Transformación de los procesos markovianos aplicados a la econometría

por
FRANCISCO JAVIER URBELZ IBARROLA

I INTRODUCCIÓN

1.1. En el trabajo de Laurence J. LAU (1) titulado "Funcions of Forms in Econometric model Bulding" trata de las relaciones funcionales lineales de forma que la variable depende y pueda expresarse en términos como la siguiente igualdad:

$$y_i = \Sigma\alpha_h f_h(x_{hi}) \qquad (1.1.1)$$

o aun con mas generalidad

$$g(y_i) = \Sigma\alpha_h f_h(x_i) \qquad (1.1.2)$$

1.2 Casos particulares de la expresión anterior son cuando

$$f_h(x_{hi}) = x_{hi} \qquad (1.2.1)$$

También otro caso particularmente importante es cuando hacemos

$$f_h(x_{hi}) = I_n(x_{hi}) \qquad (1.2.2)$$

Hemos modificado para generalizar las transformaciones de J. Lau incluyendo los subíndices de las variables explicativas.

Tenemos numerosos ejemplos de transformaciones logarítmicas que especialmente se estudian en Econometría.

Citemos expresamente el modelo económico de Cobb y Douglas (2) y que puede expresarse como una función del capital y del trabajo:

$$y = AK^{\alpha}L^{\beta} \tag{1.2.3}$$

donde A, α, β son parámetros y K y L son las variables capital y trabajo. Generalmente α, β están sujetas a las restricción lineal $\alpha+\beta=1$ porque así es una función homogénea de primer grado los parámetros expresan las correspondientes elasticidades.

Pero la expresión anterior ha sido minuciosamente estudiada por Arraw, Minhas y Soley (3), en 1961 propuso otra función más compleja y que puede expresarse así

$$y = \lambda\,[(1-\delta)K^{\rho}+\delta L^{\rho}]^{1/\rho} \tag{1.2.4}$$

pero la expresión anterior no es linealizable

Tukey (4) fue el primero que realizó un estudio detallado de las Transformaciones para que las variables transformadas cumplieran los objetivos de linealizar prácticamente el modelo y para que los errores fueren más homocedásticos.

Es importante estudiar el problema fundamental para que la variable transformada persiga los siguientes objetivos:

a) La Transformación sea una variable linealizada.

b) Los errores sean homoscedásticos, o al menos reducir al mínimo la posible heteroscedasticidad; y

c) La variable transformada siga una distribución prácticamente normal.

Aun cuando no todas las variables transformadas siguen la distribución normal, recordemos el Teorema Central del Límite para todas las variables que sigan la misma distribución y no haya ningún cambio de estructura en el tiempo.

No todas las variables transformadas siguen una distribución normal sino que depende de la estructura del fenómeno representativo y si el Teorema Central del Límite nos indica que la suma de variables que sigan la misma distribución, podríamos encontrar alguna variable transformada que siga la ley normal.

Este trabajo la varianza depende de la diferencia de tiempos porque impondremos a los procesos markovianos siguen las leyes de Wiener.

Pero aun cuando no es la varianza constante si tiene una estructura sencilla y fácilmente de analizar y se asemeja a la homocedasticidad cuando el conjunto paramétrico es de números enteros.

1.3 Una transformación es buena cuando consigue estos objetivos simultáneamente

Tukey estudia la gama de transformaciones dependiente de dos parámetros:

$$(X+\alpha)^{\lambda} \tag{1.3.1}$$

y mediante esta transformación pretende linealizar el modelo transformado.

Box y Tidwell (5) describen un método basado en el desarrollo de Taylor y Brown, Caves y Christensen (6) introdujeron un modelo aplicado a los servicios ferroviarios basados en las transformaciones logarítmicas.

Posteriormente, en 1963, Box y Cox (7) de las variables dependientes e independientes efectúan transformaciones que como casos particulares, están las transformaciones logarítmicas.

Como nuestro estudio se basará fundamentalmente en este tipo de transformaciones (que aplicaremos a variables de tipo económico), la transformación de Box-Cox es la siguiente:

$$y(\lambda) = \begin{cases} (y^{\lambda}-1)/\lambda & \text{para } \lambda \neq 0 \\ L_n y & \text{para } \lambda = 0 \end{cases} \tag{1.3.2}$$

donde la transformación de la variable $y(\lambda)$ depende del parámetro λ y que impondremos las condiciones exigidas por Tukey para $y(\lambda)$ especialmente que siga una distribución aproximadamente normal. Es decir

$$y(\lambda) \Rightarrow N(\beta x, \sigma) \tag{1.3.3}$$

donde $y(\lambda)$ es un valor concreto para un valor dado de x.

Esta variable transformada podría también ampliarse a más de una variable; es decir a una regresión de k dimensiones y, en este caso, la distribución del vector

$$y(\lambda) = N(x\beta, \sigma^2 I) \tag{1.3.4}$$

En cursos normales de Econometría la regresión múltiple se estudia con detenimiento prescindiendo del parámetro λ.

En este **Trabajo de Investigación** no haremos, en principio, aplicación generalizada de la regresión múltiple lo que perfeccionaría indiscutiblemente la finalidad concreta de las buenas transformaciones expuestas por Tukey. Pero sí haremos brevemente algunas transformaciones de Tukey aunque por brevedad prescindiremos deL parámetro c.

Estudiaremos especialmente esta transformación como si fuera un proceso estocástico siguiendo, en algunos de los conceptos a Barttlet (8), Blanc-Lapierre & Fortet (9), Paul Levi (10), Kemeny and Snell (11) Gokman-Skorokhod (12), Doob (13) referente a las leyes de probabilidad de las distribuciones de los procesos estocásticos y muy especialmente de las cadenas de Markov.

II CONCEPTOS FUNDAMENTALES

2.1 Es preciso conocer perfectamente el concepto de ***variable aleatoria*** para poder entender el de función aleatoria o mejor dicho, proceso estocástico aunque remitimos al lector a la bibliografía para ampliar estos conceptos.

Una variable aleatoria representativa de un fenómeno de naturaleza real viene asociada con una función de probabilidad para cada valor particular de la misma. Si la variable es discreta tenemos la denominada función de cuantía pero si la variable es continua está asociada con una función de densidad. Las relaciones de las funciones de distribución son bien conocidas.

2.2 Al *cambiar o transformar una variable aleatoria, transformamos igualmente su ley de probabilidad* y existe una relación entre ambas, así como las leyes de distribución. Es muy importante para este estudio deducir deun tipo sencillo de ley de probabilidad, por ejemplo del ***tipo normal*** a otro más complejo. Este problema lo utilizaremos para determinar el intervalo o nivel de confianza de la variable primitiva **conociendo por hipótesis la ley de probabilidad de la variable transformada.**

2.3 Pero las sencillas y elementales fórmulas de la Ley de distribución normal, podemos de la distribución unidimensional pasar a la multidimensional y a la función característica de la distribución normal. Este hecho, cuando tenemos una suma de variables aleatorias independientemente distribuidas nos conduce a una distribución normal y en consecuencia aplicar esta distribución a la variable aleatoria suma.

Recordemos que la suma de un número cualquiera de términos independientes (finito o infinito) converge a una distribución normal.

Es importante señalar aquí el Teorema de H. Cramer que dice que dadas dos variables aleatorias independientes y si su suma difiere muy poco de la normal, cada una de ellas, después de la adición de una constante elegida convenientemente, la distribución difiere muy poco de la normal.

Este hecho nos revela que si efectuamos transformaciones de variables aleatorias podemos encontrar una transformación apropiada cuya distribución sea prácticamente la normal.

El Teorema de Liapounoff se ocupa de variables no acotadas. Los dos primeros momentos pueden ser infinitos y no deben intervenir en una Teoría General y este Teorema reducido a su simple expresión nos dice que para que la suma de variables dependa de una ley generalizada (muy poco diferente a la normal), es suficiente que el mas grande de sus términos sea despreciable.

Para que la ley de la suma de variables aleatorias sean de un tipo generalizado aproximadamente normal, es necesario y suficiente, que cada término no individualmente despreciable sea de un tipo muy poco diferente a la distribución normal: y el mayor de los términos individualmente despreciable sea él mismo despreciable.

2.4 El concepto de ***proceso aleatorio***, se describe como una función aleatoria que depende de un parámetro. A este parámetro lo designaremos por t. Para $t=t_i$, $X(t_i)$ es una variable aleatoria.

Dado el conjunto **T** se denomina ***Conjunto índice*** de parámetro discreto cuando **T**$=\{t:t=-\infty,...-3,-2,-1,0,+1,+2,...\infty\}$ y es parámetro continuo cuando **T**$=\{t-\infty<t<\infty\}$ donde t es el parámetro de la variable aleatoria.

2.5 Se denomina **proceso estocástico** más expresamente se define como el conjunto vectorial aleatorio que varía con los valores de los parámetros definidos en el conjunto índice. Una muestra es la representación de una trayectoria de un proceso estocástico. Al conjunto infinito de trayectorias muestrales se denomina también proceso aleatorio. En consecuencia: proceso aleatorio significa una familia de variables aleatorias dependiente del parámetro (en nuestro caso el tiempo t).

2.6 No definiremos la distribución conjunta –n– dimensional de n variables aleatorias pertenecientes al conjunto índice, así como tampoco las condiciones de compatibilidad y consistencia de Kolmogorov de las leyes de probabilidad de un vector x de n dimensiones.

Indicaremos no obstante el **concepto de estacionariedad** de un proceso. Si a un vector **x(t)** el conjunto paramétrico lo desplazamos en el tiempo de forma que la componente t_i para todo i a cada componente del conjunto paramétrico le sumamos la misma cantidad h, el nuevo vector paramétrico si tiene la misma distribución de probabilidad se dice es **estacionario.**

2.7 Como utilizaremos el proceso estocástico que sigue una distribución normal, fijaremos nuestra atención en la definición de familia de variables aleatorias que siguen la distribución normal.

Recordemos que dado el vector **X** = (x1, x2, x3, ... xn) es normal cuando su función característica la podemos escribir

$$\varphi(t_1, t_2, ...t_n) = \exp(i\, t'\mu - 1/2t'\Sigma^{-1}t) \qquad (2.6.1)$$

siendo, como sabemos μ el vector de las medias y Σ la matriz de varianzas-covarianzas.

2.8 El ***movimiento Browniano*** goza un papel importante en nuestro trabajo y por eso solamente indicaremos algunas de sus propiedades que aplicaremos dejando para que el lector pueda complementar estos conocimientos con la bibliografía citada u otra análoga:

a) es un proceso de *incrementos independientes*.

b) el proceso es de *incrementos estacionarios*, y la matriz de covarianzas

$$B(t,s)=B(t-s) \qquad (2.7.1)$$

depende tan solo de la diferencia de tiempos en los instantes t y s.

c) la media

$$m\{\xi(s)-\xi(t)\} = m(s-t) \qquad (2.7.2)$$

igualmente depende de la diferencia de tiempos.

d) la varianza

$$B\{\xi(s)-\xi(t)\} = \sigma^2(s-t) \qquad (2.7.3)$$

No demostraremos estaspropiedades por considerarlas conocidas y pueden verse en cualquier libro sobre procesos estocásticos y especialmente en las citadas obras de P. Levy, Doob, Gikhman y la célebre y clásica de Wiener (14) y en mi artículo *"Introducción a los Procesos Estocásticos"*.

2.9 Unos conceptos básicos adicionales son los *Procesos y Cadenas de Markov. El lector que conozca este tema puede pasar adelante.*

Un proceso finito estocástico es independiente si para cualquier estado E su certeza depende del comportamiento de los estados precedentes.

Los estudios de crecimiento económico en un país pueden, con ciertas reservas, por su cierta subjetividad, **compararse** analógicamente con las situaciones físicas de un estado en un momento dados. Se puede deducir el estado de un sistema físico en un instante cualquiera t2 a partir del conocimiento de su estado en un instante anterior (posterior) t1 y no depende del historial del sistema anterior (posterior) del instante t1.

Para los sistemas físicos que obedecen leyes probabilísticas, fenómenos aleatorios, en vez de deterministas, la probabilidad de que el sistema físico (crecimiento-decrecimiento) esté en un estado dado t2 se puede deducir de conocimiento de su estado en un estado cualquiera y no depende del historial del sistema antes del instante t1.

Los procesos estocásticos que representan observaciones de sistemas físicos que satisfacen esta condición se denominan procesos de Markov.

Una clase especial de los procesos de Markov son las *cadenas de Markov* y que definiremos como un proceso estocástico cuyo desarrollo puede considerarse como una serie de transiciones entre valores determinados (llamados "estados" del proceso) y que tienen la propiedad de que la ley de probabilidad del desarrollo futuro del proceso, una vez que está en un estado dado, **depende solo del estado y no de cómo llegó el proceso a dicho estado.** El número de estados posibles es finito o infinito.

Si suponemos el estado Et_0 para pasar al Estado Et_t conocidas las leyes de probabilidad, a estas probabilidades (condicionadas) se denominan probabilidades de transición.

2.10 Otro concepto fundamental es **la función de densidad de la distribución normal bidimensional.** Así, en todos los libros de estadística, citaremos a Cramer (15) como la más clásica y que a partir de la distribución normal bidimensional conjunta deduce las distribuciones marginales y define la probabilidad condicionada de una de las variables respecto a la otra.

Dadas las variables aleatorias x,y con función de densidad f(x,y) y funciones marginales f1(x) y f2(y), la función de densidad condicionada se define por el cociente

$$f\,(y/x) = \frac{f(x,y)}{f_1(x)} \tag{2.10.1}$$

donde en el numerador aparece la función de: densidad conjunta y denominador la de densidad marginal de la x. Este cociente es la función de densidad marginal de la x. Este cociente es la función de densidd de y condicionada a un valor determinado de la x; es decir: *una función de densidad de probabilidad unidimensional* de la variable aleatoria de y pero que depende del valor de la x como un parámetro condicionante de esta f. de d.

