ENSAYOS ECONOMICOS
Nº 6 – Junio 1978

DISTRIBUCION LINEAL DE SERIES ECONOMICAS (o)

Por Martha Blanco de Dieguez*
y Jorge L. Cortigiani *

I. INTRODUCCION

En un trabajo clásico sobre relaciones lineales entre series económicas de diferente periodicidad, Friedman /7/ puntualizó el concepto estadístico de distribución: dados los valores de una serie anual para T años y los valores trimestrales de una serie relacionada, el problema de distribución consiste en estimar la primera serie para los 4xT trimestres, de modo que las sumas (o promedios) anuales coincidan con el valor anual observado.

Numerosos procedimientos, la mayoría de naturaleza informal, han sido propuestos para resolver este problema, particularmente en el contexto de la estimación de las Cuentas Nacionales Trimestrales a partir de las Cuentas Nacionales Anuales y de un conjunto de indicadores de periodicidad trimestral.

(o) Presentado en el 3º Simposio Nacional de Probabilidade e Estatística, San Pablo, Brasil, Julio de 1978. (*) Gerencia de Investigaciones Económicas.

En este trabajo se presenta un tratamiento general y unificado de la distribución lineal de series, que permite la comparación, en términos de sus propiedades, de los estimadores de uso más frecuente.

En la Sección II se derivā la solución general del problema de distribución; en la Sección III se analizan las soluciones particulares más importantes, y se obtienen expresiones explícitas para el caso circular. Los problemas de estimación se discuten en la sección siguiente. Por último, se presenta un ejemplo en la Sección V, y la Sección VI contiene algunas conclusiones preliminares y futuras líneas de análisis.

II. SOLUCION GENERAL

Sea $\{y_t\}$, $t = 1,2, ..., kT$, una serie de tiempo, donde T representa el número de períodos en estudio y k es un entero que indica la cantidad de subperíodos en que se divide cada uno de los T períodos.

Sólo se dispone de la serie de datos agregados, $y_{t.}$ 1/, donde

$$y_{t.} = \sum_{t=k(t.-1)+1}^{kt.} y_t \qquad , t. = 1,2, ..., T, \tag{1}$$

de modo que se plantea el problema de estimar los valores de la serie $\{y_t\}$.

Se denomina Y al vector de los kT elementos de la serie $\{y_t\}$, Y. al vector de los T elementos de la serie $\{y_{t.}\}$, y C a la matriz de orden TxkT que transforma Y en Y., es decir

$$Y. = CY \tag{2}$$

C es simplemente la matriz que suma, para cada período, las k observaciones de los subperíodos, esto es, los elementos de la fila t. de C son iguales a 1 si $k(t.-1)+1 \leq t \leq kt.$ e iguales a cero, de otra manera.

Además, sea $\{x_{it}\}$, $i = 1,2, ..., p-1$; $t = 1, 2, ..., kT$, un conjunto de series relacionadas linealmente con $\{y_{t.}\}$, y observables para todo i y t, de modo que

$$Y. = CX\beta + CU = X.\beta + U., \tag{3}$$

donde $X = \{x_{it}\}$ es la matriz de datos de orden kTxp, de rango p, β es un vector de p parámetros desconocidos, y U. es un vector aleatorio con media 0 y matriz de varianzas covarianzas V., definida positiva.

El problema de la elección de las series relacionadas no será considerado en este trabajo. Sólo se destaca que pueden ser variables asociadas por razones teóricas o empíricas, en unidades originales o transformaciones de ellas, definidas en el mismo período o en períodos precedentes, y aun posteriores, o una combinación lineal de variables provenientes, por ejemplo, de un análisis de componentes principales. En cualquier caso, las series relacionadas son consideradas como fijas en el modelo lineal (3).

Un supuesto básico, presente en todo el desarrollo de este tema, es que existe una relación lineal entre las variables en los kT subperíodos similar a la relación postulada para los T períodos en que se dispone de la información completa, de modo que

$$Y = X\beta + U, \tag{4}$$

donde U es un vector aleatorio, con media 0 y matriz de varianzas y covarianzas V que satisface

$$CVC' = V. \tag{5}$$

La estimación de β en el modelo 4) a partir de (3) ha sido extensamente considerada en la literatura sobre el tema. Una excelente síntesis puede encontrarse en /10/. Cabe recalcar que, si bien existe una pérdida de eficiencia debida a la agregación, y que incluso el cuadrado del coeficiente de correlación de los datos agrupados puede ser sensiblemente superior al que correspondería a los datos sin agrupar, en el problema de distribución estos inconvenientes son insuperables dada la naturaleza de los datos disponibles.

De todos modos, se desea estimar el vector aleatorio Y a partir de una estimación de β en (3), es decir, a partir de

$$\hat{\beta} = (X.'V.^{-1}X.)^{-1}X.'V.^{-1}Y. \quad . \tag{6}$$

La solución intuitivamente obvia

$$\hat{Y} = X\hat{\beta}, \tag{7}$$

en general no satisface la relación de agregación, es decir,

$$Y. - C\hat{Y} = e. \neq 0 \quad . \tag{8}$$

Este inconveniente, aparentemente formal, resulta particularmente importante en la estimación de las Cuentas Nacionales, donde los valores estimados trimestrales, o aun mensuales, deben coincidir con el valor anual, por lo que es necesario encontrar un estimador, $\tilde{Y}$, tal que $C\tilde{Y} = Y.$.

