

# Modelos estocásticos dinámicos en las ciencias experimentales \*

Dynamical stochastic models in experimental sciences

MARIANO J. VALDERRAMA BONNET

Departamento de Estadística e Investigación Operativa. Facultad de Farmacia. Universidad de Granada. 18071 Granada. España.

## RESUMEN

Para que un modelo matemático represente de forma adecuada un fenómeno natural de carácter dinámico, debe recoger la aleatoriedad inherente al mismo, así como la variable de temporalidad. Tras repasar procedimientos alternativos en la formulación de modelos, en el presente trabajo proponemos una metodología para la elaboración de un modelo predictivo de naturaleza estocástica, a partir de observaciones muestrales temporales, basado en el Análisis de Componentes Principales del proceso objeto de estudio. El modelo desarrollado se aplica a predecir la tasa de incidencia de SIDA en países europeos.

**Palabras clave:** Ecuación diferencial, proceso estocástico, componente principal, predicción dinámica.

## ABSTRACT

For a mathematical model to be a suitable representation of a dynamical phenomenon in Nature, it must include its random characteristics as well as the time variable. After reviewing alternative procedures in model formulation, this paper presents a methodology for elaborating a predictive stochastic model from sample observations depending upon the time, on the basis of the Principal Component Analysis of the studied stochastic process. The developed model is then applied to forecast the AIDS incidence rate in European countries.

**Key words:** Differential equation, stochastic process, principal component, dynamical forecasting.

Recibido: 8-3-1994.

Aceptado: 15-3-1994.

BIBLID [0004-2927(1994) 35:2; 277-287]

---

\* Este artículo recoge el contenido de la conferencia impartida por el autor el día 8 de diciembre de 1993, en la Facultad de Farmacia de Granada con motivo del acto académico celebrado por su Patrona, La Inmaculada Concepción.

## 1. INTRODUCCIÓN AL PROBLEMA DE FORMULACIÓN DE MODELOS

Una de las cuestiones prioritarias para el investigador experimental ha sido siempre el buscar las relaciones de causalidad que rigen un cierto fenómeno, de manera que, una vez localizadas, pueda éste interpretarse de forma rigurosa y su desenvolvimiento futuro pueda ser parcialmente pronosticado. Así, el objetivo de la modelización matemática es formular ecuaciones que permitan describir la conexión entre causa y efecto del suceso que se observa.

Cabe distinguir dos enfoques alternativos, aunque complementarios, a la hora de formular modelos: El enfoque determinista y el estocástico. El primero de ellos, usual en la Física y en las Matemáticas, plantea sistemas de ecuaciones que recogen como variables independientes aquéllas que causan o explican el fenómeno en estudio, y como variables dependientes o de respuesta las que miden su efecto. Evidentemente, mientras más realista se pretende sea modelo, más difícil será su formulación. Sin embargo, dado que las leyes que regulan la Naturaleza son extremadamente complejas, nunca podrán integrarse en el mismo la totalidad de causas de variabilidad que intervienen, por lo que siempre quedará una parte no explicada por el modelo, siendo esta la principal limitación del enfoque determinista.

Por el contrario, el enfoque estocástico da un paso más, tratando de medir la incertidumbre del fenómeno mediante modelización de la parte no explicada por el esquema determinista, disminuyendo así la rigidez del modelo planteado.

En el presente trabajo, concebido basicamente de un punto de vista divulgativo, se presentan ambos enfoques en su vertiente dinámica, es decir la modelización de fenómenos que evolucionan en el tiempo, acompañados de numerosas aplicaciones, y se presenta finalmente un procedimiento de construcción de modelos estocásticos dinámicos en base a la información muestral, que utiliza como herramienta estadística básica el análisis de componentes principales de un proceso estocástico de segundo orden. Así mismo, se presenta una aplicación epidemiológica con datos de Sida de países europeos.

