

## LAS MATEMÁTICAS EN LAS CIENCIAS SOCIALES

*Daniel Peña*

*Universidad Carlos III de Madrid*

### INTRODUCCIÓN

En el mundo en continua y rápida transformación en que vivimos las *partes descriptivas* de las disciplinas en la *Ciencias Sociales* están perdiendo peso frente a las *partes normativas*. Por ejemplo, en la enseñanza, cada vez tiene menos interés poner el énfasis en describir con detalle una realidad social, por ejemplo la estructura económica de un país, que sabemos va a modificarse en poco tiempo en profundidad, que enseñar principios generales de *análisis social* que puedan aplicarse en contextos muy distintos. En lugar de memorizar datos y hechos conviene transmitir métodos de aprendizaje, o, utilizando una frase popular, *enseñar a aprender*.

Las Matemáticas en las Ciencias Sociales, como en otros campos científicos, pueden ser la herramienta fundamental para adquirir y consolidar el conocimiento. Hay tres razones principales que justifican este carácter singular.

- (i) En primer lugar, las *matemáticas obligan a definir claramente las variables de interés en cada problema*, a *establecer las hipótesis* sobre su comportamiento y a *definir las relaciones* entre ellas.
- (ii) En segundo lugar, el lenguaje matemático permite *importar* a las Ciencias Sociales *modelos de relación entre variables* que han tenido éxito en otras ciencias, ofreciendo nuevas posibilidades de explicación de los fenómenos sociales y enriqueciendo el conjunto de modelos disponibles para investigar la realidad social.
- (iii) En tercer lugar, la creciente disponibilidad de datos, debido a la difusión de los ordenadores y la automatización en todas las actividades humanas, permite contrastar con mayor rigor los modelos sociales en la práctica mediante los *métodos estadísticos* y *generar predicciones y reglas de comportamiento* verificables con los datos.

*Los modelos matemáticos aportan el lenguaje y la estructura conceptual necesaria para expresar reglas generales de comportamiento y obtener predicciones de validez general.* Su utilización facilita que los conocimientos adquiridos en las investigaciones sociales puedan transmitirse con precisión, estimulando la comunicación entre investigadores de distintas áreas. Por ejemplo, los mismos modelos de crecimiento se han utilizado con éxito para explicar la evolución de una economía, de un tumor, de la población de un país o de la concentración de un reactivo en un proceso químico. Además, los contrastes estadísticos de adecuación entre el modelo y los datos pueden sugerir como modificarlo para adaptarlo mejor a las peculiaridades del problema a analizar. De esta manera, se consigue combinar la unidad de método, por la permanencia de los principios matemáticos, con la adaptación al problema concreto. La revisión de las hipótesis del modelo al disponer de nuevos datos permite mejorarlos y hacerlos más flexibles, más dinámicos y con mayor complejidad.

Podríamos esperar que cuanto mayor sea la posibilidad en una disciplina de medir sus variables de interés y de establecer reglas generales sobre las relaciones entre ellas, mayor será su contenido matemático. Esta progresión comenzó a mediados del siglo XVII en la Física, a finales del XIX en la Biología, a principios del XX en la Economía, y en la segunda mitad del siglo XX se ha ido

estableciendo en Sociología y Psicología, y pensamos que en el siglo XXI lo hará en la Historia, la Lingüística, y en general, en las Humanidades.

En las *Ciencias Físicas* las variables clásicas (movimiento, trabajo y energía) pueden medirse con precisión y sus relaciones son, principalmente, deterministas. Próximas a ellas se sitúan las *Ciencias Tecnológicas*, que complementan los modelos físicos añadiendo el error de medida, debido a la presencia de factores incontrolables. Esta mayor incertidumbre hace necesario el uso de la Estadística para estimar las relaciones entre las variables, y *Gauss*, a principios del siglo XIX, inventó la distribución normal para los errores y el método de mínimos cuadrados para estimar las relaciones entre variables astronómicas obtenidas mediante telescopios. Sin embargo, el componente aleatorio tiene en muchas aplicaciones tecnológicas un efecto pequeño. Por ejemplo la figura 1a) presenta la relación entre el consumo de energía y el número de unidades producidas en un proceso de fabricación. Aunque hay una pequeña variabilidad, la relación puede suponerse determinista y representarse bien por una parábola.

En las *Ciencias Naturales* y de la vida el componente estocástico de las relaciones entre las variables no puede casi nunca despreciarse. Este fenómeno fue comprendido por *Galton* a finales del siglo XIX, al estudiar la transmisión de características genéticas entre padres e hijos. Galton estudió la relación entre la estatura de los padres y sus hijos y encontró la relación que se presenta en la figura 1b): los padres altos tienen hijos en promedio altos, y los bajos en promedio bajos, pero no podemos decir nada del descendiente de un padre en particular. En esta relación se observa un fenómeno de *regresión a la media*, que es lo que mantiene unidas a las poblaciones biológicas: padres con estatura lejos de la media tienen descendientes con estaturas promedio más cerca de la media. Por esta razón se denomina de *regresión* a estas relaciones que no nos permiten afirmar con certeza nada respecto a una situación particular, sino que establecen lo que ocurrirá por término medio en un conjunto dado de situaciones similares.

