



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

FACULTAD DE PSICOLOGÍA

NIVELES COGNITIVOS Y CREENCIAS EN JUEGOS

T E S I S

QUE PARA OBTENER EL GRADO DE:

LICENCIATURA EN PSICOLOGÍA

PRESENTA:

JAIME OSVALDO ISLAS FARIAS

DIRECTOR DE TESIS:

DR. ARTURO BOUZAS RIAÑO

REVISOR:

DR. ÓSCAR ZAMORA ARÉVALO

SINODALES:

MTRO. MIGUEL HERRERA ORTÍZ

LIC. LIGIA COLMENARES

LIC. JOSÉ LUIS BAROJA MANZANO

Con el apoyo de:

Proyecto PAPIIT IN307214



CIUDAD DE MÉXICO

5 DE MAYO DE 2018

Niveles Cognitivos y Creencias en Juegos

por

Jaime Osvaldo Islas Farias

Tesis presentada para obtener la

Licenciatura en Psicología

en la

Facultad de Psicología

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad de México, 5 de mayo de 2018

Resumen

Niveles Cognitivos y Creencias en Juegos

por Jaime Osvaldo Islas Farias

En Teoría de Juegos, se entiende como razonamiento iterado al proceso mediante el cual los jugadores eligen sus respuestas incorporando las creencias que tienen sobre los otros jugadores. Este fenómeno ha sido estudiado con el juego p-Beauty contest, donde se ha encontrado que las elecciones de los participantes tienden al equilibrio tras jugar repetidamente. Los modelos propuestos para dar cuenta de este fenómeno asumen que los número elegidos son reflejo del número de pasos de razonamiento iterado que los participantes son capaces de realizar (nivel cognitivo), así como de las creencias que tienen sobre el nivel de los demás. En el presente trabajo se evaluó la consistencia en las elecciones y las creencias entre participantes con y sin experiencia. Para ello se realizaron dos subjuegos compuestos por cuatro repeticiones de p-beauty contest, donde un mismo jugador se enfrentaba a oponentes distintos, esperando encontrar el efecto de reset reportado en la literatura al enfrentarse a jugadores nuevos en el segundo subjuego. Todos los participantes debían señalar, además del número elegido, los números que creían elegirían sus contrincantes, encontrando que aún presentando el efecto de reset, la consistencia entre las creencias y las elecciones de los participantes se mantiene en los jugadores con experiencia.

Abstract

Cognitive Levels and Beliefs in games

by Jaime Osvaldo Islas Farias

Iterated reasoning, as defined within Game Theory, refers to the process in which players take into account the beliefs they have about other players in order to choose their own responses. This phenomenon has been studied with the game P-beauty contest, where it has been found that participants' choices tend to the Nash equilibrium as they play repeatedly. The models that have been proposed to account for this phenomenon assume that participants' choices are an indicator of the number of iterated reasoning steps that they are capable of take (cognitive level) and the beliefs they have about their opponents' levels. In this thesis, the consistency between choices and beliefs is assessed across participants with and without experience. Two sets with four repetitions of a p-beauty game were conducted, so that a single participant had to play against a different pair of opponents per set, in hopes of getting the so-called reset effect reported in the literature when faced against new players in the second set. Every participant had to indicate both the number they chose and the numbers they predicted the other participants would choose. Evidence obtained in this study shows that, even after presenting the reset effect, consistency gained between beliefs and choices in the first set of games, remained across the second.

Agradecimientos

A Jaime Islas Curiel, mi papá, por siempre apoyar mis decisiones, incluso si estas no siempre son las más acertadas. Es gracias al esfuerzo con el has salido a trabajar todos los días durante todos estos años que he podido llegar hasta este punto en mi vida académica, profesional y personal. Gracias por enseñarme (con el ejemplo, por si fuera poco) a ser una persona con principios y valores.

Póstumamente a Beatriz Farias Orozco, mi mamá, sin quien, literalmente, yo no estaría aquí. Es también gracias a tu ejemplo (y tus regaños) que me he convertido en el hombre que soy hoy. Quiero pensar que estarías muy feliz por este logro, y orgullosa de mí.

A Adriana Felisa Chávez De la Peña, mi princesa, el amor de mi vida. Gracias a tu ayuda y apoyo (académico y emocional) esta tesis es una realidad. El amor que te tengo y la expectativa de hacer mi vida contigo me hacen mejorar continuamente como persona. Gracias por ser mi mejor amiga, compañera y el amor de mi vida.

Si los amigos son la familia que uno escoge, entonces la familia son los amigos que no escogemos, pero que están ahí incondicionalmente y durante toda nuestra vida. Gracias a toda mi familia por su cariño y apoyo permanente; a mi abuelita y a mis tíos por ser como mis segundos papás, y a mis primos por ser como mis hermanos.

A mis amigos, a quienes no enlisto, porque ellos saben quiénes son (y así evito momentos incómodos si me olvido de mencionar a alguno). La vida está llena de giros inesperados y nunca sabemos dónde terminaremos el día de mañana. Pero aunque no siempre somos una presencia constante en nuestras vidas, y coincidimos solo de forma intermitente, sabemos que nuestra amistad se mantiene. Gracias por todos los momentos divertidos y entrañables.

A todos los masoquistas del Lab 25, que pudiendo hacer cosas más sencillas, siempre eligen ir más lejos y aceptar los mayores desafíos. Gracias por poner la vara tan alta en la calidad de todos sus proyectos; estar rodeado de estudiantes

tan brillantes y trabajadores inevitablemente me impulsó a esforzarme más para intentar alcanzar los estándares establecidos por todos ustedes. Espero que la consideren esta tesis una entrada digna al catálogo de investigaciones realizadas en el Lab 25.

A todos los profesores que participaron en mi formación; para bien o para mal, esta tesis y lo que venga después son un producto parcial de sus enseñanzas. Particularmente agradezco a los miembros de mi comité: A Ligia Colmenares, por sus excelentes comentarios y observaciones al borrador de esta tesis. A Miguel Herrera, porque fueron sus clases en los primeros semestres de la carrera lo que me hizo acercarme al área de la psicología experimental, y en los últimos semestres a teoría de juegos. A José Luis Baroja, con quien he tratado como profesor así como compañero en el Lab 25, y en ambos casos he aprendido muchísimo; considero que José Luis es el modelo ejemplar de cómo debe ser un investigador. A Oscar Zamora, cuyas excelentes clases fueron también un pilar en mi formación y contribuyeron a mi amor por el área. Y finalmente, de manera particular, al “Doc” Arturo Bouzas. Gracias por sus consejos siempre pertinentes y concretos, que permitieron volver una realidad esta tesis, desde que comenzó como solo una idea vaga y difusa. Admiro muchísimo su brillantez y sabiduría, producto de una vida de experiencia académica. Gracias a todos por compartirme su conocimiento y experiencia en mi formación como investigador. A mis compañeras de trabajo en el CEPE. Por abrirme las puertas al atemorizante mundo de los adultos de forma relativamente indolora, y enseñarme que algunas de las habilidades que nos enseñan en la escuela son, de hecho, utilizadas en la vida real. Son todas un ejemplo de responsabilidad, esfuerzo y pasión por el trabajo, y mi punto de referencia sobre cómo hay que desenvolverse en el mundo profesional.

A ti, misterioso lector. Que estés leyendo mi tesis justifica su redacción más allá de su función como requisito para titularme.

Índice general

| | |
|---|------------|
| Resumen | III |
| Agradecimientos | VII |
| 1. Introducción | 1 |
| 2. Marco Teórico | 7 |
| 2.1. Modelo de nivel-k | 7 |
| 2.2. Relación entre creencias y elecciones | 9 |
| 2.3. El efecto de la experiencia | 12 |
| 2.4. Objetivo | 13 |
| 3. Método | 15 |
| 3.1. Participantes | 15 |
| 3.2. Diseño Experimental | 15 |
| 3.3. Procedimiento | 16 |
| 4. Resultados | 19 |
| 4.1. Análisis Bayesiano | 19 |
| 4.2. Consistencia entre creencias y elecciones en el primer subjuego . | 19 |
| 4.3. Efecto de reset | 29 |
| 4.4. Consistencia entre creencias y elecciones entre subjuegos | 35 |
| 4.5. ¿Las creencias se vuelven más precisas con la experiencia? | 40 |
| 5. Discusión | 57 |
| 6. Conclusión | 61 |
| A. Instrucciones | 67 |

X

B. Formato de registro **71**

C. Análisis de datos con pruebas frecuentistas **73**

Índice de figuras

| | |
|--|----|
| 4.1. Exploración visual de la consistencia entre creencias y elecciones | 21 |
| 4.2. Diferencias Normalizadas entre creencias y elecciones en el Subjuego 1 (Factor de Bayes) | 23 |
| 4.3. Diferencias Normalizadas entre creencias y elecciones en el Subjuego 1, omitiendo la multiplicación por p (Factor de Bayes) | 25 |
| 4.4. Diferencias Relativas entre creencias y elecciones en el Subjuego 1 (Factor de Bayes) | 28 |
| 4.5. Diferencias Relativas entre creencias y elecciones en el Subjuego 1 sin la multiplicación por p (Factor de Bayes) | 30 |
| 4.6. Elecciones registradas en ensayos consecutivos | 32 |
| 4.7. Elecciones de todos los participantes | 34 |
| 4.8. Participante atípico: Comparando el desempeño del Participante A de la Sesión 3 | 43 |
| 4.9. Promedio de los números elegidos por los participantes A y no-A en cada uno de los periodos jugados | 43 |
| 4.10. Comparación entre las creencias y elecciones registradas en el Subjuego 2 | 44 |
| 4.11. Evaluación de las Diferencias Normalizadas entre creencias y elecciones en los participantes A en el Subjuego 2 (Factor de Bayes) | 45 |
| 4.12. Evaluación de las Diferencias Normalizadas entre creencias y elecciones en los participantes D y E en el Subjuego 2 (Factor de Bayes) | 46 |
| 4.13. Evaluación de las Diferencias Normalizadas entre creencias (sin multiplicar por p) y elecciones en los participantes A en el Subjuego 2 (Factor de Bayes) | 47 |
| 4.14. Evaluación de las Diferencias Normalizadas entre creencias y elecciones en los participantes D y E en el Subjuego 2 (Factor de Bayes) | 48 |

| | |
|---|----|
| 4.15. Evaluación de las Diferencias Relativas entre creencias y elecciones en los participantes A en el Subjuego 2 (Factor de Bayes) . . . | 49 |
| 4.16. Evaluación de las Diferencias Relativas entre creencias y elecciones en los participantes D y E en el Subjuego 2 (Factor de Bayes) . | 50 |
| 4.17. Evaluación de las Diferencias Relativas entre creencias (omitiendo la multiplicación por p) y elecciones en los participantes A en el Subjuego 2 (Factor de Bayes) | 51 |
| 4.18. Evaluación de las Diferencias Relativas entre creencias (omitiendo la multiplicación por p) y elecciones en los participantes D y E en el Subjuego 2 (Factor de Bayes) | 52 |
| 4.19. Precisión en la predicción de los números a elegir por los otros participantes (se incluye la multiplicación por p de las creencias y las elecciones observadas) | 53 |
| 4.20. Precisión en la predicción de los números a elegir por los otros participantes (se omite la multiplicación por p) | 54 |
| 4.21. Diferencias en la precisión de las creencias registradas por los participantes A y no A (Factor de Bayes) | 55 |

Índice de cuadros

| | |
|--|----|
| 4.1. Diferencias Normalizadas en el Subjuego 1 (prueba t de una muestra) | 22 |
| 4.2. Diferencias Normalizadas en el Subjuego 1, omitiendo la multiplicación por p (prueba t de una muestra) | 24 |
| 4.3. Diferencias Relativas en el Subjuego 1 (prueba t de una muestra) . | 27 |
| 4.4. Diferencias Relativas en el Subjuego 1, omitiendo la multiplicación por p (prueba t de una muestra) | 29 |
| 4.5. Prueba binomial bayesiana para evaluar la proporción de casos en que los participantes A aumentan y reducen su número elegido | 33 |
| 4.6. Diferencias Normalizadas en el Subjuego 2; Participante A (Pruebas t de una muestra) | 36 |
| 4.7. Diferencias Normalizadas en el Subjuego 2; Participante A (Pruebas t de una muestra) | 37 |
| 4.8. Diferencias Normalizadas en el Subjuego 2 omitiendo la multiplicación por p ; Participante A (Pruebas t de una muestra) | 38 |
| 4.9. Diferencias Normalizadas en el Subjuego 2, omitiendo la multiplicación por p ; Participantes D y E (Pruebas t de una muestra) . . . | 38 |
| 4.10. Diferencias Relativas en el Subjuego 2; Participantes A (Pruebas t de una muestra) | 39 |
| 4.11. Diferencias Relativas en el Subjuego 2; Participantes D y E (Pruebas t de una muestra) | 39 |
| 4.12. Diferencias Relativas en el Subjuego 2, omitiendo la multiplicación por p ; Participantes A (Pruebas t de una muestra) | 40 |
| 4.13. Diferencias Relativas en el Subjuego 2, omitiendo la multiplicación por p ; Participantes D y E (Pruebas t de una muestra) | 40 |
| 4.14. Comparación de la precisión de las predicciones hechas por los participantes A y no A en todos los periodos jugados | 42 |

| | |
|--|----|
| C.1. Prueba t de una muestra: Diferencias normalizadas en el Subjuego 1 | 73 |
| C.2. Prueba t de una muestra: Diferencias normalizadas en el Subjuego 1 sin la multiplicación por p | 73 |
| C.3. Prueba t de una muestra: Diferencias relativas en el Subjuego 1 . . | 74 |
| C.4. Prueba t de una muestra: Diferencias relativas en el Subjuego 1 sin la multiplicación por p | 74 |
| C.5. Prueba t de una muestra: Diferencias Normalizadas en el Subjue- go 2 (Participante A) | 74 |
| C.6. Prueba t de una muestra: Diferencias Normalizadas en el Subjue- go 2 (Participantes D y E) | 74 |
| C.7. Prueba t de una muestra: Diferencias Normalizadas en el Subjue- go 2 sin la multiplicación por p (Participante A) | 75 |
| C.8. Prueba t de una muestra: Diferencias Normalizadas en el Subjue- go 2 sin la multiplicación por p (Participantes D y E) | 75 |
| C.9. Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 (Participante A) | 75 |
| C.10. Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 (Participantes D y E) | 75 |
| C.11. Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 sin la multiplicación por p (Participante A) | 76 |
| C.12. Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 sin la multiplicación por p (Participantes D y E) | 76 |

Capítulo 1

Introducción

Teoría de Juegos es la rama de las matemáticas que estudia la toma de decisiones en situaciones de interacción. Dentro de teoría de juegos, se estudian situaciones como la que representa el juego del ciempiés (estudiada por primera vez por Rosenthal, en 1981). Se trata de un juego donde dos jugadores, A y B, toman turnos para elegir si quedarse con una ganancia acumulada o pasar la ganancia al otro jugador. Si algún jugador elige quedarse con la ganancia acumulada, el otro jugador no recibe nada y el juego termina. En cambio, si el jugador elige pasar, el tamaño de la ganancia acumulada aumenta. El juego dura 10 turnos, y la ganancia acumulada aumenta su valor en cada uno. Si en el último turno, que le corresponde al jugador B, este elige pasar, el jugador A se queda con todas las ganancias.

De acuerdo con las reglas del juego, si los jugadores quisieran maximizar sus ganancias, deberían elegir ‘pasar’ para que la ganancia acumulada siga aumentando, pero al mismo tiempo, dado que ambos jugadores quieren quedarse con la ganancia, deben elegir quedarse con ella antes de que lo haga el otro jugador. ¿En qué turno del juego sería una buena elección (para cada jugador) quedarse con la ganancia acumulada?

El juego establece que en el último turno, cuando la ganancia acumulada tiene su valor más alto, le corresponde elegir al jugador B. Si este jugador elige pasar, todas las ganancias serán para el jugador A, mientras que si decide quedarse con la ganancia acumulada, B obtendrá la mayor ganancia posible. Por lo tanto, el jugador B debería elegir quedarse con la ganancia en el turno 10. Sin embargo, este razonamiento no es un secreto para el jugador A. Si el jugador A sabe que al

jugador B le conviene quedarse con la ganancia en el turno 10 (en cuyo caso A se quedaría sin nada), entonces la mejor estrategia para el jugador A consiste en quedarse con la ganancia acumulada en el turno 9. Pero así como el jugador A es capaz de anticipar que el jugador B elegirá quedarse con la ganancia en el turno 10, el jugador B puede deducir que el jugador A elegirá quedarse con la ganancia en el turno 9. Si este es el caso, el jugador B debería elegir quedarse con la ganancia en el turno 8. Si ambos jugadores repiten este razonamiento para todos los turnos de juego, eventualmente se llega a la conclusión de que la mejor estrategia para cualquiera de los jugadores en cualquier turno es elegir quedarse con la ganancia, sin importar el turno en que se encuentren. A este proceso de repetir un razonamiento para llegar a la mejor estrategia disponible tomando en cuenta que el otro jugador también puede hacerlo, se conoce como razonamiento iterado, y es uno de los conceptos básicos en teoría de juegos.

A las situaciones donde los jugadores eligen la estrategia que maximice su ganancia esperada, tomando en cuenta las estrategias de los otros jugadores, se les conoce como el equilibrio de Nash (Nash, 1951), otro concepto básico en teoría de juegos. En el juego del ciempiés, el equilibrio de Nash supondría que el jugador A se quede con la ganancia acumulada desde el primer turno del juego, incluso si en este punto su valor es muy pequeño.

En muchas situaciones de interacción, el razonamiento requerido para llegar al equilibrio es demasiado complejo para ser plausible en términos conductuales, por lo que cuando personas reales se encuentran en dicho tipo de situaciones, sus elecciones suelen ser distintas de lo que supondría el equilibrio. Sin embargo, existe evidencia de que el aprendizaje (jugar un juego de forma repetida) genera una tendencia a converger al equilibrio y solo en situaciones donde no se permite que haya aprendizaje, se puede atribuir el equilibrio al pensamiento estratégico por sí sólo (Crawford, Costa-Gomes Iriberry, 2013).

El equilibrio se ve reflejado tanto en las elecciones de los participantes del juego, como en sus creencias sobre los otros participantes: jugadores que son racionales (en términos de teoría de decisión) tienen creencias correctas sobre los otros jugadores, si estos también son racionales (Crawford, Costa-Gomes Iri-

berri, 2013).

Durante el proceso de razonamiento iterado, los jugadores incorporan las creencias que tienen sobre la conducta de los otros jugadores en su toma de decisiones. Keynes (1956) ilustró el proceso de razonamiento iterado con una analogía que se conoce como Beauty contest: Un concurso en el que los participantes deben elegir de entre cien fotografías de rostros, cuáles piensan que los demás participantes considerarán que son los más atractivos. Tomando en cuenta que todos los participantes se enfrentan al mismo problema, para ganar no basta con elegir solamente aquellos rostros que piensen que son los más atractivos, o cuáles piensan que los demás participantes piensan que son más atractivos, sino aquellos que piensen que los demás participantes pensarán que los demás participantes piensan que son los más atractivos. Esto implica tres pasos de razonamiento iterado.

Un agente totalmente racional debería realizar tantos pasos de razonamiento iterado como fueran necesarios para llegar a la solución por dominancia del juego (el equilibrio de Nash). En la realidad, las personas no se comportan de forma perfectamente racional, y la cantidad de pasos de razonamiento iterado que realizan es limitada (Stahl Wilson, 1995, Ho, Caremer Weigelt, 1998).

Experimentalmente el juego p-Beauty contest, llamado así a partir de la analogía de Keynes, ha sido utilizado para estudiar el razonamiento iterado. En este juego, en el que participan varios participantes (por lo menos 3), cada uno debe elegir un número entero dentro el rango $[0 - 100]$, de manera simultánea y sin revelarlo a los otros jugadores. Posteriormente, se calcula la media de todos los números elegidos y este valor se multiplica por un parámetro p que es un número positivo y diferente de 1, conocido de antemano por todos los jugadores, (generalmente se utiliza $p = \frac{2}{3}$). Al valor resultante de este cálculo se le llama el número objetivo, y el ganador del juego será el participante que haya elegido el número más cercano a este número.

Si los jugadores fueran perfectamente racionales, creyeran que los demás jugadores también lo son y además pudieran realizar una cantidad infinita de pa-

sos de razonamiento iterado, llegarían a la solución por dominancia del juego mediante el siguiente razonamiento: El valor más grande que puede alcanzar el número objetivo (con $p = \frac{2}{3}$) es $100 \cdot p$; (donde $p = ,666$), así que cualquier número arriba de este valor es dominado por 66,66. Jugadores racionales obedecerán la dominancia y creerán que los demás jugadores lo harán también, por lo que todos elegirán un número menor a 66,66. Por lo tanto, el valor máximo del número objetivo será $100 \cdot p^2$; (siendo $p^2 = 44,44$) y elegir cualquier número por arriba de este será una estrategia dominada. Aplicando este razonamiento una y otra vez (de ahí el nombre de razonamiento iterado), el valor del número objetivo se reduce con cada iteración, hasta que se llega al equilibrio de Nash del juego, que todos los jugadores elijan 0 (Nagel, 1995; Ho, Camerer, Weigelt, 1998).