## 2. ELABORACIÓN DE UN MODELO DETERMINISTA DINÁMICO

Existen dos formas básicas de plantear un modelo determinista, también llamado modelo matemático en sentido amplio:

- 1.<sup>a</sup> Si se tiene conocimiento de la ley teórica que rige el fenómeno que se estudia, se planteará una ecuación diferencial o un sistema de ecuaciones diferenciales, acompañado de ciertas condiciones iniciales deducidas a partir de la observación del mismo.
- 2.<sup>a</sup> Si la información disponible es solo muestral, es decir procedente de

observaciones, se ajustará un modelo mediante alguna técnica numérica: interpolación, ajuste por mínimos cuadrados, ...

Veamos a continuación algunos ejemplos de la primera vía de modelizar, esto es, mediante planteamiento y resolución de ecuaciones diferenciales:

### Ejemplo 1.

En un cultivo de bacterias, la velocidad de crecimiento es directamente proporcional al número de bacterias presentes en cada instante, habiéndose observado que dicho número se duplica al cabo de cuatro horas.

Si denotamos entonces por  $y(t)$  el número de bacterias en el instante  $t$ , la velocidad instantánea de crecimiento en  $t$  verificará, a partir de las hipótesis iniciales, la siguiente ecuación diferencial de variables separables:  $y'(t) = k \cdot y(t)$ , siendo  $k$  una constante positiva. La resolución de esta ecuación viene dada por  $y(t) = y(0) \cdot e^{k \cdot t}$ , donde  $y(0)$  denota el número inicial de bacterias (en el instante cero). Aplicando ahora la condición inicial de duplicación en cuatro horas:  $y(4) = 2 \cdot y(0)$ , concluimos que el modelo de crecimiento para esta población bacteriana es:  $y(t) = 2^{t/4} y(0)$ .

### Ejemplo 2.

Una sustancia B se forma a partir de otra sustancia A con velocidad de reacción  $v_1 = 8.3$  y, a su vez, B se transforma en C con velocidad  $v_2 = 27.5$ . Se parte de una concentración inicial de A igual a 5 mientras que de B y C es 0, y del hecho de que la velocidad de transformación de sustancia es proporcional a la cantidad presente.

Con objeto de deducir las ecuaciones secuenciales de transformación de A en B y de B en C, denotemos por  $[A]_t$  y  $[B]_t$  las concentraciones de A y B respectivamente en el instante  $t$ . Así podemos escribir para la primera transformación la siguiente ecuación:  $[A]_t' = -8.3[A]_t$ , cuya solución, tras aplicar la condición inicial  $[A]_0 = 5$ , viene dada por:  $[A]_t = 5 e^{-8.3 t}$ . Para la segunda reacción, podemos plantear la siguiente ecuación lineal de primer orden:  $[B]_t' = 8.3[A]_t - 27.5[B]_t$ , o equivalentemente  $[B]_t' + 27.5[B]_t = 41.5 e^{-8.3 t}$ , que integrada, junto a la condición  $[B]_0 = 0$ , nos proporciona la segunda ecuación de transformación:  $[B]_t = 2.16 e^{-27.5 t} (e^{19.2t} - 1)$ .

A continuación desarrollamos otros dos ejemplos sobre formulación de modelos mediante ajuste matemático a partir de series de datos experimentales:

### Ejemplo 3.

A presión de 1 atmósfera, un volumen de agua disuelve las siguientes cantidades de  $\text{CO}_3\text{H}^- + \text{H}^+$ :

t	0	5	10	15
V	1.80	1.45	1.18	1.00

Se pretende construir, a partir de dichos datos, un modelo de tipo polinómico que represente la relación existente entre volumen y temperatura. Para ello ajustaremos una parábola de segundo grado  $\phi(t) = a + bt + ct^2$ , aplicando la técnica de mínimos cuadrados, que consiste en minimizar la suma de los cuadrados de las diferencias entre los valores observados de V y los teóricos a través de dicha función de ajuste, es decir  $\sum_i [V_i - \phi(t_i)]^2$ . El resultado de dicho ajuste nos proporciona el siguiente modelo explicativo:  $V = 1.8005 - 0.0789t + 0.0017t^2$ , siendo el error típico asociado igual a 0.0011.

