Rev.R.Acad.Cienc.Exact.Fís.Nat. (Esp)
Vol. 101, Nº. 1, pp 35-58, 2007
VII Programa de Promoción de la Cultura Científica y Tecnológica

EL AZAR EN FÍSICA Y BIOLOGÍA Y LAS MATEMÁTICAS DEL AZAR

DARIO MARAVALL CASESNOVES *

* Real Academia de Ciencias Exactas, Físicas y Naturales. Valverde 22. 28004 Madrid.

El Cálculo de Probabilidades es una de las grandes ramas del árbol de las Matemáticas, se le puede definir como la ciencia del azar y aunque esta definición no es exacta del todo, si es muy aproximada. El Cálculo de Probabilidades permite el tratamiento matemático y cuantitativo de las posibilidades de realización de un hecho o fenómeno, que todavía no ha sucedido pero que puede suceder; es útil para la previsión de ciertos hechos y fenómenos posibles, pero que no tienen por qué suceder necesariamente, como puede ser la predicción del tiempo meteorológico, predecir si va a haber una buena cosecha o no, si hay peligro de inundaciones o cualquier otra calamidad pública, etc. También es útil para aconsejarnos sobre las apuestas en un juego de azar, sobre las primas que se pueden pagar por un seguro de vida o de accidentes, el comportamiento económico de las personas, empresas o del Estado, etc.

El Cálculo de Probabilidades tiene dos campos de actuación, uno es el de las Matemáticas puras, y el otro es el de las Matemáticas aplicadas a la Física, la Biología, la Economía, la Investigación Operativa e incluso la Sociología. Es la base de la Estadística.

Su origen se puede fechar en el siglo XV, por tanto es anterior a la Geometría Analítica que nace con Descartes (siglo XVII) al Cálculo Infinitesimal que nace con Newton y Leibniz (siglos XVII y XVIII), pero es posterior a la Geometría como ciencia deductiva que se inicia con Tales de Mileto (siglo VII a C), y que alcanza su primera axiomatización con Euclides (siglos IV y III a.C.); a la Trigonometría que se inicia con Arquímedes (siglo II a C) como ciencia auxiliar de la Astronomía, y que en el siglo XIII se convierte en una ciencia independiente en Bagdad con Nasir al Diny; y al algebra debida al árabe Al Khwarizmi (siglos VIII y IX).

La palabra española azar que es muy parecida en otros idiomas como el italiano, el francés o el inglés, aunque en este último idioma también se usa una palabra distinta random, es muy probable que proceda de la palabra árabe azzahr, que según el literato Amin Maalouf en su libro “Las Cruzadas vistas por los árabes” era el nombre de un juego de dados del Egipto faraónico, al que jugaban los árabes y los turcos seldjucidas en el tiempo de las cruzadas y que allí aprendieron los franys (los cruzados) y lo trajeron a Europa.

Hay ciertas palabras que usamos en lenguaje coloquial, que han pasado al lenguaje científico, pero exacta y perfectamente definidas, porque es la única forma en que pueden utilizarse para que los Científicos se entiendan entre sí. Entre estas figuran las palabras contrapuestas: azar y necesidad, indeterminismo y determinismo, incertidumbre y certidumbre; las tres primeras de los tres pares de palabras anteriores, pertenecen al campo sobre el que opera la probabilidad. Como veremos en el desarrollo de la conferencia, el azar no es único sino múltiple, los mecanismos de actuación del azar son varios, y no tener esto en cuenta puede dar origen a paradojas consistentes en que un mismo problema tenga distintas soluciones, pero eso no es así, porque en el enunciado de un problema no se puede decir simplemente "escogido al azar" sin más, porque hay que enunciar el mecanismo según el cual actúa el azar en este problema, y entonces lo que aparenta ser un solo problema, se desdobla en varios, cada uno con su solución distinta.

Haremos un resumen de la Historia del Cálculo de Probabilidades, desde sus orígenes en el siglo XV hasta nuestros días pasando por la definición de probabilidad de Laplace (1812) hasta la axiomatización actualmente aceptada de Kolmogorov (1929-1933). Haremos mención especial del papel que los juegos de azar han desempeñado, así como de ciertas teorías matemáticas, tales como la teoría de conjuntos, las algebras de Boole, los integrales de Lebesgue y de Stieltjes y muy especialmente de las cadenas de Markov y sobre todo de los procesos estocásticos y del movimiento browniano. Todo ello dentro del campo de las Matemáticas puras.

En cuanto a las aplicaciones haremos un resumen histórico de la Física Estadística clásica de fines del siglo XIX, aún vigente y de la Física Estadística cuántica del siglo XX. Así como de la aplicación de los procesos estocásticos al estudio de los rayos cósmicos, descubiertos en el siglo XX, y a su aplicación a la Física Nuclear.

Haremos también un breve resumen de las aplicaciones del Cálculo de Probabilidades a la Biología en materias tales, como la Dinámica de Poblaciones, la Epidemiología y la Genética de Poblaciones, Teorías todas que salvo algunas escasas aportaciones del siglo XIX, caen de lleno dentro del siglo XX.

Mi aportación personal en este campo se extiende a medio siglo, comenzaron en los años cincuenta y mi última publicación aparecida es del 2004.

Aunque los fenómenos físicos y biológicos a los que se aplican los Procesos Estocásticos son de naturaleza muy distinta, sin embargo el planteamiento matemático y su resolución son iguales, y así por ejemplo conceptos como natalidad, mortalidad, inmigración o contagio tan específicos de la Biología han pasado a la Física.

Curiosamente las palabras aleatorio y estocástico de tan frecuente uso en la Teoría de la Probabilidad de nuestros días, son derivadas del latín y del griego. Aunque en el mundo grecorromano la probabilidad científica no existía, ya entonces e incluso en tiempos más antiguos los hombres sufrían o se beneficiaban de los efectos del azar, de la casualidad y de la suerte. La frase "alea jacta est" de Julio Cesar al cruzar el río Rubicón y marchar sobre Roma, clara alusión al azar y a la suerte, cambió el curso de la Historia.

Concluiremos con una breve referencia a mis investigaciones sobre un nuevo tipo de variables aleatorias que he denominado fractales, por estar definidas en conjuntos de "medida nula" y que no se ajustan a la axiomática de Kolmogorov. Son consecuencia de mi extensión de la Geometría de Masas, la Dinámica, la Integración y las Probabilidades a los modernos conjuntos fractales tan distintos de los clásicos. En estos conjuntos no existen las integrales de Riemann o Lebesgue (son nulas) pero si existe el cociente de dos integrales, como el valor de una función indeterminada (cero dividido por cero) cuyo cálculo es posible, eliminando su indeterminación mediante una generalización de la vieja regla de L'Hôpital.

En la Física, la Biología, la Economía, así como en la Meteorología y en otras actividades humanas, si se dan unas condiciones determinadas se presentan tres clases de sucesos que son:

1) ocurre necesariamente un suceso S

2) un suceso S no puede realizarse, es imposible.
   Estos dos primeros casos, caen en el terreno de la necesidad. El porvenir se dice que está determinado.

3) el suceso S puede o no realizarse
   Este caso se subdivide en dos.
   - 3.a es posible asignar al suceso S, que llamamos aleatorio, un número real positivo que mide el grado de posibilidad de realización del suceso, al que llamamos su probabilidad. El porvenir se dice que es probabilizable.
   - 3.b es imposible asignar una probabilidad al suceso, y entonces se dice que el provenir es indeterminado. Estos sucesos se dan en Economía y en Meteorología. Puede ocurrir que exista una probabilidad, pero que seamos incapaces de descubrirla, o puede que no exista tal probabilidad.

Se llama variable aleatoria, una variable que puede tomar diversos valores y tal que a cada valor se le pueda atribuir una probabilidad, las cuales forman la distribución de probabilidad de la variable aleatoria.

Un ejemplo sencillo es el de arrojar un dado, el suceso aleatorio que puede ocurrir es que salga desde un uno a un seis, y a cada uno de estos valores se le asigna la probabilidad un sexto, que expresa que existe la misma probabilidad (o posibilidad) de que salga uno cualquiera de dichos valores. Si arrojamos dos dados, entonces la variable aleatoria puede tomar un valor que va desde dos hasta doce, y a cada uno de estos valores se le puede asignar una probabilidad (calculable) distinta. Es mucho más difícil que salga 2 ó 12 que 6 ó 7.

La Ciencia que trata de las variables aleatorias y de sus probabilidades, es una parte de las Matemáticas llamada Cálculo de Probabilidades muy amplia. Comprende desde problemas y cuestiones muy elementales hasta otras de una grandísima dificultad, e incluso contiene problemas abiertos, todavía no resueltos. Por tanto los métodos matemáticos que emplea van desde unos muy sencillos hasta otros sumamente difíciles y complicados. Para resolver algunos de estos problemas basta con conocer las Matemáticas del Bachillerato, para comprender y resolver otros se precisan los conocimientos de un Ingeniero o de un Licenciado en Ciencias; el planteamiento y resolución de algunos problemas pueden ser la base de una tesis doctoral o de una investigación de muchísima altura. El Cálculo de Probabilidades es una ciencia viva y en pleno desarrollo, con un gran porvenir por delante.

El Cálculo de Probabilidades, como ya hemos dicho es muy reciente, no comienza hasta finales del siglo XV, mientras que la Geometría, la Astronomía matemática, el Algebra, la Trigonometría y la Mecánica son mucho más antiguas. Como ya dijimos la Geometría como ciencia es del siglo VII a C.

Los primeros problemas de Probabilidades, seguramente aparecen en un libro de Luca Pacioli (1445-1517) publicado en 1494, que lleva por título “Summa de arithmética y geometría proportioni y proportionalita”.

Poco después Cardano (1501-1576) se ocupó de probabilidades, es otro precursor y publicó en 1545 su libro “Ars Magna”, pero lo más importante de su obra fue la resolución de las ecuaciones de tercer y cuarto grado, que es debida a él y a Tartaglia (1499-1557). Cardano realizó un invento muy importante, que aun se usa en nuestros días, que es la junta de Cardan, un mecanismo que permite el desplazamiento angular en todas direcciones de dos árboles, cuyos ejes son concurrentes, mecanismo que permitió librar a la brújula de un barco, de sufrir los efectos de su movimiento.

Se consideran como cofundadores del Cálculo de Probabilidades a Fermat (1601-1665) y a Pascal (1623-1662) que mantuvieron entre si la más importante correspondencia en 1654, donde se puede decir que se encuentran los principios del Cálculo de Probabilidades como ciencia.

Huygens (1629-1695) conoció estas cartas y las utilizó junto a otras aportaciones suyas como fueron la introducción de la esperanza matemática y la resolución de algunos problemas de probabilidades de su época, para escribir un libro en 1657 titulado “De ratiociniis in ludo aleae”.

El concepto de esperanza matemática o valor medio de una variable aleatoria ξ, que se representa por $\bar{\xi}$ es muy importante, se define como la suma de los productos de los valores de la variable aleatoria ξ por sus probabilidades. Da una medida de la posición de ξ, si por ejemplo se trata de la altura de los individuos de un colectivo escogido al azar dentro de una gran ciudad, el valor medio nos indica si la población en general es alta o baja.

Otro concepto importante es el de varianza, que se representa por σ^2 y se define como la suma de los productos de los cuadrados de las diferencias entre los valores de ξ y de su valor medio $\bar{\xi}$ por sus probabilidades, la varianza da una medida de la dispersión y por tanto de la concentración de ξ respecto a un valor medio; si σ^2 es grande en el ejemplo anterior significa que la estatura de los individuos está repartida muy uniformemente.

En general se llama momento de orden r a la suma de los productos de las potencias de exponente r de los valores que puede tomar la variable aleatoria por sus probabilidades.

Hay variables aleatorias discretas y continuas. Las primeras son aquellas que pueden tomar un valor discreto finito o infinito. Por ejemplo la suma de los puntos de un dado que se arroja un número entero de veces o infinitas veces. Los segundos, también llamadas geométricas, son aquellas que toman un valor cualquiera en el interior de un segmento, de un área o de un volumen. Por ejemplo los impactos de una bala de fusil o de pistola sobre un blanco, o los impactos de un proyectil de cañón en el frente o en un campo de tiro. En algunos casos el segmento, área o volumen puede ser infinito.

Más adelante describiré un nuevo tipo de variables aleatorias que no encajan dentro de las anteriores, que he encontrado en mis investigaciones y que he denominado fractales; a los que ya anteriormente hice referencia.