Sustituyendo en estas expresiones las funciones de densidad conjunta y marginal normales tendremos para la condicionada normal de y respecto de x la expresión:

$$f(y/x)=[\sigma_2(2n(1-\rho^2)^{-1/2}]\exp(-1/2(y-m_2-\rho\sigma_2/\sigma_1(x-m_1)]/(1-\rho^2) \quad (2.10.2)$$

Como es bien conocido el exponente (prescindiendo del signo) es una forma cuadrática definida positiva que puede condensarse en forma matricial.

En la fórmula anterior los parámetros m_1 m_2 son las medias de las x y de las y respectivamente σ_1 σ_2 las desviaciones típicas y ρ el coeficiente de correlación lineal.

En esta fórmula recordando las varianzas residuales podemos sustituir $\sigma 2(t-\rho^2)^{1/2}$ por $\sigma_{2.1}$ y $\rho\sigma_2/\sigma_1$ por $\beta_{2.1}$ recordando las notaciones de Yule y Kendall (16) y también de mi estudio sobre el "Espacio de Hilbert aplicado a la Estadística" (17). Aunque debieran ampliarse los subíndices cuando se trata de más de una variable explicativa no lo haremos e inclusive prescindiremos de los subíndices 2.1 pero debe tenerse presente su significado.

III PROCESOS ESTOCÁSTICOS CON DISTRIBUCIÓN NORMAL

3.1. **Hipótesis fundamentales**

1ª. La variable transformada $y(\lambda, t)$ de Box-Cos definida en (1.3.2) y (1.3.2) sigue una distribución normal a la cual **introducimos el parámetro t para representar un proceso estocástico markoviano.**

2ª La variable aleatoria $y(\lambda, s)$ condicionada a que en el estado t tome el valor $y(\lambda, t)$, en un tiempo anterior ($s\geq t$) es, según lo indicado procedentemente considerando la $y(\lambda,s)$ normal en el tiempo s

$$y(\lambda,s) = y(\lambda,t)+\beta\,(s-t)+ \mu_i \quad (3.1.1)$$

μ_i es un proceso estocástico $N[o, \sigma(\lambda)^2(t_i-t_{i-1})]$ de incrementos incorrelacionados.

De las hipótesis se deduce que la esperanza matemática de la variable $y(\lambda,s)$, condicionada a que en el momento t sea $y(\lambda,t)$ es precisamente

$$E[y(\lambda,s)] = y(\lambda,t)+\beta\,(s-t) \quad (3.1.2)$$

3ª La varianza del proceso estocástico es

$$E[y(\lambda,s)]-y(\lambda,t)-\beta(t_i-t_{i-1})^2 = \sigma^2(\lambda)(s-t) \quad (3.1.3)$$

donde λ es el parámetro de la transformación y dependiendo la varianza de referido parámetro.

3.2 Las anteriores expresiones las podríamos haber calculado partiendo de la función de densidad de la variable aleatoria y(λ.s) que por ser normal su función de densidad condicionada es:

$$f[y(\lambda,s)]/y(\lambda,t)] = \frac{e^{-[y(\lambda,s)-y(\lambda,t)-\beta(s-t)]^2/[2\ \delta(\lambda)^2(s-t)]}}{\delta(\lambda)\sqrt{(2\pi(s-t))}} \qquad (3.2.1)$$

siendo su función característica

$$\varphi s(\lambda,r) = \exp.(ir[y(\lambda,t)+\beta\ (s-t)]-\sigma^2(\lambda)(\sigma-t)r^2/2] \qquad (3.2.2)$$

Hemos empleado el parámetro r para no confundirlo con el parámetro t del proceso estocástico.

Derivando n veces la anterior con respecto al parámetro r y haciendo r= 0 y dividiendo el resultado de cada derivación por i^n tenemos los momentos condicionados de la variable aleatoria y(λ.s) condicionada a su valor que toma en el estado inicial y(λ,t).

Para n=1 tendremos la media de y(λ,s) condicionada a y(λ,t). E igualmente para el momento de segundo orden respecto al origen y consecuente la varianza.

De esta forma obtendríamos las mismas fórmulas de la media condicionada del proceso estocástico y la varianza condicionada y que hemos partido de las hipótesis y cuyas fórmulas eran las (3.1.1) y la (3.1.2).

IV ESTIMACION POR PUNTO DE LOS PARAMETROS

4.1 *Formación de la Función de Verosimilitud*

Las propiedades de las funciones de verosimilitud y, sobre todo, recordando las asintóticas de los estimadores hallados por este método son normales nos da mayor claridad de exposición y rigor en nuestras demostraciones ya que podemos asegurar que mencionados estimadores siguen asintóticamente distribuciones normales.

Supondremos los instantes $t=t_0,t_1,t_2,t_3,..t_n=s$ donde $t_i<t_{i+1}$. Estimaremos los parámetros σ(λ) y ß por máxima verosimilitud. Por las hipótesis iniciales y siendo las u_i normales de incrementos incorrelacionados supone por las propiedades en la distribución normal el caso especial de independencia.

La f. de d. es según sabemos por la (3,2,1) y aplicando para dos instantes del tiempo consecutivos t_{i-1},t_i es

$$f[y(\lambda,t_l)/[y(\lambda,t_{l-1}) = \frac{\exp-[y(\lambda,t_l)-y(\lambda,t_{l-1})-ß(t_l-t_{l-1})]^2/2\sigma 2(\lambda)(t_l-t_{l-1})}{\sigma(\lambda)\sqrt{[2\pi(t_l-t_{l-1})]}} \quad (4.1.1)$$

La función de verosimilitud es el producto de n factores independientes de funciones de densidad (4.1.1) porque u_i hemos postulado la independencia e incorrelación por ser normales. Luego la función de verosimilitud es:

$$L[y(\lambda),ß,\sigma(\lambda)^2) = \frac{\exp(-\Sigma(y(\lambda,t_i)-y(\lambda,t_{l-1})-ß(t_l-t_{l-1}))^2/2\sigma(\lambda)^2(t_i-t_{l-1}))}{[2\pi\sigma(\lambda)^2]^{n/2}\prod(t_l-t_{l-1})} \quad (4.1.2)$$

4.2. *Logaritmos y derivados parciales primeras de la función de verosimilitud.*

La función de verosimilitud depende (para una muestra determinada) de los valores poblacionales $[ß,\sigma(\lambda)^2]$ y esta función de densidad varía con los referidos parámetros poblacionales. Es lógico pensar que si se verifican las hipótesis para el proceso estocástico $y(\lambda,t)$ si se maximizara la función de verosimilitud (que depende de los parámetros poblacionales) entre todos los estimadores formados la muestra obtenida será la más probable si estos estimadores hacen máxima la función de densidad conjunta.

Así pues tomando logarimos neperianos de esta f. de v. y simplificando la notación tenemos:

$$1_nL(.)=C-n/2\ln\sigma(\lambda)^2-\Sigma[Y(\lambda,t_i)-y(,t_{i-1})-ß(t_i-t_{i-1})\}^2/2\sigma(\lambda)^2(t_l-t_{l-1}) \quad (4.2.1)$$

Derivemos parcialmente respecto a ß y $\sigma(\lambda)^2$

$$\delta L(.)/\delta ß - \Sigma[y(\lambda,t)-y(\lambda,t)-ß(t-t)]/\sigma(\lambda)^2 \quad (4.2.2)$$

$$\partial 1_nL(.)/\delta\sigma(\lambda)^2=-n/2/\sigma(\lambda)^2+\Sigma[(Y(.t_i)-y(.t_{i-1})-ß(t_i-t_{i-1})]^2/2\sigma(\lambda)^4(t_l-t_l) \quad (4.2.3)$$

4.3 *Derivadas Logarítmicas parciales segundas de la F. de V.*

Las derivadas segundas son:

$$\partial^2L_n(.)/\partial^2ß = \Sigma(t_i-t_{i-1})/\sigma(\lambda)^2 = (s-t)/\sigma(\lambda)^2 \quad (4.3.1)$$

porque en la sumatoria el extremo inferior de t, t_0=t; y el extremo superior de t, t_n = s

Por el teorema de Swartz la derivada segunda mixta, el orden de derivación es el mismo. Luego:

$$\partial^2 L_n(.)/\partial ß\partial\sigma(\lambda)^2 = -\Sigma[y(\lambda,t)-y(\lambda,t)-ß(t_i-t_{i+1})]^2/\sigma(\lambda)^2(t_i-t_{i+1})] \tag{4.3.2}$$

$$\delta^2 L_n L(.)/\delta^2\sigma(\lambda)^2 =$$

$$-n/2\sigma(\lambda)^4-\Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1})-ß(t_i-t_{i-1})]^2/\sigma(\lambda)^6(t_i-t_{i-1})] \tag{4.3.3}$$

4.4. *Estimadores máximo verosímiles*

De las ecuaciones (4.2.2) y (4.2.3) los valores de losparámetros poblacionales que maximizan la función de verosimilitud serán aquellos que sus primeras derivadas parcialese sean nulas.

Así pues, tenemos:

$$0 = \Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1})-\hat{ß}(t_i-t_{i-1})] =$$

$$\hat{ß} = 1/(s-t)\Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1}) = [y(\lambda,s)-y(\lambda,t)]/(s-t) \tag{4.4.1}$$

De forma semejante, igualando a cero para un valor que anule la derivada con respecto a $\sigma(\lambda)^2$ tenemos el estimador de la varianza que es

$$\hat{\sigma}(\lambda)^2 = 1/n\ \Sigma[y(\lambda,t)-y(\lambda,t)-\hat{ß}(t-t)]^2/(t_i-t_{i-1})] \tag{4.4.2}$$

4.5 *Propiedades de los estimadores*

<u>Propiedades de ß</u>

1ª. Es un estimador centrado.

En efecto: De la (4.4.1) deducimos

$$E\hat{ß} = E\Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1})]/(s-t) =$$

$$=E\Sigma[ß(t_i-t_{i-1})+\mu_i]/(s-t) = ß \tag{4.5.1}$$

según expusimos en 3.1.2

$y(\lambda,t_i)=y(\lambda,t_{i-1})+ß(t_i-t_{i+1})+\mu_i$ (4.5.2)

y por las hipótesis de μ_i de la cual se deduce la anterior.

2ª Es un estimador consistente.

La varianza del estimador $\hat{ß}$ es

$$\mathrm{Var}(\hat{\beta})=E(\hat{\beta}-\beta)^2 = E\{\Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1})]/(s-t)-\beta\}^2 =$$

$$E\{\Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1})-\beta(t_i-t_{i-1})]/(s-t)^2 = E\{\Sigma\mu_i\}^2/(s-t)^2$$

$$\mathrm{Var}(\hat{\beta}) = \sigma(\lambda)^2/(s-t) \qquad (4.5.3)$$

De la anterior expresión si $s-t \to \infty$ se incrementa la (4.5.4) tiende a cero lo que implica la consistencia que siendo ß centrado y la varianza que depende de la diferencia s–t tenderá a concentrarse en un solo punto con probabilidad la certeza.

Estudiaremos posteriormente si es o no eficiente.

<u>Propiedades de $\hat{\sigma}(\lambda)^2$</u>

1ª *Es un estimador sesgado*

Así recordando que $E\mu_i^2 = \sigma(\lambda)^2(t_i-t_{i-1})$ y por la (4.4.2) tenemos:

$$\hat{\sigma}(\lambda)^2 = 1/n\Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1})-\hat{\beta}(t_i-t_{i-1})]^2/(t_i-t_{i-1})] \qquad (4.5.4)$$

Sumando y restando $\beta(t_i-t_{i-1})$ y recordando el valor de μ_i que se deduce de la (4.5.2) la podemos escribir:

$$\hat{\sigma}(\lambda)^2 = 1/n\Sigma\{[\mu_i-(\hat{\beta}-\beta)(t_i-t_{i-1})]2/(t_i-t_{i-1})\} =$$

$$= 1/n\{\Sigma[\mu_i^2/(t_i-t_{i-1})+(\hat{\beta}-\beta)^2(t_i-t_{i-1}) - 2\mu_i(\beta-\beta)]\}$$

Tomando esperanzas matemáticas hallamos la esperanza del estimador de la varianza estimada por máxima verosimilitud:

$$E\hat{\sigma}(\lambda)^2 = \sigma(\lambda)^2 + \sigma(\lambda)^2/n-2\Sigma E\mu_j(\Sigma[y(\lambda,t_k)-y(\lambda.t_{k-1})]/(t_k-t_{k-1})-\beta =$$

$$= \sigma(\lambda)^2 + \sigma(\lambda)^2/n - 2/nE\Sigma\mu_i\Sigma\mu_k =$$

$$= \sigma(\lambda)^2 - \sigma(\lambda)^2/n + 2\sigma(\lambda)^2/n =$$

$$E\hat{\sigma}(\lambda)^2 = \sigma(\lambda)^2(1-1/n) \qquad (4.5.5)$$

que demuestra no es centrado pero lo es asintóticamente.