Podría intentarse un ajuste mediante polinomios de grado superior a dos, siendo necesario entonces contrastar, mediante un test de hipótesis estadísticas, la significación de los coeficientes asociados a términos de grado elevado. Cabría también ajustar otro tipo de funciones no polinómicas, indicando el error típico cuál es el ajuste más preciso.

**Ejemplo 4.**

En Ecología se utiliza el muestreo de áreas contiguas para contar el número de especies distintas de plantas por área, que consiste en que cada área contigua tiene doble extensión a la anterior. Si se comienza por una superficie de 1 m<sup>2</sup> y se obtienen los siguientes tipos de especies por área:

S (en m <sup>2</sup> )	1	2	4	8	16	32	64
N (n.º esp.)	2	4	7	11	16	19	21

se pretende decidir cuál es el más adecuado para representar la relación entre número de especies y área entre los tres modelos siguientes: a) un modelo semilogarítmico  $N = a + b \cdot \ln S$ , b) un modelo exponencial  $N = a \cdot e^{bS}$ , y c) un modelo potencial  $N = a \cdot S^b$ .

Mediante ajuste mínimo cuadrático obtenemos las siguientes ecuaciones y errores típicos  $\hat{\alpha}_s$  para los tres modelos anteriores:

a)  $N = 1.13 + 4.95 \ln S$ ,      b)  $N = 5.26 e^{0.028 S}$ ,      c)  $N = 2.69 S^{0.57}$   
 $\epsilon_s = 0.095$                                        $\epsilon_s = 0.556$                                        $\epsilon_s = 0.210$

Así, claramente elegimos el modelo semilogarítmico a) para representar dicha relación.

### 3. MODELOS ESTOCÁSTICOS DINÁMICOS: PROCESOS ESTOCÁSTICOS

La palabra estocástico procede del vocablo griego  $\sigma\tau\omicron\chi\acute{\alpha}\zeta\omega$  (que se pronuncia stojatso) y que significa "formulación de modelos mediante hipótesis". Por otra parte, el término proceso, según el diccionario Espasa de Lengua Española, se define como "conjunto de fases sucesivas de un fenómeno natural o de una operación artificial". De tal forma, al hablar de procesos estocásticos nos referimos a la elaboración de modelos que evolucionan en el tiempo y cuya naturaleza es incierta, en el sentido de que las hipótesis sobre las que se construyen no son completamente fijas sino que están afectadas de un cierto grado de variabilidad o incertidumbre.

Desde un punto de vista más matemático, el concepto de estocástico es sinónimo de aleatoriedad o azar, y de hecho podemos definir los Procesos Estocásticos como la parte dinámica de la teoría de la Probabilidad, en cuanto que ésta se centra en el estudio de variables que representan un cierto fenómeno en un instante dado, mientras que los procesos estocásticos formulan leyes sobre situaciones aleatorias en movimiento.

Sin embargo, conceptualmente no está del todo definida la barrera entre ambas teorías. Si consideramos un fenómeno dinámico que evoluciona en los instantes  $t=1,2,\dots,N$ , el vector aleatorio obtenido  $(X_1, X_2, \dots, X_N)$  puede considerarse por un lado como una variable aleatoria  $N$ -dimensional, y por otro lado como un proceso estocástico en tiempo finito. La existencia de correlaciones entre las variables  $X_i$  no influye en la clasificación del mismo en una u otra categoría.

Si el número de variables que componen el vector aleatorio fuese infinito sí podríamos considerar a éste como un verdadero proceso estocástico. Sin embargo, también se ha desarrollado una teoría consistente sobre variables aleatorias de dimensión finita. Lo que sí parece claro es que si el conjunto de variables aleatorias consideradas fuese infinito no numerable, las herramientas del Cálculo de Probabilidades resultarían insuficientes para su tratamiento, y el problema encajaría dentro de la teoría de Procesos Estocásticos.