Pascal utilizó el artificio de los triángulos aritméticos para la resolución de problemas de probabilidades. Un triángulo aritmético es una figura geométrica que se represente a continuación, que se prolonga indefinidamente

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
\end{array}
$$

....

que se obtiene escribiendo las dos primeras filas como están en la figura y las siguientes se obtienen escribiendo en los dos extremos 1, y los restantes números se obtienen sumando los dos de la fila superior entre los que está. Así se tiene que la tercera fila tiene dos 1 en los extremos, y en el centro 2 suma de los dos 1 de la segunda fila entre los que está; la cuarta fila tienen dos 1 en los extremos y dos 3 en el centro, suma de 1 y 2 de la tercera fila entre los que está. Y así sucesivamente para las infinitas filas que siguen.

Para el lector matemático vamos a dar la ley de formación de los triángulos aritméticos. Numeramos las filas n = 0, n = 1, n = 2,... para las tres primeras y así sucesivamente. Los números de la fila n son los coeficientes bimoniales $\binom{n}{0},\binom{n}{1},\ldots,\binom{n}{n}$ que son en número de n + 1, los cuales numeramos desde p = 0 hasta p = n. Se pueden calcular los números de la fila n + 1, en función de los de la fila n por la fórmula:

$$\binom{n+1}{p}=\binom{n}{p}+\binom{n}{p-1} \qquad (1)$$

válida desde p = 1 hasta p = n. Para p = o, cualquiera que sea n el primer número (p = o) siempre vale 1.

Pascal estudió detalladamente estos triángulos aritméticos, de los que hizo un gran uso; hoy llevan su nombre, pero el matemático persa Al Kachi que vivió en Samarcanda entre los siglos XIV y XV ya los utilizó.

El primer libro sobre probabilidades muy importante fue escrito por Jacobo Bernoulli (1654-1705) se titula “Ars Conjectandi”, fue publicado en 1713 después de su muerte.

En él figura la primera demostración rigurosa de la ley débil de los grandes números en el juego de cara y cruz, hoy es conocido como el teorema de Bernoulli y se hace un gran uso de él, establece un puente de comunicación entre frecuencia relativa y probabilidad, entre Realidad práctica y Teoría abstracta, entre Matemática pura y Matemática aplicada.

Bernouilli introdujo la distribución de probabilidad que lleva su nombre, que es la suma de n variable aleatorias binomiales independientes, y de la misma probabilidad. La distribución binomial es la de una variable aleatoria que solamente puede tomar dos valores con probabilidades complementarias p y 1-p. El lanzamiento de una moneda es una variable binomial y el de n monedas es una variable de Bernoulli. Un ejemplo de aplicación del teorema de Benoulli es el siguiente: si se arrojan n dados y el número de veces que sale el 5 es N(5), la frecuencia relativa del 5 es el cociente de dividir N(5) por n, y cuando n tiende a infinito, se cumple que el límite de N(5)/n tiende a un sexto (1/6); pero como además esta frecuencia relativa es también una variable aleatoria se cumple también que

$$\lim_{n\to\infty} \Pr ob\left(\left|\frac{N(5)}{n}-\frac{1}{6}\right|=0\right)=1 \qquad (2)$$

es decir que el límite de la probabilidad de que N(5)/n cuando n tiende a infinito es un sexto vale 1. En lenguaje más técnico se dice que la frecuencia relativa N(5)/n converge en probabilidad a la probabilidad un sexto de que salga el número 5. Esta es la ley débil de los grandes números, sobre la que volveremos más adelante.

Jacobo Bernouilli destacó mucho en otras partes de las Matemáticas. La familia Bernouilli tuvo ocho matemáticos importantes distribuidos en tres generaciones. La primera estaba formada por Jacobo y su hermano Juan (1667-1748) que también fue un matemático muy importante.

Daniel Bernoulli (1700-1782) hijo de Juan, ha sido uno de los físicos matemáticos más importantes, entre otras cosas es autor del teorema o principio que lleva su nombre fundamental en la Hidrodinámica que relaciona el aumento de la velocidad de un fluido con la disminución de la presión. En un libro reciente de Michael Guillen que lleva por título “Cinco ecuaciones que cambiaron el mundo” se afirma que estas cinco ecuaciones fueron: La ley de la gravitación universal de Newton, el antecitado principio de Bernoulli, la ley de inducción electromagnética de Faraday, el principio de la entropía de Clausius y la ecuación de Einstein de que la energía es igual al producto de la masa por el cuadrado de la velocidad de la luz ($E = mc^2$) en el vacío.

Daniel Bernouilli hizo una aportación muy importante de la aplicación de las probabilidades a las actividades humanas, muy adelantado para su tiempo, es la introducción de lo que el llamó esperanza moral, fue el iniciador de la teoría matemática de la utilidad, que ha sido extraordinariamente desarrollada por los matemáticos y economistas del siglo XX. La hipótesis de Bernouilli establece que la riqueza R proporciona al que la posee una utilidad U, tal que el incremento de la última ΔU es directamente proporcional al incremento de la riqueza ΔR e inversamente proporcional a la riqueza poseída antes de su incremento. Se expresa por la fórmula:

$$\Delta U_n = k\,\frac{\Delta R_n}{R_n} \qquad (3)$$

siendo

$$\Delta U_n = U_{n+1} - U_n; \Delta R_n = R_{n+1} - R_n \qquad (4)$$

siendo K una constante.

De esta ecuación se sigue que si la riqueza crece en progresión geométrica, la utilidad solo crece en progresión aritmética. La ecuación anterior también es aplicable si hay decrecimiento de riqueza y utilidad en vez de crecimiento.

La hipótesis de Bernouilli está relacionada con las ideas del filósofo y economista inglés. Bentham (1748-1832) autor de la teoría del utilitarismo que es un proyecto de relacionar matemáticamente la riqueza con la felicidad y el bienestar. En nuestros días estas ideas se han puesto de modo con la Sociedad del bienestar y la teoría matemática de las funciones del bienestar de la matemática moderna.

Para el lector matemático incluimos este párrafo, que se puede saltar. La ecuación (3) tiene la solución

$$U_n = U_o + hn; R_n = R_0\left(1+\frac{h}{k}\right)^n \qquad (5)$$

ya que

$$U_{n+1} - U_n = h = k\,\frac{\left(1+\frac{h}{k}\right)^{n+1} - \left(1+\frac{h}{k}\right)^n}{\left(1+\frac{h}{k}\right)^n} = k\,\frac{h}{k} \qquad (6)$$

para u = 0, U_o y R_o son la utilidad y la riqueza mínimas y para que se note el aumento de utilidad h, el incremento de la riqueza ha de ser la fracción h/k de R_n.

Ha de existir una riqueza mínima R_o y una utilidad U_o mínima, para que sea válida la ecuación (3), porque en caso contrario bastaría con ganar un céntimo, si no se tuviese absolutamente nada de dinero para obtener una utilidad infinita. Por otra parte ha de haber también un incremento mínimo de utilidad h que sea apreciable, por debajo del cual no se aprecie la utilidad obtenida, de modo que incrementos de riqueza inferiores a la fracción h/k de la que se posee no incrementa la utilidad.

Si en la ecuación (3) se sustituye el incremento Δ por la diferencial se obtiene una ley exponencial que liga a la riqueza con la utilidad.

Es curioso que la hipótesis de Bernouilli tiene una relación matemática muy estrecha con la ley fisiológica de Weber (1795/1878) y Fechner (1801-1887). La ecuación (3) y sus consecuencias son válidas si se sustituye el significado de las letras, U y R por el de las nuevas letras S y E sensación y estímulo

$$\Delta S = k\,\frac{\Delta E}{E} \tag{7}$$

que es la expresión matemática de la ley de Weber-Fechner, la cual ha sido experimentada en muchos casos, como por ejemplo la sensación S que siente una persona al aumentar el peso E que ha de soportar con la mano, o la sensación S que siente una persona al aumentar la iluminación E de un local en cuyo interior está.

Las ecuaciones (3) y (7) un también las de la evolución de un capital en el tiempo, colocado al interés compuesto.

En resumen la sensación o la utilidad son variables discretas que varían a saltos, aumentando de h en h, son siempre un número entero de unidades h. Para que el incremento del estímulo o de la riqueza produzcan un incremento de h en la sensación o en la utilidad, es preciso que el estímulo o la riqueza se incrementen en una fracción h/k de sus valores; cualquier incremento menor no produce aumento de la sensación o de la utilidad.

Existen varias clases de sucesos aleatorios, entre ellos están los sucesos raros, son aquellos que tienen una probabilidad p muy pequeña de realizarse, pero que no lo es si se realiza un gran número n de pruebas o ensayos, de modo que el límite del producto np, cuando p tiende a cero y n a infinito, tiene un valor λ; entonces se realizan de acuerdo con una distribución de probabilidad. Tal es el caso de la distribución de Poisson, que se obtiene como el límite de una distribución de Bernoulli cuando la probabilidad p tiende a cero y el exponente n de la misma (número de ensayos) tiende a infinito, de modo que el límite del producto np es finito. Tiene muchas aplicaciones.

Dos sucesos aleatorios E_1 y E_2 se dice que son independientes, si la probabilidad de realización de uno cualquiera de ellos es independiente de que se haya realizado o no el otro. Para los dos sucesos E_1 y E_2 que no son independientes la probabilidad de que se realicen ambos $P(E_1 \cap E_2)$ es igual a la probabilidad de que se realice uno de ellos $P(E_1)$ por la probabilidad $P(E_2/E_1)$ de que se realice E_2 habiéndose realizado E_1. La $P(E_2/E_1)$ se llama probabilidad condicional y es un concepto muy importante.

Me he ocupado de lo que he llamado casi probabilidades, que es cuando no existen las probabilidades, pero si existen las probabilidades condicionales, como es el caso de un plano ilimitado donde no existe la probabilidad de elegir un punto al azar dentro de un círculo C_1, pero si existe la probabilidad de elegir un punto al azar dentro de un círculo C_1 sabiendo que el punto está dentro de un círculo C_2 que rodea al C_1. La probabilidad condicional de este último caso es el cociente de dividir el área del círculo interior C_1 por la del círculo exterior C_2.

Así como Daniel Bernouilli se adelantó a su tiempo con la introducción de la esperanza moral, también Bayes (1702-1761) se adelantó al suyo, con el teorema o principio que lleva su nombre, que da origen a las probabilidades de las causas. Sobre este teorema se ha polemizado y se han cometido muchos errores, por una mala interpretación del mismo. Vamos a dar un ejemplo: supongamos que tenemos tres urnas U_1, U_2, U_3, la primera con 1 bola blanca y 1 negra, la segunda con 2 blancas y 1 negra y la tercera con 2 blancas y 3 negras; se extrae al azar una bola que es blanca, sin saber de que urna se ha extraído, el tratar de averiguar la probabilidad de que haya sido extraída de una determinada, por ejemplo la U_1, sin conocer a priori la probabilidad de escoger cada una de las tres urnas, es imposible, el problema no tiene solución; a veces se ha resuelto erróneamente el problema, atribuyendo la misma probabilidad de elección a cada una de las tres urnas y se ha calculado entonces la probabilidad pedida por la fórmula que vamos a dar; pero esta solución es incorrecta, porque la hipótesis de equiprobabilidad en la elección de las urnas carece de fundamento, es gratuita. Sin embargo si se conocen las probabilidades de elección $P(U_1)$, $P(U_2)$, $P(U_3)$ de las tres urnas, entonces el problema si tiene solución. Supongamos que estas tres sean 1/2, 1/3 y 1/6, entonces si se puede calcular la probabilidad pedido $P(U_1 1/B)$ que es una probabilidad condicional, que vale:

$$P(U_1/B)=\frac{P(U_1)P(B/U_1)}{P(U_1)P(B/U_1)+P(U_2)P(B/U_2)+P(U_3)P(B/U_3)} \qquad (8)$$

que en este caso particular vale:

$$P(U_1/B)=\frac{1/2x1/2}{1/2x1/2+1/3x2/3+1/6x2/5} \qquad (9)$$

Actualmente el teorema de Bayes o de las probabilidades de las causas tiene una gran importancia, y así la Estadística de nuestros días se divide en Bayesiana y no Bayesiana. En el ejemplo anterior la causa es la urna escogida y el efecto la bola extraída.

