Introducción a la teoría de procesos estocásticos

Por

Doctor Fco. JAVIER URBELZ IBARROLA

Catedrático de Estadística

Este trabajo está dedicado a un conocimiento elemental de los procesos estocásticos y fundamentalmente trato de los procesos estocásticos estacionarios de parámetro t discreto (definido sobre un conjunto índice cuyos elementos están en progresión aritmética) o de parámetro t continuo.

En el primer capítulo comentamos brevemente la evolución histórica de las series cronológicas consideradas como trayectorias sin basarse en conceptos de probabilidad y después, con Wiener, Slusky, Wold, etc., empiezan a iniciarse los modelos de procesos basados en la Teoría de Probabilidades, según la Axiomática de Kolmogorov.

En la historia ha influido especialmente el análisis armónico, funciones de correlación, covarianza, ergodicidad, etc., cuestiones que perfeccionan los conceptos de procesos estocásticos. Finalizo este capítulo primero con unas ideas generales sobre su estado actual y las representaciones espectrales de los procesos.

La importancia del estudio de las series temporales en Economía, actuarial, etc., es tal que huelga todo comentario.

El segundo capítulo lo dedico íntegramente a definiciones de las distintas clases de espacios abstractos de Hilbert (1).

Las series cronológicas estudiadas en libros de Estadística de más de cuarenta años, ocupaban la mayor parte de su contenido. No obstante su

(1) F. J. Urbelz: «Aplicaciones del espacio de Hilbert a la estadística». Anales del Instituto de Actuarios, 1976.

importancia, las personas que no deseen conocer su historia y tengan conocimientos del espacio de Hilbert, pueden omitir estos dos primeros capítulos.

En el capítulo tercero tratamos de los procesos estocásticos completado con las funciones de distribución y las condiciones de consistencia de Kolmogorov para que el conjunto de estas funciones de distribución formen un sistema y definan un proceso estocástico.

Insisto en las funciones de covarianza señalando sus propiedades y condiciones que deben cumplir para su derivabilidad e integración si el proceso es de parámetro continuo.

Termino el capítulo con el estudio y propiedades de dos procesos muy importantes en este trabajo: el de proceso de incrementos ortogonales y el de proceso de incrementos independientes.

El cuarto y último capítulo lo dedico a los procesos estacionarios en el sentido de Khintchine; su derivabilidad y desarrollos.

El concepto de ergodicidad equivale a la Teoría de la Ley de los grandes números y nos permite obtener las características poblacionales de los procesos como pueden ser: la función de distribución, media, covarianza, etc., pero calculados sobre el espacio de tiempos y no sobre el espacio de probabilidades.

La existencia, según demostramos, de una relación biunívoca entre las funciones características y las funciones definidas no negativas sobre determinados espacios, nos permite establecer las representaciones espectrales de las funciones de covarianzas de los procesos estocásticos estacionarios con solo probar que la función de covarianza es una función definida no negativa.

Las funciones de densidad espectrales las deducimos calculando las transformadas de Fourier de las funciones de covarianza estacionaria.

Las funciones de distribución y densidad espectrales permiten el conocimiento de la estructura del proceso y su contribución a la varianza de cada frecuencia.

Paralelamente a la representación espectral de la función de covarianza, existe una representación espectral de los procesos. Para una comprensión clara damos una demostración heurística de la representación del proceso estocástico estacionario por mecio de combinaciones lineales de procesos cíclicos y una más completa y formal basándonos en la geometría hilbertiana.

Se le supone al lector con conocimientos básicos de análisis a un nivel de la Teoría de Funciones de Rey Pastor y de la Estadística Teórica que se estudian en las Facultades de Económicas.

Completamos este trabajo con un breve Apéndice y una extensa bibliografía.

INDICE

CAPITULO PRIMERO

BREVE INTRODUCCION HISTORICA

CAPITULO SEGUNDO

ESPACIOS DE HILBERT

— Espacio de sucesiones de cuadrado sumable.
— Espacios funcionales.
— Espacio estocástico de Hilbert.
— Dimensionalidad del espacio de Hilbert.
— Propiedades del producto interno.
— Ortogonalidad.

CAPITULO TERCERO

PROCESOS ESTOCASTICOS

— Definición.
— Funciones de distribución. Condiciones de consistencia de Kolmogorov.
— Procesos multivariantes.
— Funciones de covarianza. Propiedades.
— Diferenciación e integración.
— Procesos ortogonales e incorrelacionados.

CAPITULO CUARTO

PROCESOS ESTOCASTICOS ESTACIONARIOS REPRESENTACION ESPECTRAL

— Procesos estacionarios de orden dos.
— Derivabilidad y desarrollo de procesos estacionarios.
— Ergodicidad. De la media, covarianza y función de distribución.
— Funciones características y funciones definidas no negativas.
— Representaciones espectrales de la función de covarianza estacionaria.
— Funciones de distribución y de densidad espectrales.
— Aliasamiento. Frecuencia de Nyquist.
— Representaciones espectrales de procesos estocásticos estacionarios.

APENDICE

CAPITULO PRIMERO

BREVE INTRODUCCION HISTORICA

Sección 1.ª

Serie cronológica

1. Definimos *serie cronológica* a las variables observadas en distintos momentos del tiempo. Los problemas más difíciles de la Estadística son analizar las series temporales.

2. Definimos *sucesión o secuencia a los valores observados de la serie* temporal

$$\ldots x_{-3}\, x_{-2}\, x_{-1}\, x_0\, x_1 \ldots x_t \qquad (1)$$

cuando t pertenezca al conjunto de los números enteros.

Si t pertenece a un conjunto finito o numerable, la serie es igualmente una secuencia de parámetro discreto.

En consecuencia es una secuencia la serie observada en momentos equidistantes del tiempo.

3. Definiremos una serie temporal de parámetro t continuo

$$x_t = x(t) \qquad (2)$$

si t pertenece a un conjunto no numerable, tales como un intervalo o el conjuto de los números reales.

4. Las expresiones (1) o (2) son observaciones empíricas en diversos momentos del tiempo t discreto o continuo.

Sección 2.ª

Problemas clásicos

1. Exponemos la metodologia siguiente:

a) Recogida de datos.

2.ª Las trayectorias de los planetas son elipses y el Sol ocupa uno de los focos.

3.ª Los cuadrados de los tiempos de las revoluciones de dos planetas son proporcionales a los cubos de los semiejes mayores (II).

A partir de las leyes de Kepler halló Newton que aquéllas no eran sino consecuencia de su nueva Ley: «Que dos masas distintas se atraen directamente proporcional a las masas e inversamente proporcional al cuadrado de su distancia».

3. La regularidad de los fenómenos físicos ha sido la razón por la que a las Ciencias que los estudian se las denominen impropiamente exactas.

No indicamos más series físicas porque sería recopilar todos los avances de la Ciencia y Tecnología.

Sección 4.ª

Series económicas

1. Las series económicas de precios, producciones, demandas, comercio interior y exterior, circulación fiduciaria, Renta Nacional, etc., evolucionan en el tiempo irregularmente y dependen de fenómenos sociopolíticos y de variables interdependientes.

2. La población es una serie regular que su conocimiento siempre ha interesado a los gobernantes para fines muy diversos.

Esta serie pertenece a la Demografía y si la incluimos aquí es por su influencia en todas las series económicas.

Existen modelos que estudian la evolución de la población y la denominada Logística se adapta perfectamente en la práctica [Sec. 5.ª (2)].

3. Del estudio de la población nació la necesidad de construir tablas de mortalidad. Desde Ulpiano en 170 d. C., quien construyó la primera tabla, hasta la Teoría General de V. A. Quiquet (III) pasaron muchas centurias para construir las hipótesis sobre la supervivencia. Citemos las de Gompertz, Makehan, Lazarus y Risser.

Actualmente los Actuarios de Seguros utilizan las tablas ajustadas por la ley de Makeham:

$$l_x = k s^x g^{c^x} \tag{1}$$

(II) Gustav Jäger: «Física teórica», pág. 28. E. Labor, 1932.
(III) Antonio Lasheras Sanz: «Matemáticas del seguro», págs. 221 y sig. Dosat, 1948.

b) Representaciones gráficas.

c) Elección de una función con determinados parámetros.

d) Estimación de los mismos.

e) Estudio de sus componentes (tendencia, variaciones estacionales y cíclicas).

2. Otros problemas consecuencia de los citados previamente son:

a) Interpolación si se quiere intercalar datos o elementos calculados entre los de una serie observada o estimada.

b) Extrapolación, para completar la serie o predecir.

c) Ajuste. La elección de un modelo sencillo que se adapte a las observaciones empíricas y explique e interprete el fenómeno estudiado.

3. El tratamiento de las series depende de la clase del fenómeno observado porque el estudio teórico del mismo sugiere con frecuencia la posibilidad de su representación por una función más o menos complicada.

Existen dos tipos característicos: las series físicas y las series económicas.

Sección 3.ª

Series físicas

1. Desde la antigüedad el hombre se ha preocupado por el estudio de los astros. Las observaciones hechas con gran paciencia y constancia, precisaron tres centurias (I) para lograr el control que el hombre tiene ahora sobre los movimientos planetarios.

Las dificultades de la gravitación universal por la influencia de las masas del Sol sobre los planetas, pudo resolverse con relativa exactitud por extrapolación.

2. Entre los que marcaron las directrices del conocimiento de las Leyes naturales de tipo astronómico, recordaremos a Tycho Brahe (1546-1601), quien recopiló admirablemente una serie de datos que, analizados estadísticamente por Juan Kepler (1571-1630), dedujo sus tres leyes fundamentales:

1.ª El radio vector que va del Sol al planeta describe áreas iguales en tiempos iguales.

(I) Harol T. Davis: «The Analysis of Economic Time Series», Indiana, 1941.

siendo x la edad, l_x el número de personas vivientes a esa edad. Las constantes k, s, g y c, se determinan estadísticamente.

4. En las series económicas se estudian separadamente sus componentes: tendencia, variaciones estaciones y cíclicas.

5. La metodología para el estudio de estas series es la siguiente: su eliminación de la serie original y análisis de las variaciones estacionales o cíclicas de las series libres de la tendencia.

6. El análisis estacional se ha estudiado de diferentes formas. De las series libres de tendencia con un número de años que abarque un ciclo completo se hacían operaciones muy simples:

1.ª Medias mensuales (de los eneros, febreros, etc.).

2.ª Dividir la serie de los doce números por la media de ellos (tomada como unidad) y se iguala a 100, resultando 12 índices, cuya media es 100 y la suma de desviaciones es nula.

3.ª Estas divisiones o tantos por cientos se utilizan como sustraendo o divisor según el criterio.

7. Existen otros métodos que no describiremos y entre ellos el de Warren, H. Pearson (I); método de los cocientes de los datos y la tendencia; el de Richard A. Robb; de las segundas diferencias en cadena; el de Bordin y, finalmente, el de las series de Fourier, que por su importancia merece atención especial.

Obras clásicas importantes como las de O. W. Blackett y P. Wilson (II), W. C. Mitchell (III), la de Arthur F. Burns (IV) y H. L. Rietz (V), dedican capítulos enteros o la totalidad de su contenido a analizar las series económicas.

(I) Fernández Baños, Olegario: «Tratado de estadística». Consejo Superior de Investigaciones Científicas, 1945, pág. 227.

(II) O. W. Blackett y W. P. Wilson: «A Method of Isolating Sinusoidal Components in Economic Time Series», University Michigan Press, 1938.

(III) W. V. Mitchell: «Busines Cycles the Problem and its Settign». National Bureau of Economic Research. N. York, 1936.

(IV) Arthur F. Burns: «Business Cycles». Vol. II. Analysis of cyclican Behavioour. National Bureau of Economic Research. N. York, 1936.

(V) H. L. Rietz: «Hand Book of Mathematical Statistics». Mifflin Co. Bostón.

Sección 5.ª

ANALISIS CLASICO DE LAS COMPONENTES DE LAS SERIES ECONOMICAS

I. Tendencia

1. El movimiento regular de las series se estudió primariamente por el método gráfico que prescinde de toda relación analítica.

2. Por el contrario, otros métodos son exclusivamente analíticos. Indicamos solamente dos:

El parabólico representado por la función:

$$g(t)=a_0+a_1 t+\cdots+a_p t^p \tag{1}$$

y la curva del crecimiento poblacional (I):

$$g(t)=\frac{k}{1+e^{-ct}} \tag{2}$$

denominada logística estudiada en los libros de Estadística.

De la (1) deduce Kendall (II) diversas fórmulas por simple ajuste a 2m+1 términos consecutivos elegidos de la serie:

$$x_{-m} \ldots x_0 \ldots x_{m-1}\, x_m$$

Para resolver el problema de la estimación de los coeficientes de (1) por mínimos cuadrados se deriva con respecto a los parámetros:

$$\frac{\partial}{\partial a_j} \sum_{t=-m}^{m} (x_t-a_0-a_1 t-\cdots-a_p t_j^p)^2=0,\ 1,\ 2,\ \ldots\ p$$

y se iguala a cero obteniendo las ecuaciones normales.

El término central x_L se sustituye por otro teórico (1) dando a $t=0$; es decir por $\hat{a}_0$.

(I) G. G. Bos: «A Logistic approach to the demand for private cars», Tilburg University Press, Netherlands, 1970.

(II) Maurice G. Kendall y A. Stuart: «The Advanced Theory of Statistics», vol. 3. Griffin-London, 1968, 2.ª edic. pág. 366, Cap. 46.

La estimación de a_0 es la tendencia y es una expresión lineal de los términos impares elegidos:

$$\hat{a}_0 = x_0^* = c_0 + c_1 x_{-m} + c_2 x_{-m+1} + \cdots + c_{2m+1} x_m \tag{3}$$

donde las constantes c son independientes de los valores de la serie.

Estas fórmulas son promedios móviles y dependen de p y de m. Para $m=3$ y $p=3$ (una cúbica) el valor central estimado es:

$$\hat{a}_0 = x_0^* = \frac{1}{21}\left[7 \sum_{t=-3}^{3} x_t - \sum_{t=-3}^{3} x_t t^2\right] =$$

$$= \frac{1}{21}[-2x_{-3} + 3x_{-2} + 6x_{-1} + 7x_0 + 6x_1 + 3x_2 - 2x_3] \tag{4}$$

Con independencia de los valores de la serie la (4) puede escribirse:

$$\hat{a}_0 = x_0^* = \frac{1}{21}[-2,\ 3,\ 6,\ 7,\ 6,\ 3,\ -2]\, x_0 \tag{4'}$$

Para estimar el central observemos que (4) utiliza siete valores consecutivos de la serie y los coeficientes son simétricos y suman la unidad.

La expresión (4') se escribe por su simetría:

$$\hat{a}_0 = x_0^* = \frac{1}{21}[-2,\ 3,\ 6,\ \underline{7}]\, x_0 \tag{4''}$$

El último término subrayado es el coeficiente central de (4').

3. Otro método sencillo (3) es el de promedios móviles de Wittsein. Atenúa las irregularidades y perturbaciones que presentan las series cronológicas, y sustituye los datos observados por una nueva serie de términos centrales formados por medias de cinco términos.

Así, el término $\bar{x}_t$ de la nueva serie es:

$$\bar{x}_t = \frac{x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2}}{5} \tag{5}$$

Este método es un caso particular de (1) si hacemos $m=2$ y la parábola es una recta.

El suavizado de las series irregulares también se obtiene con todas las fórmulas de Kendall.

4. Para mayor regularidad, Filaison repitió con los valores $\bar{x}_t$ la misma operación (5).

5. A veces estas fórmulas se expresan por operadores. Si indicamos el numerador de (5) por:

$$[5]x_t = \sum_{n=-2}^{+2} x_{t+n} \tag{6}$$

la (5) puede escribirse:

$$\bar{x}_t = \frac{[5]x_t}{5} = \frac{x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2}}{5} \tag{5'}$$

donde [5] es el operador suma de cinco términos y que aplicado sobre x_t representa (6).

Introduciendo otro operador:

$$r_n x_t = x_{t-n} + x_{t+n} \tag{7}$$

la (5') puede escribirse:

$$\bar{x}_t = \frac{[5]x_t}{5} = [1 + r_1 + r_2]x_t \tag{5''}$$

6. La serie de Filaison formada con las series de Wittsein, por repeticiones sucesivas, es:

$$\bar{x}_t = \frac{[5]\bar{x}_t}{5} = \frac{[5][5]x_t}{25} = \frac{1}{25}[5 + 4r_1 + 3r_2 + 2r_3 + r_4] \tag{8}$$

porque el operador r_j cumple las condiciones:

$$r_j r_k = r_{j+k} + r_{j-k} \qquad \text{y} \qquad r_k = r_k \qquad r_0 = 2 \tag{7'}$$

7. Más científico que los dos citados precedentemente, aunque también es una expresión lineal del tipo (3), es el método de Woolhouse. Precisa 15 valores para formar el central y agrupa los valores de las series en las ternas equidistantes:

$$(x_{t-7},\ x_{t-2}\, x_{t+3}),\ (x_{t-6},\ x_{t-1},\ x_{t+4}),$$
$$(x_{t-5},\ x_t,\ x_{t+5}),\ (x_{t-4},\ x_{t+1},\ x_{t+6})$$
$$\text{y} \quad (x_{t-3},\ x_{t+2},\ x_{t+7}) \tag{9}$$

Por cada terna hace pasar una parábola y determina el valor en el punto t.

La parábola general en t es:

$$u_t(p)=x_p+\frac{t-p}{5}\Delta x_p+\left(\frac{(t-p)}{5}\frac{(t-p)}{5}-1\right)\frac{\Delta^2 x_p}{2} \tag{10}$$

$$p=t-7,\ t-6,\ ...,\ t-3$$

La media de los valores $u_t(p)$ de las cinco parábolas expresadas en función de los operadores (6) y (7) es:

$$y_t=\frac{[5]^3}{125}(10-3[3])x_t \tag{11}$$

8. La expresión lineal (3) es muy general y los coeficientes c_i en cada fórmula utilizada simpre son los mismos para formar la nueva serie.

9. Citaremos las fórmulas de Spencer de 15 y 21 puntos e indiquemos finalmente las fórmulas que utilizan operadores de las diferencias de Newton y centrales, el de traslación, etc., que tienen propiedades lineales importantes e incluso los polinomios ortogonales.

10. El método de diferencias se utiliza para determinar el grado de la parábola si la serie procede de un modelo parabólico con una perturbación aleatoria $\varepsilon_t \Rightarrow$

$$E(\Delta^h \varepsilon_t)=\binom{2h}{h}\sigma^2 \tag{12}$$

σ^2 es la varianza de la perturbación aleatoria y para $h \geqslant p$ permite estimar el valor de σ^2 (excepto fluctuaciones muestrales) y el grado de la parábola cuando practicamente σ^2 sea igual para un valor de $h \geqslant p$. (El lector interesado puede consultar a Tintner (I) y Quenouille (II) sobre esta material.)

(I) G. Tintner: «The Variate Difference Method», Principia Press, Bloomington, Indiana, 1940.

(II) M. H. Quenouille: «Some Results in the Testing of Serial Correlation-Coefficients», Biometrika, vol. 35. págs. 261-267.

ANALISIS CLASICO DE LAS COMPONENTES DE LAS SERIES ECONOMICAS

II. Componentes cíclicos

1. Los ciclos económicos observados no pueden establecerse con rigurosidad ni sus amplitudes ni sus períodos como en las leyes físicas. Los ciclos distintos de los estacionarios observados son: El ciclo de los negocios de duración aproximada de tres años y medio. En la obra citada de W. C. Mitchell, después de estudiar 127 años de datos de negocios, llegó a la conclusión que habían sido 32 ciclos, con un promedio de longitud de cuatro años.

Este ciclo parece normal en los fenómenos económicos americanos, aunque los europeos tengan una duración aproximada de cinco años (I).

Otros como el ciclo agrícola (influye las manchas solares); el de la edificación y el gran ciclo son estudiados en Economía.

El estudio del ciclo de los negocios (término aplicado a las alteraciones periódicas de los negocios entre la prosperidad y la depresión), posiblemente tuvo su origen en la obra de J. C. L. de Sismondi (II) y en la de Clement Juglar (III).

2. Muchos de los métodos para el análisis de las series económicas emplean los números índices, según indicamos en 6 de la Sección 4.ª

Recordemos la obra de Jevons (IV). Citemos los importantes y clásicos de Laspeyres, Paasche, Bradstreet y Dulot, entre los numerosos examinados por Fisher para que cumplieran ciertas condiciones de inversión en el tiempo e inversión de los factores y obtener su índice ideal.

Beveride (V) recopiló los índices de precios durante los años 1500 a 1869.

3. E. Slutzky (VI), al estudiar los índices de negocios en Inglaterra, durante los años 1855 a 1877, *los comparó con una serie artificial de promedios móviles* formada por diez términos de números aleatorios. *La consecuencia de su importante estudio fue que a partir de números aleatorios puede obtenerse una serie artificial de tipo cíclico* y de características simila-

---

(I) G. Tintner: «Price in the Trade Cycle», Viena, 1935.

(II) J. C. L. Sismondi: «Nouveaux Principes d'Economic Politique», Paris, 1919.

(III) Clement Juglar: «Des Crisis Commerciales et leur Retour Periodique, 1860.

(IV) W. S. Jevons: «On the Study of Periodic Comercial Fluctuations».

(V) M. G. Kendall, o. c. pág. 403.

(VI) E. Slutzky: «The Summations of Randon Causes as the Source of the Cyclic Processes». Econométrica, vol. 5 de 1937, pág. 105.

res a las series económicas observadas. También Yule (I) observó el mismo fenómeno formado con procesos de medias móviles y autorregresivos.

4. En resumen: El método clásico de las series cronológicas consistía en descomponer la serie en cuatro componentes independientes: tendencia secular, variaciones estacionales, variaciones cíclicas y finalmente las variaciones erráticas.

Las series sin tendencia se forman por las desviaciones de las series iniciales menos las ajustadas por cualquier método expuesto para el cálculo de la tendencia (*).

Con estas series sin tendencia o se calculan números índices trimestrales, mensuales, semanales, etc., o se efectúa un ajuste trigonométrico de períodos estacionales.

Si de la serie original eliminamos la tendencia y la estacionalidad, obtenemos nueva serie y de ésta se estudian las variaciones cíclicas ajustando términos trigonométricos.

Cuando los ciclos eran conocidos, se representaban por términos trigonométricos:

$$\alpha(k_j)\cos\frac{2\pi k_j t}{T}+\beta(k_j)\operatorname{sen}\frac{2\pi k_j t}{T}$$

Si k_j era divisor de T, se ajustaba por mínimos cuadrados a la serie y se estimaban los parámetros $\alpha(k_j)$ y $\beta(k_j)$. Este ajuste se hace en las series económicas sin tendencia y con ciclos.

Si el período fuese desconocido se determina previamente utilizando el Periodograma de A. Schuster que, por su importancia, hablaremos con cierto detalle en la Sección 9.ª

5. Pero, en general, el valor científico del estudio de las series económicas era dudoso o de escaso valor porque no se planteaba en términos de probabilidad.

Sección 6.ª

Procesos estocásticos

1. En 1933 Kolmogorov (II), escribe su célebre axiomática Probabilidades, basada en los conjuntos de Borel de -*n*- dimensiones de los espa-

---

(I) G. U. Yule: «On a Method of Investigating Periodicities in Disturbed Series». J. Roy, Staat. Soc. pág. 89, 1926.

(*) Exponemos ideas generales. Podría también ser el cociente.

(II) A. N. Kolmogorov: «Foundations of the Tehory of Probability. 2.ª edic. inglesa, Chelsea, 1956. N. Y.

cios abstractos. Contribuyeron, en principio, Lévy y Feller; y a partir de 1936, la teoría de los procesos estocásticos se desarrolla rápidamente.

2. Los procesos se han utilizado para resolver numerosos problemas prácticos y describir fenómenos empíricos.

3. El primero de los procesos estocásticos tuvo su origen en el estudio de la distribución de Bernouilli, aplicada a los «paseos aleatorios o de azar». Karl Pearson, en 1905, fue quien por vez primera resolvió este problema.

4. En 1827, el botánico inglés Robert Brown observó que las pequeñas partículas suspendidas en un fluido inmerso siguen movimientos aleatorios (I). Este fenómeno se manifiesta en los gases y se denomina movimiento browniano. La causa es porque el movimiento de las moléculas de tamaño microscópico están sometidas a choques estocásticos independientes.

Desde 1905 el movimiento browniano se ha tratado estadísticamente en las obras de Einstein y Smoluchowsky (II), contribuyendo numerosos matemáticos y físicos: Fokker, Flanck, Ornsteins, Chandrasekhar (III). Deduce este autor el movimiento browniano a partir del proceso de «paseos aleatorios»; lo extiende al plano y al espacio y lo relaciona con otros procesos físicos. Wiener, en 1923, demuestra la continuidad del proceso con «probabilidad uno». Por los trabajos de Lévy (IV) también se denomina a este proceso Wiener-Lévy.

5. Los trabajos de Khitchine, Kolmogorov, Slutsky, Cramer, Karhunen, Loéve, Doob, Hanann, etc., evolucionan el concepto de trayectoria o serie cronológica a una función aleatoria dependiendo de una variable ω estocástica. Para cada valor de esa variable se obtiene una trayectoria y todas las trayectorias imaginables representan el proceso:

$$\xi(\omega,\ t)=\{\xi(t),\ t\in T\}$$

Para cada $\omega=\omega_0$ la función $\xi(\omega_0 t)$ (t variable) es una realización experimental del proceso o la serie cronológica definida en la sección 1.ª Y si $t=t_0$. $\xi(t_0)$ es una variable aleatoria.

(I) S. Chandrasekhar: «Stochastic Problems in Physis and Astronomy. Nelson Wax, 1954, pág. 22.

(II) J. L. Doob: «The Brownian Movement and Stocastic Ecuations», Trabajo publicado en Selecter Papers & Noise & Stochastic Processes. Nelson Wax, 1954, pág. 319 a 337.