En cualquier caso, no es objetivo nuestro entrar en disquisiciones taxonómicas conceptuales, y consideramos que un mismo problema puede ser abordado desde distintas facetas. Nos limitaremos a considerar un proceso estocástico como una familia uniparamétrica de variables aleatorias  $\{X(t,w), t \in T\}$  definidas todas ellas sobre un mismo espacio probabilístico, y que toman valores en un conjunto denominado espacio de estados o espacio fásico. Generalmente, el parámetro  $t$  es el tiempo.

Si el espacio paramétrico es de dimensión mayor que uno se dice que el proceso es un campo aleatorio; y si las variables que constituyen el proceso son vectoriales, se dice que el proceso es vector-valorado.

La clasificación de los procesos estocásticos puede realizarse de diversas

formas. Una primera forma es atendiendo al carácter del espacio de estados y del espacio paramétrico, pudiendo presentarse cuatro modalidades de procesos, según éstos sean discretos y/o continuos. Una segunda clasificación considera que los procesos pueden ser estacionarios o evolutivos, según se mantenga o no su distribución en el tiempo. Pero quizás, el procedimiento más completo es el propuesto por Karlin, que se basa en las relaciones de dependencia existentes entre las variables del proceso; así, se obtienen las siguientes clases de procesos:

- \* Procesos con incrementos independientes: las variables de incremento temporal sobre intervalos de tiempo disjuntos son independientes.
- \* Procesos de Markov: la distribución de cada variable del proceso depende sólo del pasado inmediato.
- \* Martingalas: la esperanza de una variable del proceso dado el pasado hasta un cierto instante es igual a la variable en dicho instante.
- \* Procesos estacionarios: la distribución de cualquier vector de variables del proceso se mantiene si éstas se desplazan en el tiempo de forma constante.

Existen, no obstante, otras muchas formas de clasificar los procesos estocásticos.

Históricamente, podemos considerar que el punto de partida en el estudio de procesos estocásticos se debe al botánico inglés Robert Brown, quien en 1827 observó que los granos de polen en suspensión acuosa mostraban un movimiento continuo y caótico en todas las direcciones. Este comportamiento era, en una escala grande, exactamente lo que cabría esperar de una molécula si la teoría cinética fuese correcta. Dicho desplazamiento errático, denominado **movimiento Browniano**, y que inicialmente se atribuyó al hecho de que las partículas tenían vitalidad propia, se debe al bombardeo continuo al que están sometidas las partículas en suspensión por parte de las moléculas del medio que las rodea, ya sea líquido o gaseoso.

Posteriormente, Wiener (1863), William & Ord (1879), Ramsey (1879) y Einstein (1905) trataron de buscar una explicación más científica para este fenómeno. Asimismo, la evolución histórica de la teoría dinámica del movimiento Browniano está perfectamente recogida en la monografía de Nelson (1967).

El físico francés Perrin (1909) supuso que la energía cinética del movimiento Browniano de una partícula microscópica debe coincidir con la de una molécula, y preparó un experimento consistente en una suspensión de gomaguta y almáciga, obteniendo partículas esféricas de magnitud uniforme de  $13 \times 10^{-4}$  cm de diámetro, mediante un proceso de centrifugación fraccionada. Así, si denotamos por  $m$  y  $r$ , respectivamente, a la masa y radio de una partícula, por  $d$  y  $d'$  las densidades de la partícula y del medio, y por  $n_0$  y  $n$  el número de partículas existentes a niveles separados por una distancia vertical  $h$ , se obtiene la ecuación:

$$\frac{RT}{N} \ln \frac{n_0}{n} = \frac{4}{3} \pi r^3 gh(d - d')$$

A partir de esta ecuación, Perrin y sus colaboradores, aplicando la ley de Stokes, lograron determinar por varios métodos el número de Avogadro N a partir del desplazamiento lineal medio  $\bar{x}$  de una partícula en movimiento Browniano:

$$N = \frac{RTt}{3\pi \nu r \mu^2}$$

siendo t el tiempo desde el origen y  $\nu$  la viscosidad del medio. Así mismo, encontraron que los desplazamientos reales se ajustaban a una distribución de Maxwell, y con objeto de obtener un valor medio preciso era necesario efectuar un gran número de medidas sucesivas. Las investigaciones de Juan Perin fueron recompensadas en 1926 con el Premio Nobel de Física.