He investigado por los métodos del análisis de árboles, el caso más general en que supongo que existe una cadena de causas en vez de una sola, de modo que hay una causa primera, cuyo efecto es una causa segunda y así sucesivamente hasta una causa última, efecto de la anterior y cuyo efecto es la extracción de una bola. El ejemplo más sencillo es aquel en el que hay dos urnas, cada una de las cuales contiene dos urnas y cada una de las cuales tiene una composición conocida de bolas blancas y negras. Para poder resolver el problema hay que conocer las probabilidades a priori de elección de las dos urnas primeras; las demás probabilidades son calculables y se disponen como las ramas de un árbol.

En el Cálculo de Probabilidades tiene mucho interés cálcular límites inferiores de la probabilidad de realización de un suceso deseable y superiores cuando el suceso es indeseable, por lo que es muy grande el número de esta clase de límites calculados matemáticamente; se llaman desigualdades. Nos vamos a limitar a una de ellas, la llamada de Tchebycheff por la mayoría de los autores, y de Bienaymé-Tchebycheff por otros, sobre todo los autores franceses.

Bienaymé (1796-1876) de vida muy agitada, expulsado de la Escuela Politécnica por Luis XVIII en 1816, acusado de ser contrario a la monarquía, cesado como Inspector de Hacienda, acusado de ser contrario a la república y de la Sorbona en 1851. Fue uno de los iniciadores de las variables aleatorias, formuló por primera vez esta desigualdad en 1853.

Tchebycheff (1821-1894) fue un matemático muy importante por sus aportaciones a la Teoría de Números, por los polinomios e integrales que llevan su nombre; redescubrió y le dio mas generalidad a esta desigualdad, que afirma que “la probabilidad de que el valor absoluto de la diferencia entre una variable aleatoria y su valor medio, sea mayor que el producto de la raíz cuadrada de la varianza por un número positivo t, es menor o igual que $1/t^2$“. Esta desigualdad es válida para toda distribución de probabilidad, cuya varianza sea finita, y es muy importante para establecer criterios de convergencia estocástica y para demostrar las leyes de los grandes números.

Una de las consecuencias prácticas más importante de esta desigualdad es que si las varianzas $\sigma_1^2,\ldots,\sigma_n^2,\ldots$ de una sucesión de variables aleatorias $\xi_1,\ldots,\xi_n,\ldots$ de valor medio nulo, converge a cero, entonces también la sucesión de las ξ converge a cero.

Entre las distintas clases de convergencia estocástica figuran como muy importantes la convergencia en probabilidad y la convergencia casi cierta. La primera se define así, se dice que una sucesión de variables aleatorias $\xi_1,\ldots,\xi_n,\ldots$ converge en probabilidad a cero si

$$\lim_{n\to\infty} \Pr ob\left(\xi_n=0\right)=1 \qquad (10)$$

Se dice que una sucesión de variables aleatorias $\xi_1,\ldots,\xi_n,\ldots$, converge en probabilidad a una variable aleatoria μ, que eventualmente puede ser un número cierto, si la sucesión $\xi_1-\mu_1,\ldots,\xi_n-\mu_n$ converge en probabilidad a cero.

La convergencia casi cierta se define así la sucesión de variables aleatorias $\xi_1,\ldots,\xi_n,\ldots$, converge casi ciertamente a cero si

$$\Pr ob\left(\lim_{n\to\infty}\xi_n=0\right)=1 \qquad (11)$$

La sucesión de variables aleatorias $\xi_1,\ldots,\xi_n,\ldots$, converge casi ciertamente a la variable aleatoria μ que eventualmente puede ser un número cierto si la sucesión $\xi_1-\mu,\ldots,\xi_n-\mu,\ldots$ converge casi ciertamente a cero.

La convergencia casi cierta es más fuerte que la convergencia en probabilidad o lo que es lo mismo esta última es más débil que la primera. Ser más fuerte significa que si una sucesión converge casi ciertamente, también converge en probabilidad, es decir que la (11) implica la (10). Ambas convergencias son muy importantes para la demostración de las leyes de los grandes números.

Se dice que una sucesión de variable aleatorias $\xi_1,...,\xi_n,...$, del mismo valor medio m sigue la ley débil de los grandes números, cuando la sucesión de las medias aritméticas

$$\xi_1, \frac{\xi_1+\xi_2}{2},..., \frac{\xi_2+\xi_2+...+\xi_n}{n}, \tag{12}$$

converge en probabilidad a su valor medio m, cuando n tiende a infinito. Como dijimos anteriormente Jacobo Bernoulli obtuvo este resultado para las variables aleatorias que siguen la distribución de probabilidad que lleva su nombre, en particular en el juego de cara y cruz de una moneda.

Si en la definición anterior se sustituye la convergencia en probabilidad por la convergencia casi cierta, entonces se dice que la sucesión de las ξ cumple la ley fuerte de los grandes números.

Obsérvese que la diferencia entre (10) y (11) está en que en la primera el límite está antes que la probabilidad, se trata del límite de una probabilidad y que en (11) el límite está detrás de la probabilidad, se trata de la probabilidad de un límite.

La diferencia esencial entre la convergencia en probabilidad y la convergencia casi cierta, es que la primera la probabilidad de que $\xi_n - \mu$ sea inferior a cualquier número por pequeño que sea, se hace tan próxima a la unidad como se quiera, con tal de escoger n suficientemente grande, que es lo que expresa la (10). Por el contrario en la convergencia casi cierta la riqueza de resultados es mayor, la probabilidad de que conjuntamente todas las diferencias $\xi_{n+p} - \mu$ (p cualquier número natural), o lo que es lo mismo, la máxima de estas diferencias, sean inferiores a cualquier números por pequeño que sea, se hace tan próxima a la unidad como se quiera, con tal de escoger n suficientemente grande, que es lo que expresa la (11). Es decir que la probabilidad de que la sucesión de los valores particulares de todas las ξ converja al de μ es igual a la unidad en cualquier experiencia o prueba a la que están asociadas las ξ.

Si nos fijamos en el juego de cara y cruz, en virtud de la convergencia en probabilidad, la frecuencia del número de veces que sala cara difiere de 1/2 tan poco como se quiera, con tal de echar un número n de veces la moneda suficientemente grande. Mientras que en la convergencia casi cierta, la sucesión de dichas frecuencias (a medida que varía n) converge a 1/2 cuando n tiende a infinito, lo que establece un lazo entre la teoría de la frecuencia de Von Mises (1883-1953) y la axiomática de la probabilidad que veremos más adelante.

Las dos convergencias estocásticas anteriores dan valor práctico a las probabilidades y contribuyen al fundamento científico de los sondeos y muestreos estadísticos. En mi libro “Cálculo de Probabilidades y Procesos Estocásticos” (edit. Paraninfo 1974) pueden verse ejemplos de variables aleatorias que convergen en probabilidad a otras, pero que no convergen casi ciertamente.

Laplace (1749-1827) en 1812 publicó su “Teoría analítica de las Probabilidades” y en 1814 su “Ensayo filosófico de las Probabilidades”. El primero de estos libros es una obra maestra que marca una época, desde su aparición se puede hablar en Probabilidades de antes y después de Laplace, es un libro difícil por el aparato matemático que emplea, es de gran altura científica. El Ensayo está escrito en forma muy clara y didáctica y no utiliza las matemáticas.

La definición de probabilidad de Laplace, que para muchos casos prácticos sigue teniendo valor, es la siguiente: “La probabilidad es el cociente entre el número de casos favorables y el número de casos posibles en la realización de un suceso”.

En realidad es el método que habían seguido los matemáticos anteriores a él (Pascal, Fermat, Bernouilli, De Moivre, Bayes, Buffon, etc.) y que se sigue utilizando en muchos casos.

El nombre de probabilidad tiene apellidos y así se habla de probabilidades directas o a posteriori y de probabilidades inversas o a priori. Las probabilidades directas son las que pueden calcularse a partir de probabilidades conocidas, cuando se trata de sucesos aleatorios que son el efecto de causas conocidas, como por ejemplo la probabilidad del número de veces que

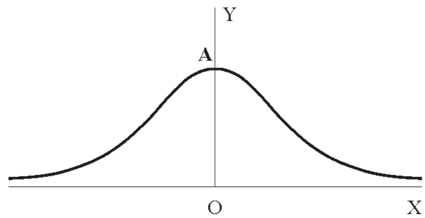

Figura 1

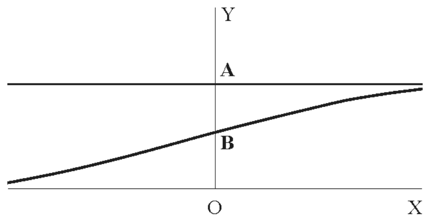

Figura 2

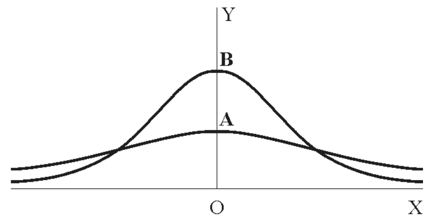

Figura 3

sale caro cuando se echan n monedas; en estos problemas se trata de prever el futuro. Las probabilidades inversas o a posteriori son las ligadas al teorema de Bayes, se trata de conocidos los efectos aleatorios calcular las probabilidades de las causas que los han producido, como en el caso antes explicado de la extracción de bolas de diversas urnas, que a su vez han sido elegidas de acuerdo con probabilidades dadas a priori.

El juego de azar se llama equitativo y noble, si el coste de cada partida es igual a la ganancia esperada. Si dos jugadores A y B juegan una partida con probabilidades p y q de ganarla (p+q = 1), sus apuestas a y b han de ser tales que las esperanzas matemáticas de ganancia de A y B sean iguales, estas son pb y qa, luego las apuestas de cada uno han de ser proporcionales a sus probabilidades de ganar.

En los textos se comenta que si dos jugadores 1 y 2 disponen de cantidades de dinero A y B distintas, si juegan hasta la ruina de uno de ellos, el que tiene más dinero tiene mayor probabilidad de arruinar al otro, por tanto tiene ventaja y el juego no es justo, pero no es así, porque el que tiene más dinero arriesga más en el juego, porque si gana, gana menos que el otro, si es éste el que gana, y si pierde, pierde mas que el otro si es éste el que pierde.

De Moivre (1667-1754) fue un hugonote francés, cuya familia huyó de Francia a Inglaterra, cuando se revocó el edicto de Nantes. Hizo importantes aportaciones a las probabilidades, introdujo la ley normal o de Gauss, como límite de la distribución de Bernouilli. La distribución de probabilidad de Gauss es la más importante de todas, su función de densidad de probabilidad f(x) es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \qquad (13)$$

que se representa en la figura 1

La ordenada y = f(x) por dx da la probabilidad de que la variable aleatoria x esté comprendida entre x y x+dx. La figura 2 representa la función de distribución de probabilidad, en la que la ordenada F(x) representa la probabilidad de que x esté comprendida entre $-\infty$ y x.

El valor medio es cero, y σ es la raíz cuadrada de la varianza. La ordenada máxima OA es inversamente proporcional a σ, y en la figura 3 se muestra que cuando A está más alto sobre OX, la variable aleatoria está más concentrada alrededor de su valor medio O, o lo que es lo mismo que cuando está más baja, la variable aleatoria está más dispersa. Las variables aleatorias que siguen esta distribución de probabilidad se llaman normales y también gaussianas. Observese que la curva de la figura 1 tiene forma de campana.

Se conocen como teoremas centrales del límite, los que dan las condiciones que han de cumplir las variables aleatorias, cuya suma, cuando viene expresada como medida estandarizada, y el número de sumandos tiende a infinito, tiende a la distribución de Gauss.

Son muchos estos teoremas y muchos los matemáticos que han contribuido a ellos. El primero en el orden cronológico, y aplicado a la distribución de Bernoulli, como ya hemos dicho es debido a De Moivre y de forma más general se debe a Laplace que no lo demostró de modo riguroso. Fué Liapounov en 1901 quien le dio una forma muy precisa y rigurosa. Otras expresiones de este teorema son debidos a Levy y Cramer, que lo enunciaron independientemente uno de otro en 1925.

Lindeberg y Levy demostraron que si $\xi_1, \ldots, \xi_n, \ldots,$ son variables aleatorias de valor medio nulo, varianza finita y de la misma función de distribución de probabilidad, se cumple que el límite

$$\lim_{n\to\infty} \frac{\xi_1 + \ldots + \xi_n}{\sqrt{n}} \tag{14}$$

converge a la ley normal de valor medio nulo y varianza igual a la de los sumandos aleatorios.