Teniendo en cuenta el supuesto básico mencionado y las consideraciones expuestas parece razonable escoger $\tilde{Y}$ de modo que las diferencias con $\hat{Y}$ sean mínimas, en algún sentido.

Si se define una función que penalice la distorsión a introducir en $\hat{Y}$, $f(\tilde{Y}, \hat{Y})$, el problema queda expresado en los siguientes términos: elegir $\tilde{Y}$ de modo que minimice la función $f(\tilde{Y}, \hat{Y})$, sujeta a la restricción $C\tilde{Y} = Y.$.

Como es frecuente en este tipo de problemas, se considera la familia de funciones representada por

$$f(\tilde{Y}, \hat{Y}) = (\tilde{Y} - \hat{Y})' A (\tilde{Y} - \hat{Y}), \tag{9}$$

es decir, una forma cuadrática en la diferencia de los dos vectores, donde A es una matriz simétrica, definida positiva 2/, de orden kT, a especificar.

Se puede demostrar que

$$\inf_{C\tilde{Y} = Y.} f(\tilde{Y}, \hat{Y}) = (Y. - C\hat{Y})' (CA^{-1}C')^{-1} (Y. - C\hat{Y}), \tag{10}$$

y que este ínfimo se obtiene en

$$\begin{aligned} \tilde{Y} &= \hat{Y} + A^{-1}C' (CA^{-1} (Y. - C\hat{Y}) \\ &= \hat{Y} + A^{-1}C' (CA^{-1}C')^{-1} e. = \hat{Y} + Be., \end{aligned} \tag{11}$$

es decir que el estimador resultante, $\tilde{Y}$, es igual al estimador inicial, $\hat{Y}$, más una combinación lineal de las discrepancias entre los valores observados y los estimados en los datos agrupados, de modo que la magnitud de la discrepancia entre $\tilde{Y}$ e $\hat{Y}$ dependerá de la magnitud del error e., y de la elección del criterio de penalización.

Además de satisfacer la propiedad de que $C\tilde{Y} = Y.$, el estimador $\tilde{Y}$ es insesgado y la matriz de varianzas y covarianzas de los errores de estimación es

$$\begin{aligned} \text{Cov}\left[\tilde{Y}-Y\right] &= (I-BC)\,\text{Cov}\left[\hat{Y}-Y\right]\,(I-BC)' \\ &= (I-BC)\,(HC-I)\,V(HC-I)'\,(I-BC)', \end{aligned} \tag{12}$$

donde

$$H = X\,(X.'V.^{-1}X.)^{-1}X.'V.^{-1} \;. \tag{13}$$

III. ALGUNAS SOLUCIONES PARTICULARES

Una elección sencilla para la función de penalización es $A = I$, la matriz identidad. En este caso el estimador (11) se reduce a

$$\tilde{Y}_I = \hat{Y} + \frac{1}{k}C'e., \tag{14}$$

es decir, si $\tilde{y}_t$ denota los elementos del vector $\tilde{Y}$, $\hat{y}_t$ los del vector $\hat{Y}$, y $e_{t.}$ los del vector e. ,

$$\tilde{y}_t = \hat{y}_t + \frac{1}{k}e_{t.}\,, \tag{15}$$

lo que implica asignar a cada estimación inicial $\hat{y}_t$, una proporción igual del error en el período correspondiente. Esta distribución del error puede introducir un salto artificial, o discontinuidad, entre el último subperíodo de un período t. y el primer subperíodo de t.+1, si la diferencia entre $e_{t.}$ y $e_{t.+1}$ es grande, alterando arbitrariamente el perfil de la serie de los subperíodos.

En general las series económicas no son estacionarias, por lo que una alternativa relevante se deduce al considerar como criterio característico de la regularidad de una serie el perfil de las primeras diferencias de la misma, lo que induce a definir A en función de estas diferencias, esto es, $A = \Delta'\Delta$, donde

$$\Delta = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ . & . & . & \dots & . & . \\ 0 & 0 & 0 & \dots & -1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}, \tag{16}$$

y el estimador resultante es

$$\tilde{Y}_\Delta = \hat{Y} + (\Delta'\Delta)^{-1}C' \left[C(\Delta'\Delta)^{-1}C'\right]^{-1}e. \quad . \tag{17}$$

Este estimador es de uso muy frecuente en la estimación de las Cuentas Nacionales Trimestrales (véase por ejemplo, /11/), y ha sido propuesto en numerosas oportunidades y obtenido por diferentes derivaciones (/1/,/4/,/7/,/9/,/10/,/11/ y/14/). De acuerdo con su definición requiere la inversión de dos matrices, una de orden kTxkT y la otra de TxT, por lo que en el Apéndice se consideran expresiones que facilitan su cálculo.