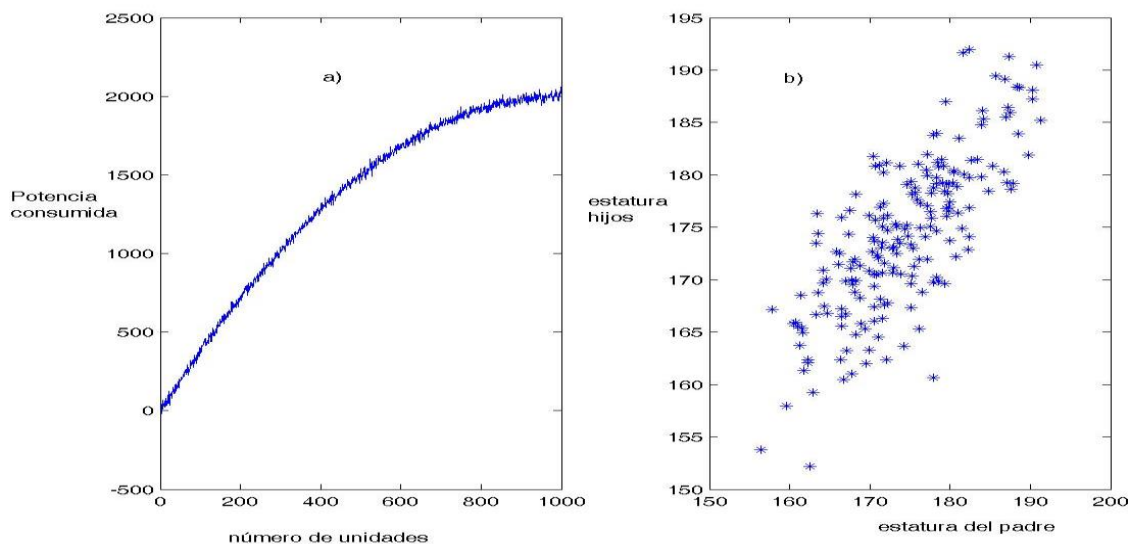


Figura 1. En el caso 1a) la relación puede suponerse determinista, mientras que en el caso 1b) es estadística, o de regresión.

En las *Ciencias Sociales* prácticamente la totalidad de las relaciones con las que nos encontramos en un libro de *Economía*, o *Sociología* son relaciones estadísticas, o relaciones de regresión o en promedio. Por ejemplo, si decimos que el aumento del precio de un producto hará disminuir su demanda si todas las demás variables permanecen constantes (*ceteris paribus*), estamos haciendo una afirmación que nunca podremos verificar en la práctica, porque nunca en la vida social dos situaciones van a ser idénticas: siempre habrá en la vida social muchas variables que pueden tener efecto en la relación y que no podemos controlar y, con frecuencia, ni incluso medir. Para que sea

operativa esta regla debe de reformularse como una relación estadística: no esperamos que vaya a cumplirse siempre, sino sólo en promedio, como en la figura 1b). Por ejemplo, en la bolsa tenemos ejemplos constantes de que el aumento del precio de una acción puede aumentar su demanda y se ha comprobado empíricamente en muchos casos que aumentos o disminuciones en el precio de un bien pueden no tener efectos apreciables sobre su demanda.

Estas características son todavía más extremas en la Sociología o la *Ciencia Política*. Cuando la Sociología establece una pauta de comportamiento en una clase de individuos no esperamos que esta pauta se aplique a cada individuo, sino esperamos que ocurra en promedio. Por ejemplo, si decimos que los miembros de la clase obrera tenderán más a votar a partidos de izquierda de nuevo sólo podemos aplicar esta regla en promedio. La situación es más extrema en las *Humanidades*, por su carácter más descriptivo que normativo, aunque la tendencia a la medición y sistematización es creciente en disciplinas como la *Historia*, la *Lingüística*, e incluso en disciplinas como la *Ética*, donde hay autores que defienden una concepción estadística de las normas éticas, transformado en parte la discusión ideológica de valores en un problema de medición (véase *Mosterín*, 2005).

El resto de este trabajo se estructura como sigue:

En primer lugar comentamos el papel de las Matemáticas en las ciencias económicas, incluyendo la Economía y la Administración de empresas. A continuación, comentamos el papel de la Estadística en la Sociología y la Ciencia Política. Finalmente, haremos unas breves reflexiones sobre el papel de la matemática en las Humanidades.