(III) O. c. Págs. 319 y siguiente.

(IV) Pauel Lévy: «Processus Stochastiques et Mouvement Brownien», Gautier-Villars, 1965.

6. La familia de variables aleatorias -*n*- dimensional o vector:

$$\xi = \{\xi_1, \xi_2, ..., \xi_n\}$$

es un punto que pertenece a un espacio de probabilidad y es un suceso elemental. Con todos los sucesos elementales posibles formamos un espacio Ω. Un suceso estocástico es un subconjunto de Ω y pertenece a cierta clase de conjuntos de Borel $\mathscr{B}_n$. Si formamos conjuntos borelianos de $\mathscr{B}_n$, por ejemplo, del tipo:

$$S = \{\xi_{t_1} \leqslant u_1, \xi_{t_2} \leqslant u_2, ..., \xi_{t_n} \leqslant u_n\}$$

a cada uno le corresponderá una función de distribución:

$$P(S) = F_{t_1 t_2 \ldots t_n}(u_1, u_2, ..., u_n)$$

con su espacio de probabilidad: $(\Omega, \mathscr{B}_n, P)$ y se cumplen las condiciones de Kolmogorov.

7. En este sentido una observación empírica es una realización; es decir, y siguiendo a Khintchine (I) el proceso estocástico puede considerarse como una función abstracta del parámetro ω asignando a cada suceso elemental una función real perteneciente a Ω cada una de las cuales representa una realización.

El proceso puede representarse por:

$$\{\xi_t, t \in T\}$$

donde T es el conjunto índice.

8. Las dificultades surgen, como dice Kari Karhunen (II), cuando se analiza el proceso: si es una función medible o no; si el conjunto paramétrico T no debe restringirse al tiempo, etc.

Doob fundamenta su teoría general del proceso considerándolo como una función abstracta.

Finalmente diremos que Karhunen desarrolla su teoría sistemática concibiendo los «procesos como funciones cuyos valores son puntos de un espacio abstracto».

(I) A. Khintchine: «Korrelation Theorie der Stationären Sochastichen Prozesse», 1934.
(II) Kari Karhunen: «Métodos lineales en el cálculo de probabilidades», Monografía número 36. C. Sup. Invest. Científicas, pág. 10, 1952.

Sección 7.ª

Espacios abstractos

1. El estudio de las Ciencias aplicadas exige de los investigadores la creación de nuevos métodos de cálculo.

2. La Teoría de la Medida fue posible con la introducción de los conceptos de conjuntos de Borel y la integral de Lebesgue.

3. Hilbert sintió la necesidad de efectuar un estudio detenido de los valores propios de los operadores lineales. Primero formuló Hilbert su espacio como extensión del euclídeo al caso en que el número de dimensiones fuera infinito.

Más tarde Hilbert formuló su axiomática de su espacio de un modo abstracto. Los elementos pertenecientes al espacio han de cumplir los axiomas de grupo, linealidad y medida (I).

4. Son espacios de Hilbert las funciones de cuadrado sumables; los sistemas ortogonales completos; las series de Fourier, etc.

En 1907, Frechet, y F. Richsz, demostraron que el espacio de funciones de cuadrado sumable posee una geometría totalmente analógica con el espacio de Hilbert, siendo completo e isomorfo.

Son importantes las obras de M. H. Stone; las contribuciones de F. Richsz, de Banach (II) y la importante obra de Béla Sz. Nagy (III).

Sección 8.ª

Teoría ergódica

1. La teoría ergódica trata del importante problema de la determinación de las características estocásticas de un proceso aleatorio mediante una sola o muy pocas realizaciones.

2. Esta teoría se basa en la transformación de espacios. Brevemente indicaremos su historia.

(I) J. R. Pastor: «Los problemas lineales de la Física», Intact. 1955, pág. 37 y sig.

(II) I. M. Gulfand; G. E. Chilov: «Les distributions», tomo 2. «Espaces Fundamentaux», Dunod, 1964.

(III) Béla Sz. Nagy: «Analyse Harmonique des Operateurs de l'Espace de Hilbert», Masson 1967. Impreso en Ungría.

3. El período clásico (1930-1944) trata de transformaciones preservando la medida, una transformación T_1 de un espacio métrico en correspondencia biunívoca sobre sí mismo con su espacio de probabilidad: $(\Omega, \mathscr{A}, \mathscr{P})$.

Si x es una variable aleatoria y definimos

$$x_n \text{ por } x_n(\omega) = x(T_1^{n-1}\,\omega)$$

La primera relación básica:

según von Newman si $x \in L_2$ entonces $\frac{1}{n}\sum_{k=1}^{n} x_k$

converge en media cuadrática. Y según Birkhoff en 1931, si

$$x \in L_1 \text{ entonces } \frac{1}{n}\sum_{k=1}^{n} x_k$$

converge casi seguramente al verdadero valor (I).

4. Khintchine redujo las hipótesis hechas por Birkhoff y, al mismo tiempo, simplifica su demostración. Hopf extiende la teoría a los cocientes de sumas de secuencias estocásticas; Yosida y Kahutani extienden el teorema de Newman a las transformaciones sobre el espacio de Banach y la aplica a las cadenas de Markov.

5. De 1944 en adelante, Loeve (II) llama el período moderno y se caracteriza en reducir las condiciones a la de convergencia en media como lo hace Dunford y Miller. Entonces F. Riesz dá una demostración muy simple del teorema de Birkhoff. Doob aplica el teorema de Birkhoff a las cadenas de tipo estacionario.

Khintchine se apoya en la representación espectral de la función de correlación y Karhunen (III) toma la idea de F. Riesz sin ayuda de la representación espectral, basándose en las proyecciones.

6. *Las leyes de un proceso pueden alterarse con el transcurso del tiempo. La teoría ergódica es la clave que nos permite el entendimiento de estas fluctuaciones.*

7. *El teorema ergódico tiene semejanza con la ley de los grandes*

(I) Véase Apéndice 1.

(II) M. Loéve: «Probability Theory». Van Nostrand, 1963, págs. 410 y sig.

(III) K. Karhunen: «Métodos lineales en el cálculo de probabilidades». C. Sup. Inv. Científicas, 1952, pág. 66.

números y nos permite obtener las características más importantes de un proceso estocástico, como pueden ser: la media, función de covarianza, función de distribución, etc., *con una o muy pocas realizaciones* que de no ser así sería, prácticamente, imposible su estimación.

Sección 9.ª

Análisis armónico

1. El análisis armónico es el fundamento del cálculo espectral. Comienza por descubrir periodicidades en las series con Lagrange hacia 1772 (I). Euler representó una función $f(t)$ por una serie de la forma:

$$f(t)=\frac{a_0}{2}+\sum_{n\in N} a_n\cos\frac{n\pi t}{a}+b_n\operatorname{sen}\frac{n\pi t}{a} \tag{1}$$

Si t fuera un punto de discontinuidad de $f(t)$ el segundo miembro de (1) valdría:

$$f(t)=\frac{f(t-0)+f(t+0)}{2}$$

donde $2a=p$ es el periodo $f(t)=f(t+p)$ (II).

2. Toda función que cumpla las condiciones de Dirichlet (III) puede aproximarse por un desarrollo trigonométrico (1) cuyos coeficientes son:

$$a_n=\frac{1}{a}\int_{-a}^{a} f(t)\cos\frac{n\pi t}{a}dt;\quad b_n=\frac{1}{a}\int_{-a}^{a} f(t)\operatorname{sen}\frac{n\pi t}{a}dt \tag{2}$$

$$\frac{a_0}{2}=\frac{1}{2a}\int_{-a}^{a} f(t)dt$$

3. Un término del desarrollo (1) como es:

$$a_n\cos\frac{n\pi t}{a}+b_n\operatorname{sen}\frac{n\pi t}{a} \tag{3}$$

(I) J. L. Lagrange: «Recherches sur la maniere de former des tables des planetes», obra citada por Davis, pág. 28.

(II) Ver Apéndice 3.

(III) Ver Apéndice 2.

se llama el armónico —n— u onda armónica, siendo $\frac{2a}{n}$ el período de tal armónico: y la semiamplitud es:

$$R_n=\sqrt{a_n^2+b_n^2} \tag{4}$$

Los períodos $\frac{2a}{n}$ forman la sucesión de Fourier:

$$2a,\ \frac{2a}{2},\ \frac{2a}{3},\ \ldots\ \frac{2a}{n}\cdots n\in N \tag{5}$$

4. Los extraordinarios resultados obtenidos por los físicos y astrónomos indujeron a los economistas a representar las series económicas por series de Fourier.

5. Si $f(t)$ es desarrollable en serie de Fourier la media es $\frac{a_0}{2}$ y la dispersión es:

$$\sigma_F^2=\frac{1}{2a}\int_{-a}^{a}\left[f(t)-\left(\frac{a_0}{2}\right)\right]^2 dt \tag{6}$$

$$=\frac{1}{2}\sum_{n=1}^{\infty}(a_n^2+b_n^2)=\frac{1}{2}\sum_{n=1}^{\infty}R_n^2$$

por las condiciones de ortogonalidad de las funciones trigonométricas.

6. Si a la función $f(t)$ se ajusta una aproximación de m armónicos (3):

$$f_m(t)=\frac{a_0}{2}+\sum_{n=1}^{m}\left(a_n\cos\frac{n\pi t}{a}+b_n\operatorname{sen}\frac{n\pi t}{a}\right) \tag{7}$$

sus coeficientes los obtendremos por mínimos cuadrados y de la (1) y (7)

$$\min\frac{1}{2a}\int_{-a}^{a}\left[f(t)-f_m(t)\right]^2 dt$$

estimamos los valores a_0, a_n, b_n, $(n\in N)$ que son precisamente las fórmulas dadas en (2).

7. La varianza residual de $f(t)$ si tomamos la aproximación (7) es:

$$\sigma_{ry}^2=\frac{1}{2a}\int_{-a}^{a}\left[f(t)-f_m^x(t)\right]^2 dt=\sigma_F^2-\frac{1}{2}\sum_{n=1}^{m}R_n^2 \tag{8}$$

porque σ_F^2 es la (6) y también:

$$\sigma_F^2 = \frac{1}{2a}\int_{-a}^{a} f(t)^2\,dt - \frac{a_0^2}{4} \tag{9}$$

Si $f(t)$ fuera desarrollable en serie de Fourier y tomásemos solamente m armónicos, la varianza residual según (6) y (8):

$$\sigma_{r\mu}^2 = \frac{1}{2}\sum_{m+1}^{\infty} R_n^2 \tag{10}$$

8. En ocasiones cuando desconocemos el período $p=2a$ hay que determinar el mínimo valor de p tal que:

$$f(t) = f(t+p) \quad \forall t \in T \tag{11}$$

así np será también un período si n es entero.

9. A Schuster (I) dedujo un método científico para resolver el problema de encontrar el período de las series observadas y al cual denominó «periodograma».

John W. Tukey (II) llama al periodograma la prehistoria del análisis espectral y Kendall (III) lo denomina espectro.

10. El periodograma es un método de análisis que tiene por finalidad descubrir componentes periódicos.

El esquema de Buys-Ballot y A. Schuster para encontrar un período p cuando tenemos N observaciones disponibles es elegir subgrupos —p variable— de los N datos; N' *observaciones consecutivas* de forma que

$$\frac{N'}{i} = p \quad \text{sea entero}$$

despreciando los últimos términos de la serie.

(I) Sir Arthur Schuster: «On Hidden Periodicities, Terrestrial Magnetism», 3. 1897 y 1898, 24: «The Periodiodograma and its Optical Analogy». Proc. Roy. Soc. London, 77. A. 1906, 133-140 y otros. Trabajos citados por Davis, pág. 31, o.c.

(II) J. W. Tukey: «An Introduction to the Calculations of numerical Spectrum Analysis», pág. 25. Advanced Seminar on Spectral Analysis of time series. John Wiley, 1967.

(III) Kendall: «The Advanced Theory if Statistics», vol. 3, pág. 412, 2.ª edición.

Realización práctica. Se forman cuadros semejantes al que exponemos variando p:

	x_1	x_2	x_3	$\ldots x_p$	
	x_{p+1}	x_{p+2}	x_{p+3}	$\ldots x_{2p}$	(12)
	$x_{(i-1)p+1}$	$x_{(i-1)p+2}$	$x_{(i-1)p+3}$	$\ldots x_{ip}$	
Sumas	S_1	S_2	S_3	$\ldots S_p$	
Medidas	M_1	M_2	M_3	$\ldots M_p$	

$$\text{Siendo } N' = ip; \; S_k = \sum_{j=1}^{i} x_{k+(j-1)p} \tag{13}$$

$$M_k = \frac{S_k}{i} \tag{14}$$

La disposición práctica de la tabla es debida a Buys Ballot, 1847. En todo cuadro se verifica:

1.° Que las observaciones elegidas son correlativas.

2.° Deberá elegirse N' para cada valor de p que sea el múltiplo más próximo a las N observaciones disponibles.

3.° Variando p obtendremos cuadros distintos.

Para cada valor de p existe una media máxima y otra mínima (14). El método de A. Schuster consiste en:

1.° Determinar los coeficientes del desarrollo de Fourier por las expresiones como si p fuese un período:

$$a(p) = \frac{2}{N'} \sum_{x=1}^{p} S_x \cos \frac{2\pi x}{p} \tag{15'}$$

$$b(p) = \frac{2}{N'} \sum_{x=1}^{p} S_x \operatorname{sen} \frac{2\pi x}{p} \tag{15''}$$

2.° Formar para cada valor de p el cuadrado de la semiamplitud:

$$R(p)^2 = a(p)^2 + b(p)^2 \tag{16}$$

3.° El gráfico cuando R_p^2 son las ordenadas y p las abscisas ($p \in N$) se llama Periodograma.

4.° El periodograma tiene la propiedad que cuando k coincide con un período p, entonces R_p alcanza su valor máximo.

Esta demostración equivalente a que tome el valor mínimo la expresión:

$$\sum_{t=1}^{N}\left(x_t - a\cos\frac{2\pi ti}{N} - b\,\mathrm{sen}\frac{2\pi ti}{N}\right)^2 \qquad (17)$$

cuando $p=\frac{N}{i}$ sea un período de (17), la anterior es igual a:

$$\sum_{t=1}^{N} x_t^2 - R_p^2\frac{N}{2} \qquad (18)$$

porque los parámetros estimados de (17) a y b valen:

$$a_p=\frac{2}{N}\sum_{t=1}^{N} x_t\cos\frac{2\pi ti}{N} \qquad (19)$$

$$b_p=\frac{2}{N}\sum_{t=1}^{N} x_t\,\mathrm{sen}\frac{2\pi ti}{N} \qquad (19')$$

equivalentes a las (15') y (15'').

Si entre todos los valores de R_k^2 existe uno p que hace mínima la varianza (18), éste será el máximo de R_x: $\underset{k\in N}{\text{Máx.}}\ R_k = R_p$.

11. En resumen: Después de hacer los cuadros (12) y calculados los valores $a(p)$ y $b(p)$ por las fórmulas (15') y (15'') y dibujado el periodograma de (16), las abscisas de p que maximicen $R^2(p)$ serán los valores de los períodos de la serie cronológica.

12. Si cambiamos la escala haciendo

$$\text{Frecuencia} = \lambda = \frac{1}{p} \qquad (17)$$

la gráfica de las semiamplitudes de $R^2(\lambda)$ en función de la frecuencia λ la denomina Anderson (I) «espectrograma».

El periodograma y el espectrograma tienen relación con la función de densidad espectral.

La frecuencia definida en (17) está medida en circunferencias o ciclos. También podría haberse definido la frecuencia como:

$$\lambda=\frac{2\pi}{p}$$

(I) T. W. Anderson: «The Statistical Analysis of Time Series». Wiley 1958, pág. 109.

Sección 10.ª

Funciones de correlación y covarianza

1. El primero que estudió la teoría de la correlación en forma embrionaria fue Gauss en su memoria fundamental de mínimos cuadrados (I) aunque desde el punto de vista de la teoría de Gauss consideraba a las variables de forma distinta a como se conciben actualmente.

2. En 1846 aparece un trabajo interesante de August Bravais titulado «Sur les probabilités des erreurs de situation d'un point».

Pearson escribe: «Los teoremas fundamentales de correlación fueron discutidos por primera vez por Bravais tratando de dos y tres variables». Y añade: «Es indudable que el coficiente de correlación de Galton aparece en la obra de Bravais pero en una terminología no usada por él.» En 1892 el profesor Edgeworth (ignorando el trabajo de Bravais), trata de la media correlacionada y obtiene resultados idénticos a los que obtendría con Bravais, aunque expresados en términos de las funciones de Galton (funciones de correlación).

3. Karl pearson y Yule (II) contribuyeron a la obtención de la distribución del coeficiente de correlación lineal y a la generalización del concepto de correlación múltiple.

4. Taylor (III) aplicó el concepto de función de correlación a las series cronológicas.

5. Norbert Wiener (IV) utiliza este concepto para predicción y filtrado cuando las series temporales son estacionarias.

6. La esperanza matemática en el tiempo t de la variable aleatoria x_t (t constante) es:

$$\alpha(t) = E(x_t) \tag{1}$$

o su equivalente:

$$\alpha(t) = E(x_t) = \int_R x f(x,\ t)\ dx = \int_R x dF(t,\ x) \tag{2}$$

(I) Karl Pearson: «Notes on the History of Correlations», Studies in the History of Statistics and Probability. Griffin, 1970, págs. 185 y siguientes.

(II) G. U. Yule: «On the Time-Correlation Problem, with Special Reference to the Variate-Diference Correlation Method». J. Roy, Stat. Soc. pág. 84, 1921.

(III) G. T. Taylor: «Diffusion by Continuos Mouvements». Prof. London. Math. Soc. págs. 196-212, 1920.

(IV) Norbert Wiener: «Generalized Harmoni Analysis». Math. vol. 55, págs. 117-258, 1930. Extrapolación: «Interpolation and Smoothing of Stationary Time-Series».

donde $F(x, t)$ es una función de distribución paramétrica con la condición

$$\int_s dF(t, x) = 1 \qquad \begin{matrix} x \in S \\ t \in T \end{matrix} \tag{3}$$

7. Definimos coeficiente de correlación del proceso $\{x_t, t \in T\}$ a la expresión:

$$\rho(t, s) = \frac{E[(x_t - \alpha(t))\overline{(x_s - \alpha(s))}]}{\sigma(x_t)\ \sigma(x_s)} \tag{4}$$

8. Definimos función de covarianza al numerador de (4) donde $\overline{x_s - \alpha(s)}$, indica el conjugado de $x_s - \alpha(s)$. Si $t = s$, entonces

$$\sigma_t^2 = E/x_s - \alpha(t)/^2 \tag{5}$$

se denomina la varianza del proceso complejo en el punto t.

La función de covarianza, numerador de la (4), la representaremos en lo sucesivo así:

$$B(t, s) = E[(x_t - \alpha(t)]\overline{(x_s - \alpha(s))}] \tag{6}$$

La covarianza (o autocovarianza) (I) es función de s y de t; si hacemos s-t, llamando r a la diferencia:

$$B(t+r, t) = E(x_{t+r} - \alpha(t+r))\overline{x_t - \alpha(t)}$$

$$= E(x_{t+r}\bar{x}_t) - \bar{\alpha}(t)\,\alpha(t+r) \tag{7}$$

9. Un caso particular importante es si se verifica:

$$B(t+r, t) = B(r) \tag{7'}$$

o sea cuando la covarianza sea función de la diferencia de tiempos.

Las condiciones suficientes para que la función de covarianza dependa de la diferencia de tiempos es que el proceso x_t tenga características poblacionales constantes e independientes de t:

$$\begin{aligned} Ex_t &= \alpha \\ \sigma_t^2 &= \sigma^2 \end{aligned} \tag{8}$$

(I) Algunos autores la denominan autocovarianza. Nosotros la llamaremos covarianza simplemente siempre que no haya posibilidad de confusión.

Estas condiciones son consecuencia de la invarianza de la función de distribución.

10. En el caso (7') la función de correlación (4) también será función de la diferencia de tiempos $t-s=r$ y vendrá dada por:

$$\rho_r=\frac{B(r)}{B(0)} \tag{9}$$

porque según las expresiones (6), (5) y (7')

$$B(0)=\sigma^2 \tag{10}$$

El conjunto índice $t \in T$ puede ser los números enteros o continuos; y x_t puede ser de naturaleza real o compleja.

En las series económicas, en general, x_t es real y si se cumple (8), se verifica:

$$\rho_r=\rho_{-r}$$

A la representación gráfica de la totalidad de ρ_r denominó Wold (I) correlograma y permite analizar en el dominio del tiempo las características de las series temporales.

En el trabajo de Wold, altamente meritorio, determina modelos estocásticos estacionarios con claridad, estudia sus correlogramas y presentan estructuras análogas a las series económicas.

Sección 11.ª

Representaciones espectrales

1. Los problemas físicos como el estudiado por Bernoulli de «las cuerdas vibrantes»; el descubrimiento por Newton de que la luz al incidir sobre un prisma se descomponía en una gama de colores, visibles e invisibles, sugirieron ideas importantes: ciertos hechos experimentales pueden descomponerse o componerse por *medio de oscilaciones propias.*

2. Estos fenómenos se estudiaron por desarrollos en series de Fourier o funciones ortogonales.

(I) Hermann O. A. Wold: «Series cronológicas estacionarias». Consejo Superior de Investigaciones Científicas, 1951, pág. 13.

3. Los problemas de Física atómica moderna se basan sobre fundamentos rigurosos de la mecánica cuántica y la estadística.

4. En 1929 Von Newman (I), Wintner (II) y en 1932 Stone, estudiaron la métrica del espacio de Hilbert y las propiedades de los operadores unitarios U, y obtienen su representación espectral:

$$U^n = \int_{-\pi}^{\pi} e^{i\lambda n} dE(\lambda) \qquad \lambda \in (-\pi, \pi) \tag{1}$$

Pero el mismo Newman dice que la integral de Stieltjes carece en rigor de sentido porque está definida para números y no para operadores.

5. Herglotz (III) en 1911 dedujo la representación de los coeficientes de correlación y en consecuencia la función de covarianza, cuando el proceso es de tipo discreto.

6. En 1934 Khintchine (IV) define la estacionariedad de segundo orden demostrando sus propiedades y la representación espectral para la función de covarianza de un proceso estocástico estacionario de parámetro t continuo:

$$B(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda) \qquad \lambda \in R \tag{2}$$

siendo $F(\lambda)$ una función monótona no decreciente denominada distribución espectral que estudiaremos en su momento. La semejanza de (2) con la función característica de Lévy, 1934, es importante.

7. Apoyándose en la representación espectral de la función de correlación, Khintchine (1934) deduce el teorema ergódico aplicado a la determinación de la media.

8. H. Wold en 1938 (V) ofrece la representación espectral del correlograma de un proceso estocástico discreto estacionario real deducido ya por Herglotz:

$$\rho_k = \int_0^{\pi} \cos k\lambda dF(\lambda) \qquad k \in N \tag{3}$$

(I) J. Von Newman: «Fundamentos matemáticos de la mecánica cuántica», Cons. Sup. Invest. Científicas, 1949, págs. 24 y siguientes.

(II) A. Wintner: «Zur Theorie der Beschrankten Bilinearformen», Math. 1929. Zeitsch. 30, ob. cit. por Doob, pág. 636.

(III) Doob, o. c. pág. 637.

(IV) A. Khintchine: «Korrelations Theorie der Stationaren Stochastichen Processe». Mathematische Annalen, 109, pág. 608 (1934).

(V) Hermann O. A. Wold: «Series cronológicas estacionarias». Cons. Sup. Inv. Científicas, 1951, ref. págs. 7 y 27.

e igualmente deduce que todo proceso estocástico estacionario puede descomponerse en suma de dos procesos: uno de naturaleza «determinista» y otro «no determinista» ambos independientes.

9. En 1941 Kolmogorov (I) demostró la descomposición armónica de la función de covarianza basada en las propiedades de los operadores lineales unitarios.

10. Existe otra descomposición espectral: la del proceso estocástico.

En 1937 Slutsky (II) obtuvo una primera descomposición armónica. Doob (III) y Lévy (IV) atribuyen a Cramer (V) en 1942 la representación espectral de un proceso estocástico de tipo estacionario que puede escribirse como una integral estocástica del tipo:

$$\xi(t) \stackrel{2}{=} \int_{-\infty}^{\infty} e^{i\lambda t} d\zeta(\lambda) \qquad \begin{array}{l} t \in R \\ \lambda \in R \end{array} \tag{4}$$

La función aleatoria $\zeta(\lambda)$ tiene propiedades estocásticas importantes y por su trascendencia la exponemos en el n.° 12 de esta sección.

11. Los trabajos de Newman (1929) y Wintner sobre operadores lineales fueron el punto de partida para que Kolmogorov (VI) en 1941 demostrase el análisis armónico y la representación espectral de las sucesiones estocásticas estacionarias.

El estudio de los procesos estocásticos estacionarios de segundo orden puede desarrollarse por la geometría del espacio de Hilbert.

12. Karhunen (VII) obtuvo descomposiciones más generales que la (4). Igualmente Loéve (VIII) en 1945/46 formuló otra descomposición que contenía los procesos estacionarios de segundo orden.

(I) A. Kolmogorov: «Sucesiones estacionarias en los espacios de Hilbert». Trabajos de estadística, dos números, 1957.

(II) M. Loéve: «Probability Theory». Van Nostrand, 1963, pág. 464.

(III) J. L. Doob: «Stochastic Processes». 7.ª edic., pág. 636. Wiley.

(IV) P. Lévy: «Processes Stochastiques et Mouvement Brownien». 3.ª edic., 1965, Gauthier Villars, págs. 123 y 298.

(V) H. Cramer: «On the Theory of the Stationary Processes. Annals of Mathematics», 41. 1940.

(VI) A. N. Kolmogorov: «Stationary Sequences in Hilbert's Space». Bulletin Moskovskogo Gosudarstvenngo Universiteta, Matematika 2 (1941).

