Revista ENERGEIA
Sociedad Ibero-Americana de Metodología Económica
ISSN: 1666-5732
Vol. VII, Núm. 1, 2020, 07-22

La Estructura Lógica de la Econometría

ADOLFO GARCÍA DE LA SIENRA
Instituto de Filosofía, Facultad de Economía
Universidad Veracruzana

Resumen
La finalidad de este artículo es formalizar de una manera detallada la doctrina metodológica de Spanos (1986) desde la perspectiva de la concepción estructuralista de las teorías (CET). Después de proporcionar los fundamentos probabilísticos de la econometría, se procede a discutir y enunciar su ley fundamental y la forma general de su aserción empírica.
Palabras clave: concepción estructuralista de las teorías, fundamentos de econometría, Spanos
Clasificación JEL: B4, C1.

The Logical Structure of Econometrics

Abstract
The aim of the present paper is to formalize in a detailed way Spanos' (1986) econometric doctrine within the perspective of the structuralist view of theories (CET). After providing the probabilistic foundations of econometric theory, it discusses and formulates its fundamental law, as well as the general form of its empirical claim.
Keywords: structuralist view of theories, foundations of econometrics, Spanos
JEL Classification: B4, C1.

1.Motivación

La finalidad de este artículo es formalizar de una manera detallada la doctrina metodológica de Spanos (1986) desde la perspectiva de la concepción estructuralista de las teorías (CET). Como habremos de ver, CET proporciona una reconstrucción muy natural de la doctrina, una que integra la misma en un marco filosófico y metodológico general. Este artículo se propone llenar una laguna que había sido observada hace ya mucho tiempo por Bruce J. Caldwell:

> "Un enfoque que a mi conocimiento ha sido completamente ignorado es la integración de la metodología económica y la filosofía con la econometría. Los metodólogos generalmente han eludido el tema de los fundamentos metodológicos de la teoría econométrica, y los pocos econometristas que han abordado los temas filosóficos a duras penas han ido más allá de referencias gratuitas a figuras tales como Feigel [Feigl] y Carnap.[1]" (Caldwell 1984: 216)

Nota. Todas las traducciones son del autor.

En un artículo publicado en 2000, George C. Davis señalaba:

> "Desde 1982, el interés en la metodología económica se ha incrementado, pero la mayoría de los metodólogos se han enfocado en los temas metodológicos más importantes, tales como el estatus científico o el realismo de la economía, más que en los "fundamentos metodológicos" de la econometría. De modo similar, la mayoría de los econometristas que han abordado los "temas filosóficos" han desarrollado sus propias metodologías y meramente actualizado las "referencias gratuitas" sea ya a Lakatos (1973), o Kuhn (1970) (v.g., Darnell y Evans, 1990; Hendry, 1993, 1995; Spanos, 1986). Consecuentemente, parece que la observación de Caldwell todavía es aplicable en un grado significativo y digno de consideración." (Davis 2000: 206)

En el artículo anteriormente mencionado, Davis presenta una interesante reconstrucción de los fundamentos metodológicos de la econometría de Haavelmo desde el punto de vista de la concepción semántica de las teorías de Suppe (cfr. Suppe 1989), la cual no debiera ser confundida con CET. Mi meta es aplicar esta última metateoría a la doctrina de Spanos, la cual es un desarrollo sistemático y detallado de la concepción de Haavelmo, formulada en un lenguaje más contemporáneo.

2. PGDs reales y concretos

Lo que los econometristas llaman 'proceso generador de datos' es un sistema o fenómeno tomado como objetivo de la investigación científica, a partir del cual se espera que fluyan datos empíricos (o sean penosamente obtenidos). Desde un punto de vista ontológico, es un sistema o fenómeno económico real y concreto que se comporta de manera aleatoria. La definición funcional de PGD de Spanos es "el mecanismo que subyace a los fenómenos observables de interés" (Spanos 1986: 20). El término es también

> "usado para designar el fenómeno de interés que la teoría se propone explicar. El concepto es usado para enfatizar el alcance pretendido de la teoría, así como la fuente de los datos observables. Definido de esta manera, el concepto del PGD bajo estudio podría ser un fenómeno real observable o una situación experimental, dependiendo del dominio de aplicaciones pretendidas de la teoría.
> Por ejemplo, en el caso de la teoría de la demanda, si el alcance pretendido de la teoría es determinar una relación entre un hipotético rango de precios y las correspondientes intenciones de comprar de un grupo de agentes económicos en un momento particular del tiempo, el PGD real podría ser usado para designar la situación experimental donde tales datos podrían ser generados. Por otro lado, si el alcance pretendido de la teoría es explicar cambios observados de cantidades y/o precios a lo largo del tiempo, entonces el PGD bajo estudio debiera referirse al proceso real de mercado que da lugar a los datos observados. Debiera notarse que el alcance pretendido de la teoría es usado para determinar la elección de los datos observables a ser utilizados." (Spanos 1986: 661-2)

El PGD es un *subjekt* real concreto que permanece fuera de la cabeza, independiente de ella y de cualquier teoría; es un "mecanismo aleatorio" que da lugar a datos observables. El pensamiento preteórico puede aprehender su existencia y algunas de sus propiedades y efectos, proporcionando el punto de partida para una descripción del mismo en un lenguaje que contiene términos generales pero no idealizados. Esta descripción es suficiente para *referirse* al PGD y empezar a sondearlo.

Por ende, la más bien extraña doctrina (idealista) de la "dependencia teórica" de la observación es rechazada como innecesaria y confusa. Esto no es negar que las teorías pueden conducir a descubrir nuevos aspectos del proceso o la entidad; solamente que su existencia no depende de ninguna conceptualización. Simplemente no es verdad que no tengamos acceso al mundo independientemente de nuestras sistematizaciones teóricas. Si ello fuese el caso nunca seriamos capaces de darnos cuenta de que los axiomas de una teoría idealizada son de hecho falsos o inexactos.