Los teoremas del límite central son la causa de que en la Ciencia, la Técnica y la Naturaleza tenga tanta importancia la ley normal, y que bajo condiciones muy generales la suma de variables aleatorias iguales e independientes, dividida por la raíz cuadrada del número n de sumandos, tienda a comportarse asintoticamente como una variable normal. Concretamente los errores en las medidas físicas, tienden a distribuirse según la ley normal alrededor de cero, y cuanto menor es su varianza, mayor es la precisión en las medidas.

Las variables aleatorias para las que son válidos los teoremas anteriores, se dice que pertenecen al dominio de atracción de la ley normal. Sin embargo existen otras variables aleatorias, cuyas sumas (14) dividida por otra potencia de n, no pertenecen al dominio de atracción de la ley normal sino al de otras leyes, llamadas estables. Si en (14) se sustituye la raíz cuadrada de n por n, y la expresión converge lo hace a la distribución de Cauchy, y si se sustituye por n^2 , si converge lo hace a la distribución de Levy. Puede verse mi antecitado libro "Cálculo de Probabilidades y Procesos estocásticos".

Un ejemplo de distribución de probabilidad, cuyo teorema central del límite conduce a la distribución de Levy es el número de lanzamientos necesarios para conseguir un empate en el juego de cara y cruz. Otro ejemplo investigado y desarrollado por mi es la evolución de una población física o biológica en un proceso de natalidad y mortalidad, cuando éstas tienen la misma probabilidad, es decir son iguales en un instante dado la probabilidad de nacimiento que de muerte de un individuo.

Cauchy en 1853 en los "Comptes rendus" de la Academia de Ciencias de Paris introdujo el concepto de función característica $\varphi(t)$, íntimamente ligado a la función de distribución de probabilidad F(x) que ha mostrado ser uno de los métodos más fecundos en el estudio de las convergencias estocásticas, de las leyes de los grandes números y de los teoremas centrales del límite. Levy dio una nueva definición de la función característica como la esperanza matemática de lo exponencial e^{itx} y la utilizó para estudiar las distribuciones de probabilidad estables y que no pertenecen al dominio de atracción de la ley normal.

Ya he utilizado (véase mi libro antecitado) la función característica para estudiar el producto y el cociente de variables aleatorias independientes o no, con lo que se encuentran propiedades de las funciones trascendentes superiores. Estas operaciones con variables aleatorias presentan diferencias con el caso de variables ciertas y así por ejemplo el producto de una variable aleatoria por una suma de variables aleatorias no es distributivo, propiedad rara, porque desde los tiempos de Hamilton (1805-1865) se conocen objetos matemáticos para los que la multiplicación no es conmutativa (el orden de factores altera el producto), como es el caso de los cuaternios y de las matrices, pero son raros los objetos matemáticos, para los que la multiplicación no es distributiva.

El primer problema de probabilidades geométricas lo formuló Buffon (1707-1788), que consiste en calcular la probabilidad de que al echar al azar una aguja de longitud ℓ sobre un papel lleno de rectas paralelas equidistantes a distancia d ($d > \ell$) unas de otras, la aguja corte a una raya) esta probabilidad vale $2\,l/\pi d$.

Siglos más tarde, durante la segunda guerra mundial, la resolución de este problema fue el inicio del método de Monte Carlo, muy usado en Cálculo Numérico, basado en la obtención, usando las probabilidades, del valor de integrales, de la solución de sistemas de ecuaciones diferenciales, determinación de puntos extremales, etc. El problema de la aguja de Buffon se puede usar para calcular el valor de π a partir del cociente entre el número de veces que se echa la aguja y el número de veces que corta la aguja a una raya.

Lo problemas sobre probabilidades geométricas experimentaron un gran aumento en el siglo XIX; entre ellos figura el de Cesàro, que consiste en calcular la probabilidad de que una cónica sea elipse o hipérbola escogiendo al azar tres de los seis coeficientes de la ecuación de una cónica, técnicamente los que componen su A_{33} que son los que determinan si la cónica es elipse o hipérbola. Naturalmente la probabilidad de que la cónica sea parábola vale cero.

Laplace generalizó el problema de Buffon. Poincaré introdujo el concepto de medida cinemática y especialmente Crofton destacó mucho en este campo. Blaschke (1885-1962) elevó esta disciplina a la categoría de una ciencia nueva que llamó Geometría Integral, que durante el siglo XX alcanzó un gran esplendor. Dos anécdotas de Blaschke: Rey Pastor le preguntó porque había escogido el nombre de Geometría Integral y le contestó que no sabía como llamarla y como tenía que ponerle un nombre, le puso ese porque habían muchos integrales. Cuando Hazidakis le presentó en Atenas, le llamo el creador de la Geometría Integral, y le replicó que el creador había sido su compatriota Arquímedes.

Bertrand (1822-1900) aparte de sus trabajos sobre Geometría y Teoría de Números, hizo importantes aportaciones a las Probabilidades y en su libro sobre esta materia planteó varias paradojas que en su época causaron sensación entre los matemáticos. La más conocida es la de averiguar la probabilidad de que al trazar una cuerda en un círculo al azar, sea de longitud mayor que el lado del triángulo equilátero inscrito. Según que se fije la cuerda por su dirección, por uno de los dos puntos en que corta a la circunferencia, o por su punto medio, se obtienen tres soluciones distintas que valen respectivamente 1/2, 1/3 y 1/4. La paradoja consiste en que un mismo problema admite tres solucione distintas y los matemáticos se plantearon cual era la correcta de estas tres soluciones.

Mucho y muy dispar se escribió sobre este tema, sin haber consenso. A mi me parece que esta pregunta no tiene sentido, porque es necesario agregar a la pregunta la correcta con relación a que hipótesis o leyes, porque según sean éstos, la solución correcta será distinta. En los tres casos el azar actúa según un mecanismo distinto, por tanto no se trata de un problema sino de tres problemas distintos con tres soluciones distintas. He aquí, a mi juicio, como pueden proyectarse las experiencias, o lo que es lo mismo como actúa el azar en cada uno de los tres casos. En el primero dejando constante la dirección de la cuerda, la comprobación experimental sería: echar una moneda circular sobre un papel rayado, siendo la distancia entre paralelas consecutivas igual al diámetro de la moneda. En el segundo problema dejando constante uno de los dos punto de intersección de la cuerda y la circunferencia, la comprobación experimental sería sujetar una varilla (la cuerda) de longitud superior al diámetro de la circunferencia, por uno de sus extremos a un punto cualquiera de la circunferencia dibujada en un papel y dejando caer al azar sobre el papel. En el tener problema fijando la cuerda para su punto medio, la comprobación experimental sería escoger al azar un punto en el interior de la circunferencia dibujada en un papel y trazar la cuerda perpendicular a la recta que une el citado punto con el centro de la circunferencia.

Hay que tener en cuenta que la naturaleza no puede dar la respuesta a una pregunta mal hecha. Lo primero que hay que hacer en Cálculo de Probabilidades, como en cualquier otra ciencia experimental es saber preguntar a la realidad.

Bertrand planteó otras muchas paradojas interesantes todas ellas, de algunas de las cuales me he ocupado en mis libros “Filosofía de las Matemáticas” y “Cálculo de Probabilidades y Procesos Estocásticos”.

La Teoría de las Probabilidades tardó mucho tiempo en formular una Axiomática consensuada; hubo a fines del XIX a pesar del gran desarrollo que había alcanzado esta Teoría, mucho confusionismo sobre importantes cuestiones de esta Ciencia. En relación con las leyes de los grandes números y de los teoremas centrales del límite, se decía mitad en broma y mitad en serio que eran las verdades científicas más firmes, porque los matemáticos creían que eran fenómenos físicos bien comprobados y los físicos creerán que eran teoremas matemáticos correctamente demostrados.

Von Mises (1883-1953) nació en Lemberg (hoy Lvov) en la Polonia que formaba parte de Austria-Hungria, vivió en Berlín hasta 1933 en que se exilió a los Estados Unidos, por la llegada al poder de Hitler, a pesar de haber sido aviador en la primera guerra mundial. Hizo importantes aportaciones a las Matemáticas y a la Física, y en lo que nos ocupa, fue el primero en intentar una axiomatización de las probabilidades y aunque no lo logró, su ensayo ha sido muy importante. La definición de Von Mises de la probabilidad, aunque no es rigurosa es suficiente desde un punto de vista práctico, su definición es: “La probabilidad es el valor límite de la frecuencia relativa del numero de realizaciones de un suceso cuando el número de experiencias tiende a infinito”. Lo anterior permite definir la probabilidad condicional como “el cociente de dividir la frecuencia con que ocurren simultáneamente dos sucesos por la frecuencia con que ocurren uno de ellos, cuando el número de experiencias tiende a infinito". Esto tuvo lugar en 1919, Von Mises es el fundador de la teoría frecuentista de la probabilidad, tomó como modelo la Mecánica Racional, es decir parte de la realidad para poder extraer de la misma el conocido concepto de frecuencia, lo idealiza, admitiendo que toma un valor límite, cuando el número de experiencias tiende a infinito. Su teoría está expuesta en su libro "Probabilidad, Estadística y Verdad" que alcanzó varias ediciones, escrito sin recurrir al simbolismo matemático. Más elevado es su libro "Cálculo de Probabilidades y sus Aplicaciones a la Estadística y a la Física Teórica" de 1931.

Kolmogorov (1903-1987), es uno de los matemáticos más importantes de nuestro tiempo y quien ha desarrollado la Axiomática más satisfactoria y completa de las Probabilidades en su libro escrito en alemán "Ideas fundamentales del Cálculo de Probabilidades" en 1933, de lo que es una anticipación su libro "Teoría general de la medida y de las probabilidades" de 1929. El tratamiento que da a esta importante cuestión es muy complicado, requiere un gran conocimiento de la teoría de conjuntos y de la teoría de la medida, así como de las σ-algebras, llamadas por algunos autores algebras de sucesos de Borel.

La nueva clase de variable aleatorias que he llamado fractales, a las que he aludido antes y sobre las que volveré no siguen la axiomática de Kolmogorov.

Existe también una escuela partidaria de la probabilidad subjetiva a la que pertenecen Keynes (Tratado de Probabilidad 1921) y Jeffreys (La Inferencia Científica 1931).

Para dar fin a este aspecto, señalaré que Harald Cramer (1893-1985) sueco, en su libro "Métodos Matemáticos de la Estadística" de 1945, establece una fundamentación rigurosa de la Estadística basada sobre el Cálculo de Probabilidades. Dedica las cien primeras páginas del libro a las Teorías de Conjuntos, Teoría de la Medida, integrales de Lebesgue, de Stieltjes-Lebesgue y de Fourier.

Markov (1856-1922), fue alumno de Tchebycheff (los nombres rusos termminados en v en español, a veces se terminan en dos f); después de hacer importantes aportaciones en otras ramas de las Matemáticas, inició en 1906 sus trabajos de probabilidades. Se le puede considerar como el creador de los procesos estocásticos. Introdujo las cadenas y el proceso estocástico que llevan su nombre.

Sea un sistema S que en instantes sucesivos 0, 1, 2,..n,... se encuentra en uno cualquiera de los m estados E_2 , ..., E_m, de modo que la probabilidad a_{ij} de que se encuentre en el instante n en el estado E_j, encontrándose en el instante n-1 en el E_i, es independiente de n, es un cadena de Markov. Los a_{ij} forman una matriz estocástica M, en la que la suma de todos los elementos de una fila vale 1. La cadena se puede resolver si se da el vector inicial de las probabilidades de los estados de S en el instante cero, que eventualmente puede estar en uno cualquiera de los m citados.

Si el tiempo en vez de ser una variable discreta es una variable continua, la cadena de Markov es un proceso estocástico de Markov, que se caracteriza (así como las cadenas) porque la distribución de probabilidad en el futuro depende solamente del presente; el pasado no influye.

Los procesos estocásticos en general son la evolución de una distribución de probabilidad en el tiempo. Se llaman estacionarios si las causas de las variaciones son independientes del tiempo, los cuales fueron introducidos por Khintchine (1894-1950) y en ellos el valor medio y la varianza con constantes y finitos y la función de correlación en dos instantes t_1 y t_2, depende solo de t_1-t_2.

Un concepto muy importante que ha pasado de la Estadística a otras materias, es el de los errores de tipo I y de tipo II, introducidos por Neyman y Pearson en la teoría del contraste de hipótesis. Un error de tipo I es cuando una hipótesis estadística es rechazada siendo verdadera y de tipo II si es aceptada siendo falsa. En la teoría del Riesgo un error de tipo I es cuando es un error para el riesgo del productor y del tipo II cuando es un error para el riesgo del consumidor. En la teoría de la Fiabilidad un error del tipo I es cuando un sistema (una alarma) no funciona cuando debe funcionar y del tipo II cuando funciona no debiendo funcionar. Por mi cuenta he opinado que los referées y los comité editoriales de una revista científica cometen un error de tipo I al rechazar un artículo bueno y un error de tipo II al aceptar un artículo malo.