(VII) Kari Karhunen: «Métodos lineales en el cálculo de probabilidades». Consejo Sup. Inv. Científicas, 1952.

(VIII) M. Loéve, o. c.

Kari Karhunen (I) determina las condiciones para que el proceso $\xi(t)$ se representable en forma espectral:

$$\xi(t)=\int_R f(t,\ a)d\zeta(a) \qquad a\in R \tag{5}$$

siendo $\zeta(S)$ una función de conjunto aleatoria definida en R denominada función espectral con las propiedades fundamentales siguientes:

1.ª Para cualesquiera conjuntos disjuntos $\forall S_i,\ S_j \in R$ se verifica:

$$\zeta(S_i+S_j)=\zeta(S_i)+\zeta(S_j) \tag{6}$$
$$S_i \cap S_j=\phi$$
$$i\neq j$$

2.ª Si en el conjunto R y respecto a él definimos la medida:

$$E\zeta(S_i)\overline{\zeta(S_j)}=M(S_i\cap S_j) \tag{7}$$

La (7) es válida para $\forall S_i$ y si fuesen los conjuntos disjuntos

$$S_i\cap S_j=\phi$$

la medida sería nula. La función aleatoria $\zeta(S_i)$ es en este caso ortogonal a $\zeta(S_j)$.

Si $S_i=S_j=S$ (los dos conjuntos coinciden) la medida es:

$$E\zeta(S)\overline{\zeta(S)}=M(S)=//\zeta(S)//^2 \tag{8}$$

denominándose $//\zeta(S)//\ \frac{1}{2}$ norma.

Si los conjuntos $S_1=(-\infty,\ \lambda)$ y $S_2=(\lambda,\ \lambda+d\lambda)$ la (8) sería la f. de distribución espectral.

13. El ruso Yu A. Rozanov (II) deduce representaciones similares a las de Karhunen para procesos estocásticos lineales, siendo $f(t,\ a)$ una función ponderativa no aleatoria.

(I) Kari Karhunen: «Zur Spektraltheorie Stochasticher Prozese». Aun. Acad. Sei. Fenicar, Se. A. I. Math. Phys. 34, año 1946. Ver obra «Métodos lineales en el cálculo de probabilidades», pág. 49, año 1952.
(II) Yu A. Rozanov: «Procesos aleatorios», Edit. Mir. Moscú 1973, págs. 217 y 218.

14. Muchos procesos estocásticos pueden ser representados por la (5) y todos los procesos estacionarios de orden dos son casos particulares de esta expresión.

15. Los avances técnicos (particularmente en telecomunicación) impulsaron a variar el método clásico de estudio de las series cronológicas.

16. En el año 1950 John W. Tukey, en Estados Unidos y M. S. Barttet, en Inglaterra, propusieron esta técnica esencialmente en su forma actual.

El mismo J. M. Tukey (I) escribe:

«En 1948 ó 1949 ocurrió que H. T. Burdenbom desarrolló mejor los cálculos de las señales de radar en el Bel Labs Whippany Laboratory. Había tomado algunos registros de datos y deseó mostrar su potencia espectral en una reunión de California. Así, él pidió al grupo de matemáticos de Murray Hill Laboratory le calcularan un espectro. Fui inocente —dice—, pero conocía que, en la población, la covarianza era la transformada de la potencia espectral. Así calculamos un número razonable de muestras de covarianzas; hicimos la transformada del coseno y dibujamos el resultado. Richard Harning vio el dibujo e inmediatamente dijo que suavizando con pesos 1/4, 1/2 y 1/4 produciría una curva mucho más fina.»

17. Hemos transcrito, por su importancia, estas manifestaciones de Tukey porque indican, además de la nueva era del estudio de las series cronológicas, las sucesivas etapas a partir de los resultados experimentales: primero recogida y elaboración de datos del proceso; después los cálculos correspondientes de medias muestrales, covarianzas y, a continuación, efectuar la transformada del coseno añadiendo complementariamente ciertas ponderaciones para la uniformidad de la gráfica.

Los cálculos no fueron otros que los siguientes:

$$f(\lambda)=\frac{1}{2\pi}\sum_{r=-\alpha}^{+\alpha} B(r)\cos r\lambda \qquad (9)$$

siendo $B(r)=B(-r)$ en el caso de ser estacionaria y real.

Al tomar las covarianzas experimentales $\hat{B}(r)$ introdujeron unas ponderaciones para estimar mejor $\hat{f}(\lambda)$. La función $\hat{f}(\lambda)$ es la derivada de lo que llamamos la función de distribución espectral.

(I) Bernard Harris: «Advanced Seminar on Spectral Analysis of time Series». Trabajos monográficos University Wiscousin, 1966, pág. 27. (Trabajo de J. W. Tikey: «An Introduction to the Calculations of Numerical Spectrum Analysis).

18. Así se inicia una nueva metodología para el estudio de las series de tiempo. Se publican artículos en revistas técnicas y se presentan en congresos y simposios numerosos trabajos de investigación sobre analisis espectral.

19. En estos trabajos se estudia el proceso $\xi(t)$, *no en el dominio del tiempo sino en el dominio de las frecuencias*; es decir: las *partes fundamentales del espectro que más influyen en el comportamiento del proceso.*

20. Quienes primeramente estudian la predicción y filtraje de procesos estocásticos fueron Kolmogorov (I) y Wiener (II) basados en las funciones de densidades espectrales, y las funciones de covarianzas.

El trabajo de Wiener apareció en circulación limitada en febrero de 1942, pero ya Kolmogorov había publicado su trabajo en el Bulletin de l'Academie des Sciences U.R.S.S. en 1941.

El autor (III) ha cooperado en el estudio de esta materia deduciendo fórmulas generales importantes.

21. En resumen: todo lo expuesto en esta Sección trata de una composición o descomposición del proceso o de la función de covarianza. Actualmente a esta teoría se la denomina «Teoría Múltiple y Descomposición Canónica» y puede representarse el proceso en forma de una integral estocástica, como:

$$x(t) = \int_{-\alpha}^{t} g(t, u) dZ(u) \tag{10}$$

o inclusive para una clase más amplia de procesos, puede representarse en la forma:

$$x(t) = \sum_{n=1}^{N} \int_{-\alpha}^{t} g_n(t, u) dZ_n(u) \tag{11}$$

permitiendo estas representaciones por las propiedades de conceptos superiores del Espacio de Hilbert (IV). El número N de términos requerido es el orden de la multiplicidad.

La integral estocástica es del tipo de Stieltjes; $g_n(t, u)$ una función no aleatoria y $Z_n(u)$ una función aleatoria espectral con propiedades expuestas en el número 12 de esta sección.

(I) A. Kolmogorov: «Interpolation und Extrapolation von Stationaren Zuflligen Folgen». Bull. Acad. Sc. (URSS) Ser. Maths. págs. 3-14, 1941.

(II) Norbert Wiener: «Extrapolation, Interpolation and Smoothing of Stationary Time Series». The M.I.T. Press. Massachussetts. Institute of Technology. Cambridge, 1970 (8.ª edic., la 1.ª en 1949).

(III) F. J. Urbelz: «Interpolación, extrapolación y filtraje», Anales del Instituto de Actuarios, 1978.

(IV) Anthony Efhremides and Tomas: «Random Processes. Multiplicity. Theory and Canonical Descompositions». Douden, Hutchinson, Pensilvania. Trabajo de H. Carmer, págs. 7 y 8.

CAPITULO SEGUNDO

NOCIONES DE ESPACIOS DE HILBERT

Sección 1.ª

Introducción

1. En el Capítulo Primero, Sección 7.ª, expusimos la importancia que tienen los conceptos de los espacios abstractos y especialmente hicimos referencia al espacio de Hilbert.

Los elementos que forman los espacios abstractos, pueden ser puntos, vectores, funciones, variables y funciones estocásticas que gozan determinadas propiedades.

2. Un espacio abstracto se define mediante un conjunto de axiomas que enuncian las propiedades básicas de los elementos pertenecientes al espacio.

Merece especial atención nuestro estudio, los espacios de Hilbert, cuya definición y axiomática estudiamos en las Secciones siguientes. También citamos las propiedades que hemos de utilizar prescindiendo de sus demostraciones (I).

Sección 2.ª

Definición del espacio de Hilbert

Definiremos espacio de Hilbert (que representaremos por H), a todo conjunto de elementos $x, y, \ldots \in H$ que cumplen los axiomas que se enuncian en la sección siguiente.

(I) Para aclaraciones remitimos al lector a la bibliografía consultada y a mi trabajo «Aplicaciones del espacio de Hilbert a la estadística». Anales del I. Actuarios, 1975.

CAPITULO SEGUNDO

NOCIONES DE ESPACIOS DE HILBERT

Sección 1.ª

Introducción

1. En el Capítulo Primero, Sección 7.ª, expusimos la importancia que tienen los conceptos de los espacios abstractos y especialmente hicimos referencia al espacio de Hilbert.

Los elementos que forman los espacios abstractos, pueden ser puntos, vectores, *funciones*, *variables y funciones estocásticas* que gozan determinadas propiedades.

2. Un espacio abstracto se define mediante un conjunto de axiomas que enuncian las propiedades básicas de los elementos pertenecientes al espacio.

Merece especial atención nuestro estudio, los espacios de Hilbert, cuya definición y axiomática estudiamos en las Secciones siguientes. También citamos las propiedades que hemos de utilizar prescindiendo de sus demostraciones (I).

Sección 2.ª

Definición del espacio de Hilbert

Definiremos espacio de Hilbert (que representaremos por H), a todo conjunto de elementos $x, y, ... \in H$ que cumplen los axiomas que se enuncian en la sección siguiente.

(I) Para aclaraciones remitimos al lector a la bibliografía consultada y a mi trabajo «Aplicaciones del espacio de Hilbert a la estadística». Anales del I. Actuarios, 1975.

Sección 3.ª

Axiomática del espacio de Hilbert

1.º *Axioma de grupo*

En el espacio H está definida una operación interna que denominaremos suma (indicada $+$) que goza de las propiedades siguientes:

Para $\forall x,\ y \in H$.

G.1. $x+y \in H$

G.2. $x+\phi=x \in H$ (ϕ elemento neutro)

G.3. $\forall x \in H$ existe $\exists x' \in H \Rightarrow x+x'=\phi$. A x' se representa por $-x$

G.4. $x+y=y+x$ propiedad conmutativa.

G.5. $x+y+z=x+(y+z)$ propiedad asociativa.

Es decir: H es un grupo abeliano.

2.º *Axioma de linealidad*

En el espacio H está definida una operación externa respecto al cuerpo C de números complejos (a los que denominaremos escalares) y que representamos por ax tal que:

$$ax \in H \qquad \forall x \in H \qquad \text{y} \qquad \forall a \in C$$

Esta operación goza de las propiedades siguientes:

Para $\forall x,\ y \in H$ y $\forall a,\ b \in C$

$$\left.\begin{array}{ll} \text{L.1.} & a(x+y)=ax+ay \\ \text{L.2.} & (a+b)x=ax+bx \end{array}\right\} \text{ propiedades distributivas}$$

L.3. $a(bx)=(ab)x$ propiedad asociativa.

L.4. $0 \cdot x=\phi$ (ϕ elemento neutro de H).

L.5. $1 \cdot x=x$.

Notas.

1.ª El símbolo $+$ en L.2. se utiliza en dos sentidos:

a) Para la suma de escalares; y
b) Para la suma definida en el axioma 1.º

2.ª De las propiedades de la operación lineal L.2. y de la suma G.3.$\Rightarrow x'=-x$.

3.ª Como $\forall a,\ b \in C$ $ax \in H$ y $by \in H$ y la definición de suma, dada en 1.ª, tenemos:

$$ax+by \in H$$

y si $a=1$, y $b=-1$, se cumple $x-y \in H$.

4.ª Un espacio E en el que se ha definido la ley de composición interna denominada adición asociado con otro espacio S de números (o escalares) con la operación de linealidad, se denomina espacio vectorial.

Si $S=R$ el espacio vectorial es real; si $S=C$ el espacio E es complejo.

3.º *Axioma métrico* (hermítico o hilbertino)

1. En el espacio H se define una tercera operación a cuyo resultado denominamos indistintamente producto interno o escalar. El producto interno es una aplicación de $H \times H$ en S que indicamos: (x, y) tal que para $\forall x, y, z \in H$ satisface las siguientes propiedades:

M.1. $(x, y)=\overline{(y, x)}$ propiedad hermítica (1)

M.2. $(ax, y)=a(x, y)$ $(a \in C)$

M.3. $(x+y, z)=(x, z)+(y, z)$ propiedad distributiva.

M.4. $(x, x)\geqslant 0$

Suele leerse (x, y), «x escalar y».

2. Se denomina «norma» o longitud de x indicada por $//x//$ a la raíz cuadrada positiva del producto (x, x):

$$//x//=(x, x)^{\frac{1}{2}}$$

3. Se llama distancia del elemento x al elemento y, indicada por $//x-y//$, a la norma de $x-y$.

4.º *Axioma de Cauchy*

Para toda sucesión $\{x_n\}$ de elementos de H, dado un $\varepsilon>0$ arbitrariamente pequeño si existe un $N(\varepsilon)$ tal que para $\forall n, m>N(\varepsilon)$ es

$$//x_n-x_m//<\varepsilon \tag{1}$$

existe un elemento $X \in H$ tal que se verifica

$$\lim_{n\to\infty} //X_n-X//=0 \tag{2}$$

(1) El conjugado de un número lo indicamos con una barra encima.

Este axioma equivale a admitir que el espacio H es completo. Cuando se cumple (2) diremos que la sucesión $\{X_n\}$ de elementos de H converge hacia el límite X y escribiremos

$$\lim_{n\to\alpha} = X \qquad \text{o bien} \qquad X_n \underset{n\to\alpha}{\to} X \tag{3}$$

Sección 4.ª

Espacios especiales de Hilbert

En las secciones siguientes se citan cuatro tipos de gran interés en diversas teorías.

Sección 5.ª

Espacio euclídeo -n- .dimensional

1. Este espacio está constuido por todas las -n- eplas:

$$x=(x_1 \ x_2 \ \dots \ x_n) \tag{1}$$

2. En el espacio $\{x\} \in H_n$, definiremos las siguientes operaciones:

a) Para $\forall x, \ y \in H_n \Rightarrow x+y=(x_1+y_1, \ x_2+y_2, \ \dots, \ x_n+y_n)$ (2)

b) Para $x \in H_n$ y $a \in C \Rightarrow ax=(ax_1, \ ax_2, \ \dots, \ ax_n)$ (3)

c) Para $\forall x, \ y \in H \Rightarrow (x, \ y)=\sum_{1=1}^{n} x_i \bar{y}_i$ (4)

Inmediatamente se prueba que el conjunto $\{x\}$ (1) satisface los axiomas de los espacios de Hilbert citados en la Sección 3.

Sección 6.ª

Espacio de las sucesiones de cuadrados sumables

1. Diremos que la sucesión $\{x_n\}$ de números complejos

$$x=(x_1 \ x_2 \ \dots \ x_n \ \dots) \tag{1}$$

es de cuadrado sumable cuando

$$\Sigma /x_n/^2 < \infty$$

2. El espacio formado por todas las sucesiones (1) de cuadrado sumable es un espacio de Hilbert con las operaciones definidas en 2. de la Sección 5, extendidas indefinidamente.

Inmediatamente se prueba que el conjunto (1) satisface los cuatro axiomas de los espacios de Hilbert.

Sección 7.ª

Espacios funcionales

1. En esta Sección nos referiremos a funciones $f(\cdot)$ definidas en el intervalo (a, b) (I) que sean de cuadrado integrables, es decir, que cumplan:

$$\int_a^b /f \quad (x)/^2 \, dF(x) < \infty \tag{1}$$

donde $F(x)$ es una función no decreciente (II).

En lo sucesivo nos referiremos al conjunto de todas estas funciones escribiendo $L_2(a, b)$ o simplemente L_2.

2. En el conjunto $L_2(a, b)$ definiremos el producto interno (f, g) para $\forall f, g \in L_2$ del siguiente modo:

$$(f, g) = [f(\cdot), g(\cdot)] = \int_a^b f(x)\overline{g(x)}\, dF(x) \tag{2}$$

o abreviadamente:

$$(f, g) = \int_a^b f\bar{g}\, dF \tag{2'}$$

3. El espacio formado por todas las funciones de L_2 que satisfagan

(I) El intervalo puede ser cerrado, abierto, semiabierto, e inclusive infinito. Puede extenderse este concepto a conjuntos borelianos de puntos $S \in R$ (de una o más dimensiones). Véase Apéndice n.º 1.

(II) Un caso particular es cuando $F(x) = x$. En nuestro estudio tienen especial interés estos espacios.

(1) con la operación producto interno definida en (2) es un espacio de Hilbert.

Inmediatamente se comprueba que en $L_2(a, b)$ satisfacen los cuatro Axiomas del espacio de Hilbert (I).

Sección 8.ª

Espacios estocásticos de Hilbert

1. En esta Sección nos referiremos al conjunto H de variables aleatorias $\{x\}$ definidas conjuntamente en el espacio probabilizable $\{\Omega, \mathscr{B}\}$ que admitan momento de segundo orden finito; es decir, a las variables aleatorias x de segundo orden:

$$E\{/x/^2\} < \infty \tag{1}$$

donde E es el operador esperanza matemática.

Si Ω es el espacio de sucesos $\omega \in \Omega$ siempre podremos identificar a x como una aplicación $f(\cdot)$ de Ω en C, es decir, una función numérica $f(\cdot)$ de $\omega \in \Omega$ representativa de la *v.a.* y con una medida de probabilidad $m(d\omega)$. La (1) puede escribirse (II):

$$E\{/x/^2\} = \int_\Omega /f(\omega)/^2\, m(d\omega) \tag{1'}$$

y dos v.a. se considerarán idénticas si:

$$E\{/x-y/^2\} = 0 = \int_\Omega /f(u) - g(u)/^2\, m(d\omega) \tag{2}$$

2. En el conjunto de variables aleatorias $\{x\}$ definiremos el producto interno (x, y) para $\forall x, y \in H$ del siguiente modo:

$$(x, y) = E x \bar{y} \tag{3}$$

3. El espacio de las v.a. consideradas en 1. y con el producto interno definido en (3), es un espacio de Hilbert.

(I) Si fuera un conjunto $L \in L(a, b)$ y no cumplieran las funciones pertenecientes a L el Axioma 4.º formaría un espacio denominado «prehilbertiano». (Introducción al Espacio de Hilbert»: S. K. Berberian, pág. 32. Teide.)

(II) R. Fortet: «Théorie des Probabilités». Eyrolles, pág. 71, 1974.

Inmediatamente se comprueba que en H se satisfacen los cuatro Axiomas del espacio de Hilbert, con las convenciones expuestas y a veces se representa al espacio H de las v.a. de segundo orden por $(L_2, \Omega\ \mathscr{P})$.

Sección 9.ª

Dimensionalidad del espacio de Hilbert

1. Definición I. Cuando en H existen n elementos (pero no más de n) linealmente independientes, diremos que H es de n dimensiones. En este caso es usual representar el espacio por H_n.

Se demuestra que para los espacios H_n el Axioma 4.° es consecuencia de los Axiomas 2.° y 3.° de la Sección 3.ª (I).

Es espacio euclídeo definido en la Sección 5.ª es un espacio -n- dimensional.

2. Definición II. Cuando, cualquiera que sea n, existen en H más de n elementos independientes, diremos que H es de dimensión infinita.

Los espacios definidos en las Secciones 6.ª, 7.ª y 8.ª son de dimensionalidad infinita.

Sección 10.ª

Propiedades del producto interno

De los axiomas de la Sección 3, se deducen las siguientes propiedades:

Para $\forall x,\ y,\ z \in H$ y $\forall a,\ b, \in C$

1.ª
$$(x,\ ay) = \bar{a}(x,\ y) \tag{1}$$

2.ª
$$(ax+by,\ z) = a(x,\ z) + b(y,\ z) \tag{2}$$

3.ª Propiedad distributiva:
$$(ax+by,\ ax+by) = /a/^2(x,\ x) + /b/^2(y,\ y) + \bar{a}b(y,\ x) + \bar{b}a(x,\ y) \tag{3}$$

4.ª Desigualdad de Schwartz:
$$(x,\ x)(y,\ y) \geqslant (x,\ y)(y,\ x) \tag{4}$$

5.ª Desigualdad triangular:
$$//x+y// \leqslant //x// + //y// \tag{5}$$

(I) J. von Neuman, o. c. pág. 31.

Sección 11.ª

Ortogonalidad

1. Dos elementos $x,\ y \in H$ se dicen ortogonales cuando

$$(x, y)=0 \tag{1}$$

Para indicar que x, y son ortogonales, se escribe también:

$$x/y$$

2. Si los elementos de un conjunto son ortogonales dos a dos, diremos que forman un sistema ortogonal.
3. Diremos que $x \in H$ está normalizado cuando su norma es la unidad; es decir:

$$(x,\ x)=1 \tag{2}$$

4. Diremos que un conjunto de elementos $\{e_h\} \in H$ $(h=1,\ 2,\ 3,\ ...)$ forma un sistema ortonormal S cuando se cumplan las siguientes relaciones:

$$(e_h,\ e_k)=\delta_{hk}=\begin{cases}1 & k=h\\ 0 & k\neq h\end{cases} \tag{3}$$

δ_{hk} es el símbolo de Kronocker; es decir: los elementos distintos del sistema son ortogonales entre sí y su módulo es la unidad.

5. Diremos que un sistema ortonormal $S(H)$ es completo cuando el único elemento de H ortogonal al sistema S es el elemento neutro ϕ de H.

6. Llamaremos base de un espacio de Hilbert a un conjunto de elementos de H que forma un sistema completo ortonormal.

7. Se demuestra que, salvo un caso muy particular, todo espacio de Hilbert tiene infinitas bases.

8. Dos elementos e_i, e_j, cualesquiera de una base son linealmente independientes.

9. Si $x_1\ x_2\ ...\ x_n$ son elementos de H y forman un sistema ortogonal, se tiene:

$$//x_1+x_2+\cdots+x_n//^2=//x_1//^2+//x_2//^2+\cdots+//x_n//^2 \tag{4}$$

Esta propiedad se generaliza para $n\to\infty$ cuando la sucesión de normas $\{//x_n//\}$ sean de cuadrado convergente:

$$//x//^2=\lim_{n\to\infty}//\sum_{i=1}^{n}x_i//^2=\lim_{n\to\infty}\sum_{i=1}^{n}//x_i//^2 \tag{5}$$

10. Si $x \in H$ es ortogonal a todo elemento de un subconjunto L de H diremos que x es ortogonal a L y escribiremos:

$$\underline{x/L}$$

Si L_1 y L_2 son dos subconjuntos de H tales que todo elemento de L_1 es ortogonal a todo elemento de L_2 decimos que ambos subconjuntos son ortogonales y escribiremos:

$$\underline{L_1/L_2}$$

11. Dado un subconjunto de elementos $\{x_1\ x_2\ \ldots\ x_k\} \in H$ linealmente independientes y para $\forall a_1\ a_2\ \ldots\ a_k \in C$ (cuerpo de números complejos) se denomina multiplicidad lineal M al conjunto de elementos de H engendrados por todas las combinaciones lineales:

$$\sum_{i=1}^{k} a_i x_i \in M \subset H$$

12. Dos multiplicidades lineales M_1 y M_2 son ortogonales cuando todo elemento de M_1 es ortogonal a todo elemento de M_2. En este caso escribiremos:

$$\underline{M_1/M_2}$$

CAPITULO TERCERO

PROCESOS ESTOCASTICOS

Sección 1.ª

Definición

En el Capítulo Primero, Sección 6.ª, someramente tratamos de la evolución histórica y conceptual.

Los procesos estocásticos han sido definidos de diferentes formas.

1. Khintchine (I) lo define como un sistema uniparamétrico de variables aleatorias.

2. La concepción de Doob (II) es eminentemente abstracta; todo proceso regido por leyes de probabilidad o una familia de variables aleatorias:

$$\{\xi(t),\ t \in T\} \tag{1}$$

dependiendo de un parámetro t que si representa el tiempo se denomina proceso y si no función aleatoria.

Por el conjunto índice:

$$\{t\} = T \tag{2}$$

la familia puede ser de tipo discreto, numerable o continuo.

Para todo conjunto de determinaciones de $\xi(t)$ se puede asociar una variable simbólica $\omega \in \Omega$ lo que permite a Doob considerar $\xi(t)$ para un valor particular de ω como una función cierta $\xi(\omega, t)$ de la variable t, pudiendo representar una trayectoria para cada valor de $\omega \in \Omega$.

(I) A. Khintchine: «Korrelationstheorie der Stationaren Stochastischen Processe. Mathematische Annalen, 109, 1934.

(II) J. L. Doob: «Stachastic Processes», Wiley & Sons, pág. 46, 2.ª edición, 1967.

3. Para Lévy (I) un proceso estocástico es, en principio, un procedimiento de definición de una función aleatoria $\xi(t)$ en el tiempo t en el que el azar interviene en cada instante, cualquiera que sea el valor $t \in T$ y su incremento τ; el conocimiento de $\xi(t)$ del instante inicial t_0 hasta el instante t_1 no determina los valores de esta función en el intervalo $(t, t+\tau)$ y estos valores son aleatorios, evolucionando en el tiempo. Teóricamente puede definirse dando un conjunto Ω de determinaciones posibles de $\xi(t)$ y definiendo una distribución de probabilidad en Ω.

4. Citaremos a Karhunen (II) quien concibe los procesos estocásticos «como funciones cuyos valores son puntos de un espacio abstracto». Establece condiciones restrictivas: son funciones integrables de cuadrado (L_2); tienen dispersión finita, por lo que pueden considerarse como elementos de un espacio de Hilbert.

5. La definición de A. Fernández de Trocóniz (III) supone una previa definición de un espacio de probabilidades:

$$\{\Omega, A, Pr\} \tag{3}$$

y el conjunto de índices (espacio paramétrico):

$$T = \{t\} \tag{4}$$

en el que para cada $t \in T$ se supone definida una variable aleatoria $\xi_t(\omega)$ o $\xi(t, \omega)$.

