



UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
DEPARTAMENTO DE COMPUTACIÓN

Mercados de predicción y profecías autocumplidas

Tesis de Licenciatura en Ciencias de la Computación

Tomás Schitter

Director: Sergio Abriola
Buenos Aires, 2023

MERCADOS DE PREDICCIÓN Y PROFECÍAS AUTOCUMPLIDAS

En el contexto del estudio de pronosticadores que realizan predicciones sobre la probabilidad de que ciertos eventos ocurran, un problema ampliamente estudiado es el de construir reglas para puntuar estas predicciones en base a los resultados observados. Estas reglas se llaman *scoring rules*, y justamente se encargan de asignar un puntaje a una predicción dada, basado en el evento que finalmente ocurre. En la práctica, existe una subclase de scoring rules muy utilizada, llamadas *strictly proper scoring rules*, cuya característica principal es que, si el pronosticador intenta maximizar el puntaje obtenido, estas reglas lo incentivan a realizar predicciones cuidadosas y honestas sobre la probabilidad asociada al evento que está prediciendo.

En principio, podría considerarse que las predicciones que realiza el pronosticador están separadas del mundo, es decir, que no influyen en este. Sin embargo, en la práctica la predicción es observada y puede influir en el evento sobre el cual predice. De esta forma, el pronosticador podría intentar manipular el mundo para aumentar la probabilidad de que se cumpla su propia predicción, y así obtener un puntaje más alto. Esto es un comportamiento sumamente indeseado que se conoce como *profecía autocumplida*, y que ha sido estudiado también dentro de la subárea de la inteligencia artificial que se encarga del análisis del comportamiento de *oráculos*, cuya única funcionalidad consiste en responder preguntas sobre eventos concretos.

Por otro lado, los *mercados de predicción* son mercados abiertos donde se intercambian activos cuyo valor depende del resultado de algún evento particular. Una utilidad de estos mercados es entonces recolectar predicciones de muchas fuentes individuales de forma de generar una única predicción colectiva. Se ha encontrado que generar predicciones colectivas a partir de predicciones individuales resulta muchas veces más preciso que otras instituciones con el mismo propósito.

En esta tesis estudiaremos entonces el problema de las *profecías autocumplidas* en el contexto de *mercados de predicción*. Esto es, el fenómeno que se da cuando la predicción final obtenida en un mercado de predicción puede influir en el mundo sobre el cual se predice. Para eso, tomaremos distintas formalizaciones simplificadas de mercado de predicción de forma de poder realizar un análisis formal de ellos utilizando herramientas de la *teoría de juegos*.

Palabras claves: Pronosticadores, Scoring rules, Mercados de predicción, Profecías autocumplidas, Teoría de juegos.

PREDICTION MARKETS AND SELF-FULFILLING PROPHECIES

When studying forecasters that make predictions on the probability associated to certain events, a widely studied problem is that of building rules to grade the predictions based on the observed outcomes. These rules are called *scoring rules*, and they assign a score to a given prediction, based on the outcome of the event that is being predicted. In practice, there is a subclass of scoring rules which are commonly used, called *strictly proper scoring rules*, whose main property is that, if the forecaster aims to maximize its score, these rules encourage it to make careful and honest predictions about the probability associated to the event which it is predicting.

Initially, we could consider that the forecaster's predictions are separated from the world, which means that they cannot influence it. However, in practice, its predictions are observed in the real world and can actually influence the very event it is trying to predict. Thus, the forecaster could try to manipulate the world to make its own predictions more likely and therefore receive a higher score. This is a highly undesired behavior, known as a *self-fulfilling prophecy*, that has also been studied within the field of artificial intelligence that analyses the behaviour of *oracles*, whose only functionality is that of answering questions about specific events.

On the other hand, a *prediction market* is an open market where assets are exchanged, and where the value of those assets depends on the outcome of a certain event. Therefore, these markets can be used to collect predictions from many different sources, and then combine them to generate a single collective prediction. Generating collective predictions based on individual ones has been found to be more accurate than other institutions with the same purpose.

In this thesis, we will study the problem of *self-fulfilling prophecies* in the context of *prediction markets*. That is, the phenomenon that arises when the final prediction that is output by the market can influence the world. To do that, we will consider several simplified models for prediction markets, so as to analyze them formally using concepts from *game theory*.

Keywords: Forecasters, Scoring rules, Prediction markets, Self-fulfilling prophecies, Game theory.

Índice general

1.. Introducción	1
2.. Preliminares	5
2.1. Scoring rules	5
2.1.1. Strictly proper scoring rules	6
2.1.2. Funciones convexas y caracterización de strictly proper scoring rules	7
2.1.3. Ejemplos de strictly proper scoring rules	9
2.1.4. Profecías autocumplidas	9
2.2. Mercados de predicción	10
2.2.1. Nuestro formalismo	12
2.3. Conceptos de teoría de juegos	12
3.. Resultados	18
3.1. Predicciones simultáneas iteradas	19
3.2. Mercado de suma constante con predicciones únicas y simultáneas	23
3.3. Predicciones únicas y secuenciales	25
3.3.1. Resultados analíticos	25
3.3.2. Resultados experimentales	26
4.. Conclusiones y trabajo futuro	35
5.. Apéndice	38
A. Implementación	38
B. Experimento 1: la importancia del orden	38
C. Experimento 2: la importancia del orden al aumentar la influencia de las predicciones	39
D. Experimento 3: la importancia del peso de los jugadores	40
E. Datos completos regla de Brier	40
F. Datos completos regla logarítmica	47

1. INTRODUCCIÓN

Un problema importante en la computación ampliamente estudiado en la actualidad es el de construir *pronosticadores* que realizan predicciones sobre ciertos eventos. Un caso clásico de estudio en el área es el de predicción del clima [1, 2, 3]. Inevitablemente, estas predicciones llevan asociadas un cierto grado de incertidumbre, por lo que resulta natural que tengan un carácter probabilístico. Junto con el desarrollo de estos pronosticadores, surge el problema de especificar una forma de clasificar o puntuar la precisión de las predicciones realizadas, es decir, de verificar la correcta calibración del pronosticador. En este contexto, la *calibración* se refiere a la consistencia entre la predicción probabilística realizada por el pronosticador sobre un cierto evento, y el evento finalmente observado. Es con esta motivación que surgen las *scoring rules*, herramientas que nos permiten asignar un puntaje numérico a una predicción realizada por un pronosticador de acuerdo al evento que finalmente sucede, de forma de medir la calidad o la precisión de la predicción. En el caso de predicción del clima, podemos pensar en un pronosticador muy simple que predice la probabilidad de lluvia en una cierta ciudad. Si puntuamos las predicciones de dicho pronosticador usando una scoring rule, esperaríamos que, por ejemplo, si se predijo un 90 % de probabilidad de lluvia y efectivamente llueve, el puntaje asignado sea uno alto. Por el otro lado, si con esa misma predicción finalmente no llueve, esperamos que el puntaje asignado sea bajo.

Una característica deseable en una scoring rule es que, si el pronosticador intenta maximizar el puntaje obtenido, la scoring rule lo incentive a realizar predicciones cuidadosas y honestas, es decir, que lo incentive a predecir su verdadera creencia respecto a la distribución probabilística asociada al evento sobre el cual predice. Por ejemplo, volviendo al caso del pronosticador de la probabilidad de lluvia, no sería muy útil considerar una scoring rule que dé mayor puntaje cuanto más lejana esté la probabilidad predicha del hecho que finalmente se observa, ya que podría incentivar al pronosticador a predecir lo opuesto a su verdadera creencia. De la misma forma, no sería útil tomar una scoring rule constante, ya que la predicción realizada resultaría irrelevante a la hora de calcular su puntaje. Para formalizar este concepto de características deseables para las scoring rules es que surgen las *strictly proper scoring rules*. Éstas son scoring rules cuyo objetivo es justamente incentivar al pronosticador a predecir su verdadera creencia. La forma en que buscan conseguir esto es garantizando que, dada una predicción del pronosticador, su puntaje esperado sea máximo cuando dicha predicción coincida con la verdadera distribución de probabilidad asociada al evento sobre el cual predice. Así, si el pronosticador predice un 90 % de probabilidad de lluvia, su puntaje esperado será máximo si eso coincide con la verdadera probabilidad de lluvia. Las strictly proper scoring rules son ampliamente utilizadas en la práctica [4], y más adelante presentaremos algunas de ellas muy estudiadas y utilizadas con particular frecuencia [5, 1].

Veamos ahora otro ejemplo. Consideremos un pronosticador que realiza predicciones sobre el estado de la economía de cierto país, y cuyas predicciones son conocidas por la población general y tienen un cierto poder de influencia sobre sus acciones (por ejemplo, por reputación o confianza general en sus predicciones). En este caso, si el pronosticador intenta maximizar la probabilidad de que su predicción sea precisa, es posible que intente *manipular* el mundo a través de sus predicciones, es decir, que realice sus predicciones con-

siderando su efecto en el mundo sobre el cual predice. Por ejemplo, el pronosticador podría predecir la próxima caída de los bancos, lo cual influiría en la población, generando temor y potencialmente una corrida bancaria que podría finalmente ser la causa de la caída de los bancos. De esta forma, la predicción realizada por el pronosticador sería precisa, pero esto se debería a la propia influencia de la predicción en el mundo, y no necesariamente a una buena calibración del pronosticador.