Neyman y Pearson pertenecen al grupo de los fundadores de la estadística matemática moderna que usa como herramienta matemática el Cálculo de Probabilidades. Neyman (1894-1981) fue un ruso que cuando la revolución se fue a Polonia donde trabajó en la Universidad de Varsovia, después emigró a Inglaterra donde trabajó en la Universidad de Londres, y por último emigró a los Estados Unidos donde trabajó en la Universidad de Berkeley. Pearson (1895-1980), aparte de su colaboración con Neyman en la teoría de la estimación y de la comprobación de las hipótesis estadísticas, hizo importantes aportaciones a las teorías del control de la calidad y de la investigación operativa, nacidas en el siglo XX. Su padre Karl (1875-1936) junto a Galton (1822-1911) fueron pioneros de la estadística matemática y de la biometría.

La hipótesis ergódica, los teoremas ergódicos y la teoría ergódica están muy relacionadas entre sí y han invadido muchas áreas de las Matemáticas puras y aplicadas, así como de la Física teórica. La hipótesis ergódica fué formulada por Maxwell (1831-1879) en la teoría cinética de los gases, en la mezcla de líquidos y en los fenómenos de difusión, afirma que cualquiera que sea la distribución inicial de las partículas, al final se reparten uniformemente. En la teoría ergódica se sustituyen las medidas instantáneas realizadas por los Físicos experimentales por valores medios que pueden obtenerse en las diferentes posiciones que ocupa una partícula a lo largo del tiempo, o por las que corresponden en un instante dado a un gran número de partículas y lo que afirma el teorema ergódico es que ambos valores medios son iguales.

En los años veinte del siglo pasado, nace una nueva rama de las Probabilidades, que es la teoría de los juegos de estrategia, en los que interviene también la habilidad del jugador; tiene un desarrollo muy rápido y en los años cuarenta y cincuenta se puso muy de moda. En 1928 Von Neumann (1903-1957) publicó un artículo que inicia en grande esta teoría, en él figura un célebre teorema del minimax, y en 1944 Von Neumann y Morgenstern (1902-1977) su libro "Teoría de juegos y comportamiento económico" revolucionó la Economía. Borel (1872-1956) en 1921 había señalado la importancia que en problemas económicos, militares y en general en las actividades humanas tienen esta clase de juegos e inició el estudio de los mismos; según Borel, con anterioridad a él, solo Bertrand en su Cálculo de Probabilidades había tratado un problema en el juego de bacarrá; en 1938 Borel publicó un libro sobre este tema. En 1953, Frechet publicó en Econométrica una nota en la que daba cuenta de los trabajos de Borel en forma muy elogiosa, y señalando la anterioridad de los mismos respecto a los de Von Neumann; en 1955 volvió a insistir sobre lo mismo en su libro "Las Matemáticas y lo concreto", en el mismo número de la revista Econometrica, Von Neumann publicó una réplica a la nota de Fréchet.

Relacionada con los juegos de estrategia está la Teoría de la Decisión en la que Wald en 1939 introdujo el principio del mimimax, que expresa que el criterio seguido en la toma de decisiones, es que el riesgo de tomar una decisión equivocada sea mínimo. En la moderna teoría de la Decisión se llama solución de Bayes a una función de decisión estadística que minimice el riesgo medio relativo a una distribución de probabilidad. Se suele llamar postulado de Bayes, que es muy discutible, a que cuando en un problema de probabilidades de causas, no existe información sobre cuales son las probabilidades a priori, y se desconocen cuales son sus valores, se adopte por principio el postulado de que todas las causas tienen la misma probabilidad.

Los juegos de que se ocupaba Von Neumann se llaman cooperativos, pero existen otros no cooperativos, relacionados con el problema de la negociación, en los que el fundamento es un teorema de John Nash (nacido en 1928). Este personaje es muy interesante, porque muy joven en 1950, sorprendió al mundo científico con su tesis doctoral en Princeton y su primera publicación, seguidas de unas pocas publicaciones más en los años cincuenta sobre los juegos no cooperativos, que le valieron el premio Nobel en Economía en 1994, cincuentenario del libro antecitado de Von Neumann y Morgenstern. Lo obtuvo conjuntamente con Hansanyi (húngaro) que había trabajado sobre juegos con información incompleta y con Selten (alemán) que había trabajado sobre la manera de distinguir los resultados razonables de los irrazonables en los juegos.

En los años cincuenta Nash descubrió algunos teoremas muy difíciles y avanzados sobre variedades algebraicas reales, y sobre inmersiones de variedades riemamianas en espacios euclideos de mayor número de dimensiones. Estos logros sorprendieron al mundo matemático y no hubiera sido extraño que dada su juventud se le hubiera concedido la medalla Field; pero no se le concedió.

La vida de Nash fue muy trágica, su época activa vino a ser de una década, de 1948 a 1958, en que cayó gravemente enfermo, enfermedad que duró hasta finales de los años ochenta, salvo unos pocos años a finales de los sesenta, en que también realizó algunos trabajos más. Su vida ha sido llevada al cine en la película "Una mente prodigiosa".

Además de Nash existen otros dos grandes científicos contemporáneos nuestros, que realizaron una grandísima labor estando gravemente enfermos o inválidos. Son Pontriagine (1908-1988) ruso ciego, profesor en la Universidad de Moscú y en el Instituto Steklov que hizo grandes aportaciones a los procesos de optimización, la topología y los grupos topológicos. El otro es Hawking (nacido en 1942) titular de la cátedra Lucasiana de Matemáticas de la Universidad de Cambridge, que ocupó antes Newton, autor de dos bestseller científicos. "La Historia del tiempo" y "El Universo en una cáscara de nuez", algunos han llegado a compararle con Einstein.

La Mecánica Racional es la ciencia que con el empleo del método matemático estudia el equilibrio y el movimiento de los puntos y sistemas materiales ocasionados por las fuerzas, de modo que la causa son las fuerzas y el equilibrio o el movimiento el efecto. Fue iniciada por Galileo (1564-1642) desde un punto de vista moderno y establecido de modo riguroso y unificado por Newton (1643-1727) en sus "Principia" (1687). Durante el siglo XVIII se avanzó mucho en el conocimiento de la misma, por obra de muchos sabios entre los que figura de manera sobresaliente. Euler (1707-1783) matemático muy prolífico. Lagrange (1736-1818) publicó su "Mecánica Analítica" (1788) que constituye un tratado teórico ya moderno y completo, en el que desarrolló las ecuaciones que llevan su nombre, con lo que se comienza la "mecanización" o "automatización" de la solución de los problemas de la Mecánica. Laplace en 1812 publica su "Mecánica celeste" en la que aplica los nuevos métodos de la Mecánica al estudio del movimiento de los planetas y satélites del sistema solar. Hamilton (1805-1865) obtiene unas nuevas ecuaciones que llevan su nombre, que dan la solución de los problemas de la Mecánica. Desde entonces se dice que existen tres enfoques de esta Ciencia que son el newtoniano el lagrangiano y el hamiltoniano. Jacobi (1804-1851) con su célebre ecuación culmina la obra que hace que la Mecánica puede considerarse a partir de entonces como un capítulo del Análisis Matemático. Hasta este momento Mecánica y Probabilidades siguen caminos distintos, sin interferencias; para saber Mecánica no hace falta saber Probabiliddes, aunque en algunos casos como Laplace, nos encontramos con un científico que sabe muchísimo de ambas materias, pero ese caso es una excepción.

En la segunda mitad del siglo XIX, Maxwell (1831-1879) Boltzmann (1844-1906) y Gibbs (1839-1903) fundan una nueva ciencia: la Mecánica Estadística en la que se conjugan las leyes de la Mecánica y las de las Probabilidades. Como dijimos antes los dos primeros establecieron la hipótesis ergódica. Se aplica a los sistemas con un número enorme de partículas, billones de millones en el caso de un cm^3 de un gas en condiciones normales de presión y de temperatura. Maxwell obtuvo la distribución de probabilidad de las velocidades de las moléculas de un gas, con lo que hace dar un paso de gigante a la Teoría cinética de los gases.

Clausius (1822-1888) definió en 1865 una nueva magnitud física la Entropía, muy importante en Termodinámica y que había usado antes de definirla. La define por su diferencial

$$dS = \frac{dQ}{T} \tag{15}$$

donde S es la entropía, Q la cantidad de calor y T la temperatura absoluta.

Boltzmann en 1875, más tarde definió la entropía mediante las probabilidades

$$\mathrm{S} = \mathrm{k} \log \mathrm{W} \tag{16}$$

donde k es la constante que lleva su nombre y W lo que él llamó la probabilidad termodinámcia. Vamos a explicar lo que hizo Boltzmann. Todo sistema formado por un gran número de partículas, está definido por su energía total y por la naturaleza de sus partículas. Cada una de las partículas está representada por un punto en el espacio de las fases, este punto viene dado por las coordenadas de su posición (tres) y por los componentes de su velocidad (tres). Este espacio está dividido en dominios de la misma extensión, y todas las partículas de un mismo dominio tienen la misma energía, estos son los microestados. Todos los microestados que tienen la misma energía forman un macroestado, entonces el cociente entre el número de microestados que forman un macroestado y el número total de microestados del sistema es su probabilidad termodinámica, Boltzmann calculó las probabilidades termodinámicas y a partir de ellas se calculan las funciones de la Termodinámica.

Los macroestados pueden ser observados los microestados no, pero se les puede manejar teóricamente y hacer cálculos con ellos. A partir de entonces las Teorías del Calor y de la Termodinámica se pueden desarrollar bajo dos puntos de vista: el fenomenológico que corresponde a los macroestados y sus magnitudes, y el estadístico basado en una concepción atomística y en los microestados no observables. Para seguir este segundo camino si hace falta que los Físicos sepan probabilidades.

A principios del siglo XX entra en crisis la Física Clásica y surge la Mecánica Cuántica, y con ella vuelven a entrar las probabilidades en la Física, y ésta cambia de ser determinista a ser indeterminista; de nuevo los Físicos van a necesitar saber probabilidades. Vamos a aclarar lo anterior: la nueva Mecánica Cuántica admite varias interpretaciones, la más aceptada es la de Copenhague, según la misma sus postulados se pueden resumir brevemente de la siguiente manera:

a) La información sobre una partícula en un punto dado y en un instante dado, está contenida en su función de onda.

b) La función de onda es la solución de una ecuación de ondas.

c) La información que se puede obtener en una experiencia es probabilista.

d) Es imposible medir exactamente dos magnitudes conjugadas (posición y momento). Son las relaciones de incertidumbre de Heisenberg.

e) Toda medida realizada cambia la función de onda.

f) La evolución de la función de onda entre dos medidas está bien determinada.

De modo que por © hay probabilidad en las medidas efectuadas y por (f) hay determinismo en la evolución de la función de onda entre dos medidas.

El apartado © nos indica que solo se puede calcular la probabilidad de encontrar la partícula en un punto dado y en un instante dado (principio de localización); y solo se pueden calcular las probabilidades de las que las magnitudes mecánicas (como el momento) tengan un valor (principio de la descomposición espectral).

Al apartado (e) se le conoce con el nombre de reducción o contracción del paquete de ondas.

Al nacer la Mecánica Cuántica se le plantea a los científicos lo que pudiéramos llamar una “regla de tres científica”; hay que encontrar una Mecánica Estadística Cuántica que sea a la Mecánica Cuántica, lo que la Mecánica Estadística clásica es a la Mecánica clásica.

La solución de esta “regla de tres científica” no es una, sino doble porque hay dos Mecánicas Estadísticas Cuánticas que son la de Bose-Einstein y la de Fermi-Dirac; ello es debido a que al rehacer el cálculo de las probabilidades termodinámicas, hay que hacerlo sobre nuevas bases.

En la Mecánica Estadística Clásica, el enorme número de partículas que integran un sistema, son distinguibles entre sí, se les puede numerar, poner un nombre y seguirlas en su evolución, aunque sea con el pensamiento.

En las Mecánicas Estadísticas Cuánticas, las partículas de un sistema no pueden distinguirse unas de otras (principio de indiscernibilidad). Este principio rige para todas las partículas.