El profesor Fernández de Trocóniz define «proceso aleatorio o estocástico» a una familia de variables aleatorias,

$$\xi(t, \omega) \qquad t \in T \tag{5}$$

definidas conjuntamente sobre el espacio de probabilidades (3) de modo que a cada índice $t \in T$ corresponda una variable aleatoria.

En un proceso estocástico cada experiencia muestral determina una trayectoria a la que identificamos mediante una variable simbólica aleatoria $\omega \in \Omega$ es decir, para cada experiencia (valor concreto de ω) define una función del parámetro $t \in T$:

$$\{\xi(t), \omega'\}_T \tag{6}$$

(I) P. Lévy: «Processus Stochastianes et Mouvement Prownien», Gauthier-Villar, 1065, pág. 27, y «Theorie de L'Addition des Variables Aleatoires», G. Villars, 1954, pág. 360.

(II) Karl Karhunen: «Métodos lineales en el cálculo de probabilidades». Consejo Superior de Investigaciones Científicas, 1952, pág. 10.

(III) A. Fernández de Trocóniz: CAP. 38. «Conceptos generales de procesos estocásticos. (Probabilidades, estadísticas y procesos aleatorios».)

Al conjunto de las infinitas trayectorias posibles se le suele llamar proceso estocástico.

6. De acuerdo con cuanto antecede, de un proceso estocástico $\{\xi(\omega, t)\ t \in T\}$ se tiene:

1.ª Una familia de funciones de $t(t$ y ω variables). Es el proceso propiamente dicho:

$$\{\xi(t),\ t \in T\}_\Omega \tag{7}$$

2.ª Una función de t (t variable y ω fijo) o también una muestra experimental, llamada realización o trayectoria.

$$\{\xi(t),\ \omega \in \Omega\}_T \tag{6'}$$

3.ª Una variable aleatoria (t fijo, ω variable); y, por último.

4.ª Un dato numérico (t y ω fijos).

La teoría de los procesos estocásticos puede considerarse como la parte dinámica de la Estadística Teórica.

Sección 2.ª

Funciones de Distribución. Condiciones de consistencia de Kolmogorov

1. Sea $(\Omega, \mathscr{A}, \mathscr{P})$ un espacio de probabilidad y sea el proceso

$$\{\xi(t), \qquad t \in T\} \tag{1}$$

definido sobre referido espacio.

Si elegimos cualquier subconjunto finito

$$\{t_1\ t_2\ \dots\ t_n\} \tag{2}$$

del conjunto índice T, los valores que tome (1) en el conjunto (2) nos da unas variables aleatorias ordenadas:

$$\{\xi(t_1), \quad \xi(t_2)\ \dots\ \xi(t_n)\} \tag{3}$$

denominadas vector aleatorio.

Con el vector aleatorio (3) perteneciente a Ω formaremos sucesos S como lo hicimos en (6) de Sección 6.ª del Capítulo Primero.

$$S = \{\xi(t_1) \leqslant x; \quad \xi(t_2) \leqslant x_2,\ \dots,\ \xi(t_n) \leqslant x_n\} \tag{4}$$

y que pertenece al álgebra de sucesos $\mathscr{A}$. La aplicación unívoca de cada suceso con la medida perteneciente a $\mathscr{P}$ la denominamos función de distribución de orden -n- y la representaremos así:

$$F_n(x_1, x_2 \dots x_n; t_1, t_2 \dots t_n) = \\ = P(\xi(t_1) \leqslant x_1, \xi(t_2) \leqslant x_2 \dots \xi(t_n) \leqslant x_n) \quad (5)$$

Si existe la función (5) para todo subconjunto (2) perteneciente a T y n arbitrario, habremos definido el proceso estocástico (1) en sentido de Slutsky (I).

2. Para que la función (5) sea una función de distribución tiene que cumplir las siguientes condiciones de consistencia de Kolmogorov (II).

a) *Condiciones de simetría*

La función de distribución (5) correspondiente al proceso estocástico (1) para $n \in N$ arbitrario y

$$\forall (t_1, t_2 \dots t_n) \in T \quad (7)$$

es invariante para toda permutación de los pares (x_i, t_i), (x_j, t_j); esto es:

$$F_n(x_1 \dots x_i \dots x_j \dots x_n; t_1 \dots t_i \dots t_j \dots t_n) = \\ = F_n(x_1 \dots x_j \dots x_i \dots x_n; t_1 \dots t_j \dots t_i \dots t_n) \quad (8)$$

b) *Condiciones de compatibilidad*

Toda función de distribución multidimensional (5) de orden $n+m$ deberá cumplir siempre:

$$\lim_{\substack{x_{n+1} \to \infty \\ \dots \\ x_{n+m} \to \infty}} F_{n+m}(x_1 \dots x_{n+m}; t_1 \dots t_{n+m}) = F_n(x_1 \dots x_n; t_1\ t_2 \dots t_n) \quad (9)$$

3. En resumen: Las funciones de distribución de n dimensiones generan funciones de distribución de órdenes inferiores. El número de funciones de distribución de orden k obtenidas de una de orden $n (n > k)$ son las combinaciones $\binom{n}{k}$ y cada una de ellas deberá igualmente cumplir las condiciones de consistencia indicadas.

(I) G. Arnaiz Vellando: Introducción a la Estadística Teórica, I, Lex Nova, 1965, página 351.

(II) A. N. Kolmogorov: «Foundations of the Theory of Probability». Chelsea, 1956, página 29.

4. Se llama sistema de funciones de distribución de una -n- dimensional a todas las funciones de distribución obtenidas a partir de la dada. Así, el número total será:

$$\sum_{k=1}^{n} \binom{n}{k} = 2^n - 1 \tag{10}$$

5. Kolmogorov (I) demuestra que todo sistema de funciones de distribución que cumplan las condiciones de consistencia expuestas define una función de probabilidad en el espacio H_n y que puede extenderse ilimitadamente, satisfaciendo los axiomas del cálculo de probabilidades, lo que determina la existencia unívoca de un proceso estocástico $\{\xi(t),\ t \in T\}$ definido sobre Ω (II).

Sección 3.ª

Procesos multivariantes

A un conjunto de procesos estocásticos definidos sobre el mismo espacio, le llamaremos proceso multivariante o también vector proceso y lo indicaremos de cualquiera de estas dos formas:

$$\xi(t) = \{\xi_k(t), \quad t \in T\} \qquad (k = 1,\ 2,\ \ldots,\ n)$$

$$\xi(t) = \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \\ \ldots \\ \ldots \\ \xi_n(t) \end{bmatrix} \qquad t \in T$$

Estos procesos pueden ser de parámetro discreto o continuo, según sea el conjunto índice.

Sección 4.ª

Funciones de covarianza

1. Prescindiremos de definir las características de procesos, como pueden ser: momentos, momentos centrados, función característica, etcétera, aunque alguno de ellos ha sido expuesto de forma elemental en el Capítulo I, Sección 10.

(I) A. Kolmogorov: «Foundations of the Theory of Probability», o. c. pág. 29.
(II) J. L. Doob: «Stochastic Processes», o. c., pág. 609 y sig.

2. La función de covarianza del proceso $\{\xi(t),\ t \in T\}$ fue definida en la Sección 10.ª del Capítulo Primero. No obstante, algunos autores llaman función de covarianza a la expresión:

$$B(t,\ s) = E\{\xi(t)\quad \overline{\xi(s)}\} \tag{1}$$

A la (1) se la denomina también «autocovarianza». Nosotros la denominaremos sencillamente función de covarianza.

3. Cuando intervengan dos procesos:

$$\{\xi_j(t),\quad \xi_k(t),\quad t \in T\} \tag{2}$$

definiremos función de covarianza mutua a la expresión:

$$B_{jk}(t,\ s) = E\,\xi_j(t)\,\overline{\xi_k(s)} \tag{3}$$

Si son -n- los procesos ($j,\ k = 1,\ 2,\ 3,\ \ldots,\ n$) puede formarse el vector proceso $\xi(t)$ y la esperanza de la matriz $n \times n$, $\xi(t)\,\overline{\xi(s)}$ está formada por funciones de covarianza mutua semejante a la (3) excepto para la diagonal principal que serían las «autocovarianzas de los procesos componentes del vector proceso $\xi(t)$:

$$B(t,\ s) = E\,\xi(t)\,\overline{\xi(s)}^{\,1} = \{B_{jk}(t,\ s)\} \tag{4}$$

4. *Propiedades:*

4.1. La función de covarianza $B(t,\ s)$ de un proceso cualquiera $\{\xi(t),\ t \in T\}$ de segundo orden, es siempre una función de covarianza centrada, puesto que existe otro proceso de esperanza nula con igual función de covarianza.

4.2. Toda función de covarianza es hermítica:

$$B(t,\ s) = \overline{B(s,\ t)} \tag{5}$$

4.3. En todo proceso real:

$$B(t,\ s) = B(s,\ t) \tag{6}$$

4.4. En los puntos diagonales $(t,\ t)$ la función de covarianza coincide con la varianza del proceso:

$$\sigma^2\{\xi(t)\} = B(t,\ t) \tag{7}$$

la cual es siempre real y no negativa.

4.5. Desigualdad de Schwartz:

$$/B(t,\ s)/^2 \leqslant B(t,\ t) \quad B(s,\ s) \tag{8}$$

4.6. La continuidad de $B(t, s)$ en todo punto diagonal (t, t) implica la continuidad de la función de covarianza $B(t, s)$ en $\forall(t, s) \in T$.

Precisamente la continuidad de la f. de covarianza implica la única clase de continuidad de los procesos estocásticos que estudiaremos y por eso la definiremos en la Sección siguiente.

4.7. Para que una función compleja $B(t, s)$ sea una función de covarianza en $T \times T$, es necesario y suficiente que sea de tipo no negativo (I).

4.8. Las funciones de covarianza son una clase cerrada bajo las operaciones de adición, multiplicación y paso al límite (I).

4.9. La matriz de la covarianza (4) es definida no negativa.

$$\sum_{j,\ k} B_{jk}(t,\ t) u_j \bar{u}_k = E/ \sum_{j=1}^{n} \xi_j(t) u_j/^2 \geqslant 0 \tag{9}$$

y es de tipo hermítico.

Sección 5.ª

Diferenciación e integración

1. De los distintos tipos de convergencia se aplica en procesos especialmente el de media cuadrática (m.c.)

Decimos que la sucesión ξ_n de v.a. que cumpla $E\{/\xi_n/^2\} < \infty$ tiende a la variable ξ en m.c. si:

$$\lim_{n \to \infty} E(/\xi_n - \xi/^2) = //\xi_n - \xi//^2 = 0 \tag{1}$$

La condición necesaria y suficiente de Cauchy es que:

$$\lim_{n,\ m \to \infty} //\xi_n - \xi_m// = 0 \tag{2}$$

(I) A. Fernández de Trocóniz: o.c. pág. 38.11.

2. Un criterio para conocer si la sucesión de variables ξ_n converge en m.c. a ξ es muy útil y es si, y solamente si:

$$\lim_{m,\ n\to\infty} E\,\xi_n\overline{\xi_m}=c \tag{3}$$

siendo c constante.

3. Un proceso

$$\{\xi(t), \quad t\in T\} \tag{4}$$

cuyo espacio paramétrico es continuo, es continuo en m.c. si

$$\lim_{t'\to t} E\{/\xi(t')-\xi(t)/^2\}=0 \tag{5}$$

4. La convergencia en m.c. del proceso $\xi(t')$ a $\xi(t)$ se escribe así:

$$\xi(t')\underset{t'\to t}{\overset{2}{=}}\xi(t) \tag{6}$$

5. Se demuestra que la razón aleatoria de incrementos:

$$\frac{\xi(t+h)-\xi(t)}{h} \tag{7}$$

tiende a un límite aleatorio m.c. (que se denomina derivada en t) si y solamente si existe, y es finita la derivada segunda de la función de covarianza:

$$\left.\frac{\partial^2 B(t,\ s)}{\partial t\,\partial s}\right]_{t=s} \tag{8}$$

Es consecuencia de la continuidad m.c. aplicando el criterio de Loeve, donde $B(t,\ s)$ es la función de covarianza del proceso.

6. Se prueban mediante límites m.c. las siguientes relaciones:

$$E\,\xi'(t)\,\overline{\xi(s)}=\frac{\partial B(t,\ s)}{\partial t} \tag{9}$$

y

$$E\,\xi(t)\,\overline{\xi'(s)}=\frac{\partial B(t,\ s)}{\partial s} \tag{9'}$$

7. El resultado se generaliza para las derivadas de órdenes superiores.

$$E\xi^{(n}(t)\overline{\xi^{(m}(s)}=\frac{\partial^{n+m}B(t,\ s)}{\partial^{n}t\,\partial^{m}s} \tag{10}$$

y la (10) nos indica la condición para que el proceso $\xi(t)$ tenga derivada n-m.c.:

$$\left.\frac{\partial^{2n}B(t,\ s)}{\partial^{n}t\,\partial^{n}s}\right]_{t=s}<C \tag{11}$$

8. La integral estocástica de Riemann representada por el símbolo:

$$\int_{\Omega}\xi(t)m(dt) \tag{12}$$

donde previamente se ha fijado el espacio de medida (Ω, $\mathscr{A}$, $\mathscr{P}$) si Ω es el eje real y m es la medida σ finita. Algunos representan $m(dt)$ por $dm(t)$ (I).

Si la notación de la I como límites de las sumas de Darboux en el eje real es:

$$I\stackrel{2}{=}\int_{a}^{b}\xi(t)m(dt) \tag{13}$$

Para la existencia de convergencia m.c. a una integral de Riemann, es condición suficiente que:

$$\iint_{a}^{b}B(t,\ s)m(dt)m(ds)<\infty \tag{14}$$

9. La integración estocástica m.c. de Riemann-Stieljes, de los procesos

$$\{\xi(t),\ \zeta(t),\qquad t\in T\} \tag{15}$$

de función de covarianza $B_{\xi}(t,\ s)$, $B_{\zeta}(t,\ s)$ siendo referidos procesos independientes, se define como el límite de las sumas aleatorias de Darboux, y se representa por:

$$I\stackrel{2}{=}\int_{a}^{b}\xi(t)\,d\,\zeta(\lambda) \tag{16}$$

(I) M. Loéve: o. c. págs. 120 a 140 y criterio de convergencia m.c. págs. 454 y siguientes.

Las condiciones para la convergencia m.c. suficiente es que:

$$\int\int_a^b B_\xi(t,\ s)d_\xi^2\, B(t,\ s) < \infty \tag{17}$$

Y la función de covarianza $B_\zeta(t,\ s)$ debe ser de variación acotada en $T \times T$.

$$\Sigma / \Delta_{hk} B(t,\ s) / < \infty \tag{18}$$

para cualesquiera subdivisión del intervalo de integración.

Sección 6.ª

Procesos incorrelacionados y ortogonales

1. Se denomina al proceso $\{\xi(t),\ t \in T\}$ de incrementos independientes (I) si para toda cuaterna de índices $t_1 < t_2 \leqslant t_3 < t_4$ las variables aleatorias de los incrementos $\xi(t_2) - \xi(t_1)$ y $\xi(t_4) - \xi(t_3)$ son independientes.

2. Se denomina al proceso $\{\xi(t),\ t \in T\}$ de incrementos $\xi(t_{i+1}) - \xi(t_i)$ ortogonales (II) si para toda cuaterna de indices $t_1 < t_2 \leqslant t_3 < t_4$ se cumple:

$$E\{[\xi(t_4) - \xi(t_3)]\, \overline{[\xi(t_2) - \xi(t_1)]}\} = 0 \tag{1}$$

restringiendo la finitud del momento de segundo orden (del proceso (III) o de su diferencia).

En estos procesos ortogonales la esperanza de la variable $/\xi(t) - \xi(s)/^2 (t > s)$ puede escribirse como una función de t y de s:

$$E\{/\xi(t) - \xi(s)/^2\} = F(t) - F(s) \tag{2}$$

siendo $F(t)$ monótona no decreciente y acotada.

La (2) desarrollada puede reducirse:

$$F(t) - F(s) = B(t,\ t) - B(t,\ s) - B(s,\ t) + B(s,\ s) \tag{3}$$

(I) A. Fernández de Trocóniz: «Probabilidades, estadistica, procesos aleatorios», página 39.2. Estudios Grafos.

(II) A. Fernández de Trocóniz: o. c. pág. 39.3.

(III) H. Cramer: «Stationary and Related S5ochastic Processes», Wiley, 1968, pág. 109.

y que si hacemos $t=t+dt$, y $s=t$, tenemos:

$$dF(t)=[d^2 B(t,\ s)]_{s=t} \tag{4}$$

3. Se denomina al proceso $\{\xi(t),\ t\in T\}$ de incrementos incorrelacionados si para toda cuaterna de índices $t_1<t_2\leqslant t_3<t_4$ los incrementos de las variables están incorrelacionados (IV).

$$\begin{aligned} &E\{[\xi(t_4)-\xi(t_3)]\,[\overline{\xi(t_2)-\xi(t_1)}]\}= \\ &=E[\xi(t_4)-\xi(t_3)]\,E\overline{\xi(t_2)-\xi(t_1)} \end{aligned} \tag{5}$$

Este proceso es más amplio que el indicado en 1, por cuanto aquél se refiere más bien a las leyes de distribución y está contenido en éste pero no al contrario. Como generalmente se opera con procesos de esperanza matemática nula, se comprueba que la (5) es igual a cero y en consecuencia el proceso sería también de incrementos ortogonales.

(IV) J. L. Doob: «Stochastic Processes». John Wiley, N. Y., pág. 97.

CAPITULO CUARTO

PROCESOS ESTOCASTICOS ESTACIONARIOS REPRESENTACIONES ESPECTRALES

Sección 1.ª

1. Procesos débilmente estacionarios de orden dos

Un proceso estocástico es débilmente estacionario de orden dos (o también en sentido amplio) cuando teniendo los dos primeros momentos finitos cumple las propiedades siguientes:

1.ª La esperanza matemática $E\xi(t)$ es independiente de t.

2.ª La función de covarianza $B(t, s)$ depende de la diferencia de tiempos.

3.ª Si el proceso es de parámetro t continuo, la función de covarianza $B(h)$ es continua en el origen e implica la continuidad en R.

Nuestro trabajo se concreta expresamente a este tipo de procesos. Demostramos las propiedades en los puntos siguientes:

2. Función de distribución de los dos primos órdenes

1. La función de distribución de primer orden no se altera por la estacionariedad del proceso:

$$F(x, t) = F(x, t+h) \qquad \forall t, t+h \in T \Rightarrow \tag{1}$$

$$F(x, t) = F(x, 0) \qquad (\text{para } h = -t) \tag{1'}$$

La (1') nos indica que $F(x)$ es independiente del tiempo y también:

$$dF(x, t) = dF(x) \tag{2}$$

2. La función de distribución de segundo orden cumplirá:

$$F(x_1\ x_2, t, s) = F(x_1\ x_2; t+h, s+h) \tag{3}$$
$$\forall t, s, t+h, s+h \in T$$

Si en (3) hacemos $h = -s$ tendremos:

$$F(x_1\ x_2;\ t,\ s) = F(x_1\ x_2;\ t-s) = F(x_1\ x_2;\ z) \qquad (3')$$

$$z = t - s$$

indicando la dependencia de la diferencia de tiempos.

3. **Media y varianza**

Por la (2) deducimos que la media es constante:

$$\alpha(t) = E\xi(t) = \int_s \xi(t)dF(x) = \int_s dF(x) = \alpha \qquad (4)$$

Igualmente la varianza:

$$\sigma_\xi^2(t) = E\{/\xi(t) - \alpha^2/\} = \int_s /x-\alpha/^2\, dF(x)$$

$$= \int_s /x/^2\, dF(x) - /\alpha/^2 \qquad (5)$$

4. **Función de covarianza**

La función de covarianza es función de la diferencia de tiempos:

$$B(t,\ s) = B(t-s) = \int_{R_2} x_1\ \bar{x}_2 dF^2(x_1\ x_2;\ t-s) \qquad (6)$$

$$\forall t,\ s \in T$$

5. **Función de covarianza mutua (I)**

Para dos procesos $\{\xi(t),\ t \in T\}$ y $\{\eta(s),\ s \in T\}$ que supondremos con esperanzas nulas, llamaremos esperanza mutua (I) a la expresión:

$$B_{\xi\eta}(t,\ s) = E\xi(t) \quad \overline{\eta(s)} \qquad (7)$$

Si el proceso es de tipo estacionario, sería:

$$B_{\xi\eta}(t,\ s) = B_{\xi\eta}(t-s) \qquad (8)$$

(I) Julio García Villalón: «Análisis espectral de series temporales en economía». Instituto de Actuarios, 1967, págs. 19 a 113. Denominación adoptada a propuesta del profesor Fernández de Trocóniz.

Sección 2.ª

Derivabilidad y desarrollos de procesos estacionarios

1. *Continuidad m.c.*

En la Sección 5.ª del Capítulo 3.° dijimos que el proceso $\{\xi(t),\ t \in T\}$ es continuo en m.c. si cumple:

$$\lim_{h\to 0} E\{/\xi(t+h)-\xi(t)/^2\}=0 \tag{1}$$

Desarrollando (1) por estacionariedad, tendremos que si para $h\to 0$ se verifica:

$$2B(0)-B(h)-B(-h)\to 0 \tag{2}$$

el proceso estocástico estacionario es continuo en $\forall t \in T$ si y solamente si $B(h)$ es continua en el origen.

2. *Derivada m.c.*

De forma semejante a como obtuvimos en la Sección 5.ª del Capítulo 3.°, el proceso estocástico estacionario $\{\xi(t),\ t \in T\}$ tendrá derivada m.c. si existe $\{\xi'(t),\ t \in T\}$ tal que:

$$\lim_{h\to 0} E\left\{\frac{/\xi(t+h)-\xi(t)}{h}-\xi'(t)/^2\right\}\to 0 \tag{3}$$

o su equivalente:

$$\lim_{h,\ k\to 0} E\left\{\frac{/\xi(t+h)-\xi(t)}{h}-\frac{\xi(t+h)-\xi(t)/^2}{k}\right\}=0 \tag{4}$$

La existencia del límite (3) implica por (8−) del Capítulo 3.°, Sección 5.ª

$$\lim_{n\to 0} E\frac{/\xi(t+h)-\xi(t)/^2}{h^2}=-B''(0) \tag{5}$$

y deberá estar acotada la:

$$\frac{\partial^2 B(t-s)}{\partial t\,\partial s}=-B''(h) \qquad h=t-s \tag{6}$$

Desarrollando (4) y recordando la (2) tenemos:

$$\frac{2B(0)-B(h)-B(-h)}{h^2} \tag{7}$$

La (7) es indeterminada y para deshacer la indeterminación aplicaremos dos veces L'Hôpital.

Si el proceso es real por ser $B(h)=B(-h)$ (8)

de (7), deducimos la condición necesaria para la existencia de derivada m.c.:

$$B'(0)=0 \tag{9}$$

pues caso contrario la (7) no sería finita cuando $h\to 0$.

3. *Derivadas sucesivas*

Para que existan las derivadas sucesivas m.c. de un proceso estocástico estacionario de órdenes m, y n, la función de covarianza $B_{\xi(m\,\xi(n}(t-s)$ aplicado a este caso la (10) de la Sección 5.ª

$$B_{\xi(m\,\xi\,n}(t-s)=E\,\xi^{(m}(t)\,\xi^{(n}(s)=$$

$$=\frac{\partial^{m+n}B(t-s)}{\partial t^n\,\partial s^n}=(-1)^n B^{(m+n)}(h) \tag{10}$$

siendo $h=t-s$.

4. *Desarrollos en serie de procesos*

Un proceso estocástico estacionario $\{\xi(t),\ t\in T\}$ decimos es desarrollable en serie m.c. cuando su función de covarianza $B(h)$ es analítica.

Decir que $B(h)$ es analítica quiere decir que todas sus derivadas existen por lo que $B(h)$ puede desarrollarse en serie Mac-Laurin:

$$B(h)=\sum_{n=0}^{\alpha}\frac{B^{(n}(0)}{\lfloor n}h^n \tag{11}$$

De forma semejante

$$B^{(m}(h)=\sum_{h=0}^{\alpha}B^{(m+n}(0)\frac{h^n}{\lfloor n} \tag{12}$$

La existencia $\xi^{(n}(t)$ en sentido m.c. viene impuesta por ser analítica

$B(h)$. Llamando $\hat{\xi}(t+h)$ a la aproximación del proceso $\xi(t)$ en $t+h$, tenemos:

$$\hat{\xi}(t+h)=\sum_{n=0}^{\infty}\xi^{(n}(t)\frac{h^n}{\underline{|n}} \qquad (13)$$

Papoulis (I) demuestra que en las hipótesis indicadas:

$$E/\xi(t+h)-\hat{\xi}(t+h)/^2=0 \qquad (14)$$

por lo que el proceso $\xi(t)$ puede aproximarse m.c. por un desarrollo de la forma expuesta en (13) y puede escribirse:

$$\xi(t+h)\overset{2}{=}\sum_{n=0}^{\infty}\xi^{(n}(t)\frac{h^n}{\underline{|n}} \qquad (15)$$

Demostremos primeramente la ortogonalidad entre la variable $\{\xi(t+h)-\hat{\xi}(t+h)\}$ y la variable $\xi^{(n}(t)$:

$$E\{[\xi(t+h)-\hat{\xi}(t+h)]\overline{\xi^{(n}(t)\}}=E\xi(t+h)\,\overline{\xi^{(n}(t)}-\sum_{m=0}^{\infty}E\frac{\xi^{(m}(t)h^m}{\underline{|m}}\overline{\xi^{(n}(t)}=$$

$$=(-1)^nB^{(n}(h)-\sum_{m=0}^{\infty}(-1)^n\frac{B^{(m+n)}}{\underline{|m}}(\text{o})h^m=(-1)^nB^{(n}(h)-(-1)^nB^{(n}(h)$$

$$=0 \qquad (6)$$

recordando (12).

