APRENDIZAJE E INFORMACIÓN SOBRE LOS PARÁMETROS DE PREFERENCIAS

*Miguel Ángel Sámano Rodríguez**
*Francisco Venegas-Martínez***

Resumen. El presente trabajo utiliza el enfoque bayesiano de aprendizaje para medir la cantidad de información que un individuo gana sobre su función de utilidad después de consumir una cantidad inicial de algún bien genérico. Para ello, se desarrolla un modelo de aprendizaje bayesiano sobre los parámetros que determinan las preferencias de un individuo racional. En este marco, la información de Fisher (1950) se utiliza como una medida de la cantidad de información que contienen los parámetros del índice de satisfacción del agente después de que éste ha iniciado su consumo. A partir de la información de Fisher se construye la distribución *a priori* de los parámetros siguiendo la regla de Jeffreys (1961), lo cual permite tratar a la información como una variable.

Abstract. This paper uses a Bayesian learning model to measure the amount of information that an individual gains on his/her utility function after consuming an initial amount of a generic good. To do so, a learning model on the parameters determining the preferences of a rational individual is developed. In this framework, Fisher's information (1950) is used to measure the amount of information contained in the parameters of the agent's satisfaction index once consumption has started. By using Fisher's information and Jeffreys' rule (1961) *a priori* distribution on the parameters is constructed, which allows to treat information as a variable.

---

* Profesor de la Universidad Autonoma Metropolitana, Unidad Cuajimalpa. Correo electrónico: masr73@yahoo.com.

** Profesor de la Sección de Estudios de Posgrado e Investigación de la Escuela Superior de Economía del Instituto Politécnico Nacional. Correo electrónico: fvenegas@ipn.mx.

1. Introducción

El estudio de la información en teoría económica es indudablemente extenso. Al respecto, es importante mencionar las investigaciones seminales de Stigler (1961), Spence (1974) y Akerlof (1970), así como las más recientes sobre el proceso del aprendizaje de Wieland (2000) y Kelly y Kolstad (1999). Sin embargo, la mayor parte de estas investigaciones se han concentrado en los efectos que causa la ausencia de información en los mercados y, más aún, en estos trabajos la información es tratada como un supuesto y no como una variable.

En la literatura especializada se encuentran muy pocos estudios que tratan a la información como una variable, como ha sido el trabajo de Arrow (2000). En dicha investigación se señala que si bien se ha avanzado en el estudio de la información como un supuesto, al no tratarla como una variable es imposible observar directamente las consecuencias de la escasez de la misma. Además, el enfoque que considera a la información como supuesto no concede suficiente importancia a las alteraciones en la información debidas a cambios en las decisiones económicas. Evidentemente, la única forma de capturar dichas alteraciones es tratando a la información como una variable. Si bien los avances logrados con el tratamiento de la información como supuesto son innegables, los mismos son limitados, pues evitan la posibilidad de observar en acción a la información. Al respecto, sólo unos cuantos trabajos se han enfocado a dicho tratamiento al estudiar el aprendizaje de los agentes sobre sus funciones de utilidad, como el trabajo de Venegas-Martínez (2005). No obstante, falta mucho para avanzar en la formalización de la información como variable.

El presente trabajo utiliza el enfoque bayesiano de aprendizaje para cuantificar la información que un individuo puede ganar al consumir una cantidad inicial de un bien genérico, lo cual, a su vez, le revela información sobre sus preferencias. Con el propósito de aplicar el análisis bayesiano, se considera que el aprendizaje surge de la experimentación y que la información provista por el experimento puede ser cuantificada, véanse al respecto, por ejemplo: Bernardo (1979), Carlin y Polson (1991), Good (1960), Lindley (1956), Mallows (1959) y Polson (1992). Por otro lado, el consumo puede ser visto como un proceso de experimentación de acuerdo con: Aghion, Bolton y Jullien (1991), Grossman, Kihlstrom y Mirman (1977) y Venegas-Martínez (2000).