Empíricamente, esto no suele ocurrir, pero cuando el mismo grupo de participantes juegan repetidamente (más de un periodo), se ha reportado consistentemente que sus elecciones se acercan paulatinamente al equilibrio (elegir 0) con cada periodo (Nagel, 1995, Ho, Camerer Weigelt, 1998). También se ha observado que dicha tendencia se interrumpe cuando se agregan nuevos participantes al juego, siendo que los jugadores con experiencia incrementan el número elegido al enfrentarse a estos jugadores novatos, lo que se conoce como efecto de reset (Slonim, 2005).

Se han propuesto varios modelos que dan cuenta de la forma en la que las personas eligen sus números en el juego. Estos modelos capturan la noción de que el número elegido por las personas es un reflejo del número de pasos de razonamiento iterado que son capaces de realizar, que en la literatura de juegos se conoce como “nivel cognitivo”, así como de las creencias o expectativas que tienen sobre el nivel cognitivo de los demás jugadores (Crawford, Costa-Gomes Iriberry, 2013).

Algunos estudios (Agranov et al., 2012 y Slonim, 2005) han explorado el efecto que tienen las creencias sobre el desempeño de los otros jugadores en las elecciones de cada participante, evaluando esta relación de forma indirecta y encontrando evidencia a favor de una relación positiva. En contraste, en otros estudios que han intentado tener un acercamiento más directo, se han encon-

trado inconsistencias entre las creencias sobre lo que harán los otros jugadores y las elecciones reales; Lahav (2015) utilizó un método en el que le pidió a los jugadores hacer una estimación aproximada de las elecciones de los otros jugadores y comparó directamente estas creencias con sus elecciones reales, encontrando una falta de consistencia entre las dos.

El presente trabajo de investigación pretende estudiar de manera directa la relación entre las elecciones de las personas y sus creencias sobre las elecciones de los demás, y adicionalmente explora si la experiencia que se adquiere en juegos repetidos influye en la relación de estas dos variables.

El diseño experimental consiste en juegos repetidos de p-beauty contest (organizando cada sesión experimental en 2 subjuegos compuestos por 4 repeticiones cada uno). En cada sesión experimental, sólo un jugador participa en el juego durante los dos subjuegos, siendo que al término del Subjuego 1 los demás jugadores son reemplazados por nuevos participantes que no han jugado previamente. Además de registrar los números elegidos por cada participante en cada periodo, se les solicita a los jugadores que reporten directamente los números que creen que los demás elegirán. Este método de explicitar creencias permite comparar directamente las elecciones y las creencias de los jugadores, de una forma más evidente a la utilizada por Lahav (2015), si bien igual que en ese estudio, acercarse a los números que eligen los otros jugadores es recompensado con ganancias en el juego para agregar motivación a la expresión de las creencias.

Se encontró que la entrada de nuevos jugadores en el segundo subjuego incrementa el número elegido de los jugadores con experiencia en el primer periodo de este nuevo subjuego, pero la diferencia entre sus creencias y elecciones no incrementa junto con esta, lo que aporta evidencia de que la adquisición de experiencia incrementa la consistencia.

El resto de la tesis está dividida en cinco apartados: En el primero se presenta el marco teórico, que describe el modelo de nivel-k usado para explicar la conducta de las personas en el juego y revisa la relación empírica entre las eleccio-

nes de las personas y sus creencias, y la forma en que la experiencia en juegos repetidos influye en las elecciones de los jugadores. En esta misma sección se detallan los objetivos concretos del trabajo de investigación y las estrategias para alcanzarlos. En el segundo apartado se describe el método utilizado, incluyendo información sobre los participantes, el procedimiento y el diseño experimental. En el tercer apartado se presentan los resultados del experimento. Se reporta el grado de consistencia que existe entre creencias y elecciones de los jugadores en el primer subjuego, el efecto de introducir a participantes sin experiencia en el segundo subjuego, y las diferencias en consistencia entre creencias y elecciones que hay entre los dos subjuegos. Por último, en el cuarto apartado se elabora la discusión a partir de los resultados. Las conclusiones se presentan en el quinto apartado.

Capítulo 2

Marco Teórico

2.1. Modelo de nivel-k

Este modelo fue propuesto por Nagel (1995), para dar cuenta de la conducta de las personas en juegos con solución por dominancia, como ocurre con textitp-beauty contest. El modelo define niveles cognitivos que describen el número de pasos de razonamiento iterado que realiza una persona en el juego y difiere respecto de los modelos de equilibrio clásicos, en que las creencias que se tienen sobre las elecciones de los otros jugadores no están basadas en la definición del equilibrio (Nash, 1951, Crawford, Costa-Gomes Iriberry, 2013), es decir, admite la posibilidad de que los otros jugadores no son perfectamente racionales.

Según este modelo, los jugadores con un nivel cognitivo 0 serían aquellos que no realizan ningún paso de razonamiento iterado, es decir, que no toman en consideración que las elecciones de los otros participantes influyen en el cálculo del número objetivo. Estos jugadores eligen un número con base en alguna regla arbitraria, (por ejemplo, su número de la suerte o favorito), por lo que podrían elegir cualquier número dentro del rango establecido con una probabilidad similar.

Un jugador de nivel 1 es aquel que sí considera que las elecciones de los otros jugadores influyen en el cálculo del número objetivo, pero supone que los otros jugadores no han tomado esto en consideración. El jugador de nivel 1 asume que los demás jugadores son de nivel 0 y elige el número que es la respuesta óptima contra este tipo de jugadores, asumiendo que la media de los números

elegidos por todos los jugadores estará cerca de 50 (el mejor predictor de la media de un conjunto de números aleatorios en el rango $[0 - 100]$) y multiplicará este número por p para acercarse lo más posible al número objetivo.

Por su parte, un jugador de nivel 2 no sólo considera que las elecciones de otros jugadores influyen en el número objetivo, sino que también asume que los otros jugadores saben esto. El jugador de nivel 2 elegirá el número que es la respuesta óptima contra una población de oponentes de nivel 1, y como éstos eligen números cercanos a $50 \cdot p$, el jugador de nivel 2 debe multiplicar por p nuevamente para acercarse al que piensa que será el número objetivo, esto es $50 \cdot p^2$.

En general, un jugador de nivel k elegirá la respuesta óptima contra una población de jugadores de nivel $k - 1$, esto es $50 \cdot p^k$. Con base en esta regla, el modelo computa el nivel cognitivo de los jugadores en función de a cuál de los intervalos de elección establecidos por el modelo pertenece su número elegido.

Una variación más sofisticada del modelo de nivel- k es el modelo de jerarquía cognitiva, propuesto por Camerer, Ho, y Chong, (2004). El modelo propone que un jugador de nivel k no sólo da la mejor respuesta contra jugadores $k - 1$, sino a una combinación de todos los tipos de jugador desde el nivel 0 hasta $k - 1$, a partir de una distribución Poisson que se actualiza de forma bayesiana. El modelo mantiene el supuesto de que las personas consideran que su nivel está por arriba de los demás jugadores, y predice un aprendizaje sobre los niveles k de los otros jugadores.

En general, de acuerdo con la evidencia obtenida experimentalmente, los modelos de nivel- k explican mejor el comportamiento de las personas en juegos con solución por razonamiento iterado que los modelos de equilibrio (Crawford, Costa-Gomes Iriberry, 2013).

En los dos modelos previamente descritos, la elección de los jugadores depende de tres elementos: 1) sus creencias sobre cómo juegan los participantes de nivel 0; 2) sus expectativas sobre el nivel cognitivo de los oponentes; y 3) el número de pasos de razonamiento que son capaces de hacer en el juego (Agra-

nov et al., 2012).

En la siguiente sección se ahonda sobre el segundo elemento: las expectativas (textiti. e. creencias sobre el nivel cognitivo de sus oponentes, y la evidencia que se ha encontrado sobre su relación con la elección.

2.2. Relación entre creencias y elecciones

Para aportar evidencia empírica de la influencia de las creencias acerca de la sofisticación de los otros jugadores sobre las elecciones de los jugadores en p-beauty contest, Agranov et al. (2012), manipularon las creencias que los participantes tenían sobre sus oponentes, informando a cada participante que jugaría contra siete estudiantes graduados de Economía con conocimiento sobre este tipo de juegos, o bien, contra siete computadoras programadas para seleccionar aleatoriamente, con la misma probabilidad, números dentro del rango $[0 - 100]$. En este estudio se encontró que los números registrados por los participantes corresponden con un nivel cognitivo significativamente mayor cuando se les decía que se enfrentarían a estudiantes graduados que cuando creían competir contra las computadoras. Este resultado parece sugerir que el nivel cognitivo que muestran las personas en juegos de p-beauty contest depende no únicamente de su sofisticación cognitiva, sino también de sus creencias sobre la sofisticación de los otros jugadores.

Para estudiar de forma más directa la relación entre las creencias y las elecciones de los participantes, Lahav (2015) utilizó un método con *elicited beliefs* (la mejor forma de traducirlo sería “*que hace explícitas las creencias*”) en sesiones experimentales compuestas por 5 periodos de p-beauty contest con hasta 20 participantes. En cada periodo, además de elegir su propio número, se les pidió a los participantes que estimaran cuántos de los otros participantes elegirían un número dentro de cada uno de 10 posibles intervalos en el rango $0 - 100$ ($0 - 10$, $11 - 20$, $21 - 30$, \dots , $91 - 100$) y con dichas estimaciones, se calcularon las creencias de los participantes sobre el número promedio en cada periodo del juego. Al implementar este método, en contraste con investigaciones previas,

Lahav concluye que las elecciones no son un reflejo preciso de las creencias de los participantes, pues encuentra diferencias significativas entre el número objetivo computado de acuerdo a las creencias de los participantes acerca de las tiradas de los demás jugadores y el número que de hecho eligen en el juego.

Otra investigación con resultados similares fue realizada por Costa-Gomes y Weizsäcker en el 2008, utilizando un método para explicitar creencias en juegos sencillos de 3x3 (donde dos jugadores tienen que elegir entre tres estrategias posibles), en los que es posible llegar al equilibrio del juego mediante razonamiento iterado. En dicha investigación se encontró que en la mayoría de los casos, de acuerdo con el modelo de nivel-k, las creencias registradas por los jugadores acerca de sus oponentes los situaban en el nivel 2, mientras que sus elecciones correspondían al nivel 1. Ya que no parece que las elecciones observadas habrían sido una respuesta óptima ante las creencias explicitadas que se registraron, este resultado pone en duda que estas últimas sean la base a partir de la cual los participantes emiten sus respuestas. A la luz de estos hallazgos, los autores concluyen que los jugadores basan sus decisiones en reglas de elección que influyen en ambas, creencias y decisiones.

Debido a que el presente trabajo de tesis incorpora parte del método de Lahav (2015) para recopilar las creencias de los jugadores acerca de las tiradas de sus oponentes en juegos de p-beauty contest, se enfatizan los siguientes puntos respecto del estudio citado:

1. El método propuesto no permite conocer con exactitud la creencia de los participantes sobre el número objetivo. En su lugar, se obtuvo un cálculo aproximado de esta a partir del número de jugadores que se creyó elegirían un número dentro de cada intervalo, tomando la media de cada intervalo como valor de referencia.
2. Con alrededor de 20 personas participando en el juego, parece inverosímil, dada la demanda cognitiva, que los jugadores puedan calcular con precisión el número objetivo derivado de sus creencias acerca de las elecciones

del resto de los participantes para emitir su respuesta.

3. Se contó con dos grupos control. En uno, no se realizó la explicitación de creencias, y en el otro, dicho procedimiento se realizó después de que los participantes hubieran elegido su número. Comparado la elección de los participantes en los grupos control con el grupo experimental, los resultados indican que solicitar a los participantes que registraran sus creencias no cambia significativamente su número elegido.
4. El último periodo del juego fue el único en el que no se encontraron diferencias significativas entre creencias y elecciones, lo que podría sugerir que dicha discrepancia disminuye con la experiencia.
5. Dado que se sabe que los participantes tienden a elegir números cada vez más pequeños en cada periodo y que esto reduce invariablemente la magnitud de cualquier diferencia entre elecciones y creencias (ya que las diferencias entre números más pequeños son, por definición, más pequeñas), Lahav implementó un método de normalización con el que ponderó las diferencias entre las creencias y elecciones de cada participante en cada periodo por el promedio de los números elegidos por todos los participantes en dicho periodo. Sin embargo, como la medida de normalización depende de la tirada de todos los jugadores, la magnitud de la diferencia normalizada es afectada por el nivel cognitivo promedio, y “castiga” (incrementa la magnitud) las diferencias entre creencias y elecciones cuando estas no son tan sofisticadas como las del promedio. Debido a esto, podría no ser la mejor forma de compensar la tendencia al equilibrio.

El resultado mencionado en el punto 4 permite cuestionar si la discrepancia entre creencias y elecciones se ve afectada por la experiencia que tienen los participantes en el juego. En la siguiente sección se revisa el efecto de la experiencia en juegos repetidos de p-beauty contest.

2.3. El efecto de la experiencia

Para estudiar el efecto de la experiencia en juegos repetidos de p-beauty contest, Slonim (2005), realizó sesiones experimentales compuestas de 12 periodos (repeticiones del juego), distribuidos equitativamente en tres Subjuegos con tres jugadores en cada uno.

En una primera condición, al terminar cada subjuego se reemplazaba a dos de los tres participantes por jugadores nuevos, siendo que sólo un jugador permanecía en el experimento durante los 12 periodos completos. En una segunda condición, los dos participantes retirados al término de cada subjuego eran sustituidos por jugadores con la misma experiencia que el jugador que se mantenía en el juego (es decir, que habían jugado la misma cantidad de periodos). En ambas condiciones, los participantes tenían información acerca del nivel de experiencia de los demás jugadores (el número de periodos jugados).

Slonim (2005), reportó que en el primer periodo de los Subjuegos 2 y 3, los jugadores con más experiencia mostraron un mayor nivel cognitivo (es decir, que eligieron números más cercanos a 0) cuando sabían que jugaban contra oponentes con la misma experiencia que ellos, que cuando jugaban contra oponentes que no habían jugado previamente. Este resultado aporta evidencia a favor de que las creencias sobre el nivel cognitivo de los otros jugadores influyen en las elecciones individuales. Por su parte, los jugadores sin experiencia no mostraron diferencias significativas en sus elecciones cuando jugaron con oponentes experimentados o no experimentados.

En cuanto al efecto de la experiencia en el desempeño de los jugadores, también se observó que los jugadores experimentados ganan el juego con mayor frecuencia cuando juegan con jugadores sin experiencia. Sin embargo, esta ventaja parece reducirse periodo a periodo, conforme los otros jugadores adquieren experiencia.

Un último resultado reportado por Slonim (2005), y probablemente el de mayor relevancia para efectos de la presente tesis, corresponde a un efecto de reset

en la tendencia a elegir un número cada vez más pequeño entre cada periodo, cuando se inicia un nuevo Subjuego. En otras palabras, los jugadores con experiencia presentan una reversión en la tendencia a elegir números cada vez más cercanos al equilibrio cuando entran nuevos jugadores.

Con base en estos hallazgos, y los reportados en las secciones anteriores, se procede a plantear formalmente el objetivo de este trabajo de investigación, así como las estrategias metodológicas empleadas para llevarlo a cabo.

2.4. Objetivo

El presente trabajo de investigación busca evaluar si la reducción en las diferencias entre creencias y elecciones reportada por Lahav (2015), es verdaderamente el resultado de la experiencia adquirida al participar repetidas veces en un juego.

Para responder a esta pregunta, se propone una versión modificada del método de Lahav (2015) que contempla la participación de un grupo más pequeño de jugadores, lo que facilita preguntar a los participantes directamente por las creencias que tienen sobre los números que elegirán los demás participantes, permitiendo una estimación más precisa del número objetivo que corresponda a las creencias registradas, además la demanda cognitiva para los jugadores. De esta forma, resulta más verosímil esperar una correspondencia directa entre las elecciones de los participantes y el cómputo del número objetivo de acuerdo a sus creencias.

Para atenuar el peso que tiene la tendencia a elegir números cada vez más pequeños entre cada periodo, se buscó replicar el efecto de Reset reportado por Slonim (2005) en las elecciones de los jugadores con experiencia. Para ello, algunos participantes permanecieron durante más de un subjuego y se substituyó al resto por jugadores nuevos en un juego p-beauty contest repetido. Se espera que el efecto de reset opere no sólo en las elecciones, sino también en las creencias del jugador experimentado, lo que permitiría evaluar si la diferencia entre estas

sigue reduciéndose, como sugeriría lo reportado por Lahav (2015), aún cuando se abandone la tendencia a reducir las elecciones en el último periodo del primer Subjuego jugado. Si a pesar de observarse el efecto de Reset, la diferencia entre sus elecciones y sus creencias disminuye, esto aportaría evidencia a favor de que es la experiencia lo que reduce las inconsistencias entre creencias y elecciones.

En conjunto, el diseño experimental aquí propuesto permite determinar si las creencias de los jugadores que participan en los dos subjuegos se acercan más a las elecciones reales de los otros jugadores en el segundo subjuego gracias a que cuentan con mayor experiencia.

Capítulo 3

Método

3.1. Participantes

Se contó con la participación de 50 estudiantes de los primeros semestres de la carrera de Psicología en la Facultad de Psicología de la Universidad Nacional Autónoma de México. A cambio de su participación, todos los estudiantes recibieron medio punto extra sobre la calificación de uno de sus exámenes, y los participantes que obtuvieron mayores ganancias al final de cada Subjuego recibieron otro medio punto extra. Su participación en el experimento fue voluntaria, podían abandonarlo si lo deseaban, y se mantuvo el anonimato entre los jugadores durante el experimento y en el análisis de datos, asignando una clave de identificación única a cada participante.

3.2. Diseño Experimental

Se realizaron 10 sesiones experimentales con 5 participantes diferentes en cada una. El experimento se llevó a cabo en un aula sin distracciones externas. Al inicio de cada sesión, se asignó aleatoriamente un rol a cada participante (identificado con las letras A, B, C, D y E).

Las sesiones experimentales estuvieron compuestas por ocho periodos de p-beauty contest, repartidos en dos Subjuegos de cuatro periodos. En cada Subjuego participaron solo tres personas: En el Subjuego 1, los participantes A, B y C jugaron por cuatro periodos, mientras los participantes D y E esperaban en un

aula diferente sin poder hablar entre ellos. Terminado el Subjuego 1, los participantes B y C se retiraban y eran reemplazados para el Subjuego 2 por los participantes D y E. En el Subjuego 2, los participantes A, D y E jugaron por cuatro periodos, al final de los cuales se dio por terminado el experimento.

Además de elegir su número en el juego, se solicitó a cada participante que en cada periodo escribieran dos números que creyeran que estarían lo más cerca posible de los números elegidos por los otros dos jugadores, dentro de un margen de ± 5 .

Se decidió utilizar únicamente dos subjuegos ya que los efectos de la experiencia reportados por Slonim (2005) no mostraron ser significativamente diferentes al agregar un tercer subjuego.

Es importante mencionar que el Subjuego 1 es una replicación parcial del experimento de Lahav (2015), dado que es un juego repetido con explicitación de creencias.

3.3. Procedimiento

Para cada sesión, se citó a los 5 participantes en un aula donde se les leyeron las instrucciones del experimento: se les explicó cómo jugar p-beauty contest, se les dio la instrucción adicional de intentar adivinar los números a elegir por los otros jugadores (registrando dos creencias por periodo) y se les explicó que la sesión se dividiría en dos Subjuegos con tres participantes en cada uno. Finalmente, se hizo de su conocimiento las ganancias que podrían acumular en caso de dar el número que más se acercara al número objetivo computado a partir de todas los números emitidos (6 puntos para el jugador ganador, que en caso de empate se dividirían equitativamente entre los ganadores), o bien, de acertar en la predicción de las tiradas de sus compañeros (1 punto por cada número que estuviera dentro de un rango de ± 5 del número elegido por otro jugador). Se les pidió a los participantes que repitieran las reglas del juego para garantizar que habían entendido las instrucciones (las cuales se anexan como Apéndice).

Los participantes asignados mediante sorteo a los roles D y E esperaron en un aula separada, en la que se cuidó que no hablaran entre ellos. Mientras tanto, a los otros tres participantes (roles A, B y C) se les entregaron cuatro formatos de respuesta para que en cada periodo escribieran su número elegido y sus creencias. Los formatos incluían una clave de identificación única para cada participante, el número de periodo y los espacios para que los participantes escribieran su número elegido y su estimación de los números a elegir por los otros dos participantes. Un ejemplo de estos formatos se puede consultar en los Apéndices.