Para la econometría, el PGD es un “mecanismo” aleatorio pero la aleatoriedad debe ser tomada como una propiedad objetiva de las cosas reales y no como una medida subjetiva de certidumbre o certeza. Como señalara Patrick Suppes (1984:10): “las leyes fundamentales de los fenómenos naturales son esencialmente probabilísticas más que deterministas en carácter”. Esta misma proposición puede ser extendida para incluir fenómenos sociales y no solamente naturales. Al adoptar la interpretación frecuentista de la probabilidad, Spanos en realidad acepta esta proposición, la cual también daré por sentada.

Spanos utiliza ‘teoría’ para referirse a cierta interpretación de la teoría de probabilidad que pretende explicar un cierto fenómeno de interés. En la jerga de CET, una teoría es mucho más que eso. Es la descripción general de una clase de estructuras conjuntistas, más una familia de aplicaciones pretendidas que incluyen descripciones cualitativas de su estructura. En el uso de Spanos, es la aplicación del concepto de espacio de probabilidad a un sistema objetivo particular, junto con varios métodos de estimación diseñados para determinar los parámetros T-teóricos. Por ende, parece ser solamente natural proponer una definición de econometría como teoría de probabilidades junto con una clase más bien grande de sistemas aleatorios de naturaleza económica, con métodos correspondientes para estimar los parámetros que los caracterizan.

La anterior concepción implica que el elemento teórico básico de la econometría está dado por los axiomas de probabilidad de Kolmogórov pero hay mucho más para la econometría que tales axiomas. En la sección final presentaré una caracterización detallada si bien compacta de la teoría econométrica.

3. Experimentos aleatorios

El punto de partida de la doctrina econométrica de Spanos es el concepto de experimento aleatorio ε, al cual caracteriza como una representación idealizada del PGD que satisface las tres condiciones de la siguiente definición.

DEFINICIÓN 1. Un *experimento aleatorio* es un experimento $\boldsymbol{\varepsilon}$ que satisface las siguientes condiciones:

(1) Todos los posibles resultados distintos son conocidos a priori;
(2) En cualquier ensayo particular el resultado no se conoce a priori;
(3) Se puede repetir bajo condiciones idénticas.

Ejemplos de experimentos aleatorios son: el lanzamiento de una moneda equilibrada, o una lotería repetida bajo circunstancias similares. La anterior definición de experimento aleatorio —más aún, la misma palabra ‘experimento’— sugiere que el mecanismo que produce los resultados se halla bajo el control del observador. Pues de otra manera sería imposible repetir el proceso “bajo condiciones idénticas”. Es por ello que algunos autores prefieren caracterizar un experimento aleatorio como un proceso por el cual algo incierto se observa. A veces el observador es solamente un observador pasivo de un proceso que se comporta de manera aleatoria sin ser capaz de influenciar en lo más mínimo el resultado del proceso. Acordemente, introduciré el más general concepto de proceso aleatorio modificando la cláusula (3) como sigue: si se repite bajo condiciones análogas cuando estas condiciones se hallan bajo el control del observador el proceso es un experimento.

Por otro lado, el proceso aleatorio (o su descripción) no tiene que ser idealizado. Por ejemplo, el conjunto de resultados de arrojar una moneda equilibrada es descrito exactamente como consistiendo de dos elementos, a saber, águila y sol. Así, el espacio aleatorio usualmente no es un constructo idealizado, especialmente cuando es finito. Las operaciones conjuntistas tampoco son idealizadas, ya que tienen un significado intuitivo muy claro.

En realidad, la Definición 1 es apropiada para la interpretación de probabilidades de la frecuencia relativa, la cual requiere la repetición del mismo experimento. El problema es que no es poco común encontrar en economía que las probabilidades tienen que ser obtenidas a partir de ocurrencias de un solo caso de eventos singulares, sin ninguna posibilidad de repetición. La razón por la que filósofos como Karl Popper y Patrick Suppes se apartaron de la frecuencia relativa para considerar la interpretación de propensidad de la probabilidad de eventos singulares como fundamental y primaria es el problema del caso único.

Puesto que encontramos en el ámbito económico muchos fenómenos aleatorios que solamente ocurren una vez, la interpretación de propensidad parece más apropiada para la econometría. Las estructuras de probabilidad condicional son más apropiadas para las representaciones de probabilidad de propensidad, y más fundamentales que los espacios de probabilidad usuales definidos por los axiomas de Kolmogorov. No obstante, será conveniente empezar recordando la definición de estos espacios.

4. Espacios de probabilidad

Un espacio de probabilidad representa los resultados posibles de un experimento o proceso aleatorio, llamados *eventos*, cada uno de los cuales puede ocurrir con una cierta probabilidad. Por ejemplo, el resultado de arrojar dos dados puede ser el evento "sale siete" o "sale un número par". El evento "sale siete" está compuesto de los eventos elementales (1,6), (2,5), (3,4), (4,3), (5,2), (6,1), y está representado por el conjunto

$$\{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}.$$

Por ende, se puede ver que los eventos de un espacio de probabilidad pueden ser elementales o compuestos de eventos elementales. Un evento elemental en el ejemplo anterior, es "(2,5) resulta"; este evento es un constitutivo del evento complejo "sale siete". En general, cada evento complejo es identificado con el conjunto de todos los eventos elementales que lo constituyen, siendo suficiente la ocurrencia de cualquiera de éstos para la ocurrencia del evento complejo. Los eventos elementales son llamados *puntos muestrales* y el conjunto de todos los puntos muestrales es llamado *espacio muestral*.

Desde un punto de vista lógico, 'espacio aleatorio' es un término primitivo que es caracterizado axiomáticamente meramente como un conjunto no vacío S, y 'punto muestral' es definido meramente como cualquier elemento del espacio muestral. El espacio de eventos, el cual es la familia de los eventos complejos es también introducido como un término primitivo y caracterizado como un álgebra de conjuntos sobre S. El tercer primitivo es 'medida de probabilidad', el cual se caracteriza axiomáticamente como una medida sobre el álgebra de eventos. La definición canónica de 'espacio de probabilidad' es la siguiente.