Sin embargo, la formulación matemática del movimiento Browniano se debe principalmente a Wiener (1923) quien lo definió como un proceso estocástico centrado con incrementos independientes estacionarios Gaussianos. Posteriormente, Barnes & Silverman (1934), Doob (1942), Kac (1947) y Lévy (1948) completaron su estudio riguroso. Los trabajos realizados por Uhlenbeck & Ornstein (1930) y Wang & Uhlenbeck (1945) trataron el movimiento Browniano desde el punto de vista de las difusiones.

Paralelamente al estudio del movimiento Browniano se fue desarrollando una teoría general de procesos estocásticos, abarcando tanto el estudio de propiedades estructurales básicas, como el tratamiento de clases concretas.

Actualmente, los procesos estocásticos se utilizan como modelo probabilístico de fenómenos de naturaleza muy diversa, siendo usual encontrar aplicaciones en campos tales como la Medicina, Economía, Psicología, etc., y existiendo grupos de investigación en los que participan especialistas de diversas áreas.

#### 4. ELABORACIÓN DE UN MODELO ESTOCÁSTICO DINÁMICO

Sea  $\{X(t), 0 \leq t \leq T\}$  un proceso estocástico de segundo orden, es decir con momentos de segundo orden finitos, que supondremos continuo en media cuadrática y cuyas trayectorias son funciones del espacio  $L^2 [0, T]$ . El Análisis de Componentes Principales (ACP) del proceso permite descomponerlo como una serie de funciones ortogonales ponderadas por variables aleatorias incorreladas. Estas funciones  $f_i(t)$ , denominadas factores principales, son las funciones propias del operador de covarianza, de manera que

$$\lambda_i f_i(t) = \int_0^T R(t,s) f_i(s) ds \quad (1)$$

siendo el núcleo  $R(t,s)$  de esta ecuación tipo Fredholm la función de covarianza del proceso. Los estimadores de los factores principales mediante una muestra de  $N$  trayectorias independientes del proceso son las funciones propias asociadas al operador de covarianza muestral. La descomposición del proceso vendrá dada por (Todorovic, 1992):

$$X(t) = \mu(t) + \sum_{i=1}^{\infty} f_i(t) \cdot b_i \quad (2)$$

donde  $\mu(t)$  es la función media y  $\{b_i, i=1,2,\dots\}$  una sucesión de variables aleatorias incorreladas que pueden expresarse como una cierta integral del propio proceso.

En la práctica, resolver una ecuación integral del tipo (1) puede llegar a ser muy complicado, siendo posible obtener soluciones explícitas solo para núcleos muy concretos. Este problema se complica aún más si tenemos en cuenta que, en la práctica, la covarianza muestral no puede calcularse analíticamente en términos de funciones elementales para todo punto de  $[0,T] \times [0,T]$ . Para solventar estos problemas, hemos desarrollado dos métodos eficientes para la aproximación de los factores principales muestrales:

- \* **Método del trapecio** (Aguilera *et al.*, 1992): Consiste en resolver la ecuación (1) a nivel muestral mediante la fórmula de cuadratura compuesta del trapecio, eligiendo convenientemente los nodos de la partición del intervalo  $[0,T]$ .
- \* **Método de proyección ortogonal** (Aguilera *et al.*, 1995a): Consiste en elegir la mejor aproximación de los factores principales muestrales en un subespacio finito-dimensional de  $L^2[0,T]$  generado por un sistema ortonormal, cuya elección en cada caso particular dependerá del carácter de la trayectorias observadas.

En la práctica aparece la dificultad adicional de que normalmente las trayectorias son conocidas solo en un conjunto discreto de instantes del intervalo, aunque tengan carácter continuo. Proponemos entonces interpolar las trayectorias mediante B-splines cúbicos (Aguilera *et al.*, 1995b), generalizando estudios realizados por otros autores.