## MATEMÁTICAS Y CIENCIAS ECONÓMICAS

El núcleo central de la Economía moderna es la *Teoría Económica*, que consiste en la explicación mediante una firme base matemática de los fenómenos económicos. Consideramos uno de los primeros modelos que se introducen en un libro de economía: la formación del precio de un bien como punto de equilibrio entre las curvas de oferta y demanda. Se supone que la cantidad de un bien que los consumidores están dispuestos a adquirir depende del precio, según una *curva de demanda*, y que la cantidad que los productores ofertan es también función del precio, según la *curva de oferta*. El punto de encuentro o de equilibrio entre ambas curvas (véase la figura 2) es el precio. Si consideramos un mundo estático, la existencia de un *punto de equilibrio* en este tipo de situaciones es central en la teoría económica.

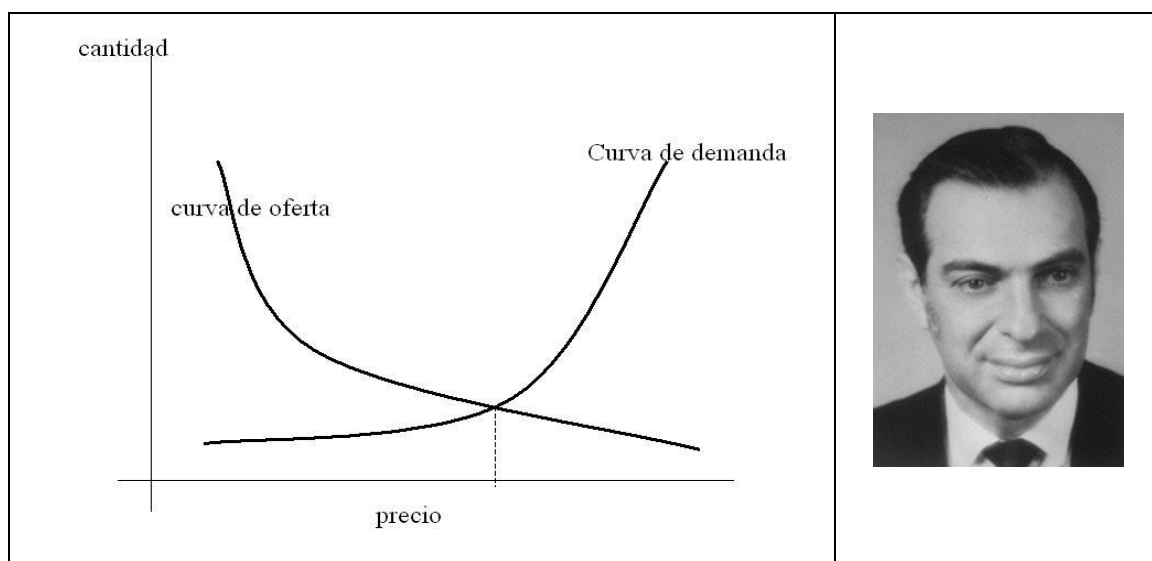


Figura 2. La formación del precio como equilibrio entre oferta y demanda y el premio Nobel de Economía *Kenneth J. Arrow*

La existencia y caracterización del equilibrio es un problema matemático de interés que ha sido extensamente estudiado en la Economía Matemática. Si pensamos en el conjunto de bienes de una economía, aparece el problema del equilibrio general, donde la situación es más compleja al existir la posibilidad de sustituir unos bienes por otros. La *teoría del equilibrio general* fue planteada matemáticamente por *León Walras* (1834-1910), profesor de Economía en Lausana, que conjeturó la existencia de equilibrio en la economía y luchó infructuosamente por convencer a los economistas de las ventajas de un enfoque matemático para este problema. La importancia de su aportación comenzó a ser reconocida varios años después de su muerte y *Gerard Debreu* (1921-2004), matemático y economista francés, obtuvo el Premio Nobel de Economía en 1983 por su rigurosa reformulación de la teoría del equilibrio general, generalizado y ampliando el enfoque de Walras.

Debreu trabajó con *Arrow*, que había obtenido previamente el Premio Nobel de Economía en 1972 por su análisis de los *problemas de elección social*. Arrow comenzó su tesis en Estadística con *Hotelling* en Columbia, pero luego la reorientó hacia la decisión social. En su tesis doctoral en 1951 demostró su famoso *teorema de imposibilidad*, que establece que no es posible construir una función de bienestar social congruente con las preferencias individuales. Es decir, si suponemos que los individuos tienen una función coherente de preferencias, no es posible establecer una función social de preferencias que tenga en cuenta las preferencias de más de una persona y que sea a su vez coherente con las funciones individuales, en el sentido de que si todos los individuos de la sociedad prefieren A a B entonces la sociedad también debe preferir A a B.