Los participantes A, B y C jugaron cuatro periodos de p-beauty contest, durante los cuales registraron tanto su propia elección como sus predicciones acerca de las elecciones de sus oponentes. Luego de que los participantes llenaran sus formatos en cada periodo, estos se recogían y se anotaba en el pizarrón del aula todos los números elegidos, sin indicar qué participante había elegido cada uno. Inmediatamente, y de manera explícita, se calculaba el promedio de estos números y el número objetivo, que también se anotaban en el pizarrón. Una vez hecho esto, se anunciaba qué jugador o jugadores habían elegido el número más cercano al número objetivo (con la clave de identificación asignada), y se les anotaba una ganancia de 6 puntos de juego distribuidos entre todos los ganadores. También se estableció qué jugadores habían estimado acertadamente los números elegidos por los otros jugadores con un margen de error de 5 números de diferencia y se les anotó una ganancia de 1 punto por cada acierto. Una vez realizado este proceso, se devolvía su formato de respuesta a cada candidato, incluyendo en el reverso de los mismos el total de puntos de juego obtenidos en ese periodo (que podía oscilar entre 0 y 8). Luego de que los participantes conocieran los puntos que obtuvieron en cada periodo, se borraba el pizarrón y se daba paso al siguiente periodo.

Al terminar los cuatro primeros periodos del Subjuego 1, se contabilizaron los puntos obtenidos por cada participante y se anunciaba al participante ganador. Si el jugador con más puntos era el participante B o C, se le anunciaba que obtendría otro medio punto extra sobre la calificación de su examen, además del medio punto que ya había obtenido por participar en el experimento. Si la

persona con más puntos era el participante A, se le decía que para obtener el medio punto extra adicional, debía ganar también en el siguiente Subjuego.

Tras anunciar el ganador del Subjuego 1, se despedía a los participantes B y C, y los participantes D y E ingresaban al aula. Antes de comenzar con el Subjuego 2, se dio un breve recordatorio de las reglas del juego, haciendo hincapié en que el participante A contaba con más experiencia, al haber jugado los cuatro periodos anteriores.

El Subjuego 2 se llevó a cabo de la misma forma que el Subjuego 1. Al final de cuatro periodos se contabilizaron los puntos de juego conseguidos por cada participante, y el que obtuvo la mayor cantidad recibió el segundo medio punto extra sobre su examen.

Con el término del subjuego 2 y el anuncio del participante ganador, se despedía a los participantes y se daba por terminada la sesión.

Capítulo 4

Resultados

Para el análisis de los resultados obtenidos, se utilizaron tanto pruebas estadísticas bayesianas como pruebas clásicas o frecuentistas. En el presente apartado se presentan únicamente los resultados de los análisis bayesianos, mientras que los resultados de los análisis clásicos se incluyen en los Apéndices. En general, ambas aproximaciones apuntan hacia las mismas conclusiones.

4.1. Análisis Bayesiano

Se decidió presentar los resultados del análisis bayesiano porque se considera que el factor de Bayes, que funciona como una alternativa al *p-value* como medida de significancia estadística, refleja de forma mucho más clara la relación entre los datos obtenidos y las hipótesis nula y alternativa, al señalar cuántas veces es más probable que la evidencia haya sido producida bajo la hipótesis alterna respecto a la hipótesis nula (BF_{10}), o viceversa (BF_{01}).

4.2. Consistencia entre creencias y elecciones en el primer subjuego

El primer subjuego del presente experimento presenta una replicación parcial del experimento de Lahav (2015). En ambos casos se juega p-beauty contest por varios periodos y se utiliza un método para explicitar creencias. Las diferencias radican en el número de jugadores que participan en un periodo (3 en el

presente experimento, y hasta 20 en el de Lahav), la cantidad de periodos jugados (4 en el presente experimento y 5 en el de Lahav), y el método utilizado para explicitar las creencias (en el presente experimento se preguntó directamente por el número que se creía elegirían los otros jugadores, mientras que en el experimento de Lahav se estimaba el número de jugadores que elegirían un número dentro de ciertos rangos).

En la Figura 4.1 se contrasta la elección de cada participante con el promedio de las creencias registradas para cada uno de los periodos del Subjuego 1, tomando en cuenta la multiplicación por p de éstas y omitiéndola (paneles izquierdos y derechos, respectivamente). Se observa que en ambos casos la diferencia entre creencias y elecciones se reduce periodo a periodo, pues en periodos posteriores los puntos se acercan más a la línea de identidad. Sin embargo, como ya se ha mencionado, esto puede deberse únicamente a que los valores elegidos se acercan más al límite inferior del espacio de elección.

Para evaluar la consistencia entre las creencias de los participantes y sus números elegidos tomando en cuenta el efecto de suelo, se emplearon dos métodos: el primero de ellos, computa la *Diferencia Normalizada* entre las creencias y las elecciones de los participantes de acuerdo con las elecciones promedio observadas en cada periodo (Lahav, ?); el segundo, calcula la *Diferencia Relativa* entre creencias y elecciones a partir del punto medio entre ambos valores (método empleado por Slonim, 2005, para calcular el cambio relativo de los números elegidos por los jugadores de un periodo al siguiente). A su vez, tal y como lo reporta Lahav, estos métodos fueron aplicados con dos variantes para evaluar el cómputo realizado por los participantes, incluyendo u omitiendo la multiplicación por el parámetro p .

El procedimiento sugerido por Lahav (2015) para calcular las diferencias normalizadas entre las creencias y las elecciones de cada participante en cada periodo, fue incorporado a partir de la siguiente ecuación:

$$DN_i^t = \frac{(\frac{2}{3}B_i^t - C_i^t)}{C_i^t}$$

Donde DN_i^t es la Diferencia Normalizada entre las creencias y elecciones de

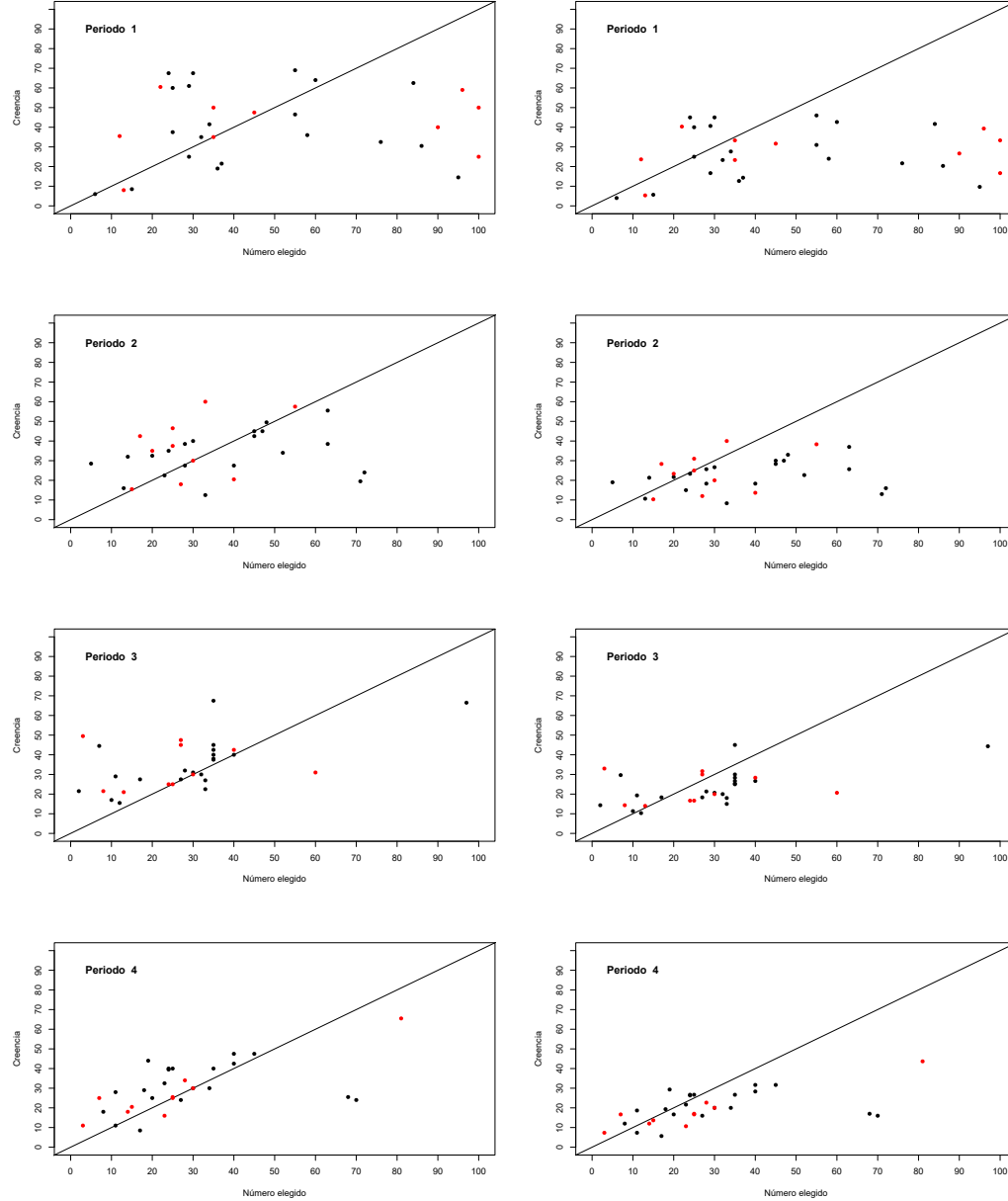


FIGURA 4.1: Comparación entre los números elegidos por los participantes y el promedio de las creencias registradas. Los puntos más cercanos a la línea de identidad señalan una mayor consistencia. En las figuras del lado izquierdo las creencias promedio son multiplicadas por p y en el lado derecho se omite dicha operación.

cada participante i en el periodo t , computada a partir de la diferencia entre la media de los números que el participante i estimó que elegirían los otros dos jugadores en el periodo t multiplicado por $\frac{2}{3}$, $(\frac{2}{3}B_i^t)$ y el número elegido por el propio participante i para ese periodo t (C_i^t), dividida por el promedio de los números elegidos por todos los participantes en el periodo t (\bar{C}_t).

Una vez computadas las diferencias por cada participante y periodo, se calculó el promedio de las mismas para poder someterlas a un análisis estadístico que permitiera evaluar si estas fueron significativamente diferentes de 0. Para ello, se realizaron pruebas-T bayesianas de una sola muestra para cada periodo de juego. En la Tabla 4.1 se reportan los Factores de Bayes obtenidos en dicho análisis, que permiten estimar qué tantas veces es más probable que la evidencia corresponda con la Hipótesis Alterna ("¿hay diferencia entre creencias y elecciones") respecto a la Hipótesis Nula ("no hay diferencia entre creencia y elecciones"). Como puede verse, sólo se encontraron diferencias significativas entre creencias y elecciones en los primeros dos periodos del juego.

CUADRO 4.1: Diferencias Normalizadas en el Subjuego 1 Prueba t bayesiana de una sola muestra que compara contra 0 el promedio de las Diferencias Normalizadas entre las creencias y las elecciones de los jugadores en cada periodo del Subjuego 1.

| | BF_{10} | $error\%$ | Diferencia promedio |
|-----------|-----------|---------------|----------------------------|
| Periodo 1 | 19.300 | $1.823e^{-6}$ | -0.366 |
| Periodo 2 | 34.545 | $3.137e^{-4}$ | -0.342 |
| Periodo 3 | 0.281 | $2.840e^{-5}$ | -0.097 |
| Periodo 4 | 0.652 | 0.015 | -0.147 |

En la Figura 4.2 se presenta de manera gráfica la relación entre las distribuciones prior y posterior computadas en cada periodo. Las distribuciones prior representan la Hipótesis Nula (asume diferencias cercanas a 0) y las distribuciones posteriores presentan la magnitud de la diferencia estimada a la luz de los datos. La forma más sencilla de interpretar estas figuras es como una razón de densidades de probabilidad: si la densidad de probabilidad es mayor en la distribución prior que en la distribución posterior en el punto de "no diferencias" (que señala un tamaño del efecto 0, $\delta = 0$), quiere decir que la evidencia favo-

rece la hipótesis alterna, ya que a la luz de la evidencia es "*muy poco probable*" (menos de lo que se esperaba de acuerdo a la distribución prior) que el tamaño de efecto tenga un valor cercano a 0.

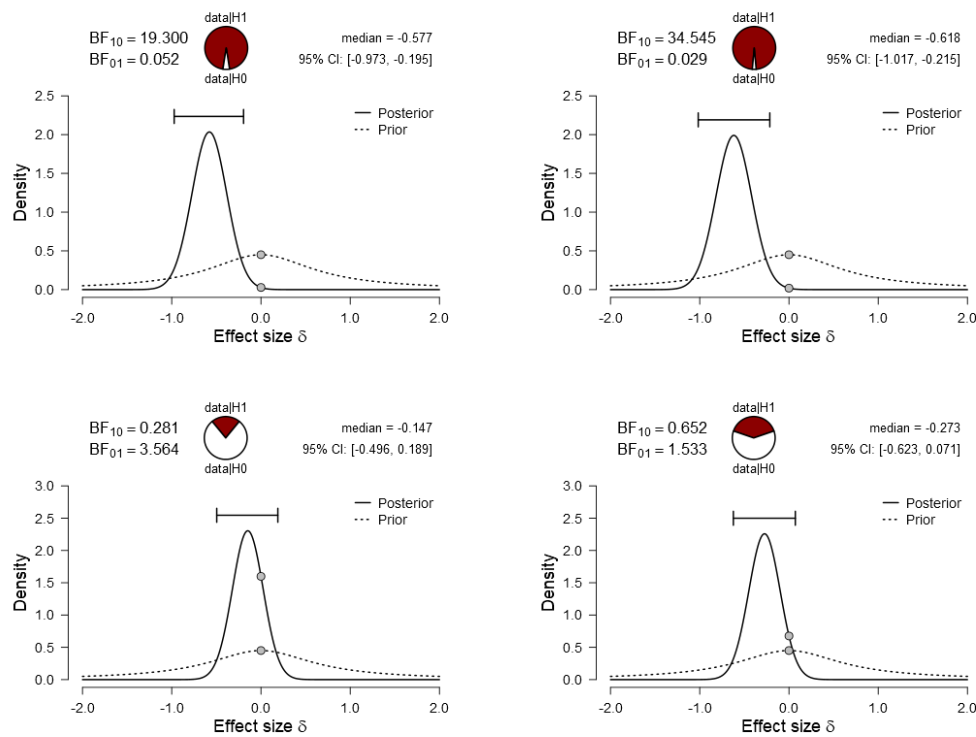


FIGURA 4.2: Para la evaluación con pruebas t de una muestra de las Diferencias Normalizadas entre las creencias y las elecciones registradas en los cuatro periodos del Subjuogo 1, se presenta la relación entre las distribuciones prior y posteriores computadas por cada periodo. La razón de probabilidad entre estas distribuciones en el punto $\delta = 0$ señala qué tan probable es que no hayan diferencias según los datos recabados (la distribución posterior), en comparación con la hipótesis nula (la distribución prior).

Estos resultados son consistentes con lo que reporta Lahav (2015): en los primeros periodos no hay consistencia entre las creencias y elecciones, pero ésta parece adquirirse conforme avanzan los periodos. Así mismo, en todos los periodos se encontraron diferencias negativas, sugiriendo que en promedio las creencias de los participantes estuvieron por debajo de sus elecciones reales.

En el estudio presentado por Lahav (2015), el cómputo de la Diferencia Normalizada entre las creencias y las elecciones se realizó también omitiendo la multiplicación de las creencias por p , en un intento por evaluar la posibilidad de que las inconsistencias halladas entre éstas y las elecciones se deban a que los participantes no habían considerado dicha operación. El presente trabajo también incorporó dicha variación del análisis, llevada a cabo de acuerdo a la siguiente ecuación, en la que ya no se incluye la multiplicación por $\frac{2}{3}$ en B_i^t :

$$DN_i^t = \frac{(B_i^t - C_i^t)}{C_t} \bar{C}_t$$

Nuevamente, las diferencias promedio computadas en cada periodo asumiendo que los participantes no multiplicaron sus creencias por p , fueron evaluadas con pruebas-T bayesianas de una sola muestra. Este análisis arrojó resultados inversos a los encontrados cuando la multiplicación por p fue tomada en cuenta: los primeros periodos no muestran diferencias significativas y los últimos, sí. Aunado a ello, las diferencias en los periodos 3 y 4 se vuelven positivas (indicando que las creencias cayeron por encima de las elecciones). Estos resultados se presentan en la Tabla 4.2.

CUADRO 4.2: Diferencias Normalizadas en el Subjuego 1 omitiendo la multiplicación por p : Prueba t bayesiana de una sola muestra que compara contra 0 el promedio de las Diferencias Normalizadas entre las creencias, omitiendo la multiplicación por p , y las elecciones de los jugadores en cada periodo del Subjuego 1.

| | BF_{10} | $error\%$ | Diferencia promedio |
|-----------|-----------|---------------|----------------------------|
| Periodo 1 | 0.207 | 0.010 | -0.049 |
| Periodo 2 | 0.196 | 0.013 | -0.012 |
| Periodo 3 | 3.811 | $3.017e^{-5}$ | 0.355 |
| Periodo 4 | 1.861 | $4.032e^{-4}$ | 0.280 |

De acuerdo con el Factor de Bayes, aunque la Hipótesis Alterna es más probable en los periodos 3 y 4 (es decir, que sí hay diferencias entre las creencias y las elecciones) respecto de la Hipótesis Nula, la evidencia acumulada es relativamente pequeña, (particularmente en el periodo 4, donde podría considerarse

anecdótica). En la Figura 4.3 se muestra la relación entre las distribuciones prior y posteriores de cada periodo.

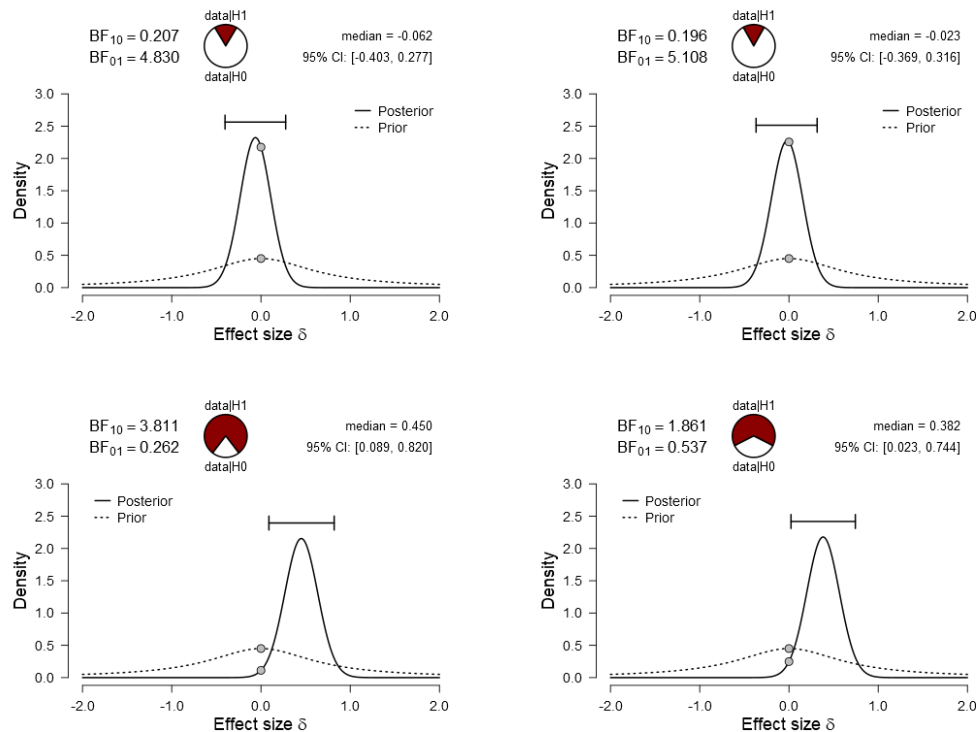


FIGURA 4.3: Para la evaluación con pruebas t de una muestra de las Diferencias Normalizadas entre las creencias (omitiendo la multiplicación por p) y las elecciones registradas en los cuatro periodos del Subjeto 1, se ilustra la relación entre las distribuciones prior y posteriores computadas por cada periodo. La razón de probabilidades en el punto $\delta = 0$ indica qué tan probable es, a la luz de los datos (la distribución posterior), que no hayan diferencias en comparación con lo que nuestra Hipótesis Nula supondría.

Este resultado difiere considerablemente de los hallazgos reportados por Lahav (2015), quien reportó que al excluir la multiplicación por $\frac{2}{3}$, las diferencias en los cuatro periodos se volvieron positivas, significativamente diferentes de 0, y en general, más grandes que cuando la multiplicación por p era tomada en cuenta. En el estudio conducido por Lahav, dichos resultados fueron interpretados como un indicador de que los participantes sí toman en cuenta la multiplicación

por p .

Cuando se incluye la multiplicación por p en el cálculo de las Diferencias Normalizadas entre creencias y elecciones en el presente estudio, se encuentra que éstas fueron significativas en los primeros periodos, y al excluir dicha multiplicación, las diferencias significativas aparecen sólo en los últimos periodos, aunque la evidencia parece ser débil. Estos resultados podrían estar sugiriendo a que los participantes comienzan el juego sin considerar la multiplicación por $\frac{2}{3}$, pero la incorporan en sus decisiones al avanzar en los periodos (o por lo menos, aprenden que el número objetivo siempre está por debajo del número promedio).