Luego: $\xi(t+h)-\hat{\xi}(t+h)$ *será ortogonal a toda combinación lineal de* $\xi^{(n}(t)$ con coeficientes cualesquiera no aleatorios y en consecuencia a $\hat{\xi}(t+h)$:

$$E\{\xi(t+h)-\hat{\xi}(t+h)\}\overline{\hat{\xi}(t+h)\}}=0 \qquad (17)$$

Probemos que también es:

$$E\{\xi(t+h)-\hat{\xi}(t+h)\}\overline{\xi(t+h)}=B(0)-E\hat{\xi}(t+h)\overline{\xi(t+h)}=0 \qquad (18)$$

y habremos demostrado la (14) porque no es más que la diferencia entre la (18) y (17).

(I) Athanasios Papoulis: «Probability, Random Variables and Stochastic Processes». McGraw-Hill, 1965, págs. 318 y siguientes.

Recordando (13) y sustituyendo en (18) hallando esperanzas, tenemos:

$$B(0)=\sum_{n=0}^{\infty}\frac{h^n}{\lfloor n}E\,\xi^{(n}(t)\,\overline{\xi(t+h)}=$$

$$=B(0)-\sum_{n=0}^{\infty}\frac{(+h)^n}{\lfloor n}E^{(n}(t)\,\overline{\xi(t+h)}=$$

$$=B(0)-\sum_{n=0}^{\infty}\frac{(+h)^n}{\lfloor n}B^{(n}(-h)=0$$

de acuerdo con (10) y por ser:

$$B(0)=B(h-h)=\sum_{n=0}^{\infty}\frac{B^{(n}(-h)(+h)^n}{\lfloor n} \tag{19}$$

Luego si el proceso estacionario $\{\xi(t),\ t\in T\}$ tiene covarianza estacionaria analítica puede desarrollarse por Taylor en el punto t y la variable aleatoria $\xi(t+h)$ puede expresarse estocásticamente en función de $\xi(t)$ y de todas sus derivadas m.c. en t (I).

5. *Derivabilidad de procesos mutuos*

Recordando la Sección 3.ª, se puede determinar la función de covarianza mutua de los procesos $\{\xi(t),\ t\in T\}$ y en el de la derivada del proceso $\{\eta(s),\ s\in T\}$ de forma semejante a como estudiamos en el Capítulo 3.º, Sección 5.ª:

$$B_{\xi\eta'}(t,\ s)=E\,\xi(t)\,\overline{\eta'(s)}=\lim_{h\to 0}E\,\xi(t)\,\overline{\frac{\eta(t+h)-\eta(t)}{h}}=$$

$$=\frac{\partial B_{\xi,\eta}(t,\ s)}{\partial s} \tag{20}$$

e igualmente:

$$B_{\xi',\eta'}(t,\ s)=\frac{\partial^2 B_{\xi\eta}(t,\ s)}{\partial t\,\partial s} \tag{20'}$$

(I) Belyaev, Y. K., en su trabajo «Analistic Random Processes», págs. 81 a 88, de «Multiplicity Theory and Canonical Decomposition», publicada en Ed. Anthony Ephremides and John Tomas Dowden Hutchinson, 1973. Investiga los procesos analíticos..

En el caso de estacionariedad, tendremos:

$$B_{\xi' \eta'}(t-s) = \frac{\partial^2 B_{\xi,\eta}(t-s)}{\partial t\, \partial s} = -B''_{\xi,\eta}(h) \tag{21}$$

donde: $h = t - s$.

La covarianza mutua de los procesos derivados m.c.

$$\{\xi^{(n}(t), \quad t \in T\}, \ \{\eta^{(n}(s), \ s \in T\}$$

puede extenderse a órdenes superiores análogamente a lo ya expuesto en citada Sección 5.ª del Capítulo 3.º, para obtener la fórmula analítica de la función de covarianza mutua:

$$B_{\xi^{(n}\eta^{(m}}(t, \ s) = \frac{\partial^{n+m} B(t, \ s)}{\partial^n t\, \partial s^m} \tag{22}$$

y si son estacionarias:

$$B_{\xi^{(n}\eta^{(m}}(t-s) = (-1)^m B^{(n+m)}_{\xi,\eta}(h) \tag{23}$$

siendo $h = t - s$.

Sección 3.ª

Ergodicidad

1. *Introducción*

En esta Sección aplicaremos el concepto de ergodicidad para la determinación de las características poblacionales de la media, función de covarianza y funciones de distribución de los procesos estocásticos estacionarios.

Para no ser reiterativos supondremos al proceso de parámetro t continuo. Si fuese discreto las integrales se transformarían en sumas.

2. *Ergodicidad en media*

De una muestra del proceso estacionario $\{\xi(t), \quad t \in T\}$ una estimación de la media poblacional es la integral (no aleatoria):

$$\eta_T = \frac{1}{2T}\int_{-T}^{T} \xi(t,\ \omega)dt \qquad \omega \in \Omega \tag{1}$$

Si consideramos η_T un estimador, entonces (1) es una v.a. pues la función integrando es un proceso $\{\xi(t),\ -T \leqslant t \leqslant T\}$, y si tiene media constante $E\xi(t) = m$ (2)
el estimador (1) es centrado, ya que:

$$E\eta_T = \frac{1}{2T}\int_{-T}^{T} E\xi(t) \cdot dt = m \tag{3}$$

Es importante analizar el significado de las (1) y (3).

El valor de la media del proceso debe considerarse en el espacio de Ω aunque lo hayamos estimado en el espacio del tiempo (1).

La (3) nos indica que el estimador es centrado.

Para determinar si es o no consistente, emplearemos la varianza del estimador (1):

$$\sigma^2_{\eta T} = \frac{1}{4\ T^2} E\left\{ \Big/ \int_{-T}^{T} \Big/ \xi(t) - m/^2\, dt \right\} = \frac{1}{4\ T^2} \iint_{-T}^{T} B(t-s)\, dt\, ds \tag{4}$$

donde $B(t-s)$ es la covarianza de Kernel.

La (4) podemos reducirla a una integral. Por el cambio $u = t - s$ e integrando respecto a s tenemos:

$$\sigma^2_{\eta_T} = \frac{1}{2T}\int_{-2T}^{2T} B(u)\left(1 - \frac{/u/}{2T}\right)du \tag{4'}$$

Para que sea ergódico es preciso que la (4') tienda hacia cero cuando $T \to \infty$ y se cumplirá siempre que se verifique:

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-2T}^{2T} B(u)du = 0 \tag{5}$$

Una condición suficiente para la ergodicidad es que el límite de la función de covarianza de Kernel sea cero:

$$\lim_{h \to \pm\infty} B(h) = 0 \tag{6}$$

En el supuesto de que exista límite de la covarianza:

$$/\lim_{h\to\pm\infty} B(h)/=K \tag{7}$$

aplicando el criterio de sumas de Cèsareo podremos encontrar un $\varepsilon>0$ tal que

$$/B(h)-K/<\varepsilon\,/h/>N(\varepsilon)$$

Luego:

$$/\frac{1}{2T}\int_{-T}^{T}/B(h) \quad -K/dh\leqslant\frac{1}{2T}\int_{-N(\varepsilon)}^{N(\varepsilon)}/B(h)-K/dh+\varepsilon\left(1-\frac{N(\varepsilon)}{T}\right) \tag{8}$$

siendo $N(\varepsilon)$ finito, el primer sumando del segundo miembro tiende a cero cuando $T\to\infty$ y el primero tiende a ser menor que cualquier número, es decir que:

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} B(h)dh=K=\lim_{n\to\infty} B(h) \tag{9}$$

Luego ambos límites coinciden y en consecuencia si se cumple la (6) se cumplirá la (5) y será una condición suficiente para que (4') tienda a cero mostrando que el proceso estacionario $\{\xi(t),\quad t\in T\}$ es ergódico en media.

3. *Ergodicidad de la función de covarianza estacionaria*

Para la función de covarianza de un proceso estocástico estacionario indicaremos dos estimadores:

$$\hat{B}_T(h)=\frac{1}{2T}\int_{-T}^{T}\xi(t+h)\,\overline{\xi(t)}\,dt \tag{10}$$

siendo el otro:

$$\hat{B}_T(h)=\frac{1}{T}\int_{0}^{T-h}\xi(t+h)\,\overline{\xi(t)}\,dt \tag{10'}$$

dependiendo de la longitud T. El estimador (10') es asintóticamente centrado, ambos del mismo parámetro poblacional $B(h)$.

Para conocer la consistencia por la desigualdad de Tchebycheff tenemos:

$$\text{Prob.}\,\{/\hat{B}_T(h)-B(h)/<\varepsilon\}>1-\frac{D\hat{B}_T(h)}{\varepsilon^2} \tag{11}$$

y si la varianza del estimador $\hat{B}_T(h)$ tiende a cero, será consistente. La varianza del estimador (10) es:

$$\sigma^2_{B_T}=E\left\{\left/\frac{1}{2T}\int_{-T}^{T}\xi(t+h)\overline{\xi(t)}\,dt-B(h)\right/^2\right\} \tag{12}$$

Esta varianza precisa momentos de orden 4.° y es complicada excepto para procesos sencillos. Igualmente para la varianza de (10') son necesarios los momentos de orden cuarto.

4. *Ergodicidad de la función de distribución*

Si de un proceso estacionario $\{\xi(t),\quad t\in T\}$ pretendemos estimar la función de distribución, en la (1') de la Sección 1 vimos que la distribución de primer orden es independiente del tiempo; es decir:

$$F(x)=P(\,\xi(t)\leqslant x\,) \tag{13}$$

Si formamos el proceso dicotómico:

$$\eta(t)=\begin{matrix}1 & \text{si} & \xi(t)\leqslant x\\ 0 & \text{si} & \xi(t)>x\end{matrix} \tag{14}$$

los parámetros media y covarianza de $\eta(t)$ son:

$$E\eta(t)=1\cdot P\big(\xi(t)\leqslant x\big)=F(x) \tag{15}$$

$$B(t+h,\ t)=E\eta(t+h)\eta(t)=1\cdot P\big(\xi(t+h)\leqslant x,\ \xi(t)\leqslant x\big)=F(x,\ x;\ h) \tag{16}$$

siendo $F(x,\ x;\ h)$ la función de distribución de segundo orden dependiente de la diferencia de tiempos.

Si formamos el estimador del proceso $\{\eta(t),\quad t\in T\}$:

$$\eta_T(t)=\frac{1}{2T}\int_{-T}^{T}\eta(t)\,dt \tag{17}$$

se comprueba que en las condiciones de estacionariedad es centrado siendo la esperanza $F(x)$.

Para determinar la consistencia utilizaremos la varianza, que es:

$$\sigma^2_{\eta_T} = E\{/\eta_T(t) - F(x)/^2\} =$$

$$= \frac{1}{2T}\int_{-2T}^{2T}\left(1 - \frac{/h/}{2T}\right)[F(x,\ x,\ h) - F(x)^2]\,dh \qquad (18)$$

y para que sea consistente tiene que cumplir:

$$\lim_{h\to\alpha} F(x,\ x,\ h) = F(x)^2 \qquad (19)$$

indicativa que las variables $\xi(t)$ y $\xi(t+h)$ tienden a la independencia cuanto mayor sea la diferencia de tiempos.

Sección 4.ª

Funciones características y Funciones definidas no negativas

1. *Introducción*

Las funciones características definidas en los libros de Estadística están relacionadas con las funciones definidas no negativas, continuas y acotadas.

Existen también funciones definidas no negativas sobre conjuntos discretos que carecen de la propiedad de ser continuas para los valores pertenecientes a referidos conjuntos.

El concepto usual de función característica está definida sobre conjuntos reales. Nada impide que estén definidas estas funciones sobre conjuntos discretos

$$D_T = \{\ \ldots -2d,\ -d,\ 0,\ +d,\ +2d\ \ \ldots\} \qquad (1)$$

y que denominaremos funciones características impropias o también función característica definida sobre D_T (para $t \in D_T$).

Existe una correspondencia entre las funciones definidas no negativas y acotadas y las funciones características (en el mismo dominio de definición).

La unicidad de esta correspondencia nos la proporcionan dos teoremas importantes: el de Herglotz basado sobre conjuntos (1) y el de Bochner-Kintchine, definidos sobre R por lo que estudiaremos ambos tipos de funciones definidas no negativas y las funciones características sobre los mismos conjuntos.

Seguiremos fundamentalmente a los autores Lèvy (I), Loève (II) y Cramer (III) dando unas ideas generales para estudiar los dos casos concretos indicados: cuando los conjuntos son de tipo D_T y el eje real.

2. *Funciones características*

Llamaremos clase $\mathscr{C}$ de funciones características sobre $\mathscr{R}$ a las funciones de la variable real t definida sobre un conjunto T y que puede expresarse en la forma:

$$\varphi(t)=c\varphi_0(t) \qquad c>0 \qquad t\in T \tag{2}$$

definiendo el conjunto T de forma que contenga el elemento 0; y si contiene cualquier número real t también contiene su opuesto $-t\in T$.

Y tal que $\varphi_0(t)-y$ en consecuencia φ sea representable en la forma:

$$\varphi_0(t)=\int_S e^{itx}dF(x) \qquad x\in S \tag{3}$$

donde $F(x)$ es monótona no decreciente y acotada y $S\subset R$.

Los casos particulares importantes de (3) son dos, para demostrar los dos teoremas que hemos indicado:

1.° Cuando $t\in D_T$ (1) y

$$x\in S=\left\{x\Big/\frac{-\pi}{d}<x\leqslant+\frac{\pi}{d}\right\} \tag{4}$$

en cuyo caso la (3) la escribiremos:

$$\varphi(t)=\int_{\frac{-\pi}{d}}^{\frac{\pi}{d}} e^{itx}dF(x) \qquad t\in D_T \tag{3'}$$

2.° Si $t\in R$ y $x\in R$ entonces la (3) es la tan conocida y estudiada en Estadística, pero recordemos la (2) por lo que diferirá en una constante multiplicativa:

(I) P. Lèvy: «Processus Stochastiques», Gauthier-Villars, pág. 103 y sig. 1965, 2.ª ed.
(II) M. Loève: «Probability Theory». Van Nostrand, 1962, 3.ª ed., págs. 207 y siguientes.

(III) H. Cramer: «Stationary and Related Stochastic Processes», Wiley, 1968, págs. 168 y siguientes.

$$\varphi(t)=\int_{-\infty}^{\infty} e^{itx}\, dF(x) \qquad t \in R \tag{3"}$$

En ambos casos $F(x)$ es una función monótona no decreciente y acotada.

3. *Funciones definidas no negativas*

Definiremos función definida no negativa en T a la función $\varphi(t)$ acotada en $t \in T$ tal que para:

$$\forall\, t_1\ t_2\ \ldots\ t_n \in T$$

$$\forall\, \xi_1\ \xi_2\ \ldots\ \xi_n \in C \qquad (C \text{ conjunto de números complejos}) \tag{5}$$

se cumple:

$$\sum_{k=1}^{n}\sum_{j=1}^{n} \varphi(t_k-t_j)\ \xi_k\, \xi_j \geqslant 0 \tag{6}$$

$$\forall\, t_k,\ t_j,\ t_k-t_j \in T$$

Siempre $0 \in T$. Los casos interesantes en nuestro estudio son los ya indicados. Si $T=D_T(1)$ o si $T=R$.

4. *Propiedades generales de las funciones características definidas sobre T.*

1.ª Toda función característica 2 definida sobre T es una función definida no negativa sobre T y acotada.

Cumple la condición de ser definida no negativa: En efecto. Sustituyendo (3) en el primer miembro de (6):

$$\sum_{k=1}^{n}\sum_{j=1}^{n} \varphi(t_k-t_j)\ \xi_k\, \overline{\xi}_j = \int_S \Big/ \sum_{k=1}^{n} e^{it_k x} \Big/^2 dF(x) \tag{7}$$

y por ser el integrando no negativo igualmente lo será la integral y se cumplirá (6) para todas las funciones características definidas en 2.

2.ª acotación:

La (3) está acotada al estarlo $F(x)$:

$$/\varphi(t)/ \leqslant \int_S /e^{itx}/ dF(x) = \int_S dF(x) < c > 0 \tag{8}$$

Luego es una función definida no negativa por cumplir las condiciones impuestas en la definición, 3, para esta clase de funciones.

5. *Propiedades generales de las funciones definidas no negativas sobre T*

1.ª Propiedad:

$$\varphi(0)\geqslant 0 \tag{9}$$

Se deduce inmediatamente haciendo en (6): $n=1$, $t=0$, $\xi=1$

2.ª Propiedad: La función $\varphi(t)$ es hermítica (I); es decir:

$$\varphi(-t)=\overline{\varphi(t)} \tag{10}$$

La demostración es simple porque haciendo en (6):

$$n=2 \quad t_1=0 \quad t_2=t \qquad \{0,\ t,\ -t\}\in T$$

y para $\forall \xi_1\ \xi_2 \in C$, tenemos:

$$0\leqslant\varphi(0)(/\xi_1/^2+/\xi_2/^2)+[\varphi(-t)\xi_1\ \overline{\xi}_2+\varphi(t)\overline{\xi}_1\ \xi_2] \tag{11}$$

Por ser el primer sumando real (1.ª propiedad) el paréntesis necesariamente debe ser real.

Si hacemos:

$$\begin{aligned}\varphi(t)\ \ &=\alpha_1+i\beta_1\\ \varphi(-t)&=\alpha_2+i\beta_2\end{aligned} \tag{12}$$

y

$$\xi_1\ \overline{\xi}_2=\gamma+i\delta\Rightarrow\overline{\xi}_1\ \xi_2=\gamma-i\delta \tag{13}$$

Si son arbitrarios $\xi_1\ \xi_2$ lo serán igualmente γ y δ. Sustituyendo (12) y (13) en (11) tendremos un número complejo y para que sea real, la parte imaginaria:

$$i\gamma(\beta_1+\beta_2)+i\delta(\alpha_2-\alpha_1)$$

necesariamente debe ser idénticamente nula para $\forall\gamma,\ \delta\in \mathrm{R}\Rightarrow$

$$\alpha_1=\alpha_2 \quad \beta_1=-\beta_2\Rightarrow\varphi(-t)=\overline{\varphi}(t)$$

C.h.d.

(I) P. Lévy señala esta condición innecesaria en su definición por cuanto es demostrable, o. c., pág. 103.

3.ª Propiedad:

$$/\varphi(t)/\leqslant\varphi(0) \tag{14}$$

Formemos el determinante no negativo en $\{0, t\}\in T$

$$0\leqslant\begin{vmatrix}\varphi(0) & \varphi(t)\\ \varphi(-t) & \varphi(0)\end{vmatrix}=\varphi(0)^2-\varphi(t)/^2\geqslant 0\Rightarrow/\varphi(t)/\leqslant\varphi(0)$$

Estas y otras muchas propiedades gozan las funciones definidas no negativas sobre los conjuntos particulares $D_T(1)$ y R.

6. *Continuidad de la función definida no negativa sobre R.*

Si $\varphi(t)$ es una función definida no negativa sobre R continua en el origen, lo es para todo $t\in R$.

Formemos el determinante de la forma haciendo:

$$t_1=0,\ t_2=t \text{ y } t_3=t'$$

y que será no negativo:

$$0\leqslant\begin{vmatrix}\varphi(0) & \varphi(t) & \varphi(t')\\ \overline{\varphi(t)} & \varphi(0) & \varphi(t'-t)\\ \overline{\varphi(t')} & \overline{\varphi(t'-t)} & \varphi(0)\end{vmatrix}=\varphi(0)^3+\varphi(t)\overline{\varphi(t')}\varphi(t'-t)+$$

$$+\varphi(t')\overline{\varphi(t)}\,\overline{\varphi(t'-t)}-/\varphi(t'-t)/^2\varphi(0)-/\varphi(t')/^2\varphi(0)-/\varphi(t)/^2\varphi(0)$$

Pasando al primer miembro los dos últimos términos y restando después a ambos miembros de la desigualdad la expresión:

$$\varphi(0)\left[\varphi(t)\overline{\varphi(t')}+\overline{\varphi(t)}\varphi(t')\right]$$

tendremos:

$$\varphi(0)/\varphi(t')-\varphi(t)/^2\leqslant\varphi(0)^3-\varphi(0)//\varphi(t'-t)/^2-$$

$$-\varphi(t)\overline{\varphi(t')}\left[\varphi(0)-\varphi(t'-t)\right]-\overline{\varphi(t)}\varphi(t')\left[\varphi(0)-\overline{\varphi(t'-t)}\right] \tag{15}$$

Por la continuidad en el origen de $\varphi(t)$, $\varphi(t'-t)$ es continua si $t'\to t$ y en consecuencia el segundo miembro de la (15) se anula, lo que implica que:

$$\lim_{t'\to t}\varphi(t')=\varphi(t)\qquad\forall t\in R \tag{16}$$

c.h.d.

7. *Propiedad de las funciones características sobre R.*

En cualquier tratado de Estadística pueden consultarse pero especialmente en Luckas (I).

Por haber demostrado en (4) que las funciones características definidas sobre T eran funciones definidas no negativas y acotadas, tienen todas las propiedades indicadas en (5) para estas funciones, también cuando $T=R$.

Demostraremos solamente que cuando $T=R$ esta clase de funciones características goza de la continuidad en todo $t \in R$.

De (3") deducimos:

$$/\varphi(t+h)-\varphi(t)/\leqslant \int /e^{ith}-1/\ dF(x)=2\int_{-\alpha}^{\alpha}\left/\operatorname{sen}\frac{hx}{2}\right/ dF(x) \qquad (17)$$

Si elegimos A de forma que:

$$\int_{/x/>A} dF(x)\leqslant\varepsilon$$

La (17) la podemos escribir:

$$/\varphi(t+h)-\varphi(t)/=2\int_{-A}^{A}\left/\operatorname{sen}\frac{hx}{2}\right/ dF(x)+2\int_{/x/>A} dF(x)$$

Para $h\rightarrow 0$, fijo A, el primer sumando siempre puede hacerse menor que cualquier número, por lo que:

$$/\varphi(t+h)-\varphi(t)/<\varepsilon \qquad \text{c.h.d.} \qquad (18)$$

En síntesis: Toda función característica sobre R es una función continua y acotada, definida no negativa (1) incluyéndose en el segundo caso del Epígrafe 2.

8. *Teorema de Herglotz*

«La clase de funciones características coincide con la clase de funciones **definidas no negativas**, ambas definidas sobre D_T» (II).

1. El teorema directo que toda función característica era una función definida y acotada se ha demostrado con las propiedades generales estudiadas de las funciones características en el parágrafo 4.

(I) Eugène Luckas: «Characteristic Functions», Hafner, New York, 1970, 2.ª ed.

(II) Véase párrafo 1.º (I).

2. Demostraremos el recíproco: «Toda función definida no negativa y acotada sobre D_T, es una función característica» y en consecuencia representable por:

$$\varphi(t)=\int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{itx}dF_n(x) \qquad t \in D_T=\{0, \ \pm d, \ \pm 2d \ldots\} \tag{19}$$

siendo $F(x)$ una función monótona no decreciente y acotada.

Hagamos:

$$t=jd \qquad s=kd, \qquad \forall t, \ s \in D_T \tag{20}$$

siendo j y k números enteros y elijamos:

$$\xi_j=e^{-idjx}\sqrt{\frac{dx}{2\pi n}} \qquad \xi_k=e^{-idkx}\sqrt{\frac{dx}{2\pi n}} \tag{21}$$

Sustituyendo estos valores en la (6), siendo $T=D_T$ y las sumatorias será:

$$dF_n(x)=\frac{1}{2\pi n}\sum_{j=1}^{n}\sum_{k=1}^{n}\varphi[(j-k)d]e^{-ix(j-k)d}\,dx \tag{22}$$

gozando de la propiedad de ser no negativa. Haciendo el cambio:

$$j-k=h \in Z$$

y sumando respecto a k teniendo en cuenta que para un valor de $|h|$ hay $n-|h|$ sumandos iguales, y siendo la variación de h: $-n+1 \leqslant h \leqslant n-1$, tendremos:

$$dF_n(x)=\frac{1}{2\pi}\sum_{-n+1}^{n-1}\left(1-\frac{/h/}{n}\right)\varphi(hd)\,e^{-ixhd}\,dx \tag{22'}$$

multiplicando la (22') por e^{ixdr} ($dr \in D_T$) e integrnado respecto a x entre

$$\left(-\frac{\pi}{d}, \ +\frac{\pi}{d}\right):$$

$$\int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{ixdr}dF_n(x)=\frac{1}{2\pi}\sum_{-n+1}^{n-1}\left(1-\frac{/h/}{n}\right)\varphi(hd)\int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{ixd(r-k)}dx=$$

$$=\left(1-\frac{/r/}{n}\right)\varphi(rd)$$

porque todas las integrales cuando $r-k\neq 0$ son nulas y únicamente no se anula cuando $r=k$.

Luego:

$$\int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{ixdr}dF_n(x)=\left(1-\frac{/r/}{n}\right)\varphi(rd) \tag{23}$$

El primer miembro es una función característica sobre D_T y el segundo una función definida no negativa sobre D_T. Estas funciones características en D_T convergen monótonamente hacia una función definida positiva:

$$\lim_{n\to\infty}\left(1-\frac{/r/}{n}\right)\varphi(rd)=\varphi(rd) \tag{24}$$

por lo que en el límite y poniendo $rd=t$, tenemos: (25)

$$\varphi(t)=\int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{ixt}dF(x) \qquad t\in D_T \tag{26}$$

completando la demostración del teorema.

Cuando $D_T=Z$ (conjunto números enteros) la (26) se convierte:

$$\varphi(t)=\int_{-\pi}^{\pi} e^{ixt}dF(x) \qquad t\in Z \tag{27}$$

que utilizaremos más frecuentemente que la fórmula anterior. c.h.d.