DEFINICIÓN 2. $\mathfrak{C}$ es un *espacio de probabilidad* syss existen *S*, *F* y *P* tales que

(1) $\mathfrak{C}$ = ‹*S*, *F*, *P*›

(2) *S* es un conjunto vacío.

(3) *F* es un σ-álgebra de conjuntos sobre *S*.

(4) P es una función *P*:*F* → [0;1] que es aditiva; es decir, para todo $A \epsilon F$ y para cualquier descomposición finita $A_1, \ldots, A_n \ \in \ F$ de A:

$$P(A) = \sum_{k=1}^{n} P(A_k).$$

(5) *P*(S) = 1.

Recordemos que un σ-álgebra de conjunto sobre *S* es un anillo de conjuntos, con *S* como unidad, cerrado bajo uniones numerables. Un modo compacto de formular el axioma 3 consiste en decir que *P* es una medida sobre *F*. Como dijimos arriba, el conjunto *S* es llamado *espacio muestral*, sus elementos *eventos elementales,* y *F* es el conjunto de los eventos. *P* es llamado *medida de probabilidad* sobre *F*. Los términos conjuntistas admiten una interesante interpretación probabilista (véase la Tabla 1).

Tabla 1: Correlaciones entre nociones conjuntistas y probabilísticas.

Teoría de conjuntos	**Teoría de probabilidades**
$A \in F$	A es un evento.
$A \cap B = \emptyset$	Los eventos A y B son incompatibles.
$A_i \ \cap \ A_j \ = \ \emptyset$	Los eventos $A_1, \ldots, A_n$ con $i \ \neq \ j$ y $1 \leq i; j \leq n$ son incompatibles por pares
$\prod_{j=1}^{n} A_i = B$	*B* es el evento de que todos los eventos $A_1, \ldots, A_n$ ocurren simultáneamente.
$\coprod_{i=1}^{n} A_i = B$	*B* es el evento que ocurre si ocurre al menos uno de los eventos $A_1, \ldots, A_n$.
$A' = S \setminus A$	El evento que ocurre cuando *A* no ocurre.
$A = \emptyset$	*A* es el imposible evento de que no ocurra nada.
$A = S$	*A* es el seguro evento de que algo ocurre
$B \subseteq A$	La ocurrencia del evento *B* acarrea la ocurrencia del evento *A*

Las siguientes son identidades elementales que valen en todo espacio de probabilidad. Su demostración es fácil y es dejada al lector.

TEOREMA 1. Sea $\langle S, F, P \rangle$ un espacio de probabilidad. Entonces:

(1) Para todo $A \epsilon F$: $P(S \setminus A) = 1 - P(A)$.

(2) $P(\emptyset) = 0$.

(3) Para todo $A, B \in F$, si $A \subseteq B$ entonces $P(B \setminus A) = P(B) - P(A)$.

(4) Para todo $A, B \in F$, si $A \subseteq B$ entonces P(A) ≤ P(B).

(5) Para todo $A \in F : 0 \leq P(A) \leq 1$

(6) Para todo $A, B \in F$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

El espacio muestral de un espacio de probabilidad puede ser finito, contablemente infinito, o continuo. Si S es a lo sumo numerable, al menos algunos eventos elementales deben obtener probabilidades positivas. Esto se hace a través de una función llamada 'distribución de probabilidad'.

DEFINICIÓN 3. Sea S un conjunto a lo sumo numerable. Una *distribución de probabilidad* sobre S es una función $p: S \to [0; 1]$ tal que

$$\sum_{x \in S} p(x) = 1$$

Una distribución de probabilidad asigna probabilidades positivas al menos a algunos eventos elementales. Si asigna probabilidades positivas a todos ellos es llamada una *distribución simple*.

Es fácil ver que una distribución de probabilidad sobre un conjunto S genera de manera natural un espacio de probabilidad si, para cada subconjunto A de S, $P(A)$ es definido mediante la igualdad

$$P(A) = \sum_{x \in A} p(x)$$

Como espacio de eventos usualmente se toma el conjunto potencia de S.

Un σ-álgebra particularmente importante en la teoría de probabilidades es el mínimo conjunto que contiene todos los intervalos de $\mathbb{R}$ el cual es cerrado bajo uniones numerables. Esta álgebra es llamada *álgebra de Borel*, denotada por $\boldsymbol{B}$¸y sus elementos son llamados *conjuntos de Borel*. Estas nociones son esenciales para la definición del concepto de variable aleatoria.

DEFINICIÓN 4. $\mathfrak{C} = \langle S, F, P \rangle$ es un *espacio de probabilidad contablemente aditivo* syss $\mathfrak{C}$ es un espacio de probabilidad tal que, para toda familia $\{A_n\}_{n<w}$

$$P\left(\coprod_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Las medidas de probabilidad de un espacio de probabilidad contablemente aditivo tienen las propiedades expresadas en el siguiente teorema.

TEOREMA 2. Sea $\langle S, F, P\rangle$ un espacio de probabilidad. Las siguientes aserciones son equivalentes:

(1) $\langle S, F, P\rangle$ es contablemente aditivo;

(2) Si $\{A_n\}_{n\in\boldsymbol{\omega}}$ es una secuencia creciente de elementos *F*:
$A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$,
tal que $\cup_{n\in\boldsymbol{\omega}} A_n = S$, entonces $lím_{n\to\infty} P(A_n) = 1$.

(3) Si $\{A_n\}_{n\in\boldsymbol{\omega}}$ es una secuencia decreciente de elementos *F*:
$A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$,
tal que $\cap_{n\in\boldsymbol{\omega}} A_n = \emptyset$, entonces $lím_{n\to\infty} P(A_n) = 0$.