En general, a pesar de las diferencias en los métodos empleados para la elicitación de creencias, el presente estudio presenta hallazgos similares a los reportados por Lahav (2015) al emplear el método propuesto para calcular las Diferencias Normalizadas entre las creencias y las elecciones de los participantes:

- Existen discrepancias entre las creencias y las elecciones de los participantes en los primeros periodos de juego, pero no en los últimos.
- Las elecciones de los participantes suelen encontrarse entre el promedio de sus creencias respecto de los otros jugadores y el número objetivo computado de acuerdo a su multiplicación por p .

La diferencia más importante entre lo hallado en el presente estudio y lo reportado por Lahav (2015), es que, en promedio, los participantes no parecen incorporar la multiplicación por $\frac{2}{3}$ en su elección, al menos en los primeros periodos.

Además de replicar el método de Diferencias Normalizadas utilizado por Lahav (2015), las diferencias entre creencias y elecciones fueron evaluadas con un segundo método que no dependía de la elección promedio de todos los participantes en cada periodo para ponderarlas. La medida utilizada fue la Diferencia

Relativa entre las creencias y elecciones de cada participante i en cada periodo t , calculada de la siguiente manera:

$$DR_i^t = \frac{(\frac{2}{3}B_i^t - C_i^t)}{0,5(\frac{2}{3}B_i^t + C_i^t)}$$

El promedio de las Diferencias Relativas computadas por cada periodo fueron evaluadas en términos de qué tanto se alejaron de 0, mediante la realización de pruebas t bayesianas de una sola muestra, que se presentan en la Tabla 4.3, encontrando que en tres de los cuatro periodos hubo diferencias significativas entre las creencias y elecciones. En la Figura 4.4 se presenta la comparación entre las distribuciones prior y posterior computadas por cada periodo.

CUADRO 4.3: **Diferencias Relativas en el Subjuego 1** Prueba t bayesiana de una sola muestra que compara contra 0 el promedio de las Diferencias Relativas computadas entre las creencias y las elecciones de los jugadores en cada periodo del Subjuego 1.

| | BF_{10} | $error\%$ | Diferencia promedio |
|-----------|-----------|---------------|----------------------------|
| Periodo 1 | 72.283 | $7.775e^{-5}$ | -0.457 |
| Periodo 2 | 12.797 | $2.047e^{-6}$ | -0.328 |
| Periodo 3 | 0.214 | 0.008 | -0.052 |
| Periodo 4 | 1.871 | $4.022e^{-6}$ | -0.212 |

Tal y como se encontró con el método de Diferencias Normalizadas, todas las diferencias son negativas, indicando que consistentemente las creencias estuvieron por debajo de las elecciones reales. Sin embargo, contrario a lo que se esperaba con base en los resultados hallados con el método de Diferencias Normalizadas donde los participantes reducen la inconsistencia entre sus elecciones y las creencias explicitadas conforme adquieren experiencia, con el método de Diferencias Relativas se observan inconsistencias (diferencias estadísticamente significativas) entre creencias y elecciones en el periodo 4, aunque no en el periodo 3. De cualquier forma, de acuerdo con el factor de Bayes, la evidencia a favor de la hipótesis alterna en el periodo 4 es anecdótica.

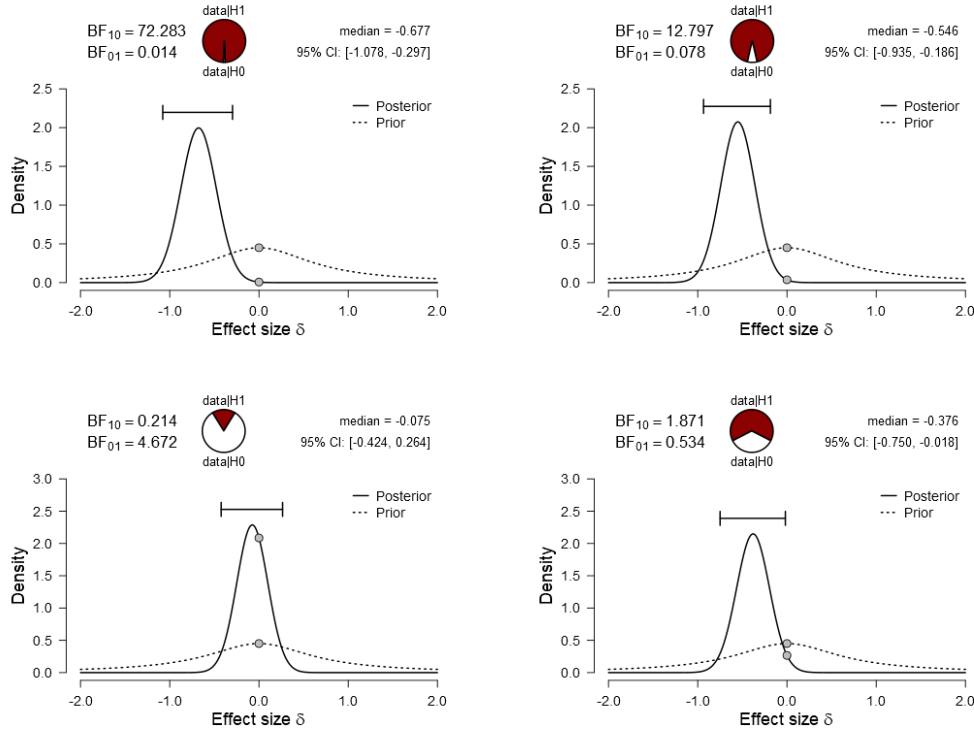


FIGURA 4.4: Para la evaluación con pruebas t de una muestra de las Diferencias Relativas encontradas entre las creencias y las elecciones de los participantes a lo largo del Subjuego 1, en el punto de no diferencias, se presenta la razón de probabilidades en el punto $\delta = 0$ que indica si a la luz de los datos (la distribución posterior) es más o menos probable (y qué tanto) que no hayan diferencias entre estas, de lo que supondría la Hipótesis Nula.

Posteriormente, se computaron las Diferencias Relativas omitiendo la multiplicación por p , de acuerdo con la siguiente ecuación:

$$DR_i^t = \frac{(B_i^t - C_i^t)}{0,5(B_i^t + C_i^t)}$$

Se realizaron pruebas t de una sola muestra para determinar si las Diferencias Relativas promedio en cada periodo son significativamente diferentes de 0, cuando la multiplicación por p no es tomada en cuenta. En la Tabla 4.4 se presentan los resultados obtenidos y en la Figura 4.5, la relación entre las distribuciones prior y las posteriores en cada periodo.

CUADRO 4.4: Diferencias Relativas en el Subjuego 1 omitiendo la multiplicación por p Prueba t bayesiana de una sola muestra que compara contra 0 el promedio de las Diferencias Relativas entre las creencias y las elecciones de los jugadores en cada periodo Subjuego 1.

| | BF_{10} | $error\%$ | Diferencia promedio |
|-----------|-----------|---------------|---------------------|
| Periodo 1 | 0.269 | $5.246e^{-4}$ | -0.101 |
| Periodo 2 | 0.211 | 0.009 | 0.044 |
| Periodo 3 | 8.393 | $2.307e^{-6}$ | 0.315 |
| Periodo 4 | 0.810 | $5.669e^{-6}$ | 0.167 |

Similar a lo observado cuando se omitió la multiplicación por p en el cómputo de las Diferencias Normalizadas propuesto por Lahav (2015), se encontró una reversión en la significancia reportada en todos los periodos, (aunque nuevamente, la evidencia en el periodo 4 es anecdótica), encontrando diferencias positivas en tres de los cuatro periodos. Esto último sugiere que con el uso del método de Diferencias Relativas, en promedio, las creencias están más cercanas y ligeramente por arriba de las elecciones reales de los participantes, cuando no se toma en cuenta la multiplicación por p .

Aunque cada uno de los dos métodos empleados para evaluar la consistencia entre las creencias y elecciones de los participantes intenta compensar la tendencia hacia el equilibrio y el problema de suelo resultante de forma diferente, ambos mostraron resultados muy similares: en los primeros periodos las diferencias entre creencias y elecciones son grandes, pero esta se reduce en los periodos posteriores, sugiriendo que los jugadores se vuelven consistentes conforme adquieren experiencia y aprenden también que para acercarse al número objetivo necesitan elegir números por debajo del promedio de sus creencias.

4.3. Efecto de reset

De acuerdo con los resultados reportados por Slonim (2005), los jugadores presentan un efecto de "Reset" en la tendencia a elegir números cada vez más pequeños cuando los otros jugadores son reemplazados por jugadores nuevos,

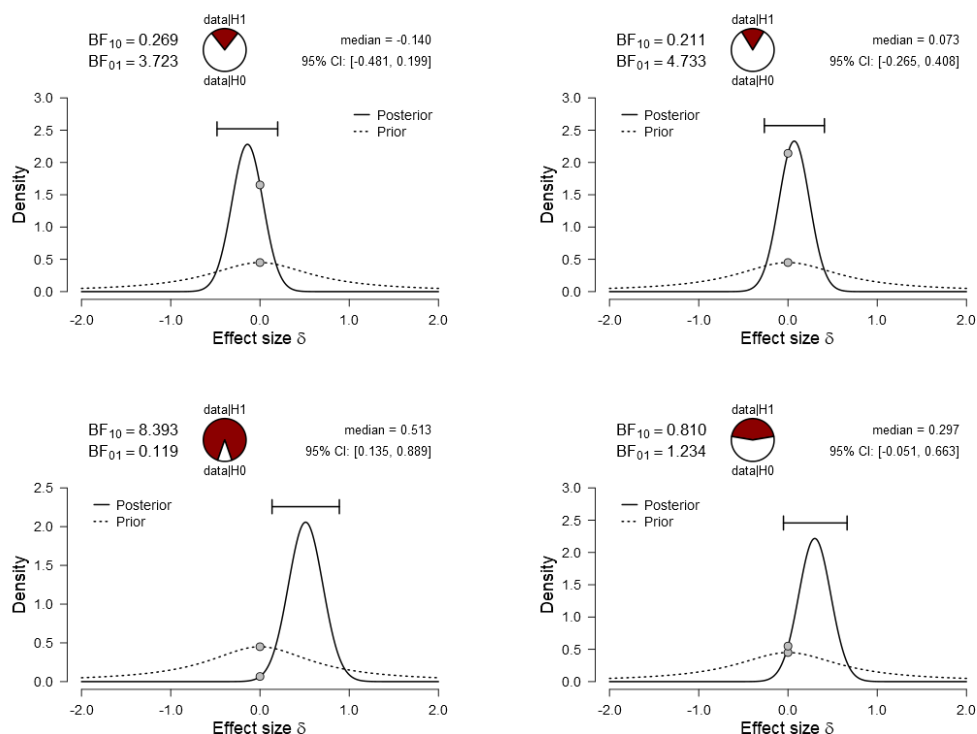


FIGURA 4.5: Para la evaluación con pruebas t de una muestra de las Diferencias Relativas encontradas entre las creencias (sin multiplicar por p) y las elecciones de los participantes a lo largo del Subjuego 1, se señala la razón de probabilidades en el punto $\delta = 0$ que indica si a la luz de los datos (la distribución posterior) es más o menos probable (y qué tanto) que no hayan diferencias entre estas, de lo que supondría la Hipótesis Nula.

que no tienen experiencia en el juego. Tomando este hallazgo en cuenta, el presente estudio incorporó un Subjuego 2, donde sólo uno de los jugadores del Subjuego 1 permaneció jugando por otros cuatro periodos, mientras que el resto fue reemplazado por jugadores nuevos. Esta manipulación se realizó para evaluar la tendencia que en el estudio de Lahav (2015) lleva a asumir que los jugadores se vuelven más consistentes conforme adquieren experiencia. En otras palabras, agregar un segundo Subjuego permite evaluar, con base en las respuestas del participante con experiencia, si la consistencia entre las elecciones y las creencias es algo que se adquiere con la experiencia o si es sólo el resultado del efecto de suelo asociado a la tendencia típicamente reportada en cualquier serie de

juegos p-beauty contest repetido a elegir números cada vez más pequeños, (Ho, 1988).

Para poder comparar los resultados del Subjuego 1 con el Subjuego 2, fue necesario comprobar que la incorporación de nuevos jugadores en el Subjuego 2 interrumpió la tendencia a elegir números cada vez más pequeños en los jugadores que se mantuvieron en el juego (los participantes A). Es decir, para corroborar que el diseño experimental propuesto permite responder a la cuestión de si las diferencias entre creencias y elecciones se reducen como reflejo de una consistencia adquirida o como producto del efecto de suelo, es necesario evaluar la presencia del efecto de Reset reportado por Slonim (2005).

En la Figura 4.6 se muestran los cambios en las elecciones realizadas en periodos consecutivos, por los participantes A. Los primeros tres cuadros muestran los cambios dentro del Subjuego 1 y en el cuarto panel se evalúa directamente el Efecto de Reset, al comparar la elección registrada en el último periodo del Subjuego 1 y el primer periodo del Subjuego 2. En estas gráficas, los puntos que caen por debajo de la línea de identidad indican que se eligieron números más pequeños de un periodo a otro, con lo que se observa que hay más participantes A que eligen números más pequeños entre los periodos del Subjuego 1 (por ejemplo, el 80 % reduce su elección entre el periodo 3 y el periodo 4), pero que esta tendencia se revierte al iniciar el subjuego 2 (el 70 % incrementa su número).

Se realizó una prueba Binomial bayesiana de una cola para comparar contra el azar la proporción de participantes A que aumentaban o disminuían sus elecciones entre cada uno de los primeros cinco periodos que jugaron (ver Tabla 4.5), encontrando evidencia anecdótica. La falta de robustez en los resultados obtenidos en esta prueba puede deberse a que solamente se realizaron 10 sesiones experimentales.

Para determinar si, en promedio, el número elegido por el participante A en el último periodo del Subjuego 1 es menor que el número elegido en el primer periodo del Subjuego 2, se realizó una prueba t de una cola, donde se encon-

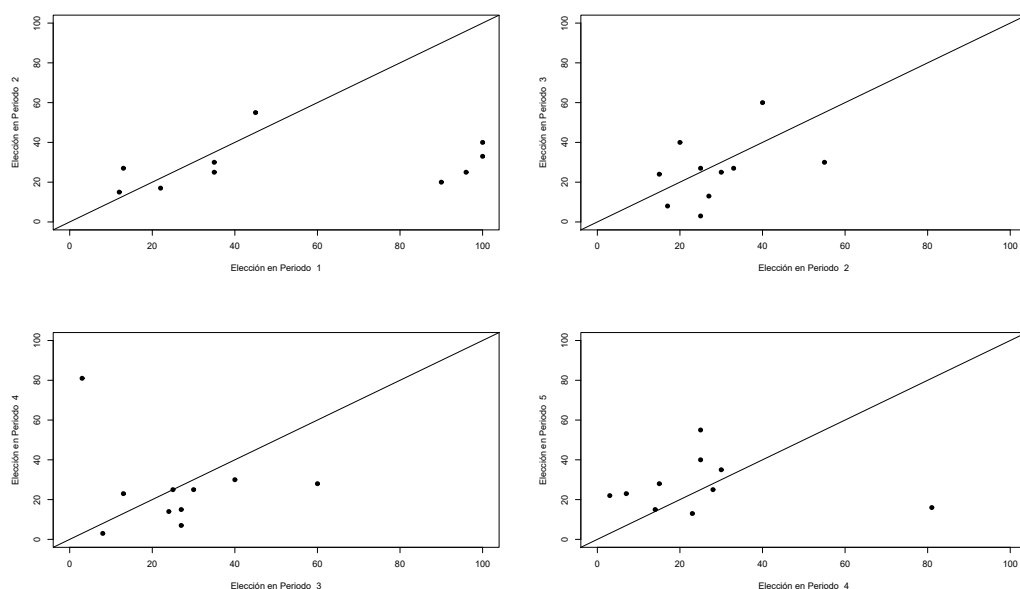


FIGURA 4.6: Se presenta la relación entre los números elegidos en periodos consecutivos por los participantes A, quienes no sólo jugaron los cuatro periodos del Subjuego 1, sino que permanecieron a lo largo del Subjuego 2. En los primeros tres pánels se aprecia que en el Subjuego 1 las elecciones tendían a disminuir, mientras que en el último panel, se observa un incremento en las mismas (el Efecto Reset).

tró que si bien el promedio de los números elegidos en el primer periodo del Subjuego 2 es mayor que el promedio de las elecciones registradas al final del Subjuego 1, la diferencia parece ser pequeña y no significativa.

No obstante, al revisar los datos de cada sesión experimental se encontraron anomalías en la ejecución del participante A de la Sesión 3, cuyo comportamiento difiere considerablemente de lo que se reporta en la literatura y en el resto de las sesiones realizadas en la presente investigación.

En la Figura 4.7 se presentan las elecciones de los participantes para cada uno de los ocho periodos jugados en los dos Subjuegos, en cada una de las diez sesiones experimentales. Se puede apreciar que el participante A de la Sesión 3 no presenta la tendencia a elegir números más pequeños periodo a periodo y por

CUADRO 4.5: Cambios en los números elegidos en periodos consecutivos Prueba binomial bayesiana que compara contra el azar el número de veces que el número elegido en periodos consecutivos aumentaba o disminuía.

| Periodos | Cambio | Casos | Proporción | BF_{10} |
|----------|-----------|-------|------------|-----------|
| 1 vs 2 | aumentó | 3 | 0.3 | 0.176 |
| | disminuyó | 7 | 0.7 | 1.376 |
| 2 vs 3 | aumentó | 4 | 0.4 | 0.243 |
| | disminuyó | 6 | 0.6 | 0.643 |
| 3 vs 4 | aumentó | 2 | 0.2 | 0.135 |
| | disminuyó | 8 | 0.8 | 4.002 |
| 5 vs 6 | aumentó | 7 | 0.7 | 1.376 |
| | disminuyó | 3 | 0.3 | 0.176 |

el contrario, elige números muy grandes en el último periodo de cada subjuego, una estrategia que no aporta ningún tipo de ventaja en el juego. Este patrón de respuesta sugiere que el participante A de la Sesión 3 pudo no haber entendido la dinámica del juego.

Para evaluar qué tan saliente es la tirada del participante A de la sesión 3 en el primer periodo del Subjuego 2, en la Figura 4.8 se presentan dos diagramas: El primero, de caja y bigotes, muestra que la elección realizada por el jugador A de la Sesión 3 en el primer periodo se encuentra por arriba del rango intercuadrático multiplicado por 1.5, lo que permite clasificarla como un valor atípico. El segundo diagrama, de violín, complementa la información del diagrama de caja y bigotes con un diagrama de densidad a los lados; la forma externa del diagrama representa todos los resultados posibles, y el grosor indica qué tan comunes son.