Hay un segundo principio, el de exclusión debido a Pauli (1900-1958) que solo rige para ciertas partículas (los fermiones) según el cual dos partículas iguales no pueden tener todos sus números cuánticos iguales.

En resumen en la Mecánica Estadística clásica no rige ninguno de los principios anteriores. En la Mecánica Estadística cuántica para los basones solo rige el primer principio, es la de Bose-Einstein, y para los fermiones rigen los dos principios es la de Fermi-Dirac.

Un movimiento aleatorio importante, en el que se conjugan las leyes de la Mecánica clásica y las de las probabilidades, es el movimiento browniano, en el que lo que evoluciona en el tiempo es una densidad de probabilidad y no una función de onda como en la Mecánica cuántica. La evolución está regida por una ecuación como la de propagación del calor y la densidad de probabilidad es la de la ley de Gauss (ley normal) como la (13) en la que la varianza σ^2 crece linealmente con el tiempo.

El movimiento browniano fue observado por primera vez en 1827 por Brown (1773-1858) como un movimiento aleatorio de granos de polen suspendidos en el agua; en general es el de partículas sólidas muy finas en suspensión en un líquido. En 1908 el duque Mauricio de Broglie (1875-1960), hermano del premio Nobel Luis fundador de la Mecánica ondulatoria, observó este movimiento en partículas sólidas muy finas en suspensión en un gas. Perrin (1870-1942) obtuvo el premio Nobel, entre otros trabajos por sus experiencias sobre el movimiento browniano el cual ha sido considerado como una de las pruebas indirectas más importantes de la existencia de las moléculas y con ello contribuyó al triunfo final del atomismo sobre la energética en Física, a principios del siglo XX. En la explicación teórica del mismo es fundamental un artículo de Einstein de 1905. Cuando se descubrió no se sabía a que era debido y hoy se sabe que es debido a la agitación térmica de los moléculas del líquido o gas, que colisionan con las partículas sólidas en suspensión.

En un libro "Líneas de investigación en los procesos estocásticos y el movimiento browniano" editado en 1975 por el Instituto de España. Me he ocupado del movimiento browniano generalizado y de otros temas relacionados como los paseos al azar, y la adición, proyección y teoremas centrales del límite de vectores aleatorios isótropos en los espacios euclídeos de cualquier número de dimensiones y en el espacio de Hilbert.

En resumen he intentado escribir una breve exposición de lo que han sido y son las Matemáticas del azar, desde los comienzos modestos del Cálculo de Probabilidades en los siglos XV y XVI, su elevación a la categoría de Ciencia en el XVII, su desarrollado acelerado en los XVIII y XIX sobre todo en este último, hasta su gran esplendor en el XX, en el que se consigue ya su axiomatización rigurosa y se amplia enormemente su campo como Ciencia pura y aplicada. La Estadística Matemática basada en el Cálculo de Probabilidades se inicia a fines del XIX y en el XX obtiene un gran desarrollo. También en el siglo XX nacen nuevas ciencias muy relacionadas con las probabilidades, que antes eran desconocidas, como son el Control de Calidad, la Investigación Operativa y la Teoría de la Fiabilidad.

Las Matemáticas del azar entran de lleno en la Física a fines del siglo XIX con la Mecánica Estadística clásica, y en el XX con la Mecánica Cuántica, las Mecánicas Estadísticas Cuánticas, el Movimiento Browniano y lo procesos estocásticos de la radiactividad, la radiación cósmica y la Física Nuclear adquieren una gran importancia las probabilidades.

También en el siglo XX nace una nueva Ciencia relacionada con las probabilidades que es la Teoría de la Información que introduce una nueva Entropía con un parecido a la antigua de la Física del XIX que se mide en unas nuevas unidades los bits.

El azar en la Economía entra en el siglo XX con la moderna teoría de los juegos de estrategia.

En la Biología a fines del XIX se inicia la Biometría y se enuncian las leyes de la herencia de Mendel (1822-1884) que se expresan en lenguaje probabilista. En el XX la penetración de las Matemáticas del azar es muy profunda en muchas partes de la Biología, que se hallan ya enormemente matematizadas; en los procesos estocásticos de natalidad, mortalidad, inmigración, propagación de enfermedades (curaciones y contagio), en la Genética cuantitativa, la selección natural y la lucha por la existencia, la Ecología, etc.; en general lo que constituye la moderna Dinámica de Poblaciones.

Para concluir quiero destacar un hecho que me parece curioso e interesante y que no he visto referencias del mismo en muchos textos que he consultado. Es el siguiente: en 1900 en el 2º Congreso Internacional de Matemáticas de París, Hilbert quizás el matemático más importante de su época, pronunció una conferencia de más de treinta páginas en el que enunciaba y analizaba una lista de 23 problemas abiertos de las Matemáticas, que él y otros muchos científicos estuvieron de acuerdo con él, consideró esenciales. Con posterioridad algunos de estos problemas han sido resueltos total o parcialmente, positiva o negativamente, y siempre las soluciones obtenidas se consideraron logros muy importantes de las Matemáticas. Si se lee la lista de los 23 problemas en ninguno se hace referencia explícita a las Probabilidades. Es claro que ya entonces era esencial la axiomatización de las mismas, por lo que extraña la falta de mención explícita de este problema. El problema sexto es "¿se pueden matematizar los axiomas de la Física?"; con posterioridad se han axiomatizado algunas partes de la Física. Al analizar este sexto problema se hace mención de pasada a las Probabilidades. A mi me parece que Hilbert debería de haber agregado un problema nº 24 que fuese "axiomatizar las Probabilidades", lo que conseguiría Kolmogorov en los años 1929 á 1933, aunque como ya he dicho antes las que he llamado variables aleatorias fractales escapan a la axiomática de Kolmogorov.

Para el lector especializado que esté interesado por los aspectos matemáticos incluyo unas notas entresacadas de mis investigaciones publicadas y algunas aún inéditas. Su lectura no es necesaria.

Notas 1.

TEOREMA DEL LÍMITE DE UNA VARIABLE ALEATORIA CUYA PROBABILIDAD DE VALER INFINITO NO ES NULA

Sea una variable aleatorio (v.a) ξ, cuya probabilidad de valer infinito es c, y de ser una v.a. de función característica (fc) φ (t) 1-c. Sea ξ_1 una v.a. cuya probabilidad de valer infinito es c_1, y de ser una v.a. de fcφ(t) es $1-c_1$ entonces la suma de las dos v.a. ξ y ξ_1 tiene la probabilidad (1-c) $(1-c_1)$ de ser una v.a. de fc$\varphi(t)^2$ y la complementaria $1-(1-c)(1-c_2)$ de valer infinito.

Si $\xi_1,..,\xi_n$, son v.a. que tienen las probabilidades $c_1,\ldots,c_n$ de valor infinito y las probabilidades de ser v.a. de fc$\varphi(t)1-c_1,\ldots,1-c_n$, entonces la v.a. η:

$$\eta = \frac{\xi_1 + \ldots + \xi_n}{n} \tag{1}$$

supuestas las ξ independientes tiene la probabilidades

$$(1-c_1)\ldots(1-c_n) \tag{2}$$

de tener la fc

$$\varphi\left(\frac{t}{n}\right)^n \tag{3}$$

y la probabilidad complementaria

$$1-(1-c_1)\ldots(1-c_n)=a \tag{4}$$

de valer infinito.

Si ahora hacemos tender n a infinito y es convergente el producto infinito

$$\prod_{k=1}^{\infty}(1-c_k) \tag{5}$$

entonces la v.a. η tiene la probabilidad (5) de converger en probabilidad al valor medio m de η(t) y la probabilidad complementaria de valer infinito; si pertenecen la ξ al dominio de atracción de la ley normal.

Si las $\xi_1,\ldots,\xi_n$ son de valor medio nulo, entonces la v.a. χ:

$$\chi = \frac{\xi_1 + \ldots + \xi_n}{\sqrt{n}} \tag{6}$$

cuando n tiende a infinito, tiene la probabilidad (5) de ser una v.a. normal (gaussiana) de varianza igual a ξ y valor medio nulo, y la probabilidad complementaria de valer infinito.

Nota 2.

LA LEY DE LOS SUCESOS ALEATORIOS RAROS. GENERALIZACIÓN DE LA LEY DE POISSON

La distribución de Poisson resulta de la distribución de Bernoulli cuando la probabilidad de realización de un suceso tiende a cero y el número de pruebas, tiende a infinito, de modo que el producto de la anterior probabilidad por el número de pruebas, tiene un límite.

La distribución de Poisson se conoce también como ley de los sucesos raros. Vamos a obtener una nueva distribución de probabilidad que la generaliza y que podemos calificar de ley de los sucesos aleatorios raros, se distingue de la de Poisson porque ésta es una ley de los sucesos ciertos raros.

Sea una v.a. que tiene una probabilidad λ muy próxima a 1 de valer cero y la probabilidad complementaria 1-λ de ser una v.a. de fcφ(t). La fc$\theta_1(t)$ de ξ es:

$$\theta_1(1) = 1 - \lambda + \lambda\varphi(t) = 1 + \lambda(\varphi(t) - 1) \qquad (1)$$

La suma de n $\phi_{\text{v.a.}}$ $\xi_1, \ldots, \xi_n$ de fc(1) e independientes tiene la fc$\theta_n(t)$:

$$\theta_n(t) = \left(1 + \lambda(\varphi(t) - 1)\right)^n \qquad (2)$$

Si hacemos tender λ a cero y n a infinito de modo que

$$\lim_{n\to\infty} \lambda n = a \qquad (3)$$

entonces es

$$\theta_\infty(t) = \lim_{n\to\infty} \theta_n(t) = e^{a(\varphi(t)-1)} \qquad (4)$$

Si la comparamos con la de Poisson de fc

$$\psi(t) = e^{a(e^{it}-1)} \qquad (5)$$

observamos que (4) resulta de hacer en (5) la sustitución:

$$e^{it} \to \varphi(t) \qquad (6)$$

en la que e^{it} es la fc del número cierto 1 y φ(t) la fc de una v.a.; de aquí las denominaciones que proponemos de llamar leyes de los sucesos raros ciertos y de los sucesos raros aleatorios a (5) y a (4). La (5) es un caso particular de la (4).

Se tiene que

$$\theta_\infty = \varphi\left(\frac{\log \varphi(t)}{i}\right) \qquad (7)$$

que muestra que la v.a. de fc (4) es la suma de un número aleatorio de fc(5) de v.a. independientes ξ de fc φ(t) con el convenio de que si el número de sumandos es cero, la suma vale cero. El problema de la adición de v.a. en número aleatorio la obtuve en los años cincuenta.

Notas 3.

EL MOVIMIENTO BROWNIANO CONSECUENCIA DE UN TEOREMA CENTRAL DEL LÍMITE. UN EXTRAÑO MOVIMIENTO EN EL QUE EL MÓVIL NO CAMBIA NUNCA DE POSICIÓN

Un vector aleatorio isótropo es un vector cuyo módulo es una v.a. de función de densidad de probabilidad o de frecuencia (ff) f(r) siendo r el módulo del vector (que puede ser constantes) y cuya dirección está repartida uniformemente al azar. Puede estar en un espacio euclídeo de p dimensiones E_p

En memorias publicadas en la Revista de la Real Academia de Ciencias en los años cincuenta, planteé y resolví las ecuaciones integrales que ligan entre sí su ff: f_p (r) y la fcφ_q (z) de la proyección del vector sobre un subespacio euclídeo E_q de E_p (q$\angle$p), que puede ser una recta E_1. Estas ecuaciones permiten calcular una de las φ_q (z) o f_p (r), conocida la otra.

Ello me ha permitido resolver el problema de la adición de n vectores aleatorios isótropos de E_p, calculando la φ_1(z) de la proyección del vector sobre la recta E_1 y a continuación calcular la f_p (r) que corresponde a la $\varphi_1(z)^n$. De ello se sigue que la proyección de un movimiento browniano de E_p sobre E_q es también un movimiento browniano, y que todo movimiento browniano sobre E_q es la proyección sobre E_q de un movimiento browniano sobre E_p.

Si un punto se mueve sobre una recta R_1 dando un paso de longitud l hacia la derecha o la izquierda con la misma probabilidad 1/2, su desplazamiento es una v.a. de fc:

$$\varphi(z) = \frac{1}{2}\left(e^{izl} + e^{-izl}\right) = \cos zl = 1 - \frac{z^2 l^2}{2} + \ldots \qquad (1)$$

Si da n pasos en la unidad de tiempo, la fc de un desplazamiento en el tiempo t es:

$$\varphi_n(z) = (\cos zl)^{nt} \tag{2}$$

si ahora hacemos tender l a cero y n a infinito de modo que

$$\lim_{n\to\infty} nl^2 = \sigma^2 \tag{3}$$

entonces

$$\lim_{n\to\infty} (\cos zl)^{nt} = e^{-\sigma^2 tz^2/2} \tag{4}$$

que es la fc de una v.a. gaussiana (normal) de valor medio cero y varianza $\sigma^2 t$; es un movimiento browniano.