En estos casos, la elección de la matriz A no depende de V. ¿Existe alguna elección de A para la cual (11) es mejor, en el sentido de menor varianza?.

Si V es conocida, excepto quizás por un factor de escala, resulta que la elección de A que minimiza la traza de la matriz $\mathrm{Cov}\left(\tilde{Y}-Y\right)$ es $A = V^{-1}$, y el estimador resultante

$$\tilde{Y}^* = \hat{Y} + VC' V.^{-1}e., \tag{18}$$

es el mejor lineal insesgado.

El primer componente de (18) consiste en la aplicación de los coeficientes de la regresión de los datos agregados a las variables explicativas desagregadas. El segundo es una estimación de los residuos de la regresión de los datos sin agregar (esto es, una estimación del vector aleatorio U), obtenido por aplicación de la matriz $VC'V.^{-1}$, de orden kTxT, a la estimación de los residuos de la regresión de los datos agregados (esto es, a la estimación del vector aleatorio U.).

Para (18), la matriz de varianzas y covarianzas de los errores de estimación, obtenida reemplazando A^{-1} por V en (12) es:

$$\mathrm{Cov}\left[\tilde{Y}^{*}-Y\right]= V-VC'V.^{-1}CV + (X-VC'V.^{-1}X.)(X.'V.^{-1}X.)^{-1}(X'-X.'V.^{-1}CV) \tag{19}$$

Este mismo resultado ha sido obtenido por Chou y Lin/5/, por aplicación directa del método propuesto por Goldberger/9/ en predicción en mínimos cuadrados generalizados. Las diferencias con el enfoque presentado en este trabajo, además del procedimiento de derivación y algunas expresiones explícitas, radican en el análisis del problema de estimación, que se discute en la sección siguiente.

Con el propósito de comparar la perfomance de $\tilde{Y}^{*}$ con respecto a $\hat{Y}$, se obtiene una expresión para la matriz de varianzas y covarianzas de los errores de estimación correspondiente a $\hat{Y}$,

$$\mathrm{Cov}\left[\hat{Y}-Y\right] = V + X(X.'V.^{-1}X.)^{-1}X' - X(X.'V.^{-1}X.)^{-1}X'C'V.^{-1}CV - VC'V.^{-1}CX(X.'V.^{-1}X.)^{-1}X \tag{20}$$

De (19) y (20) surge que

$$\mathrm{Cov}\left[\hat{Y}-Y\right] = \mathrm{Cov}\left[\tilde{Y}^{*}-Y\right] + VC'\left[V.^{-1}-V.^{-1}X.(X.'V.^{-1}X.)^{-1}X.V.^{-1}\right]CV$$

$$= \mathrm{Cov}\left[\tilde{Y}^{*}-Y\right] + D, \tag{21}$$

donde D es una matriz definida no negativa. Si se considera una función escalar de Cov $[\hat{Y} - Y]$, por ejemplo la traza de la matriz, se sigue que tr D puede interpretarse como la ganancia en eficiencia asociada con el uso del estimador $\tilde{Y}^*$ en lugar de $\hat{Y}$.

Alternativaménte, se puede reemplazar la diferencia de las trazas por el cociente de las mismas, esto es,

$$e(\hat{Y}) = \frac{\text{traza Cov}\left[\tilde{Y}^* - Y\right]}{\text{traza Cov}\left[\hat{Y} - Y\right]} \quad . \tag{22}$$

En ambos casos se deduce que $\tilde{Y}^*$ puede considerarse como una mejora con respecto a $\hat{Y}$.

A continuación y con el propósito de interpretar estos resultados, se considera el caso en que $V = \sigma^2 I$, I la matriz identidad de orden kTxkT. En este caso, $\hat{\beta} = (X.'X.)^{-1} X'.Y.$, es decir que $\hat{\beta}$ es el estimador por mínimos cuadrados ordinarios para la regresión lineal múltiple de los datos agregados. En consecuencia,

$$\tilde{Y}^* = (I - \frac{1}{k} C'C) X\hat{\beta} + \frac{1}{k} C'Y., \tag{23}$$

ó sea que cada elemento del vector $\tilde{Y}^*$ se puede expresar como:

$$\tilde{y}_t^* = \frac{y_{t.}}{k} + \sum_{i=1}^{p-1} \hat{\beta}_i (x_{it} - \frac{x_{it.}}{k}), \tag{24}$$

que para el caso de una sola variable se reduce a:

$$\tilde{y}_t^* = \frac{y_{t.}}{k} + \hat{\beta}(x_t - \frac{x_{t.}}{k}) \quad , \tag{25}$$

ó

$$\frac{\tilde{y}_t^* - (y_{t.}/k)}{x_t - (x_{t.}/k)} = \hat{\beta} \tag{26}$$

Esto es, el estimador para los subperíodos es tal que su diferencia con el promedio del período a que pertenece es proporcional a la diferencia correspondiente de la variable independiente, y la constante de proporcionalidad es precisamente $\hat{\beta}$.