Tomando en cuenta dicho resultado, se repitieron las pruebas t de una cola para comprobar que el número elegido por los participantes A en el último periodo del Subjuego 1 son más pequeños que el número elegido en el primer periodo del Subjuego 2, omitiendo al participante A de la Sesión 3. Al realizar dicho análisis, el número elegido en promedio por los participantes A en el último periodo del subjuego 1 fue significativamente menor que la elección promedio registrada en el primer periodo del subjuego 2, encontrando evidencia

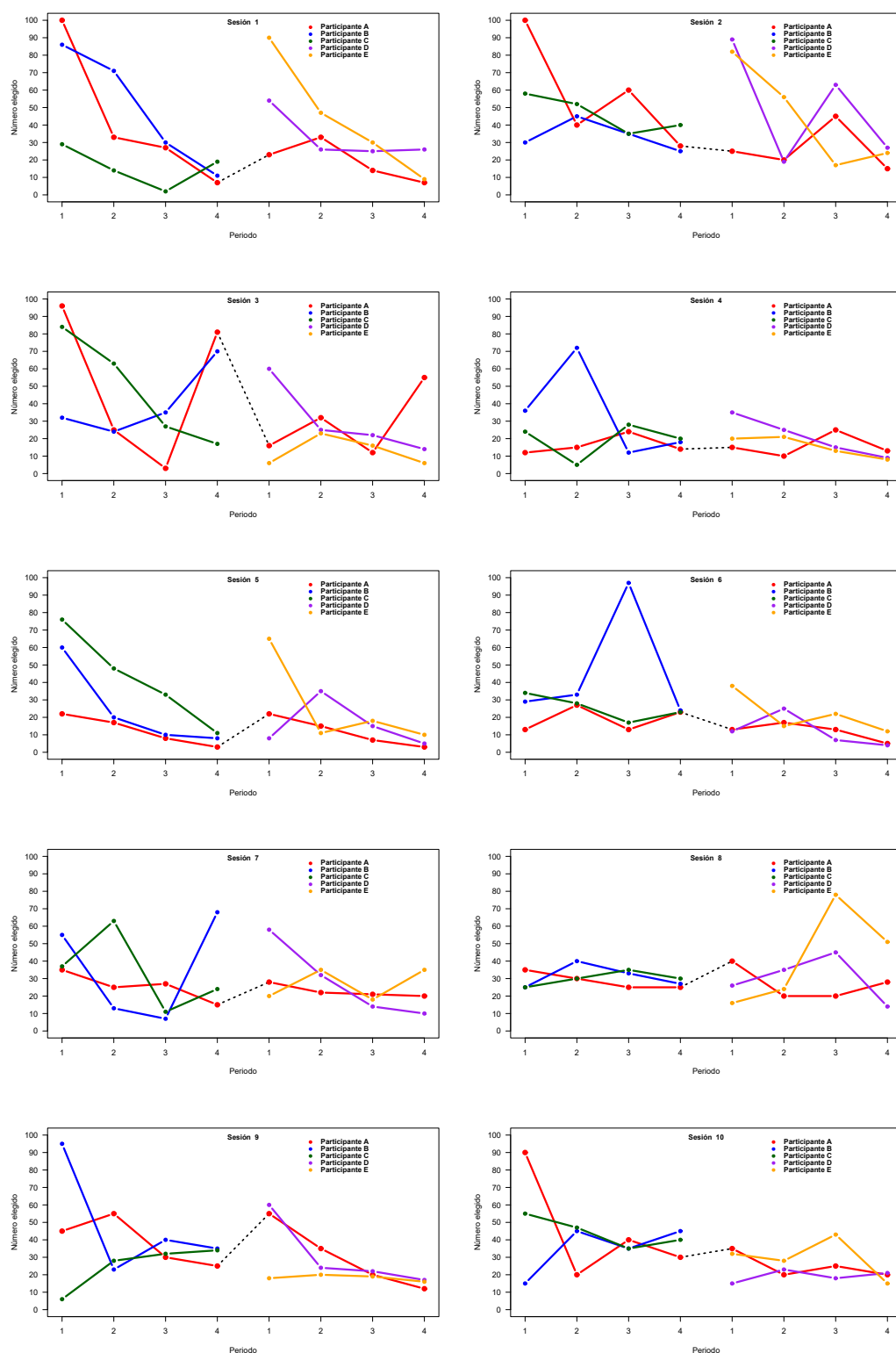


FIGURA 4.7: Se presentan todas elecciones registradas en las diez sesiones experimentales realizadas, en los Subjuegos 1 y 2. Se puede apreciar que el participante A de la sesión 3 tiene un patrón atípico respecto del resto de las sesiones y de lo que se esperaría de acuerdo a la literatura.

del efecto de reset (prueba clásica $t = -2,317$, $p = ,025 < 0,5$, prueba bayesiana $BF_{10} = 3,57$, $error = 1,377e - 4$).

En la Figura 4.9 se presenta el promedio de los números elegidos por los participantes A y los participantes no-A (participantes B y C, D y E para los Subjuegos 1 y 2, respectivamente) en cada periodo. Se presentan en gráficos diferentes los promedios computados cuando se consideran los datos de la Sesión 3 (panel izquierdo) y cuando no (panel derecho).

Con base en estos resultados, se puede afirmar que se logró replicar el efecto de reset reportado por Slonim (2005), omitiendo los datos recopilados de la Sesión 3. Con ello, se comprueba la validez en la comparación de la consistencia entre las creencias y elecciones de los participantes A en cada periodo, a lo largo de los dos subjuegos.

4.4. Consistencia entre creencias y elecciones entre subjuegos

Las creencias y elecciones de los participantes en el Subjuego 2 fueron sometidos a los mismos análisis reportados para el subjuego 1 en la **sección 4.2**: se computaron las Diferencias Normalizadas y las Diferencias Relativas, incluyendo y excluyendo la multiplicación por p . Pero en esta ocasión, se hizo una distinción entre la ejecución de los jugadores con experiencia (participantes A) y los jugadores que participaban en el juego por primera vez (participantes D y E).

En la figura 4.10 se contrasta la elección de cada participante con su creencia promedio para cada uno de los periodos del subjuego 2, cuando se incluye la multiplicación por p y cuando esta se omite. Los participantes A están marcados en rojo. En general, se observa que estos participantes fueron más consistentes en sus creencias y elecciones comparados con los participantes D y E.

Se calculó la Diferencia Normalizada entre las creencias y elecciones registradas en cada periodo del Subjuego 2, por los participantes A y los participantes D y E. Posteriormente, se realizaron pruebas t bayesianas para determinar si estas diferencias fueron estadísticamente diferentes de 0.

En las Tablas 4.6 y 4.7 se presentan los resultados obtenidos en las pruebas t bayesianas de una sola muestra que evalúan el promedio de las Diferencias Normalizadas computadas entre creencias y elecciones contra 0. De acuerdo con la Tabla 4.6, los participantes A no mostraron diferencias significativas (el factor de Bayes en este caso indica qué tantas veces es más que no haya diferencias), lo que indica una mayor consistencia entre sus creencias y elecciones en todos los periodos del subjuego 2. En contraste, en la Tabla 4.7 se puede observar que los participantes D y E mostraron diferencias significativas en los primeros tres periodos del subjuego, y sólo parecieron adquirir consistencia entre sus creencias y elecciones en el último periodo. Este último resultado es muy similar a lo reportado en los participantes del Subjuego 1, donde todos los jugadores compartían el mismo nivel de experiencia. Además, en todos los casos, las creencias estuvieron por debajo de las elecciones reales. En las Figuras 4.11 y 4.12 se incluyen las distribuciones prior y posterior de las Diferencias Normalizadas por periodo para los participantes A y los participantes D y E, respectivamente.

CUADRO 4.6: Diferencias Normalizadas en el Subjuego 2 (Participantes A) Se presentan los resultados de una prueba t de una muestra que compara contra 0 las Diferencias Normalizadas computadas para los participantes A en el Subjuego 2

| | BF_{01} | $error\%$ | Diferencia promedio |
|-----------|-----------|-----------|----------------------------|
| Periodo 5 | 3.010 | 0.006 | -0.049 |
| Periodo 6 | 3.212 | 0.007 | -0.011 |
| Periodo 7 | 2.012 | 0.003 | -0.133 |
| Periodo 8 | 3.205 | 0.007 | -0.028 |

Se repitió el cálculo de las Diferencias Normalizadas en los cuatro periodos del Subjuego 2 omitiendo la multiplicación por p , con el propósito de evaluar si los jugadores tomaron en cuenta este cálculo en la elección de su número. Tras la realización de las pruebas t bayesianas de una sola muestra correspondientes,

CUADRO 4.7: Diferencias Normalizadas en el Subjuego 2 (Participantes A) Se presentan los resultados de una prueba *t* de una muestra que compara contra 0 las Diferencias Normalizadas computadas para los participantes A en el Subjuego 2

| | BF_{01} | <i>error</i> % | Diferencia promedio |
|-----------|-----------|----------------|---------------------|
| Periodo 5 | 0.315 | 0.003 | -0.349 |
| Periodo 6 | 0.102 | 6.001e-4 | -0.243 |
| Periodo 7 | 0.063 | 2.519e-4 | -0.307 |
| Periodo 8 | 4.002 | 0.022 | -0.047 |

sólo se encontraron diferencias significativas entre las elecciones y las creencias de los participantes A en los dos primeros periodos, (ver Tabla 4.8). En general, la magnitud de las diferencias parece ser mayor en todos los periodos cuando se excluye la multiplicación por p que cuando esta sí se incluye. Este resultado sugiere que los jugadores con experiencia previa sí tomaron en cuenta la multiplicación por p , o por lo menos, que aprendieron desde el Subjuego 1 que el número objetivo siempre está por debajo del promedio de los números elegidos.

Por otra parte, en el caso de los participantes sin experiencia (D y E) solo se encontraron diferencias significativas respecto de 0 en el último periodo (ver Tabla 4.9), resultado que coincide con lo reportado en el Subjuego 1, sugiriendo nuevamente que los participantes incorporaron la multiplicación por p (o comprendieron la tendencia del juego hacia el equilibrio) sólo después de varias repeticiones. Así mismo, en las Figuras 4.13 y 4.14 se presentan las distribuciones prior y posterior computadas con las pruebas *t* realizadas por cada periodo, para los participantes con y sin experiencia, respectivamente.

Posteriormente, se computaron las Diferencias Relativas entre las creencias y elecciones registradas en los cuatro periodos del Subjuego 2 y se realizaron pruebas *t* bayesianas para determinar si estas eran significativamente diferentes de 0. Los resultados de este análisis se presentan en la Tabla 4.10 para el participante A y en la Tabla 4.11 para los participantes D y E.

Los resultados que se obtienen son muy similares a lo que se observa al utilizar el método de Diferencias Normalizadas: el participante A muestra ser con-

CUADRO 4.8: Diferencias Normalizadas en el Subjuego 2 (Participantes A) Se presentan los resultados de una prueba t de una muestra que compara contra 0 las Diferencias Normalizadas computadas para los participantes A en el Subjuego 2, cuando las creencias registradas no son multiplicadas por p .

| | BF_{01} | $error\%$ | Diferencia promedio |
|-----------|-----------|-----------|---------------------|
| Periodo 5 | 0.537 | 0.001 | 0.333 |
| Periodo 6 | 0.046 | 9.917e-5 | 0.414 |
| Periodo 7 | 1.274 | 0.004 | 0.238 |
| Periodo 8 | 0.643 | 0.003 | 0.438 |

CUADRO 4.9: Diferencias Normalizadas en el Subjuego 2 (Participantes D y E) Se presentan los resultados de una prueba t de una muestra que compara contra 0 las Diferencias Normalizadas computadas para los participantes D y E en el Subjuego 2, cuando las creencias registradas no se multiplican por p .

| | BF_{01} | $error\%$ | Diferencia promedio |
|-----------|-----------|-----------|---------------------|
| Periodo 5 | 4.265 | 0.022 | 0.024 |
| Periodo 6 | 1.417 | 0.004 | 0.170 |
| Periodo 7 | 3.541 | 0.021 | 0.071 |
| Periodo 8 | 0.115 | 7.221e-4 | 0.440 |

sistente en los cuatro periodos y los participantes D y E son inconsistencias en los primeros tres. En las Figuras 4.15 y 4.16 se incluyen las distribuciones prior y posterior computadas como resultado de las pruebas t bayesianas realizadas por periodo, por cada tipo de participante.

Finalmente, se repitió el cálculo de las Diferencias Relativas entre las creencias y las elecciones de los participantes, omitiendo la multiplicación por p . Se realizaron pruebas t bayesianas para comparar las diferencias computadas en cada periodo contra 0. En la Tabla 4.12 se presentan los resultados de las pruebas realizadas para evaluar las diferencias calculadas para los participante A en cada periodo, y en la Tabla 4.13 para los participantes D y E. De acuerdo a estos análisis, en los participantes A se observa una reversión de las significancias estadísticas reportadas cuando la multiplicación por p es incluida y diferencias, en promedio, más grandes. Esto sugiere que las creencias de los participantes

CUADRO 4.10: Diferencias Relativas en el Subjuego 2 (Participantes A) Se presentan los resultados de una prueba t de una muestra que compara contra 0 las Diferencias Normalizadas computadas para los participantes D y E en el Subjuego 2.

| | BF_{01} | error % | Diferencia promedio |
|-----------|-----------|---------|---------------------|
| Periodo 5 | 3.085 | 0.006 | -0.042 |
| Periodo 6 | 3.231 | 0.007 | -0.007 |
| Periodo 7 | 1.831 | 0.002 | -0.162 |
| Periodo 8 | 3.165 | 0.007 | 0.038 |

CUADRO 4.11: Diferencias Normalizadas en el Subjuego 2 (Participantes D y E) Se presentan los resultados de una prueba t de una muestra que compara contra 0 las Diferencias Relativas computadas para los participantes D y E en el Subjuego 2.

| | BF_{01} | error % | Diferencia promedio |
|-----------|-----------|---------------|---------------------|
| Periodo 5 | 0.272 | 0.002 | -0.407 |
| Periodo 6 | 0.044 | $3.044e^{-4}$ | -0.283 |
| Periodo 7 | 0.080 | $4.394e^{-4}$ | -0.341 |
| Periodo 8 | 4.295 | 0.022 | -0.007 |

A son más consistentes con sus elecciones cuando se asume que tomaron en cuenta la multiplicación por p . Por su parte, los participantes D y E también presentan una reversión en las significancias encontradas cuando la multiplicación por p es tomada en cuenta, pero las diferencias promedio parecen ser más pequeñas, excepto en el último periodo. Este último hallazgo es consistente con la idea sugerida por lo reportado al evaluar las diferencias considerando la multiplicación por p : los participantes aprenden a multiplicar por p en los últimos periodos del Subjuego. En las Figuras 4.17 y 4.18 se presentan las distribuciones prior y posterior computadas en cada prueba t realizada por cada periodo para los participantes A y para los participantes D y E, respectivamente.

En general, los resultados obtenidos en términos de la evaluación de la consistencia con que los jugadores con experiencia responden en el segundo sub juego en comparación con los otros jugadores sin experiencia, sugieren que la reducción reportada por Lahav (2015) en las diferencias entre las elecciones y las creencias de los participantes a lo largo una serie de periodos de p-beauty

CUADRO 4.12: Diferencias Normalizadas en el Subjuego 2 (Participantes A) Se presentan los resultados de una prueba t de una muestra que compara contra 0 las Diferencias Relativas computadas para los participantes A en el Subjuego 2, cuando no se asume que multiplican sus creencias por p .

| | BF_{01} | $error\%$ | Diferencia promedio |
|-----------|-----------|-----------|---------------------|
| Periodo 5 | 3.883 | 4.510e-4 | 0.346 |
| Periodo 6 | 25.537 | 1.528e-4 | 0.386 |
| Periodo 7 | 0.896 | 0.008 | 0.225 |
| Periodo 8 | 2.753 | 1.237e-4 | 0.413 |

CUADRO 4.13: Diferencias Normalizadas en el Subjuego 2 (Participantes D y E) Se presentan los resultados de una prueba t de una muestra que compara contra 0 las Diferencias Relativas computadas para los participantes D y E en el Subjuego 2, cuando las creencias no son multiplicadas por p .

| | BF_{01} | $error\%$ | Diferencia promedio |
|-----------|-----------|-----------|---------------------|
| Periodo 5 | 4.076 | 0.022 | -0.056 |
| Periodo 6 | 1.932 | 0.010 | 0.108 |
| Periodo 7 | 4.066 | 0.022 | 0.039 |
| Periodo 8 | 0.048 | 3.096e-4 | 0.372 |

contest, ocurre como producto de la experiencia adquirida por los participantes y no como co-producto del efecto de suelo asociado a la tendencia identificada en este tipo de juegos a elegir números cada vez más pequeños.

4.5. ¿Las creencias se vuelven más precisas con la experiencia?

Finalmente, dado el efecto que la experiencia demostró tener sobre la ejecución de los participantes en juegos repetidos de *p-beauty* contest, tanto en la elección de sus números de acuerdo a la experiencia de los otros jugadores, como a la adquisición de una mayor consistencia entre sus propias creencias y elecciones, se valoró una última pregunta de investigación derivada del tipo

de datos obtenidos con el diseño propuesto. Dicha pregunta estuvo orientada a evaluar el efecto que pudo haber tenido la experiencia sobre la precisión en las predicciones que los jugadores experimentados hacen sobre las elecciones de los otros jugadores.

En las Figuras 4.19 y 4.20 se presenta la comparación de las creencias de cada jugador en cada periodo contra los números que de hecho eligieron los otros jugadores. Se observa una alta falta de precisión en los primeros periodos de cada subjuego (periodos 1 y 5). La precisión del participante A (marcado en rojo) parece ser similar a la de los otros participantes. En la Figura 4.19 se toma en cuenta la multiplicación por p y en la Figura 4.20, no.

Se realizó un último conjunto de pruebas t de dos muestras para comparar el número de veces que los participantes lograron acercarse en su predicción, dentro de un margen de error de $\pm \frac{1}{5}$, a los números elegidos por sus oponentes en cada periodo jugado.

Primero, para evaluar el efecto de la experiencia en la precisión de las predicciones registradas, se comparó el número de aciertos cometidos por los participantes A en los Subjuegos 1 y 2, (antes y después de haber adquirido experiencia). Se encontró que el número de predicciones acertadas obtenidas en el Subjuego 2 (41,25 %) es mayor que en el Subjuego 1 (32,50 %), sin embargo, esta diferencia no se encontró significativa (prueba clásica $t = -1,312$, $p = 0,197 > 0,05$, prueba bayesiana $BF_{10} = 0,377$, $error = 4,241e - 6$).

Tampoco se observaron diferencias significativas en la precisión con que los participantes no-A de cada Sesión consiguieron predecir las elecciones de los otros jugadores. Los participantes B y C, que participaron en el Subjuego 1, obtuvieron un 33,125 % de los aciertos, mientras que los participantes D y E alcanzaron un 35 % en el Subjuego 2 (prueba clásica $t = -0,349$, $p = 0,728 > 0,05$, prueba bayesiana $BF_{10} = 0,131$, $error = 2,907e - 5$).

Finalmente, aunque la proporción de aciertos obtenidos por los participantes A no aumentó significativamente entre los Subjuegos 1 y 2, se evaluó la me-

jora en las predicciones registradas periodo a periodo (que se vería reflejado en una reducción en las diferencias entre las creencias registradas y las elecciones realizadas), comparando el desempeño de los participantes A con el resto mediante el uso de pruebas t bayesianas de muestras independientes. En la Tabla 4.14 se reportan los resultados de este análisis. La evidencia indica que no hay diferencias en el desempeño del participante A con respecto a los demás, en ningún periodo del juego, y que las diferencias entre las creencias reportadas y las elecciones observadas en los otros jugadores no cambian de forma sistemática entre periodos. En las Figura 4.21 se presentan las distribuciones prior y posterior computadas como resultado de dicho análisis.

CUADRO 4.14: Diferencias en la precisión de las predicciones hechas por los participantes con y sin experiencia Se presentan los resultados de una prueba t muestras independientes para comparar los aciertos conseguidos en la predicción de las jugadas de sus oponentes por los participantes A y no A, en cada periodo jugado.

| | BF_{10} | $error\%$ | Media (A) | Media (A') |
|-----------|-----------|-----------|-----------|------------|
| Periodo 1 | 0.468 | 1.064e-5 | -3.500 | -9.400 |
| Periodo 2 | 0.408 | 1.065e-5 | -1.900 | -0.150 |
| Periodo 3 | 0.428 | 6.613e-5 | 4.350 | 7.525 |
| Periodo 4 | 0.526 | 0.001 | -2.400 | 4.050 |
| Periodo 5 | 0.424 | 5.433e-5 | -1.800 | 2.475 |
| Periodo 6 | 0.403 | 3.276e-6 | 5.700 | 6.525 |
| Periodo 7 | 0.434 | 7.814e-5 | -1.950 | 1.000 |
| Periodo 8 | 0.409 | 1.241e-5 | 6.150 | 5.025 |

Con base en los análisis realizados para la evaluación de los posibles efectos en la precisión de las predicciones realizadas por cada jugador acerca de las elecciones de sus oponentes, se concluye que la experiencia no parece tener un efecto significativo sobre la habilidad de los participantes de anticipar las tiradas de sus contrincantes. Este resultado hace sentido con la manipulación experimental propuesta en el presente estudio: los participantes D y E que juegan con los participantes A en el segundo subjuego, son independientes de los participantes con los que jugaba en el primero (B y C), y no habría razón para esperar que sigan las mismas estrategias en la elección de sus tiradas.

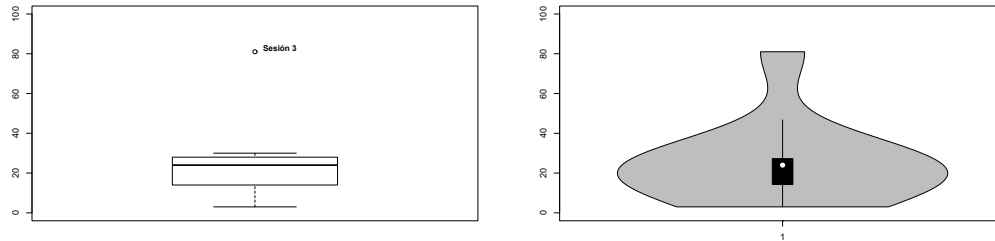


FIGURA 4.8: Comparación del desempeño de los participantes A con el desempeño observado en la Sesión 3. Del lado izquierdo se presenta un diagrama de caja y bigotes, donde se puede apreciar que el desempeño registrado en la Sesión 3 está muy alejado del resto. Del lado derecho, un diagrama de violín, que complementa esta información con un diagrama de densidad.