El camino recorrido en la recta E_1 en el tiempo t es nlt que es infinito por (3). De modo que el móvil tiene una velocidad infinita y en cualquier intervalo de tiempo t, recorre un camino infinito, pero su desplazamiento sobre la recta R_1 es una v.a. gaussiana finita de varianza $\sigma^2 t$. El valor medio del desplazamiento es

$$\sigma\sqrt{\frac{2t}{\pi}} \tag{5}$$

Los caminos recorridos hacia la derecha y la izquierda son infinitos pero su diferencia (el desplazamiento real) es finito.

Si en vez de cumplirse la (3) se cumple la:

$$\lim_{n\to\infty} nl = m \tag{6}$$

entonces es

$$\lim_{n\to\infty} \varphi_n(z)^n = 1 \tag{7}$$

el móvil por (7) no se desplaza durante todo el movimiento, cualquiera que sea el tiempo transcurrido, a pesar de que en un tiempo t tiene una velocidad m, y recorre un camino mt, debido a que la mitad del camino lo recorre a la derecha y la otra mitad a la izquierda, cualquiera que sea el valor de t; y hay infinitos cambios de sentido en el movimiento durante el tiempo t, se obtiene un extraño movimiento aleatorio en que el móvil está constantemente en movimiento nunca cambia de posición, está aparentemente en reposo, pero sin embargo se mueve.

Recordando los versos de un famoso poeta español Machado "Caminante no hay camino, se haces camino al andar" en este extraño movimiento podríamos decir "Caminante no haces camino aunque no paras de andar". Parece una paradoja no desplazarse de su posición inicial y sin embargo moverse.

Si el móvil se mueve en E_p, dando n pasos de longitud l en la unidad de tiempos, repartidos uniformemente al azar sus direcciones, si se cumple la (3) también se obtiene un movimiento browniano.

Se obtiene también un movimiento browniano si se sustituye la (1) y la (2) por la

$$(\psi_1(z))^n = \left(1 - \frac{\mu_2 l^2}{2} + \frac{\mu_4 z^4 l^4}{4!} + \dots\right)^n \tag{8}$$

cuando l tiende a cero y n a infinito cumpliéndose la (2).

Notas 4.

VARIABLES ALEATORIAS CONJUGADAS. UNA RELACIÓN DE INCERTIDUMBRE FUERTE

En esta nota vamos a exponer un caso particular de una teoría general que hemos desarrollado para v.a. cualesquiera, aunque no pertenezcan al dominio de atracción de la ley normal. Nos vamos a limitar a v.a. gaussianas (normales).

Sea $f_1(x_1)$ la ff de una v.a. normal ξ_1 de valor medionulo y varianza σ_1^2. La fórmula (1), que es una transformada de Fourier:

$$\sqrt{f_2(x_2)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ix_1x_2} \sqrt{f_1(x_1)}\, dx_1 \tag{1}$$

define una función $f_2(x_2)$ que es la ff de una v.a. normal ξ_2 de valor medio 0 y varianza σ_2^2 porque:

$$\sqrt{f_1(x_1)} = \frac{1}{\sigma_1^{1/2}(2\pi)^{1/4}} e^{-x_1^2/4\sigma_1^2} \tag{2}$$

con lo que (1) se escribe:

$$\sqrt{f_2(x_2)} = \frac{\sigma_1^{1/2}}{(2\pi)^{1/4}} \int_{-\infty}^{\infty} \frac{e^{-x_1^2/4\sigma_1^2}}{\sigma_1\sqrt{2\pi}} e^{ix_1x_2} dx_1 =$$

$$\frac{(2\sigma_1)^{1/2}}{(2\pi)^{1/4}}\int_{-\infty}^{\infty}\frac{e^{-x_1^2/4\sigma_1^2}}{\sqrt{2}\sigma_1\sqrt{2\pi}}e^{ix_1x_2}dx_1=\frac{(2\sigma_1)^{1/2}}{(2\pi)^{1/4}}e^{-x_2^2 2\sigma_1^2} \quad (3)$$

y de aquí

$$f_2(x_2)=\frac{2\sigma_1}{\sqrt{2\pi}}e^{-x_2^2 4\sigma_1^2/2} \quad (4)$$

que es una v.a. normal de valor medio nulo y varianza σ_2^2:

$$\sigma_2^2=\frac{1}{4\sigma_1^2}\Rightarrow\sigma_1\sigma_2=1/2 \quad (5)$$

La (5) es una relación de incertidumbre que llamamos fuerte porque se expresa por una igualdad, a diferencia de las relaciones de incertidumbre de Heisenberg que se expresan por desigualdades y por eso las llamamos débiles.

A ξ_1 y ξ_2 les llamamos v.a. conjugadas y gozan de la propiedad de que cuando mayor es la varianza de una, menor es la varianza de la otra, de modo que si σ_2 vale infinito (o cero) entonces σ_1 vale cero (o infinito) que implica que si ξ_1 es una v.a. repartida uniformemente al azar sobre una recta (o es el número cierto cero) a ξ_2 le sucede lo contrario y viceversa.

A una v.a. ξ_1 no sólo le corresponde una v.a. conjugada, porque si en (1) efectuamos la sustitución

$$\sqrt{f_1(x_1)}\to e^{ix_1a}\sqrt{f_1(x_1)} \quad (6)$$

siendo a un número real cualquiera, entonces el primer miembro de (1) valer

$$\sqrt{f_2(x_2-a)} \quad (7)$$

y ξ_2 es una v.a. gaussiana (normal) de valor medio a y varianza σ_2^2. Por lo que en el párrafo anterior a la (6) se puede sustituir "el número cierto cero" en el paréntesis con "el número cierto a", Obsérvese que

$$\left|e^{ix_1a}\sqrt{f_1(x)}\right|^2=f_1(x) \quad (8)$$

como la (1) es invertible el papel de $f_1(x_1)$ y $f_2(x_2)$ es simétrico.

En la última (5) si las dimensiones de σ_1 y σ_2 son d_1 y d_2, la dimensión de 1/2 es la del producto d_1d_2. Por ejemplo si las dimensiones de σ_1 y σ_2 son L, la de 1/2 es L^2; y si la de σ_1 es una longitud (L) y la de σ_2 un momento (cantidad de movimiento) MLT^{-1}, la de 1/2 es ML^2T que son las de una acción (energía por tiempo).

Los resultados de las notas, 3ª, 4ª y 5ª, como ya lo dijimos en la 3ª son generalizables a v.a. de cualquier número de dimensiones.

Notas 5.

EL MOVIMIENTO ANTIBROWNIANO Y EL PRINCIPIO ANTIERGÓDICO. FENÓMENOS DE DIFUSIÓN Y DE CONCENTRACIÓN CONJUGADOS

Sea ξ_1 (t) una v.a. normal de valor medio cero y varianza σ_1^2t,t es el tiempo. Su ff y su fc son:

$$f_1(x;t)=\frac{1}{\sigma_1\sqrt{2\pi t}}e^{-x^2/2\sigma_1^2t};\varphi_1(z;t)=e^{-\sigma_1^2tz^2/2} \quad (1)$$

La ff está representada en al figura 4, la ordenada es la primera (1)

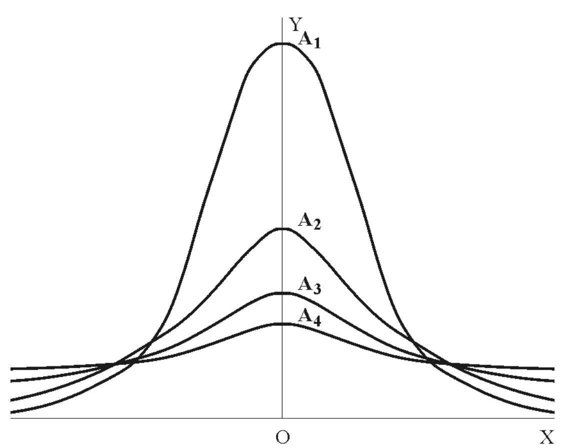

Figura 4

Cuando t = 0, la v.a. ξ (t) vale cero, y cando t = ∞ es una v.a. repartida uniformemente al azar. En los instantes $t_1<t_2<t_3<t_4$ las ordenadas máximas de $f_1(x,t)$ son A_1, A_2, A_3 y A_4. A medida que decrece el tiempo la

curva de la figura 4 se va aplastando sobre el eje OX. La $f_1(x;t)$ gobierna la propagación del calor en una barra, la densidad de probabilidad en el movimiento browniano, y en general un fenómeno de difusión.

Si filmasemos una película de la evolución en el tiempo de $f_1(x,t)$ y luego la pasasemos al revés que es a lo que equivale cambiar de signo el tiempo, entonces en los instantes $t_4 - t_3$, $t_3 - t_2$, $t_2 - t_1$ la curva se iría levantando desde la posición A_4 a la A_1 pasando por A_3 y A_2 hasta A_1 se iría aplastando sobre el eje OY.

Por lo visto en la nota 4ª la v.a. ξ_2 de ff y fc:

$$f_2(x,t) = \frac{\sqrt{t}}{\sigma_2\sqrt{2\pi}} e^{-x^2t/2\sigma_2^2}; \varphi_2(z,t) = e^{-\sigma_2^2 z^2/2t} \qquad (2)$$

y también la ff igual a f_2 $(x\text{-}a_i\ t)$, siendo a un número cualquiera, son v.a. conjugadas de la (2) si se cumple la:

$$\sigma_1\sqrt{t}\,\frac{\sigma_2}{\sqrt{t}} = \sigma_1\sigma_2 = \frac{1}{2} \qquad (3)$$

vese nota 4ª

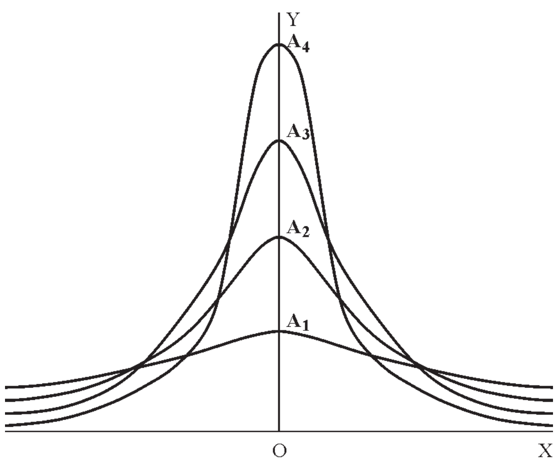

Figura 5

En la figura 5 se ha representado la $f_2(x,t)$. Para $t = \infty$, $\xi_2(t)$ está repartida uniformemente al azar. En los instantes $t_1<t_2<t_3<t_4$ se va aplastando sobre el eje OY, hasta ser $\xi_2(0)$ para $t = 0$, igual a cero.

Esta función $f_2(x,t)$ gobierna un fenómeno de concentración que llamamos conjugado del fenómeno de difusión de la figura (4). Por lo dicho después de la (2), la curva de la figura (5) se puede trasladar paralelamente a OX en el plano XOY, de modo que el origen de coordenadas sea cualquier punto de ordenada cero y abscisa a cualesquiera. En estos fenómenos de concentración rige un principio antiergódico, según el cual cualquiera que sea la distribución inicial $(t = 0)$ de f_2 (x,t), finalmente la ξ_2 (t) converge a un número cierto a.

La rapidez de convergencia en el principio ergódico es tanto mayor cuanto mayor sea σ_1 y en el principio antiergódico cuando menor sea σ_2 en virtud de (3).

Para pasar de (1) a (2) hay que cambiar t por 1/t o lo que es lo mismo ln t por $\ln 1/t = -1/t$, hay que hacer una inversión logarítmica del tiempo.

Notas 6.