De la anterior podemos deducir una fórmula para el estimador centrado:

$$\tilde{\sigma}(\lambda)^2 = \hat{\sigma}(\lambda)^2/(n-1) \qquad (4.5.6)$$

donde $\hat{\sigma}(\lambda)^2$ es el estimador (4.5.4)

es decir

$$\hat{\sigma}(\lambda)^2 = 1/(n-1)\ (\Sigma[y(\lambda,t_i)-y(\lambda,t_{i-1})-\hat{\beta}(t_i-t_{i-1})]^2 \qquad (4.5.6b)$$

4.6 Distribuciones de los estimadores

Distribución de $\hat{\beta}$

Por ser centrado y combinación de variables normales es normal. Luego

$$\hat{\beta} \Rightarrow N(\beta,\ \sigma(\lambda)^2/(s-t) \qquad (4.6.1)$$

Distribución de $\hat{\sigma}(\lambda)^2$

Para conocer esta distribución recordaremos las hipótesis de μ_i y formemos una X_n de la siguiente forma:

$$\Sigma\mu_i^2/\ \sigma(\lambda)^2(t_i-t_{i-1}) =$$

$$x_n^2 = \Sigma\,\frac{(y(\lambda,t1)-y(\lambda,t1-1)-\beta(t1-t1-1))^2}{\sigma(\lambda)^2(t1-t1-1)}$$

Sumando y restando $\hat{\beta}(t_i-t_{i-1})$ sacando factor común a t_i-t_{i-1}, tenemos:

$$x^2_n=\Sigma(y(\lambda,t_i)-y(\lambda,t_{i-1})-\tilde{\beta}(t_i-t_{i-1})+(\tilde{\beta}-\beta(t_i-t_{i-1}))2/\sigma(l)2(t_i-t_{i-1})$$

$$=n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2+(\hat{\beta}-\beta)\Sigma(t_i-t_{i-1})/\sigma(\lambda)^2=n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2+(\hat{\beta}-\beta)^2(s-t)/(\sigma(\lambda)^2) \qquad (4.6.2)$$

El doble producto se anula debido a la fórmula (4.4.1) obtenida por máxima verosimilitud para el estimador $\hat{\beta}$.

En efecto:

$$\Sigma(\hat{\beta}-\beta(y(\lambda,t_i)-y(\lambda,t_{i-1})-\hat{\beta}(t_i-t_{i-1}))=(\hat{\beta}-\beta)\{\Sigma(y(\lambda,t_i)-y(\lambda,t_{i-1})-\hat{\beta}\Sigma(t_i-t_{i-1})\}$$

y como lo encerrado entre corchetes es cero (recordando la condición para estimar β) hemos demostrado la descomposición de la X_n^2 en dos X^2 de Pearson.

La primera es

$$x_1^2=\{(\hat{\beta}-\beta)/\sigma(\lambda)/\sqrt{(s-t)}\}^2 \qquad (4.6.3)$$

recordando las propiedades del estimador $\hat{\beta}$ (4.6.1) y por ser el paréntesis una variable tipificada que, al cuadrado es una X_1^2

Por el Teorema de Cochran (18) de la descomposición de la X^2 tenemos que la otra sumatoria es otra X^2 pero con n–l g. de l.:

$$X_{n-1}^2 = n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2 \qquad (4.6.4)$$

siendo $\hat{\sigma}(\lambda)^2$ el estimador obtenido en (4.5.4)

Por las propiedades de la X2 tenemos:

$$E(X_{n-1}^2)=E\{n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2\}=n-1 \Rightarrow E\{n\hat{\sigma}(\lambda)^2\}=n\sigma(\lambda)^2(1-1/n)$$

como ya la obtuvimos anteriormente.

La varianza de este estimador sesgado es:

$$Var(X_{n-1}^2)=2(n-1)=Var(n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2)=n^2/\sigma(\lambda)^2 \Rightarrow$$

$$Var(\hat{\sigma}(\lambda)^2=2(n-1)\sigma(\lambda)^4/n^2 \qquad (4.6.5)$$

Distribución del ~estimador centrado de la $\tilde{\sigma}(\lambda)^2$

Este estimador centrado lo obtuvimos de (4.5.6) y, en consecuencia, tenemos:

$$E\tilde{\sigma}(\lambda)^2=En\hat{\sigma}(\lambda)^2/(n-1)=\sigma(\lambda)^2$$

La varianza de este estimador centrado es:

$$Var(\tilde{\sigma}(\lambda)^2)=Var(n\hat{\sigma}(\lambda)^2/(n-1))=n2/(n-1)^2 Var\ \hat{\sigma}(\lambda)^2 \Rightarrow$$

$$Var(\tilde{\sigma}(\lambda)^2)=2\sigma(\lambda)^4/(n-1) \qquad (4.6.6)$$

La distribución del estimador $\tilde{\sigma}(\lambda)^2$. Como la de $\hat{\sigma}(\lambda)^2$se relaciona con la X_n^2. Así que es sencillo encontrar su ley de distribución.

$$\text{Así la de } n\hat{\sigma}(\lambda)^2/n\sigma(\lambda)^2=X_n^2 \Rightarrow$$

$$\tilde{\sigma}(\lambda)^2=n\hat{\sigma}(\lambda)^{2/}(n-1)=\sigma(\lambda)^2X_{n-1}^2/(n-1) \qquad (4.6.7)$$

4.7. *Matriz de información y Cota de Cramer Rao*

Para calcular la Matriz de Información nos basamos en las fórmulas obtenidas en 4.3 de las derivadas segundas parciales del logaritmo de la función de verosimilitud (4.3.1), (4.3.2) y (4.3.3) calculando sus esperanzas matemáticas y cambiando de signo

$$E\{\delta^2 LnL(.)/\delta 2\beta\}=-(s-t)/\sigma(\lambda)^2 \quad (4.7.1)$$

por ser una constante.

$$E\{\delta 2LnL(.)/\delta(\beta)\delta\sigma(\lambda)2\}=-\Sigma E\{((y(\lambda,t_i)-y(\lambda,t_{i-1})-\beta(t_i-t_{i-1})/(\sigma(\lambda)^4\}=0$$

Finalmente la

$$E\delta 2LnL(.)/\delta 2\sigma(\lambda)2=n/2\sigma(\lambda)^4$$

$$E\{y(\lambda,t_i)-y(\lambda,t_{i-1})-\beta(t_i-t_{i-1}))/(\sigma(\lambda)^6(t_i-t_{i-1}\}=$$

$$=n/(2\sigma(\lambda)^2)-\Sigma u_i^2/\sigma(\lambda)^6(t_i-t_{i-1})=-n/(2\sigma(\lambda)^4)$$

Luego la matriz de Información es

$$MI=\begin{cases}[(s-t)/\sigma(\lambda)2,\ 0]\\ [0,n/2\sigma(\lambda)^4]\end{cases}$$

Y la Cota de Cramer-Rao, por ser una matriz diagonal, es su inversa que es sencilla:

$$\Sigma CR=\begin{cases}[\sigma(\lambda)^2/(s-t),\ 0]\\ [0,2\sigma(\lambda)^4/n]\end{cases}$$

y cualquier estimador centrado de los parámetros mencionados las varianzas de estos estimadores no pueden ser inferiores a referidas cotas.

Pero las varianzas de los estimadores obtenidos por máxima verosimilitud la correspondiente a $\hat{\beta}$ coincide con la Cota. No así la del estimador centrado $\hat{\sigma}(\lambda)^2$ que difiere en el denominador y en una unidad.

Luego completando las propiedades de los estimadores podemos añadir que como el estimador $\hat{\beta}$ su varianza coincide con la Cota es un estimador eficiente.

Por otra parte el estimador de $\hat{\sigma}(\lambda)^2$, es consistente porque su varianza tiende a cero con $n\rightarrow\infty$. Y por otra parte es asintóticamente eficiente.

V ESTIMACION POR INTERVALO

5.1. *Errores por causa de las transformaciones*

En el capítulo anterior hemos considerado la estimación por punto utilizando las hipótesis establecidas para las variables transformadas y por ser una familia de transformaciones dependiente de parámetros, solamente nos dedicamos a la aplicación de los primeros momentos para determinar las características de las estimaciones efectuadas. Pero si partimos de unas hipótesis, los errores que cometemos, dependen de los parámetros de la población. Y evidente si los objetivos perseguidos por Tukey basados en las observaciones empíricas al hacer variar los parámetros conseguimos el fin de hacer que las transformaciones de las variables cumplan las hipótesis establecidas *los errores por estas transformaciones son menores y estos errores típicos* se reducirán si las variables transformadas siguen prácticamente una distribución normal, y sin ninguna dificultad aplicaremos a estas variables las distribuciones de los errores.

Dos problemas trataremos aquí. *El primero de ellos*, los intervalos de confianza de las variables transformadas con las hipótesis establecidas e incluso también (y a forma un poco de introducción) las transformaciones de Tukey y posteriormente la de Box-Cos complementando la teoría de los Procesos de Markov como venimos haciéndolo.

El segundo problema cuando conozcamos los intervalos de confianza para las variables transformadas *(en las hipótesis que sigan rigurosamente la distribución normal) es deshacer el cambio* y pasar a las variables originales; es decir: a las primitivas. Es *aquí donde se producen los errores por dos causas fundamentales: La 1ª si no siguen exactamente* las distribuciones transformadas las hipótesis establecidas *los errores se multiplican al deshacer el cambio de variable y la 2ª por los errores de muestreo en las estimaciones de los parámetros poblacionales.*

5.2. *Transformación de Tukey*

En (1.3.1) vimos cómo fundamentalmente esta transformación depende rigurosamente de dos parámetros α y λ y que para simplificar denominaremos $y_i(\alpha,\lambda)$ a la expresión:

$$y_i(\lambda,\alpha)=(y_i+\alpha)^\lambda \qquad (5.2.1)$$

Si y(i) lo relacionamos con alguna variable, x por ejemplo, y partimos de las hipótesis de regresión clásicas, incorrelación, normalidad de los errores, homocedasticidad, etc. la anterior expresión podremos escribirla así

$$y_i(\lambda_1,\alpha)=\beta_1+\beta_2 x_{2i}(\lambda_2,\alpha)+u_i \qquad (5.2.2)$$

donde x_2 es una variable exógena.

Los parámetros α y λ pueden ser distintos para cada variabla y así damos *generalidad a las expresiones anterior.*

Si fueren un conjunto de k variables exógenas la expresión anterior la podríamos poner en forma matricial y cada una de las variables podría tener transformaciones distintas.

La perturbación aleatoria ui la consideraremos cumple las condiciones clásicas del modelo de regresión cuya varianza de μ_i es $\sigma(\lambda,\alpha)^2$ o simplemente $\sigma(\lambda)^2$ como venimos haciéndolo aunque dependa de dos parámetros la transformación de la x puede modificarse dando valores distintos de λ y α que a los de la variable y.

En el caso de ser un proceso de Markov la varianza sería igual a

$$\sigma(\lambda,\alpha)^2(t_j-t_{j-1}) \tag{5.2.3}$$

Para cada par de parámetros (α, λ) obtendremos intervalos de confianza de la transformada de y cuyo cálculo es sencillo.

5.3. *Intervalos de confianza para las y para un nivel de significación determinado en las transformaciones de Tukey conocidos los parámetros.*

Este caso no se dará nunca en la práctica excepto para datos construidos artificialmente.

Tomando el nivel de significación ε (para distinguirlo del parámetro α usualmente utilizado como nivel de significación y que en nuestro caso tiene una significación diferente), y con las hipótesis establecidas tendremos

$$\text{Prob}(\beta_1+\beta_2 x_{2i}(\lambda)-t_{\alpha/2}\sigma(\lambda)<y_i(\lambda,\alpha)<\beta_1+\beta_2 x_{2i}(\lambda)\}=1-\in \tag{5.3.1}$$

donde la t la buscaremos en la tabla de la normal. Si elegimos el 5% de nivel de significación como sabemos t=1.96, y conocidos los regresores tendremos el intervalo correspondiente.

<u>*De la expresión precedente deducimos fácilmente el intervalo de confianza para la variable primitiva deshaciendo la transformación.*</u>

De la anterior el intervalo de confianza de $1-\in$ para y_i: es

$$-\alpha(\beta_1+\beta_2 x_{2i}(\lambda,\alpha)-t_{\in/2}\sigma(\lambda))^{1/\lambda}<y_i<-\alpha+(\beta_1+\beta_2 x_{2i}(\lambda)+t_{\in/2}\sigma(\lambda))^{1/\lambda} \tag{5.3.2}$$

Hemos supuesto que conocemos los parámetros poblacionales para determinar estos intervalos de confianza de la variable transformada y de la primitiva. En la práctica esto no ocurre y habremos de estimarlos y aplicaremos en estas fórmulas la t de Student.

5.4 *Funciones de probabilidades primitivas basadas en las hipótesis de Tukey*

Si las hipótesis se cumplen con rigurosidad los dos intervalos de confianza (5.3.1) y (5.3.2) son correctos y la probabilidad que la variable y_i (variable original) tome el valor de y_i es sencillamente por las hipótesis efectuadas:

$$f(y)=f(y(\lambda,\alpha)|dy(\lambda,\alpha)/dy| \qquad (5.4.1)$$

donde

$$dy(\lambda,\alpha)=\lambda(y+\alpha)^{\lambda-1}dy \qquad (5.4.2)$$

Por las hipótesis de normalidad de la variable transformada obtendremos la (5.4.1) sustituyendo los dos factores:

$$f(y(\lambda,\alpha)=\exp\{-(y(\lambda,\alpha)-\beta_1-\beta_2x_2)^2/2\sigma(\lambda,\alpha)^2\}/\sigma(\lambda,\alpha)\sqrt{2\pi} \qquad (5.4.3)$$

Sustituyendo esta expresión en la (5.4.1) tenemos para la función de densidad de la y

$$f(y)=\exp\{-((y+\alpha)^{\lambda}-\beta_1-\beta_2x_2)^2/2\sigma(\lambda,\alpha)^2\}|\lambda(y+\alpha)^{\lambda-1}|\sigma(\lambda,\alpha)\sqrt{2\pi} \qquad (5.4.4)$$

En la expresión anterior hemos suprimido los subíndices i para facilitar la comprensión y por razones tipográficas.