Puede resultar entonces indeseable que un pronosticador pueda manipular el mundo sobre el cual predice. Un área en la cual este tema ha sido estudiado es en *inteligencia artificial*. Al construir pronosticadores utilizando técnicas de inteligencia artificial, que llamaremos *oráculos*, una preocupación importante es que un oráculo suficientemente avanzado pueda justamente intentar manipular el mundo con sus predicciones, de forma de volver sus propias predicciones más precisas o incluso de transformar al mundo para que se vuelva más sencillo de predecir. Sin embargo, se han estudiado algunas propuestas para contrarrestar o al menos paliar esta situación, de forma de poder desarrollar oráculos poderosos de forma más segura [6].

Por otro lado, los *mercados de predicción* son mercados abiertos donde se intercambian activos cuyo valor depende del resultado de cierto evento [7]. Los precios de los diferentes activos del mercado en un momento dado se pueden interpretar como la predicción del mercado sobre la probabilidad de dicho evento [8]. Así, los mercados de predicción son instrumentos sumamente útiles para recoger predicciones de diversas fuentes y realizar una única predicción basada en las predicciones individuales [9, 10]. Se han observado muy buenos resultados al utilizar predicciones colectivas y mercados de este tipo para predecir eventos, incluso mejor que otras instituciones con el mismo fin, lo cual vuelve particularmente interesante su estudio [11]. Los mercados de predicción pueden ser abordados desde muchos puntos de vista, y tienen conexiones con diversas áreas del conocimiento, como la economía [12], la computación [5, 10] o la matemática [13].

Si bien las reglas de los mercados de predicción pueden ser difíciles de definir de manera precisa en ciertos casos, se puede considerar un formalismo simplificado de mercado de predicción que resulte útil para demostrar propiedades sobre ellos. En nuestro caso haremos justamente eso, tomaremos varias formalizaciones simplificadas de mercado de predicción, lo cual nos permitirá analizarlas formalmente utilizando herramientas de la teoría de juegos [14].

Finalmente, el mismo fenómeno que describimos para un pronosticador u oráculo individual puede darse en el caso de mercados de predicción. Como antes, la predicción final emitida por el mercado puede potencialmente influir en el evento sobre el cual se realiza la predicción, llevando a resultados indeseados de índole similar a los descritos para el caso más simple de un único pronosticador. Un ejemplo análogo al descrito anteriormente sería un mercado de predicción que goce de cierta reputación e influencia sobre la población, y cuyos activos dependan de la situación económica de cierto país. Esto podría llevar no solo a comportamientos como los descritos anteriormente considerando al mercado en su totalidad, sino también a comportamientos indeseados de participantes individuales que buscan con sus predicciones alterar la predicción final del mercado de forma de influir en el evento sobre el cual realiza su predicción y aumentar su propia ganancia.

En este trabajo analizaremos pequeñas variaciones sobre un modelo simplificado de mercado de predicción, en el cual cada uno de los participantes del mercado (ya sea una persona real o un pronosticador informático) puede considerarse como un jugador de un *juego* [14] y la predicción final del mercado puede influir en el mundo sobre el cual predice [15].

Estudiaremos distintas estrategias posibles, tanto conjuntas como desde el punto de vista individual de cada jugador, y los efectos que potencialmente puede tener el hecho de que la predicción del mercado pueda influir sobre el evento finalmente observado.

El trabajo se organiza de la siguiente forma: en la sección 2 presentamos toda la notación y los conceptos previos necesarios para estudiar estos modelos de mercado de predicción, lo cual incluye nociones asociadas a la predicción de eventos y su puntuación, definiciones básicas sobre mercados de predicción, y conceptos básicos sobre teoría de juegos. En la sección 3 presentamos los resultados obtenidos para diversas variaciones sobre el modelo de base, tanto analíticos como experimentales. Finalmente, en la sección 4 se presentan algunas conclusiones generales y posibles trabajos futuros relacionados.

2. PRELIMINARES

2.1. Scoring rules

Intuitivamente, dado un pronosticador que realiza una predicción probabilística sobre un cierto evento, una *scoring rule* es una función que le asigna un puntaje a la predicción de acuerdo al evento que finalmente se materializa, de forma de poder dar una evaluación numérica sobre la precisión del pronosticador. Así, el objetivo del pronosticador será maximizar el puntaje obtenido.

Por ejemplo, dado un pronosticador que realiza predicciones sobre un evento con dos posibilidades (lo que llamaremos predicciones binarias), si la predicción realizada es que el evento ocurrirá con 80 % de confianza, entonces el pronosticador debería obtener mayor puntaje si el evento sucede que si no sucede. De la misma forma, otro pronosticador que le asigne 90 % de probabilidad al evento debería obtener mayor puntaje que el anterior si este ocurre, y obtener menos en caso de que no ocurra. Esto da lugar a la siguiente definición, que presentamos por completitud para el caso general de predicciones sobre medidas de probabilidad arbitrarias, pero que en este trabajo nos interesará para predicciones sobre espacios muestrales finitos, y más específicamente para predicciones binarias.

Formalmente, sea Ω un espacio muestral, \mathcal{A} una σ -álgebra y \mathcal{P} una clase convexa de medidas de probabilidad sobre (Ω, \mathcal{A}) , llamaremos *predicción probabilística* a cualquier $P \in \mathcal{P}$, y diremos que una función definida sobre Ω que toma valores en $\overline{\mathbb{R}} = [-\infty, +\infty]$ es *\mathcal{P} -cuasi-integrable* si es medible con respecto a \mathcal{A} y es cuasi-integrable con respecto a toda $P \in \mathcal{P}$. Se define entonces una **scoring rule** como cualquier función $S : \mathcal{P} \times \Omega \rightarrow \overline{\mathbb{R}}$ tal que $S(P, \cdot)$ es \mathcal{P} -cuasi-integrable para toda $P \in \mathcal{P}$. Sin embargo, para los alcances de este trabajo, será suficiente la siguiente definición simplificada para predicciones binarias:

Definición 1. Una **scoring rule** es cualquier función $S : [0, 1] \times \{0, 1\} \rightarrow \overline{\mathbb{R}}$ que toma una predicción $p \in [0, 1]$ y un resultado $q_o \in \{0, 1\}$, y le asigna un puntaje a la predicción.

Consideremos un pronosticador que realiza predicciones sobre la temperatura en un momento dado para un cierto lugar (esto es, la predicción es una función que da una probabilidad para cada posible temperatura). En este caso, el espacio de probabilidades es continuo, y es útil considerar entonces la definición general mencionada. Sin embargo, para este trabajo, más allá de la definición formal, es importante el concepto que, dada una predicción probabilística P y el evento posteriormente observado q_o , una scoring rule asigna el puntaje $S(P, q_o)$ al pronosticador. Es este concepto el que nos interesará para scoring rules sobre espacios muestrales finitos, e incluso, como dijimos, nos restringiremos principalmente al caso de predicciones binarias.

Por ejemplo, en lugar de considerar el caso donde se predice la temperatura para un cierto lugar, consideraremos el caso donde se predice si la temperatura será mayor al los 20 grados o no. Esto quiere decir que la predicción del pronosticador será simplemente un número $p \in [0, 1]$.

En la práctica, hay algunas scoring rules que se utilizan con mucha frecuencia, y que tienen características distintas entre sí. En la sección 2.1.3 veremos dos ejemplos típicos: la regla logarítmica y la regla de Brier. La primera diferencia que se puede observar entre ambas

es que, con la regla de Brier, lo que se busca es minimizar el puntaje obtenido, mientras que con la logarítmica se busca maximizarlo. Esto da lugar a las siguientes definiciones:

Definición 2. Una scoring rule se dice **orientada positivamente** si valores más altos corresponden a mejores puntajes. Análogamente, se dice **orientada negativamente** si valores más bajos corresponden a mejores puntajes.

En nuestro caso vamos a asumir, sin pérdida de generalidad, que las scoring rules con las que trabajamos son orientadas positivamente, es decir, vamos a considerar que los pronosticadores intentan maximizar el puntaje obtenido.

2.1.1. Strictly proper scoring rules

En principio, las scoring rules simplemente asignan puntajes a predicciones basadas en el resultado finalmente observado, sin ninguna restricción sobre el significado de este puntaje o los comportamientos que puede incentivar en predictores que busquen maximizarlo. Sin embargo, en la práctica es útil y necesario pedirles a las scoring rules que cumplan ciertas propiedades. Por ejemplo, un efecto deseable de las scoring rules es que incentiven a los pronosticadores a predecir su verdadera creencia sobre la distribución de probabilidad asociada al evento sobre el cual están prediciendo. Por ejemplo, podría suceder que una scoring rule le dé al pronosticador puntaje mínimo si la distribución real asociada al evento coincide con su predicción, lo cual sería sumamente contraproducente, ya que incentivaría al pronosticador a no predecir su verdadera creencia. Podría suceder también que la scoring rule dé puntajes muy altos a predicciones extremas si coinciden con lo observado (por ejemplo, que de un puntaje desproporcionadamente alto si se predijo un 90 % de probabilidades para un cierto evento y el evento sucede), incentivando al pronosticador a realizar predicciones extremas aunque no reflejen realmente la distribución verdadera del evento.