El análisis del equilibrio y los problemas de las *preferencias o utilidades* están presentes también en la que es probablemente la mayor contribución a la matemática proveniente de la necesidad de resolver un problema económico: La *teoría de juegos*. Fue desarrollada por *Morgenstern y John Von Neuman* (1944) y supuso un nuevo planteamiento de las decisiones económicas donde los agentes, en lugar de reaccionar a unos precios que se fijan externamente, eligen su estrategia en un mundo competitivo frente a otros agentes económicos. Los autores argumentaban que la teoría de juegos podía ser para la Economía lo que el Cálculo fue para la Física. En su libro encontraron la solución de equilibrio, es decir la situación donde ninguno de los jugadores tiene nada que ganar por cambiar de estrategia, para juegos entre dos jugadores *de suma cero*, donde lo que gana uno necesariamente lo pierde el otro.

La teoría de juegos tiene un avance fundamental por el trabajo del matemático *John Nash* (1950), que estudió *juegos cooperativos* de suma no nula y demostró la existencia de equilibrio en estos juegos permitiendo *estrategias mixtas*, es decir, donde cada jugador selecciona su decisión con una probabilidad dada. Esta situación se ilustra en la Figura 3. En el juego 1 la matriz representa los beneficios para cada jugador de tomar la estrategia A o B. El jugador 1 puede tomar la estrategia A1, y obtener 3 ó 0, ó la B1 y obtener 0 ó 1. En este juego existen dos puntos de equilibrio, el (A1, A2) y el (B1, B2), ya que una vez llegado a estos puntos cambiar la estrategia no supone ninguna ventaja para los jugadores. Sin embargo, en el juego 2 no existe ningún punto de este tipo: en cualquier casilla de la tabla hay un incentivo para cambiar de estrategia para alguno de los dos jugadores. Sin embargo, puede demostrarse que en estos juegos existe siempre una estrategia aleatoria que lleva al equilibrio.

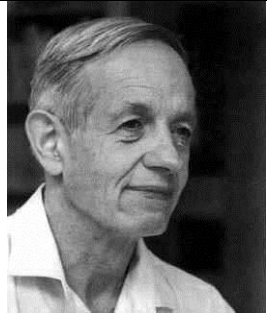
		A2	B2			A2	B2
	A1	3,1	0,0		A1	0,0	0,-1
	B1	0,0	1,2		B1	1,0	-1,3
		Juego 1			Juego 2		

Figura 3 El premio Nobel de Economía J. F. Nash (1926) y dos juegos de estrategia. El A tiene solución pura y el B mixta.

La determinación de las condiciones de equilibrio corresponden principalmente al campo de la matemática determinista y han tenido un gran impacto en la consolidación de la economía como ciencia. También, los problemas de *optimización*, estrechamente ligados a los anteriores, constituyen el centro de muchas aplicaciones en las ciencias económicas y, especialmente, la *programación dinámica* ha tenido una importante repercusión en la modelización de los sistemas dinámicos.

Los problemas de juegos pueden englobarse dentro del campo de la *teoría de la decisión*, que es otra de las contribuciones fundamentales de la matemática aplicada a la Economía. Esta teoría estudia la decisión óptima entre un conjunto de alternativas, conocidas las consecuencias de sus decisiones. En particular, el caso más importante es cuando la decisión se toma en condiciones de incertidumbre, y las consecuencias dependen de variables aleatorias de distribución conocida. Dentro de este campo tiene especial importancia el *enfoque Bayesiano*, que permite cuantificar la incertidumbre mediante probabilidades subjetivas. La teoría de la decisión ha tenido especial importancia en el campo de las *decisiones de inversión*. En 1952 *Harry Markowitz*, demostró, utilizando esta teoría, que el problema de invertir en condiciones de incertidumbre entre muchos activos puede reducirse a contraponer la varianza de la inversión, que mide el riesgo, y la rentabilidad esperada. Además, influido por su contacto con *George Dantzig* en la RAND corporation, demostró cómo calcular la cartera óptima mediante programación cuadrática. Por este trabajo obtuvo el premio Nobel en Economía en 1990, compartido con *Sharpe* y *Miller*, que hicieron también contribuciones importantes a las finanzas mediante el análisis cuantitativo. En particular, Sharpe es el creador del llamado Capital Asset Pricing Model, o CAPM, que muestra que la cartera óptima de un inversor depende sólo de las predicciones sobre la evolución de los activos y no de su actitud ante el riesgo, como se suponía previamente. Sharpe es también el introductor de los *coeficientes beta* de un activo, que no son más que los coeficientes de regresión obtenidos al correlacionar la evolución del activo frente a un índice global del mercado.