Demostración: Supongamos que $\langle S, F, P\rangle$ es un espacio de probabilidad contablemente aditivo, sea $\{A_n\}_{n\in\boldsymbol{\omega}}$ una secuencia creciente de elementos de *F* tal que $\cup_{n\in\boldsymbol{\omega}} A_n = S$, y defínase la secuencia de eventos $\{B_n\}_{n\in\boldsymbol{\omega}}$como sigue:

$$B_1 = A_1;\ B_n = A_n \setminus \coprod_{k<n} A_k$$

Claramente, para cada $n \in \boldsymbol{\omega}$, $A_n = \coprod_{k=1}^{n} B_n$, y así $\coprod_{n\in\boldsymbol{\omega}} B_n = \coprod_{n\in\boldsymbol{\omega}} A_n = S$. Como los conjuntos B_n son disyuntos por pares, $P(\coprod_{n\in\boldsymbol{\omega}} B_n) = \sum_{n\in\boldsymbol{\omega}} P(B_n) = 1$ y la secuencia

$$P(B_1), \sum_{K=1}^{2} P(B_k), \ldots,\ \sum_{k=1}^{n} P(B_k), \ldots$$

converge a 1. Pero

$$P(A_n) = P(\coprod_{K=1}^{n} B_k) = \sum_{K=1}^{n} P(B_k)$$

para cada $n \in \boldsymbol{\omega}$, y por lo tanto $lím_{n\to\infty} P(A_n) = 1$.

Suponga ahora que cualquier secuencia creciente de elementos de *F* satisface la condición (2), sea $\{B_n\}_{n\in\boldsymbol{\omega}}$ una secuencia decreciente de elementos de *F*:

$$B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots,$$

tal que $\prod_{n \in \boldsymbol{\omega}} A_n = \emptyset$, y defínase la secuencia de eventos $\{C_n\}_{n \in \boldsymbol{\omega}}$ mediante la condición "$C_n = S \setminus A_n$". Claramente, $S \setminus A_n \subseteq S \setminus A_{n+1}$, y así $\{B_n\}_{n \in \boldsymbol{\omega}}$ es una secuencia decreciente de eventos con

$$\coprod_{n \in \boldsymbol{\omega}} C_n = \coprod_{n \in \boldsymbol{\omega}} (S \setminus A_n) = S \setminus \prod_{n \in \boldsymbol{\omega}} A_n = S.$$

Finalmente, supóngase que cualquier secuencia decreciente de elementos de F satisface la condición (3), y sea $\{A_n\}_{n \in \boldsymbol{\omega}}$ una colección de elementos disyuntos por pares de F. Para cada $n \in \boldsymbol{\omega}$, póngase

$$B_n = \coprod_{n=1}^{\infty} A_n \setminus \coprod_{K=1}^{n} A_k$$

Entonces $\{B_n\}_{n \in \boldsymbol{\omega}}$ es una secuencia decreciente de elementos de F tal que $\prod_{n=1}^{\infty} B_n = \emptyset$. Por ende,

$$\begin{aligned}
0 &= \lim_{n \to \infty} P(B_n) \\
&= P\left(\coprod_{n=1}^{\infty} A_n \setminus \coprod_{K=1}^{n} A_k\right) \\
&= P\left(\coprod_{n=1}^{\infty} A_n\right) - \lim_{n \to \infty} P\left(\coprod_{K=1}^{n} A_k\right) \\
&= P\left(\coprod_{n=1}^{\infty} A_n\right) - \lim_{n \to \infty} \sum_{K=1}^{n} P(A_k).
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
P\left(\coprod_{n=1}^{\infty} A_n\right) &= \lim_{n \to \infty} \sum_{K=1}^{n} P(A_k) \\
&= \sum_{n=1}^{\infty} P(A_n).
\end{aligned}$$

Un espacio de probabilidad cuyo espacio aleatorio S es a lo sumo contable es llamado *discreto*. Es llamado *continuo* si la probabilidad de cualquier evento elemental es cero.

5. Variables aleatorias

Una variable aleatoria sobre un espacio de probabilidad es una función que asocia a todo evento elemental del espacio un cierto resultado numérico. Si S es el conjunto de todos los posibles resultados de arrojar simultáneamente dos dados, un ejemplo de una variable aleatoria sería la función que asigna a cada evento elemental el pago que resulta para el jugador si ese evento ocurre (siempre y cuando la función asigne el mismo pago a todos los diferentes modos en los cuales el evento ocurre). Por ejemplo, como vimos arriba el evento "sale siete" es representado como

$$A = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}.$$

Pero el casino debe pagar la misma cantidad al jugador no importa como resulte el evento, y es así que una variable aleatoria X que representa la regla de pagar una cantidad establecida m al siete debe mapear todo elemento de A en un número que representa esa cantidad de dinero: $X(x) = m$ para todo $x \in A$. En general, en situaciones económicas en las que la utilidad depende del azar, resultará que la función de utilidad es una variable aleatoria. Esto motiva la relevancia de la siguiente definición.

DEFINICIÓN 5. Sea $\mathfrak{C} = \langle S, F, P\rangle$ un espacio de probabilidad. Una *variable aleatoria (real)* (v.a.) sobre S es una función X que mapea cada evento elemental en el espacio medible $\langle \mathbb{R}, \boldsymbol{B}\rangle$ donde $\boldsymbol{B}$ es el σ-álgebra de los conjuntos de Borel, y X es medible, i.e., $X^{-1}(B) = \{x \in S \mid X(x) \in \boldsymbol{B}\}$ es un evento en F para todo $B \in \boldsymbol{B}$.

La mensurabilidad de la v.a. permite la inducción de una medida de probabilidad en el espacio de valores de la variable, en el sentido de la siguiente proposición.