Finalmente, una vez aproximados los estimadores en el ACP del proceso, proponemos extender la técnica de regresión mediante componentes principales, con objeto de resolver el problema de predicción lineal de la variable  $X(T+h)$ , con  $h>0$ , en función de las variables del proceso, observadas en el intervalo

[0,T]. El estimador lineal mínimo cuadrático de  $X(T+h)$  dado  $\{X(t), 0 \leq t \leq T\}$  viene dado por la expresión:

$$\hat{X}(T+h) = \mu(T+h) + \sum_{i=1}^{\infty} E\{[X(T+h) - \mu(T+h)] b_i\} \frac{b_i}{\lambda_i}$$

donde la serie infinita en la expresión (3) converge en media cuadrática. Así, proponemos el siguiente modelo predictivo

$$\hat{X}^m(T+h) = \bar{X}(T+h) + \sum_{i=1}^{\infty} S_{i, T+h} \frac{\hat{b}_i}{\lambda_i}$$

siendo  $\bar{X}$  la media muestral del proceso en el instante considerado y  $S_{i, T+h}$  la covarianza muestral entre  $b_i$  y  $X(T+h)$ . Más aún, incluiremos en el modelo solo aquellas componentes principales que tengan mayor grado de correlación con  $X(T+h)$  y mayor varianza.

## 5. APLICACIÓN A LA MODELIZACIÓN Y PREDICCIÓN DE TASAS DE INCIDENCIA DE SIDA EN EUROPA

En este apartado vamos a construir un modelo estocástico del tipo (2) para la tasa de incidencia (IR) de SIDA en países europeos, correspondiente al periodo 1985-1991, a partir de datos publicados anualmente por la O.M.S. Mediante muestreo aleatorio hemos elegido 8 países del total de 39 en Europa, y se ha realizado el ACP a partir de las series temporales que figuran en la Tabla 1 en la misma fila que el nombre del país muestreado. Se ha observado que incluyendo dos componentes principales, el modelo explica un 95% de la variabilidad total del proceso estudiado, mientras que con tres componentes, la varianza explicada supera el 99%.

Las IR estimadas al 99% y al 95% figuran así mismo en la Tabla 1 para cada uno de los países. Puede observarse que la reconstrucción de las trayectorias de IR al 99%, en general, es más precisa que al 95%, especialmente a medida que se consideran instantes cada vez más avanzados, tras un periodo inicial de estabilización.

En segundo lugar, la Tabla 2 muestra la predicción de IR para 1992 en otros cuatro países no muestreados anteriormente, habiéndose obtenido éstas mediante regresión sobre una y dos componentes principales respectivamente. Es evidente, a partir de estos resultados, que ambas predicciones son fiables,

especialmente la realizada con el modelo de regresión que incluye dos componentes.

Tabla 1.—Tasa de incidencia (IR) de SIDA reales y estimadas mediante el modelo (2) al 99 % y 95 % de varianza explicada respectivamente