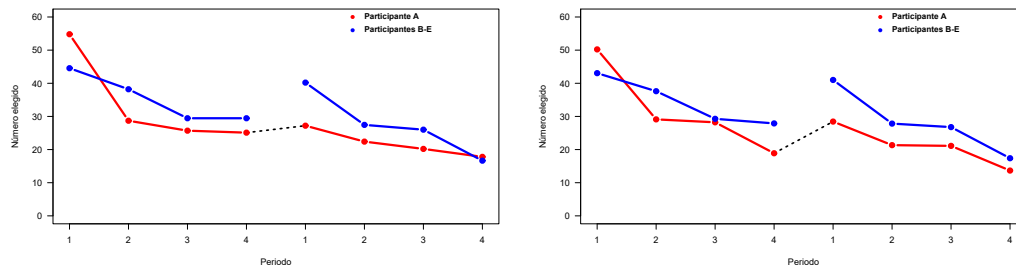


FIGURA 4.9: Se presenta el promedio de los números elegidos en cada uno de los periodos jugados por los participantes A (en color rojo) y los participantes B y C en el Subjuego 1 y los participantes D y E en el Subjuego 2 (en color azul). La gráfica de la izquierda incorpora los datos obtenidos en la Sesión 3 y la gráfica de la derecha, no.

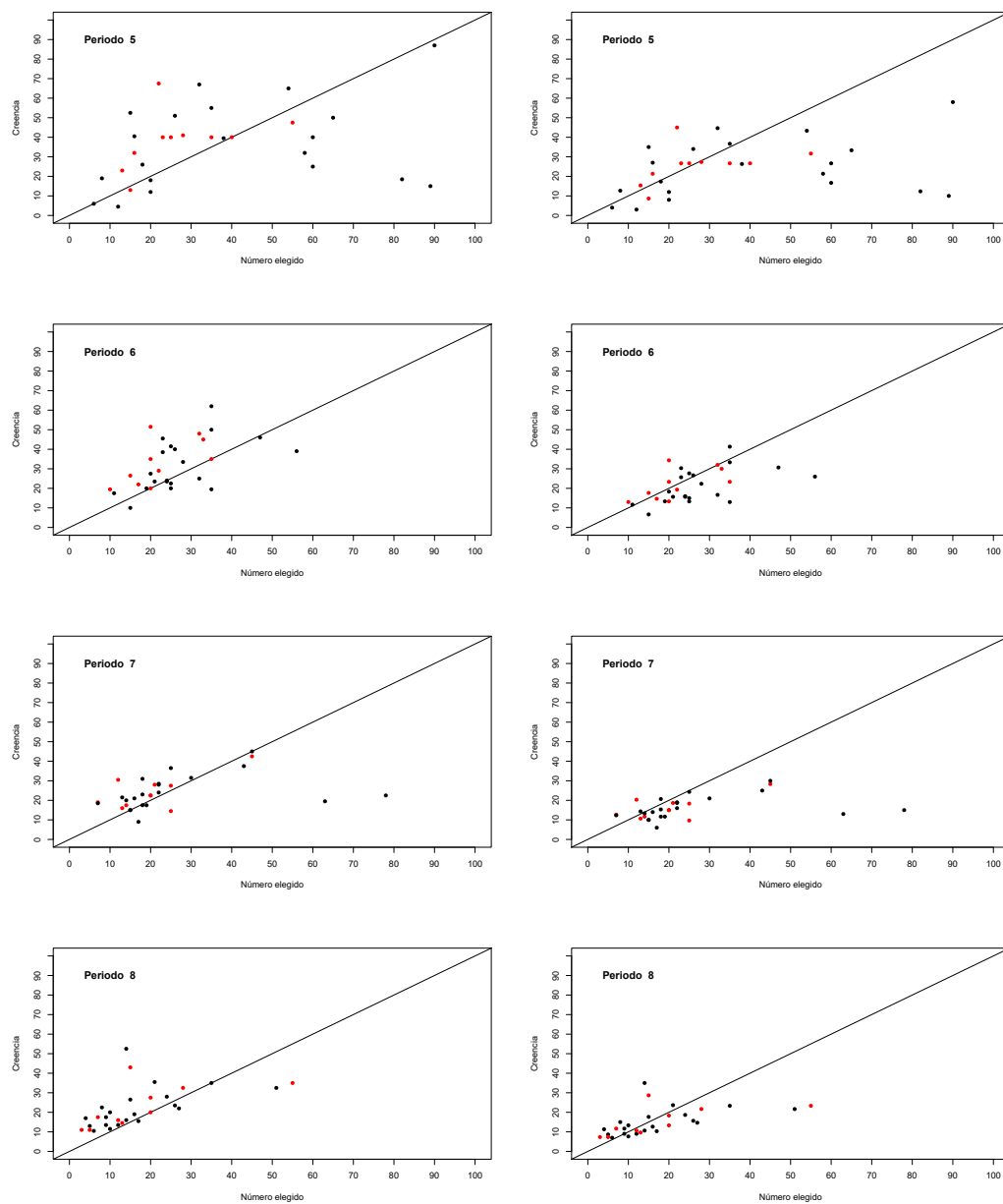


FIGURA 4.10: Se presenta la relación entre el promedio de las creencias registradas en cada periodo por los participantes del Subjuego 2 y sus elecciones, cuando sus creencias son multiplicadas por p (paneles izquierdos) o no (paneles derechos). Los participantes A, que ya tenían la experiencia de haber jugado en el Subjuego 1, se señalan en rojo.

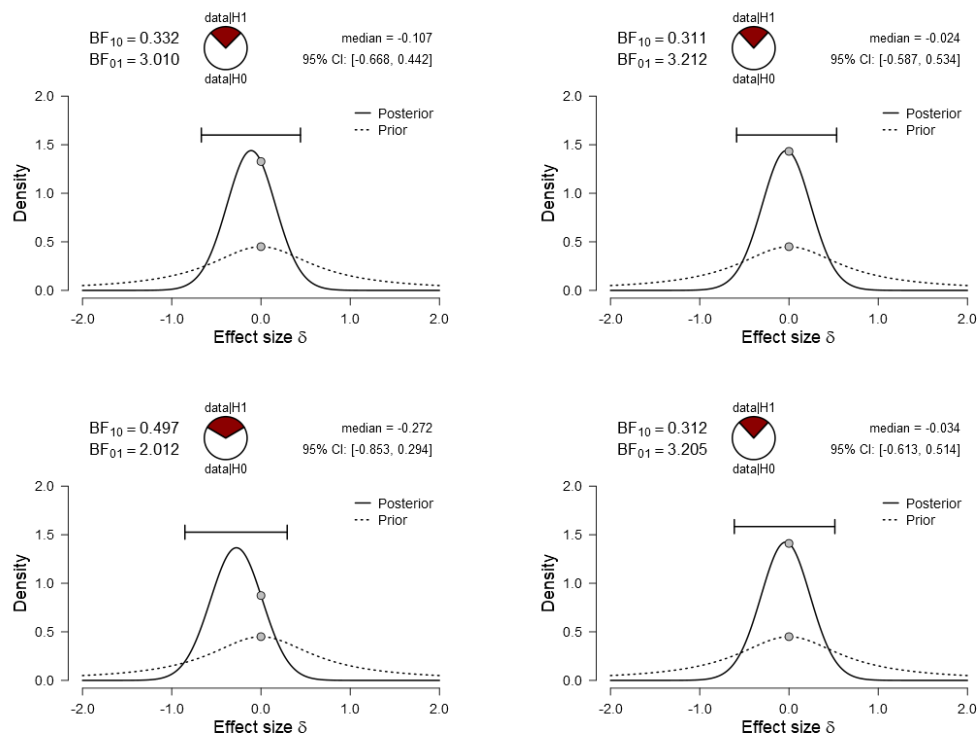


FIGURA 4.11: Como resultado de la evaluación de las Diferencias Normalizadas computadas entre las creencias y las elecciones de los participantes A en el Subjuego 2, se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

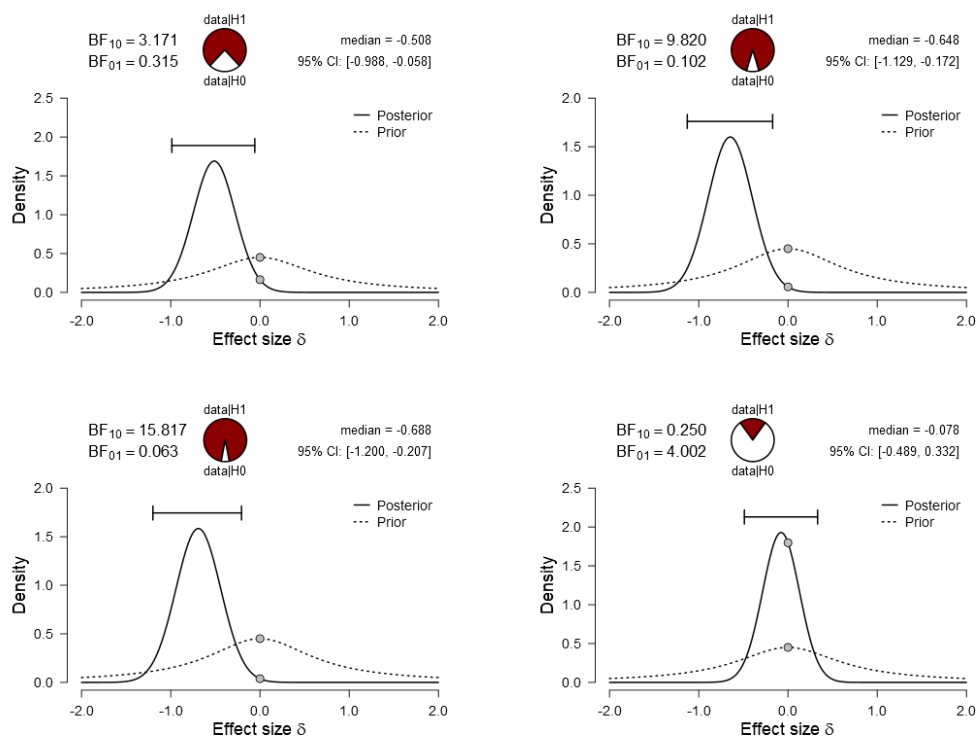


FIGURA 4.12: Como resultado de la evaluación de las Diferencias Normalizadas computadas entre las creencias y las elecciones de los participantes D y E en el Subjuego 2, se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

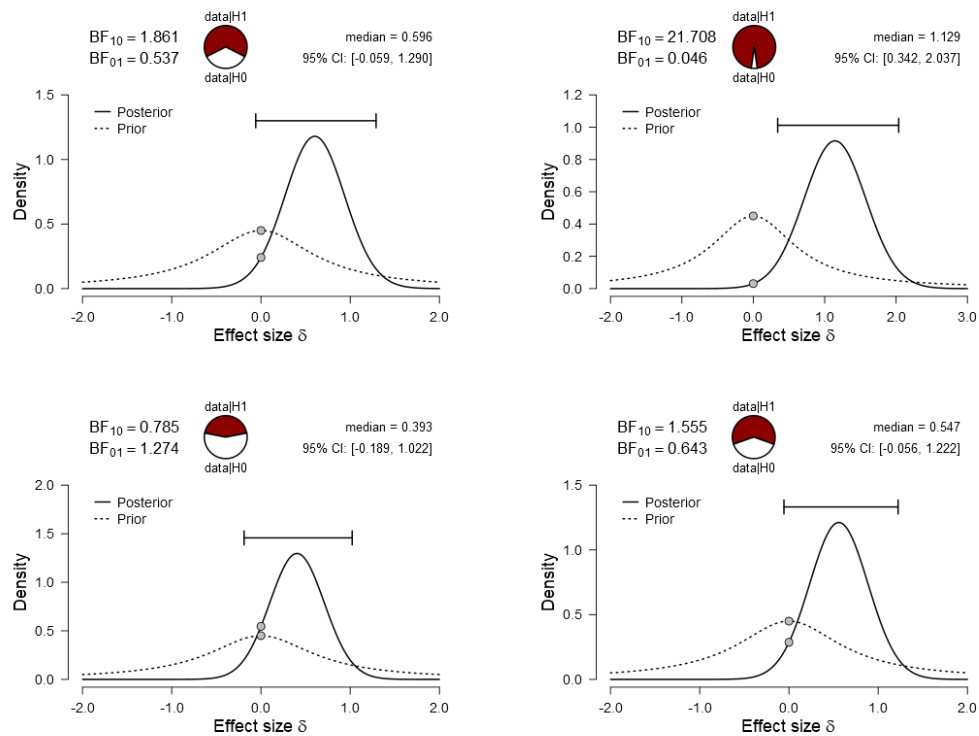


FIGURA 4.13: Como resultado de la evaluación de las Diferencias Normalizadas computadas entre las creencias (omitiendo la multiplicación por p) y las elecciones de los participantes A en el Subjuego 2, se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

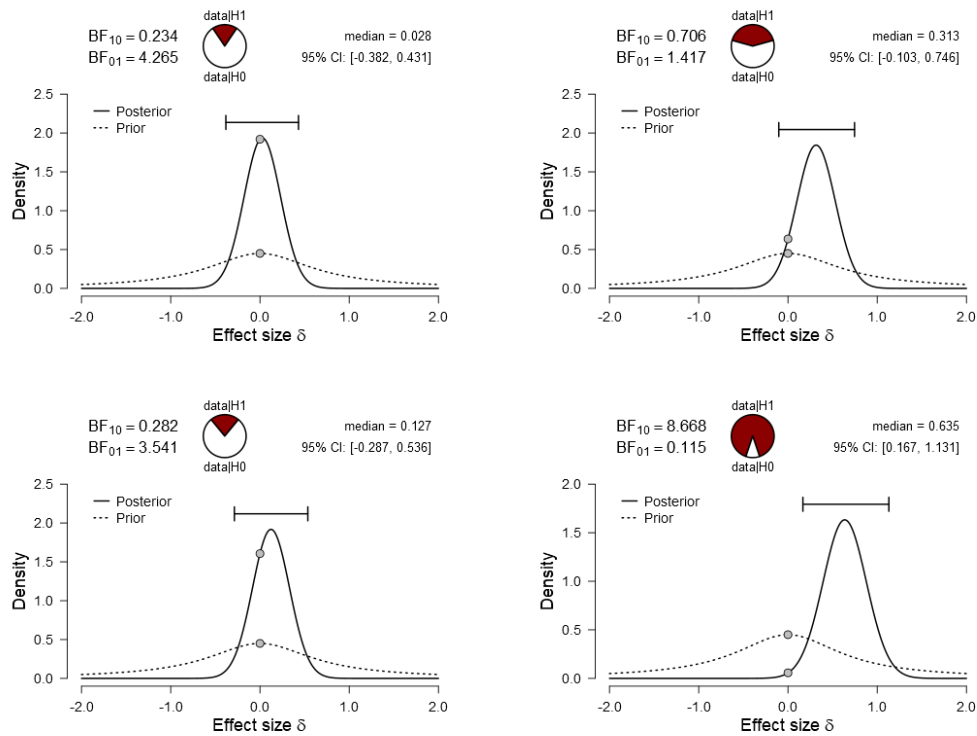


FIGURA 4.14: Como resultado de la evaluación de las Diferencias Normalizadas computadas entre las creencias no multiplicadas por p y las elecciones de los participantes D y E en el Subjuego 2, se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

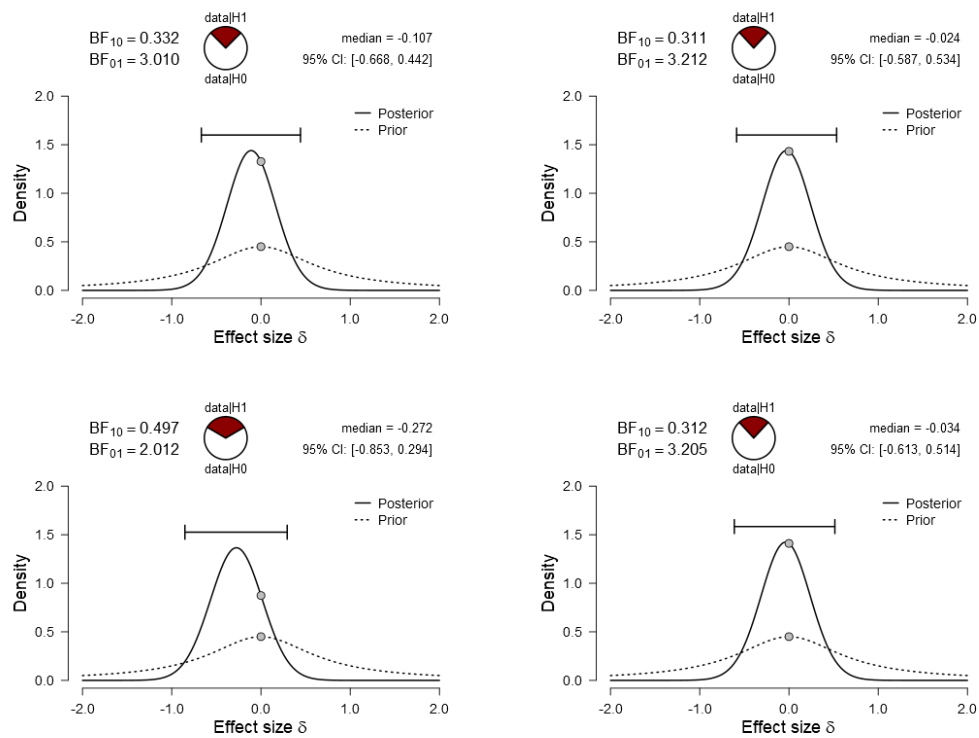


FIGURA 4.15: Como resultado de la evaluación de las Diferencias Relativas computadas entre las creencias y las elecciones de los participantes A en el Subjuego 2, se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

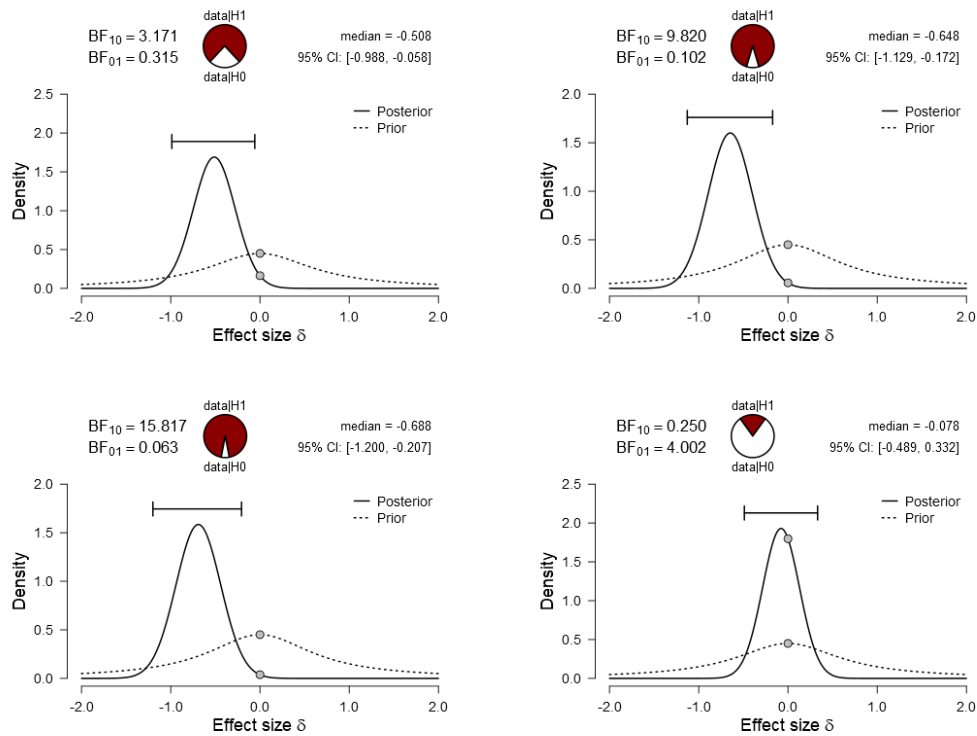


FIGURA 4.16: Como resultado de la evaluación de las Diferencias Relativas computadas entre las creencias y las elecciones de los participantes D y E en el Subjuego 2, se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

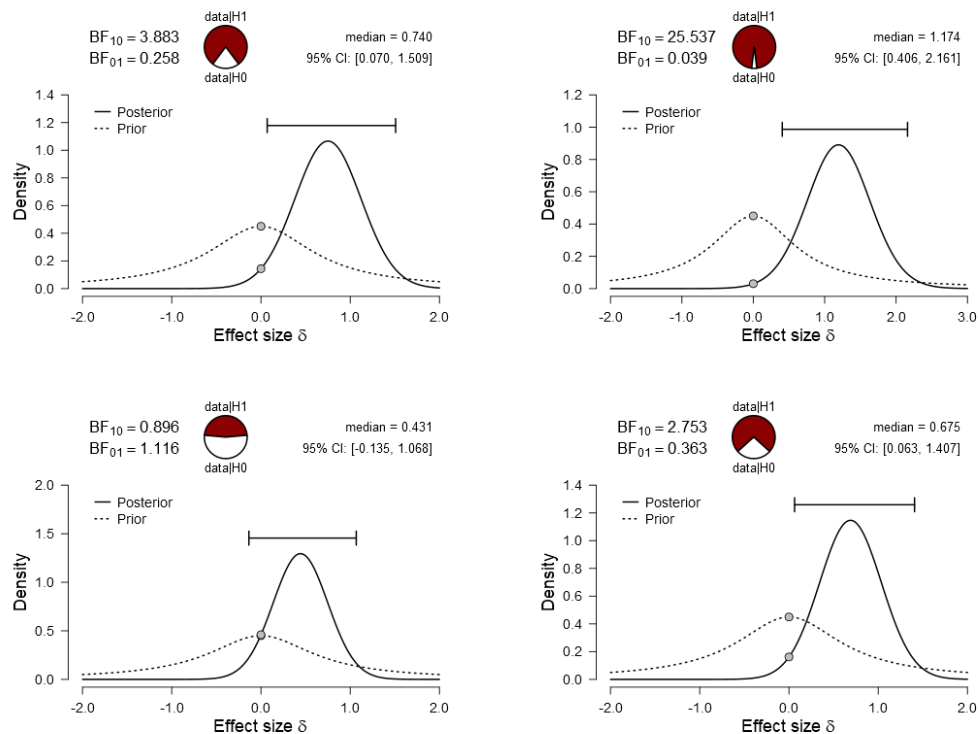


FIGURA 4.17: Como resultado de la evaluación de las Diferencias Relativas computadas entre las creencias y las elecciones de los participantes A en el Subjuego 2 cuando sus creencias no eran multiplicadas por p , se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