En el análisis bayesiano, el consumidor expresa el conocimiento ganado a través de una distribución *a priori*, la cual es actualizada cada vez que llega nueva información mediante el teorema de Bayes, lo que da paso a una nueva distribución (*a posteriori*). Cada actualización deseablemente se debería aproximar cada vez más a la distribución verdadera del parámetro, véanse, por ejemplo, Hacking (1995), Hacking (2001) y Robert (2001).

2. Aprendizaje

Uno de los primeros planteamientos sobre la cantidad de información provista por un experimento es debido a Lindley (1956), quien define la cantidad de información provista por un experimento ε, cuando se observa x, como

$$I(\varepsilon, x) = I_1(x) - I_0(x)$$

donde $I_1(x)$ es la cantidad de información una vez realizado el experimento e $I_0(x)$ es el conocimiento previo del experimentador (consumidor). El caso del experimentador que aprende sobre un parámetro de interés desde la observación de una variable ha sido también estudiado por Mallows (1959). Para evaluar qué tan informativo es un experimento, este autor utiliza la distancia de Kullback-Leibler (1951) y sugiere el empleo de la información de Fisher (1950). De esta forma, la distancia Kullback-Leibler es empleada para evaluar dos experimentos (procesos de aprendizaje) y se desea establecer cuál de ellos es más informativo. Esto es, se evalúa un experimento con n variables y otro con $n-1$ variables, y se desea saber si la consideración de la *n*-ésima variable es realmente determinante en el aprendizaje del individuo. No obstante, cuando se trata de un solo experimento, no es posible comparar las distribuciones que generan ambos experimentos, lo que constituye una limitación superable empleando la información de Fisher.

Por otro lado, Polson (1992) emplea la información de Fisher para cuantificar la información que un experimentador puede obtener acerca del parámetro de una distribución a través de la observación. Además, muestra que el experimentador expresa el conocimiento adquirido en una distribución *a priori* obtenida con la regla de Jeffreys. Este tratamiento de la información se encuentra en los trabajos de Lindley (1956), Mallows (1959), Kempthorne (1986), Polson (1992), Bernardo (1979), Carlin y Polson (1991), Good (1960) y Tanner y Wong (1987), los cuales tienen en común los siguientes aspectos: primero, determinan la cantidad de información que proveen diferentes experimentos; segundo, verifican si la cantidad de información sufre alguna variación al observarse un conjunto de datos o si no se observa; tercero, establecen si existen experimentos más informativos que otros. Este análisis puede adaptarse al estudio de los agentes económicos y ha sido tratado por Grossman, Kihlstrom y Mirman (1977), Venegas (2000), Kelly y Kolstad (1999) y Wieland (2000). Sin embargo, la cuantificación de la información aún no ha sido abordada. Para realizar esto vamos a suponer que:

a) El agente, al consumir, está llevando a cabo un experimento que le permite acumular información sobre un vector de parámetros, ω, de su función de utilidad.
b) El proceso de consumir genera información sobre dicho parámetro.
c) La información ganada puede expresarse en una distribución *a priori* del parámetro.

Posteriormente, el agente incorpora la información ganada en la distribución *a posteriori* para hacer inferencias sobre los valores de los parámetros. En lo que sigue, el enfoque bayesiano se vincula con la teoría económica tratando a la información como una variable y no como un supuesto. A su vez, esta aproximación permite evaluar la pertinencia de la aplicación de dicho enfoque en la teoría económica.

3. Aprendizaje sobre parámetros de preferencia debido al consumo inicial

En esta sección se estudia el proceso de aprendizaje sobre los parámetros de las preferencias debido al consumo inicial. La información que le es proporcionada al agente a través del consumo debe ser tal que:

a) El conocimiento previo del agente es nulo.
b) El conocimiento del individuo surge del proceso de consumo y la información puede expresarse en una distribución *a priori* del conocimiento.