Tomemos por ejemplo una predicción sobre la probabilidad que el resultado de una tirada de moneda sea cara, y una scoring rule que asigna como puntaje $S(p, 1) = 1 + p$ y $S(p, 0) = \frac{1-p}{2}$. En este caso, el puntaje esperado se maximiza prediciendo siempre probabilidad 1, a pesar de que la verdadera probabilidad subyacente del evento es 50 %.

Es por esto que surgen las *strictly proper scoring rules*, cuyo objetivo es justamente incentivar al pronosticador a dar predicciones cuidadosas y honestas, es decir, que coincidan con su verdadera creencia. La forma en que van a hacer esto es asignándole puntaje máximo a una predicción dada cuando ésta coincide con la distribución real asociada al evento predicho.

Definición 3. Denotamos al **puntaje esperado** bajo $Q \in \mathcal{P}$ dada la predicción probabilística $P \in \mathcal{P}$ como:

$$S_e(P, Q) = \int S(P, \omega) dQ(\omega)$$

Observación 4. Para el caso de predicciones binarias, que es en el que nos concentraremos, p y q serán simplemente números que indicarán la probabilidad predicha de que suceda el evento y la probabilidad real de que este suceda respectivamente. En este caso, el puntaje esperado se define simplemente como:

$$S_e(p, q) = qS(p, 1) + (1 - q)S(p, 0) \tag{2.1}$$

A partir de este punto, siempre que nos refiramos a una *scoring rule* nos referiremos al puntaje esperado, como está definido en la ecuación 2.1, y la llamaremos simplemente S en lugar de S_e .

Definición 5. Sea S una scoring rule orientada positivamente, decimos que S es una **proper scoring rule** si

$$S(Q, Q) \geq S(P, Q) \text{ para todo } P, Q \in \mathcal{P}.$$

Definición 6. Decimos que S es una **strictly proper scoring rule** si cumple la definición 5 y la igualdad vale si y solo si $P = Q$.

En esencia, las *strictly proper scoring rules* son scoring rules que, dada una predicción probabilística P , maximizan el puntaje otorgado cuando ésta coincide con la verdadera distribución probabilística del evento predicho. De esta forma, como mencionamos antes, si el pronosticador tiene la creencia que la distribución probabilística Q es la verdadera distribución asociada al evento, no tiene ningún incentivo para predecir otra distribución $P \neq Q$. Esta propiedad de incentivar al pronosticador a dar como predicción su verdadera creencia es esencial a la hora de construir predictores. Recientemente, se han analizado los efectos negativos que puede tener en ciertos contextos el utilizar scoring rules que no sean strictly proper, y se estudiaron algunos ejemplos reales de cómo esto puede generar comportamientos indeseados, principalmente reportar probabilidades deshonestas [16]. En la sección 2.1.3 daremos algunos ejemplos concretos de strictly proper scoring rules muy utilizadas en la práctica.

Para ver una strictly proper scoring rule particular, consideremos el caso simple donde se realizan predicciones sobre el resultado de tirar una moneda balanceada. Una posible strictly proper scoring rule sería $S(p, q) = -(q - p)^2$, que es orientada positivamente. S es así una scoring rule que alcanza máximo si y solo si $p = q$. En el caso de la moneda balanceada, el máximo se alcanzará cuando la predicción del pronosticador sea justamente $1/2$. Notar que si omitimos el signo menos en S , seguimos obteniendo una strictly proper scoring rule, pero orientada negativamente.

2.1.2. Funciones convexas y caracterización de strictly proper scoring rules

A continuación daremos una caracterización para las *strictly proper scoring rules* en base a funciones convexas que será muy útil durante el desarrollo del trabajo.

Definición 7. Una función $G : \mathcal{P} \rightarrow \mathbb{R}$ se dice **convexa** si vale que para todo $p_0, p_1 \in \mathcal{P}$ y para todo $\alpha \in (0, 1)$

$$G(\alpha p_0 + (1 - \alpha)p_1) \leq \alpha G(p_0) + (1 - \alpha)G(p_1)$$

Para el caso de espacios muestrales finitos, que es el que más nos interesa en este trabajo, donde p_0 y p_1 son simplemente vectores de probabilidad, el multiplicarlos por α es el producto escalar usual, y el resultado de $\alpha p_0 + (1 - \alpha)p_1$ cae dentro del dominio de G por ser combinación convexa de vectores de probabilidad.

Definición 8. Una función $G : \mathcal{P} \rightarrow \mathbb{R}$ se dice **estrictamente convexa** si cumple la definición 7 y la igualdad vale solamente con $p_0 = p_1$

Geoméricamente, una función es convexa si la recta que une dos puntos cualquiera de su gráfico siempre queda por encima del gráfico. Por otro lado, un resultado conocido en el contexto del análisis de funciones convexas es que no toda función convexa es diferenciable en todo su dominio. El ejemplo clásico es la función módulo, que es convexa (aunque no estrictamente convexa), pero no es derivable en el cero. Es por esto que surge el concepto de *subgradiente*, que es una generalización del gradiente, y que es necesario para el teorema de caracterización que presentaremos.

Definición 9. Sea $G : \mathcal{P} \rightarrow \mathbb{R}$ una función convexa, un **subgradiente** de G es alguna función $G^* : \mathcal{P} \rightarrow \mathbb{R}^n$ tal que para todo $p_0, p_1 \in \mathcal{P}$

$$G(p_0) - G(p_1) \geq G^*(p_0)(p_1 - p_0)$$

Observación 10. Siempre existe el subgradiente de una función convexa. Si además la función G es diferenciable en $p_0 \in \mathcal{P}$, entonces $G^*(p_0) = \nabla G(p_0)$. Sino, podría haber varias posibilidades para G^* .

Por otro lado, en la sección 2.1 dijimos que una scoring rule podía eventualmente tomar los valores $-\infty$ y $+\infty$. Para el teorema de caracterización necesitamos restringirnos a una subclase de scoring rules, llamadas scoring rules regulares:

Definición 11. Una scoring rule $S : \mathcal{P} \times \mathcal{P} \rightarrow \overline{\mathbb{R}}$ orientada positivamente se dice **regular** si $S(p, q)$ es un valor finito para todo $p, q \in \mathcal{P}$, excepto tal vez que $S(p, q) = -\infty$ si $p \neq q$.

Generalmente, a menos que se indique lo contrario, trabajaremos con scoring rules que siempre toman valores reales, pero es interesante notar que ese no es siempre necesariamente el caso, y que existe scoring rules que no lo cumplen, como por ejemplo la regla logarítmica que presentaremos más abajo para el caso de predicciones binarias 2.3. Finalmente, presentamos una caracterización dada en [17] sumamente útil para *strictly proper scoring rules*:

Teorema 12. (Gneiting y Raftery 2007) [17] Una scoring rule regular $S : \mathcal{P} \times \Omega \rightarrow \mathbb{R}$ es una *proper scoring rule* si y solo si existe una función convexa $G : \mathcal{P} \rightarrow \mathbb{R}$ con subgradiente G^* tal que para todo $p, q \in \mathcal{P}$

$$S(p, q) = G(p) + G^*(p)(q - p) \tag{2.2}$$

El teorema también vale si se reemplaza *proper scoring rule* con *strictly proper scoring rule* y *convexa* con *estrictamente convexa*.

Observación 13. Si existe G diferenciable además de convexa tal que S puede escribirse como describe la fórmula 2.2, entonces, además de ser *strictly proper*, S cumple que:

$$S(p, q) = G(p) + \nabla G(p)(q - p)$$

Esto restringe la clase de *strictly proper scoring rules* consideradas, ya que no todas pueden ser escritas con una G diferenciable (el Teorema 12 solo garantiza convexidad), pero aporta propiedades que serán útiles, principalmente que S resulta derivable.

2.1.3. Ejemplos de strictly proper scoring rules

Podemos dar ahora algunos ejemplos de strictly proper scoring rules habituales que se suelen usar en la práctica [4]. Supongamos por simplicidad que el espacio muestral es finito (de hecho, en esta tesis nos limitaremos a un caso más simple aún, que es el caso de predicciones binarias). La primera regla que presentamos es la *regla logarítmica* [5], que puede definirse como:

$$\text{logscore}(P, q) = \log P(q)$$

Donde q es el evento observado y P la predicción probabilística del pronosticador. Aunque se suele usar el logaritmo decimal, la regla logarítmica resulta una strictly proper scoring rule para cualquier base $b \geq 2$. Otra strictly proper scoring rule muy utilizada es la *regla cuadrática*, o *regla de Brier* [1]. Ésta se define como:

$$B(P, o) = \frac{1}{N} \sum_{i=1}^N (P_i - o_i)^2$$

donde P_i es la probabilidad asignada al evento i por el pronosticador y o_i es 1 si se observó el evento o_i y 0 si no. Podemos observar que la regla logarítmica es una regla orientada positivamente, mientras que la regla de Brier está orientada negativamente.