5.5 *Intervalo de confianza para el proceso estocástico de la transformada de Cox-Box y(λ,s) cuando se conozcan los parámetros poblacionales.*

Las razones expuestas para la transformación de Tukey son válidas para estas transformaciones de Box-Cox excepto que la varianza en este caso concreto es de

$$\sigma(\lambda)^2(t_i-t_{i-1}) \qquad (5.5.1)$$

si consideramos la transformación como un proceso estocástico.

Conocidos los parámetros poblacionales de las transformadas, por la normalidad la perturbación aleatoria que se encuentre en el intervalo es

$$Prob(-t_{\epsilon/2}\sigma(\lambda)\sqrt{(s-t)}<u_i<t_{\epsilon/2}\sigma(\lambda)\sqrt{(s-t)})=1-\epsilon \qquad (5.5.2)$$

siendo ϵ el nivel de significación elegido y los valores de $t_{\epsilon/2}$ *son tomados de la normal.*

Sustituyamos μ_i por su relación con el proceso transformado y despejando $y(\lambda,s)$ la probabilidad de que el proceso estocástico transformado esté entre los extremos que a continuación se relacionan es

$$y(\lambda,t)+\beta(s-t)-t(\alpha/2)\sigma(l)\sqrt{(s-t)}<(y(\lambda,s)<$$

$$y(\lambda,t)+\beta(s-t)+t(\alpha/2)\sigma(l)\sqrt{(s-t)} \qquad (5.5.3)$$

siendo esta probabilidad $1-\in$

5.6 *Intervalo de confianza de la variable original cuando las hipótesis de Box-Cos son ciertas y los parámetros conocidos.*

Igualmente que hemos hecho para el intervalo de y si hacíamos la transformación de Tukey haremos para el proceso estocástico $y(t_i)$ pero supondremos el punto s conocido el valor del proceso en su anterior situación en el punto t.

Efectuando operaciones y sustituyendo $y(\lambda,t)$ por la variable primitiva tenemos la siguiente desigualdad con un nivel de significación $\in$.

Sustituyendo en la (5.5.3) la variable aleatoria μ_i por su valor y después de calcular el valor de y(s) cuando se conoce su estado en el tiempo t es decir conocida y(t) tenemos la siguiente expresión

$$(y(t)^\lambda+\lambda(\beta(s-t)-t_{\in/2}\sigma(\lambda)\sqrt{(s-t)}))^{1/\lambda}<y(s)<$$

$$(y(t)^\lambda+\lambda(\beta(s-t)+t_{\in/2}\sigma(\lambda)\sqrt{(s-t)}))^{1/\lambda} \qquad (5.6.1)$$

con probabilidad $1-\in$

Para valores concretos de λ,β y $\sigma(\lambda)$ los intervalos precedentes varían en función **de la raíz cuadrada de s-t y a medida que esta diferencia sea mayor, la amplitud del intervalo es mayor**. El menor intervalo es cuando s=t porque en este caso **coinciden los extremos inferior y superior del intervalo y este intervalo es degenerado porque se convierte en un punto** cualquiera que sea el valor de $t_{\in/2}$ y si damos un valor suficientemente grande a $t_{\in/2}$ toda la masa de probabilidad está concentrada precisamente en y(t) por lo que el límite de la (5.6.1) se transforma en la distribución en un solo punto y su probabilidad es la certeza.

5.7. *Funciones de probabilidad de las variables primitivas de las transformaciones de Box-Cox*

Por un camino semejante a las transformaciones de Tukey determinamos las leyes de probabilidad de las variables primitivas supuestas siempre que sus transformadas siguen leyes normales.

$$f(y(s))=f(y(\lambda,s)|dy(\lambda,s)/dy(s)| \qquad (5.7.1)$$

donde $f[y(\lambda,s)]$ es la función de densidad normal del proceso estocástico en el punto s y

$$dy(\lambda,s)=y^{\lambda-1}d_y \qquad (5.7.2)$$

Luego la f. de d. de y(s) es

$$f(y(s))=$$

$$y(s)^{\lambda-1}[\exp(-\{y(s)^{\lambda}-y(t)^{\lambda})/\lambda-ß(s-t))^2/2\sigma(\lambda)^2(\sigma-t)\}\}/\sigma(\lambda)(\sqrt{}(2\pi(\sigma-t)))] \qquad (5.7.8)$$

Esta función de densidad depende fundamentalmente del parámetro de la transformación de su desviación típica $\sigma(\lambda)$ y de ß.

Un caso particularmente importante es cuando $\lambda \to 0$.

Decimos particularmente importante porque demostraremos en el siguiente capítulo que sigue **la distribución logarítmico normal del proceso y(s).**

5.8. *Intervalos de confianza de la transformación de Box-Cox cuando son desconocidos los parámetros poblacionales.*

1. Por la importancia de nuestro estudio en las aplicaciones prácticas recordaremos que las estimaciones de los parámetros poblacionales son las siguientes:

Estimador $\hat{ß}$.

Es centrado. Varianza de $\hat{ß}$

Es $var(\hat{ß}) = \sigma(\lambda)^2/(s-t)$ donde s es el extremo superior de las observaciones del tiempo y t es el inferior. Como nota importante el tiempo puede estar medido en años, meses, días no siendo por qué ser entero necesariamente. No obstante examinaremos este caso concreto más adelante.

Estimador $\hat{\sigma}(\lambda)^2$

1. Es sesgado $E\hat{\sigma}(\lambda)^2=(n-1/n)\sigma(\lambda)^2$

2. El estudio de los intervalos de confianza lo haremos **para el predictor con respecto a su media teórica y después con respecto a un valor individual.**

Así pues el predictor con respecto a la media poblacional de $y(\lambda,t_i)$ es decir con respecto

$$E(y(\lambda,t_i)=y(\lambda,t_{i-1})+ß(t_i-t_{i-1}) \quad (5.8.1)$$

El caso más importante es cuando tomamos el tiempo de predicción fuera del intervalo del ajuste y llamando w > s (extremo superior del tiempo de estimación de los *parámetros y siendo s el último valor del proceso para la estimación de los parámetros* y W el tiempo de predicción. Tenemos para el **predictor por punto** con respecto a la (5.8.1)

$$Py(\lambda,w)=y(\lambda,s)+ß(w-s) \quad (5.8.2)$$

En el supuesto que el proceso transformado no cambie su estructura poblacional.

Y por las propiedades del estimador $\hat{ß}$ tenemos

1. *Esperanza matemática: del predictor respecto a su media poblacional*

$$E(Py(\lambda,w))=y(\lambda,s)+ß(w-s) \quad (5.8.3)$$

y que nos indica su insesgadez.

2. *Varianza del Predictor del proceso transformado con respecto a su media teórica*

De las (5.8.2) y (5.8.3) tenemos

$$E\{Py(\lambda,w)-y(\lambda,s)-ß(w-s))^2=E((\hat{ß}-ß)£w-s))^2=Var(\hat{ß})(w-s)^2$$

$$Var(P(y(\lambda,w)=\sigma(\lambda)^2(w-s)^2/(\sigma-t) \quad (5.8.4)$$

recordando la varianza del estimador $\hat{ß}$ expuesta al principio de este párrafo.

Analizando el significado de esta varianza observamos que depende no solamente del parámetro de la transformación λ sino de las dos diferencias de tiempos: la del tiempo de predicción w menos el tiempo efectuado en la última observación s y de la diferencia de tiempos del ajuste (s–t).

Esta varianza por ser w ≥ s (siendo el denominador s-t una diferencia entre los extremos del tiempo de los valores empíricos observados del proceso estocástico) varía con w por cuanto el denominador es constante. Un caso particularmente importante es cuando w→s y se comprueba **cómo disminuye la varianza del predictor es decir la función de densidad de probabilidad y como la varianza tiende a cero** y significa que **su ley de probabilidad es la distribución en un punto siendo la función de densidad en s la función delta de Dirac como límite de la normal degenerada.**

3. *Intervalo de confianza para el Predictor del proceso transformado cuando no se conocen los parámetros poblacionales, cuando se mide respecto a la media poblacional del proceso transformado.*

Elegido un nivel de significación $\in$ y como las estimaciones de los parámetros son conocidas formaremos una t de Student con n-1 grados de libertad y por su sencillez omitimos ya que el numerador es una variable tipificada y en el denominador una raíz cuadrada de $x_{n-1}^2/(n-1)$

Conocidas la esperanza y varianza del PREDICTOR respecto de su media poblacional tipificaremos para obtener una normal (0,1).

Recordando la (5.8.2) y (5.8.3) la variable tipificada que decimos es:

$$(P(y(\lambda,w)-(y(\lambda,s)+\beta(w-s))/(\sigma(\lambda)(w-s):\sqrt{}(s-t))$$

Por otra parte recordemos la distribución de la varianza estimada que se relaciona con la X^2 y que es

$$x_{n-1}^2=n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2$$

Para formar la t de Student con n-1 grados de libertad pondremos en el numerador la variable normal (0,1) del Predictor y dividiremos por la raíz cuadrada de la media de la X^2 de Pearson (dividida por sus grados de libertad) siendo ambas variables independientes como puede comprobarse en cualquier Tratado de Econometría.

La t de Student con n-1 g. de l. es

$$\text{La } t_{n-1} = \frac{(P(y(\lambda,w)-y(\lambda,s) - \beta(w-s)) / (\ \sigma(\lambda)\ (w-s):\sqrt{}(s-t)))}{\sqrt{}(n\hat{\sigma}(\lambda)^2 /\ \ \sigma(\lambda)^2)/n-1}$$

Simplificando la anterior y recordando el estimador centrado $\tilde{\sigma}(\lambda)^2$ de la varianza la anterior puede escribirse sencillamente:

$$t_{n-1} = \frac{Py(\lambda,w)-y(\lambda,s) - \beta(w-s)}{\tilde{\sigma}(\lambda)\ (w-s)/\sqrt{}(s-t)} \qquad (5.8.5)$$

Esta t_{n-1} si elegimos un nivel de significación a dos bandas $\in/2$ (probabilidad de cada banda) la probabilidad que la t_{n-1} de la (5.8.5) caiga entre $-t_{n-1}(\in/2)$ y $+t_{n-1}(\in/2)$ es lo mismo que **la esperanza matemática del predictor** P(y(λ,w) desconocida caiga dentro del intervalo

$$y(\lambda,s)+\hat{\beta}(w-s)-\hat{\sigma}(\lambda)(w-s)/\sqrt{(s-t)}<$$

$$E(Py(\lambda,w)\}=y(\lambda,s)+\beta(w-s)<$$

$$y(\lambda,s)+\hat{\beta}(w-s)+\hat{\sigma}(\lambda)(w-s)/\sqrt{(s-t)} \qquad (5.8.6)$$

donde $\hat{\sigma}(\lambda)^2$ como sabemos es un estimador centrado de $\sigma(\lambda)^2$ y que depende de λ.

Ahora los extremos del intervalo (5.8.6) son conocidos y esta expresión significa que la esperanza de la predicción del proceso transformado $y(\lambda,w)$ **en w**: es decir $E\{y(\lambda,w)\}$ caiga en el intervalo formado por los extremos de la predicción tiene una fiabilidad de $1-\epsilon$ donde ϵ es el nivel de significación elegido.

4. *Intervalo del proceso estocástico y(w) cuando conocemos el intervalo del proceso estocástico de la predicción respecto de su esperanza matemática del proceso transformado y(λ,w) desconociendo sus parámetros poblacionales.*

A la **esperanza matemática del proceso transformado** en el punto w le corresponderá **una variable promedio del proceso original que representaremos por** Ry(w). Existe una correspondencia biunívoca. Variando w tenemos el valor promedio Ry(w) correspondiente al proceos estocástico transformado $y(\lambda,w)$.

Deshaciendo el cambio de variable $y(\lambda,t)$ y sustituyéndolo por su valor **en el supuesto que el proceso transformado cumpla las hipótesis establecidas** obtenemos el intervalo para el valor con respecto a la variable primitiva del promedio del proceso transformado y con probabilidad $1-\epsilon$.

Así después de efectuar operaciones y simplificar intervalo que buscamos es:

$$[y(s)^\lambda+\lambda\{\hat{\beta}(w-s)-\hat{\sigma}(\lambda)(w-s)/\sqrt{(s-t)}\}]^{1/\lambda}<Ry(w)<$$

$$<[y(s)^\lambda+\lambda\{\beta(w-s)+\hat{\sigma}(\lambda)(w-s)/\sqrt{(s-t)}\}]^{1/\lambda} \qquad (5.8.7)$$

donde $\hat{\sigma}(\lambda)^2$ estimador centrado de $\sigma(\lambda)^2$ su valor es según sabemos

$$\hat{\sigma}(\lambda)^2=1/(n-1)\sum\{(y(t_i)-y(t_{i-1})-\hat{\beta}(t_i-t_{i-1}))^2/(t_i-t_{i-1})\}$$

fórmula que demostramos cuando tratamos de las estimaciones de los parámetros poblacionales, sus distribuciones y propiedades.

El valor de Ry(w) (repetimos nuevamente) es el punto promedio de la regresión del proceso estocástico en w correspondiente a la de la esperanza matemática de la variable transformada.

La fórmula (5.8.7) nos permite encontrar intervalos de confianza para los puntos del proceso estocásticos correspondientes a la de la regresión de los valores del proceso estocástico original conociendo su transformado y aunque en el proceso transformado eran rectas en el del proceso original no tiene necesariamente que ser recta. Esto se ve fácilmente por la naturaleza de la fórmula del intervalo mencionada (5.8.7).

5. *Intervalo del PREDICTOR del proceso estocástico transformado $y(\lambda,w)$ que contenga un valor individual del proceso con probabilidad $1-\epsilon$ cuando se desconocen los parámetros poblacionales.*

El predictor representado por $Py(\lambda,w)$ coincide con el de la esperanza matemática. Y según vimos en (5.8.2) es

$$Py(\lambda,w)=y(\lambda,s)+\hat{\beta}(w-s)$$

Pero ahora como el valor individual del proceso transformado es

$$y(\lambda,w)=y(\lambda,s)+\beta(w-s)+u_w \tag{5.8.8}$$

si no ha cambiado de estructura en el proceso estocástico transformado.