La consideración *a*) supone que el agente está en su primer día en el mercado, por lo cual no tiene conocimiento previo sobre el mismo. La consideración *b*) supone que dicho agente invierte en consumo para conocer el comportamiento del mercado. De hecho, no existe otra forma de aprender sobre el mercado. Además, el consumidor es un agente maximizador de utilidad que dispone de ciertos activos los cuales constituyen su riqueza real. Por lo cual, el individuo seguirá un patrón de consumo óptimo, el cual le provee la información que requiere para conocer el mercado; en cierto sentido se tiene un proceso generador de datos. Posteriormente, con base en el patrón óptimo de consumo, se cuantificará la información con que cuenta el agente.

3.1 Modelo de aprendizaje

En esta sección se desarrolla el modelo de aprendizaje, el cual está asociado a un proceso óptimo de consumo, mismo que constituye la fuente de información desde la cual el agente acumula conocimiento sobre los parámetros de sus preferencias. Para saber más sobre el conocimiento que el agente alcanza después de consumir, se empleará el análisis bayesiano, el cual permite cuantificar la información y representarla a través de una distribución *a priori*. Dicha distribución se construye a partir de la información acumulada siguiendo la regla de Jeffreys.

3.1.1 Descripción del agente

El agente, el cual tiene vida infinita y es racional, es un individuo representativo de una economía que produce y consume un bien genérico. El agente posee un activo (un bono) que paga una tasa de interés constante a todos los plazos y libre de riesgo de incumplimiento, lo cual constituye todo su ingreso. Este individuo experimenta, al consumir, lo que le retribuye conocimiento sobre las características de sus preferencias. La utilidad total, descontada del individuo, es:

$$V = \int_0^\infty u(c,\theta)e^{-\rho t}\,dt \tag{1}$$

donde ρ es la tasa subjetiva de descuento. La restricción presupuestal del individuo está dada por:

$$\frac{da}{dt} = ar - c$$
$$\text{con } a(0) = a_0 \tag{2}$$
$$a_0 \text{ conocido}$$

En lo que sigue se supone que el comportamiento del consumidor queda expresado en la función de utilidad con la forma funcional

$$u(c,\theta) = \frac{c^{1-\theta}}{1-\theta}, \ \theta \neq 1$$

Se supone también que el agente no tiene conocimiento previo sobre θ, así que aprenderá sobre dicho parámetro. El individuo maximiza utilidad total descontada en un horizonte de planeación infinito, específicamente el intervalo $[0,\infty]$, sujeto a su restricción presupuestal. Por lo tanto, el problema del agente es maximizar

$$V = \int_0^\infty u(c,\theta)e^{-pt}\,dt$$

sujeto a: (3)

$$\frac{da}{dt} = ar - c$$
$$\text{con: } a\,(0) = a_0$$
$$a_0 \text{ conocido}$$

El activo a en posesión del agente, puede pensarse como un bono que paga la tasa *r*, c es el consumo y ρ es la tasa subjetiva de descuento también llamada tasa de sustitución intertemporal del agente.

El conocimiento debe ser reflejado por la trayectoria óptima de consumo c^*. En este caso, la información provista es determinada empleando la información de Fisher sobre la trayectoria óptima de consumo, para lo cual es necesario normalizar dicha trayectoria. Observe que las trayectorias ordinaria y normalizada conducen exactamente a la misma tasa de crecimiento del consumo, siendo ésta la variable relevante. La trayectoria óptima de consumo es el proceso que provee de información al agente de forma óptima; sin embargo, la información que describe requiere ser cuantificada, dicha cuantificación podría tomar como base el intervalo [0,1]. Esto permitirá la homogeneización de la medida y la comparación con otros procesos informativos. Evidentemente, existen diversos procesos que le pueden proveer de información al agente, uno de ellos puede ser el consumo, pero puede existir algún otro, como la tenencia de bonos (que usualmente crece a la misma tasa que el consumo); ambos procesos proveerán información, pero en términos de niveles no es posible reconocer cuál de ellos es más o menos informativo, la normalización ofrece una salida favorable, pues por un lado, permite que los procesos puedan expresarse informativamente en un rango bien delimitado, que vaya de nulidad informativa, es decir, cero, hasta *totalidad* informativa, la unidad. Además, con esta normalización se supera la necesidad de establecer una unidad de medida para la información, con lo cual cualquier proceso informativo se hace comparable con otro, independientemente de que la información surja de diferentes variables, ya que, cuando un agente observa una variable, está observando una señal informativa que toma un valor en las unidades que la describen a ella, y lo mismo ocurre con otra variable cualquiera, así que la normalización permite situar la informatividad de las variables en una rango bien definido. Además, la normalización conduce a que el proceso se exprese como si fuera una función de densidad, como si fuera un modelo de probabilidad, que finalmente es el que genera la información (Shannon, 1948) y permite emplear la teoría bayesiana del aprendizaje (Bernardo, 2002), así como la matemática de la teoría de la información.