La aplicación de los modelos generados por la teoría económica a la realidad requiere la consideración de la incertidumbre, debida a las variables que no podemos medir y al comportamiento aleatorio resultante de la agregación de las decisiones de los agentes económicos. Por esta razón, la Estadística es imprescindible en la modelización cuantitativa en las ciencias económicas. Como ilustración de la estrecha interconexión entre Estadística y Economía, en el año 2000 la Sociedad Americana de Estadística (ASA) publicó un conjunto de artículos para evaluar la contribución de la Estadística a distintos campos científicos. No se dedicó ninguno a la Economía ya que, según Becker (2000): “*the interdigitation of economics and statistics is so complete that explicating the role of statistics in economics would essentially necessitate an almost complete accounting of the field of statistics itself*.” No es sorprendente, por tanto, que en la lista de premios Nobel de Economía haya una amplia representación de estadísticos y que, por ejemplo, entre los siete primeros premiados seis hicieron investigación en Estadística (*Frisch, Tinbergen, Samuelson, Kuznets, Arrow* y *Leontieff*), siendo varios de ellos presidentes o Fellows de las sociedades más prestigiosas de Estadística.

En particular, el primer premio Nobel de Economía fue otorgado en 1969 a los estadísticos/económetras *R. Frish* (1895-1973) y *J. Tinbergen* (1903-1994) por su contribución al desarrollo de los modelos econométricos. Frisch fue pionero en el análisis de los ciclos económicos y en los contrastes de hipótesis en las relaciones entre variables económicas. Tinbergen, formado inicialmente como físico, comenzó a trabajar en la Oficina Central de Estadística de Holanda y concibió aplicar los modelos de regresión utilizados por *Gauss* en astronomía y por *Galton* en Biología para modelar las relaciones entre las variables económicas. Una economía puede representarse por un sistema de ecuaciones que establece las relaciones lineales entre cada variable económica y todas las demás y que puedan estimarse a partir de los datos con herramientas estadísticas. Su trabajo supuso un importante avance en la predicción económica.

La siguiente revolución en la utilización de los modelos estocásticos en Economía provino de la aplicación de los *procesos estocásticos*, esto es, el estudio de variables dinámicas que evolucionan con

estructura temporal, a la predicción económica. Hasta los años 70 la mayoría de los modelos econométricos utilizados se basaban en los modelos de regresión clásicos, concebidos para situaciones estáticas. En estos años *Box y Jenkins* (1970) desarrollan una metodología para construir modelos lineales de series temporales, los modelos ARIMA (*autoregressive integrated moving average*), que suponen un avance fundamental en la modelización de series temporales en cualquier área, y, en particular, en las ciencias económicas. La teoría de los procesos estocásticos estaba construida alrededor de variables estacionarias, que son aquellas que evolucionan en el tiempo con dependencia, pero con estabilidad en la media y en la variabilidad. Por ejemplo, si consideramos la temperatura media anual en la tierra durante el siglo XIX podemos ver que aunque exista dependencia (años más calientes que la media son seguidos con frecuencia por años calientes), las temperaturas anuales oscilan alrededor de unos valores fijos que podemos suponer constantes, y la variabilidad también permanece constante en el tiempo. Esta estabilidad no es adecuada para modelar las variables económicas, que tienen frecuentemente crecimiento con tendencias y ciclos. La solución adoptada durante mucho tiempo fue introducir tendencias deterministas, pero esto es claramente insatisfactorio tanto desde el punto de vista práctico como conceptual.

La contribución principal de Box y Jenkins fué introducir procesos sin media constante, pero con derivada de algún orden constante. Por ejemplo, la media de la serie puede variar en el tiempo pero el crecimiento de la serie, es decir su primera derivada, puede oscilar alrededor de un valor constante. Decimos entonces que tenemos un *proceso integrado* de orden uno. De la misma forma, podemos considerar procesos con segunda derivada constante, o procesos integrados de orden dos. Por ejemplo, la figura 4 presenta la serie de precios del trigo en Valladolid desde Julio de 1880 a Diciembre de 1890 y su primera derivada (diferencia). Se observa que aunque la serie no es estacionaria su primera diferencia evoluciona alrededor de un valor constante y puede suponerse estacionaria.