9. *Teorema de Bochner*

«La clase $\mathscr{C}$ de funciones características coincide con la clase de funciones definidas no negativas continuas y acotadas.»

En este teorema ambas funciones están definidas sobre R.

I. Directo: Ha sido demostrado en (4), pues probamos que toda función característica era una función definida no negativa de acuerdo con la definición dada en (3).

II. Recíproco: «Toda función definida no negativa, continua y acotada, es una función característica».

En consecuencia estas funciones definidas no negativas, continuas y acotadas en R pueden representarse por (3").

1. Formación de integral Riemann:

Elijamos los conjuntos $\forall t_j$ y $\forall \xi_k$ de la siguiente forma:

$$u=t_k \qquad \xi_k=\frac{e^{-ixu}}{\sqrt{2\pi T}}du\sqrt{dx}$$

$$v=t_j \qquad \xi_j=\frac{e^{-ixv}}{\sqrt{2\pi T}}dv\sqrt{dx} \tag{28}$$

sustituyendo estos valores en la (6) y las sumas límites se transforman en una integral de Riemann:

$$dF_T(x)=\frac{1}{2\pi T}\int\int_0^T \varphi(u-v)\,e^{-ix(u-v)}du\,dv\cdot dx \tag{29}$$

siendo (29) no negativa.

2. Integrabilidad en el eje R de $dF_T(x)$. Por el cambio:

$$t=u-v$$

e integrando respecto a v, tenemos:

$$dF_T(x)=\frac{dx}{2\pi}\int_{-T}^{T}\varphi(x)\left(1-\frac{/t/}{T}\right)e^{-ixt}\;dt \tag{30}$$

Examinemos si $dF_T(x)$ es integrable en R.

Introduzcamos la función:

$$\Gamma\left(\frac{t}{T}\right)=\begin{matrix}1-\dfrac{/t/}{T} & \qquad /t/\leqslant T\\ 0 & \qquad /t/>T\end{matrix} \tag{31}$$

Con esta función auxiliar podemos escribir la (30):

$$dF_T(x)=\frac{dx}{2\pi}\int_{-\infty}^{\infty}\varphi(t)\Gamma\left(\frac{t}{T}\right)e^{-itx}dt \tag{32}$$

Los límites de la integral son aparentemente formales y si multiplicamos la (32) por $\Gamma\left(\dfrac{x}{2M}\right)$ e integrando respecto a x (desde $-\infty$ a $+\infty$), el orden de integración puede invertirse por la circunstancia de la definición de $\Gamma(\,\cdot\,)$ y ser los límites finitos:

$$\int_{-\infty}^{\infty} \Gamma\left(\frac{x}{2M}\right) dF(x) = \int_{-\infty}^{\infty} \varphi(t)\,\Gamma\left(\frac{t}{T}\right) dt \int_{-\infty}^{\infty} \Gamma\left(\frac{x}{2M}\right) e^{-itx} dx = \tag{33}$$
$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \varphi(t)\,\Gamma\left(\frac{t}{T}\right) \frac{\operatorname{sen}^2 Mt}{Mt^2} dt$$

La función producto $\varphi(t)\ \Gamma\left(\frac{t}{T}\right)$ es continua y acotada al serlo cada una de ellas; y además, $\Gamma\left(\frac{t}{T}\right)$ es una función característica (I). La (33) está acotada y es integrable sobre todo intervalo finito (II).

El primer miembro de (33) es no negativo pues no lo son ni $dF_T(x)$ ni $\Gamma\left(\frac{x}{2M}\right)$. La suma está acotada:

$$\int_{-\infty}^{\infty} \Gamma\left(\frac{x}{2M}\right) dF_T(x) \leqslant \frac{\varphi(0)}{\pi} \int_{-\infty}^{\infty} \frac{\operatorname{sen}^2 t}{t^2} dt = \varphi(0) \tag{34}$$

La (34) se cumple para $M>0$ y el integrando del primer miembro crece monótonamente a $dF_T(x)$ cuando $M \to \infty$ y, en el límite:

$$\int_{-\infty}^{\infty} dF_T(x) \leqslant \varphi(0) \tag{35}$$

que nos indica que tanto $dF_T(x)$ como $\Gamma\left(\frac{t}{T}\right)\varphi(t) = \varphi_T(t)$ (36)

son integrables sobre R.

3. Formaciones de sucesiones características:

Si multiplicamos (32) por:

$$\Gamma\left(\frac{x}{2M}\right) e^{iux} \tag{37}$$

e integrando con respecto a x en R, tenemos:

$$\int_{-\infty}^{\infty} \Gamma\left(\frac{x}{2M}\right) e^{iux} dF_T(x) =$$

(I) Kai Lai Chung: «A Cours in Probability Theory», pág. 157. Academic Press, New York, 2.ª ed. 1974.

(II) Tichmarsch: «Theory of Fourier Integral Teorema 21», pág. 37. Oxford U. Press, 1967.

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi_T(t)dt\int_{-\infty}^{\infty}\Gamma\left(\frac{x}{2M}\right)e^{ix(t-u)}dx \tag{38}$$

recordando (31) y (36).

El orden de integración ha podido realizarse por las mismas razones expuestas anteriormente.

Realizando la integración, tenemos:

$$\int_{-\infty}^{\infty}\Gamma\left(\frac{x}{2M}\right)e^{iux}dF_T(x)=\frac{1}{\pi}\int_{-\infty}^{\infty}\varphi_T(t)\frac{\operatorname{sen}^2 M(t-u)}{M(t-u)^2}dt \tag{38'}$$

El primer miembro de la (38') es una función característica, pues es la integral de una función no negativa por e^{iux}. El segundo miembro converge a $\varphi_T(t)$ cuando $M\to\infty$ como puede verse en Titchmarsh (I).

Luego $\varphi_T(t)$ es el límite de una sucesión de funciones características.

Y también de (36) recordando (31):

$$\varphi_T(t)\to\varphi(t) \qquad T\to\infty \tag{39}$$

que nos demuestra el teorema, siendo $\varphi(t)$ una función característica.

Luego por criterios de convergencia de función de distribución (II) y el criterio de convergencia de integrales de Helly-Bray (III):

$$\lim_{T\to\infty}\int_{-\infty}^{\infty}e^{iux}dF_T(x)=\int_{-\infty}^{\infty}e^{iux}dF(x)=\varphi(u) \quad \text{c.h.d.} \tag{40}$$

Sección 5.ª

Representaciones espectrales de la función de covarianza estacionaria

1. *Representación espectral de la función de covarianza estacionaria sobre D_T*

1.1. Teorema: Toda función de covarianza $B(t)$ correspondiente a un proceso estocástico estacionario $\{\xi(t), \quad t\in D_T\}$ puede representarse en la denominada forma espectral:

(I) E. C. Tichmars: «Theory of Fourier Integral», 1967, págs. 36 y 37.

(II) Gonzalo Arnaiz Vellando: «Introducción a la estadística». Lex Nova, 1965, página 285, ref. 1.ª

(III) M. Loéve: «Teoría de la probabilidad». Tecnos. Traduc. Española, 1976, págs. 179 y siguientes.

$$B(t)=\int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{it\lambda}dF(\lambda) \qquad t\in D_T=\{0,\ \pm d,\ \pm\cdots\pm nd...\} \tag{1}$$

donde $F(\lambda)$ es una función monótona no decreciente y acotada tal que:

$$F\left(-\frac{\pi}{d}\right)=0$$

A $F(\lambda)$ se denomina función de distribución espectral asociada al proceso estacionario $\{\xi(t),\quad t\in D_T\}$.

1.2. Demostración: Si demostramos que $B(t)$ es semidefinida y acotada en D_T cumplirá las condiciones del teorema de Herglotz y será representable por (1). Que es no negativa se deduce:

$$\sum_j\sum_k B(t_j-t_k)\xi_j\,\overline{\xi}_k= \qquad \forall t_j,\ t_k\in D_T$$

$$\sum_j\sum_k E\xi(t_j)\overline{\xi(t_k)}\eta_j\overline{\eta}_k=E\left|\sum_{j=1}^{n}\xi(t_j)\eta_j\right|^2\geqslant 0 \qquad \begin{matrix}\forall t_j\in D_T\\ \forall\eta_j\in C\end{matrix} \tag{2}$$

Y que es acotada se cumplirá porque el proceso $\xi(t)$ es de segundo orden; es decir:

$$E\{/\xi(t)/^2\}<\infty \tag{3}$$

Luego $B(t)$ es una función definida, no negativa, en D_T correspondiente a un proceso de segundo orden, y de acuerdo con el Teorema de Herglotz es representable por (1). c.h.d.

1.3. El momento de segundo orden podemos calcularlo por la función de distribución espectral.

En efecto. Haciendo en (1) $t=0$, tenemos:

$$B(0)=E\{/\xi(t)/^2\}=F\left(\frac{\pi}{d}\right)-F\left(\frac{-\pi}{d}\right) \tag{4}$$

y por ser $dF(\lambda)\geqslant 0\Rightarrow F(\lambda)$, que, salvo una constante se elige de forma $F(\lambda)$ sea siempre positiva o nula:

$$F\left(\frac{-\pi}{d}\right)=0 \tag{5}$$

1.4. Si la esperanza del proceso fuese nula, la función de covarianza sería de Kernel y para $t=0$ nos dará la varianza y es el valor máximo de la función de distribución espectral:

$$\sigma^2 \xi(t) = E\{/\xi(t)/^2\} - [E\xi(t)]^2 = B(0) = F\left(\frac{\pi}{d}\right) \tag{6}$$

1.5. Si el conjunto D_T fuere el de los números enteros, la representación espectral de la función de covarianza es:

$$B(n) = \int_{-\pi}^{\pi} e^{in\lambda} dF(\lambda) \qquad n \in \{0, \ \pm 1, \ \pm 2 \ ...\} \tag{7}$$
$$\pi < \lambda \leqslant \pi$$

por ser $d=1$.

A la función $F(\lambda)$ se la denomina función de distribución espectral asociada al proceso estacionario $\{\xi_n, \ \ n \in Z\}$ y es monótona, no decreciente y acotada. Por ser un caso particular, las (4) y (5), se convierten:

$$B(0) = E/\xi(t)/^2\} = F(\pi) \tag{8}$$

Las representaciones precedentes se han considerado para cualquier proceso estocástico estacionario (real o complejo).

1.6. No obstante, cuando el proceso estocástico estacionario fuese real, la sucesión de funciones de covarianzas $\{B(n), \ n \in Z\}$ es también real y su representación es:

$$B(n) = \int_{-\pi}^{\pi} \cos \lambda n dF(\lambda) \qquad n = 0, \ \pm 1, \ \pm 2 \ ... \tag{9}$$

En efecto. Por ser:

$$B(-n) = B(n) \tag{10}$$

y recordando $e^{xi} = \cos x + i \operatorname{sen} x$ sustituyendo en (7) tenemos para los valores (10):

$$B(n) = \int_{-\pi}^{\pi} [\cos \ \lambda n + i \ \operatorname{sen} \ \lambda n] \ dF(\lambda) \tag{10'}$$

$$B(-n) = \int_{-\pi}^{\pi} [\cos \lambda n - i \operatorname{sen} \lambda n] dF(\lambda) \tag{10''}$$

por ser par la función coseno e impar la función seno. Sumando ambas y recordando (10), obtenemos la (9).

c.h.d.

1.7. La representación espectral de la función de covarianza estacionaria en D_T correspondiente a un proceso real es:

$$B(t) = \int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} \cos t\lambda\, dF(\lambda) \qquad t \in D_T = \{0, \pm d, \pm 2d \ldots\} \tag{11}$$

La demostración es idéntica a la precedente, ya que se cumple la (10) y excepto en los límites, las (10') y (10'').

2. *Representación espectral de la función de covarianza estacionaria* (parámetro continuo)

2.1. Toda función de covarianza continua en el origen correspondiente a un proceso estocástico estacionario de segundo orden de parámetro t continuo $\{\xi(t),\ t \in R\}$ puede representarse en la denominada forma espectral de Kintchine:

$$B(t) = \int_{-\infty}^{\infty} e^{it\lambda}\, dF(\lambda) \qquad \begin{array}{l} t \in R \\ \lambda \in R \end{array} \tag{12}$$

2.2. A la función $F(\lambda)$ se la denomina función de distribución espectral asociada al proceso $\{\xi(t),\ t \in R\}$ y es una función monótona, no decreciente y acotada, tal que $F(-\infty) = 0$.

2.3. La demostración de ser definida no negativa y acotada es semejante a la (2) y (3) y la omitimos.

2.4. Si la función de covarianza es continua m.c. tal como indicamos en la Sección 2.1 (y se ha admitido como hipótesis) es representable en la forma espectral (12).

2.5. Igualmente a como vimos en el caso de $t \in R$, para $t = 0$ tendremos:

$$B(0) = E\{/\xi(t)/^2\} = F(\infty) - F(-\infty) \tag{13}$$

eligiendo —salvo constante— $F(\lambda)$ tal que sea siempre positiva y que para el extremo inferior

$$F(-\infty) = 0 \tag{14}$$

según hemos indicado en 2.2.

2.6. Si el proceso estocástico estacionario es real por las razones expuestas en 1.6. para las sucesiones de covarianza, tendremos para las covarianzas estacionarias $\{B(t), \quad t \in R\}$ la representación espectral:

$$B(t) = \int_{-\infty}^{\infty} \cos t\lambda \, dF(\lambda) \qquad \begin{matrix} t \in R \\ \lambda \in R \end{matrix} \tag{15}$$

Sección 6.ª

Funciones de Distribución y de Densidad Espectrales

I. *Introducción*

1. En la Sección anterior hemos representado la función de covarianza estacionaria espectralmente asociando una función $F(\lambda)$ que denominamos función de distribución espectral asociada a un proceso.

Cuando $B(t)$ estaba definida sobre D_T, la función de distribución espectral variaba sobre el intervalo $\lambda \in \left(\frac{-\pi}{d}, \frac{\pi}{d} \right]$.

2. Es de suma importancia cuando $D_T = Z$ en cuyo caso el intervalo de variación de λ es de $(-\pi, \ +\pi]$.

3. Para no ser reiterativos estudiaremos sólo los casos de ser $D_T = Z$ y el de parámetro continuo, de la función de covarianza.

II. *Función de distribución espectral de un proceso estacionario* $\{B(n), \quad n \in Z\}$ *si* $B(-n) = \overline{B(n)}$

Teorema 1.º

1. La representación espectral de la covarianza estacionaria para $\forall n \in Z$ de un proceso estocástico estacionario complejo, según vimos en 1.5 es:

$$B(n) = \int_{-\pi}^{\pi} e^{i\lambda n} dF(\lambda) \qquad -\pi < \lambda \leqslant \pi \tag{1}$$

siendo $F(\lambda)$ una función monótona no decreciente y acotada denominada función de distribución espectral y tal que $F(-\pi) = 0$.

Si $B(n)$ es representable por (1) la función de distribución espectral normalizada es:

$$\frac{F(\lambda_2+0)+F(\lambda_2-0)}{2}-\frac{F(\lambda_1+0)+F(\lambda_1-0)}{2}=$$

$$=\frac{1}{2\pi}\lim_{N\to\infty}\sum_{n=-N}^{+N}\frac{e^{-i\lambda_2 n}-e^{-i\lambda_1 n}}{-in}B(n) \tag{2}$$

donde el término de la sumatoria para $n=0$ vale:

$$\frac{B(0)(\lambda_2-\lambda_1)}{2\pi} \tag{3}$$

2. *Demostración*

2.1. Fundamentalmente seguiremos a Doob (I) aunque modificaremos la terminología por unidad de criterio, siguiendo a los principales tratadistas.

2.2. Formemos la función:

$$\beta(\lambda)=\begin{cases} 1 & \lambda_1<\lambda<\lambda_2 \qquad (\lambda_1,\ \lambda_2)\in(-\pi,\ \pi) \\ \frac{1}{2} & \lambda=\lambda_1 \qquad \lambda=\lambda_2 \\ 0 & \text{en otro caso} \end{cases} \tag{4}$$

y desarrollemos $\beta(\lambda)$ en serie de Fourier de términos complejos, que coincidirá con $\beta(\lambda)$ en todo el intervalo $(-\pi,\ +\pi)$ y en los puntos de discontinuidad de $\beta(\cdot)$ por cuanto, según (4), se han elegido de forma que

$$\beta(\lambda)=\frac{\beta(\lambda+0)+\beta(\lambda-0)}{2} \tag{5}$$

2.3. El desarrollo es:

$$\beta(\lambda)=\lim_{N\to\infty}\sum_{n=-N}^{+N}b_n e^{i\lambda n} \qquad n\in Z \tag{6}$$

(I) J. Doob: o. c. pág. 474 y sig. El integrando de la (1) en Doob es:

$$e^{i2\pi\lambda t}dF(\lambda)$$

y modifica esencialmente las expresiones y límites de la (1) y (2).

siendo b_n el coeficiente de Fourier determinado por:

$$b_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \beta(\lambda)\, e^{i\lambda n}\, d\lambda \tag{7}$$

Sustituyendo los valores de (4) en (7) tendremos para los coeficientes b_n:

$$b_n = \frac{1}{2\pi} \cdot \frac{e^{-i\lambda_2 n} - e^{-i\lambda_1 n}}{-itn}; \ y \ b_0 = \frac{(\lambda_2 - \lambda_1)}{2\pi} \tag{8}$$

2.4. Si analizamos el segundo miembro de (2) los coeficientes de $B(n)$ son, precisamente, los que acabamos de obtener en (8).

2.5. Formemos la suma finita con (7):

$$\sum_{n=-N}^{+N} b_n B(n) \tag{9}$$

y sustituimos $B(n)$ por su valor (1). La (9) la escribiremos:

$$\sum_{n=-N}^{N} b_n B(n) = \sum_{n=-N}^{+N} b_n \int_{-\pi}^{\pi} e^{i\lambda n} dF(\lambda) =$$

$$\sim \int_{-\pi}^{\pi} \left[\sum_{n=-N}^{N} b_n e^{i\lambda n} \right] dF(\lambda) \tag{10}$$

El paréntesis del integrando tiende a $\beta(\lambda)$ según (6). Luego:

$$\sum_{n=-N}^{N} b_n B(n) = \int_{-\pi}^{\pi} \left[\sum_{n=-N}^{N} b_n e^{i\lambda n} \right] dF(\lambda) \rightarrow \int_{-\pi}^{\pi} \beta(\lambda) dF(\lambda) \tag{11}$$

La última integral es el primer miembro de la (2) como se prueba descomponiendo la integral en las tres integrales no nulas, recordando los valores de $\beta(\lambda)$ de (4):

$$\int_{-\pi}^{\pi} \beta(\lambda)\, dF(\lambda) = \int_{\lambda_1-0}^{\lambda_1+0} \beta(\lambda)\, dF(\lambda) + \int_{\lambda_1+0}^{\lambda_2-0} \beta(\lambda)\, dF(\lambda) +$$

$$+ \int_{\lambda_2-0}^{\lambda_2+0} \beta(\lambda)\, dF(\lambda) = \frac{1}{2}[F(\lambda_1+0) - F(\lambda_1-0)] + [F(\lambda_2-0) - F(\lambda_1+0)] +$$

$$+ \frac{1}{2}[F(\lambda_2+0) - F(\lambda_2-0)] = \frac{F(\lambda_2+0) + F(\lambda_2-0)}{2} - \frac{F(\lambda_1+0) + F(\lambda_1-0)}{2} \tag{12}$$

Sustituyendo los coeficientes b_n obtenidos en (8) en el primer miembro de la (11) y (12) en el segundo, tenemos en el límite la (2). c.h.d.

Teorema 2.°

1. *Si la representación espectral de la covarianza estacionaria* $\forall n \in Z$ de un proceso real $\{\xi(n),\ n \in Z\}$ según (9) de la Sección anterior, es:

$$B(n)=\int_{-\pi}^{\pi} \cos n\lambda\, dF(\lambda) \tag{13}$$

la función de distribución espectral es:

$$\frac{F(\lambda_2+0)+F(\lambda_2-0)}{2}-\frac{F(\lambda_1+0)+F(\lambda_1-0)}{2}=$$

$$=B(0)\frac{(\lambda_2-\lambda_1)}{2\pi}+\frac{1}{\pi}\lim_{N\to\infty}\sum_{n=1}^{N}\frac{\operatorname{sen} n\lambda_2-\operatorname{sen} n\lambda_1}{n}B(n) \tag{14}$$

2. *Demostración*

En este caso particular, por ser el proceso real, $B(-n)=B(n)$ de la (2) deducimos:

$$F(\lambda_2)-F(\lambda_1)=\frac{B(0)(\lambda_2-\lambda_1)}{2\pi}+\frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{e^{-i\lambda_2 n}-e^{-i\lambda_1 n}}{-in}B(n)+$$

$$+\frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{e^{i\lambda_2 n}-e^{i\lambda_1 n}}{+in}B(n)=$$

$$=B(0)\frac{(\lambda_2-\lambda_1)}{2\pi}-\frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{\left(e^{-i\lambda_2 n}-e^{i\lambda_2 n}\right)-\left(e^{-i\lambda_1 n}-e^{+i\lambda_1 n}\right)}{in}B(n)=$$

$$=B(0)\frac{(\lambda_2-\lambda_1)}{2\pi}-\frac{1}{\pi}\sum_{n=1}^{\infty}\frac{-\operatorname{sen} n\lambda_2+\operatorname{sen} n\lambda_1}{n}B(n) \qquad \text{c.h.d.} \tag{15}$$

3. Particularmente notables son los casos de la (14) si $\lambda_2=\lambda$ y $\lambda_1=0$ ó $\lambda_1=-\pi$. Tenemos las expresiones:

a) Para $\lambda_1 = 0$:

$$F(\lambda) - F(0) = B(0)\frac{\lambda}{2\pi} + \frac{1}{\pi}\lim_{N\to\infty}\sum_{n=1}^{N} B(n)\frac{\operatorname{sen}\lambda n}{n} \tag{16}$$

b) Para $\lambda_1 = -\pi$:

$$F(\lambda) = B(0)\frac{[\lambda+\pi]}{2\pi} + \frac{1}{\pi}\lim_{N\to\infty}\sum_{n=1}^{N} B(n)\frac{\operatorname{sen}\lambda n}{n} \tag{16'}$$

4. Decimos que $F(\lambda)$ está normalizada en λ si es:

$$F(\lambda) = \frac{F(\lambda+0) + F(\lambda-0)}{2} \tag{17}$$

por lo que la (16) y (16') deben entenderse en este sentido.

Teorema 3.°

1. Las diferencias de la función de distribución espectral de un proceso estocástico estacionario real $\{\xi_n,\ n \in Z\}$ en intervalos simétricos respecto al origen, son iguales; es decir:

$$F(\lambda_2) - F(\lambda_1) = F(-\lambda_1 - 0) - F(-\lambda_2 - 0) \tag{18}$$

2. El primer miembro de la izquierda corresponde al intervalo semiabierto por la izquierda y cerrado por la derecha: $(\lambda_1, \lambda_2]$, ... y el segundo es el simétrico y deberá ser cerrado por la izquierda y abierto por la derecha: $[-\lambda_2, -\lambda_1)$.

Determinando en este último intervalo la diferencia de las funciones de distribución de (18) según (14), tendremos:

$$F(-\lambda_1 - 0) - F(-\lambda_2 - 0) = \frac{B(0)(\lambda_2 - \lambda_1)}{2\pi} +$$

$$+ \frac{1}{\pi}\sum_{n=-1}^{\infty} \frac{\operatorname{sen}\lambda_2 n - \operatorname{sen}\lambda_1 n}{n} B(n) \tag{19}$$

El segundo miembro de (18) no es sino la (14); luego hemos demostrado la (18).

Teorema 4.°

1. En todo proceso real estacionario $\{\xi_n, \quad n \in Z\}$ cuya función de covarianza se representa por (13), también puede representarse por:

$$B(n) = \int_0^{\pi} \cos n\lambda dG(\lambda) \tag{20}$$

donde $G(\lambda)$ se define como una función monótona, no decreciente y acotada, denominada igualmente función de distribución espectral y tal que

$$\frac{G(\lambda+0)+G(\lambda-0)}{2} - G(0) = \frac{B(0)\lambda}{\pi} + \frac{2}{\pi}\sum_{n=1}^{\infty} B(n)\frac{\operatorname{sen} n\lambda}{n} \tag{21}$$

2. *Demostración:*

Formemos, como Cramer (I) la función:

$$G(\lambda) = F(\lambda) - F(-\lambda-0) =$$

$$= F(\lambda) - [F(\pi) - F(\lambda)] = 2\,F(\lambda) - B(0) \tag{22}$$

$$F(\lambda) = \frac{G(\lambda)+B(0)}{2} \qquad 0 \leqslant \lambda \leqslant \pi \tag{23}$$

$$dF(\lambda) = \frac{dG(\lambda)}{2} \tag{24}$$

Recordando el Teorema 2.°, la representación de las sucesiones de covarianza (13) $\{B(n), \quad n \in Z\}$ con esta función toma la expresión (20).