La teoría bayesiana emplea la información de Fisher como la medida que cuantifica la información que provee un modelo asociado a un conjunto de datos y trabaja directamente sobre la función de densidad. Esto constituye una gran ventaja, pues permite cuantificar directamente la información del proceso, además de las ventajas indicadas antes con respecto de su normalización, la información de Fisher está expresada, en forma general, por

$$I(\omega) = -E_x\left[\frac{\partial^2 \ln \hat{c}(x \mid \omega)}{\partial \omega^2}\right] \tag{4}$$

donde $x = (r,t)$ y $\omega = (\theta, \rho)$. La trayectoria de consumo normalizada $\hat{c}$, debe cumplir con la condición $\int \hat{c}\, dx = 1$. El vector de parámetros sobre los que se acumula información es ω y la información es provista por el vector de variables x.

La información acumulada por el agente se expresa en una distribución *a priori* sobre el vector de parámetros ω, la cual se obtiene mediante la distribución *a priori* de Jeffreys, $\pi^*(\omega) \propto [\det I(\omega)]^{1/2}$, para el caso multiparamétrico. Así pues, el agente desea acumular información sobre los parámetros ρ y θ y reflejar el conocimiento acumulado a través de la distribución de Jeffreys $\pi^*(\rho,\theta) \propto [\det I(\rho,\theta)]^{1/2}$.

3.1.2 Trayectoria óptima de consumo

Para conocer la trayectoria óptima de consumo que le proporciona información al individuo se tiene que resolver el problema de maximización. El hamiltoniano (en valor presente) está dado por

$$H_A = \frac{c^{1-\theta}}{1-\theta} + e^{\rho t}\gamma\left(ra - c\right) \tag{5}$$

Si se escribe $\lambda = \gamma e^{\rho t}$, entonces el hamiltoniano anterior se puede expresar como

$$H_A = \frac{c^{1-\theta}}{1-\theta} + \lambda(ra - c) \tag{6}$$

Para obtener las condiciones de primer orden de una solución interior se recurre al principio de Pontryagin, el cual conduce a

$$\frac{\partial H_A}{\partial c} = 0, \quad \frac{da}{dt} = \frac{\partial H_A}{\partial \lambda} \quad \text{y} \quad \frac{d\lambda}{dt} = \frac{-\partial H_A}{\partial a} + \rho\lambda$$

La primera condición

$$\frac{\partial H_A}{\partial c} = 0$$

implica que

$$c = \frac{1}{\lambda^{1/\theta}} \tag{7}$$

Asimismo, la segunda condición

$$\frac{da}{dt} = \frac{\partial H_A}{\partial \lambda}$$

lleva a

$$\frac{da}{dt} = ra - c \tag{8}$$

Por último, la tercera condición

$$\frac{d\lambda}{dt} = \frac{-\partial H_A}{\partial a} + \rho\lambda$$

produce

$$\lambda = \lambda_0 e^{t(\rho - r)} \tag{9}$$

Al sustituir (9) en la ecuación (7), se sigue que

$$c = \frac{1}{\lambda_0^{(1/\theta)} e^{t(\rho - r)/\theta}} \tag{10}$$

A partir de (8) se tiene una ecuación diferencial por factor integrante, la cual conduce a

$$\int_0^\infty e^{-rt} \frac{da}{dt} dt - \int_0^\infty e^{-rt} \; ra\, dt = -\int_0^\infty e^{-rt} \; c\, dt \tag{11}$$