TEOREMA 3. Sea $S = \langle S, F, P\rangle$ un espacio de probabilidad, y X una v.a. sobre S. La función $P_X : \boldsymbol{B} \to \mathbb{R}$, definida por la condición

$$P_X(B) = P[X^{-1}(B)],$$

es una medida de la probabilidad sobre $\boldsymbol{B}$. Si P es σ - aditiva, también lo es P_X.

Demostración: Claramente, el codominio de P_X es [0;1], así que bastará mostrar que P_X es aditiva. Sea $B_1, \ldots, B_n \in \boldsymbol{B}$ una descomposición finita de cualquier conjunto de Borel $B \in \boldsymbol{B}$. Es fácil ver que

$$X^{-1}\left(\coprod_{k=1}^{n} B_k\right) = \coprod_{k=1}^{n} X^{-1}(B_k)$$

Y así

$$P_X\left(\coprod_{K=1}^{n} B_k\right) = P\left[X^{-1}\left(\coprod_{k=1}^{n} B_k\right)\right]$$

$$= P\left(\coprod_{k=1}^{n} X^{-1}(X_k)\right)$$

$$= \sum_{K=1}^{n} P_X(B_k).$$

En realidad, si $\{B_k\}_{k \in \omega}$ es una familia de conjuntos disjuntos por pares en **B**,

$$X^{-1}\left(\coprod_{k \in \omega} B_k\right) = \coprod_{k \in \omega} X^{-1}(B_k)$$

y un argumento análogo establece que $P_X\ (\coprod_{k \in \omega} B_k) = \sum_{k \in \omega} P_x(B_k)$.

La medida de probabilidad P_X es llamada *la función de probabilidad de la v.a.* X. La *función de distribución* de P_X es la aplicación $F : \mathbb{R} \to \mathbb{R}$ definida por la condición

$$F(x) = PX\{X \leq x\} = P_X(\{\xi \in \mathbb{R} \mid \xi \leq x\}):$$

Es fácil ver que *F* es continua no decreciente desde la izquierda. *F* es *discreta* si hay un conjunto de valores distintos x_i tales que $p_i = P_X\{X = x_i\} > 0\ (i = 1, 2, 3, \ldots)$ y $\Sigma_i\, p_i = 1$. Si $P_X\{X = x\} = 0$ para todo número real *x*, se dice que *F* es *continua*.
Si hay una función no negativa $f(x)\ (x \in \mathbb{R})$ tal que

$$F(x) = \int_{-\infty}^{x} f(\xi)\, d\,\xi,$$

f es llamada *la densidad de probabilidad* de F. Hay ejemplos de v.a. que son continuas pero carecen de una densidad de probabilidad. No obstante, la teoría econométrica considera sólo aquellos que la tienen.

6. La probabilidad condicional como fundamental

El econometrista en activo toma el concepto de espacio de probabilidad como una formalización del concepto de un experimento aleatorio. Esto está bien si la interpretación adoptada de probabilidad es la de la frecuencia relativa. No obstante, es más natural en economía adoptar la interpretación de propensidad, así que debiera ser adoptada como básica la noción de probabilidad condicional (cualitativa). Introduciré en lo que sigue una axiomatización más bien natural de este concepto, originalmente debida a Luce (1968).

La idea básica es la de la factibilidad de un evento *A* dado un evento *B*, denotado (como es usual) por *'A | B'*. La comparación elemental es entre la factibilidad de ciertos eventos dados otros eventos, a saber, si la ocurrencia del evento *A* condicional a la ocurrencia de *B* es juzgada como cualitativamente al menos tan probable como que ocurra *C* condicionado a que *D* haya ocurrido. Esto se expresa simbólicamente mediante la notación:

$$A \mid B \succsim C \mid D:$$

Axiomas más bien naturales sobre estas nociones implican la existencia de una representación numérica de probabilidad condicional, la cual es una función P sobre un álgebra F de eventos tal que

$$A \mid B \ \gtrsim\ C \mid D \ syss\ P(A \cap B)/\,P(B)\ \geq\ P(C \cap D)\ /\ P(D).$$

para cualesquiera $A, B, C, D\ \in\ F$. Como la representación requiere que $P(B), P(D)\ > 0$, la factibilidad condicional de un evento no puede ser tomada relativa a un evento imposible. Es por ello que es necesario recoger los eventos imposibles en un conjunto N; obviamente todos los eventos en N son equivalentes en factibilidad a $\emptyset$.

Estos axiomas se resumen en la siguiente definición.

DEFINICIÓN 6. $\boldsymbol{\mathcal{E}}$ es una *estructura arquimediana de probabilidad condicional cualitativa* syss existen $S, F, N, y\ \gtrsim$, tales que, para todo $A, B, C, D, A', B', y\ C'\ en\ F$ (o $F\ \setminus N$ siempre que el símbolo aparezca a la derecha de |) valen los siguientes axiomas:

(0) $\boldsymbol{\mathcal{E}} =\ \langle S, F, N, \gtrsim \rangle$.
(1) $\langle F\ \times (F\ \setminus N)\ \gtrsim \rangle$ es un orden débil.
(2) $S\ \in\ F\ \setminus N;\ y\ A\ \in N\ syss\ A \mid S\ \sim\ \emptyset \mid S.$
(3) $S \mid S\ \sim\ A \mid A\ y\ S \mid S\ \gtrsim A \mid B.$
(4) $A \mid B\ \sim\ A\ \cap B \mid B.$
(5) Supóngase que $A\ \cap\ B\ =\ A' \cap\ B' = \emptyset$. Si $A \mid C\ \gtrsim\ A' \mid C'\ y\ B \mid C\ \gtrsim\ B' \mid A$, entonces $A \cup B \mid C\ \gtrsim\ A' \cup B' \mid C'$; más aún, si cualquier hipótesis es $\succ$, entonces la conclusión es $\succ$.
(6) Supóngase que $C\ \subset\ B\ \subset\ A\ y\ C'\ \subset\ B' \subset A'$. Si $B \mid A\ \gtrsim\ C' \mid B'\ y\ C \mid B\ \gtrsim\ B' \mid A'$, entonces $C \mid A\ \gtrsim\ C'\ \gtrsim\ A'$; mas aún, si cualquier hipótesis es $\succ$, entonces la conclusión es $\succ$.
(7) Toda secuencia estándar es finita, donde $\{A_i\}$ es una secuencia estándar syss, para todo $i, A_i\ \in\ F\ \ \setminus\ N, A_i\ \subset\ A_i + 1, y\ S \mid S\ \succ\ A_i \mid A_i + 1\ \sim A_1 \mid A_2.$