Si bien estas definiciones son generales y pueden utilizarse tanto para predicciones binarias como para no binarias sobre espacios muestrales finitos, en nuestro caso trabajaremos con sus versiones simplificadas para predicciones binarias. Para el caso de la regla logarítmica, su versión simplificada es:

$$\text{logscore}(p, q) = q \log p + (1 - q) \log (1 - p) \quad (2.3)$$

y para la regla de Brier es:

$$B(p, q) = q(p - 1)^2 + (1 - q)p^2 \quad (2.4)$$

En particular, para mantener coherencia con el resto del trabajo y sin pérdida de generalidad, utilizaremos la versión de la regla de Brier modificada para que sea orientada positivamente, es decir:

$$B'(p, q) = -(q(p - 1)^2 + (1 - q)p^2) \quad (2.5)$$

2.1.4. Profecías autocumplidas

Como dijimos antes, la ventaja de las strictly proper scoring rules es que incentivan al pronosticador a predecir la distribución probabilística que coincide con su verdadera creencia. Sin embargo, esto puede no ser así cuando la predicción del pronosticador influye en el resultado del evento sobre el cual está prediciendo, lo cual puede llevar a resultados indeseados. Para modelar esta situación, introducimos la siguiente definición:

Definición 14. Consideramos que existe una función $f : [0, 1] \rightarrow [0, 1]$, llamada **función de influencia**, que describe la influencia de la predicción p del pronosticador en la verdadera distribución asociada al evento q .

De esta forma, tenemos que $q = f(p)$ (Notar que esta función f puede pensarse como una f_{q_I} , que modifica la probabilidad a priori q_I de acuerdo a la predicción p). Así, la función que va a intentar maximizar el pronosticador será $S(p, f(p))$, con S alguna strictly proper scoring rule. Se puede observar entonces que, dada una predicción p , el puntaje obtenido por dicha predicción será máximo si p es un punto fijo de f , y notar que, por el teorema de punto fijo de Brower, si f es continua tendrá al menos un punto fijo. Este modelo fue analizado en [18], donde se demostraron varios resultados importantes. Por un lado, se demostró que ante la presencia de varios puntos fijos, al pronosticador le convendrá favorecer puntos fijos más extremos (más cercanos a 0 o a 1). Si además la strictly proper scoring rule utilizada es simétrica (esto es, $S(p, p) = S(1 - p, 1 - p)$ para todo $p \in [0, 1]$), entonces el pronosticador favorecerá aquellos puntos fijos de menor entropía de Shannon [19]. Esto puede entenderse como que el pronosticador intentaría volver el mundo lo más predecible posible, de forma que resulte más fácil predecir eventos sobre él. Por otro lado, se demostró que hay casos donde, incluso aunque existan puntos fijos de f , el valor óptimo para el pronosticador no será un punto fijo y que, bajo ciertas condiciones, casi seguramente la predicción óptima no es un punto fijo. Este resultado, que puede resultar anti intuitivo, se debe a que el pronosticador no solo intentará manipular el evento para mejorar su precisión, sino también para extremalizar ¹ la probabilidad resultante. Finalmente, se probaron ciertas cotas sobre la distancia que puede llegar a haber entre la predicción óptima para el pronosticador y puntos fijos de f . Además del análisis teórico de este fenómeno [18, 20], han surgido algunas propuestas en los últimos años para solucionar o al menos paliar este fenómeno en diversos contextos [21], como los oráculos contrafácticos [6, 22] o los *stop gradients* [23, 24].

2.2. Mercados de predicción

Los *mercados de predicción* son mercados abiertos donde se intercambian activos cuyo valor depende del resultado de algún evento particular. Los precios de los diferentes activos del mercado en un momento dado se pueden interpretar como la predicción del mercado sobre la probabilidad de dicho evento [8]. De esta forma, los mercados de predicción son una forma de recoger el conocimiento de expertos o incluso el conocimiento popular de forma de generar predicciones sobre eventos futuros. Se ha observado que en muchos casos este tipo de modelos de predicción basados en la recolección de muchas predicciones individuales generan predicciones tan buenas o mejores que las de otras instituciones con el mismo objetivo [11]. Este fenómeno ha sido estudiado en diversas áreas como psicología y economía, y ha dado lugar al concepto de *wisdom of crowds*, desarrollado por Francis Galton en el siglo XX y estudiado también en la actualidad [25]. Hay asimismo actualmente varios ejemplos conocidos de mercados y plataformas de predicción sobre diversos temas [26, 27, 28], con resultados generalmente buenos [29, 30, 31].

¹ Anglismo que viene del verbo *extremize*.

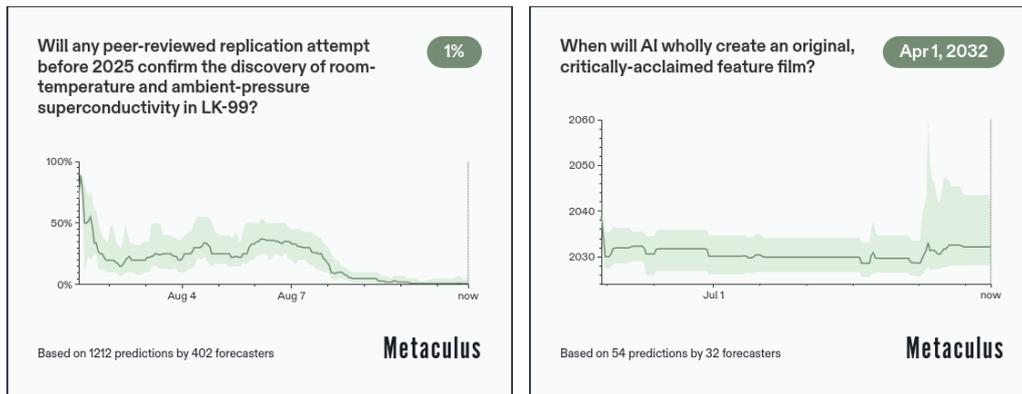


Fig. 2.1: Ejemplos de la plataforma de predicción Metaculus. En el primer caso las predicciones son binarias, por lo que se muestra la evolución de las predicciones a través del tiempo. En el segundo caso las predicciones ya no son binarias, por lo que se muestra la evolución de la función de densidad que resulta de las predicciones de los participantes.

En particular, los mercados de predicción han sido estudiados desde diversos puntos de vista, incluyendo la economía [12], la computación [10, 5] y la matemática [13, 8]. La complejidad del estudio de este tipo de mercados puede ser muy grande, ya que en muchos casos es difícil formalizar sus reglas, o es difícil conocer las estrategias de cada agente que participa en el mercado o estrategias conjuntas entre ellos. Es por esto que nos limitaremos a estudiar una formalización simplificada de mercado de predicción que facilitará su estudio analítico y experimental. Nos concentramos en estudiar uno de los principales problemas que pueden tener los mercados de predicción, que es el de las *profecías autocumplidas* (o una versión más general del concepto, que es el de *performative predictions* [21]), similar a lo descrito en la sección 2.1.4. Este problema ha sido analizado en los últimos tiempos [32, 15], y, análogamente a lo que sucedía en el caso de un solo pronosticador, consiste en que la predicción generada por el mercado influye en el resultado del evento sobre el cual se realizan las predicciones, lo cual puede llevar a comportamientos indeseados.

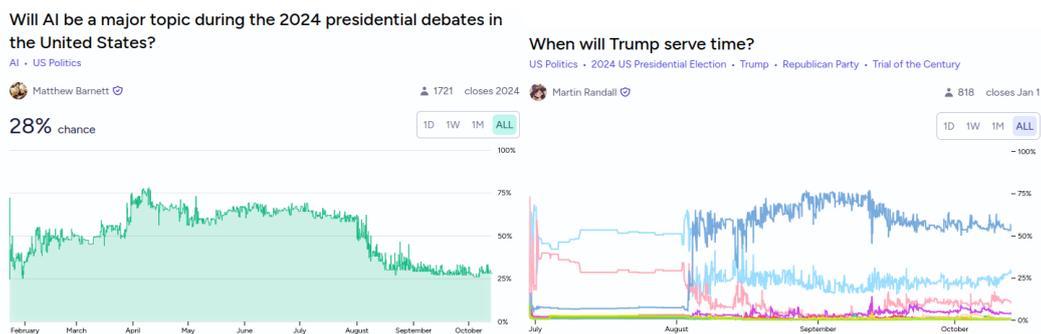


Fig. 2.2: Ejemplos de mercados de predicción de la página Manifold. Al igual que en el caso anterior, en el primer caso las predicciones son binarias, mientras que en el segundo no lo son.

2.2.1. Nuestro formalismo

Tomaremos un formalismo similar al presentado en [15]. Partimos en principio de un mercado con n participantes que realizan simultáneamente una única predicción binaria sobre un cierto evento, y que son puntuados con una strictly proper scoring rule (notar que estas son potencialmente distintas para cada jugador). Además, cada participante tendrá asociado un peso $w_i \in (0, 1)$, que típicamente representará la fracción del capital que aporta el jugador i al capital total del mercado, de forma que $\sum_{i=1}^n w_i = 1$, y que determinará el peso de su predicción a la hora de decidir la predicción final del mercado. Finalmente, consideraremos que la predicción final del mercado será un promedio de las predicciones individuales de cada participante teniendo en cuenta su peso, es decir, dadas las predicciones (p_1, \dots, p_n) , la predicción final \hat{p} será:

$$\hat{p} = \sum_{i=1}^n w_i p_i \quad (2.6)$$

lo cual está bien definido ya que una combinación convexa de medidas de probabilidad da como resultado una medida de probabilidad. Además, como en el modelo presentado en la sección 2.1.4, asumiremos que existe una función $f : [0, 1] \rightarrow [0, 1]$ que modela la influencia de la predicción final del mercado en la distribución de probabilidad asociada al evento sobre el cual trata el mercado. De esta forma, dicho evento ocurrirá con probabilidad $f(\hat{p}) = f(\sum_{i=1}^n w_i p_i)$ tras las predicciones, de forma que cada participante del mercado intentará maximizar su puntaje esperado, es decir, la función $S(p_i, f(\sum_{i=1}^n w_i p_i))$. Asumiremos en principio que tanto la función f como las scoring rules utilizadas para asignar puntajes a cada jugador son conocidas por todos los participantes, así como el peso asignado a cada uno (al estudiar el formalismo como un juego, esto querrá decir que este es de información completa).