Al sustituir (10) en la ecuación (11), se sigue que

$$\int_0^\infty e^{-rt} \frac{da}{dt} - \int_0^\infty e^{-rt} \; ra\, dt = -\int_0^\infty \frac{e^{-rt}}{(\lambda_0^{1/\theta})\, e^{\,t(\rho - r)/\theta}} dt \tag{12}$$

Al resolver la ecuación anterior, se llega a

$$a_0 = \frac{\theta}{\lambda_0^{1/\theta}(\rho + r\theta - r)} \tag{13}$$

Al despejar $\lambda_0^{1/\theta}$ de (13) y sustituir en (10) , se obtiene la trayectoria óptima de consumo

$$c^* = \frac{a_0(\rho + r\theta - r)}{\theta} e^{-t(\rho-r)/\theta} \tag{14}$$

Se requiere normalizar la trayectoria óptima de consumo, para lo cual se multiplica (14) por

$$\frac{\rho - r}{a_0(\rho - r + r\theta)}$$

En este caso, se obtiene

$$\hat{c} = \left(\frac{\rho - r}{\theta}\right) e^{-t(\rho-r)/\theta}, \quad \text{para} \quad 0 \le t \le \infty \text{ y } 0 \le r \le 1, \tag{15}$$

con lo que verifica,

$$\int_0^1 \int_0^\infty \left(\frac{\rho - r}{\theta}\right) e^{-t(\rho-r)/\theta}\, dt\, dr = 1, \text{ para } 0 \le t \le \infty \text{ y } 0 \le r \le 1$$

Otra posibilidad para obtener una trayectoria de consumo óptima ya normalizada es suponer que

$$a_0 = \frac{\rho - r}{(\rho - r + r\theta)}$$

Después de sustituir este valor de a_0 en (14), se obtiene inmediatamente (15).

3.1.3 Información adquirida

La información provista por el patrón óptimo de consumo puede ser cuantificada por la información de Fisher. Con la ventaja siguiente: la información de Fisher puede emplearse para determinar la distribución *a priori* que refleja el conocimiento acumulado por el individuo sobre un conjunto de parámetros. La variable que proporciona información al agente es la tasa de interés, r, misma que tomará valores entre cero y uno. Este supuesto es restrictivo, pero aceptable en el sentido que en una economía estable los episodios con tasas de interés superiores son transitorios y no permanentes.

El agente requiere de información para fijar los valores de los parámetros $\rho\ y\ \theta$ para finalmente tomar decisiones. La información de Fisher, en este caso biparamétrico, está determinada por

$$I(\rho,\theta)=-E_{r,t}\begin{pmatrix}\dfrac{\partial^2 \ln\hat{c}}{\partial\rho^2} & \dfrac{\partial^2 \ln\hat{c}}{\partial\rho\,\partial\theta}\\ \dfrac{\partial^2 \ln\hat{c}}{\partial\theta\,\partial\rho} & \dfrac{\partial^2 \ln\hat{c}}{\partial\theta^2}\end{pmatrix}$$

Por lo tanto, la información provista por la trayectoria de consumo es

$$I(\rho,\theta)=-E_{r,t}\begin{pmatrix}\dfrac{-1}{(\rho-r)^2} & \dfrac{t}{\theta^2}\\ \dfrac{t}{\theta^2} & \dfrac{1}{\theta^2}-\dfrac{2t(\rho-r)}{\theta^3}\end{pmatrix} \tag{16}$$

Al calcular el valor esperado, se tiene

$$I(\rho,\theta)=\begin{pmatrix}\dfrac{1}{\rho\,(\rho-1)} & \dfrac{1}{\theta}(\ln\rho-\ln(\rho-1))\\ \dfrac{1}{\theta}(\ln\rho-\ln(\rho-1)) & \dfrac{1}{\theta^2}\end{pmatrix} \tag{17}$$