1. *La esperanza matemática de $y(\lambda,w)$* coincide con la esperanza matemática del predictor y ambas son

$$E\{y(\lambda,w)=EPiy(\lambda,w)=y(\lambda,s)+\beta(w-s) \tag{5.8.9}$$

2. *La varianza del proceso estocástico respecto del predictor es:*

$$E(y(\lambda,w)-y(\lambda,s)-\hat{\beta}(w-s)\}^2=$$

$$E(y(l,s)+\beta(w-s)+u_w-y(\lambda,s)-\hat{\beta}(w-s)\}^2=$$

$$E\{u_w-(\hat{\beta}-\beta)(w-s)\}^2=\mathrm{Var}u_w+(w-s)^2\mathrm{var}(\hat{\beta})$$

ya que el término doble producto su esperanza es nula.

Recordando las varianzas tenemos, según la (4.5.6) y la (5.8.4)

$$\mathrm{Varp}_1(y(\lambda,w)=\sigma(\lambda)^2(w-s)+(w-t)^2\sigma(\lambda)^2/(s-t) \tag{5.8.10}$$

recordando la varianza de $\hat{\beta}$.

Formaremos una t_{n-1} grados de libertad. En el numerador estará la variable tipificada normal donde y(λ,w) - P(y(λ,w) (diferencia entre el valor individual respecto del predictor) que dividida por su desviación típica –la raíz cuadrada de la (15.8.10)– es una variable tipificada. Y el denominador pondremos la raíz cuadrada de una $x_{n-1}^2/n-1$ y como la x^2 con n-1 grados de libertad es

$$n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2$$

la t de Student con n-1 grados de libertad es:

$$t_{n-1} = \frac{y(\lambda,w)-y(\lambda,s) - \hat{\beta}(w-s)}{\sigma(\lambda)\ ((w-s) + (w-s)\ ^2/(s-t))^{1/2}};\sqrt{(n\hat{\sigma}(\lambda)^2 / \sigma(\lambda)^2) /(n-1))}$$

y simplificando queda

$$t_{n-1} = \frac{y(\lambda,w)-y(\lambda,s) - \hat{\beta}(w-s)}{\tilde{\sigma}(\lambda)\ ((w-s) + (w-s)\ ^2/(s-t))^{1/2}} \qquad (5.8.11)$$

donde $\tilde{\sigma}(\lambda)^2$ es el estimador centrado de $\sigma(\lambda)^2$

Elegido ε como nivel de significación (a dos bandas) para que t_{n-1} esté comprendido entre $\pm t_{n-1}(\in/2)$ equivale a que el proceso estocástico transformado esté comprendido en el intervalo de confianza

$$y(\lambda,s)+\tilde{\beta}(w-s)-t_{n-1}(\in/2)\tilde{\sigma}(\lambda)(w-s+(w-s)^2)^{1/2}/(s-t)\}^{1/2}<y(\lambda,w)<$$

$$y(\lambda,s)+\hat{\beta}(w-s)+t_{n-1}(\in/2)\tilde{\sigma}(\lambda)(w-s+(w-s)^2)^{1/2}/(s-t)\}^{1/2} \qquad (5.8.12)$$

que nos da el intervalo para el proceso estocástico transformado individual en el punto w con probabilidad fiducial de 1-∈ cuando se conoce el valor del proceso en el punto s.

5.9. *Intervalo de confianza para el proceso estocástico y(w) cuando se conoce el valor del proceso estocástico en el punto y(s) desconociéndose sus parámetros y el proceso transformado sigue las hipótesis establecidos.*

Para encontrar este intervalo partiremos de la (5.8.12) y recordaremos la transformación de Box and Cox.

Así sustituyendo en la (5.8.12) los valores de y(λ,s) e y(λ,w) por los valores originales del proceso recordando las fórmulas de Box-Cox de estas transformaciones **tendremos que la probabilidad será la misma y, por otra parte, el intervalo se pondrá en relación con el proceso estocástico markoviano. →**

$$(y(s)^{\lambda}-1)/\lambda+\hat{\beta}(w-s)-t_{n-1}(\in/2)\tilde{\sigma}(\lambda)\sqrt{(w-s+(w-s)^2/(\sigma-t)}<(y(w)^{\lambda}-1)/\lambda<$$

$$(y(s)^{\lambda}-1)/\lambda+\hat{\beta}(w-s)+t_{n-1}(\in/2)\tilde{\sigma}(\lambda)\sqrt{(w-s+(w-s)^2/(\sigma-t)}$$

Luego multiplicando por λ pasando el menos -1, simplificando, y extrayendo la raíz λ el intervalo para el proceso original de un punto como valor original y(w) es:

$$\{(y(s)^{\lambda}+\lambda(\hat{\beta}(w-s)-t_{n-1}(\in/2)\tilde{\sigma}(\lambda)\sqrt{(w-s+(w-s)^2/(s-t))}\}^{1/\lambda}<y(w)$$

$$\{(y(s)^{\lambda}+\lambda(\hat{\beta}(w-s)+t_{n-1}(\in/2)\tilde{\sigma}(\lambda)\sqrt{(w-s+(w-s)^2/(s-t))}\}^{1/\lambda} \qquad (5.9.1)$$

Los extremos del intervalo son conocidos {excepto y(w)} nos permite construir el intervalo con una probabilidad fiducial de $1-\in$ de que contenga el valor individual del proceso y(w).

5.10. *Función característica de la transformación de Cox-Box para el proceso estocástico y(λ,s) conocido el valor de y(λ,t)*

Sabemos que la función característica es la **esperanza** de la variable aleatoria (en el punto s) de la expresión:

$$e^{iry(\lambda,s)}$$

siendo r el parámetro que normalmente en los tratados de Estadística se representa por t y aquí hacemos esta distinción sencilla de notación para evitar confusión. **Para un valor concreto del parámetro tiempo s, la variable aleatoria por las hipótesis iniciales es normal.**

Luego

$$E(e^{iry(\lambda,s)})=e^{iry(\lambda,t)+\beta(s-t)-r^2\sigma(\lambda)^2(s-t)/2} \qquad (5.10.1)$$

Analicemos esta expresión.

Conocidos el valor del proceso en t y(t) para s>t la función de densidad de probabilidad es la ya conocida y si en el eje de tiempos trazamos una perpendicular al mismo en el punto s tendremos el campo de la variable aleatoria tome el valor y(s) en la perpendicular, con su probabilidad elemental es la función de densidad (5.7.8) multiplicada por la diferencial de la variable aleatoria dy(λ,s).

Ahora bien cuando s⟶t la función característica toma precisamente el valor

$$e^{iry(\lambda,t)} \qquad (5.10.2)$$

El exponente no es sino la f.c. **de la distribución de un punto** es la certeza como hemos comentado en otras ocasiones puesto que hemos partido de la hipótesis que en el estado inicial en t conocemos el valor del proceso estocástico.

De la f. de d. observemos cuando sucede esto que si $s \rightarrow t$ la función de densidad tiende a la función δ de Dirac con la propiedad siguiente:

$$\int_{t-E}^{t+E} \delta(s,t)ds \qquad \text{entre } t-\in<s<t+\in \tag{5.10.3}$$

siendo $\in$ es un número muy pequeño (cualquiera que sea), la integral definida tiende a la unidad donde

$$_{\delta}(s,t) = \begin{cases} 0 \text{ para s distinto de t} \\ \infty \text{ para } s=t \end{cases} \tag{5.8.4}$$

es la conocida función δ de Dirac.

Las propiedades de esta función nos permite relacionar distribuciones discretas con distribuciones continuas.

VI TRANSFORMACIONES DE PROCESOS MARKOVIANOS DE DISTRIBUCION LOGARITMICO NORMAL

6.1. *Caso especial de la transformación de Box-Cox*

Hemos indicado que un caso especialmente importante en las aplicaciones Econométricas es cuando se estudian transformaciones logarítmicas.

Y decimos es importante por cuanto los coeficientes de estas transformaciones de las variables logarítmicas representan los conceptos de elasticidades tan importante en Economía. Pero justificamos este caso especial por su estudio separado aunque bien pudiera tratarse como un caso particular de las transformaciones de Box-Cos. Pero también y como anteriormente estudiaremos el proceso markoviano fundamentalmente de tipo wiener.

Esta transformación merece le dediquemos un breve capítulo por sus especiales propiedades pero insistimos todas sus fórmulas se obtienen hallando límites cuando $\lambda \rightarrow 0$ de la transformación de Box-Cox. Así si calculamos el Límite de $y(\lambda,s)$ cuando

$\lambda \Rightarrow 0$- fácilmente vemos que $y(0,s) = I_n y$.

Con carácter general estudiaremos las propieddes del proceso logarítmico transformado directamente aunque a veces recurramos a las fórmulas generales de transformación ya obtenidas y tomemos su límite para $\lambda \Rightarrow 0$.

6.2 *Función de densidad de probabilidad de un proceso markoviano logarítmico normal*

La función de densidad es

$$f(l_n y(s)) = \exp\{-(l_n y(s) - \ln y(t) - \beta(s-t))^2 / 2\sigma^2(t-s)\} / \sigma\sqrt{(2\pi(t-s))}y(s)) \qquad (6.2.1)$$

Por análisis de esta fórmula observamos que el proceso y(s) está en el denominador y a su vez el exponente de e su logaritmo neperiano. Por las característica especiales el proceso y(t) {t= $-\infty+\infty$) no puede ser negativo y(t) porque Ln(y(t) no es real.

Luego el proceso y(s) como la función de densidad es siempre no negativa y necesariamente la transformación logarítmica únicamente es viable cuando y(s)≥0.

Esta función de densidad es un caso concreto aplicado al proceso estocástico y(s) de la **función de densidad logarítmico normal, muy conocida en los textos de Estadística.**

Pero nosotros tratamos a y(s) como un proceso estocástico que sigue su logaritmo neperiano esa distribución normal y con parámetros semejantes de ß y de $\sigma(.0)$ con la perturbación μ_{ti}.

En adelante la notación $\sigma(.0)$ lo escribiremos simplemente σ y entenderemos σ es el lím. $\lambda \to$ de $\sigma(\lambda)$.

6.3. *Función característica y función generatriz de momentos del proceso logarítmico estocástico en el punto s cuando se conoce su valor en el punto t.(t 16).*

La función característica –cuya notación la representaremos por $\varphi(r)$ y la función generatriz de momentos por la notación g(r)– están muy relacionadas y se definen

$$\varphi(r)=E(\exp(irl_n y(s)) \qquad (6.3.1)$$

donde i es la unidad imaginaria ($\sqrt{-1}$)

$$g(r)=E\{\exp(rl_n y(s)\} \qquad (6.3.2)$$

y recordando que Iny(s) es normal N(Iny(t)+ß(s-t), σ)

$$\varphi(r)=\exp\{ir(I_ny(t)+\beta(s-t))-r^2\sigma^2(s-t)/2\} \qquad (6.3.3)$$

y la función generatriz de momentos es

$$g(r)=esp\{r(I_ny(t)+\beta(s-t))-r^2\sigma^2(s-t)/2\} \qquad (6.3.4)$$

Los momentos ordinarios y centrados del proceso logarítmico normal y(s) se obtienen simplemente hallando las derivadas sucesivas respecto al parámetro r e igualando a cero bien de la función g.m. bien de la f.c. Si se utiliza esta para calcular el momento condicionado centrado de orden k precisamos calcular las derivadas sucesivas hasta la k inclusive, hacer r=0 y dividir por i^k. Utilizando la función generatriz de momentos ordinarios no es preciso dividir la derivada por i^k.

6.4. *Momento ordinario de orden k del proceso markoviano y(s) cuando se conoce el valor del proceso en el punto t (t < s) si se cumplen las hipótesis para el proceso logarítmico normal.*

El momento ordinario de orden k del proceso estocástico y(s) es

$$E\{y(s)^k/y(t)\}=E\{\exp(kI_ny(s)\} \qquad (6.4.1)$$

Si comparamos esta definición con la (6.3.2) vemos que la (6.4.1) es un caso particular ya que el parámetro r toma ahora valores discontinuos porque k=0,1,2,3,..n...

Luego el momento de orden k del proceso estocástico en el punto s condicionado al conocimiento del valor del proceso en el punto t es igual a la 6.3.4 cuando r=k: no utilizaremos la notación condicionada por comodidad.

$$\alpha_k(y(s))=\exp(k(I_ny(t)+\beta(s-t))+k^2\sigma^2(s-t)/2\} \qquad (6.4.2)$$

$$=g(k) \qquad (6.4.2')$$

La (6.4.2) puede escribirse:

$$\alpha_{(k}(y(s)\}=y(t)^k\exp\{k\beta(s-t))+k^2\sigma^2(s-t)/2\} \qquad (6.4.3)$$

La (6.4.3) nos da todos los momentos condicionados del proceso estocástico logarítmico normal.