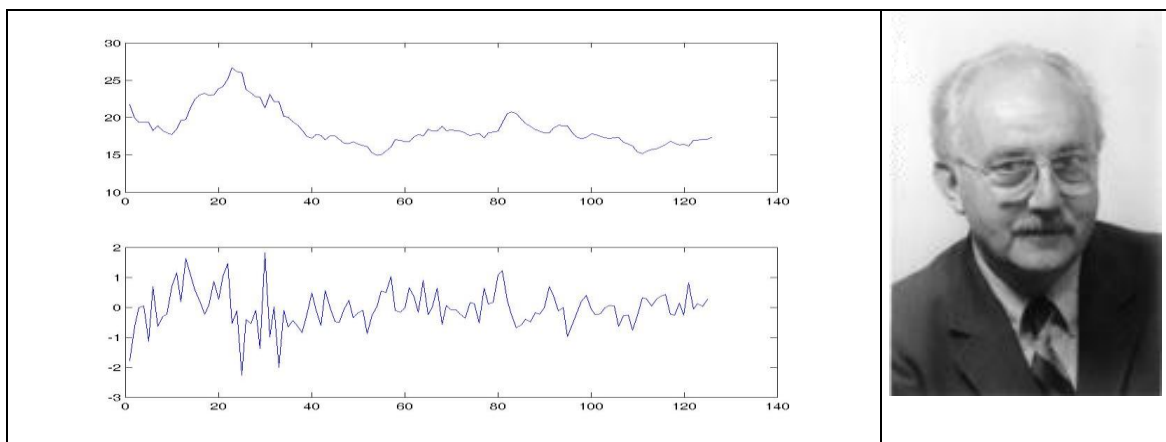


Figura 4. Serie del precio del trigo en Valladolid y su primera diferencia y el estadístico *George Box*.

Estas ideas permiten reformular la idea de equilibrio entre variables económicas: dos variables económicas decimos que están *cointegradas* si siendo ambas integradas, y por tanto no estacionarias, existe una combinación lineal que es estacionaria. Es decir, aunque las variables no sean estables en la media existe una combinación de ambas que sí lo es. La utilización de esta restricción conduce a un replanteamiento de la estimación de los modelos econométricos, y ha supuesto la concesión del premio Nobel de economía en el año 2003 a los economistas *Granger y Engle*.

Así como los modelos ARIMA de Box y Jenkins supusieron la generalización de los procesos estacionarios permitiendo pautas dinámicas en la media del proceso, las aplicaciones financieras de estos modelos han llevado a la introducción de modelos con varianza condicional no constante, donde la variabilidad futura es predecible en parte a partir de los datos observados. Los modelos ARCH (*autoregresivos condicionalmente heterocedásticos*), introducidos por Engle en 1982, tienen esta

propiedad y han sido ampliamente utilizados para modelizar variables financieras donde la predicción de la variabilidad futura, la volatilidad, es importante, al ser una medida del riesgo al invertir en un activo.

Cuando pasamos del campo de la macroeconomía al de la microeconomía y la administración de empresas encontramos que las relaciones entre las variables importantes son siempre estadísticas. Por ejemplo, en el área comercial el efecto de una campaña de promoción sobre la demanda de un producto tiene que estimarse estadísticamente, normalmente mediante técnicas de series temporales. Determinar las variables que explican (estadísticamente) las decisiones de los consumidores requiere técnicas de diseño de experimentos y clasificar y segmentar a los clientes y encontrar grupos homogéneos ha impulsado las técnicas de análisis estadístico multivariante. Además, los problemas de distribución, de redes, de canales de comercialización etc. han impulsado el desarrollo de las técnicas de investigación operativa. En el área de personal la relación entre las características personales y el rendimiento laboral es de nuevo una relación estadística y las técnicas de análisis factorial y de clasificación tienen mucha importancia en Psicometría y en el análisis del comportamiento humano en las organizaciones. En el área de producción, la caracterización del estado de control de un proceso o de la capacidad de un sistema se establece en términos estadísticos y todo el campo moderno de la mejora de la calidad y productividad está basado en estos conceptos. Finalmente, la planificación de la producción ha sido un motor en el desarrollo de la programación matemática, y la investigación operativa ha desarrollado un repertorio amplio de técnicas y modelos para la planificación y gestión de la producción y la distribución de productos.

## **LAS MATEMÁTICAS EN LA SOCIOLOGÍA Y LA CIENCIA POLÍTICA**

El conocimiento de la realidad social y política se realiza cada vez con mayor frecuencia mediante análisis de encuestas y exploraciones de campo. En una investigación sociológica al Estadística se utiliza para: 1) Diseñar la muestra; 2) Decidir el tipo de muestreo más informativo en función de la población a investigar y (3) Analizar los datos obtenidos. Además, nuevas preguntas sociológicas, resultado del análisis de encuestas, ha llevado a nuevos desarrollos de modelos estadísticos (véase *Raftery*, 2000 para un análisis más amplio de las relaciones entre Estadística y Sociología y *Beck*, 2000 para las relaciones con la Ciencia Política).

Durante la mayor parte del siglo XX los investigadores sociales, y entre ellos los estadísticos trabajando en estas aplicaciones, se concentraron en el análisis de datos de tabulación cruzada, que ha sido la presentación más habitual para resumir los resultados de una encuesta. Un avance importante en los años 60 y 70 es el desarrollo de los modelos log-lineales, que permiten explicar las frecuencias observadas de una variable en función de otras. Estos modelos permiten trasladar al análisis de tablas de contingencia muchos de los resultados de modelos lineales obtenidos en las aplicaciones económicas durante la primera mitad del siglo XX. También, en la segunda mitad del siglo XX, se desarrollaron medidas de asociación entre variables categóricas que generalizan el coeficiente de correlación para medir la dependencia en contextos más amplios. Podemos concluir que los avances en el tratamiento estadístico de las variables cualitativas han venido impulsado por las necesidades del análisis sociológico.