Además,

$$\mathrm{Var}\,(\tilde{y}_t^* - y_t) = \frac{\sigma^2}{kT\sigma^2_{x.}}\left[\sigma^2_{x.}(k-1)\,T + (kx_t - x_{t.})^2\right] \quad , \qquad (27)$$

para $t = 1, 2, ..., kT$, donde

$$\sigma^2_{x.} = \frac{1}{T}\sum_{T=1}^{T} (x_{t.} - \bar{x}.)^2, \quad \bar{x}_. = \frac{1}{T}\sum_{t.=1}^{T} x_{t.}\,, \qquad y$$

$\sigma^2_{x.}$ se puede interpretar como la variabilidad "entre los períodos".

Si $\sigma^2 x = \frac{1}{kT}\sum_{t=1}^{kT} (k\,x_t - x_{t.})^2$, que se puede interpretar como la variabilidad "dentro de los períodos", a partir de (27) se encuentra que

$$\mathrm{traza\ Cov}\left[\tilde{Y}^* - Y\right] = \frac{\sigma^2}{\sigma^2_{x.}}\left[(k-1)T\sigma^2_{x.} + \sigma^2_{x}\right], \qquad (28)$$

$$\mathrm{traza\ D} = \sigma^2\,(T-2), \qquad (29)$$

y

$$e\,(\hat{Y}) = \frac{(k-1)T\,\sigma^2_{x.} + \sigma^2_{x}}{(kT-2)\,\sigma^2_{x.} + \sigma^2_{x}} \quad . \qquad (30)$$

Se observa que $e\,(\hat{Y}) < 1$ si $T > 2$. Además, si $\sigma^2_{x} = \sigma^2_{x.}$,

$$e\,(\hat{Y}) = 1 - \frac{T-2}{kT-1} \qquad (31)$$

Para T grande e $(\hat{y}) \approx 1 - (1 / k)$. Por otra parte, si la variabilidad "dentro" de los períodos es menor que la variabilidad "entre" los períodos e $(\hat{Y}) < 1 - (T - 2) / (kT-1)$, y en el caso contrario, e $(\hat{Y}) > 1 - (T - 2) / (kT-1)$ aunque e $(\hat{Y})$ es poco sensible a cambios en la relación entre σ_x^2 y $\sigma_{x.}^2$

Estas consideraciones parecen indicar una ventaja importante de $\tilde{Y}^*$ con respecto a $\hat{Y}$, aun en el caso circular. Sin embargo, es importante observar cómo distribuye $\tilde{Y}^*$ el error de los períodos entre los subperíodos. A partir de (18) se puede escribir

$$\tilde{Y}^* = \hat{Y} + \frac{1}{k} C'e. \quad , \tag{32}$$

el cual es idéntico a (15), y ya se señalaron los inconvenientes de este reparto arbitrario del error, que destruye el perfil de $\hat{Y}$.

No obstante, la ganancia en eficiencia puede ser importante, si k es chico, y parece razonable conjeturar que no será menor si $V \neq \sigma^2 I$.

IV. ESTIMACION

El estimador (18) requiere el conocimiento de V, la matriz de varianzas y covarianzas del vector aleatorio U. En la práctica V es desconocida y debe ser estimada. En este sentido, el problema parece simplemente un caso particular del resuelto por Goldberger /9/. Sin embargo, en distribución solo se dispone de estimaciones de U., a partir de las cuales debe inferirse la estructura de varianzas y covarianzas de U, a diferencia del caso de mínimos cuadrados generalizados en que se dispone de una estimación directa de U. Es conveniente enfatizar esta situación diferencial, por lo que, antes de considerar la estimación de U, y por ende de V, se analiza la relación entre V y V., utilizando el concepto de inversa generalizada de matrices rectangulares. Si A es una matriz de orden mxn, existe una única matriz A^- tal que 1) AA^- es simétrica, 2) A^-A es simétrica, 3) $AA^-A=A$, y 4) $A^-AA^- = A^-$.

Recordando que $V_{.} = CVC'$ se observa que $C^{-} = \frac{1}{k} C'$, y que la ecuación matricial que relaciona V con $V_{.}$ es consistente, y por lo tanto, la solución general es

$$V = C^{-} V_{.} C'^{-} + G - C^{-} C G C' C'^{-}$$

$$= \frac{1}{k^2} C' V_{.} C + G - \frac{1}{k^2} C' C G C' C \quad , \qquad (33)$$

donde G es cualquier matriz del mismo orden de V

Algunas soluciones para V serán no singulares, y otras, singulares. Por ejemplo, si $G = 0$, $V = \frac{1}{k^2} C' V_{.} C$, que en el caso $V_{.} = \sigma^2 I$ se reduce a $V = \frac{\sigma^2}{k^2} I \otimes 1$, donde 1 es una matriz de orden $k \times k$ cuyos elementos son todos unos.

En cambio, si G es no singular y tal que $CGC' = V_{.}$, $V = G$ y existen infinitas matrices no singulares que satisfacen este supuesto de agregación.

