ortogonales de variables estadísticas.
II. Polinomios ortogonales definidos en un conjunto de números naturales.
III. Polinomios ortogonales centrados.
Sección 2.ª *Teorema Fundamental.*
Sección 3.ª *Transformación de Polinomios.*
Sección 4.ª *Fórmula general de un polinomio.*
Sección 5.ª *Fórmulas para los primeros polinomios.*

CAPITULO II

Modelo de la tendencia.
Sección 1.ª *Problemas preliminares.*
Sección 2.ª *Estudio del modelo.*
Sección 3.ª *Axiomática del Modelo.*

CAPITULO III

Distribuciones de probabilidad.
Sección 1.ª *Perturbación aleatoria.*
Sección 2.ª *Modelo estocástico de la tendencia.*

CAPITULO IV

Estimación por punto
Sección 1.ª *Estimaciones de los parámetros del modelo.*
Sección 2.ª *Estimación del modelo y de su varianza.*
Sección 3.ª *Estimadores Lineales óptimos regresores.*

INDICE

Introducción

CAPITULO I

Polinomios ortogonales.
Sección 1.ª *Definiciones*
I. Funciones ortogonales de variables estadísticas.
II. Polinomios ortogonales definidos en un conjunto de números naturales.
III. Polinomios ortogonales centrados.
Sección 2.ª *Teorema Fundamental.*
Sección 3.ª *Transformación de Polinomios.*
Sección 4.ª *Fórmula general de un polinomio.*
Sección 5.ª *Fórmulas para los primeros polinomios.*

CAPITULO II

Modelo de la tendencia.
Sección 1.ª *Problemas preliminares.*
Sección 2.ª *Estudio del modelo.*
Sección 3.ª *Axiomática del Modelo.*

CAPITULO III

Distribuciones de probabilidad.
Sección 1.ª *Perturbación aleatoria.*
Sección 2.ª *Modelo estocástico de la tendencia.*

CAPITULO IV

Estimación por punto
Sección 1.ª *Estimaciones de los parámetros del modelo.*
Sección 2.ª *Estimación del modelo y de su varianza.*
Sección 3.ª *Estimadores Lineales óptimos regresores.*

CAPITULO V

Teoría de la estimación por intervalo

Sección 1.ª *Distribución de los estimadores.*
Sección 2.ª *Estimaciones por intervalo.*
Sección 3.ª *Análisis del grado de la tendencia.*

CAPITULO VI

Predicción de la tendencia

Sección 1.ª *Posibilidad de modificaciones del modelo.*
Sección 2.ª *Estimador ELIO del predictor.*
Sección 3.ª *Predictor de la tendencia.*
Sección 4.ª *Varianzas del predictor.*
Sección 5.ª *Estimación por intervalo para la tendencia Teórica no aleatoria.*
Sección 6.ª *Estimación por intervalo del proceso del modelo estocástico de la tendencia.*

CAPITULO VII

Ejercicio de aplicación.