Las condiciones (1)-(7) son necesarias pero no suficientes para la existencia de una representación. KLST (1971) introduce un axioma que completa los axiomas y arroja la representación. El axioma es una especie de axioma de resolubilidad que afirma que la estructura es suficientemente rica para que, siempre que $A \mid B\ \succ\ C \mid D$, sea posible agregar suficiente de C para hacer que $C' \mid D$ sea equivalente a $A \mid B$. En términos más precisos,

(8) Si $A \mid B\ \gtrsim\ C \mid D$, entonces existe $C' \in F$ tal que $C \cap D\ \subset C'\ y\ A \mid B\ \sim C' \mid D$.

Un análisis inicial del PGD puede revelar intuitivamente alguna de las dependencias entre los eventos, pero, como es usual en la aplicación de cualquier teoría científica, los métodos de determinación de los términos de una teoría pueden ser más bien indirectos (son llamados, en econometría, 'inferencia estadística'). Partiendo de las estructuras de probabilidad condicional cualitativa, mostraré cómo se construye el aparato de la econometría. A grandes rasgos, los pasos son los siguientes:

(1) Un sistema económico aleatorio es identificado como objeto de estudio, y convertido en un sistema objetivo por el científico.

(2) El objetivo —también conocido como PGD— es descrito de un modo cualitativo, no teórico, usando el lenguaje ordinario y la noción de probabilidad condicional cualitativa.
(3) Usando los teoremas de representación, se reconoce que el sistema objetivo puede ser representado mediante un espacio de probabilidad estándar.
(4) Dentro del espacio de probabilidad estándar, se introduce una variable aleatoria X, y el espacio de probabilidad original es reemplazado por el espacio de Borel inducido por X.
(5) Las distribuciones de probabilidad de las variables aleatorias se suponen integrables a partir de una familia de funciones de densidad paramétricas con parámetros desconocidos $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. El conjunto de estas funciones es llamado el *modelo de probabilidad*; es denotado como $\Phi = \{D(y;\boldsymbol{\theta}), \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$.
(6) Suponiendo que las leyes (probabilísticas) de la teoría son satisfechas por el sistema objetivo, se da por sentado que los datos observados (y potencialmente infinitamente muchos más) son generados por el mecanismo aleatorio representado por la densidad "verdadera". La meta de la teoría es encontrar el parámetro verdadero que define esta densidad.
(7) Finalmente, un modelo muestral se introduce definiendo una muestra de la densidad $D(y;\boldsymbol{\theta_0})$, para algún parámetro "verdadero" $\boldsymbol{\theta_0} \in \boldsymbol{\Theta}$.

Para encontrar el parámetro verdadero, se supone que los datos observados constituyen una realización $\boldsymbol{Z} \equiv (\boldsymbol{Z}, \ldots, \boldsymbol{Z}_T)$ de una secuencia de vectores aleatorios, o procesos estocásticos $\{\boldsymbol{Z}_t, t \in T\}$. De acuerdo con Spanos (1986: 665):

> Esta suposición proporciona el vínculo necesario entre el PGD real y la teoría de probabilidad. Nos permite postular una estructura probabilística para $\{\boldsymbol{Z}_T, t \in T\}$ en la forma de una función de distribución conjunta $D(\boldsymbol{Z}, \boldsymbol{\psi})$. Esto proporcionará la base para la especificación del modelo estadístico.

7. De la probabilidad condicional a la absoluta

Aun cuando las probabilidades condicionales son más intuitivas, es necesario transitar de ellas a las probabilidades incondicionales para obtener el modelo de probabilidad relevante, requerido para determinar efectivamente los valores de las medidas de probabilidad. Los axiomas que definen las estructuras de probabilidad condicional cualitativa, junto con el axioma (8), implican la siguiente proposición.

TEOREMA 4. Supóngase que $\boldsymbol{Q} = \langle S, F, N, \succsim \rangle$ es una estructura de probabilidad condicional cualitativa arquimediana para la cual vale el axioma 8. Entonces existe una función de valores reales P sobre F tal que para todo $A, C \in F$ y todo $B, D \in F \setminus N$:

i. $\langle S, F, P \rangle$ es un espacio de probabilidad finitamente aditivo;
ii. $A \in N$ syss $P(A) = 0$;
iii. $A \mid B \succsim C \mid D$ syss $P(A \cap B)/P(B) \geq P(C \cap D)/P(D)$.

Mas aún, si F es un σ-álgebra y $\boldsymbol{Q}$ satisface la condición

(iv) Si $B \mid S \succsim A_i \mid S$ entonces $B \mid S \succsim \coprod_{i \in \omega} A \mid S$

Para todo $B \in F$ y cualquier colección $\{A_i\}_{i \in \omega}$ de elementos de F, entonces $\langle S, F, P\rangle$ es contablemente aditivo.

Demostración: Por el argumento en la sección 5.7 de (KLST: 228–232), existe una función P que satisface las condiciones (i)-(iii) del Teorema 4. Defínase la relación $\succsim'$ por la condición

$$A \succsim' B \text{ syss } A \mid S \succsim B \mid S.$$

Por el teorema 9 en (KLST: 232), $\boldsymbol{Q} = \langle S, F, \succsim'\rangle$ es una estructura arquimediana de probabilidad cualitativa. Por ende, por el teorema 2 (KLST: 208), existe una única función preservadora del orden P' de F en [0;1] tal que $\langle S, F, P'\rangle$ es un espacio de probabilidad finitamente aditivo. Como P' es única, $P' = P$.