Si la (23) la sustituimos en la (14) y la ponemos en forma normalizada y hacemos $\lambda_1 = 0$ y $\lambda_2 = \lambda$ tendremos:

$$\frac{G(\lambda)+B(0)}{2} - \frac{G(0)+B(0)}{2} = \frac{B(0)\lambda}{2\pi} + \frac{1}{\pi}\lim_{N\to\infty}\sum_{n=1}^{N}\frac{\operatorname{sen} n\lambda}{n}B(n) \quad \Rightarrow$$

$$\Rightarrow G(\lambda) - G(0) = \frac{B(0)\lambda}{\pi} + \frac{2}{\pi}\lim_{N\to\infty}\sum_{n=1}^{N}\frac{\operatorname{sen} n\lambda}{n}B(n)$$

c.h.d. (25)

(I) Cramer: o. c., pág. 136. Doob, pág. 476, o. c. hace la sustitución $G(\lambda) = 2[F(\lambda) - F(+0)] + F(+0) - F(-0) \Rightarrow$

$$F(\lambda) = \frac{G(\lambda)}{2} + \frac{F(+0)+F(-0)}{2} \quad \text{y para } G(0) = 0$$

donde:

$$G(\lambda)=\frac{G(\lambda+0)+G(\lambda-0)}{2}$$

III. *Función de distribución espectral correspondiente a un proceso estocástico estacionario de parámetro continuo*

Teorema 1.°

1. La función de covarianza correspondiente a un proceso estacionario $\{\xi(t), \quad t\in R\}$ representable según Kintchine por:

$$B(t)=\int_{-\infty}^{\infty} e^{i\lambda t}dF(\lambda) \tag{26}$$

le corresponde por función de distribución espectral la expresión:

$$\frac{F(\lambda_2+0)+F(\lambda_2-0)}{2}-\frac{F(\lambda_1+0)+F(\lambda_1-0)}{2}=$$

$$=\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-i\lambda_2 t}-e^{-i\lambda_2 t}}{-it}B(t)dt \tag{27}$$

2. *Demostración*

2.1. Seguiremos a Doob (I) y seremos concisos porque es muy conocida esta fórmula de inversión en los libros de Estadística.

2.2. Formaremos la función $\beta(\cdot)$ (4) excepto que el dominio es el eje real y $(\lambda_1, \lambda_2)\in R$.

La transformada de Fourier $\beta^x(\cdot)$ de la función $\beta(\cdot)$ es:

$$\beta^x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\beta(\lambda)e^{-i\lambda t}d\lambda=\frac{e^{-i\lambda_2 t}-e^{-i\lambda_1 t}}{-2\pi it} \tag{28}$$

y fijémonos que los coeficientes de $B(t)$ en (27) son precisamente (28).

2.3. Por las propiedades de las transformadas de Fourier la inversa de $\beta^x(t)$ es:

$$\beta(\lambda)=\lim_{T\to\infty}\int_{-T}^{T}\beta^x(t)e^{it\lambda}d\lambda \tag{29}$$

(I) Doob: o. c., pág. 520.

Esta expresión es convergente, pues para que exista $\beta^x(t)$ es suficiente:

$$\int_{-\infty}^{\infty} /\beta(\lambda)/\, d\lambda < \infty \qquad (30)$$

Por la razón de estar acotada $B(t)$ sustituyendo $B(t)$ por (26) en la expresión siguiente, tenemos:

$$\int_{-T}^{T} \beta^x(t)\, B(t)\, dt = \int_{-T}^{T} \beta^x(t) dt \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda)$$

$$\int_{-\infty}^{\infty} \mathrm{dF}(\lambda) \int_{-T}^{T} \beta^x(\mathrm{t}) \mathrm{e}^{i\lambda t} \mathrm{dt} \to \int_{-\infty}^{\infty} \beta(\lambda)\, dF(\lambda) \qquad (31)$$

El valor de la última integral es idéntico a la obtenida en (12) pues los valores de $\beta(\,\cdot\,)$ según se han establecido en 2.2. son los mismos que la función (4) excepto en los límites fuera del intervalo λ_1, λ_2 que la hicimos $\beta(\lambda)=0$.

En consecuencia (31) se descompone en las mismas tres integrales no negativas de (12) y el resultado es el primer miembro de la (27). Sustituyendo $\beta^x(t)$ de (31) por los valores que obtuvimos en (28) en el límite, demostramos (27).

Teorema 2.º

1. Si la función de covarianza estacionaria $B(t)$ es real, puede representarse por

$$B(t) = \int_0^{\infty} \cos \lambda t d\, G(\lambda) \qquad t \in R \quad (32)$$

y la función de distribución espectral $G(\lambda)$ es monótona, no decreciente y acotada, tal que:

$$\frac{G(\lambda+0)+G(\lambda-0)}{2} - G(0) = \frac{2}{\pi} \int_0^{\infty} \frac{\operatorname{sen} t\lambda}{t} B(t) dt$$

$$\lambda > 0 \qquad (33)$$

2. *Demostración*

2.1. La (32) se deduce inmediatamente recordando que $B(-t)=B(t)$

igual que hicimos en (9) excepto límites y haciendo sustitución semejante en (26) donde ahora:

$$F(\lambda)=\frac{G(\lambda)+B(0)}{2} \qquad \lambda>0 \tag{34}$$

tendremos la (32).

2.2. La segunda parte del teorema (33) se deduce sustituyendo los valores normalizados de $F(\lambda)$ en (27) por (34), recordando que la función de covarianza es una función par y haciendo $\lambda_2=\lambda$ y $\lambda_1=0$:

$$\frac{G(\lambda)+B(0)}{2}-\frac{G(0)+B(0)}{2}=\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-\lambda t}-1}{-it}B(t)\,dt=$$

$$\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{\cos\lambda t-i\operatorname{sen}\lambda t-1}{-it}B(t)\,dt=$$

$$=\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{\operatorname{sen}\lambda t}{t}B(t)\,dt=\frac{1}{\pi}\int_{0}^{\infty}\frac{\operatorname{sen}\lambda t}{t}B(t)dt\Rightarrow$$

$$\Rightarrow G(\lambda)-G(0)=\frac{2}{\pi}\int_{0}^{\infty}\frac{\operatorname{sen}\lambda t}{t}B(t)\,dt \qquad \text{c.h.d.}$$

porque la función:

$$\frac{1-\cos\lambda t}{t}\cdot B(t)$$

es impar y su integral en $(-T, T)$ es nula.

IV. *Continuidad y Discontinuidad de $F(\lambda)$:*

4.1. La función de distribución en un punto de discontinuidad satisface la relación:

$$F(\lambda-0)\leqslant F(\lambda)\leqslant F(\lambda+0) \qquad \forall\lambda \tag{35}$$

Seguiremos a Cramer, Doob (I), etc., y tomaremos usualmente la continuidad a la derecha, poniendo:

$$F(\lambda)=F(\lambda+0) \tag{35'}$$

(I) H. Cramer: «Stationary Processes», c. o., pág. 128, y J. Doob, o. c., págs. 521 y 522.

4.2. La función $F(\lambda)$ nos permite conocer, como veremos, la estructura del proceso. En particular *si en el conjunto de variabilidad de* λ, $F(\lambda)$ crece solamente a saltos, en esos puntos aislados de λ, en los que varía bruscamente $F(\lambda)$ son los «valores espectrales» del proceso de *frecuencias propias* y que influyen en su realización armónica; es decir: tiene un carácter cíclico.

4.3. La función de distribución espectral $F(\lambda)$ (para $F(-\infty)=0$) nos indica la contribución a la varianza total del proceso hasta la frecuencia λ. Es de señalar la fórmula de inversión que proporciona directamente la función $F(\lambda)$ obtenida por el autor, que es:

$$F(\lambda)=\frac{B(0)}{2}+\frac{1}{2\pi}\int_0^\infty \frac{e^{i\lambda t}B(-t)-e^{-i\lambda t}B(t)}{it}dt \qquad (36)$$

Si $B(0)=1$ entonces $B(t)$ es una función característica cuando $F(\lambda)$ es una distribución de probabilidad y la fórmula indicada para este caso la dedujo Gil Peláez (I). La (36) la deducimos de manera similar a la de Gil Peláez, pero teniendo presente que $F(\infty)=B(0)$ puede no ser igual a la unidad, empleando funciones de covarianza $B(t)$. Omitimos su demostración por ser prácticamente similar.

V. *Función de densidad espectral*

5.1. Si $F(\cdot)$ es absolutamente continua, esto es: la integral de su derivada, el proceso se dice tiene función de densidad espectral y escribiremos para $n \in Z$:

$$f(\lambda)=F'(\lambda) \qquad \text{caso } B(t) \text{ complejo} \qquad (37)$$

$$g(\lambda)=G'(\lambda)=2F'(\lambda) \qquad \text{caso real} \qquad (38)$$

5.2. Función de densidad de un proceso $\{\xi_n, \quad n \in Z\}$:

Para que exista una función de densidad espectral, es necesario que converja absolutamente:

$$\Sigma/B(n)/<\infty \qquad (39)$$

(I) L. Gil Peláez: «Biométrica», 38, 1951, pág. 481. Citan a este autor P. Lévy, en Théorie de l'Addition des variables aleatoires, Gautier-Villars, 1954, pág. 159, y J. Koerts y Abrahame: «On the application of the general lineal model», Rotterdam University Press, 1971, pág. 77 y sig. y también Luckacs, o. c. pág. 33 y no consideramos correcto a Kendall, «The advanced Theory of Statistics», que deduce la fórmula de Gil Peláez, pág. 95, 1968, 3.ª ed. y no cita a su autor ni siquiera en su extensa bibliografía.

Entonces, de la (16) derivando, tenemos:

$$f(\lambda)=F'(\lambda)=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty} B(n)e^{-i\lambda n} \qquad -\pi<\lambda\leqslant\pi \tag{40}$$

y es consecuencia de las series absolutamente convergentes que pueden derivarse e integrarse término a término y tenemos la correspondencia entre las expresiones (2) y (40).

5.3. Si el proceso estacionario $\{\xi_n, \quad n \in Z\}$ fuese real, cumpliendo la condición (39), derivando (21) término a término, tenemos:

$$G'(\lambda)=g(\lambda)=2f(\lambda)=\frac{B(0)}{\pi}+\frac{2}{\pi}\sum_{n=1}^{\infty} B(n)\cos n\lambda \tag{41}$$

A $g(\lambda)$ y $f(\lambda)$ se las denomina densidades espectrales.

5.4. Si el proceso estacionario $\{\xi(t), \quad t \in R\}$ fuere complejo y la función $F(\lambda)$ fuese derivable, entonces de la (27) deduciríamos (I):

$$f(\lambda)=F'(\lambda)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\lambda t}B(t)\,dt \qquad \lambda \in R \tag{42}$$

La condición suficiente para la existencia de $f(\lambda)$ es que $B(t)$ sea absoutamente integrable, es decir:

$$\int_{-\infty}^{\infty} /B(t)/\,dt<\infty \tag{43}$$

porque si $B(t)$ es absolutamente integrable y $\varphi(t)$ una función acotada, entonces el producto es una función integrable. En nuestro caso $\varphi(t)=e^{-i\lambda t}$ es una función acotada (II).

5.5. En el caso de ser proceso real $\{\xi(t), \quad t \in R\}$ y cumpliéndose la condición (43) de (33), deducimos para la función de densidad:

$$g(\lambda)=\frac{2}{\pi}\int_{0}^{\infty} \cos\lambda t\, B(t)\,dt \tag{44}$$

(I) Muchos autores definen la función de distribución por esta transformada. Ver, por ejemplo, Anderson, o. c. pág. 380 que define $f(\lambda)$ por esta transformada de Fourier del coseno.

(II) Tolstov, Fourier Series. Prentice Hall, 1962, pág. 8.

VI. *Consecuencias importantes*

1.ª Para que exista función de densidad debe siempre cumplirse:

$$\sum_{-\infty}^{\infty} /B(n)/ < \infty \qquad \text{si} \qquad n \in Z$$

procedente de un proceso discreto; o

$$\int_{-\infty}^{\infty} /B(t)/\,dt < \infty$$

cuando el parámetro $t \in R$: $\{\xi(t), \quad t \in R\}$

2.ª Las funciones de densidad espectrales son siempre no negativas, por proceder de una función de distribución monótona no decreciente.

3.ª También se definen y son las transformadas de Fourier de la función de covarianza estacionaria $B(t)$ y la transformada del coseno de Fourier cuando $B(t) = B(-t)$ $(t \in R)$.

4.ª En los procesos de parámetro $t \in R$, la función λ varia en el campo real y el proceso es real, por ser $f(\lambda) = f(-\lambda)$ se estudia en el semieje positivo $\lambda \geqslant 0$.

5.ª En los procesos de parámetro $t \in Z$, el campo de variación de λ es el intervalo $(-\pi, +\pi)$; y si el proceso es real por ser la función de densidad una función par, $g(\lambda) = 2f(\lambda)$ se estudia en el intervalo $(0, +\pi)$.

Sección 7.ª

Aliasamiento Frecuencia de Nyquist

1. *Concepto*

Las observaciones experimentales de un proceso $\{\xi(t), t \in R\}$ pueden hacerse en intervalos equidistantes de tiempo, pero no de una manera continua.

Si las observaciones se hacen en el conjunto:

$$D_T = \{- \ldots -d, \quad 0, \quad +d, \quad +2d, \quad + \cdots + nd \ldots\} \tag{1}$$

es natural que las frecuencias del proceso $\{\xi(t), \quad t \in R\}$ difieran de las del proceso $\{\xi(t), \quad t \in D_T\}$.

La elección del intervalo $\Delta t = d$ (2)

influye en la calidad y cantidad de información que el investigador ha de disponer.

Si de una serie cronológica se recogen a intervalos equidistantes del tiempo $\Delta t = d$ la frecuencia $\pi/\Delta t$ (I) se denomina de Nyquist y se pierde información de la serie cronológica para frecuencias superiores a la de Nyquist, ya que al pretender determinar el espectro, la descomposición del proceso en armónicos, carecemos de información suficiente para los de frecuencias superiores; por ejemplo:

$$\frac{\Delta t}{4} \text{ la frecuencia sería: } \frac{8\pi}{\Delta t}$$

Expondremos sencillamente el fundamento matemático relacionando los espectros de ambos procesos y con una sencilla gráfica justificaremos las consideraciones anteriores.

2. *Funciones de covarianza y de distribución espectral.*

Si del proceso estocástico de parámetro continuo $\{\xi(t)., \ t \in R\}$ en los $dh\{h \in Z\}$ efectuamos observaciones, tenemos un proceso de tipo discreto $\xi(dh)$ y cuya representación espectral de la función de covarianza, según la (3') de la Sección 4, es:

$$B(dh) = \int_{-\frac{\pi}{d}}^{\frac{\pi}{d}} e^{+idh\lambda} dF_d(\lambda) \qquad dh \in D_T \tag{3}$$

$$= \int_{-\pi}^{\pi} e^{ivh}\, dF_d\left(\frac{V}{d}\right) \tag{3'}$$

Para la del proceso de parámetro continuo es:

$$B(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda) \tag{4}$$

siendo $F_d(\lambda)$ y $F(\lambda)$ las funciones de distribución, espectrales de los procesos discreto y continuo. Por ser un mismo proceso, las funciones de covarianza deben ser iguales, cuando $t = dh$.

(I) J. García-Villalón: «Análisis Espectral de Series Temporales en Economía», Anales del Instituto de Actuarios Españoles, Madrid, 1967, págs. 37 y siguientes.

La función de covarianza de (4) para $t=hd$

$$B(hd)=E\{\xi(t+hd)\overline{\xi(t)}\}=\int_{-\infty}^{\infty} e^{i\lambda dh}\, dF(\lambda)=$$

$$\sim \sum_{j=-\infty}^{\infty} \int_{(2j-1)\frac{\pi}{d}}^{(2j+1)\frac{\gamma\pi}{d}} e^{i\lambda dh} dF(\lambda)=$$

$$=\sum_{j=-\infty}^{\infty} \int_{-\pi}^{\pi} e^{ih(2\pi j+v)} dF\left(\frac{2j\pi+v}{d}\right)=$$

$$=\int_{-\pi}^{\pi} e^{ihv} \sum_{j=-\infty}^{+\infty} dF\left(\frac{v+2j\pi}{d}\right) \tag{5}$$

por ser $e^{ih2\pi j}=1$ y al hacer el cambio

$$\lambda=\frac{v+2\pi j}{d} \tag{6}$$

los límites de la integral se transforman $\pm\pi$, y de las (3') y (5) para que sean iguales ambas funciones de covarianza para $\forall h \in Z$, debe ser:

$$dF_d\left(\frac{v}{d}\right)=dF\left(\frac{v}{d}\right)+\sum_{j=1}^{\infty}\left[dF\left(\frac{v+2\pi j}{d}\right)+dF\left(\frac{v-2j\pi}{d}\right)\right] \tag{7}$$

$$-\pi<v\leqslant+\pi$$

La (7) nos muestra la relación entre las funciones de distribución espectrales de ambos procesos.

Si fuesen derivables, recordando (7) y (3') tenemos:

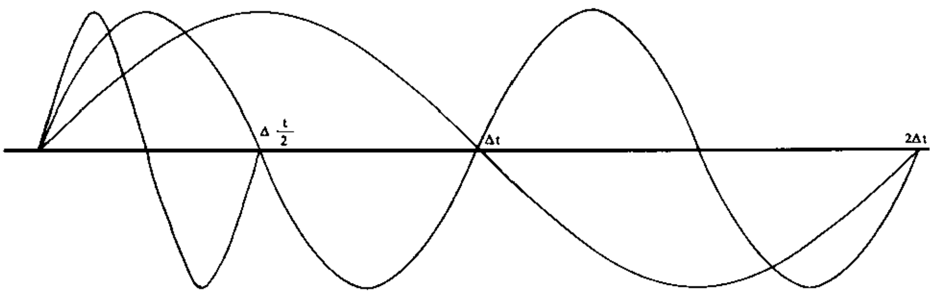
$$f_d\left(\frac{v}{d}\right)=f\left(\frac{v}{d}\right)+\sum_{j=1}^{\infty}\left[f\left(\frac{v+2\pi j}{d}\right)+f\left(\frac{v-2\pi j}{d}\right)\right] \tag{8}$$

y es la relación entre las densidades espectrales, entre los procesos discreto y continuo. Y este efecto es denominado aliasamiento.

Pero de las observaciones no podemos determinar la función de distribución espectral, pues carecemos de suficiente información para las altas frecuencias.

3. *Gráfica del aliasamiento.*

Representaremos en el eje de tiempos las distintas sinusoides, eligiendo la frecuencia de Nyquist como básica (o fundamental) para demostrar cómo se acumulan las frecuencias en los puntos Δt y superiores. Al período de $2\Delta t$ le corresponde la frecuencia angular es: $\frac{2\pi}{2\Delta t}=\frac{\pi}{\Delta t}$ que es la frecuencia de Nyquist.



En la gráfica hay otra frecuencia, de la que carecemos de datos: la correspondiente a $\frac{\Delta t}{2}$; es decir $\frac{4\pi}{\Delta t}$ y así para otras frecuencias superiores.

4. *Consideraciones prácticas del «aliasamiento»*

1. Los picos en el espectro para frecuencias inferiores a $\frac{1}{2}$ ciclo por unidad de tiempo, supone son debidas a causas superiores al intervalo de frecuencia observable y para frecuencias superiores a la de Nyquist no es fiable el espectro.

2. En Economía, si se elige como unidad de tiempo el día, puede aparecer un pico correspondiente a 30,437 días $\left(\text{frecuencia } \frac{2\pi}{30{,}437}\right)$; o si la unidad de tiempo fuese la semana sin pico de un período de 4,348 semanas $\left(\text{frecuencia angular } \frac{2\pi}{4{,}348}\right)$ e indicaría una marcada periodicidad mensual (I).

(I) Granger y Hatanaka, o. c., pág. 47.

Sección 8.ª

Representaciones espectrales de procesos estocásticos estacionarios

1. *Introducción*

En las Secciones precedentes hemos estudiado los procesos estacionarios y hemos visto que las funciones de covarianza (dependiente de la diferencia de tiempos exclusivamente) eran funciones no negativas y por tanto representables en forma de funciones características y asociadas a estos procesos estudiamos las funciones de distribución espectral.

Sin pérdida de generalidad supondremos que la media del proceso sea constante $E\xi(t)=0$.

Pretendemos encontrar un conjunto de variables aleatorias (con media cero) $\{\zeta_s : s=1, 2, \ldots\}$ *mutuamente ortogonales y tales que por combinación* lineal de las mismas, con coeficientes funcionales, pueda arbitrariamente aproximarse al proceso $\xi(t)$ sobre cualquier intervalo $(-T, T)$.

Más precisamente: Para $t \in (-T, +T)$ y $\varepsilon>0$ existe un N tal que para $\forall m>N$ se cumpla:

$$E/\xi(t)-\sum_{k=1}^{m} e^{i\lambda_k t}\zeta_k/^2<\varepsilon \qquad (1)$$

Las variables aleatorias ζ_k con coeficientes no aleatorios $e^{i\lambda_k t}$ forman m procesos estocásticos cada uno de ellos formado por procesos de tipo cíclico: $\zeta_k e^{i\lambda_k t}$.

Es natural que precisemos el grado de exactitud o aproximación y qué mejor que en términos de probabilidad. Si dados ε y δ arbitrariamente pequeños, existe un $m>N$ tal que para todo m podemos escribir:

$$P\{/\xi(t)-\sum_{k=1}^{m} e^{i\lambda_k t\zeta_k} < \varepsilon\} > 1-\delta \qquad (2)$$

nos indica que $\{\xi(t),\ t \in T\}$ puede elegirse en sentido de convergencia en probabilidad, por una combinación lineal de sinusoides con amplitudes aleatorias y mutuamente ortogonales; es decir, de un proceso:

$$\xi_1(t)=\sum_{k=1}^{m} e^{i\lambda_k t\zeta_k} \qquad (2')$$

que se aproxima a $\xi(t)$.

2. *Conceptos previos*

2.1. El proceso estocástico

$$\zeta_j e^{i\lambda_j t} = \zeta_j[\cos \lambda_j t + i \text{ sen } \lambda_j t] \tag{3}$$

es el generador del proceso sumatorio (2') donde previamente definiremos todos los elementos que intervienen en (3) para la mejor comprensión de nuestra exposición.

2.2. Las ζ_j son variables aleatorias ortogonales de las características conocidas:

$$E\zeta_j = 0 \qquad E\{/\zeta_j/^2\} = \sigma_j^2 \qquad E\zeta_j\overline{\zeta_k} = 0 \qquad j \neq k \tag{4}$$

2.3. t es el parámetro del proceso y normalmente es el tiempo. En las series económicas se toma una unidad de medida: el día, mes, semana, año, etc.

Advirtamos que el año real tiene 365, 25 ... días y existiendo meses con distintos días, a los efectos frecuenciales, el mes se considera de 30,4375 días; y medido en semanas 4,3482 semanas.

En las series económicas, la información se recoge según el calendario, pero repercute en el espectro.

2.4. El coeficiente constante λ_j se denomina frecuencia angular. Si se considera un móvil que se desplaza sobre la circunferencia de radio unidad y con velocidad uniforme λ_j la posición del mismo después de transcurrir un tiempo t es:

$$e^{i\lambda_j t}$$

λ_j indica (para un mismo t) la mayor o menor velocidad del móvil y si t es igual a la unidad, λ_j está medido en radianes e indica el recorrido que ha efectuado de circunferencia o circunferencias. Si no hubiese comenzado en el origen sino con un ángulo θ a partir del origen, entonces la posición del móvil después de transcurrir un tiempo t a la velocidad λ_j la posición en la circunferencia sería:

$$e^{i\lambda_j t + \theta i}$$

A θ se denomina «fase».

2.5. La frecuencia puede medirse en grados, en cuyo caso las fórmulas de relación son:

$$\frac{\lambda_j^\circ}{360^\circ} = \frac{\lambda_j}{2\pi} \Rightarrow \lambda_j = \frac{\lambda_j^\circ}{360^\circ} \cdot 2\pi \Rightarrow \lambda_j^\circ = \frac{\lambda_j}{2\pi} 360$$

2.6. También de la significación de (3) representativa del movimiento circular de velocidad angular λ_j constante, podemos tomar la circunferencia, como unidad de medida y entonces la frecuencia medida en circunferencias o «ciclos» es:

$$\lambda'_j=\frac{\lambda_j}{2\pi} \tag{5}$$

El numerador nos indica la longitud expresada en radianes y en consecuencia, el cociente, representa el número completo o no de circunferencias que recorre en la unidad de tiempo.

La frecuencia angular λ_j la designamos usualmente como frecuencia.

2.7. Se denomina período al inverso de la frecuencia medida en «ciclos» λ'_j

$$P=\frac{2\pi}{\lambda_j}=\frac{1}{\lambda'_j} \tag{6}$$

y es el tiempo mínimo para que el proceso (3) sea idéntico. Consecuencia de la definición, tenemos:

$$\zeta_j e^{i\lambda_j t}=\zeta_j e^{i\lambda_j\left(t+\frac{2\pi}{\lambda_j}\right)}=\zeta_j e^{i\lambda_j(t+P)} \tag{7}$$

recordando:

$$e^{2\pi k}=1 \qquad\qquad k\in Z$$

y por el carácter cíclico de las funciones trigonométricas: sen $(\lambda_j t+2\pi)=$ $=\operatorname{sen}\lambda_j t;$ $\cos(\lambda_j+2\pi)=\cos\lambda_j t$

2.8. El proceso (3) es estacionario por 2.2. y también el proceso (2') formado por sumas de procesos del tipo cíclico con las características:

$$E\xi_1(t)=0 \tag{8}$$

$$B_1(t-s)=\sum_{j=1}^{m} e^{i\lambda_j(t-s)}\sigma_j^2 \tag{9}$$

$$B_1(0)=\sigma^2=\sum_{j=1}^{m}\sigma_j^2 \tag{10}$$

3. *Deducción heurística de la representación espectral de procesos*

3.1. Ordenemos de menor a mayor las frecuencias angulares de (2) y en las líneas siguientes sus variables aleatorias ortogonales y las varianzas. Formemos el cuadro:

$$\begin{array}{ccccccc} \lambda_1 < \lambda_2 & \cdots & < \lambda_j < & \cdots & < \lambda_m \\ \zeta_1 \quad \zeta_2 & \cdots & \zeta_j & \cdots & \zeta_m \\ \sigma_1^2 \quad \sigma_2^2 & \cdots & \sigma_j^2 & \cdots & \sigma_m^2 \end{array} \tag{11}$$

por lo que cada subíndice se corresponde con componentes de procesos elementales.