En [15] se demostró que, si las predicciones de los jugadores resultan ser un equilibrio de Nash 15 (concepto que desarrollaremos en la siguiente sección), los jugadores con un suficientemente bajo peso darán una predicción cercana a $f(\hat{p})$. También se demostró que, si todos los jugadores tienen pesos bajos, entonces un equilibrio de Nash será cercano a un punto fijo de f .

A lo largo de esta tesis realizaremos diversos análisis sobre este formalismo, considerando pequeñas variaciones sobre sus reglas. Por ejemplo, estudiaremos los casos donde se permite más de una ronda de predicción para cada jugador, donde se considera al mercado como de suma constante, o donde las predicciones no se realizan simultáneamente sino secuencialmente. El estudio de diversas variaciones sobre este modelo puede llevar a una mejor aproximación de lo que sucede en un mercado de predicción real, de forma que los resultados teóricos obtenidos pueden dar una mejor intuición del efecto del fenómeno estudiado en la realidad. Por ejemplo, resulta más cercano a la realidad considerar que se permite más de una ronda de predicciones (los participantes pueden modificar su predicción original), o que las predicciones no se realizan simultáneamente, sino de forma secuencial, ambas variantes que estudiaremos en esta tesis

2.3. Conceptos de teoría de juegos

El modelo descrito en la sección anterior puede ser visto como un juego, donde la predicción emitida por cada jugador representa su *estrategia*, y la scoring rule representa su

función de pago, que determina su *ganancia* al realizar una cierta predicción. Es por eso que vamos a utilizar herramientas de la *teoría de juegos* para analizarlo.

Formalmente, un juego consiste en una cierta cantidad de *jugadores*, j_1, \dots, j_n , el conjunto de *estrategias posibles*, E_i , que son todas las posibles decisiones que un jugador puede tomar en un momento dado, y una *función de pago* S_i definida para cada jugador, que determina su ganancia. Es importante remarcar que la ganancia de un jugador no depende únicamente de su estrategia, sino de las de todos los jugadores. Notaremos entonces con el vector $e = (e_1, \dots, e_n)$ a una *estrategia conjunta*, que incluye la decisión de cada uno de los jugadores, con $e_i \in E_i$. La ganancia del i -ésimo jugador estará dada entonces por $S_i(e)$.

Hay una gran variedad de juegos clásicos que pueden ser estudiados utilizando la teoría de juegos. Un ejemplo usual es el *dilema del prisionero* [33], aunque también otros juegos conocidos como el ajedrez o las damas. Todos estos son ejemplos de juegos de dos jugadores, aunque con características distintas. En el dilema del prisionero, ambos jugadores deben tomar una decisión simultáneamente, y sus posibles decisiones son las mismas, mientras que en el ajedrez y las damas los jugadores toman decisiones por turnos, y sus posibles movimientos no son siempre iguales.

Un concepto importante en el estudio de los juegos es el de *equilibrio de Nash*. Intuitivamente, una estrategia conjunta es un equilibrio de Nash si cualquier jugador que decide unilateralmente cambiar de estrategia (es decir, si solo ese jugador modifica su estrategia), disminuye su ganancia.

Definición 15. Una estrategia conjunta $(e_1, \dots, e_i, \dots, e_n)$ es un **Equilibrio de Nash puro** si

$$S_i(e_1, \dots, e_i, \dots, e_n) \geq S_i(e_1, \dots, e'_i, \dots, e_n)$$

para todo $e'_i \in E_i$ y para todo $1 \leq i \leq n$.

Podemos además dar una definición equivalente de equilibrio de Nash que resultará muy útil. Para eso, será útil la siguiente definición:

Definición 16. Dada una estrategia conjunta $E = (e_1, \dots, e_i, \dots, e_n)$, definimos

$$E_{-i} = (e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n)$$

Es decir, E_{-i} es la estrategia conjunta E pero quitando la estrategia del i -ésimo jugador. Podemos definir entonces el concepto de *mejor respuesta*, que intuitivamente es la mejor estrategia que puede elegir el jugador i ante la estrategia conjunta E_{-i} . Formalmente:

Definición 17. Dada una estrategia conjunta $E = (e_1, \dots, e_n)$, e_i es la **mejor respuesta** del jugador i a E_{-i} si

$$S_i(e_1, \dots, e_i, \dots, e_n) \geq S_i(e_1, \dots, e'_i, \dots, e_n)$$

para todo $e'_i \in E_i$ y para todo $1 \leq i \leq n$.

A partir de la definición anterior, podemos observar que:

Observación 18. Una estrategia conjunta $E = (e_1, \dots, e_n)$ es un equilibrio de Nash sii e_i es la mejor respuesta del jugador i a E_{-i} para todo $1 \leq i \leq n$.

La idea intuitiva detrás de la observación es que, si e_i no fuera una mejor respuesta del jugador i a E_{-i} , entonces el jugador i podría decidir unilateralmente cambiar su estrategia, y su puntaje mejoraría.

Por otro lado, además de clasificar las estrategias dentro de un juego de acuerdo a sus características, también se pueden clasificar los juegos de acuerdo a sus propiedades. Por ejemplo, se pueden clasificar los juegos de acuerdo a si existe o no alguna simetría entre las funciones de pago de los jugadores. Por ejemplo, en juegos como el ajedrez, puede considerarse que lo que gana un jugador necesariamente lo pierde el otro. Es decir, si decimos que ganar da 1 punto, perder -1 y empatar $1/2$, entonces sin importar cuál sea el resultado, si sumamos el puntaje de ambos jugadores el resultado da 0. Esto da lugar a la siguiente definición:

Definición 19. Un juego de n jugadores con funciones de pago S_i se dice de **suma cero** si

$$\sum_{i=1}^n S_i(p) = 0$$

para toda estrategia conjunta $p = (p_1, \dots, p_n)$.

Más en general, pueden considerarse juegos donde al sumar los pagos de todos los jugadores el resultado es siempre el mismo, sin que sea necesariamente 0.

Definición 20. Un juego de n jugadores con funciones de pago S_i se dice de **suma constante** si

$$\sum_{i=1}^n S_i(p) = c \tag{2.7}$$

para toda estrategia conjunta $p = (p_1, \dots, p_n)$ y cierta constante c .

Un caso clásico de juego de suma constante es el de repartición de una cantidad fija de recursos entre un cierto número de jugadores. Como la cantidad de recursos es fija, sin importar cómo estos sean repartidos, la suma total de recursos que se reparten será siempre la misma, y, suponiendo que los recursos tienen el mismo valor para todos los participantes, esto implica que la suma de las ganancias de todos los jugadores será siempre la misma.

Por otra parte, podemos clasificar los juegos de acuerdo al conocimiento de cada jugador con respecto a las distintas componentes del juego. Evidentemente, todos los jugadores deben conocer las reglas básicas del juego, y sus propias funciones de pago; sin embargo, cada jugador podría o no conocer por ejemplo las funciones de pago de los demás jugadores, así como sus posibles estrategias.

Definición 21. Diremos que un juego es de **información completa** si el conocimiento acerca del juego de los otros participantes está disponible para todos los jugadores. En particular, todos los participantes conocen las funciones de pago y las posibles estrategias de todos los demás jugadores.

Por ejemplo, el formalismo para mercados de predicción presentado en 2.2.1 define un juego de información completa, ya que cada participante conoce las funciones de pago de todos los demás jugadores, que en este caso son las strictly proper scoring rules con las cuales son puntuados, y conoce sus posibles estrategias, que son siempre predicciones en el

intervalo $[0, 1]$. Otro ejemplo de un juego de información completa es el *dilema del prisionero* mencionado anteriormente, donde ambos jugadores tienen la misma función de pago y las mismas estrategias posibles y conocen este hecho. En muchas situaciones clásicas, por ejemplo en teoría económica, la información completa es una asunción necesaria para el funcionamiento de ciertos modelos, ya que, si un juego no es de información completa, los jugadores no pueden conocer el efecto de sus decisiones en los demás participantes. Notar que la información completa no dice nada sobre el conocimiento de las jugadas realizadas por los demás jugadores. Esto da lugar a la siguiente definición:

Definición 22. Diremos que un juego es de **información perfecta** si todos los jugadores conocen todo lo que podrían desear conocer sobre el juego desde que éste comenzó a la hora de elegir su próxima jugada. En particular, todos los jugadores conocen las funciones de pago y las posibles estrategias de todos los demás jugadores, así como todas las jugadas hechas hasta el momento por los demás jugadores.