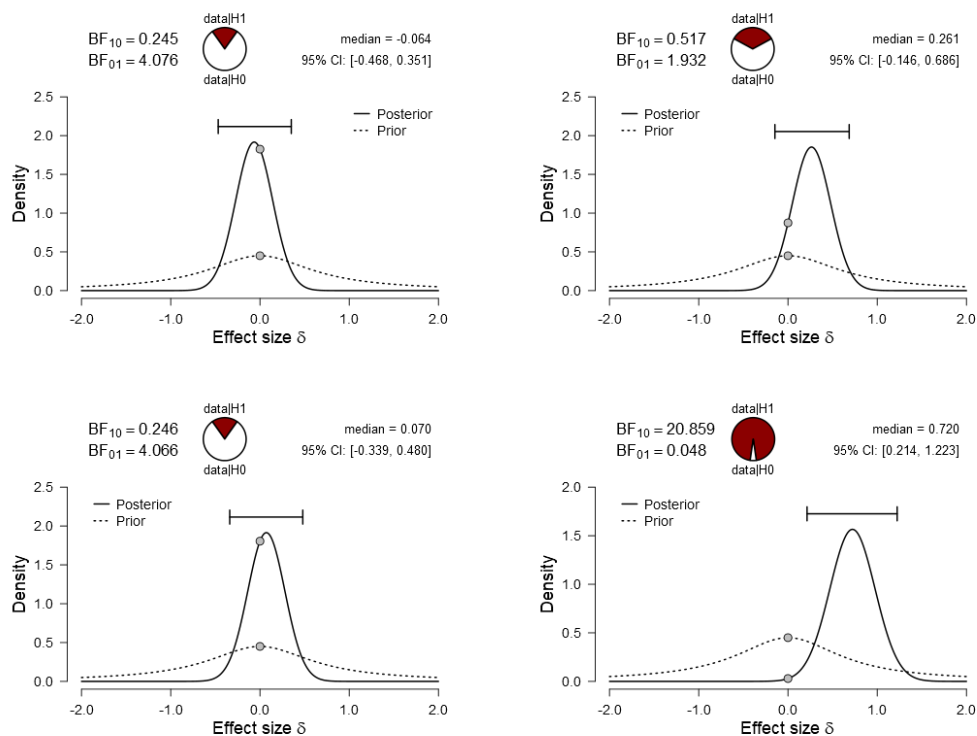


FIGURA 4.18: Como resultado de la evaluación de las Diferencias Relativas computadas entre las creencias y las elecciones de los participantes D y E en el Subjuego 2, sin multiplicar sus creencias por p , se presenta la comparación entre las densidades de probabilidad de las distribuciones prior (hipótesis nula) y posterior (hipótesis alterna) computadas por cada periodo, en el punto de no diferencia ($\delta = 0$).

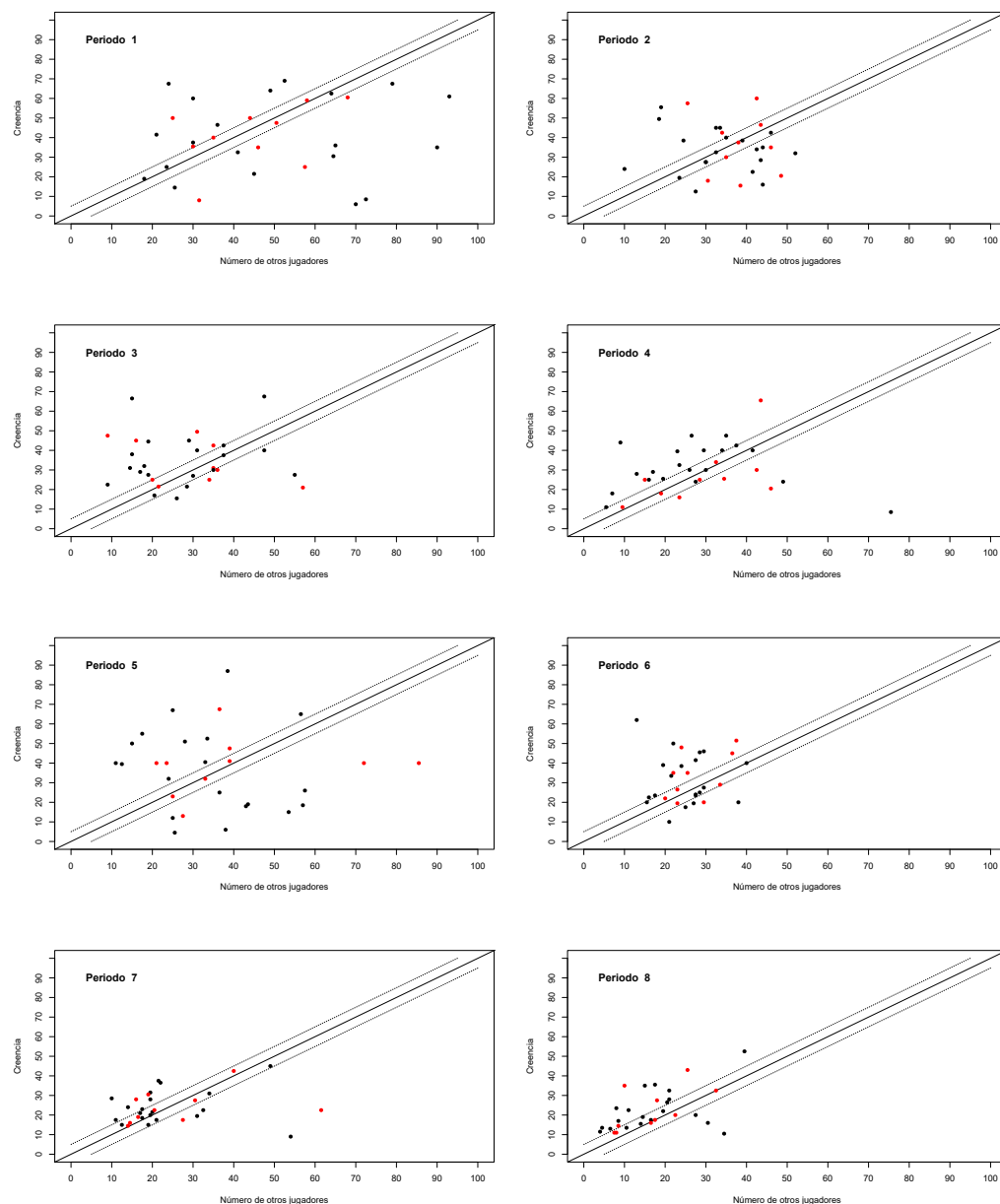


FIGURA 4.19: Contraste entre el promedio de las elecciones registradas en cada periodo por los otros jugadores y el promedio de las creencias registradas por cada jugador. Se señala en rojo los datos obtenidos por los participantes A. Los puntos cercanos a la línea de identidad indican un mayor grado de consistencia. Tanto las creencias como las elecciones observadas son multiplicadas por p .

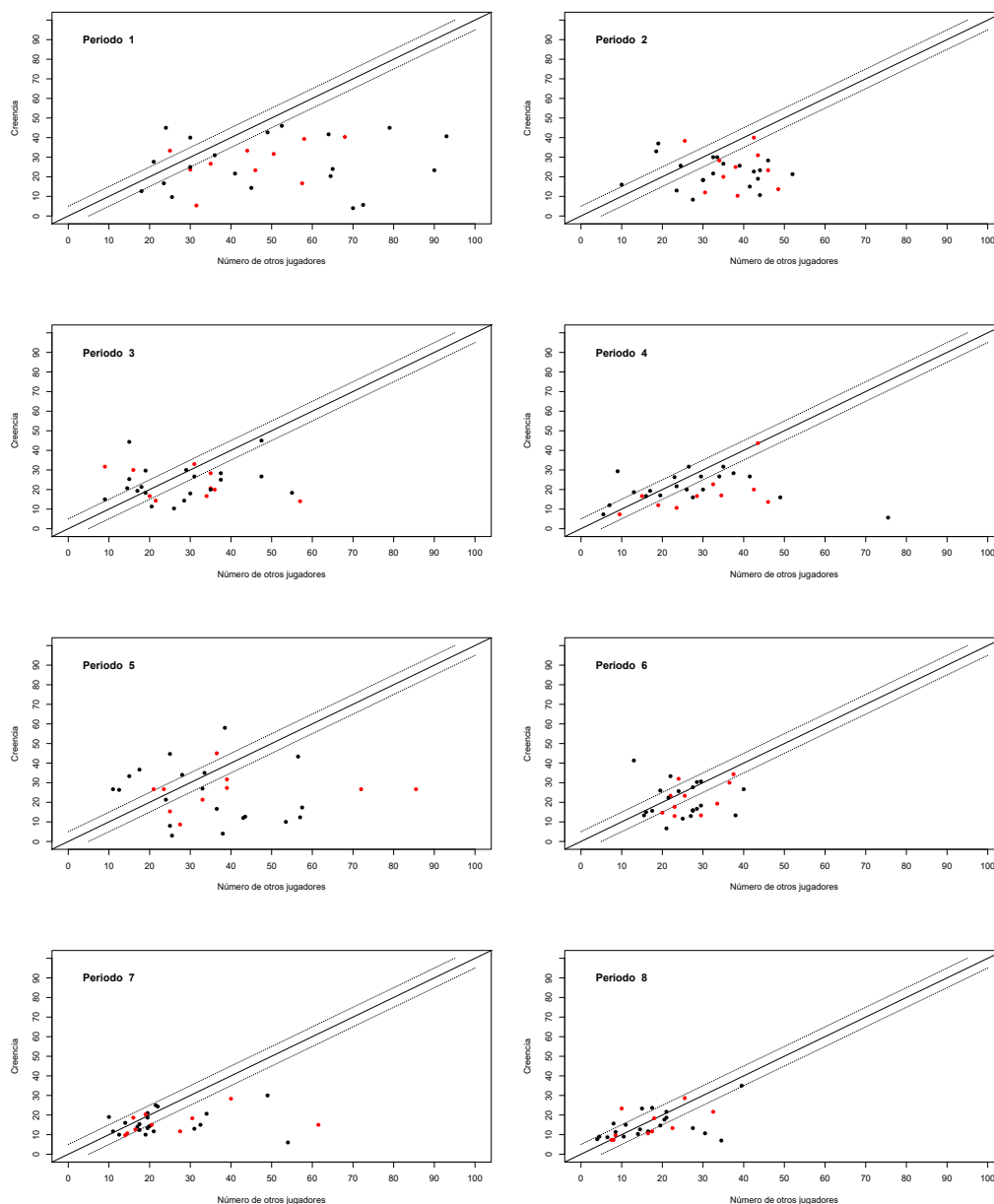


FIGURA 4.20: Contraste entre el promedio de las elecciones registradas en cada periodo por los otros jugadores y el promedio de las creencias registradas por cada jugador. Se señala en rojo los datos obtenidos por los participantes A. Los puntos cercanos a la línea de identidad indican un mayor grado de consistencia. No se incluye la multiplicación por p

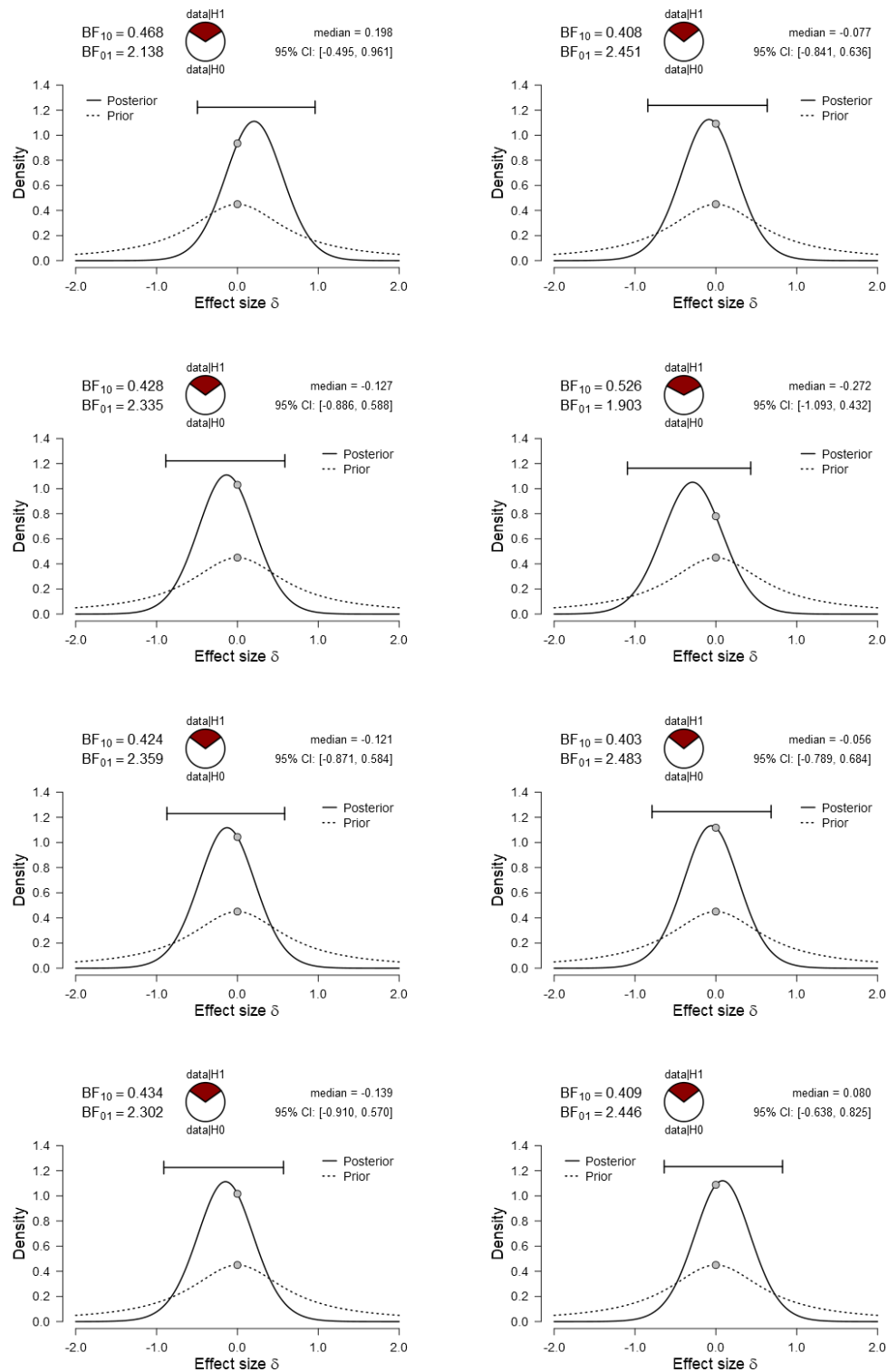


FIGURA 4.21: Como resultado de las pruebas *t* de muestras independientes que comparan los aciertos conseguidos por los participantes A y no A en sus predicciones, se presenta la relación entre las distribuciones prior y posterior computadas por cada periodo en el punto de no diferencias ($\delta = 0$)

Capítulo 5

Discusión

El experimento realizado aporta evidencia acerca de la relación que existe entre la elección de un número en el juego de p-beauty contest y el cómputo de las creencias que se tienen sobre las elecciones de los demás jugadores, mediante la incorporación de una variación del método para explicitar creencias propuesto por Lahav (2015) en un procedimiento que aprovecha los efectos de Reset reportados por Slonim (2005) para evaluar el posible impacto que tiene la experiencia sobre la consistencia con que las elecciones de los jugadores reflejan sus creencias sobre los demás participantes en el juego.

Para evaluar la diferencia entre las elecciones de los jugadores y sus creencias sobre las tiradas de sus oponentes, se utilizaron dos métodos diferentes. Primero, se tomó en cuenta la diferencia normalizada entre las creencias y las elecciones al ponderar esta por el promedio de los números elegidos por todos los jugadores en cada periodo; y después, se computó la diferencia relativa que toma como factor de ponderación el valor intermedio entre las creencias y las elecciones. Ambos métodos buscan compensar la tendencia que presentan las elecciones de los jugadores a converger en un equilibrio cercano a 0 cuando el juego se repite a lo largo de varios periodos. La diferencia sustancial entre ambos, es que la diferencia normalizada depende de la elección promedio registrada por todos los jugadores en el periodo a evaluar y la diferencia relativa se calcula únicamente a partir de la creencia y elección del jugador en cuestión.

Si bien estas dos métodos llevaron al cálculo de distintos valores por cada jugador en cada periodo, la relación entre estos se presentó de la misma forma: Los jugadores presentan inconsistencias entre sus creencias y sus elecciones cuando

no tienen experiencia, tal y como se observó en los primeros periodos jugados por los participantes sin experiencia en el subjuego 1 (participantes A, B y C) y en el subjuego 2 (participantes D y E). Dichas inconsistencias se reducen conforme los participantes adquieren experiencia, hacia el final del primer subjuego, y se mantienen a lo largo de los cuatro periodos que conforman el segundo subjuego para los participantes A, que juegan con los participantes sin experiencia D y E.

En promedio, las elecciones reales de los jugadores se situaron por encima del número objetivo hipotético, computado de acuerdo a las creencias registradas en cada periodo, y en cambio, se mantuvieron por debajo del promedio de sus creencias.

Para evaluar la posibilidad de que las inconsistencias observadas se debieran a que los jugadores no estuvieran tomando en cuenta que el promedio de sus creencias debía multiplicarse por p al momento de elegir su número, se incluyeron variaciones en el cálculo de las diferencias normalizadas y relativas que omitían la multiplicación por p . Con ello se observó que la elección de los participantes era más consistente con el promedio de sus creencias (sin incluir la multiplicación por p) en los primeros periodos, pero conforme adquirieron experiencia entre periodos, sus elecciones se fueron acercando más a la del número objetivo estimado de acuerdo a sus creencias (en los últimos periodos del primer subjuego jugado por cada participante y durante todo el segundo subjuego, en el caso de los participante A). Este resultado indica que los participantes aprenden a incluir la multiplicación por p conforme adquieren experiencia en el juego. De cualquier forma, no es posible determinar si los participantes incorporan el cálculo explícitamente, o simplemente aprenden de forma intuitiva a elegir números cada vez más pequeños, por debajo del promedio de sus creencias.

El inicio del segundo subjuego estuvo marcado por la introducción de dos nuevos jugadores (D y E) que reemplazaron a dos de los jugadores participantes en el subjuego 1 (B y C), siendo que uno de los jugadores de dicho subjuego permaneció durante cuatro periodos más. Con esta manipulación experimental se replicó exitosamente el efecto de Reset reportado por Slonim (2005), permitien-

do evaluar la consistencia entre las elecciones y las creencias de los participantes A como una función de su experiencia, sin la influencia del efecto de suelo. Los resultados obtenidos a este respecto confirman la importancia que tiene la experiencia de los participantes sobre su desempeño en el juego de p-beauty contest, en términos de la consistencia entre los números elegidos y los números que predijeron que tirarían los demás jugadores.

El experimento que se realizó para conducir este estudio se llevó a cabo en 10 sesiones experimentales compuestas de 8 periodos de juego. En cada periodo 3 jugadores elegían números e intentaban adivinar los números de los otros jugadores. Considerando la cantidad de datos generados con este diseño, así como teniendo en cuenta diseños utilizados en la literatura (Kocher et al., 2007, Slonim, 2005), se considera que 10 sesiones fueron suficientes para obtener resultados válidos, pero que replicar el experimento con una muestra más grande podría incrementar la robustez de los hallazgos reportados.