Después de seguir la regla de Jeffreys para establecer una distribución *a priori*, se sigue que $\pi^*(\rho,\theta)\propto[\det I(\rho,\theta)]^{1/2}$. Si se calcula el determinante, se obtiene

$$\det I(\rho,\theta)=\frac{1-(\ln\rho-\ln(\rho-1))^2(\rho^2-\rho)}{\theta^2(\rho^2-\rho)} \tag{18}$$

En consecuencia, se obtiene

$$\pi^*(\rho,\theta)\propto\frac{1}{|\theta|}\left[\frac{1-(\ln\rho-\ln(\rho-1))^2(\rho^2-\rho)}{(\rho^2-\rho)}\right]^{\frac{1}{2}} \tag{19}$$

La información ganada después de realizar el experimento de consumir se transforma en la distribución *a priori* sobre los parámetros que le son de interés al agente. Esta distribución muestra el conocimiento del agente.

3.2 Desempeño de la información

La información de Fisher es un número que refleja la información provista por las variables sobre los parámetros de la función de utilidad. Este número también arroja información sobre los posibles valores que pueden tomar dichos parámetros; a partir de la ecuación (18) se pueden establecer los intervalos en que cada parámetro puede tomar valores.

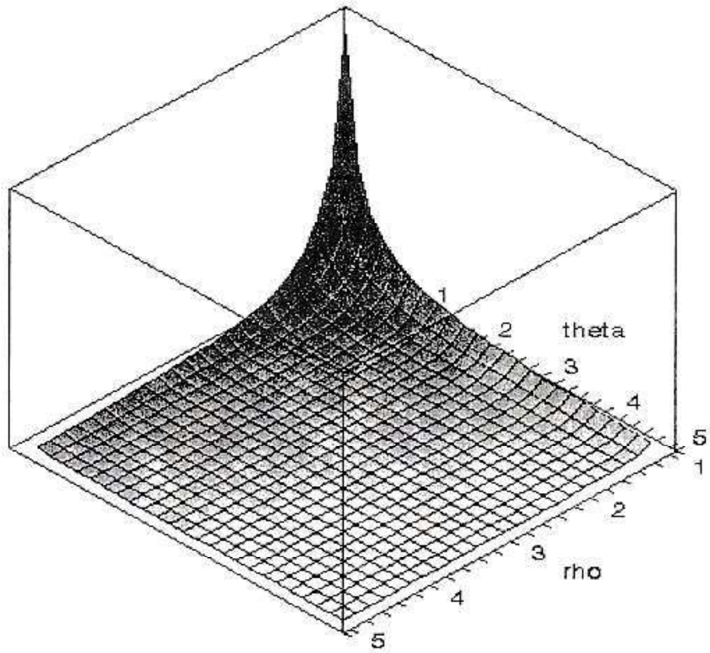
Al agente le será de interés poder asignar un valor determinado para los parámetros de su función de utilidad. La asignación de dicho valor estará predeterminada por la información adquirida y por la distribución *a priori* que describa el conocimiento inicial. Al considerar que se trata de un individuo racional, en el sentido que busca maximizar utilidad, evidentemente el conocimiento estará influido por la forma funcional que tome la utilidad.

Por lo anterior se ha establecido que el conocimiento del agente sigue la distribución determinada por la ecuación (19). Esta distribución, sin embargo, tan sólo representa el *kernel* de la distribución verdadera, para encontrar la distribución exacta una vez definido el espacio de los parámetros se tiene que calcular la constante de normalización. En la figura 1 se muestra la grafica del comportamiento del *kernel* de la distribución *a priori*, expresado en (19), para $\rho \in [0,5]$ y $\theta \in [1,5]$.