Supóngase ahora que $\boldsymbol{Q}$ satisface la condición (iv), sea $A_1, A_2, \ldots A_i, \ldots$ cualquier secuencia en F y B cualquier elemento de F tal que $A_i \subset A_{i+1}$ y $B \succsim A_i$ para todo i. Esto equivale, por la definición de $\succsim'$, a $B \mid S$ o $A_i \mid S$ para cada i y así, por (iv), $B \mid S \succsim \coprod_{i \in \omega} A_i \mid S$; $i.e., B \succsim' \coprod_{i \in \omega} A_i$. Esto es decir que $\succsim'$ es monotónicamente continua (cfr. Definición 6, KLST: 216) y por lo tanto, por el teorema 4 en (KLST: 216), $\langle S, F, P\rangle$ es contablemente aditiva.

El Teorema 4 proporciona la base para representar cualquier estructura de probabilidad condicional cualitativa mediante un modelo de probabilidad, una vez que las variables aleatorias relevantes son especificadas.

8. Inferencia estadística

Una vez que ha sido definido el modelo de probabilidad Φ, es necesario determinar los valores numéricos de sus parámetros para introducir un modelo estadístico. Éste se puede definir como sigue (*cfr*. Spanos 1986: 216-8).

DEFINICIÓN 7. Un modelo estadístico es un par $\langle \Phi, \boldsymbol{X}\rangle$ tal que

(1) Φ es un modelo de probabilidad; i.e., una familia $\{D(x; \boldsymbol{\theta}), \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ de funciones de densidad.

(2) $\boldsymbol{X}$ es un conjunto de variables aleatorias $(X_1, \ldots, X_n)$ cuyas funciones de densidad coinciden con la función de densidad "verdadera" $f(\boldsymbol{x}; \boldsymbol{\theta_0})$ tal y como está postulada por el modelo de probabilidad.

(3) La distribución de $\boldsymbol{X}$ es la distribución conjunta de las v.a.'s $X_1, \ldots, X_n$, denotada por
$f\ (x_1, \ldots, x_n)\ , \boldsymbol{\theta}) = f\ (\boldsymbol{x}; \boldsymbol{\theta})$:

Los datos observados $(x_1, \ldots, x_n)$ son interpretados como una especificación particular del modelo aleatorio, i.e., $\boldsymbol{X} = \boldsymbol{x}$, y luego procesada computando $f\ (x_1, \ldots, x_n, \boldsymbol{\theta})$. La información obtenida permite la consideración de las siguientes preguntas, las cuales delinean las áreas principales de la inferencia estadística:

(1) ¿Son los datos observados consistentes con el modelo estadístico postulado? (*mala especificación*)

(2) Suponiendo que el modelo estadístico postulado es consistente con los datos observados, ¿qué podemos inferir acerca de los parámetros desconocidos $\boldsymbol{\theta} \in \boldsymbol{\Theta}$?
(a) ¿Podemos disminuir la incertidumbre acerca de **θ** reduciendo al espacio de parámetros de **Θ** a $\mathbf{\Theta_0}$, donde $\mathbf{\Theta_0}$ es un subconjunto de **Θ**? (*estimación de la confianza*)
(b) ¿Podemos disminuir la incertidumbre acerca de **θ** eligiendo un valor particular en **Θ**, digamos $\hat{\boldsymbol{\theta}}$, como proporcionando el valor más representativo de **θ**? (*estimación puntual*)
(c) ¿Podemos considerar la pregunta de que **θ** pertenece a algún subconjunto $\mathbf{\Theta_0}$ de **Θ**? (*prueba de hipótesis*).

(3) Suponiendo que un valor representativo particular $\hat{\boldsymbol{\theta}}$ de **θ** ha sido elegido, ¿qué podemos inferir acerca de ulteriores observaciones del PGD como es descrito por el modelo estadístico postulado? (*predicción*) (Spanos 1986: 221)

Es particularmente interesante observar que responder la segunda pregunta presupone que el modelo estadístico postulado (ya no digamos los axiomas de la teoría de la probabilidad) es válido, "y considera varias formas de inferencia que se relacionan con parámetros desconocidos **θ**". Desde luego, esto es típico de los términos T-teóricos.

9. La estructura lógica de la econometría

¿Cuál es, entonces, la estructura lógica de la teoría econométrica (TE)? ¿Cuál es su elemento teórico básico, sus especializaciones, su dominio de aplicaciones pretendidas, sus términos T-teóricos?

El dominio de pretendidas aplicaciones de TE es vasto, ya que muchos fenómenos aleatorios económicos han sido exitosamente sometidos a análisis econométricos. Dejaré como un problema abierto la tarea de clasificar estos fenómenos. Mi tarea en esta sección final es meramente la de identificar la estructura lógica del elemento teórico básico de TE. No obstante, podemos decir algo relevante acerca de la naturaleza general de la estructura no-T-teórica $\boldsymbol{B_\sigma}$ que describe un PGD particular, real-concreto σ. Tomando como punto de partida los datos empíricos $\boldsymbol{D} = (x_1, \dots, x_n)$ acerca de σ, el econometrista traza un histograma y un histograma acumulativo que describen un efecto particular de las fuerzas causales aleatorias que operan en **σ**. A partir de la información proporcionada por tales histogramas, calcula características numéricas que describen la ubicación, la dispersión y la configuración del histograma. Estos son la media, la mediana, la moda, la varianza, la desviación estándar, los momentos centrales más altos, y la torcedura y los coeficientes de la kurtosis. No obstante, como Spanos (1986: 27) señala, "aunque el histograma puede ser muy útil para resumir y estudiar los datos observados, no es un descriptor muy conveniente de los datos". Mas aún, analíticamente el histograma es una torpe función de escalones de la forma:

$$h(z) = \sum_{i=1}^{m} \frac{\Phi_i}{(z_{i+1} - z_{i)}} I\,([z_i, z_{i+1}]), \quad z \in \mathbb{R}$$

donde $[zi,\ z_{i+1})$ representa el i-ésimo intervalo semicerrado e I(.) es la función indicador.