Se sigue que la denominación de mejor lineal e insesgado para el estimador (18) no es apropiada, cuando no se conoce V, cualquiera sea la elección de G(Ver/5/).

No obstante esta indeterminación, es posible obtener estimaciones óptimas si se adopta el supuesto adicional de que las series de tiempo $\{U_t\}$ y $\{U_{t.}\}$ son generadas por la familia de modelos autoregresivos - promedios móviles (ARMA) de orden (p, q), estacionarios e invertibles.

Este supuesto provee de un conjunto flexible de modelos al mismo tiempo que brinda un marco de referencia general, sin restringir las posibilidades de aplicación, de manera que es factible, a partir de la observación, identificación y estimación de la serie $\{U_{t.}\}$, identificar y estimar consistentemente el miembro de la familia (ARMA) (p, q) correspondiente a la serie inobservable $\{U_t\}$.

Este procedimiento de estimación, descripto en /17/ no solo requiere muestras grandes, ya que cuando el número de observaciones disponibles es pequeño, se dificulta la etapa de identificación y los parámetros del modelo agregado no pueden ser estimados con precisión, sino que también implica la utilización de métodos de estimación de considerable complejidad computacional.

V. EJEMPLO

En esta sección se presenta un ejemplo simple, fundamentalmente con fines de comparación de los estimadores definidos en (17) y (23).

Se denomina Indice de Precios Implícitos en el Producto Bruto Interno, a costos de factores, a la relación que se obtiene al considerar las estimaciones anuales del producto a precios corrientes y a precios constantes/16/. La necesidad de contar con índices similares para las cuentas trimestrales plantea el problema de estimar el Indice de Precios Implícitos Trimestral.

En el Cuadro I se muestran las siguientes series: Indice de Precios Implícitos Anual, Indice de Precios al por mayor no agropecuarios nacional trimestral 3/, que se utiliza como serie relacionada, ambas para el período 1950-73, y las estimaciones trimestrales por (17) y (23). Al final del cuadro aparece la regresión anual en que se basa la distribución. Como los índices anuales no son sumas sino promedios de los índices trimestrales, se introducen las modificaciones necesarias en las definiciones de las matrices C y X.

En este ejemplo, $\sigma^2_{x.} = 337{,}25$ y $\sigma_x^2 = 3{,}56$, de modo que $e(\hat{Y}) = 0{,}766$.

Se observa que las diferencias entre $\tilde{Y}$ e $\tilde{Y}^*$ son mínimas, debido a la magnitud del coeficiente de correlación entre las variables, por lo que el tamaño de los errores a distribuir es muy pequeño. No obstante, la forma diferencial de distribuir el error aparece claramente ejemplificada en el Gráfico 1.

VI. CONCLUSIONES

A pesar de la aplicación frecuente de algunos métodos de distribución lineal mediante series relacionadas, particularmente en el contexto de la estimación de las Cuentas Nacionales de periodicidad menor que la anual, no existe una evaluación coincidente de los procedimientos de estimación ni de la magnitud de la información incluida en los residuos de los datos agregados. Algunos autores sostienen que toda la información está contenida en la estimación inicial, y por ende solo se debe considerar una distribución matemática del error, basada en criterios que tratan de evitar alterar el perfil natural de las series económicas. Otros sostienen, en cambio, la existencia de mejores estimadores lineales insesgados, utilizando precisamente la información de los residuos.

La conclusión de este trabajo es que, si bien la información no está irremediablemente perdida por la agregación, solo es recuperable parcialmente cuando se dispone de muestras grandes y los datos agregados poseen una estructura de correlación interna.

En el caso circular, aun para muestras grandes, el "mejor estimador lineal insesgado" puede producir estimaciones que alteren arbitraria e inconvenientemente el perfil de las estimaciones iniciales.

Si bien no se cuenta con resultados sobre la eficiencia del método de distribución matemática del error basado en el criterio de minimización de las primeras diferencias, parece una alternativa razonable para muestras chicas, y posee la ventaja de su aplicación automática en el caso en que se deba distribuir periódicamente un número considerable de series.

