

Llamado de tesis de magíster o doctorado.

## **Extendiendo modelos y experimentos de aprendizaje por reforzamiento para la inferencia de parámetros de preferencia**

El análisis científico de las elecciones humanas es fundamental para el diseño de sistemas de ingeniería y para el diseño de políticas públicas. En transporte, salud, marketing, economía y otras áreas, se estudian las preferencias por distintos productos, medios de transporte, programas de salud, etc., asumiendo que las personas comparan los atributos de cada uno y eligen el que les reporta mayor satisfacción esperada. Para ello, lo que se ha hecho históricamente es aplicar el enfoque de Maximización de Utilidad Aleatoria (*Random Utility Maximization*, RUM) (McFadden, 1974), que es un modelo matemático que permite saber y cuantificar, por ejemplo, cual es la disposición a pagar por disminuir un minuto de tiempo de viaje en bus o metro, y con ello cuantificar los beneficios sociales de proyectos reales. En estos campos, la recolección de datos reales de personas tomando decisiones representa un desafío importante, ya que con frecuencia los investigadores buscan analizar preferencias por productos que aun no están en el mercado o por niveles de atributos que aún no se observan en la realidad.

Una forma habitual de abordar este problema es mediante Experimentos de Elección Discreta (*Discrete Choice Experiments*, DCE), en los que los individuos eligen una alternativa entre un conjunto de opciones hipotéticas, definidas por ciertos atributos (por ejemplo, precio, calidad, duración o contaminación asociada), a lo largo de 8 a 10 tareas de elección. Mediante análisis econométrico, las decisiones tomadas permiten identificar los parámetros de preferencia (cuanto la persona valora cada atributo en su toma de decisiones).

Sin embargo, los DCE sufren de un problema significativo: en la vida real, los atributos de cada opción no están dados de antemano ni son directamente observables en un papel. Las personas, independientemente de sus características, tienen que aprender el valor de cada opción por ensayo y error. Por ejemplo, conocer el tiempo de viaje real y la probabilidad de encontrar asiento en una micro requiere tomarla múltiples veces en el mismo paradero. En los DCE, los individuos no experimentan realmente las alternativas ni enfrentan un proceso de aprendizaje, el cual sí está presente en muchas (si no en todas) las decisiones de la vida real.

En contraste con los DCE, las ciencias conductuales y la neurociencia han estudiado extensamente los mecanismos cognitivos y neuronales que las personas usan para aprender el valor que les reportan distintas opciones. Para ello utilizan experimentos de aprendizaje, en los que los participantes realizan 100 o más elecciones y deben ir aprendiendo qué opción prefieren en base al refuerzo o recompensa obtenida (e.g., Erev et al., 2010; Henríquez-Jara et al., 2025). Estos experimentos permiten que los participantes reciban retroalimentación a partir de sus decisiones, actuando como una señal de reforzamiento para aprender el valor asociado a cada alternativa. Esto permite al investigador observar un proceso que se asemeja más a las condiciones de la vida real.

No obstante, los modelos de aprendizaje por reforzamiento asumen que las personas aprenden la satisfacción global de cada opción, es decir, no son los atributos específicos los que generan esa satisfacción final. En la práctica, esto impide obtener parámetros de preferencia asociados a atributos individuales, lo que limita las aplicaciones de modelos de aprendizaje a situaciones reales con impacto en políticas públicas.

Este proyecto de tesis busca responder preguntas fundamentales para extender el uso de experimentos de aprendizaje hacia aplicaciones relevantes para la política pública, como por ejemplo la inferencia de la disposición a pagar. El primer objetivo es determinar si las preferencias inferidas a partir de un DCE están correlacionadas, y en qué magnitud, con las preferencias que emergen en un experimento de aprendizaje. El segundo objetivo es identificar la relación entre las preferencias por atributos y la tasa de aprendizaje con la que las personas actualizan información durante dicho proceso. Responder ambas preguntas será una contribución relevante para campos como econometría, neuroeconomía y elecciones discretas.

Es deseable que el o la estudiante (de magíster o doctorado) tenga conocimientos previos de estadística y probabilidades y nociones de modelación de elección discreta y habilidades de programación en R, Python o lenguajes similares. Cualquier consulta o solicitud de información puede dirigirse a los profesores Omar D. Pérez (omar.perez.r@uchile.cl) y Bastián Henríquez-Jara (bastian.henriquez@uchile.cl).

**Omar D. Pérez**, Departamento de Ingeniería Industrial

**Bastián Henríquez-Jara**, Departamento de Ingeniería Civil

## Referencias

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., Hertwig, R., Stewart, T., West, R., & Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making*, 23(1), 15–47.

Henríquez-Jara, B., Guevara, C. A., Munizaga, M., & Perez, O. D. (2025). Habits and the subexploration of better transportation options: A dual-system approach. *Travel Behaviour and Society*, 38, 100877.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (ed.) *Frontiers in Econometrics*, 105-142.