Dando valores a k=1,2 tendremos los dos primeros momentos del proceso markoviano cuya transformada sigue una distribución normal y que son:

$$\alpha\{(y(s)\}=y(t)\exp\{\beta(s-t))+\sigma^2(s-t)/2\} \qquad (6.4.4)$$

y

$$\alpha_2\{y(s)\}=y(t)^2\exp\{2ß(s-t))+2\sigma^2(s-t)\} \qquad (6.4.5)$$

Obtenidos estos momentos ordinarios con respecto al origen es inmediato el cálculo de la varianza del proceso y(s) también condicionada al conocimiento del valor que toma el proceso en el momento t

$$\mathrm{Var}\{y(s)\}=\alpha_2\{y(s)\}-\alpha\{y(s)\}^2=$$

$$=y(t)^2\exp\{2ß(s-t)+2\sigma^2(s-t)-y(t)^2\exp\{2ß(s-t)+\sigma^2(s-t)\}=$$

$$=y(t)^2\{\exp\{2ß(s-t)+\sigma^2(s-t)\}(\exp\{\sigma^2(s-t)\}-1$$

Luego la media poblacional y la varianza del proceso poblacional condicionadas son:

$$\alpha\{y(s)\}=y(t)\exp\{ß(s-t)+\sigma^2(s-t)/2 \qquad (6.4.4')$$

$$\mathrm{Var}\{y(s)\}=y(t)^2\{\exp\{2(ß(s-t)+\sigma^2(s-t)\}(\exp\{\sigma^2(s-t)\}-1 \qquad (6.4.5')$$

6.5. *Estimación de los parámetros poblacionales del proceso estocástico.*

Las estimaciones de ß y de σ^2 las podemos calcular por máxima verosimilitud o como caso particular de las obtenidas en la transformación de Box-Cox. Emplearemos este método.

Así para la estimación de $\hat{ß}$ si calculamos el límite cuando $\lambda \Rightarrow 0$ tenemos

$$\hat{\beta}=\frac{\Sigma(\mathrm{Ln}y(t_i)-\mathrm{Ln}y(t_{i-1}))}{(s-t)} \qquad (6.5.1)$$

La estimación de σ^2 es:

$$\hat{\sigma}2 = 1/n\ \Sigma\frac{\mathrm{Ln}y(t_i)-\mathrm{Ln}y(t_{i-1})}{t_i-t_{i-1}} \qquad (6.5.2)$$

6.6. *Distribuciones de los estimadores (6.5.1) y (6.5.2)*

1. El estimador $\hat{ß}$ es una combinación lineal de variables normales$\rightarrow$ **sigue una distribución normal**

$$\hat{ß}\rightarrow N(ß,\sigma^2/(s-t)\} \qquad (6.6.1)$$

2. El estimador $\hat{\sigma}^2$ está relacionado con la x^2 porque según vimos y sabemos por Estadística

$$\frac{n\hat{\sigma}2}{\sigma 2} \rightarrow X_{n-1}2 \qquad (6.6.2)$$

siendo la estimación de la varianza la fórmula (6.5.2)

De este estimador obtenemos el **estimador centrado de la varianza**

$$\tilde{\sigma}^2 = n\hat{\sigma}^2/(n-1) \qquad (6.6.3)$$

6.7. *Predictor del proceso logaritmo neperiano de y(w) respecto a su media poblacional, cuando desconocemos los parámetros poblacionales.*

Como siempre supondremos se cumplen las hipótesis establecidas para esta transformación logarítmica. Los datos observados del proceso han sido en los tiempos $t_1, t_2, \ldots t_n$ e incluso en el instante inicial t_0

Llamando al predictor de acuerdo con (8.8.2) y para este caso concreto

$$PL_n y(w) = L_n y(s) + \hat{\beta}(w-s) \qquad (6.7.1)$$

su esperanza matemática es

$$E\{PL_n y(w) = L_n y(s) + \beta(w-s) \qquad (6.7.2)$$

La expresión anterior también es equivalente a la E{Lny(w)} ya que ambas expresiones coinciden.

6.8. *Varianza del predictor del logaritmo neperiano del proceso estocástico respecto de media logarítmica poblacional del predictor.*

La desviación del Plny(w) respecto de (6.7.2) es

$$PLny(w) - Lny(s) - \beta(w-s) \qquad (6.8.1)$$

y sustituyendo (6.7.1) y calculando la varianza tenemos:

$$Var(PLny(w) = E\{Lny(s) + \hat{\beta}(w-s) - Lny(s) - \beta(w-s)\}^2 =$$

$$= E\{\hat{\beta} - \beta)(w-s)\}^2 = (w-s)^2 var(\hat{\beta}) \qquad (6.8.2)$$

y recordando el valor de la varianza del estimador $\hat{\beta}$ la varianza del predictor del logaritmo transformado respecto de la media poblacional es:

$$Var(PLny(w)\} = \sigma^2(w-s)2/(s-t) \qquad (6.8.2')$$

por ser la variable aleatoria en w del Pln(w) una distribución normal y cuyos parámetros nos son conocidos (6.8.1) y (6.8.2)

6.9. *Distribución del Plny(w). (Predictor del Ln del proceso respecto de su media teórica)*

Como hemos visto y por según el teorema de adición de combinación de variables normales independientes Plny(w) es normal y sus parámetros son

$$\text{Media } EPLny(w)=Lny(s)+\beta(w-s) \quad \text{y}$$

$$\text{Varianza}=VarPlny(w)=(w-s)^2\sigma^2/(s-t) \tag{6.9.1}$$

Luego la variable tipificada es también normal:

$$\frac{PLny(w) - Ln\text{-}\int(w-s)}{\sigma(w-s)/\sqrt{(s-t)}} \rightarrow N(0,1) \tag{6.9.2}$$

Igualmente de acuerdo con la (6.5.2) la x^2 es:

$$X_{n-1}2 = \frac{n\hat{\sigma}^2}{\sigma^2} \tag{6.9.3}$$

Y dividiendo la (6.9.2) por la raíz cuadrada de la media de la $Xn\text{-}1^2$ y recordando la (6.6.3) del estimador centrado de σ^2 formamos una t de Student con n-1 g.de l.

$$t_{n-1}=\frac{PLny(w) - (Lny(s)- \beta(w-s))}{\tilde{\sigma}^2(w-s)/(s-t)} \tag{6.9.4}$$

siendo $\tilde{\sigma}^2$ el estimador insesgado de σ^2

6.10 *Intervalo de confianza para la media poblacional Lny(ω) cuando se desconocen los parámetros poblacionales.*

Conocida la distribución del predictor del logaritmo neperiano del proceso estocástico y(w) con respecto de su media poblacional expresada por la relación (6.9.4) es sencillo determinar un intervalo fiducial para referida esperanza matemática cuando hemos elegido un nivel de significación.

La distribución de la t de Student con n-1 grados de libertad nos da para el nivel a de significación ε a dos bandas por la variable (6.9.4) si es un estimador se encuentra entre

$$P\{t_{n-1}(\in/2)<t_{n-1}<t_{n-1}(\in/2)\}=1-\in \qquad (6.10.2)$$

Sustituyendo en la anterior t_{n-1} por la (6.9.4) y despejando la media poblacional E{PLny(w)}=y(s)+ß(w-s) tenemos el intervalo que buscamos:

$$L_n y(s)+\hat{ß}(w-s)-t_{n-1}(\in/2)\tilde{\sigma}(w-s)/\sqrt{(s-t)}<E\{ELny(w)=$$

$$=y(s)+ß(w-s)<$$

$$Lny(s)+\hat{ß}(w-s)+t_{n-1}(\in/2)\tilde{\sigma}(w-s)/\sqrt{(s-t)} \qquad (6.10.3)$$

con probabilidad fiducial 1-ε que contenga al proceso estocástico Iny(w).

6.1.1. *Intervalo de confianza para el proceso y(w) que se relaciona con el valor medio del logaritmo neperiano de su transformado E{Lny(w)}.*

Denominaremos Ry(w) al valor del proceso estocástico que en w corresponde biunívocamente con el de la E{Iny(w)} y que este valor pertenece a una línea de regresión en el sentido de la transformación que hemos indicado.

Así poniendo por exponente del número e los miembros de la (6.1.10) y simplificando tenemos la siguiente desigualdad:

$$y(s)\exp\{\hat{ß}(w-s)-t_{n-1}(\in/2)\tilde{\sigma}(w-s)/\sqrt{(s-t)}\}<Ry(w)$$

$$<y(s)\exp\{\hat{ß}(w-s)+t_{n-1}(\varepsilon/2)\sigma(w-s)/\sqrt{(s-t)} \qquad (6.11.1)$$

La (6.11.1) podemos obtenerla directamente a partir del intervalo general deducido de la transformación de Box-Cox en (5.8.11) la cual depende del parámetro λ cuando λ→0.

Llamando L al límite del logaritmo neperiano de referido intervalo y calculando el límite tenemos:

$$L=\mathrm{Lim}1/\lambda\ \{Lny(s)\lambda+\lambda(\hat{ß}(w-s)\pm tn\text{-}1(\in/2)\tilde{\sigma}(\lambda)(w-s)/\sqrt{(s-t)}$$

Esta indeterminación la resolvemos aplicando L'Hopital

Luego

$$L=\mathrm{Lim}.y(s)^{\lambda}L_n y(s)+\hat{ß}(w-s)\ -t_{n-1}(\in/2)\tilde{\sigma}(\lambda)(w-s)/\sqrt{(s-t)} \qquad (6.11.2)$$

y llamando σ al Lim. $\sigma(\lambda)$ cuando $\lambda \to 0$ tenemos el Logaritmo neperiano de la (6.11.1) donde el antilogaritmo es la (6.11.1) recordando el valor del Lím $y(s)^{\lambda}$ cuando $\lambda \to 0$ es la unidad.

Es decir que de las expresiones generales encontradas en 5.8 para la transformación de Box-Cox deducimos todas las transformaciones de este capítulo simplemente calculando el límite $\lambda \to 0$

6.12. *Intervalo del Predictor del Logaritmo neperiano del proceso estocástico que contenga un valor individual del logaritmo neperiano del proceso estocástico y(w) cuando estimamos sus parámetros.*

1. Por las (6.7.1) y (6.7.2) tenemos el predictor del logaritmo neperiano del proceso estocástico en el punto w y su correspondiente esperanza matemática.

2. Ahora nos interesa relacionar el predictor con el correspondiente logaritmo neperiano del proceso que tomará en el punto w del eje de tiempos cuando (como hasta ahora hemos supuesto) conocemos el valor del proceso que toma en el momento s.

Por ser

$$Lny(w)=Lny(s)+\beta(w-s)+u_w \tag{6.12.1}$$

la esperanza matemática de la (6.12.1) coincide con la (6.7.2). Así la esperanza matemática de

$$E\{L_ny(W)-PL_ny(w)\}^2=E\{L_ny(s)+\beta(w-s)+u_w-Lny(s)-\hat{\beta}(w-s)\}^2=$$

$$E\{u_w-(\hat{\beta}-\beta(w-s)\}^2=Var(u_w)+(w-s)^2Var(\hat{\beta}) \tag{6.12.2}$$

siendo la esperanza del doble producto nula.

Según la (5.8.9) y la varianza de la perturbación del proceso en el punto w cuando se conoce el valor del proceso en el punto s, por hipótesis, tenemos:

$$Var(PiLny(w)\}=\sigma^2(w-s)+(w-s)^2\sigma^2/(s-t) \tag{6.12.3}$$

3. Distribución de la diferencia del Iny (w) y el predictor Lny(w).

Para determinar esta distribución conocemos que las esperanzas son nulas y como por hipótesis son normales el proceso Iny(w) y, a su vez, el predictor está formado por combinación lineal de variables normales, la distribución de este también es normal. Luego esta diferencia es normal de media nula y varianza la (6.12.3).

Tipificando esta expresión tendremos una normal N(0,1) que es:

$$\frac{Lny(w) - Plny\,(w)}{\sigma\sqrt{(w-s + (w-s)\,^2/(s-t))}} \qquad (6.12.4)$$

Sustituyendo el numerador Plny(w) por su valor tenemos:

$$\frac{Lny(w) - Lny\,(s) - \hat{B}\,(w-s)}{\sigma\sqrt{(w-s + (w-s)\,^2/(s-t))}} \longrightarrow N(0,1)$$

Por otra parte sabemos que

$$X^2_{n-1} = \frac{n\hat{\sigma}^2}{\sigma^2} \qquad (6.12.5)$$

es una x^2 con n-1 g.de l. Luego la anterior dividida por la raíz cuadrada de la media de la — nos da una t de Student por ser ambas variables independientes.

Por la (6.6.2) $\tilde{\sigma}\,^2$ es un estimador centrado de $\sigma\,^2$

Es bien sabido por definición que la t de Student se forman por el cociente entre dos variables independientes: el numerador una normal tipificada y el denominador la raíz cuadrada de una media de una X^2. Si dividimos la normal tipificada por la raíz cuadrada de la media de la x^2 de Pearson y después de simplificar por las hipótesis establecidas tenemos:

$$t_{n-1} = \frac{Lny(w) - Lny\,(s) - \hat{\beta}(w-s)}{\tilde{\sigma}\sqrt{(w-s + (w-s)\,^2/(s-t))}} \qquad (6.12.6)$$

que es la distribución del logaritmo neperiano del proceso estocástico individual cuando conocemos el valor del proceso en el punto s desconocemos los valores de los parámetros poblacionales.

4. Intervalo de confianza para el logaritmo del proceso estocástico en el punto w cuando se cumplen las hipótesis establecidas de normalidad.

Es inmediata la obtención del Lny(w) a partir de la (6.11.6) y hemos estimado los valores de los parámetros poblacionales. Para ello tendremos que elegir el nivel de significación correspondiente á a dos bandas y que la (6.11.6) esté comprendido entre $-t_{n-1}(\varepsilon/2)$ y $+t_{n-1}(\varepsilon/2)$ es equivalente a despejar el Lny(w) –valor individual del proceso en el eje de tiempos en el punto w– con un nivel de confianza 1-ε para que el proceso estocástico Iny(w) caiga dentro del intervalo formado por la siguiente expresión

$$Lny(s)+\hat{ß}(w-s)-t_{n-1}(\in/2)\tilde{\sigma}\sqrt{\{w-s+(w-s)^2/(s-t)\}}<Lny(w)$$

$$<Lny(s)+\hat{ß}(w-s)+t_{n-1}(\in/2)\tilde{s}\sqrt{\sqrt{\{w-s)2/(s-t)\}}} \qquad (6.12.7)$$

con una probabilidad fiduycial de 1-∈ de que contenga el verdadero valor individual del proceso estocástico Lny(w).