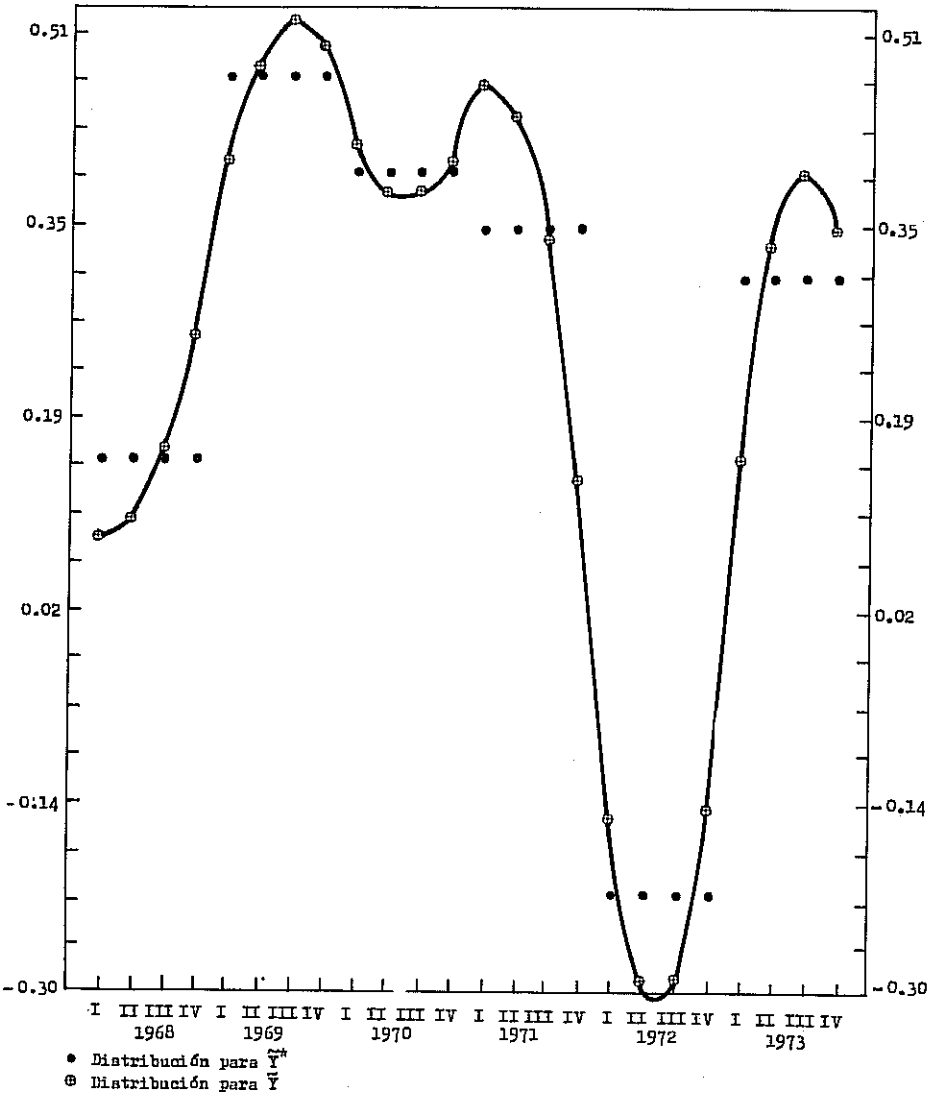
CUADRO I

Años y trimestres	Indice de precios al por mayor no agropecuarios nacional	Indice de precios implícitos en PBI a costo de factores	Estimación del índice de precios implícitos		Años y trimestres	Indice de precios al pormayor no agropecuarios nacional	Indice de precios implícitos en PBI a costo de factores	Estimación del índice de precios implícitos	
			$\bar{Y}$	$\bar{Y}*$				$\bar{Y}$	$\bar{Y}*$
1950	0,0795	0.0920			1962	1.4040	1.4100		
I	0.0738		0.0864	0.0861	I	1.2040		1.2011	1.2024
II	0.0762		0.0887	0.0886	II	1.3573		1.3609	1.3616
III	0.0802		0.0927	0.0927	III	1.4873		1.4968	1.4966
IV	0.0878		0.1002	0.1006	IV	1.5660		1.5798	1.5783
1951	0.1147	0.1240			1963	1.7730	1.8080		
I	0.0970		0.1093	0.1057	I	1.6570		1.6761	1.6876
II	0.1052		0.1162	0.1142	II	1.7280		1.7561	1.7613
III	0.1237		0.1326	0.1334	III	1.8000		1.8387	1.8360
IV	0.1327		0.1380	0.1428	IV	1.9133		1.9688	1.9537
1952	0.1587	0.1560			1964	2.2330	2.3050		
I	0.1462		0.1468	0.1420	I	2.0557		2.1325	2.1209
II	0.1558		0.1528	0.1520	II	2.1960		2.2799	2.2666
III	0,1668		0.1603	0.1623	III	2.2863		2.3613	2.3604
IV	0.1670		0.1601	0.1636	IV	2.3980		2.4506	2.4763
1953	0.1739	0.1680			1965	2.9150	2.9350		
I	0.1723		0.1652	0.1664	I	2.5850		2.6039	2.5924
II	0.1740		0.1673	0.1682	II	2.8140		2.8207	2.8302
III	0.1739		0.1682	0.1680	III	3.0923		3.1087	3.1191
IV	0.1752		0.1714	0.1694	IV	3.1640		3.2019	3.1935
1954	0.1812	0.1810			1966	3.4820	3.5990		
I	0.1740		0.1728	0.1736	I	3.2557		3.3359	3.3640
II	0.1768		0.1770	0.1765	II	3.3917		3.5039	3.5052
III	0.1852		0.1859	0.1852	III	3.5530		3.6861	3.6727
IV	0.1888		0.1886	0.1890	IV	3.7273		3.8698	3.8537
1955	0.1993	0.2000			1967	4.3400	4.4770		
I	0.1937		0.1915	0.1943	I	3.9820		4.1249	4.1054
II	0.1970		0.1952	0.1978	II	4.2677		4.4099	4.4019
III	0.2001		0.2010	0.2010	III	4.4823		4.6190	4.6248
IV	0.2060		0.2123	0.2071	IV	4.6413		4.7681	4.7898
1956	0.2342	0.2510			1968	4.7570	4.9100		
I	0.2223		0.2367	0.2387	I	4.7083		4.8194	4.8595
II	0.2313		0.2499	0.2480	II	4.7120		4.8319	4.8633
III	0.2387		0.2576	0.2557	III	4.7980		4.9569	4.9526
IV	0.2443		0.2598	0.2615	IV	4.8100		5.0321	4.9650
1957	0.2920	0.3050			1969	4.9920	5.3590		
I	0.2597		0.2683	0.2714	I	4.8613		5.1751	5.2234
II	0.2827		0.2905	0.2953	II	4.9413		5.3133	5.3064
III	0.3090		0.3218	0.3227	III	5.0523		5.4493	5.4216
IV	0.3157		0.3384	0.3296	IV	5.1137		5.4992	5.4853