Un avance importante en este campo fue encontrar la relación entre el análisis de correspondencias, herramienta introducida por el matemático francés *Benzecri* para representar datos multivariantes de frecuencias, y los modelos log-lineales. El *Análisis de Correspondencias* puede verse como una generalización para variables cualitativas de una técnica estadística clásica para reducir la dimensión de variables continuas: *Componentes Principales*. Esta técnica tiene muchas aplicaciones a campos distintos. Por ejemplo, los índices de desarrollo humano de la ONU, muchos sistemas de compresión de imágenes o video para su envío por Internet, o los factores que explican la inteligencia humana, utilizan esta herramienta introducida por *Hotelling* en 1933. Matemáticamente, se trata de buscar combinaciones lineales de un conjunto de variables,  $X_1, \dots, X_k$  que contengan la máxima



información, lo que se traduce en que tengan varianza máxima. En otros términos, se trata de encontrar una combinación lineal

$$Y = a_1 X_1 + \dots + a_k X_k$$

(o la dirección de proyección definida por el vector  $\mathbf{a} = (a_1, \dots, a_k)$ ) que permita prever cada una de las variables  $X$  con error mínimo (véase *Peña*, 2003 para una explicación más detallada de estas ideas). *Goodman* en 1985 relacionó los modelos log-lineales y el análisis de correspondencias, estableciendo una metodología unificada para el análisis de encuestas.

Otro desarrollo importante de las aplicaciones sociológicas es el modelo *LISREL*, que generaliza el *modelo factorial clásico* introducido por *Spearman* en Psicometría para analizar la inteligencia humana. Estos modelos consideran un sistema de ecuaciones estructurales, similares a los modelos econométricos multiecuacionales, pero que en lugar de relacionar variables observadas relacionan variables no observadas o factores. Además, incorporan ecuaciones similares a un modelo factorial clásico para relacionar los factores con las variables observadas. Estos modelos se han convertido en una herramienta habitual para la contrastación de teorías sociológicas.

Un tercer aspecto que ha estimulado la investigación es Estadística es el *análisis de datos faltantes*. En muchas encuestas es frecuente que los individuos entrevistados no respondan a parte de las preguntas y nos encontramos con el problema de cómo utilizar esta información parcial. Este problema ha impulsado la creación de nuevos métodos de estimación, como el *algoritmo EM* o los nuevos métodos de cadenas de *Markov* (véase por ejemplo Robert y Casella, 1999), que permiten una utilización efectiva de toda la información disponible.

## LAS MATEMÁTICAS Y LAS HUMANIDADES

Las matemáticas han tenido una reconocida influencia en la caracterización de la armonía y la belleza en el arte y la arquitectura. Un ejemplo bien conocido es la *sección áurea* o dorada de *Leonardo da Vinci* ( $a/b = b/(a+b)$ ) para las proporciones de segmentos y lados de figuras geométricas, el cuerpo humano o los templos clásicos como el Partenón (véase la Figura 5). Más recientemente, la Estadística ha sido utilizada con éxito para resolver problemas de autoría en manifestaciones literarias y artísticas, elaborar mapas de palabras e investigar la evolución y relaciones entre los idiomas y analizar y clasificar pinturas y estilos artísticos.

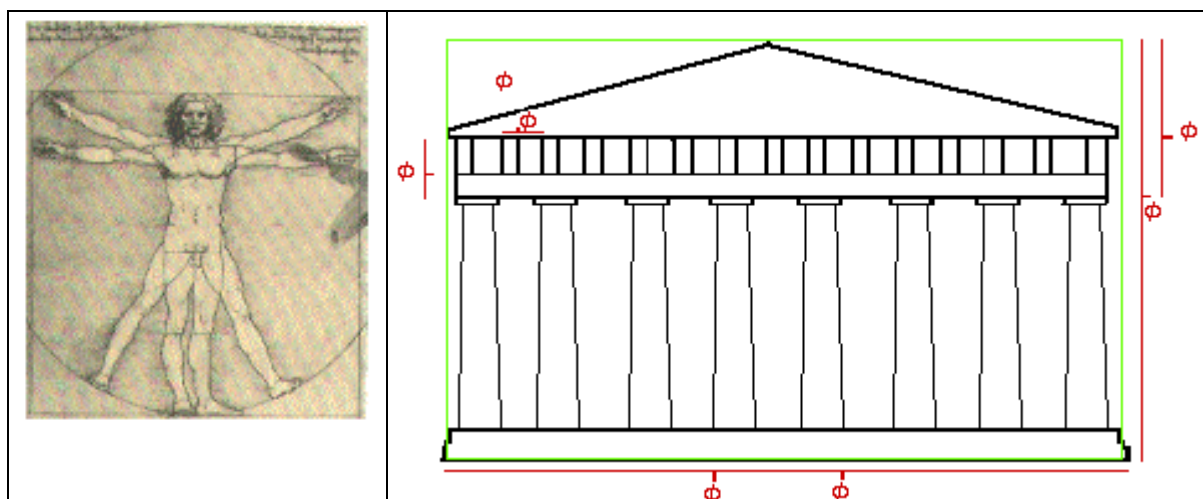


Figura 5. Dos ejemplos de la *razón áurea*: a la izquierda el hombre de Leonardo con este cociente entre el lado del cuadrado y el radio de la circunferencia con centro en el ombligo; a la derecha dimensiones que siguen la razón áurea en el Partenón. (tomado de <http://roble.cnice.mecd.es/~tvirgos/matematicas/partenon.htm>)