Notar que esta definición difiere de la de información completa en que la información completa habla solamente de conocimiento sobre la estructura del juego, mientras que la información perfecta habla también de conocimiento sobre el desarrollo del juego hasta el momento. Por ejemplo, el formalismo presentado en 2.2.1 no es de información perfecta, ya que al momento de realizar una predicción, ningún jugador conoce las predicciones realizadas por los demás participantes, ya que se realizan de forma simultánea. Por la misma razón, el juego del dilema del prisionero tampoco es de información perfecta. Algunos ejemplos de juegos de información perfecta son el ajedrez o las damas, ya que, al jugarse por turnos, al momento de realizar una jugada los participantes conocen todas las jugadas hechas por el otro jugador hasta el momento.

Otro posible ejemplo de juego con información perfecta sería una variación sobre el formalismo presentado en 2.2.1 donde las predicciones no se realizan de forma simultánea, sino que se realizan de forma secuencial. Si asumimos como antes que todos los participantes conocen las funciones de pago de todos los demás, así como sus estrategias y sus pesos, que todos conocen la función f que modela la influencia de la predicción del mercado sobre el evento predicho, y que al realizar una predicción conocen todas las predicciones realizadas por los demás jugadores hasta el momento, entonces dicho juego resulta de información perfecta.

Finalmente, distinguiremos los juegos donde todos los jugadores eligen sus estrategias simultáneamente, como es el caso del formalismo presentado en 2.2.1, de los juegos por turnos, como es el caso anterior de predicciones por turnos. Estos últimos se denominan **juegos secuenciales**.

En el caso de los juegos secuenciales, se puede extender el concepto de equilibrio de Nash para tener en cuenta que hay varias rondas donde los jugadores eligen estrategias.

Definición 23. Una estrategia conjunta es un **equilibrio de Nash perfecto en subjuegos** si genera un equilibrio de Nash puro en cada subjuego del juego original.

Intuitivamente, esto quiere decir que en cualquier momento del juego, las estrategias de los jugadores a partir de ese punto representan un equilibrio de Nash en el subjuego restante (es decir, el juego que resulta tras la estrategia elegida por el último jugador en jugar). Hay un resultado conocido que garantiza, bajo ciertas condiciones, la existencia de estos equilibrios:

Teorema 24. *Todo juego secuencial, finito y de información perfecta tiene un equilibrio de Nash perfecto en subjuegos.*

La idea para encontrar un equilibrio de Nash perfecto en subjuegos en juegos secuenciales, finitos y de información perfecta es utilizar la *inducción hacia atrás*. Partiendo desde el final, es decir, desde el punto de vista del último jugador, si tomamos como fijas las estrategias de los primeros $n - 1$ jugadores, entonces queda unívocamente determinada la estrategia que debe tomar el jugador n para maximizar su pago (suponiendo que, en caso de empate, existe algún método de desempate conocido por todos). De la misma forma, si consideramos luego al jugador $n - 1$, como conoce la función de pago del jugador n y sus posibles estrategias, puede también determinar unívocamente cuál es la estrategia que maximiza su pago, ya que para cada posible estrategia que tome puede saber cuál será la respuesta del jugador n . Siguiendo inductivamente el mismo argumento hasta llegar al jugador 1, se obtiene una estrategia que resulta un equilibrio de Nash perfecto en subjuegos, que representa una estrategia conjunta que, asumiendo que todos los jugadores juegan racionalmente para maximizar su puntaje, le garantiza puntaje máximo a cada jugador [14]. Notar que la hipótesis de información perfecta es esencial para que funcione el argumento de inducción hacia atrás.

Algunos ejemplos donde se puede aplicar esta estrategia son juegos como el ta-te-ti, donde la solución obtenida mediante inducción hacia atrás es conocida. Sin embargo, este método podría aplicarse también a juegos más complejos como el ajedrez o el Go, ya que ambos son secuenciales, finitos, y de información perfecta, aunque por su complejidad, resulta extremadamente difícil hacerlo en la práctica debido al crecimiento exponencial de la cantidad de posibilidades a verificar si se quiere hacer de forma exhaustiva.

3. RESULTADOS

A continuación presentaremos los resultados obtenidos, tanto analíticos como experimentales, para distintas variaciones sobre el formalismo para *mercados de predicción* presentado en la sección 2.2.1. Consideraremos estos mercados como juegos para estudiarlos utilizando conceptos y herramientas de la *teoría de juegos*. Nuestro juego base será el que se obtiene al considerar el formalismo mencionado en la sección 2.2.1 como un juego, es decir, uno de n jugadores donde la estrategia de cada jugador es una predicción $p_i \in [0, 1]$, y donde cada jugador tiene un peso $w_i > 0$ tal que $\sum_{i=1}^n w_i = 1$, que típicamente representa la fracción del capital que aporta el jugador i al capital total del mercado. Consideramos que cada jugador es puntuado por una *strictly proper scoring rule* S_i (donde estas reglas son potencialmente distintas entre sí), que puede considerarse como el puntaje obtenido por cada jugador en el juego, es decir, como la función de pago de cada jugador, y consideramos que la predicción final del mercado puede influir sobre el evento predicho, hecho que representamos con una función f que modela esta influencia, como fue descrito en la sección 2.1.4. Para obtener la predicción final del mercado, simplemente tomamos un promedio teniendo en cuenta los pesos de cada jugador, como indica la fórmula 2.6:

$$\hat{p} = \sum_{i=1}^n w_i p_i$$

De esta forma, asumimos que cada jugador intentará maximizar su puntaje esperado, es decir:

$$S_i(p_i, f(\hat{p})) \tag{3.1}$$

Asumimos también en principio que el juego es de *información completa*, es decir, todos los jugadores conocen las funciones de pago de todos los jugadores, así como sus posibles estrategias, sus pesos, y la función f que modela la influencia de la predicción del mercado sobre el evento predicho. Notar que la hipótesis que todos los jugadores conocen la función f de forma exacta es una hipótesis muy fuerte, ya que no solo conocen la forma de f , sino también exactamente cómo quedará la probabilidad final del evento dependiendo de la predicción \hat{p} . Como mencionaremos luego, relajar esta hipótesis sería un camino importante para investigaciones futuras.

Uno de los objetivos será entonces generalizar el Corolario 15 presentado en [15] para otras variantes de este formalismo. Este resultado dice que:

Teorema 25 ([15], Corollary 15). *Sea G la función convexa asociada a la proper scoring rule¹ S , sea la estrategia conjunta $\mathbf{p} = (p_1, \dots, p_n)$ un equilibrio de Nash, con f diferenciable en \hat{p} y Lipschitz continua, G y G' diferenciables en p_i , $G''(p_i) \neq 0$ y con $C = \sup_{p \in [0,1]} \left| \frac{G'(p)}{G''(p)} \right| < \infty$, vale que dado $\epsilon > 0$, existe $\delta > 0$ tal que si $w_i < \delta$ para todo i , entonces:*

$$\begin{aligned} |f(\hat{p}) - p_i| &< \epsilon \text{ para todo } i, \text{ y} \\ |f(\hat{p}) - \hat{p}| &< \epsilon \end{aligned}$$

¹ En la publicación original se asumía que todos los jugadores eran puntuados con la misma scoring rule.

3.1. Predicciones simultáneas iteradas

Comenzamos tomando un modelo como el planteado en la sección 2.2.1, donde todas las predicciones de los jugadores se realizan de manera simultánea, solo que ahora consideraremos que se realiza más de una ronda de predicciones. Para empezar, ahora la predicción final del mercado en la ronda k se calcula en función de las predicciones de los jugadores en esa ronda:

$$\hat{p}_k = \sum_{i=1}^n w_i p_{i_k}$$

Por otro lado, modificaremos también la forma en que puntuamos las predicciones de cada jugador. Sea S_i una strictly proper scoring rule asociada a cada jugador i (asumiremos por simplicidad que todas las scoring rules serán siempre positivas), ahora el puntaje de cada jugador i en la ronda k será

$$T_i(p_{i_k}, f(\hat{p}_k)) = S_i(p_{i_k}, f(\hat{p}_k)) - S_i(m(p_{-i_k}), f(\hat{p}_k)) \quad (3.2)$$

donde $m(p_{-i_k})$ es la mediana de las predicciones de todos los demás jugadores en la ronda k (p_{-i_k} representa las predicciones de todos los jugadores excepto el i -ésimo en la ronda k). Es decir, medimos qué tan buena es la predicción del jugador i respecto de la mediana de las predicciones de los demás jugadores (aunque, si bien es razonable utilizar la mediana, se verá que la elección de la función m no afecta a los resultados posteriores).

La intuición detrás del formalismo es que tras cada ronda, el observar las predicciones realizadas por los demás jugadores y el puntaje obtenido puede influir en la decisión de la ronda siguiente, y en cada ronda, cada jugador es puntuado considerando como afectaría la predicción del mercado en esa ronda a la probabilidad del evento predicho. Sin embargo, consideramos que las predicciones intermedias del mercado no afectan realmente al evento, solamente la predicción del mercado tras la ronda final tiene ese efecto. Es por esto que tomamos que la probabilidad a priori del evento es la misma en todas las rondas.