4. Aprender al consumir

Un agente informado, en un segundo periodo de consumo, establece sus preferencias con la información anterior y podría no acumular una cantidad de información superior a la previa. El consumo posterior irá refinando el

Figura 1
Comportamiento del *kernel* de la distribución
a priori con $\rho \in [0,5]$ y $\theta \in [1,5]$



conocimiento logrado y confirmando o rechazando los juicios formulados sobre las preferencias por los bienes.

Un tema recurrente en teoría bayesiana al considerar que los parámetros influyen sobre las variables, ha sido estudiado en Pérez y Berger (2000). Estos autores trabajan en la identificación del modelo predictivo que permita realizar las mejores estimaciones bayesianas; otro planteamiento de modelos predictivos puede encontrarse también en Bernardo y Berger (1992) y Bernardo (2002).

Alternativamente, el agente puede emplear el conocimiento de dos formas diferentes. Una de ellas es que conociendo la distribución de los parámetros, encuentre las distribuciones marginales y realice estimaciones puntuales sobre los parámetros. La otra forma de incorporar el conocimiento se relaciona con la teoría de la decisión bayesiana. En este sentido, Rafia y Schlaifer (2000) han establecido que la introducción del conocimiento previo debe realizarse mediante el valor esperado de la utilidad. La utilidad esperada y los modelos predictivos tienen la desventaja de que las integrales que resultan no siempre pueden resolverse analíticamente y el empleo de métodos numéricos para aproximar la solución privan de estudiar los efectos sobre las variables.

Conclusiones

La vinculación que se hace entre análisis bayesiano y teoría económica permite que el tratamiento de la información como una variable sea posible. A través de la información de Fisher, como una medida de información y la regla de Jeffreys, es posible establecer una distribución *a priori* que describa el conocimiento inicial. En este caso, el experimento de consumir representa la fuente de información sobre la cual se cuantifica el conocimiento.

Una primera conclusión consiste en señalar que la distribución *a priori* obtenida refleja el conocimiento acumulado sobre los parámetros de las preferencias. La conducta del consumidor estará guiada tanto por su búsqueda de maximizar la utilidad como por el conocimiento adquirido. La información ganada sobre las preferencias tendrá efectos en el mercado, es de esperarse que si todos los agentes están informados de la misma forma, entonces los precios podrían verse afectados. En ese sentido, si los precios han de reflejar la información en posesión de los agentes, la distribución que sigan los mismos debería ser similar a la distribución que representa el conocimiento del agente representativo.

Otra conclusión importante corresponde a la idea de vincular el análisis bayesiano con la teoría económica. En esta vinculación es importante destacar que las dificultades técnicas surgen cuando la distribución *a priori* obtenida no corresponde a las habitualmente empleadas en los problemas de aprendizaje bayesiano. Las distribuciones empleadas en la construcción de distribuciones *a priori* son ampliamente conocidas y los resultados suelen caer en terreno más o menos explorado. Frecuentemente, las distribuciones *a priori* que tienen como sustento el propio mecanismo de extracción de información empleado por el agente, no suelen corresponder a modelos de probabilidad comunes, lo cual conduce a problemas técnicos en su tratamiento posterior.

Por último, es relevante destacar que el tratamiento de la información como variable es incipiente y requiere de investigaciones más profundas. Algunas líneas de trabajo futuro versarán sobre la tasa de aprendizaje del agente, la información como disminución de la incertidumbre y la relación de los precios con el conocimiento *a priori* de los individuos.

Bibliografía

Aghion, H. and J. Aghion (1991). "Optimal Learning by Experimentation". *Review of Economic Studies*, núm. 58, pp. 621-654.

Akerlof, G. (1970). "The Market for Lemons: Quality Uncertainty and the Market Mechanism". *Quarterly Journal of Economics*, vol. 84, pp. 488-500.

Arrow, K. (2000). "Información técnica y estructura industrial". En Glenn Carroll y David Teece. 2000. *Empresas, mercados y jerarquías*. 1ª edición, Oxford University Press, Oxford México.

Bernardo, J. M. (1979). "Expected Information as Expected Utility". *The Annals of Statistics*, núm. I. 7, núm. 3, 686-690.

