

AN ANALYSIS OF DECISION UNDER RISK IN RATS

CHRISTINE M. CONSTANTINOPLE ALEX T. PIET
CARLOS D. BRODY

2019

En la teoría del prospecto el valor se determina por las funciones de probabilidad de la utilidad (la satisfacción o utilidad que las recompensas proveen) y la distorsión de probabilidad (se enfatizan las probabilidades bajas, y se minimizan las altas). Una alternativa son las teorías de aprendizaje, como el aprendizaje por refuerzo. El aprendizaje por refuerzo provee un marco de referencia para la toma de decisiones basada en valor, en el cual los estimados de valor se aprenden por la experiencia y se aprenden ensayo por ensayo con base en errores de predicción.

Los algoritmos de aprendizaje por refuerzo indican que los agentes aprenden el valor esperado (volumen \times probabilidad) de acciones o resultados mediante la experiencia, por lo que mostrarán funciones de probabilidad y utilidad lineales, lo que es incompatible con la teoría del prospecto. Se encontró que las ratas muestran señales de teoría de prospecto y de aprendizaje por refuerzo y se intenta integrar ambos.

En humanos se suelen estudiar “decisiones desde descripción”, mientras que en roedores se estudian “decisiones desde experiencia”, que son difíciles de reconciliar con la teoría de prospecto. Aquí se diseñó una tarea en la cual la probabilidad y cantidad de recompensa se comunican mediante evidencia sensorial, evocando decisiones desde descripción en lugar de experiencia. Esto posibilita aproximaciones de economía conductual.

Las ratas iniciaban un ensayo con un *nosepoke* en una pared de tres “puertos”. Se presentaban luces parpadeantes en los puertos laterales, y la tasa de parpadeo comunicaba la probabilidad de agua en cada puerto. A la vez, clicks audibles se presentaban desde bocinas laterales, y la tasa de clicks comunicaba el volumen de agua que había en cada puerto. Un puerto ofrecía una recompensa garantizada, y el otro una riesgosa con una probabilidad explícitamente señalada. Los puertos seguro y riesgoso variaban aleatoriamente. Uno de cuatro volúmenes de agua (6 a 48 μ L) podía ser la recompensa segura o riesgosa; y las probabilidades de la opción riesgosa iban de 0 a 1.

Las ratas mostraron aprender el significado de las claves al rechazar frecuentemente ensayos con recompensas pequeñas, abandonando el *nosepoke* a pesar de implicar un tiempo fuera. Quizá el rechazo, que persistió a pesar de tiempos fuera largos, reflejara estrategias de maximización. Las ratas favorecían prospectos con mayor valor esperado.

Los resultados indican que las ratas, como los humanos, muestran sensibilidad marginal disminuida: discriminar 24 de 48 μL es más difícil que discriminar 0 de 24.

Las funciones de probabilidad de las ratas mostraron que sobre ponderaban las probabilidades.

Parea evaluar las actitudes ante el riesgo de las ratas se utilizaron *certainty equivalents* (CE) para las apuestas de 48 μL . El CE es la recompensa garantizada que la rata evalúa como equivalente a la apuesta (como un procedimiento de ajuste). Si el CE es menor que el valor esperado de la apuesta, se intuye aversión al riesgo: el sujeto infravalora la apuesta y acepta una recompensa menor para evitar el riesgo; y viceversa. Los CE medidos se ajustaron cercanamente a los predichos.

Aunque las ratas mostraron utilidad y ponderación de probabilidad no lineales, consistentes con teoría de prospecto, también mostraron aprendizaje ensayo por ensayo, consistente con aprendizaje por refuerzo. Ajustar el modelo a ensayos posteriores a elecciones recompensadas y no recompensadas mostró cambios sistemáticos en las funciones de utilidad y ponderación de probabilidad: las funciones de utilidad se volvían menos cóncavas (es decir, con menor disminución en la sensibilidad marginal), y las funciones de ponderación de probabilidad se elevaban para reflejar la probabilidad incrementada de elecciones riesgosas después de una recompensa.

Otra característica de la conducta humana es la dependencia del referente: las personas evalúan las recompensas como ganancias o pérdidas con base en un punto de referencia interno. No está claro lo que determina al punto, pero se ha propuesto a la riqueza ya tenida, la expectativa de recompensa, heurísticos o experiencia reciente.

Las ratas mostraron dependencia del referente al tratar las recompensas pequeñas como pérdidas. Mostraron sesgos de *win-stay lose-switch*: después de ensayos no recompensados tenían mayor probabilidad de cambiar de puerto; y mostraron sesgo para cambiar después de recibir 6 o 12 μL , consistente con tratar esos resultados como pérdidas. El umbral para “ganar o perder” (*i.e.*, el punto de referencia) era dependiente de la experiencia: un grupo de ratas entrenado con recompensas dobles (12 - 96 μL) trataba a 12 y 24 μL como pérdidas.

Dado que el umbral de ganancia/pérdida dependía de la historia, se parametrizó un punto de referencia r como uno de dos valores, dependiendo de si el ensayo anterior fue recompensado. Las recompensas menores que r eran pérdidas.

Una comparación de modelos (*Akaike information criterion*) favoreció al modelo del punto de referencia. Incluir más ensayos en el punto de referencia mediante decaimiento exponencial no incrementó significativamente el ajuste del modelo.

La teoría del prospecto no explica cómo los agentes aprenden valores subjetivos de la experiencia, y aquí se incorporaron explícitamente parámetros de historia para dar cuenta del aprendizaje ensayo por ensayo.

El aprendizaje por refuerzo describe un proceso por el que los animales aprenden el valor de estados y acciones, pero implica una utilidad y ponderación de probabilidad lineales.

DISCUSSION

Como los humanos, las ratas muestran una utilidad no lineal cóncava para ganancias, distorsión de probabilidad, dependencia de referente, y, frecuentemente, aversión a las

pérdidas.

La teoría del prospecto, teoría de aprendizaje animal, y el aprendizaje por refuerzo, son marcos de referencia complementarios para estudiar toma de decisiones. El aprendizaje por refuerzo y el aprendizaje animal se centran en cómo los sujetos aprenden valores de la experiencia y usan esos valores aprendidos para tomar decisiones. La teoría del prospecto, en contraste, no se centra en aprendizaje sino que describe distorsiones no lineales que dan cuenta de la decisión. Se propone una aproximación simple para integrar los marcos de referencia en la cual los animales aprenden los valores de las acciones asociadas con estados, pero el error de predicción que dirige al aprendizaje está en unidades de valor subjetivo de acuerdo con la teoría del prospecto.