Por otro lado, es interesante notar que en una situación donde la predicción del mercado no influye en el evento sobre el cual predice, la scoring rule T_i resulta strictly proper, ya que el jugador i no puede afectar el puntaje obtenido por la mediana de las predicciones de los demás jugadores, entonces la forma de maximizar T_i es maximizando S_i , que es una strictly proper scoring rule. Sin embargo, en nuestro contexto, como la predicción del jugador i tiene un peso en la predicción final del mercado \hat{p}_k , y la predicción final del mercado influye en la distribución de probabilidad subyacente del evento, lo cual es $f(\hat{p}_k, q)$ en nuestro formalismo, con la cual se puntúa a la mediana de las predicciones de los demás jugadores, entonces el jugador i puede no solo intentar maximizar su puntaje como en el caso anterior, sino que puede también intentar afectar el puntaje de la mediana de las predicciones para disminuir el puntaje obtenido por los otros jugadores, por lo que T_i podría no resultar una strictly proper scoring rule.

Presentamos a continuación un resultado que extiende el Teorema 25 para esta generalización del formalismo básico planteado en 2.2.1.

Teorema 26. *Sea G_i la función convexa asociada a la scoring rule del jugador i , S_i , proveniente del Teorema de caracterización 12, y sea k la última ronda de predicciones realizada. Si la función G_i es de clase C^2 , la función f es diferenciable en \hat{p}_k y Lipschitz continua, la estrategia conjunta $\mathbf{p}_k = (p_{1_k}, \dots, p_{n_k})$ es un equilibrio de Nash puro, y vale que $G_i''(p_{i_k}) \neq 0$, entonces para todo $\epsilon > 0$ existe un $\delta > 0$ tal que, si el peso $w_i < \delta$, entonces $|f(\hat{p}_k) - p_{i_k}| < \epsilon$.*

Demostración. Empecemos por plantear la derivada de la scoring rule S_i del i -ésimo jugador. Por el teorema de caracterización 12, tenemos que:

$$\frac{\partial}{\partial p_i} S_i(p_i, f(\hat{p})) = \frac{\partial}{\partial p_i} (G_i(p_i) + G_i^*(p_i)(f(\hat{p}) - p_i))$$

donde G_i es una función convexa. Suponiendo que G_i es además doblemente derivable, entonces

$$\begin{aligned} \frac{\partial}{\partial p_i} S_i(p_i, f(\hat{p})) &= \frac{\partial}{\partial p_i} (G_i(p_i) + G_i'(p_i)(f(\hat{p}) - p_i)) \\ &= G_i'(p_i) + G_i'(p_i)(w_i f'(\hat{p}) - 1) + G_i''(p_i)(f(\hat{p}) - p_i) \\ &= w_i G_i'(p_i) f'(\hat{p}) + G_i''(p_i)(f(\hat{p}) - p_i) \end{aligned}$$

Y por lo tanto

$$\frac{\partial}{\partial p_i} S_i(p_i, f(\hat{p})) = w_i G_i'(p_i) f'(\hat{p}) + G_i''(p_i)(f(\hat{p}) - p_i) \quad (3.3)$$

La demostración luego es muy similar a la del Teorema 14 de [15]. Si $\mathbf{p}_k = (p_{1_k}, \dots, p_{n_k})$ es un equilibrio de Nash puro y $p_{i_k} \in (0, 1)$, entonces la derivada de la scoring rule $T_i(p_{i_k}, f(\hat{p}_k))$ respecto a p_{i_k} debe ser 0 (si $p_{i_k} = 0$ debe ser negativa y si $p_{i_k} = 1$ debe ser positiva). Entonces si $p_{i_k} \in (0, 1)$:

$$\frac{\partial}{\partial p_{i_k}} T_i(p_{i_k}, f(\hat{p}_k)) = \frac{\partial}{\partial p_{i_k}} S_i(p_{i_k}, f(\hat{p}_k)) - \frac{\partial}{\partial p_{i_k}} S_i(m(p_{-i_k}), f(\hat{p}_k)) \quad (3.4)$$

$$= G_i''(p_{i_k})(f(\hat{p}_k) - p_{i_k}) + w_i f'(\hat{p}_k)(G_i'(p_{i_k}) - G_i'(m(p_{-i_k}))) = 0 \quad (3.5)$$

donde usamos la fórmula para la derivada de S_i presentada en 3.3. Por lo tanto:

$$|f(\hat{p}_k) - p_{i_k}| = \frac{|w_i f'(\hat{p}_k)(G_i'(p_{i_k}) - G_i'(m(p_{-i_k})))|}{G_i''(p_{i_k})}$$

notar que el módulo no es necesario en el denominador de la parte derecha, ya que, como G_i es convexa y C^2 , entonces G_i'' es siempre positiva, y por hipótesis $G_i''(p_{i_k}) \neq 0$. Ahora, como G_i es C^2 , entonces G_i' y G_i'' son funciones continuas definidas en compactos, por lo que alcanzan máximo y mínimo, y la función f cumple que $|f'(\hat{p}_k)| \leq L$ por ser Lipschitz continua (con L su constante de Lipschitz). Sean entonces:

$$\begin{aligned} C &= \max_{p_1, p_2 \in [0, 1]} |G_i'(p_1) - G_i'(p_2)| \\ D &= \min_{p \in [0, 1]} G_i''(p) \end{aligned}$$

Además, por hipótesis pedimos que $w_i < \delta$ para algún $\delta > 0$. Tomando $\delta = \frac{D\epsilon}{CL}$ tenemos entonces que

$$|f(\hat{p}_k) - p_{i_k}| = \frac{|w_i f'(\hat{p}_k)(G_i'(p_{i_k}) - G_i'(m(p_{-i_k})))|}{G_i''(p_{i_k})} < \frac{\delta CL}{D} = \epsilon$$

En el caso que $p_{i_k} = 0$ tenemos que

$$\frac{\partial}{\partial p_{i_k}} T_i(p_{i_k}, f(\hat{p}_k)) = G_i''(p_{i_k})(f(\hat{p}_k) - p_{i_k}) + w_i f'(\hat{p}_k)(G_i'(p_{i_k}) - G_i'(m(p_{-i_k}))) \leq 0$$

y por lo tanto, como $p_{i_k} = 0$ y $f(\hat{p})$ es siempre positiva

$$\begin{aligned} f(\hat{p}_k) - p_{i_k} &= |f(\hat{p}_k) - p_{i_k}| \leq \frac{-w_i f'(\hat{p}_k)(G'_i(p_{i_k}) - G'_i(m(p_{-i_k})))}{G''_i(p_{i_k})} \\ &\leq \frac{|w_i f'(\hat{p}_k)(G'_i(p_{i_k}) - G'_i(m(p_{-i_k})))|}{G''_i(p_{i_k})} \end{aligned}$$

y el Teorema se sigue con la misma demostración que antes (notar que la desigualdad sigue valiendo porque $G''_i(p_{i_k}) > 0$).

Finalmente, si $p_{i_k} = 1$ entonces:

$$0 \leq G''_i(p_{i_k})(f(\hat{p}_k) - p_{i_k}) + w_i f'(\hat{p}_k)(G'_i(p_{i_k}) - G'_i(m(p_{-i_k})))$$

y por lo tanto

$$\frac{-w_i f'(\hat{p}_k)(G'_i(p_{i_k}) - G'_i(m(p_{-i_k})))}{G''_i(p_{i_k})} \leq f(\hat{p}_k) - p_{i_k}$$

donde de nuevo la desigualdad se preserva porque $G''_i(p_{i_k}) > 0$. Esta última desigualdad es equivalente a que

$$p_{i_k} - f(\hat{p}_k) \leq \frac{w_i f'(\hat{p}_k)(G'_i(p_{i_k}) - G'_i(m(p_{-i_k})))}{G''_i(p_{i_k})}$$

y como $p_{i_k} = 1$ y $f(\hat{p}) \leq 1$ para cualquier $\hat{p} \in [0, 1]$, entonces

$$\begin{aligned} p_{i_k} - f(\hat{p}_k) &= |p_{i_k} - f(\hat{p}_k)| \leq \frac{w_i f'(\hat{p}_k)(G'_i(p_{i_k}) - G'_i(m(p_{-i_k})))}{G''_i(p_{i_k})} \\ &\leq \frac{|w_i f'(\hat{p}_k)(G'_i(p_{i_k}) - G'_i(m(p_{-i_k})))|}{G''_i(p_{i_k})} \end{aligned}$$

y nuevamente el Teorema se sigue con la misma demostración que en el primer caso. \square

Lo que nos dice el Teorema es que todos los jugadores con pesos suficientemente chicos, es decir, que aportan poco capital o tienen poca influencia en el mercado, predirán $f(\hat{p})$ con alta precisión. Sin embargo, nada garantiza que \hat{p} sea un punto fijo de f o siquiera sea cercano a uno.