____________ (2002). Bayesian Statistics. Encyclopedia of Life Support Systems (EOLSS). Paris: UNESCO. Accessible on line: http://www.uv.es/bernardo/

Bernardo, J M. y J. Berger (1992). "On the Development of Reference Priors" en Bernardo, Berger, Dawid y Smith. *Bayesian Statistics* (1994), 1st edition, Oxford University press.

Carlin, B. and N. Polson (1991). "An Expected Utility Approach to Influence Diagnostics". *Journal of the American Statistical Association*, December 1991, vol. 86, vo. 416, pp. 1013-1021.

Fisher, R. A. (1950). *Contributions to Mathematical Statistics*, John Wiley.

Good, I. J. (1960). "Weight of Evidence, Corroboration, Explanatory Power, Information and Utility of Experiments". *Journal of the Royal Statistical Society. Series B(methodological)*, vol. 22, No. 2, pp. 319-331.

Grossman, S., R. Kihlstrom, and L. Mirman (1977). "A Bayesian Approach to the Production of Information and Learning by Doing". *Review of Economic Studies*, oct. 1977,No. **64**, 533-547.

Hacking, I. (1995). *El surgimiento de la probabilidad*. 1a edi., Editorial Gedisa, Barcelona, España.

_________ (2001). *An Introduction to Probability and Inductive Logic*. 1st edi., Cambridge University press.

Jeffreys, H. (1961). *Theory of Probability* (3rd edn), Oxford University Press, London.

Kelly, D. and C. Kolstad (1999). "Bayesian Learning, Growth and Pollution". *Journal of Economics Dynamics and Control*, No. **23**, pp. 491-518

Kempthorne, P. J. (1986). "Decision-Theoretic Measures of Influence in Regression". *Journal of the Statistical Society*, Series B, vol. 48, pp. 370-378.

Kullback, S. and R. A. Leibler (1951). "On Information and Sufficiency". *The Annals of Mathematical Statistics*, vol. 22, No. 1, pp. 79-86.

Lindley, D. V. (1956). "On a Measure of the Information Provided by an Experiment". *The Annals of the Statistics*, vol. 27, No. 4, pp. 986-1005.

Mallows, C. L. (1959). "The Information in an Experiment". *Journal of the Royal Statistical Society*, Series B (Methodological), vol. 21, No. 1, pp. 67-72.

Pérez, M. y J. Berger (2000). "Expected Posterior Prior Distributions for Model Selection". Working Paper, Duke University.

Polson, N. G. (1992). "On the Expected Amount of Information from a Non-Linear Model". *Journal of The Royal Statistical Society.* Series B(Methodological), vol. 54, No. 3, pp. 889-895.

Raiffa, H. and R. Schlaifer (2000). *Applied Statistical Decision Theory*. John Wiley and Sons, New York.

Robert, C. P. (2001). *The Bayesian choice*. 2nd edi., Springer Verlag, New York.

Shannon, C. E. (1948). "A Mathematical Theory of Communication", *Bell System Tech.* J., vol. 27, pp. 379-423.

Spence, M. (1974). "Job Market Signalling". *Quarterly Journal of Economics,* vol. 87, pp. 355-374.

Stigler, G. (1961). "The Economics of Information". *Journal of political economy*, vol. 69. pp. 213-225.

Tanner, M. and W. H. Wong (1987). "The Calculation of Posterior Distributions by Data Augmentation". *Journal of the American Statistical Association*, vol. 82, No. 398, pp. 528-540.

Venegas-Martínez, F. (2000). "Aprendizaje, utilidad y estabilización". *Gaceta de Economía*, vol. 5, núm. 10, pp. 153-169.

_________________ (2005). "Bayesian Inference, Prior Information on Volatility, and Option Pricing: A Maximum Entropy Approach". *International Journal of Theoretical and Applied Finance*, **8** (1), pp. 1-12.

Wieland, V. (2000). "Learning by Doing and the Value of Optimal Experimentation". *Journal of Economic Dynamics and Control*, No. 24, pp. 501-534.