1958	0.3810	0.4170			1970	5.6520	5.9930		
I	0.3187		0.3570	0.3523	I	5.2300		5.5717	5.5549
II	0.3520		0.3958	0.3869	II	5.4243		5.7474	5.7567
III	0.4027		0.4416	0.4395	III	5.7150		6.0455	6.0584
IV	0.4513		0.4742	0.4900	IV	6.2347		6.6036	6.5979
1959	0.8534	0.8430			1971	7.7290	8.1190		
I	0.6907		0.6936	0.6741	I	6.7797		7.2102	7.1335
II	0.8183		0.8067	0.8066	II	7.2713		7.7063	7.6439
III	0.9380		0.9211	0.9308	III	8.1450		8.5437	8.5509
IV	0.9770		0.9615	0.9713	IV	8.7213		9.0173	9.1492
1960	1.0000	1.0000			1972	12.9210	13.1540		
I	1.0000		0.9949	1.0000	I	10.9003		11.0994	11.0563
II	0.9950		0.9955	0.9948	II	12.4190		12.5868	12.6329
III	1.0003		1.0033	1.0004	III	13.6113		13.8260	13.8707
IV	1.0053		1.0071	1.0055	IV	14.7527		15.1031	15.0555
1961	1.1010	1.0970			1973	19.8080	20.6310		
I	1.0353		1.0331	1.0288	I	17.1913		17.8176	17.9146
II	1.0700		1.0654	1.0648	II	20.3507		21.2093	21.1944
III	1.1253		1.1205	1.1223	III	20.8063		21.7234	21.6674
IV	1.1733		1.1691	1.1721	IV	20.8857		21.7759	21.7498

Regresión con datos anuales $\hat{PI}_t = -0.007064 + 1,038134\ PM_t$
$(-0,35) \quad (285,71)$

$R^2 = 0,99973$ DW = 2,02 $\frac{\text{Error Estandar de Estimación}}{\text{Media } PI_t} \times 100 = 2,46$

GRAFICO 1



0.51
0.35
0.19
0.02
-0.14
-0.30
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
1968 1969 1970 1971 1972 1973
● Distribución para $\tilde{\bar{Y}}^*$
⊕ Distribución para $\tilde{\bar{Y}}$

APENDICE

Si $(\Delta'\Delta)^{-1} = \left[\delta_{ij}\right]$ y $C(\Delta'\Delta)^{-1}C' = \left[c_{ij}\right]$, entonces

$$\delta_{ij} = \begin{cases} kT+1-i \ , \ \text{si} \quad i \geqslant j \\ kT+1-j \ , \ \text{si} \quad i \leqslant j \ , \end{cases}$$

$$c_{ij} = \begin{cases} \dfrac{k(k+1)(2k+1)}{6} + (T-i)k^3 \ , \ \text{si} \quad i = j \\ \dfrac{k^2(k+1)}{2} + (T-i)k^3 \ , \quad \text{si} \quad i > j \\ \dfrac{k^2(k+1)}{2} + (T-j)\,k^3 \ , \ \text{si} \ i < j \ . \end{cases}$$

Esta última matriz se puede escribir como el producto de matrices triangulares superiores.

$$C(\Delta'\Delta)^{-1}C' = S'_u \ S_u \ ,$$

$$\left[C(\Delta'\Delta)^{-1}C'\right]^{-1} = S_u^{-1} \ (S_u^{-1})' \ ,$$

y

$$(\Delta'\Delta)^{-1} \ C'\left[C(\Delta'\Delta)^{-1}C'\right]^{-1} = \left[\frac{k(k+1)(2k+1)}{6} + (T-1)\,k^3\right] PS_u^{-1}(S_u^{-1})',$$

donde $P = \left[p_{ij}\right]$, $S_u = \left[s_{ij}\right]$, $S_u^{-1} \quad , \left[r_{ij}\right]$,