En un trabajo pionero *Mosteller y Wallace* (véase un resumen de este trabajo en Tanur et al. 1989) demostraron cómo utilizar métodos estadísticos simples, basados en la frecuencia de utilización de palabras, para decidir respecto a la autoría de un escrito. Estos autores estudiaron 12 artículos cuya autoría no era conocida, y que habían sido atribuidos a Hamilton o Madison, y los clasificaron de forma convincente y sin ambigüedad. Estudios similares se han realizado para decidir la autoría de poemas atribuidos a Shakespeare y a otros autores y en España recientemente para el análisis literario de Tirant le Blanc (Véase *Girón, Ginebra y Riba* 2005).

En mi opinión, y por las razones indicadas en la introducción, este siglo va a experimentar un crecimiento espectacular de las aplicaciones de las matemáticas, y en particular de la estadística, a las humanidades.

## CONCLUSIONES

En este trabajo hemos ilustrado el papel de las matemáticas como soporte para la adquisición de conocimiento científico en el mundo actual. En las ciencias sociales la Estadística es, en palabras de *Mood y Graybill* (1972), “la tecnología del método científico” y la herramienta principal para convertir los hechos observables en conocimiento e información. Esta influencia comenzó en la Economía, y la resolución de problemas económicos ha sido uno de los motores principales de desarrollo de algunas ramas de la matemática y de la estadística en el siglo XX. Este efecto es también claro en las ciencias sociológicas, si bien su desarrollo cuantitativo es menor, así como en las humanidades, donde todavía se encuentra en sus inicios. Sin embargo, es previsible que en este siglo se producirán avances importantes en la cuantificación de estas disciplinas.

Para el lector interesado en ampliar esta breve introducción sobre las aplicaciones a las ciencias sociales recomendamos el libro de Tanur y otros (1989).

## Agradecimientos

Este trabajo se ha beneficiado de los comentarios e ideas de Carlos Hervés, Francisco Maruhenda, Carmelo Núñez y Juan Pablo Rincón sobre el papel de las matemáticas en las Ciencias Económicas. Los editores han contribuido a mejorar la presentación y hacerla más clara y legible. Para todos ellos mi agradecimiento. Este trabajo ha contado con el apoyo del proyecto del MEC SEJ2004-03303.

## REFERENCIAS

- Beck, N.L. (2000), Political Methodology: A Welcoming Discipline, *The Journal of American Statistical Association*, 95, 628-629.
- Becker, M.P. (2000), Statistics in Business and Social Science, *The Journal of American Statistical Association*, 95, 628-629.
- Box, G. E. P. y Jenkins, G. M. (1970), *Time Series Analysis Forecasting and Control*, Holden-Day
- Girón, F. J., Ginebra, J. y Riba, A. (2005). Bayesian Analysis of a Multinomial Sequence and homogeneity of Literary Style, *The American Statistician* 59 (1): 19-30.
- Mood A. M. y Graybill, F. A. (1972) *Introducción a la Teoría de la Estadística*. Aguilar.
- Morgenstern, O. y von Neumann, J. (1944) *The Theory of Games and Economic Behavior* Princeton University Press
- Mosterín, J (2005), *La naturaleza humana*. Espasa Calpe.
- Nash, J. (1950), Equilibrium points in n-person games, *Proceedings of the National Academy of the USA* 36(1): 48-49.
- Peña, D. (2003), *Análisis de datos multivariantes*, Mc-Graw Hill.

Raftery, A. (2000), Statistics in Sociology, *The Journal of American Statistical Association*, 95, 654-660.  
Robert, C. y Casella, G. (1999), *Monte Carlo Statistical Methods*, Springer.  
Tanur, J. M. y otros (1989) *La Estadística, una guía de lo desconocido*. Alianza editorial.



**Daniel Peña** es Catedrático de Estadística de la Universidad Carlos III de Madrid. Ha sido Catedrático en las universidades Wisconsin-Madison, Chicago y Politécnica de Madrid. Tiene publicados 13 libros y más de 150 trabajos de investigación sobre Estadística y Econometría. Es Fellow del Institute of Mathematical Statistics (IMS), de la American Statistical Association (ASA) y miembro del International Statistical Institute (ISI).