	1985	1986	1987	1988	1989	1990	1991
DINAMARCA	7.5	13.5	19.6	24.7	33.9	38.3	40.2
IR al 99%	6.2	12.8	19.5	24.2	35.5	37.6	40.8
IR al 95%	4.0	8.3	14.9	24.3	32.9	40.5	43.3
FRANCIA	10.5	22.6	40.0	54.2	66.6	73.4	77.1
IR al 99%	10.5	22.3	38.1	53.1	70.1	70.4	81.7
IR al 95%	6.2	13.6	29.1	53.3	65.1	76.5	86.6
GRECIA	0.7	2.2	5.3	8.2	10.7	13.3	14.6
IR al 99%	1.8	3.4	4.6	8.6	9.4	13.7	15.3
IR al 95%	2.0	3.9	5.0	8.6	9.7	13.4	15.1
IRLANDA	1.4	1.7	5.7	10.6	14.2	15.7	17.7
IR al 99%	1.8	3.6	5.5	11.1	11.7	16.5	18.9
IR al 95%	2.2	4.2	6.1	11.1	12.1	16.1	18.5
ITALIA	3.4	7.9	17.7	30.6	42.2	53.0	63.4
IR al 99%	3.6	7.9	16.9	31.8	42.2	54.6	58.0
IR al 95%	4.8	10.3	19.4	31.8	43.6	52.9	56.6
LUXEMBURGO	5.0	7.5	7.5	10.0	29.2	23.7	30.0
IR al 99%	4.1	8.4	11.3	11.1	22.7	27.8	26.1
IR al 95%	3.4	7.0	9.9	11.1	21.9	28.8	26.9
HOLANDA	4.6	9.4	16.6	22.0	26.2	27.6	29.6
IR al 99%	5.0	10.2	15.7	21.7	26.5	26.7	32.3
IR al 95%	3.0	6.2	11.4	21.8	24.1	29.5	34.6
ESPAÑA	4.3	11.9	25.8	53.5	73.3	84.8	99.1
IR al 99%	5.2	11.8	27.9	54.4	69.8	87.6	94.9
IR al 95%	6.7	15.3	31.7	54.3	71.9	85.1	92.9

Tabla 2.—IR real de 1992 y estimaciones de la misma mediante regresión sobre una y dos componentes principales

	IR de 1992	IR estimada para 1992 con 1 comp. principal	IR estimada para 1992 con 2 comp. principales
AUSTRIA	23.9	25.0	25.6
BELGICA	26.0	24.9	25.4
FINLANDIA	6.2	2.8	5.5
SUIZA	90.3	108.7	103.4

## REFERENCIAS

- AGUILERA, A. M., VALDERRAMA, M. J. & DEL MORAL, M. J. (1992): "Un método para la aproximación de estimadores en ACP. Aplicación al proceso de Ornstein-Uhlenbeck". *Rev. Soc. Chilena de Estad.*, 9(2); 57-77.
- AGUILERA, A. M., GUTIÉRREZ, R. & VALDERRAMA, M. J. (1995a): "Computational approaches to estimation in the PCA of a stochastic process". *Appl. Stoch. Models & Data Anal*, en prensa.
- (1995b): "Approximation of estimators in the PCA of a stochastic process using B-splines". Sometido a *Comm. Stat.*
- BARNES, R. B. & SILVERMAN, S. (1934): "Brownian motion as a natural limit to all measuring processes". *Rev. Mod. Phys.*, 6; 162-192.
- BROWN, R. (1928): "A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies". *Philosophical Magazine N. S.*, 4; 161-173.
- DOOB, J. L. (1942): "The Brownian movement and stochastic equations". *Ann. Math.*, 43; 351-369.
- EINSTEIN, A. (1906): "Zur theorie der Brownschen Bewegung". *Ann. Phys.*, IV 19; 371-381.
- KAC, M. (1947): "Random walk and the theory of Brownian motion". *Amer. Math. Monthly*, 54; 369.
- LÉVY, P. (1948): *Processus Stochastiques et Mouvement Brownien*. Gauthier-Villars, Paris.
- NELSON, E. (1967): *Dynamical Theories of Brownian Motion*. Mathematical Notes, Princeton Univ. Press.
- PERRIN, J. (1909): "Brownian movement and molecular reality". *Annales de Chimie et de Physique*, 8<sup>me</sup> series.
- TODOROVIC, P. (1992): *An Introduction to Stochastic Processes and their Applications*. Springer-Verlag, New York.
- UHLENBECK, G. E. & ORNSTEIN, L. S. (1930): "On the theory of brownian motion". *Phys. Rev.*, 36; 823-841.
- VALDERRAMA, M. J. (1989): *Métodos Matemáticos en las Ciencias Experimentales*. Pirámide, Madrid.
- WANG, M. C. & UHLENBECK, G. E. (1945): "On the theory of brownian motion II". *Rev. Mod. Phys.*, 17; 323-342.
- WIENER, N. (1923): "Differential space". *J. Math. Phys. Mass. Inst. Tech.*, 2; 131-174.