$$p_{ij} = \begin{cases} \displaystyle\sum_{h-i-k(j-1)}^{k} h + k^2\,(T-j), \ \text{si} \ k\,(j-1) < i \leqslant kj \ , \\ k^2\,(T-j)-k(i-k_j-1), \ \text{si} \ \ kj < i \ , \\ \dfrac{k(k+1)}{2} + k^2 \ (T-j), \text{si} \ \ i \leqslant k \ \ (j-1), \ \ j > 1 \quad ; \end{cases}$$

$$
s_{ij} = \begin{cases}
s_{11} = \dfrac{k(k+1)(2k+1)}{6} + (T-1)k^3 & , \text{si } i = j = 1 \ , \\[2ex]
s_{1j} = \dfrac{k^2(k+1)}{2} + (T-j)k^3 & , \text{si } i = 1 ; J = 2, \ldots, T, \\[2ex]
s_{ii} = \sqrt{s_{11}\left(s_{1i} - \dfrac{k(k+1)}{2}\right) - \displaystyle\sum_{h=1}^{i=1} s_{hi}^{\,2}} & , \text{si } i=j=2,3, \ldots, T \\[2ex]
s_{ij} = \dfrac{1}{s_{ii}}\left(s_{11} \ s_{1j} - \displaystyle\sum_{h=1}^{i=1} s_{hi} \ s_{hj}\right) & , \text{si } i < j \ , \\[2ex]
s_{ij} = 0 & , \text{si } i > j \ ;
\end{cases}
$$

$$
r_{ij} = \begin{cases}
\dfrac{1}{s_{ij}} & , \text{si } i = j = 1,2, \ldots, T \ , \\[2ex]
= \dfrac{1}{s_{ii}} \displaystyle\sum_{h=i+1}^{j} s_{ih} \ s_{hj} & , \text{si } i < j \ , \\[2ex]
0 & , \text{si } i > j \ .
\end{cases}
$$

Se dispone de un programa en FORTRAN IV basado en este algoritmo, elaborado por B.E.L. de La Valle, del Banco Central de la República Argentina.

1/ Alternativamente, $y_{t.} = \frac{y_{t.}}{k}$. No se considerará explícitamente este caso por su similitud con el de agregación.
2/ Se pueden obtener resultados similares utilizando matrices definidas no negativas e inversas generalizadas, pero las expresiones resultantes son complejas y extensas, por lo que se prefiere limitar este trabajo al caso de matrices definidas positivas.
3/ La serie relacionada no presenta estacionalidad significativa.

Referencias Bibliográficas

/1/ Anderson, O. D. Time series analysis and forecasting. Butterworths, 1976.

/2/ Boot, C. G., Feibes, W. and Lisman, J.H.C. Further comments on the derivation of quarterly figures from annual data. Applied Statistics. (London) 16: 65-75, 1967.

/3/ Box, G.E.P. and Jenkins, G.M. Time series analysis, forecasting and control. Holden-Day, 1970.

/4/ Chang, G. and Liu, T. Monthly estimates of certain national product components, 1946-49. Review of Economics and Statistics. (Cambridge) XXXIII: 219-227, August 1951.

/5/ Chow, G.C., and Lin, A. Best linear unbiased interpolation, distribution and extrapolation of time series by related series. Review of Economics and Statistics (Cambridge) LIII: 372-375, November 1971.

/6/ Denton, F.T. Adjustment of monthly or quarterly series to annuals totals; an approach based on Quadratic minimization. JASA (Washington) 66, Nº 333: 99-102, March 1971.

/7/ Friedman, M. The interpolation of time series by related series. JASA (Washington) 57: 729-257, December 1962.

/8/ Ginsburgh, V.A. A further note on the derivation of quarterly figures consistent with annual data. Applied Statistics (London) 22: 368-373, 1973.

/9/ Goldberger, A.S. Best linear unbiased prediction in the generalized linear regression model. JASA (Washington) 57: 369-375, June 1962.

/10/ Haitovsky, Y. Regression estimation from grouped observations. Griffin, 1973.

/11/ Laroque, G., Le Galvez, B. et Nasse, P. Comptes trimestriels; méthodes statistiques et séries rétrospectives. INSEE (París) Serie C Nº 40, Nº 183: 9-12, Décembre 1975.

/12/ Lisman, J.H.C. and Sandee, J. Derivation of quarterly figures from annual data. Applied Statistics (London) 13: 87-90, 1964.

/13/ Liu, T. A monthly recursive econometric model of United States; a test of feasibility. Review of Economics and Statistics (Cambridge) LI: 1-13, February 1969.

/14/ Nasse, P. Peut-on suivre l'evolution trimestrielle de la consommation. Economice et Statistique (París) 20-22, 1970.

/15/ Vangrevelinghe, G. L'évolution à court terme de la consommation des ménages. INSEE. Etudes et conjoncture (paris) 9: 54-102, 1966.

/16/ Sistemas de cuentas del producto e ingreso de la Argentina. BCRA, 1975.

/17/ Cortigiani, J. Extrapolación y eficiencia en la estimación de las cuentas nacionales. BCRA, Gerencia de Investigaciones Económicas, 1978.