$$I([z_i, z_{i+1})) = \begin{array}{l} 1 \text{ para } z \in [z_i, z_{i+1}) \\ 0 \text{ para } z \notin [z_i, z_{i+1}). \end{array}$$

Por ende, el histograma no es un descriptor ideal especialmente en relación con la faceta de modelado de los datos observados. (Ibídem)

Aunque es fácil obtener una curva continua que trace el polígono de frecuencias, el modo canónico de representar los datos es una curva de frecuencia Pearson, la cual es una familia basada en la ecuación diferencial

$$\frac{d(\log \varphi(z))}{dz} = \frac{z+a}{b_0 + b_1 z + b_2 z^2}$$

(para mayores detalles véase Spanos 1986: 28). La estructura no-T-teórica $\boldsymbol{B_\sigma} \in M_{pp}$ que describe la aplicación pretendida en caso es una curva de Pearson que encaja los datos en un grado aceptable a la comunidad de econometristas.

Las curvas de Pearson son análogas, por ejemplo, a las elipses que describen las órbitas planetarias como las describe Johannes Kepler. Así como estas órbitas son descritas sin presuponer nociones dinámicas, las curvas de Pearson son introducidas sin presuponer en su determinación ninguna medida de probabilidad.

Una curva de Pearson describe un comportamiento factual particular de una población *S*. Por ende, $\boldsymbol{B_\sigma}$ parece ser de la forma $\langle S, f^* \rangle$, donde *f** es una curva de frecuencia, que es una función no-estocástica de los datos observados, basada enteramente en los datos a la mano (en $\boldsymbol{D}$).

La meta científica del econometrista es "explicar" $\boldsymbol{B}_\sigma = \langle S, f* \rangle$ mediante la teoría de probabilidades, mostrando que $\boldsymbol{B}_\sigma$ no es más que una posible realización de los procesos aleatorios que tienen lugar en **σ**. El punto es encontrar el espacio de probabilidad al que $\boldsymbol{B_\sigma}$ pertenece. Esto se hace postulando un modelo de probabilidades junto con un modelo muestral, derivado del modelo estadístico, que describe la relación entre el modelo de probabilidades y los datos observados. "El modelo estadístico debiera representar una buena aproximación al fenómeno real a ser explicado de un modo que tome en cuenta la naturaleza de los datos disponibles" (Spanos 1986: 341). Usando los datos se determinan los parámetros del modelo estadístico y, de esta manera, también las funciones de densidad del modelo de probabilidad. Claramente, estas funciones permiten también, por integración, la determinación de las medidas de probabilidad de los espacios de probabilidad subyacentes. En particular, *f** debe resultar ser un caso especial de la función de densidad verdadera en el espacio de probabilidad. Por ende, parece que podemos sintetizar la discusión previa mediante una definición compacta.

DEFINICIÓN 8. $\boldsymbol{A}$ es una *estructura econométrica* syss existen *S*, *f**, Φ y $\boldsymbol{X}$ tales que

(0) $\boldsymbol{A} = \langle S, f^*, \Phi, \boldsymbol{X} \rangle$.
(1) S es un conjunto no vacío.

(2) f* es una curva de Pearson sobre un conjunto **X** de variables aleatorias definidas sobre S.
(3) Φ es un modelo de probabilidades; *i.e.*, una familia $\{D(\boldsymbol{x};\boldsymbol{\theta}), \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ de funciones de densidad.
(4) La curva de Pearson sobre $(X_1, \dots, X_n)$ coincide con la "verdadera" función de densidad $f(\boldsymbol{x};\boldsymbol{\theta_0})$ tal y como es postulada por el modelo de probabilidad; *i.e.*,
$f^*(x_1, \dots, x_n) = f(\boldsymbol{x};\, \boldsymbol{\theta})$:

El axioma 8(4) es la ley fundamental de TE. La forma general de su aserción empírica es que las curvas de Pearson obtenidas a partir de un fenómeno aleatorio particular **σ** pueden encajar en una función de densidad de una variable aleatoria definida sobre un espacio de probabilidades apropiado.

Referencias

CALDWELL, B. (ed.) (1984). *Appraisal and Criticism in Economics*. Boston: Allen & Unwin.
DARNELL, A. & J. EVANS (1990). *The Limits of Econometrics,* Cheltenham: Edward Elgar Publishing.
DAVIS, G. C. (2000). "A Semantic Interpretation of Haavelmo's Structure of Econometrics". *Economics and Philosophy* 16: 205–228.
HENDRY, D. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.
KRANTZ, D. H., R. D. LUCE, P. SUPPES, AND A. Tversky (1971). *Foundations of Measurement, Volume I: Additive and Polynomial Representations*. New York: Academic Press.
KUHN, TH. S. (1970). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
LAKATOS, I. (1973). "Falsification and the Methodology of Scientific Research Programmes". In I. Lakatos and A Musgrave (eds.) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 91-196.
LUCE, R. D. (1968). "On the Numerical Representation of Qualitative Conditional Probability". *Annals of Mathematical Statistics* 39(2): 481-491.
SPANOS, A. (1986). *Statistical Foundations of Econometric Modelling*. Cambridge: Cambridge University Press.
SUPPE, F. (1989). *The Semantic Conception of Theories and Scientific Realism*. Urbana and Chicago: University of Illinois Press.
SUPPES, P. (1984). *Probabilistic Metaphysics.* Oxford: Basil Blackwell.