6.13. *Intervalo para un punto individual del proceso estocástico y(w) cuando se conoce el valor del proceso en el punto s y se desconocen los parámetros del proceso y los logaritmos neperianos cumplen las hipótesis establecidas.*

De la (6.12.7) se deduce fácilmente el intervalo de aceptación de y(w). Tenemos símplemente que hallar antilogaritmos neperianos o su equivalente que es la expresión anterior pero siendo el exponente del número e.

Así pues este intervalo para el valor individual del proceso estocástico markoviano que siga las hipótesis establecidas simplificando los valores es:

$$y(s)\exp\{\hat{ß}(w-s)-t_{n-1}(\in/2)\tilde{\sigma}\sqrt{[w-s+(w-s)^2/(t-s)]}\}<$$

$$<y(w)<$$

$$y(s)\exp\{\hat{ß}(w-s)+tn-1(\in/2)\tilde{\sigma}\sqrt{[w-s+(w-s)\in/(t-s)]}\} \qquad (6.13.1)$$

Este intervalo de la aceptación de un valor del proceso estocástico en el punto futuro del eje de tiempos w cuando se conoce la trayectoria pasada de t a s y a partir de los valores del proceso en **esa trayectoria se han estimado los parámetros poblacionales desconocidos.**

Si se cumplen las hipótesis este valor del proceso estocástico contenido dentro del intervalo obtenido para el valor individual en el tiempo futuro w **si utilizamos la transformación general de Box-Cos** es un caso particular si calculamos su límite (que nos da una expresión indeterminada pero aplicando L'Hopital deshacemos la indeterminación) y el exponente del número e es el límite hallado y que coincide con la expresión precedente.

VII CASOS ESPECIALES DE LOS PROCESOS MARKOVIANOS CUANDO EL PARAMETRO T ES ENTERO (TRANSFORMACION BOX and COX)

7.1. *Introducción general a los casos especiales de las transformaciones de Box-Cox cuando el parámetro t=0,1,2,3, ..n del conjunto índice es discreta y sigue una sucesión numerable.*

Hasta ahora nos hemos detenido en fórmulas más generales pero seguramente no muy utilizables en Econometría porque los datos de las series temporales vienen representados por series discretas y los tiempos son equidistantes. En este breve capítulo resumiremos las fórmulas aplicables especialmente a observaciones para la sucesión estocástica está definida sobre el conjunto índide T={0,1,2,3, ...n...)

Mas concretamente nuestra aplicación sobre procesos markovianos o concretamente cadena de Markov son casos particulares de los estudiado hasta el presente. Y para no repetir demostraciones, **solamente haremos recopilar las fórmulas condensadas cuando el tiempo venga expresado en la forma indicada.**

7.2. *Valores del proceso estocástico discreto transformado*

Este valor del proceso de parámetro discreto transformado markoviano es de la forma (cuando aplicamos la transformación de Box-Cox) como la siguiente:

$$y(\lambda,j)=(y(j)^{\lambda}-1)/\lambda \qquad (7.2.1)$$

$$(j=0,1,2......n)$$

es decir tenemos n+1 valores del proceso {y(j)} = {j=0,1,...n) y de su transformado.

7.3. *Hipótesis sobre las perturbaciones*

Las hipótesis sobre este proceso son semejantes a las estudiadas en la transformación de Box-Cox pero como estas perturbaciones eran condicionadas a la diferencia de tiempos de valores consecutivos, en nuestro caso esta diferencia es la unidad y el proceso en el tiempo j conocido su valor de j-1 se relaciona por la siguiente expresión:

$$y(\lambda,j)=y(\lambda,j-1)+\beta+u_j \qquad (7.3.1)$$

con las hipótesis para μ_j para las perturbaciones:

1ª La esperanza de μ_j es nula.

$$Eu_j=0 \text{ para todo } j \qquad (7.3.2)$$

2ª Incorrelación

$$Eu_j u_k = 0 \quad j \neq k \tag{7.3.3}$$

3ª Homoscedasticidad

$$Eu_i^2 = \sigma^2 \text{ Para todo i} \tag{7.3.4}$$

4ª Normalidad

$$u_i \Rightarrow N(0,\sigma)$$

7.5. *Función de densidad del proceso transformado en j conocido si valor en j-l(j=1,2,...n).*

Por las hipótesis planteadas y conocido el valor del proceso en el punto j-l, la función de densidad de $y(\lambda,j)$ es

$$f(y(\lambda,j)/y(\lambda,j-1)=$$

$$\exp\{-(y(\lambda,j)-y(\lambda,j-1)-\beta)^2/2\sigma(\lambda)^2\}]/(\sigma(\lambda)\sqrt{2\pi}) \qquad ((7.5.1)$$

Esta función de densidad es condicionada a la aparición del proceso transformado en j-1. La representaremos por $f(\lambda,j)$ pero advertimos su naturaleza para que no ofrezca ninguna duda.

En consecuencia si en lugar de conocer el proceso en el punto j-l lo conocemos en h (siendo $h < j$) la varianza no será la misma pues ahora depende de la diferencia de tiempos y ß e igualmente de la misma diferencia.

Esta probabilidad condicionada es

$$f(y(\lambda,j)/y(\lambda,h)=$$

$$=[\exp.\{-(y(\lambda,j)-(y(\lambda,h)-\beta(j-h)\}/2\sigma^2(j-h)]/(s\sqrt{2\pi}(j-h)$$

$$\textbf{para todo j>h} \tag{7.5.2}$$

Obsérvese que la anterior (7.5.1) es un caso particular cuando h=j-1.

7.6. *Forma del proceso transformado en función del conocido en referido punto. Esperanza, varianza y función característica del proceso transformado y(λ,j) cuando conocemos el proceso en un punto anterior h.*

1. Forma del proceso transformado en función del proceso en un punto anterior –no consecutivo–.

De las hipótesis iniciales tenemos:

$$y(\lambda,j)=y(\lambda,j\text{-}1)+u_j$$

$$j=h+1,h+2,\ldots n \qquad (7.6.1)$$

Dando valores a j y sumando y simplificando tenemos que la representación del proceso transformado en $y(\lambda,n)$ puede escribirse también

$$y(\lambda,n)=y(\lambda,h)+\beta(n\text{-}h)+u_{h+1}+u_{h+2}+\ldots u_n \qquad (7.6.2)$$

Un caso particular es cuando h=n-1. La (7.6.2) se convierte en

$$y(\lambda,n)=y(\lambda,h)+\beta+u_n \qquad (7.6.3)$$

2. Esperanza matemática condicionada.

La esperanza de la (7.6.2) es

$$E(y(\lambda,n)\}=y(\lambda,h)+\beta(n\text{-}h) \qquad (7.6.4)$$

3. Varianza matemática condicionada.

Y la varianza (cuando conozcamos el valor del proceso en h) es

$$Var[y(\lambda,n)/y(\lambda,h)]=E[u_{k+1}+\ldots u_n]^2=E\sigma^2X_{n\text{-}h}{}^2=$$

$$\sigma^2EX_{n\text{-}h}{}^2=\sigma^2(n=h) \qquad (7.6.5)$$

por haber n-h sumandos y por hipótesis de incorrelación entre μ_j y μ_h.

Esta varianza la representaremos usualmente sin la condicionante que se supone implícita. Así la varianza es

$$Var[y(\lambda,n)]=(n\text{-}h)\sigma^2 \qquad (7.6.5)$$

Un caso particular es cuando h=n-1. La (7.6.3) se convierte en ß2 es decir la hipótesis de partica para el proceso transformado porque consideramos una cadena consecutiva de una sucesión numerable que Markov.

4. Función característica

Esta función característica condicionada simplemente es

$$\varphi(r)+E(\exp(iry(l,j)=\exp.\{iry(\lambda,h)+\beta(j\text{-}h)\text{-}r^2\sigma^2(j\text{-}h)/2 \qquad (7.6.1)$$

y lo mismo podemos determinar la función generatriz de los momentos ordinarios condicionados.

Cuando j=n tenemos la función característica en el punto n conocido el proceso transformado en el punto h.

7.7. *Estimaciones por punto de los parámetros*

Las fórmulas de las estimaciones por punto son casos particulares de los generales obtenidos y así para

1. Estimador $\hat{\beta}$

$$\hat{\beta}= \frac{\sum_{1}^{n}(y(\lambda,j)-y(\lambda,j-1))}{n} \tag{7.7.1}$$

$$j=1,2,3,...n$$

La expresión anterior puede reducirse después de simplificar términos iguales pero de signo contrario a

$$\hat{\beta}=\frac{y(\lambda,n)-y(\lambda,0)}{n} \tag{7.7.2}$$

Esta fórmula es un caso particular de la general deducida en (4.4.1) porque ahora los valores del proceso toman una secuencia del eje de tiempos de intervalos unitarios siendo s=n y t=0.

2. Estimador $\hat{\sigma}(\lambda)^2$

Este estimador lo obtenemos de la (4.4.2) haciendo la diferencia de tiempos igual a la unidad. Así observemos que $\hat{\beta}$ no viene acompañado de esta diferencia de tiempos ni tampoco el denominador.

$$\hat{\sigma}(\lambda)^2= \frac{\sum_{1}^{n}[(y(\lambda,j)-y(\lambda,j-1)-\hat{\beta})^2]}{n} \tag{7.7.3}$$

Esta fórmula la deducimos por máxima verosimilitud recordando ahora las observaciones indicadas respecto a la diferencia de tiempos.

3. Estimador $\tilde{\sigma}(\lambda)^2$ centrado

De la anterior deduciremos el estimador centrado para $\hat{\sigma}(\lambda)^2$ según las (4.5.6) y (4.5.6b).

Así para nuestro caso concreto

$$\sigma(\lambda)^2 = \frac{\sum_{1}^{n}[(y(\lambda,j)-y(\lambda,j-1)-\hat{\beta})^2]}{n-1} \qquad (7.7.4)$$

$$(j=1,2,3...n)$$

7.8. *Varianzas de estos estimadores*

1. Varianza de $\hat{\beta}$

$$\mathrm{Var}\hat{\beta} = \frac{\sigma(\lambda)^2}{n} \qquad (7.8.1)$$

y no es sino un caso particular de la (4.5.3)

2. Varianza de $\hat{\sigma}(\lambda)^2$

El estimador (7.7.3) no es centrado porque tampoco lo era el estimdor (4.5.4) y su esperanza es igualmente:

$$E\hat{\sigma}(\lambda)^2=\sigma(\lambda)^2(1-1/n) \qquad (7.8.2)$$

que vemos es asintóticamente centrado.

Y la varianza es idéntica a la (4.6.5)

$$\mathrm{Var}\hat{\sigma}^2(\lambda)^2=2(n-1)\sigma(\lambda)^4/n^2 \qquad (7.8.3)$$

3. Varianza del estimador centrado de la varianza $\tilde{\sigma}(\lambda)^2$

Coincide con la (4.6.6) aunque el estadístico $\tilde{\sigma}(\lambda)^2$ es el (7.7.4)

$$\mathrm{Var}(\tilde{\sigma}(\lambda)^2=2\sigma(\lambda)^4/(n-1) \qquad (7.8.4)$$

7.9. *Distribuciones*

Distribución de $\hat{\beta}$

La distribución del estimador $\hat{\beta}$ es normal

$$\hat{\beta}\text{-}N(\beta,\sigma(\lambda)/n) \tag{7.9.1}$$

Luego la de su variable tipificada es

$$\frac{\hat{\beta}-\beta}{\sigma(\lambda)/\sqrt{n}}\text{-}N(0,1) \tag{7.9.2}$$

2. Distribución de la variable $(\hat{\beta}-\beta)/\tilde{\sigma}(\lambda):\sqrt{n}$

La variable $n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2$ es una X^2_{n-1} (7.9.3)

Formemos una t de Student con n-1 g.de l. dividiendo la (7.9.2) por la raíz cuadrada de la media de la (7.9.30 –al dividir por n-1–

$$t_{n-1}=\frac{(\hat{\beta}-\beta)/(\sigma(\lambda)/\sqrt{n})}{\sqrt{n\hat{\sigma}(\lambda)^2/\sigma(\lambda)^2/n-1}} \tag{7.9.4}$$

simplificando la anterior y poniendo la t de Student en función del estimador centrado de la varianza, la (7.9.4) se reduce:

$$t_{n-1}=\frac{\hat{\beta}-\beta}{\tilde{\sigma}(\lambda)/\sqrt{n}} \tag{7.9.5}$$

fórmula que nos permite obtener intervalos de confianza para el estimador del parámetro poblacional ß.

3. Distribución del predictor transformado respecto de la media poblacional transformada.

Llamando al predictor respecto de la media poblacional Py(λ,w) siendo w un número entero (w>n) tenemos

$$Py(\lambda,w)=y(\lambda,n)+\hat{\beta}(w-n) \tag{7.9.6}$$

siendo su esperanza

$$E(Py(\lambda,w)\}=y(\lambda,n)+\text{ß}(w-n) \tag{7.9.7}$$

y como el proceso transformado en w es

$$y(\lambda,w)=y(\lambda,n)+\text{ß}(w-n)+\sum_{n+1}^{w}u_j$$
$$(j=n+1,n+2,\ldots w) \tag{7.9.8}$$

también la esperanza de (y.9.8) coincide con la (7.9.7).