FRACTALES FÍSICOS Y MATEMÁTICOS. EXTENSIÓN A LOS FRACTALES DE LAS PROBABILIDADES, LA DERIVACIÓN, LA INTEGRACIÓN, LA MECÁNICA Y LAS ECUACIONES DIFERENCIALES

En 1988 en un artículo publicado en la Revista de la Real Academia de Ciencias publiqué mis primeras investigaciones sobre las fractales, del que una parte está recogida en mi Diccionario de Matemática Moderna (3ª edición de 1994 en la editorial Ra-Ma). En él definí un fractal físico para el que propuse el nombre de barra de Cantor, con un momento de inercia y una masa finitos y no nulos, el cual se puede considerar como el primer ejemplo de péndulo fractal, de longitud nula, formado por un conjunto discontinuo de puntos materiales, con la potencia del continuo. Se describe el proceso de formación del fractal por una sucesión infinita de estados intermedios, cuya masa permanece la misma, mientras que su longitud tiende a cero; el límite de este proceso de fractalización es el fractal. Señalaba que se podrían obtener otros fractales físicos a partir de fractales geométricos conocidos. No tienen representación en la Mecánica clásica, pero se les pueden aplicar las leyes de la Mecánica. Introducía en ellos las probabilidades, definía funciones fractales y la manera de extender a ellas la integración.

En el 2003 concluí un libro todavía inédito de algo más de 250 páginas en ocho capítulos, titulado “Teoría Físico-Matemática de los Fractales” donde reúno mis investigaciones, que resumo a continuación.

En la Mecánica clásica existen dos clases de puntos materiales, los que representan partículas o corpúsculos que son puntos geométricos dotados de masa finita y densidad infinita. Los otros puntos son los que constituyen los sólidos y los sistemas materiales, son puntos materiales sin masa, ni volumen, pero con densidad finita para todo el conjunto.

Por el contrario los fractales mecánicos o físicos, están formados por puntos geométricos, contenidos en un volumen finito, pero con volumen total nulo, es materia que no ocupa lugar en el espacio, pero su masa es nula y su densidad infinita, de modo que el conjunto de puntos que forman el fractal, que tiene la potencia del continuo, si tiene masa. Si existen en la naturaleza tendrían propiedades extrañas, porque estarían sometidas a la ley de la gravitación universal. Si existen, en la mente del matemático, como le sucede a la geometría de n dimensiones.

Al estar dotados de masa, centro de gravedad y momento de inercia, las ecuaciones de la Mecánica del sólido le son aplicables, a las barras, placas y sólidos fractales. La generalización más sencilla es la de los péndulos fractales. Al existir ruedas (placas circulares) fractales y esferas fractales (capítulo 5 y 7) las propiedades de los movimientos de rodadura se extienden también a los fractales. El resumen del libro es el siguiente:

En el Capítulo 1 se extiende el Cálculo de Probabilidades, la Geometría de Masas y la Integración a los fractales numéricos definidos sobre los sistemas de numeración de base cualquiera, los fractales geométricos sobre un segmento, y los fractales físicos que hemos denominado barras fractales. Definimos un nuevo tipo de integrales que llamamos integrales ramificadas, que resultan de la integración sobre un árbol en vez de sobre un segmento o un arco de curva. Extendemos la integración a los conjuntos de medida nula, en los que las integrales son nulas pero el cociente de dos integrales tiene un valor definido. Calculamos los centros de gravedad y los momentos de inercia de las barras fractales.

En el capítulo 2, definimos la que hemos llamado ley de probabilidad de los sucesos aleatorios raros que generaliza la ley de Poisson que se refiere a sucesos raros ciertos. Investigamos las variables aleatorias que tienen probabilidad no nula de valer infinito y damos la forma que toma el teorema central del límite para las mismas. Obtenemos representaciones de funciones trascendentes mediante productos infinitos convergentes, y también la expresión de variables aleatorias mediante sumas infinitas de variables aleatorias discretas en progresión geométrica decreciente. Exponemos los árboles de abscisas, probabilidades y masas, y definimos las integrales simples ramificadas sobre los mismos.

El capítulo 3, trata de los temas anteriores, pero con dos dimensiones; de la teoría de los fractales geométricos y físicos sobre un cuadrado (las placas). Calculamos la función característica de un punto repartido uniformemente al azar sobre varios ejemplo de estos fractales, entre ellos el polvo de Cantor sobre un cuadrado y la carpeta de Sierpinski. Estudiamos las que hemos llamado placas de Cantor y de Sierspinski. Definimos los árboles dobles de coordenadas y los integrales dobles, ramificadas sobre los mismos. Introducimos los pseudofractales, como conjuntos que se obtienen por procesos de fractalización como los fractales, pero que tienen dimensión entera en vez de fraccionaria. Extendemos la teoría a fractales físicos no homogéneos y a variables aleatorias no repartidas uniformemente al azar sobre fractales geométricos. Resolvemos los problemas relativos a distribuciones de probabilidad marginales.

El capítulo 4 extiende las investigaciones de los tres capítulos anteriores al espacio de tres dimensiones primero, y al de cualquier número de dimensiones. Definimos los sólidos fractales, entre los que figuran los que hemos llamado de Cantor y de Sierpinski, que son los fractales físicos sobre el polvo de Cantor en un cubo y sobre la esponja de Sierpinski. Extendemos esta teoría a los espacios de cualquier número de dimensiones. Obtenemos la representación de funciones trascendentes mediante productos infinitos convergentes. En este caso las integrales ramificadas son triples y múltiples.

Hemos propuesto llamar cifras binales, temales, cuaternales y en general pales, a las cifras a la derecha de la coma en los sistemas de numeración de bases 2, 3, 4 y p en general que generalizan los decimales.

El capítulo 5 comienza con un método nuevo para calcular la función característica de un punto repartido uniformemente al azar sobre un triángulo irósceles o equilatero, y su aplicación a la Geometría de Masas. A continuación se extiende el Cálculo de Probabilidades, la Integración y la Geometría de Masas a los fractales sobre los triángulos isósceles y equiláteros y a los estados intermedios en el proceso de fractalización. Se extiende este estudio a los triángulos rectángulos y a los triángulos en general. Se aplican los resultados anteriores a polígonos regulares y al círculo como límite de un polígono regular inscrito o circunscrito, cuando el número de lados tiende a infinito. Se introduce el concepto de fractales de segunda clase como límites de fractales ordinarios, a los que llamamos de primera clase; la rueda fractal es un ejemplo de segunda clase. También se introduce el concepto de fractales equivalentes.

El capítulo 6 extiende estas investigaciones a tetraedros y en particular al tetraedro regular y al tetraedro cuya base es un triángulo equilátero y el vértice con sus tres aristas es un tiedro trirrectangular. A continuación se investigan los fractales sobre el octaedro regular. Se extienden estas investigaciones a los politopos (poliedros en p dimensiones) que hemos llamado p+1-edros, 2^P-edros y 2p-edros (hipercubos) que generalizan el tetraedro el octaedro y el cubo en el espacio euclideo de p dimensiones. Se investiga con mucha amplitud la geometría y la Geometría de Masas de estos politopos en p dimensiones. Esta última parte es muy complicada.

El capítulo 7 está dedicado a la Geometría de cualquier número de dimensiones p, al Cálculo de Probabilidades, la Integración y la Geometría de Masas asociadas. Se hace un estudio exhaustivo de los polítopos anteriores y se investigan fractales sobre hiperesferas de cualquier número de dimensiones.

En el capítulo 8 como los fractales físicos tienen masa, centro de gravedad (cdg) y momentos de inercia, y también frontera geométrica, se pueden extender a ellos las leyes de la Dinámica del sólido. El ejemplo más sencilla es el del péndulo físico fractal, en el que se sustituye la barra ordinaria maciza por una barra fractal, por ejemplo la barra de Cantor (capítulo 1).

Se analizan las variaciones de energía y los posibles movimientos espontáneos en un proceso de fractalización de un sólido; proceso que puede ser reversible y transformar un fractal en un sólido.

Nos preguntamos si existen fractales físicos en la naturaleza o si se pueden producir en el laboratorio, en cuyo caso existiría una fuerza o interacción entre los puntos del fractal que los mantiene confinados en su frontera.

Se introduce el cálculo de cocientes, de integrales de Riemann sobre volúmenes nulos en un fractal y de integrales de Lebesgue en un conjunto de medida nula. Se utilizan estas propiedades para calcular la función de densidad de probabilidad en un fractal, que vale infinito, salvo a lo sumo en un número finito de puntos, lo que les hace inmanejables, pero sin embargo las funciones características existen, son finitas y manejables.

Se analizan las profundas diferencias entre las v.a. ordinarias del Cálculo de Probabilidades y estas nuevas v.a. fractales que son completamente distintas. Las probabilidades fractales no cumplen la axiomática de Kolmogorov, sino otra distinta que describimos.

También se extiende la derivación a los fractales, y como anteriormente se había extendido la integración, quedan definidas las ecuaciones diferenciales sobre un fractal. Resolvemos la más sencilla:

$$\frac{Df(x)}{Dx} = \lambda f(x) \tag{11}$$

donde el primer miembro es la derivada fractal. La función exponencial es la más sencilla del conjunto de soluciones de la (11) que son distintas para fractales distintos.

En el 9 del capítulo 8 se explica como la materia puede filtrarse a través de los fractales físicos.

Notas 7.

LA AXIOMÁTICA DE LAS PROBABILIDADES FRACTALES

Las probabilidades fractales no siguen la axiomática de Kolmogorov, sino otra distinta como dijimos en la nota 5ª.

En el proceso de fractalización en el que E_o es el estado inicial $E_1,\ldots,E_n,\ldots$, los estados intermedios y E_∞ el fractal; E_o es un espacio de probabilidad definido por la terna: E_o, A_o, P_o (a_o) donde A_o es una tribu de partes de E_o (subálgebra de sucesos) cuyos elementos son los sucesos y entre ellos figura $E_1,..,E_n,...$;P_o (a_o) es la probabilidad de un suceso $a_o \in A_o$. Las probabilidades son las aplicaciones de A_o en el intervalo [0,1].

Los estados intermedios $E_1,\ldots,E_n,\ldots$, son espacios de probabilidad, el E_n se define por la terna:E_n, A_n, P_n, donde n es cualquier número entero. La terna se define de la siguiente manera: E_n es el conjunto E_n, A_n es la tribu de E_n formada por las intersecciones de los elementos de A_o con E_n, que denotamos por

$$A_n = A_0 \cap E_n \tag{1}$$

de modo que a_n es la intersección de a_o con E_n. La probabilidad P_n (a_n) vale:

$$a_n \in A_n; P_n(a_n) = \frac{P_o(a_n)}{P_o(E_n)}; a_n = a_o \cap E_n \tag{2}$$

Se cumple que:

$$\lim_{n\to\infty} P_o(E_n) = 0 \Rightarrow \lim P_o(a_n) = 0 \tag{3}$$

porque las medidas de a_n y de E_n tienden a cero, cuando n tiende a infinito. La (2) es una expresión indeterminada de la forma 0/0 y si tiene un valor verdadero (levantando la indeterminación) entonces

$$\lim_{n\to\infty} P_n(a_n) = P_\infty(a_\infty) \neq 0 \tag{4}$$

y la terna E_∞, A_∞, P_∞, define un espacio de probabilidad, donde A_∞ es la tribu $A_o \cap E_\infty$ y P_∞ es igual a (4).

Esta axiomática es la que cumple la v.a. fractal definida al final de la nota 6ª.

BIBLIOGRAFÍA

1. Borel: "El Azar". Montaner y Simón, 1935.
2. Haigh: "Matemáticas y juegos de azar" Tusquets 2003.
3. Bell: "Historia de las Matemáticas". Editorial Efe. 1949.
4. Bourbaki: "Elementos de Historia de las Matemáticas". Alianza. Universidad 1976.
5. Morris Kline: "El pensamiento matemático de la antiguedad a nuestros días". Tres Tomos. Alianza Universidad 1992.
6. Gray: "El reto de Hilbert". Editorial Crítica 2003.
7. Feller: "Introducción a la teoría de probabilidades y sus aplicaciones" dos tomos, Limusa 1973 y 1978. Méjico.
8. Von Mises: "Probabilidad, Estadística y Verdad". Espasa Calpe, 1946.

Y mis libros:

9. Filosofía de las Matemáticas
10. Teoría de la Investigación Matemática
11. Didáctica y Dialéctica Matemáticas editados por Dossat en 1961, 1966 y 1969.
12. Cálculo de Probabilidades y Procesos Estocásticos!" editado por Paraninfo en 1974.
13. Diccionario de Matemática Moderna editado por Ra-Ma, 1994, 3ª edición.
14. Líneas de Investigación en los Procesos Estocásticos y el Movimiento Browniano editado por el Instituto de España en 1975.
15. Teoría físico-matemática de los fractales.