Capítulo 6

Conclusión

Se realizó un experimento de p-beauty contest repetido con una variación del método de explicitación de creencias presentado por Lahav (2015). Los resultados encontrados en el presente estudio confirman el hallazgo principal reportado por Lahav acerca de la consistencia entre las elecciones de los participantes en cada periodo y sus creencias sobre las tiradas de sus oponentes: en un comienzo los participantes eligen números poco consistentes con las creencias reportadas, pero conforme van adquiriendo experiencia al participar en más periodos, se van acercando más. El presente trabajo aporta evidencia a favor de la relación experiencia-consistencia, descartando al efecto de suelo como la causa principal de la reducción en las diferencias registradas entre elecciones y creencias en cada periodo, mediante la incorporación de un segundo subjuego donde participantes que adquirieron experiencia durante el primer subjuego fueron enfrentados a nuevos oponentes, generando un efecto de reset que llevara a los participantes con experiencia a elegir números más grandes pero consistentes con sus creencias.

Además, los resultados encontrados muestran que los participantes no sólo se vuelven más consistentes conforme adquieren experiencia, sino que también comienzan a elegir números que caen por debajo del promedio de sus creencias (lo cual podría sugerir que aprenden a incorporar la multiplicación por p al elegir el número con el que competirán en cada periodo).

Todos los juegos de p-beauty contest que se llevaron a cabo en este trabajo de tesis se realizaron con grupos de únicamente tres participantes. Esta decisión obedeció principalmente a mantener la misma cantidad de jugadores con

y sin experiencia que reporta Slonim (2005), además de facilitar la explicitación de creencias específicas al contar con grupos pequeños. Se realizaron 10 sesiones experimentales de acuerdo con diseños experimentales similares presentes en la literatura (Ho, Camerer Weigelt, 1998, Kocher, Sutter Wakolbinger, 2007, Slonim, 2005). Adicionalmente, la repetición del juego por 4 periodos en cada subjuego permitió detectar aquellos participantes que no tienen una estructura discernible en sus elecciones.

La principal fortaleza del presente diseño experimental radica en la inclusión de un segundo subjuego donde nuevos jugadores entran en el juego, permitiendo con ello:

- Un análisis de medidas repetidas de las creencias y elecciones del participante A en dos escenarios diferentes (un Subjuego sin experiencia y un Subjuego con experiencia).
- Un mayor tamaño de la muestra de jugadores sin experiencia.
- Poner a prueba, a partir del efecto de reset reportado en las elecciones de jugadores con experiencia al enfrentarse a jugadores sin experiencia, la posibilidad de que los resultados reportados por Lahav (2015), sean un co-producto de la tendencia hacia el equilibrio que se ha observado de manera consistente en juegos repetidos de p-beauty contest.
- Al enfrentar al participante A con distintos grupos de jugadores, se garantiza que el incremento observado en la consistencia entre sus creencias y elecciones no es resultado de la retroalimentación específica que obtiene sobre la conducta de sus oponentes.

En conjunto, estos elementos permiten atribuir los resultados obtenidos a la experiencia, lo que tiene implicaciones directas sobre la pregunta de investigación planteada en un inicio, y en general, la literatura en teoría de juegos y

razonamiento iterado.

Lahav (2015) concluye que parece haber poca evidencia en favor de la consistencia entre creencias y elecciones, y que el nivel de sofisticación de las personas no puede estimarse a partir de sus creencias pues estas no parecen reflejarse en sus elecciones. El autor propone como posibles explicaciones una 'anomalía conductual' o el malentendido de las instrucciones, reconociendo la necesidad de más investigaciones. Por su parte, Costa-Gomes y Weizsäcker (2008), que estudiaron juegos con creencias hechas explícitas en los que los jugadores no recibían retroalimentación y quienes también encontraron que las elecciones de los participantes no eran consistentes con sus creencias, concluyen que no es posible asumir que las elecciones de las personas son dirigidas por las creencias que tienen sobre los otros jugadores, por lo menos en juegos complejos con ausencia de retroalimentación, y que la construcción de expectativas sobre la conducta de otros jugadores implica un proceso de aprendizaje a partir de la retroalimentación recibida en interacciones repetidas.

Debido al diseño del experimento de Lahav (2015), resulta difícil asumir que los jugadores tienen expectativas individuales sobre las elecciones de cada jugador, al tener que enfrentarse a un amplio número de ellos (veinte jugadores por sesión experimental). Por su parte, aunque en el experimento de Costa-Gomes y Weizsäcker (2008), los juegos eran de solo dos participantes, se evitó la retroalimentación, lo que no permitió que se crearan expectativas sobre el otro jugador. En contraste, en el presente experimento se trabajó con grupos pequeños de jugadores y se permitió la retroalimentación al final de cada periodo de juego, dando oportunidad a que se crearan expectativas sobre los otros jugadores. Sin embargo, la posibilidad de formarse expectativas acerca del comportamiento de los otros jugadores no es suficiente para explicar el aumento en la consistencia entre creencias y elecciones por parte de los participantes, pues en el segundo subjuego donde se introducen nuevos jugadores el incremento observado en la consistencia de las elecciones se mantiene, efecto que sólo puede atribuirse a su experiencia, ya que de acuerdo con lo previamente sugerido por Slonim (2005), adquirir experiencia acelera la tasa de aprendizaje en un juego, o en juegos similares, tal y como se vio reflejado en sus resultados, donde luego del primer

subjuego, la convergencia al equilibrio fue más rápida y los jugadores con experiencia mantuvieron una mayor ventaja en el juego, aun jugando con personas nuevas. En palabras de Ho, Camerer Weigelt (1998), los jugadores “aprenden a aprender”.

El efecto de la experiencia se distingue del efecto de la formación de expectativas sobre los otros jugadores y sus estrategias, porque los cambios en la consistencia se mantienen aun cuando los demás jugadores cambian. De esta forma, aún si las expectativas que se tiene sobre los nuevos jugadores fueran similares a las que se tenían sobre los jugadores anteriores (en el periodo equivalente), y resultaran ser erróneas respecto de las elecciones observadas, esto no compromete a la consistencia que se registra entre sus propias creencias y elecciones, porque se ha adquirido experiencia sobre el juego en sí mismo, sus reglas y la forma en que las personas las aprenden. Esto se alinea con las conclusiones planteadas por Costa-Gomes Weizsäcker (2008), acerca de la influencia que tiene la situación (el contexto, el juego) sobre la formación simultánea de creencias y elecciones, que reemplaza el supuesto de que la situación influye primeramente en las creencias, que posteriormente permean las elecciones.

En su revisión de modelos de no-Equilibrio en Teoría de Juegos, Crawford, Costa-Gomes e Iriberri (2013) concluyen que aunque existe evidencia experimental de que las personas se desvían del equilibrio de Nash, lo hacen de forma sistemática, a partir de un componente estructural que se puede modelar; los modelos de nivel-k son los mejores para predecir la conducta en muchos juegos con este componente, pues permiten predecir dichas desviaciones, detectando sus causas y anticipando su frecuencia, proporcionando respuestas a distintos problemas empíricos. El presente experimento replicó el efecto de reset observado por Slonim (2005), siendo dicho resultado una evidencia importante a favor de los modelos de nivel-k (que establecen que las personas anclan sus creencias de manera intuitiva de acuerdo con el juego, ajustándolas mediante respuestas óptimas iteradas), en tanto que tras la introducción de nuevos jugadores (sin experiencia; en niveles cognitivos 0 ó 1), los jugadores experimentados se alejan del equilibrio de Nash para responder de forma óptima (a la luz de sus nuevos contrincantes; eligiendo números que corresponderían con niveles

cognitivos 1 ó 2). En otras palabras, este tipo de resultados confirman la idea de que los jugadores experimentados son racionales, pero pueden tener creencias sobre los otros jugadores que impliquen asumir que estos no lo son.

Apéndice A

Instrucciones

Instrucciones para todos los participantes al inicio de la sesión

Hola a todos y gracias por venir. Este es un experimento sobre toma de decisiones y no queremos que influyan sobre las decisiones de los demás. Por lo tanto, no está permitido que hablen o se comuniquen entre ustedes.

Si tienen alguna duda levanten la mano e iré a su lugar para resolverla.

En este experimento, van a participar en un juego que se repite cuatro veces. Llamaremos a cada repetición del juego un *Periodo*. En el juego sólo participan tres personas. Mediante un sorteo, elegiremos a tres de ustedes para que jueguen primero, mientras los otros dos esperarán en otra aula. Cuando las primeras tres personas terminen de jugar por cuatro periodos, se elegirá a una de estas tres personas para que juegue junto con las dos personas que estaban esperando. Cuando este segundo grupo termine de jugar cuatro veces, terminará el experimento.

En cada periodo podrán ganar puntos de juego. Por el hecho de participar en este experimento, todos tienen medio punto sobre su examen parcial, y al final de los cuatro periodos, el participante que haya acumulado más puntos de juego ganará otro medio punto sobre su examen parcial, por lo que pueden ganar hasta un punto completo sobre su examen. En caso de empates, el medio punto se dividirá entre los ganadores. En cada periodo, un jugador puede ganar hasta 8 puntos de juego, pero esto dependerá del desempeño de todos los participantes.

(Entregar cuatro (4) formatos de respuesta a cada participante. Cada participante debe recibir formatos con los números del 1 al 4 y con su clave personal.)

Le estoy entregando cuatro formatos de respuesta a cada uno. Noten que los formatos que cada uno recibió tienen una combinación de números y letras en la celda llamada *Clave*. Esta clave es única para cada uno de ustedes y la usaremos para identificarlos.

Los formatos también contienen una celda llamada *Periodo* que contiene un número del 1 al 4. En cada periodo de juego, usarán únicamente el formato de respuesta que corresponda al periodo que se está jugando, es decir, el formato que dice Periodo 1 en el primer juego, el formato que dice Periodo 2 en el segundo juego, y así sucesivamente.

¿Cómo se juega? En cada periodo, cada jugador debe elegir un número entero entre el 0 y el 100. Deberán escribir su número en el formato de respuesta, en la celda llamada *Mi Número Elegido*. No dejen que los otros participantes conozcan el número que eligieron. El ganador de ese periodo será el participante cuyo número elegido esté lo más cercano posible al *Número Objetivo* de ese periodo. ¿Cuál es el Número Objetivo? El Número Objetivo se calcula de la siguiente manera:

Se obtiene el promedio de los números elegidos por cada jugador, es decir, se suman los tres números y se dividen entre 3. Después, este número promedio se multiplica por $\frac{2}{3}$, es decir se multiplica por 2 y se divide entre 3. El resultado es el Número Objetivo.

En otras palabras, para ganar deberán elegir un número que crean que estará lo más cerca posible al promedio de los números elegidos por todos los participantes, multiplicado por $\frac{2}{3}$. El ganador obtendrá 6 puntos de juego. Si dos o los tres de ustedes eligen números igual de cercanos al Número Objetivo, los 6 puntos de juego se dividirán equitativamente entre todos los participantes ganadores.

Como verán, hay una celda más en su formato de respuesta, llamada *Números de los otros Jugadores* que contiene espacio para que escriban dos números. Lo que deben hacer en cada periodo después de elegir su propio número es escribir en esta celda dos números enteros que ustedes crean que estarán lo más cerca posible a los números que van a elegir los otros dos participantes. En otras palabras, deben intentar adivinar qué números elegirán los otros jugadores.

Ganarán 1 punto de juego si uno de los otros participantes elige para el juego un número hasta 5 números por arriba o por debajo de uno de los números que escribieron en la celda de *Números de los otros Jugadores*. Ganarán otro punto de juego si el otro participante elige un número hasta 5 números por arriba o por debajo de su segundo número escrito en la celda de *Números de los otros Jugadores*.

Es decir, sólo ganarán dos puntos si sus dos números se acercan a los dos números de los otros jugadores.

(Dibujar en el pizarrón: $X \pm 5$ $Y \pm 5$)

Ustedes eligen dos *Números de los otros Jugadores*, X y Y . Si ambos jugadores eligen un número que está dentro del rango de $X \pm 5$, pero ninguno de los dos entra en el rango de $Y \pm 5$, entonces sólo ganarán un punto. Para que sea posible ganar el segundo punto, el número de uno de los otros jugadores debe caer dentro de $X \pm 5$ y el otro dentro del rango de $Y \pm 5$. Si creen que los otros jugadores van a elegir números muy cercanos, es válido elegir números muy cercanos o incluso iguales. Recuerden, los números que elijan para la celda *Números de los otros Jugadores* NO influyen en el valor del Número Objetivo ni influyen en determinar qué jugador gana en cualquier periodo. Los números de esta celda únicamente sirven para ganar puntos ADICIONALES si adivinan los números que los otros participantes escribieron en la celda *Mi Número Elegido*.

Una vez que hayan llenado todas las celdas del formato de respuesta para el periodo actual, coloquen su formato boca abajo y esperen a que los otros participantes terminen y hagan lo mismo. Una vez que todos hayan terminado, pasará

a sus lugares a recoger sus formatos de respuesta para este periodo. Escribiré en el pizarrón los tres números elegidos sin indicar a quién corresponde cada número, y usaré los números elegidos para calcular el promedio, que escribiré en el pizarrón. Multiplicaré el promedio por $\frac{2}{3}$ y escribiré este número, que será el Número Objetivo, en el pizarrón.

Revisaré cuál de los tres números elegidos es el más cercano al Número Objetivo, y si los números que escribieron en la celda *Números de los otros Jugadores* acertaron a los números elegidos por sus oponentes. En función a esto registraré cuántos puntos obtuvo cada quien en este periodo y se los haré saber de forma individual.

Borraré los números escritos en el pizarrón y comenzaremos el siguiente periodo, repitiendo el proceso.

(Las personas con las claves A, B, y C juegan primero).

Instrucciones para los participantes del subjuego 2

Les repito brevemente las instrucciones. Van a repetir un juego cuatro veces. En cada repetición, o periodo, van a elegir un número entero entre 0 y 100 que escribirán en la celda *Mi Número Elegido*. Ganará 6 puntos de juego el participante que haya elegido el número más cercano al promedio de los números elegidos por los todos participantes, multiplicado por $\frac{2}{3}$.

En la celda *Números de los otros Jugadores* deben escribir dos números enteros que crean que estarán lo más cerca posible de los números elegidos por los otros participantes. Ganaran 1 punto de juego por cada número que hayan escrito en esta celda que esté 5 números por arriba o por debajo de un número elegido por los otros jugadores.

Recuerden que uno de ustedes ya ha jugado este juego, mientras que dos de ustedes nunca lo han jugado.

Apéndice B

Formato de registro

A continuación se presenta el Formato de registro de respuestas entregado a cada uno de los jugadores de cada subjuego:

| | | |
|---------|-------------------|-----------------------------------|
| Clave | Mi Número Elegido | Números de los otros jugadores |
| S1A | | |
| Periodo | | |
| 1 | _____ | _____ y _____ |

Apéndice C

Análisis de datos con pruebas frecuentistas

CUADRO C.1: Prueba t de una muestra: Diferencias normalizadas en el Subjuego 1

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 1 | -3.426 | 29 | 0.002 | -0.366 |
| Periodo 2 | -3.677 | 29 | <.001 | -0.342 |
| Periodo 3 | -0.895 | 29 | 0.378 | -0.097 |
| Periodo 4 | -1.651 | 29 | 0.110 | -0.147 |

CUADRO C.2: Prueba t de una muestra: Diferencias normalizadas en el Subjuego 1 sin la multiplicación por p

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 1 | -0.368 | 29 | 0.715 | -0.049 |
| Periodo 2 | -0.123 | 29 | 0.903 | -0.012 |
| Periodo 3 | -2.674 | 29 | 0.012 | 0.355 |
| Periodo 4 | -2.299 | 29 | 0.029 | 0.280 |

CUADRO C.3: Prueba t de una muestra: Diferencias relativas en el Subjuego 1

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 1 | -3.986 | 29 | <.001 | -0.457 |
| Periodo 2 | -3.244 | 29 | 0.003 | -0.328 |
| Periodo 3 | -0.456 | 29 | 0.652 | -0.052 |
| Periodo 4 | -2.302 | 29 | 0.029 | -0.212 |

CUADRO C.4: Prueba t de una muestra: Diferencias relativas en el Subjuego 1 sin la multiplicación por p

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 1 | -0.839 | 29 | 0.408 | -0.101 |
| Periodo 2 | 0.424 | 29 | 0.675 | 0.044 |
| Periodo 3 | 3.052 | 29 | 0.005 | 0.315 |
| Periodo 4 | 1.800 | 29 | 0.082 | 0.167 |

CUADRO C.5: Prueba t de una muestra: Diferencias Normalizadas en el Subjuego 2 (Participante A)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | -0.413 | 9 | 0.689 | -0.049 |
| Periodo 6 | -0.138 | 9 | 0.893 | -0.011 |
| Periodo 7 | -1.083 | 9 | 0.307 | -0.133 |
| Periodo 8 | -0.154 | 9 | 0.881 | -0.028 |

CUADRO C.6: Prueba t de una muestra: Diferencias Normalizadas en el Subjuego 2 (Participante D y E)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | -2.593 | 19 | 0.018 | -0.349 |
| Periodo 6 | -3.210 | 19 | 0.005 | -0.243 |
| Periodo 7 | -3.457 | 19 | 0.003 | -0.307 |
| Periodo 8 | -0.401 | 19 | 0.693 | -0.047 |

CUADRO C.7: Prueba t de una muestra: Diferencias Normalizadas en el Subjuego 2 sin la multiplicación por p (Participante A)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | 2.306 | 9 | 0.047 | 0.333 |
| Periodo 6 | 4.250 | 9 | 0.002 | 0.414 |
| Periodo 7 | 1.565 | 9 | 0.152 | 0.238 |
| Periodo 8 | 2.160 | 9 | 0.059 | 0.438 |

CUADRO C.8: Prueba t de una muestra: Diferencias Normalizadas en el Subjuego 2 sin la multiplicación por p (Participante D y E)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | 0.142 | 19 | 0.889 | 0.024 |
| Periodo 6 | 1.617 | 19 | 0.122 | 0.170 |
| Periodo 7 | 0.660 | 19 | 0.517 | 0.071 |
| Periodo 8 | 3.144 | 19 | 0.005 | 0.440 |

CUADRO C.9: Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 (Participante A)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | -0.336 | 9 | 0.745 | -0.042 |
| Periodo 6 | -0.072 | 9 | 0.944 | -0.007 |
| Periodo 7 | -1.193 | 9 | 0.264 | -0.162 |
| Periodo 8 | 0.230 | 9 | 0.823 | 0.038 |

CUADRO C.10: Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 (Participantes D y E)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | -2.677 | 19 | 0.015 | -0.407 |
| Periodo 6 | -3.646 | 19 | 0.002 | -0.283 |
| Periodo 7 | -3.334 | 19 | 0.003 | -0.341 |
| Periodo 8 | -0.067 | 19 | 0.947 | -0.007 |

CUADRO C.11: Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 sin la multiplicación por p (Participante A)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | 2.886 | 9 | 0.018 | 0.346 |
| Periodo 6 | 4.384 | 9 | 0.002 | 0.386 |
| Periodo 7 | 1.688 | 9 | 0.126 | 0.225 |
| Periodo 8 | 2.617 | 9 | 0.028 | 0.413 |

CUADRO C.12: Prueba t de una muestra: Diferencias Relativas en el Subjuego 2 sin la multiplicación por p (Participantes D y E)

| | t | df | p | Diferencia promedio |
|-----------|----------|-----------|----------|----------------------------|
| Periodo 5 | -0.347 | 19 | 0.732 | -0.056 |
| Periodo 6 | 1.361 | 19 | 0.190 | 0.108 |
| Periodo 7 | 0.355 | 19 | 0.727 | 0.039 |
| Periodo 8 | 3.599 | 19 | 0.002 | 0.372 |

Referencias

- Agranov, M., Potamites, E., Schotter, A., y Tergiman, C. (2012). Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior*, 75(2), 449-463.
- Camerer, C., Ho, T., y Chong, J. K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3), 861-898.
- Costa-Gomes, M. A., y Weizsäcker, G. (2008). Stated beliefs and play in normal-form games. *The Review of Economic Studies*, 75(3), 729-762.
- Crawford, V. P., Costa-Gomes, M. A., y Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence and applications. *Journal of Economic Literature*, 51(1), 5-62.
- Ho, T.-H., Camerer, C., y Weigelt, K. (1988). Iterated dominance and iterated best response in experimental p-beauty contests. *The American Economic Review*, 88(4), 947-969.
- Ho, T. H., Camerer, C., y Weigelt, K. (1998). Iterated dominance and iterated best response in experimental p-beauty contests. *The American Economic Review*, 88(4), 947-969.
- Keynes, J. (1956). *The general theory of employment, interest and money*.
- Kocher, M., Sutter, M., y Wakolbinger, F. (2007). The impact of naïve advice and observational learning in beauty-contest games. *Tinbergen Institute Discussion Paper*, 1-32.
- Lahav, Y. (2015). Eliciting beliefs in beauty contest experiments. *Economic Letters*, 137, 45-49.
- Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, 1312-1326.
- Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, 54(2), 286-295.
- Rosenthal, R. W. (1981). Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*, 25(1), 92-100.
- Slonim, R. L. (2005). Competing against experienced and inexperienced players.

Experimental Economics, 8(1), 55-75.

Stahl, D. O., y Wilson, P. W. (1995). On players models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(2), 218-254.