Varianza del predictor del proceso transformado respecto de su media.

$$Var\{Py(\lambda,w)\}=E\{y(\lambda,n)+\hat{\text{ß}}(w-n)-(y(\lambda,n)+\text{ß}(w-n)\}^2=$$

$$=E\{((\hat{\text{ß}}-\text{ß})(w-n))^2\}=(w-n)^2Var(\hat{\text{ß}})$$

y recordando la (7.8.1) tenemos:

$$Var\{Py(\lambda,w)=\sigma(\lambda)^2(w-n)^2/n \tag{7.9.9}$$

El predictor del proceso transformado por ser combinación lineal de variables normales, es normal con media teórica (7.9.7) y varianza la (7.9.9).

La variable normal tipificada del predictor respecto a su media poblacional es

$$\frac{Py(\lambda,w)-(y(\lambda,n)+\text{ß}(w-n))}{\sigma(l)(w-n)/\sqrt{n}}\ N(0,1)$$

Por otra parte

$$n\sigma(\lambda)2/\sigma(\lambda)2\Rightarrow x_{n-1}^{\ 2}$$

Luego formando una t de Student con n-1 grados de libertad tenemos:

$$t_{n-1}=\frac{Py(\lambda,w)-(y(\lambda,n)+\text{ß}(w-n))}{\sigma(\lambda)(w-n)/\sqrt{n}}:\sqrt{n\hat{\sigma}(\lambda)^{(2)^-}/\sigma(\lambda)^{(2)^-}(n-1)}$$

que simplificando y poniendo en función del estimdor centrado $\sigma(\lambda)^2$

$$t_{n-1}=\frac{Py(\lambda,w)-(y(\lambda,n)+\text{ß}(w-n))}{\tilde{\sigma}(\lambda)(w-n)/\sqrt{n}} \tag{7.9.10}$$

Sustituyendo el predictor del proceso transformado por su valor (7.9.6) tenemos

$$t_{n-1}=\frac{y(\lambda,n)+\hat{\beta}(w-n)-(y(\lambda,n)+\beta(w-n))}{\hat{\sigma}(\lambda)(w-n)/\sqrt{n}} \tag{7.9.11}$$

o también

$$t_{n-1}=\frac{Py(\lambda,w)-E\{py(\lambda,w)\}}{\hat{\sigma}(\lambda)(w-n)/\sqrt{n}} \tag{7.9.12}$$

donde $py(\lambda,w)$ es conocido.

4. Distribución del predictor $Py(\lambda,\omega)$ con respecto a su valor individual.

Por ser ambas esperanzas iguales {según las (7.9.7) y las (7.9.8)} la varianza es

$$VarPry(\lambda,w)/y(\lambda,w)\}=E\{Pry(\lambda,w)-y(\lambda,w)=$$

$$\sigma^2(\lambda)(w-n+(w-n)^2/n\} \tag{7.9.13}$$

porque

$$E[y(\lambda,w)-y(\lambda,n)\}^2]=\sigma(\lambda)^2(w-n)$$

y recordando la (7.7.9) y la (7.6.5) que es la varianza respecto de la media.

Formemos la variable

$$\frac{y(\lambda,n)+\hat{\beta}(w-n)-y(\lambda,w)}{\sigma(\lambda)\sqrt{w-n+(w-n)^2/n}}$$

Esta variable es una normal tipificada por ser combinación lineal de variables normales y, además, su esperanza matemática es nula y su varianza la unidad según las (7.9.6), (7.9.7) y (7.9.13).

Formemos una t de Student con n-1 g. de l.

$$t_{n-1}=\frac{y(\lambda,n)+\hat{\beta}(w-n)-y(\lambda,w)}{\hat{\sigma}(\lambda)\sqrt{(w-n+(w-n)^2)/n}} \tag{7.9.14}$$

donde hemos simplificado y hemos sustituido la estimación de la varianza por su estimador centrado.

7.10. *Estimación por intervalo*

Estudiaremos al mismo nivel de significación á (a dos bandas) los estimadores por intervalos de los procesos transformados respecto a sus medias transformadas teóricas y también con el intervalo de la predicción que contenga el valor del proceso individual transformado cuando no ha existido cambio de estructura con el tiempo.

1. Estimación por intervalo de la media teórica del proceso transformado cuando se conoce el predictor y la estimación de la varianza centrada.

Para que la (7.9.12) caiga entre $-t_{n-1}(\epsilon/2)$ y $+t_{n-1}(\epsilon/2)$ donde ε es el nivel de significación elegido es lo mismo que la **media teórica del proceso transformado** caiga en el intervalo de aceptación con probabilidad fiducial $1-\epsilon$.

$$y(\lambda,n)+\hat{\beta}(w-n)-t_{n-1}(\epsilon/2)\tilde{\sigma}(\lambda)(w-n)/\sqrt{n}<$$

$$<E\{y(\lambda,w)=y(\lambda,n)+\beta(w-n)<$$

$$<y(\lambda,n)+\hat{\beta}(w-n)+t_{n-1}(\epsilon/2)\tilde{s}(\lambda)(w-n)/\sqrt{n} \qquad (7.10.1)$$

2. Intervalo de aceptación para el predictor que contenga un valor individual del proceso transformado en w cuando se conoce su estimación hasta n y la varianza centrada.

De la (7.9.14) formaremos el intervalo pero en este caso el intervalo que contenga el valor individual transformado es

$$(\lambda,n)+\hat{\beta}(w-n)+t_{n-1}(\epsilon/2)\tilde{\sigma}(\lambda)\sqrt{w-n+(w-n)^2/n}<$$

$$<y(\lambda,w)<$$

$$(\lambda,n)+\hat{\beta}(w-n)+t_{n-1}(\epsilon/2)\tilde{s}(\lambda)\sqrt{w-n+(w-n)^2/n} \qquad (7.10.2)$$

3. Intervalo de aceptación para el proceso original en w correspondiente al transformado promedio, conocidos el proceso en el punto n y los valores estimados de los parámetros y de la varianza centrada.

Llamaremos Ry(w) a este valor por ser el correspondiente en w de la esperanza matemática del proceso transformado.

Sustituyendo $y(\lambda,w)=(y(w)^{\lambda}-1)\lambda$ en la (7.10.1) y simplificando tenemos

$$y(w)^{\lambda}+\lambda[\hat{\beta}(w-n)-t_{n-1}(\in/2)\tilde{\sigma}(\lambda)(w-n)/\sqrt{n})]<$$

$$<Py(w)^{\lambda}<$$

$$y(w)^{\lambda}+\lambda[\hat{\beta}(w-n)+t_{n-1}(\in/2)\tilde{s}(\lambda)(w-n)/\sqrt{n})] \qquad (7.10.1)$$

de la que deducimos:

$$y(w)^{\lambda}+\lambda\{\hat{\beta}(w-n)-t_{n-1}(\in/2)\tilde{\sigma}(\lambda)(w-n)/\sqrt{n}\}]^{1/\lambda}<Py(w)<$$

$$y(w)^{\lambda}+\lambda\{\hat{\beta}(w-n)-t_{n-1}(\in/2)\tilde{\sigma}(\lambda)(w-n)/\sqrt{n}\}]^{1/\lambda} \qquad (7.10.3)$$

que nos da el intervalo de confianza para el proceso y(w) del punto de regresión Py(w) de la predicción media.

4. Intervalo del predictor que contiene un valor individual del proceso conocidos los parámetros estimados y el último punto del proceso estocástico en n.

De la (7.10.2) podemos deducir el intervalo para el valor individual del proceso en el punto w.

Sustituyendo el valor transformado por el del proceso tendremos para el intervalo que contenga el valor individual del proceso y(w).

$$[y(w)^{\lambda}+\lambda\{(\hat{\beta}-\beta)(w-n)-t_{n-1}(\in/2)\tilde{\sigma}(\lambda)\sqrt{w-n+(w-n)^2/n}\}]^{1/\lambda}<$$

$$<y(w)<$$

$$[y(w)^{\lambda}+\lambda\{(\hat{\beta}-\beta)(w-n)+tn-1(\in/2)\tilde{\sigma}(\lambda)\sqrt{w-n+(w-n)^2/n}\}]^{1/\lambda}$$

(7.10.4)

5. Intervalo del proceso logarítmico normal respecto de su esperanza matemática, cuando hemos estimado los valores poblacionales de ß y de $\sigma(\lambda)^2$

La transformación en este caso ¬¤ 0 y la fórmula (7.10.1) se transforma en

$$Lny(n)+\hat{\beta}(w-n)-t_{n-1}(\in/2)\tilde{\sigma}(w-n)/\sqrt{n} < E\{Lny(w) <$$

$$Lny(n)+\hat{\beta}(w-n)+t_{n-1}(\in/2)\tilde{\sigma}(w-n)/\sqrt{n} \qquad (7.10.5)$$

Hemos puesto Lím $\hat{\sigma}(\lambda) = \hat{\sigma}$ cuando $\lambda \to 0$ y Lím $y(\lambda,w) = Lny(w)$

5. Intervalo de aceptación del proceso logarítmico transformado que contenga un valor individual del Lny(w) cuando se han estimado los valores de los parámetros y el de la varianza centrada.

Este intervalo lo obtendremos de la (7.10.2) *calculando límites para* $\lambda \to 0$.
Así calculando límites de la (7.10.2) tenemos:

$$Lny(n)+\hat{\beta}(w-n)-t_{n-1}(\epsilon/2)\tilde{\sigma}\sqrt{w-n+(w-n)^2/n} < Lny(w) <$$

$$Lny(n)+\hat{\beta}(w-n)+t_{n-1}(\epsilon/2)\tilde{\sigma}\sqrt{w-n+(w-n)^2/n} \qquad (7.10.6)$$

donde σ es la raíz cuadrada del estimador centrado de la varianza.

7. Intervalo para la regresión del proceso en y(w) cuando el logaritmo neperiano del proceso sigue una distribución normal y cuyos parámetros se han estimado y la varianza es centrada.

Este intervalo lo obtenemos de la (7.10.3) cuando $\lambda \to 0$ y en consecuencia efectuando operaciones, por ser una expresión indeterminada, aplicando L'hopital (después de tomar logaritmos neperianos límites y volviendo a números) tenemos

$$y(n)\exp\{(\hat{\beta}(w-n)-t_{n-1}(\epsilon/2)\tilde{\sigma}(w-n)/\sqrt{n})\} < Py(w) <$$

$$y(n)\exp\{(\hat{\beta}(w-n)+t_{n-1}(\epsilon/2)\tilde{\sigma}(w-n)/\sqrt{n})\} \qquad (7.10.7)]$$

con un nivel de confianza de $1-\epsilon$ de que contenga el valor de la regresión correspondiente a la ELn(y(w).

8. Intervalo del proceso logarítmico que contiene un valor individual del proceso y(w) cuando hemos estimado los parámetros $\hat{\beta}$ y el estimador centrado de la varianza $\sigma(\lambda)^2$

Este caso lo obtendremos del límite de la (7.10.4) cuando $\lambda \to 0$ por ser indeterminado, tomaremos logaritmos, calcularemos el límite y después pasaremos a antilogaritmos y nos queda, por sencillos cálculos la siguiente fórmula para el intervalo de aceptación de la predicción individual con un nivel de significación ϵ

$$y(n)\exp\{(\hat{\beta}(w-n)-t_{n-1}(\epsilon/2)\tilde{\sigma}\sqrt{w-n+(w-n)^2/n} <$$

$$< y(w) <$$

$$y(n)\exp\{(\hat{\beta}(w-n)+t_{n-1}(\epsilon/2)\tilde{\sigma}\sqrt{w-n+(w-n)^2/n} \qquad (7.10.8)$$

Las fórmulas (7.10.7) y (7.10.8) son muy útiles para predecir cuándo se utilizan el proceso logarítmico y sigue las hipótesis establecidas y hemos de distinguir dos casos bien distintos: El primero es que estas fórmulas de intervalos de confianza nos permiten determinar con un nivel de confianza 1-ε el intervalo correspondiente a la media poblacional proceso estocástico y el segundo cuando deseamos un valor individual del proceso futuro en el punto w.

CONCLUSIONES

1ª. Nuestro estudio ha sido enfocado para transformar los procesos Markovianos en otros procesos que gocen de las propiedades siguientes:

a) Posibilidad de que el proceso transformado **sea linealizable.**

b) Tenga una distribución aproximadamente normal de **media y varianza constante o más concretamente es una constante multiplicada por la diferencia de tiempos.**

c) Estimar los parámetros contrastando las hipótesis de si la distribución es normal y homocedástico aplicando los test usuales conocidos en Estadística y Econometría.

2ª. Si conseguimos este objetivo **deshacer la transformación** y calcular:

a) Estimadores y estimaciones **puntuales** del proceso original.

b) Establecer **intervalos de confianza** y

c) Utilizar los estimadores con fines **predictivos**

3ª. En todo el estudio que hemos realizado pueden introducirse cuantas **variables exógenas consideremos** necesarias por lo que esto solamente influye en los grados de libertad.

4ª. Las **aplicaciones** son numerosas porque un caso particular del expuesto está la transformación logarítmica y que puede utilizarse en diversas aplicaciones prácticas como también las otras transformaciones.

No mencionamos pero ya pensamos en la posibilidad de aplicar esta técnica :

a) Mercados de **opciones.**

b) Mercados de **futuros** que funcionan en E.U. y Europa y que en España no tenemos tan alto grado de especialización.

Si hacemos una transformación y la variable transformada la ponemos en forma

$$y^*(s)=y^*(t)+\beta_1(s-t)+\beta_2x_2+u_t$$

tenemos un modelo con una variable exógena y que podemos analizar que no es sino la (3.1.1) incluyendo el término de esta variable exógena. Particularmente es interesante cuando la diferencia de tiempos t_k-t_{k-1} de las observaciones es constante. Si la variable x representa cotizaciones por ejemplo peseta dólar, y la variable original representa la cotización de futuro de una mercancía estamos en presencia de los mercados de futuro.

Como puede apreciarse en estas teorías juega un papel sumamente importante el **predictor.**