3.2. Introduzcamos la variable λ que recorra el eje real o el intervalo $(-\pi, +\pi]$ y de (11) asociemos las funciones siguientes relacionadas con λ:

$$\zeta(\lambda) = \sum_{\lambda_j \leqslant \lambda} \zeta_j \tag{12}$$

$$F(\lambda) = \sum_{\lambda_j \leqslant \lambda} \sigma_j^2 \tag{13}$$

Por las definiciones dadas en (4) y (11) las diferenciales de (12) y (13) gozan de las siguientes propiedades:

$$dF(\lambda) = \begin{cases} \sigma_j^2 & \lambda = \lambda_j (i = 1, 2, \ldots, m) \\ 0 & \text{en otro caso} \end{cases} \tag{14}$$

$$d\zeta(\lambda) = \begin{cases} \zeta_j & \lambda = \lambda_j (j = 1, 2, \ldots, m) \\ 0 & \text{en otro caso} \end{cases} \tag{15}$$

Las relaciones (12) a (15) nos permiten conocer mejor el significado de $F(\lambda)$, $\zeta(\lambda)$ y el significado de las representaciones espectrales.

3.3. La representación de la función de covarianza (9) con la notación introducida en (14) la podemos escribir por una integral de Stieltjes:

$$B(t-s) = \int_\Lambda e^{i\lambda(t-s)} dF(\lambda) \tag{16}$$

porque para $\lambda_i < \lambda < \lambda_{i+1}\, dF(\lambda) = 0$.

El conjunto Λ representativo de un intervalo finito (o infinito) necesariamente deberán contener todas las frecuencias λ_i.

En este caso, la función de distribución $F(\lambda)$ crece a saltos y precisamente en los puntos de discontinuidad (denominado espectro) los valores son:

$$\sigma_j^2, \ \{(j=1,\ 2,\ \ldots,\ m)\}$$

Una de las componentes de la varianza total del proceso (2') es, precisamente, σ_j^2 que es la varianza del proceso (3).

3.4. El proceso (2') tiene una representación espectral semejante.

Examinemos previamente la notación introducida en (15). Recordando (4) y (12), $\zeta(\lambda)$ es un proceso de incrementos ortogonales; y cumple las condiciones para $\lambda_1 < \lambda_2 \leqslant \lambda_3 < \lambda_4$:

$$E\zeta(\lambda)=0 \Rightarrow Ed\zeta(\lambda)=0$$

$$E\{[\zeta(\lambda_2)-\zeta(\lambda_1)]\,[\overline{\zeta(\lambda_4)-\zeta(\lambda_3)}\}=0 \tag{17}$$

$$E/d\zeta(\lambda)/^2\}=dF(\lambda) \tag{18}$$

de acuerdo con (4), (13) y (15).

Con la notación introducida en (14) el proceso (2') puede escribirse:

$$\xi_1(t)=\int_\Lambda e^{i\lambda t}\,d\zeta(\lambda) \tag{19}$$

siendo Λ el conjunto R si $t \in R$ o el $(-\pi,\ +\pi)$ si $t \in Z$.

En (2') los puntos de λ eran $\{\lambda_j;\ j=1,\ 2,\ \ldots,\ m\}$ y en (19) λ se considera como una variable, siendo su campo de variabilidad el dominio de las frecuencia Λ. Cuando $\lambda_i < \lambda < \lambda_{i+1}$ por (15) $d\zeta(\lambda)=0$ y en los valores interiores al intervalo $[\lambda_i,\ \lambda_{i+1}]$ se anula la sumatoria estocástica de Stieltjes (19).

3.5. Esta representación espectral para el proceso estocástico estacionario (2') válida cuando el espectro posee frecuencias propias, induce a considerar si también puede extenderse cuando el espectro tome valores continuos y esta respuesta es afirmativa siempre que la función de distribución espectral sea continua.

Este problema de descomposición de un proceso estocástico estacionario de segundo orden en una suma de infinitos términos aleatorios de tipo armónico, mutuamente ortogonales y la convergencia de esta suma infinita al proceso estocástico estacionario se entenderá siempre en sentido de la media cuadrática.

4. *Representación espectral de un proceso estacionario de segundo orden*

1. Seguiremos en la demostración de este teorema fundamental de Cramer (I), cuando sea el parámetro t continuo y con la hipótesis de continuidad en media cuadrática.

$$\lim_{t-s\to 0} E\{/\xi(t)-\xi(s)/^2\}=0 \tag{20}$$

$$\text{y de ser:} \qquad E\xi(t)=0 \tag{21}$$

Enunciemos el teorema así:

2. A todo proceso estocástico estacionario continuo en m.c. $\xi(t)$ puede asociarse un proceso $\zeta(\lambda)$ con incrementos ortogonales tal que para cada t fijo el proceso $\xi(t)$ puede representarse espectralmente:

$$\xi(t) \stackrel{2}{=} \int_{-\infty}^{\infty} e^{i\lambda t}\, d\zeta(\lambda) \tag{22}$$

siendo $\zeta(\lambda)$ (23) un proceso de incrementos ortogonales y se define salvo una variable aleatoria aditiva si se fija tomando $\zeta(-\infty)=0$, con las características:

$$E\zeta(\lambda)=0 \qquad E\{/\zeta(\lambda)/^2\}=F(\lambda) \tag{24}$$

$$E\{/d\zeta(\lambda)/^2\}=dF(\lambda) \tag{25}$$

3. Utilizaremos las propiedades del espacio de Hilbert para la demostración de este teorema.

Consideremos dos espacios de Hilbert: uno de elementos estocásticos y el otro funcional. Estableceremos una correspondencia entre los elementos de cada espacio tal que sea idéntica la métrica entre los elementos correspondientes: es decir, los espacios sean isométricos.

Llamaremos $H(\xi)$ al espacio estocástico y $H(F)$ al espacio funcional con respecto a la medida $dF(\cdot)$.

Espacio $H(\xi)$:

Formado por variables aleatorias y por todas las combinaciones lineales finitas:

$$\alpha_1\xi_1(t_1)+\cdots+\alpha_n\xi_n(t_n) \tag{26}$$

y todos los límites en media cuadrática de secuencias de tales combinaciones con las propiedades siguientes, para $\forall \eta_1\ \eta_2 \in H(\xi)$:

(I) Cramer, o. c., págs. *128* y siguientes.

1.ª Momento de segundo orden finito:

$$//\eta//^2 < \infty \tag{27 a}$$

2.ª El producto interno se define por una esperanza matemática:

$$(\eta_1, \eta_2) = E\eta_1\bar{\eta}_2 \tag{27 b}$$

3.ª La distancia al cuadrado es:

$$//\eta_1 - \eta_2//^2 = E\{/\eta_1 - \eta_2/^2\} \tag{27 c}$$

y los elementos η_1, η_2 se consideran idénticos cuando su distancia es nula.

Espacio $H(F)$:

Formado por todas las funciones de cuadrado integrable respecto de $dF(\lambda)$ y todas las combinaciones lineales finitas de mencionadas funciones:

$$\alpha_1 g_1 + \alpha_2 g_2 + \cdots + \alpha_n g_n \tag{28}$$

y todos los límites en media cuadrática de las secuencias tales que:

Para $\forall g_1\ g_2 \in H(F)$ se verifica:

1.ª Las integrales de cuadrado respecto dF sean finitas:

$$//g//^2 = \int_{-\infty}^{\infty} /g(\lambda)/^2\, dF(\lambda) < \infty \tag{29 a}$$

2.ª Producto escalar; se define así:

$$(g_1, g_2) = \int_{-\infty}^{\infty} g_1(\lambda)\overline{g_2(\lambda)}\, dF(\lambda) \tag{29 b}$$

3.ª Distancia al cuadrado:

$$//g_1 - g_2//^2 = \int_{-\infty}^{\infty} /g_1(\lambda) - g_2(\lambda)/^2\, dF(\lambda) \tag{29 c}$$

Dos elementos se consideran idénticos cuando la distancia sea nula.

4. Examinemos en qué condiciones los espacios $H(\xi)$ y $H(F)$ son isométricos. En principio tomemos elementos de ambos espacios y hagamos la correspondencia:

$$\eta = \xi(t) \rightarrow e^{i\lambda t} = g(t) \tag{30}$$

$t \in R$ consideramos un parámetro; $\eta \in H(\xi)$ $g(t) \in H(F)$. Estos elementos cumplen las condiciones indicadas en sus espacios respectivos.

Por la representación de Kintchine para la función de la covarianza del proceso estacionario $\xi(t)$, $t \in R$, tenemos:

$$(\eta_1, \eta_2) = E\xi(t)\overline{\xi(s)} = \int_{-\infty}^{\infty} e^{i\lambda(t-s)} dF(\lambda) \tag{31}$$

El producto interno de g_1, $g_2 \in H(F)$ por 29 b) es:

$$(g_1, g_2) = \int_{-\infty}^{\infty} e^{i\lambda(t-s)} dF(\lambda) \tag{31'}$$

siendo:

$$g_1 = e^{i\lambda t} \qquad \text{y} \qquad g_2 = e^{i\lambda s} \tag{32}$$

y que implica:

$$(\eta_1, \eta_2) = (g_1, g_2) \tag{33}$$

Igualmente para las distancias al cuadrado:

$$//g_1 - g_2//^2 = //\eta_1 - \eta_2//^2 = E\{/\xi(t) - \xi(s)/^2\} = \int_{-\infty}^{\infty} /e^{i\lambda t} - e^{i\lambda s}/^2 dF(\lambda) \tag{34}$$

La correspondencia se extiende a todas las combinaciones lineales:

$$\eta = \alpha_1 \xi(t_1) + \cdots + \alpha_n \xi(t_n) \qquad \eta \in H(\xi)$$

$$g(\lambda) = \alpha_1 e^{it_1\lambda} + \cdots + \alpha_n e^{it_n\lambda} \qquad g(\lambda) \in H(F) \tag{35}$$

Los puntos correspondientes a ambos espacios de estas combinaciones lineales son isométricos pues se comprueba cumplen:

$$(\eta_1, \eta_2) = (g_1, g_2) \tag{36 a}$$

$$//g_1 - g_2//^2 = /\eta_1 - \eta_2, \eta_1 - \eta_2/^2 = \int_{-\infty}^{\infty} /g_1(\lambda) - g_2(\lambda)/^2 dF(\lambda) \tag{36 b}$$

5. Estas relaciones nos demuestran la correspondencia biunívoca entre los elementos de ambos espacios y sus combinaciones lineales. Y para que sean idénticas η_1, η_2 deben de serlo sus homólogos g_1 y g_2.

6. Si existe una sucesión $\{\eta_n\}$ de variables estocásticas que converge en media cuadrática a η existirá otra sucesión funcional $\{g_n\}$ que convergerá igualmente a $g(\cdot)$ tal que:

$$//\eta_m - \eta_n// \overset{2}{=} \int_{-\infty}^{\infty} /g_m - g_n/^2 \, dF(\lambda) \to 0 \tag{37}$$

y si existe un elemento límite η (sucesiones de Cauchy) existirá otro elemento homólogo $g(\cdot) \in H(F)$ que también será límite de sucesiones de cuadro integrable.

En consecuencia: $\eta \overset{2}{=} \lim\limits_{n\to\infty} \eta_n$ (38)

y $$g \overset{2}{=} \lim g_n \tag{38'}$$

7. Establezcamos una correspondencia entre elementos funcionales de $H(F)$ y elementos de $H(\xi)$ de forma que los elementos de $H(\xi)$ tengan la propiedad de ser de incrementos ortogonales.

Para cualquier $\lambda_i \in R$ sea la función:

$$g(\lambda - \lambda_i) = \begin{cases} 1 & \lambda \leqslant \lambda_i \\ 0 & \text{en otro caso} \end{cases} \tag{39}$$

La función (39) es de cuadrado integrable respecto de $dF(\cdot)$, luego es un elemento de $H(F)$. Representemos su correspondiente en $H(\xi)$ por $\zeta(\lambda_i)$.

Al incremento $\zeta(\lambda_2) - \zeta(\lambda_1)$ (I) le corresponde el $g(\lambda - \lambda_2) - g(\lambda - \lambda_1)$ que toma la unidad si $\lambda_1 < \lambda \leqslant \lambda_2$ y cero en otro caso.

Para dos intervalos disjuntos (λ_1, λ_2) y (λ_3, λ_4) tendremos:

$$E\{[\zeta(\lambda_4) - \zeta(\lambda_3)]\,\overline{\zeta(\lambda_2) - \zeta(\lambda_1)}] =$$

$$= \int_{-\infty}^{\infty} [g(\lambda - \lambda_4) - g(\lambda - \lambda_3)]\,[\overline{g(\lambda - \lambda_2) - g(\lambda - \lambda_1)}]\, dF(\lambda) = 0$$

La distancia al cuadrado es:

$$E/\zeta(\lambda_2) - \zeta(\lambda_1)/^2 = \int_{-\infty}^{\infty} /g(\lambda - \lambda_2) - g(\lambda - \lambda_1)/^2 \, dF(\lambda)$$

$$= F(\lambda_2) - F(\lambda_1) \tag{40}$$

(I) Estos incrementos pertenecen a los espacios por ser combinaciones lineales.

y $$E/\zeta(\lambda)/^2=F(\lambda) \text{ si } F(-\infty)=0$$

Luego el proceso $[\zeta(\lambda),\ \lambda \in R\}$ tiene las propiedades (25); es decir, es un proceso de incrementos ortogonales y para $\lambda=\lambda_i$ el homólogo en el espacio $H(F)$ es la función definida en (39).

8. Hagamos una partición del intervalo $(-A,\ A)$ cualquiera:

$$-A=\lambda_1<\lambda_2<\cdots<\lambda_n<\lambda_{n+1}=A \tag{41}$$

y con el proceso $\{\zeta(\lambda),\quad \lambda \in (-A,\ A\}$ formemos la combinación lineal:

$$\eta_n=\sum_{j=1}^{n} e^{i\lambda_j t}[\zeta(\lambda_{j+1})-\zeta(\lambda_j)] \tag{42}$$

$\eta_n \in H(\xi)$ y al elemento η_n le haremos corresponder su homólogo $g_n(\lambda)$ del espacio $H(F)$:

$$g_n(\lambda)=\begin{matrix} e^{i\lambda_j t} & \lambda_j<\lambda\leqslant\lambda_{j+1} \\ 0 & \text{en otro caso} \end{matrix} \tag{43}$$

Inmediatamente comprobamos que:

$$//\eta_n//^2=//g_n//^2=F(\lambda_{n+1})-F(\lambda_1) \tag{44}$$

Si ampliamos el intervalo (41) y las subdivisiones de los mismos de forma que $\forall(\lambda_{j+1}-\lambda_j)\to 0$ existirá una sucesión de funciones g_n que convergerá en media cuadrática (38) al elemento límite:

$$g_n \xrightarrow{2} g(\lambda)=e^{i\lambda t} \qquad \in H(F) \tag{45}$$

e igualmente la sucesión de su homólogo tenderá el límite en m.c.:

$$\eta \overset{2}{=} \int_{-\infty}^{\infty} e^{i\lambda t}\, d\zeta(\lambda) \qquad \eta \in H(\xi) \tag{46}$$

Pero de (30) el elemento correspondiente de $e^{i\lambda t}$ era $\xi(t)$; luego:

$$\xi(t) \overset{2}{=} \int_{-\infty}^{\infty} e^{i\lambda t}\, d\,\zeta\,(\lambda) \qquad \text{c.h.d.} \tag{47}$$

9. Si el proceso $\{\xi(t),\ t \in R\}$ estacionario es real, entonces existen

dos procesos $c(\lambda)$ y $s(\lambda)$ de incrementos ortogonales y la representación espectral es:

$$\xi(t) \stackrel{2}{=} \int_0^\infty \cos \lambda t dc(\lambda) + \int_0^\infty \operatorname{sen} \lambda t ds(\lambda) \tag{48}$$

Los procesos $dc(\lambda)$ y $ds(\lambda)$ tienen las propiedades siguientes:

$$E/dc(\lambda)/^2 = E/ds(\lambda)/^2 = dG(\lambda) \qquad 0<\lambda \tag{49}$$

$$Edc(\lambda) \cdot \overline{dc(\lambda')} = Eds(\lambda)\overline{ds(\lambda')} = Edc(\lambda)\, ds(\lambda) = 0$$

En efecto: Descompongamos el proceso $\zeta(\lambda)$, en los dos siguientes:

$$\zeta(\lambda) = \frac{1}{2} \; [c(\lambda) - is(\lambda)] \tag{50}$$

siendo:

$$c(-\lambda) = c(\lambda) \qquad \text{y} \qquad s(-\lambda) = -s(\lambda) \tag{51}$$

Estos procesos son de incrementos ortogonales si:

$$\zeta(-\lambda) = \overline{\zeta(\lambda)} \tag{52}$$

la diferencial de (50) es:

$$d\zeta(\lambda) = \frac{1}{2}[dc(\lambda) - ids(\lambda)] \tag{53}$$

y según (51) y (52) tendremos:

$$\overline{d\zeta(\lambda)} = \frac{1}{2}[dc(\lambda) + ids(\lambda)] \tag{54}$$

pero por (51) $\cos \lambda t \cdot dc(\lambda)$ y $\operatorname{sen} \lambda t \cdot ds(\lambda)$ son funciones pares y sen $d\lambda t \cdot dc(\lambda)$ y cos $\lambda t \cdot ds(\lambda)$ son funciones impares.

Luego en (47) sustituyendo $e^{i\lambda t}$ por la fórmula:

$$\xi(t) \stackrel{2}{=} \int_{-\infty}^{\infty} [\cos \lambda t + i \operatorname{sen} \lambda t]\, d\zeta(\lambda) =$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} [\cos \lambda t d c(\lambda) + \operatorname{sen} \lambda t d s(\lambda)] = \int_0^\infty \cos \lambda t d c(\lambda) + \operatorname{sen} \lambda t d s(\lambda)$$

Falta probar que $c(\lambda)$ y $s(\lambda)$ son procesos de incrementos mutuamente ortogonales y de igual varianza.

De las (53) y (54) deducimos:

$$dc(\lambda)=d\zeta(\lambda)+\overline{d\zeta(\lambda)} \tag{55 a}$$

$$ds(\lambda)=i[d\zeta(\lambda)-\overline{d\zeta(\lambda)}] \tag{55 b}$$

y en consecuencia:

$$E/dc(\lambda)/^2=E/d\zeta(\lambda)+\overline{d\zeta(\lambda)}/^2\}=$$

$$=E\{/d\zeta(\lambda)/^2\}+E\overline{d\zeta(\lambda)}/^2=2E/d\zeta(\lambda)/^2=dG(\lambda) \tag{56 a}$$

$$Edc(\lambda)\overline{d\zeta(\lambda')}=Eds(\lambda)d\overline{s(\lambda')}=Edc(\lambda)\overline{ds(\lambda)}=0 \tag{56 b}$$

y

$$E\{/ds(\lambda)/^2=E\{/dc(\lambda)/^2\}=dG(\lambda)=2dF(\lambda) \tag{57}$$

$\lambda>0$ c.h.d.

10. La representación espectral del proceso estacionario con media cero $\{\xi_n,\ n\in Z\}$ es:

$$\xi_n\overset{2}{=}\int_{-\pi}^{\pi}e^{i\lambda n}d\zeta(\lambda) \qquad n\in Z \tag{58}$$

siendo $\zeta(\lambda)$ el proceso de incrementos ortogonales asociado teniendo las propiedades conocidas. Prescindimos de la demostración, porque sería repetir el parágrafo 4, excepto límites.

11. Igualmente para el caso real del proceso estacionario $\{\xi_m\ n\in Z\}$ la representación espectral es:

$$\xi_n\overset{2}{=}\int_{0}^{\pi}[\cos\lambda n dc(\lambda)+\operatorname{sen}\lambda n ds(\lambda)] \qquad n\in Z \tag{59}$$

con las propiedades establecidas en (9) para los procesos de incrementos ortogonales asociadas $C(\lambda)$ y $S(\lambda)$ indicados en (55) y (57).

12. *Continuidad y discontinuidad*

Concluiremos esta Sección y el artículo analizando la función de distribución espectral. Si $F(\lambda)$ fuera discontinua es porque $\zeta(\lambda)$ tendría discontinuidad y la varianza de las variables aleatorias $\zeta(\lambda\pm0)$ serán diferentes, eligiendo la continuidad a la derecha: $\zeta(\lambda)=\zeta(\lambda+0)$, entonces $d\zeta(\lambda)=$ $=\zeta(\lambda)-\zeta(\lambda-0)$ y la varianza de esta variable aleatoria es:

$$E\{/d\zeta(\lambda)/^2\}=F(\lambda)-F(\lambda-0) \tag{60}$$

Para el punto λ_j existirá salto aleatorio determinando el límite:

$$\xi_j=\zeta(\lambda_j)-\zeta(\lambda_j-0)=\underset{T\to\infty}{\text{L.e.m.}}\frac{1}{2T}\int_{-T}^{T} e^{-i\lambda_j t}\xi(t)dt \tag{61}$$

Igualmente de la función de distribución espectral podremos determinar los saltos:

$$F(\lambda_j)-F(\lambda_j-0)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} e^{-i\lambda_j t}B(t)dt>0 \tag{62}$$

Los puntos λ_j forman (a lo máximo) una sucesión numerable y excluidos estos de $F(\lambda)$, los restantes puntos, el límite será cero; es decir: son puntos de continuidad.

El proceso $\{\xi(t),\quad t\in R\}$ puede representarse por una descomposición ortogonal (I).

$$\xi(t)=\xi_d(t)+\xi_c(t) \tag{63}$$

siendo

$$\xi_d(t)=\sum_{j=1}^{m} e^{i\lambda_j t}\xi_j \qquad \text{(proceso con dist. esp. discreta)} \tag{64 a}$$

y

$$\xi_c(t)=\int_{-\infty}^{\infty} e^{i\lambda t}d_c\zeta(\lambda) \qquad \text{(proceso con dist. esp. continua)} \tag{64 b}$$

en cuyo caso $\xi_c(t)$ es continua en todo punto.

Si las funciones de distribución de los procesos (64) son F_d y F_c la del proceso (63) es:

$$F(\lambda)=F_d(\lambda)+F_c(\lambda) \tag{65}$$

siendo ξ_j y $d_c\zeta(\lambda)$ variables mutuamente ortogonales y tales que:

$$dF_d(\lambda)=\begin{matrix}\sigma_j^2 & \lambda=\lambda_j\\ 0 & \text{en otro caso.}\end{matrix}$$

$$dF_c(\lambda)=E\{/dF_c(\lambda)/^2\}\to 0$$

(I) H. Wold.

y en consecuencia:

$$dF(\lambda) = \begin{matrix} \sigma_j^2 & \lambda = \lambda_j \\ dF_c(\lambda) & \lambda \neq \lambda_j \end{matrix} \tag{66}$$

En esta descomposición el proceso $\xi_d(t)$ es el único que tiene frecuencias propias y $\xi_c(t)$ es continuo en todo punto t y no posee frecuencias propias (II).

Las frecuencias propias son los puntos de discontinuidad de la función de distribución según hemos indicado.

(II) Por citar autores citaremos a: Kart Karhunen: «Métodos lineales en el cálculo de probabilidades», C. S. I. Científicas, 1953, págs. 76 y siguientes, que añade la componente estacionaria $\xi_0(t)$ de «clase nula» que no utilizamos en nuestro trabajo.

APENDICE

1. **Clase:** $L_p(a, b)$

Definimos (I) clase $L_p(a, b)$ al conjunto de funciones $f(\cdot) \in L_p(a, b)$ medibles (L) para las que

$$\int_a^b /f(x)/^p\, dx < \infty$$

Si $p=1$ tenemos la clase $L_1(a, b)$ o simplemente $L(a, b)$ sin subíndice. Si $p=2$ obtenemos la clase de funciones de cuadrado integrable (I).

Si $x \in R$ escribimos sencillamente L_p.

2. **Condiciones de Dirichlet**

Diremos que una función $f(\cdot)$ definida en (a, b) satisface las condiciones de Dirichlet (I) cuando:

1.ª La función $f(\cdot)$ tiene un número finito de discontinuidades en (a, b).

2.ª La función $f(\cdot)$ tiene un número finito de máximos y de mínimos en (a, b); y

3.ª La función $f(\cdot)$ pertenece a la clase $L(a, b)$ (I).

Sencillamente: Toda función que cumpla estas condiciones puede descomponerse en arcos monótonos (continuos o discontinuos) completando con segmentos paralelos al eje de las x y hemos formado funciones monótonas que difieren de una constante y son integrables.

3. **Desarrollos en serie de Fourier**

Toda función $f(\cdot)$ que cumpla las condiciones de Dirichlet en el intervalo $(-\pi \leqslant x \leqslant \pi)$ puede desarrollarse en serie:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos x + b_n \operatorname{sen} nx)$$

(I) Véase Capítulo II, Sección 7.ª

siendo

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx \qquad n=0,\ 1,\ 2,\ \ldots$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx dx \qquad n=1,\ 2,\ 3,\ \ldots$$

La función $f(\cdot)$ para que sea desarrollable en serie de Fourier, si x fuese un punto de discontinuidad se demuestra que el valor del desarrollo es:

$$\frac{f(x+0)+f(x-0)}{2}$$

4. Transformadas de Fourier

Si la función $f(\cdot)\in L$ cumple las condiciones de Dirichlet en todo intervalo

$$(-a\leqslant x\leqslant +a) \quad \text{si} \quad a\to\infty$$

se denomina transformada inversa de Fourier a la integral:

$$g(\lambda)=\int_{-\infty}^{\infty} e^{-i\lambda x} f(x)dx$$

Si desarrollamos en serie $f(\cdot)$ (ver Capítulo Primero, Sec. 9.ª) y sustituimos los coeficientes a_n, b_n de (2) en la (1) en las fórmulas allí expuestas y hacemos que $a\to\infty$ llegamos fácilmente:

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u)e^{i\lambda(x-u)}du\,d\lambda$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\lambda x}\left[\int_{-\infty}^{\infty} f(u)e^{-i\lambda u}du\right]d\lambda=$$

y por la anterior tenemos la transformada directa de Fourier.

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\lambda x}g(\lambda)d\lambda$$

que nos recuerda las f. c. salvo constantes (I).

(I) Tichmarsh: «Integrales de Fourier», cap. I.