Observación 27. *Notar que para el Teorema anterior alcanza con pedir que f sea Lipschitz continua y diferenciable en \hat{p}_k , que G_i sea doblemente diferenciable en p_i , que $G''_i(p_i) \neq 0$, y que $C = \sup_{p_1, p_2 \in [0, 1]} \left| \frac{G'_i(p_1) - G'_i(p_2)}{G''_i(p_1)} \right| < \infty$.*

Observación 28. *Notar también que la elección de T_i podría modificarse preservando la validez del Teorema. Si bien tomamos la mediana de las predicciones de los demás jugadores para comparar con la predicción de un jugador en una cierta ronda, cualquier función m que tomemos sobre las predicciones de los demás jugadores no altera la validez del Teorema, ya que no altera el resultado al derivar.*

Para el Teorema anterior, definimos cómo obtener el puntaje para cada jugador en cada ronda, pero no definimos cuál es el puntaje final obtenido por cada jugador al finalizar

el juego. Sea entonces k la cantidad total de rondas jugadas, definimos el puntaje total obtenido por un jugador tras las k rondas como:

$$F_i(p_{i_1}, \dots, p_{i_k}) = \sum_{j=1}^k T_i(p_{i_j}, f(\hat{p}_j)) \quad (3.6)$$

Es decir, el puntaje final de cada jugador es simplemente la suma de los puntajes que obtuvo en cada ronda. Esto da lugar a la siguiente propiedad:

Teorema 29. *Sea G_i la función convexa asociada a la scoring rule del jugador i , S_i , proveniente del Teorema de caracterización 12, y sea k en número total de rondas realizadas. Si la función G_i es de clase C^2 , la función f es Lipschitz continua y diferenciable en \hat{p}_j para todo $1 \leq j \leq k$, el jugador i logra su máximo puntaje alcanzable tras jugarse las k rondas, y vale que $G_i''(p_{i_j}) \neq 0$ para algún $1 \leq j \leq k$, entonces para todo $\epsilon > 0$ existe un $\delta > 0$ tal que, si $w_i < \delta$, entonces $|\sum_{j=1}^k f(\hat{p}_j) - p_{i_j}| < \epsilon$.*

Demostración. Sea $p_i = (p_{i_1}, \dots, p_{i_k})$, si el puntaje alcanzado es un máximo de F_i , entonces $\nabla F_i(p_i) = 0$. Además, sea $1 \leq l \leq k$:

$$\frac{\partial}{\partial p_{i_l}} F_i(p_i) = \frac{\partial}{\partial p_{i_l}} \sum_{j=1}^k T_i(p_{i_j}, f(\hat{p}_j)) = \frac{\partial}{\partial p_{i_l}} T_i(p_{i_l}, f(\hat{p}_l)) = 0$$

Y por la fórmula 3.4 tenemos que

$$G_i''(p_{i_l})(f(\hat{p}_l) - p_{i_l}) + w_i f'(\hat{p}_l)(G_i'(p_{i_l}) - G_i'(m(p_{-i_l}))) = 0$$

Como esto vale para todo $1 \leq l \leq k$, vale que

$$\sum_{l=1}^k G_i''(p_{i_l})(f(\hat{p}_l) - p_{i_l}) + w_i f'(\hat{p}_l)(G_i'(p_{i_l}) - G_i'(m(p_{-i_l}))) = 0$$

Por lo tanto:

$$\left| \sum_{l=1}^k f(\hat{p}_l) - p_{i_l} \right| = \frac{\left| \sum_{l=1}^k w_i f'(\hat{p}_l)(G_i'(p_{i_l}) - G_i'(m(p_{-i_l}))) \right|}{\sum_{l=1}^k G_i''(p_{i_l})}$$

donde nuevamente no es necesario el módulo en el denominador porque G_i es convexa y C^2 , entonces $G_i''(p) > 0$ para todo p , y existe algún j tal que $G_i''(i_{i_j}) \neq 0$. Procediendo de forma similar al Teorema anterior, defino:

$$C = \max_{p_1, p_2 \in [0,1]} |G_i'(p_1) - G_i'(p_2)|$$

$$D = \min_{p \in [0,1]} G_i''(p)$$

Usando que f es Lipschitz continua, con $L > 0$ su constante de Lipschitz, y que $w_i < \delta$, tenemos que:

$$\frac{\left| \sum_{l=1}^k w_i f'(\hat{p}_l)(G_i'(p_{i_l}) - G_i'(m(p_{-i_l}))) \right|}{\sum_{l=1}^k G_i''(p_{i_l})} \leq \frac{k\delta CL}{kD} = \frac{\delta CL}{D}$$

y tomando $\delta = \frac{D\epsilon}{CL}$ se deduce el Teorema. \square

Lo que nos dice el Teorema es que, si el i -ésimo jugador predice de tal forma de obtener el máximo puntaje posible, entonces, si su peso es suficientemente chico, la suma total de las diferencias entre su probabilidad declarada y la probabilidad final en cada ronda es arbitrariamente chica.

3.2. Mercado de suma constante con predicciones únicas y simultáneas

Estudiaremos ahora el mismo formalismo, pero nuevamente considerando que hay una única ronda de predicciones, y que el mercado es de suma constante, es decir, que cumple la ecuación 2.7. En [11] se plantea que los mercados de predicción tienden a ser de suma cero, y si consideramos que la strictly proper scoring rule utilizada para puntuar la predicción de cada jugador es también el pago que dicho jugador recibe, entonces resulta razonable considerar el juego asociado a este mercado como de suma constante.

El ejemplo más simple de mercado de suma 0 es aquel donde el único dinero disponible es el que ponen los predictores mismos, y lo único que se decide es de qué forma se reparte entre ellos dependiendo de sus predicciones. Un caso de suma constante con una constante positiva podría ser si alguien pone una cierta cantidad de dinero inicialmente por fuera de la que ponen los jugadores, que luego se reparte entre ellos de acuerdo a sus predicciones. Por último, un caso de mercado de suma constante con una constante negativa podría darse, por ejemplo, si realizar las predicciones tiene un costo de transacción.

Sea $\hat{p} = \sum_{i=1}^n w_i p_i$, partimos entonces de que

$$\sum_{i=1}^n S_i(p_i, f(\hat{p})) = c \quad (3.7)$$

para alguna constante c . Además, como en el modelo anterior, seguimos considerando que todos los jugadores realizan las predicciones de manera simultánea.

Consideramos el caso donde cada jugador realiza una única predicción. Por el teorema de caracterización de strictly proper scoring rules 12, sabemos que para cada regla S_i existe una función convexa G_i tal que $S_i(p_i, f(\hat{p})) = G_i(p_i) + G_i^*(p_i)(f(\hat{p}) - p_i)$. Veremos ahora que, bajo ciertas hipótesis, si el peso de un jugador es suficientemente chico, entonces su predicción será cercana a la probabilidad real del evento predicho *después* de realizadas las predicciones.

Teorema 30. *Si cada una de las funciones G_i es de clase C^2 , la función f es diferenciable en \hat{p} y Lipschitz continua, y $G_j''(p_j) \neq 0$, entonces para todo $\epsilon > 0$ existe un $\delta > 0$ tal que si $w_j < \delta$ entonces $|f(\hat{p}) - p_j| < \epsilon$.*

Lo que nos dice el teorema es que, bajo ciertas condiciones, para los jugadores cuyos pesos son suficientemente chicos, la diferencia entre su predicción y la probabilidad del evento predicho *después* de realizadas las predicciones resulta arbitrariamente chica. Esto es interesante porque, si bien no podemos garantizar que las predicciones sean puntos fijos, al menos podemos acotar qué tan lejos están de serlo.

Demostración. Como el mercado es de suma constante, cumple la ecuación 3.7, y por lo tanto:

$$\frac{\partial}{\partial p_j} \sum_{i=1}^n S_i(p_i, f(\hat{p})) = 0$$

para todo $1 \leq j \leq n$. Sabemos además que

$$\frac{\partial}{\partial p_j} S_i(p_i, f(\hat{p})) = w_j G'_j(p_j) f'(\hat{p}) + G''_j(p_j)(f(\hat{p}) - p_j) \quad \text{si } i = j \quad (3.8)$$

$$\frac{\partial}{\partial p_j} S_i(p_i, f(\hat{p})) = w_j G'_i(p_i) f'(\hat{p}) \quad \text{si } i \neq j \quad (3.9)$$

Por lo tanto

$$0 = \frac{\partial}{\partial p_j} \sum_{i=1}^n S_i(p_i, f(\hat{p})) = \sum_{i=1}^n \frac{\partial}{\partial p_j} S_i(p_i, f(\hat{p})) \quad (3.10)$$

$$= \sum_{i=1}^n \frac{\partial}{\partial p_j} (G_i(p_i) + G'_i(p_i)(f(\hat{p}) - p_i)) \quad (3.11)$$

$$= G''_j(p_j)(f(\hat{p}) - p_j) + w_j \sum_{i=1}^n G'_i(p_i) f'(\hat{p}) \quad (3.12)$$

usando que, como G es C^2 entonces $G^* = G'$ y la fórmula anterior. Por lo tanto, reordenando:

$$|f(\hat{p}) - p_j| = \left| \frac{w_j \sum_{i=1}^n G'_i(p_i) f'(\hat{p})}{G''_j(p_j)} \right|$$

El pasaje es válido asumiendo que $G''_j(p_j) \neq 0$. Procedemos ahora de forma similar al Teorema 26: como f es diferenciable en \hat{p} y Lipschitz continua, entonces $|f'(\hat{p})| \leq L$ con $L > 0$ su constante de Lipschitz. Por otro lado, como cada G_i es C^2 , tanto G'_i como G''_i son continuas para todo i , y ambas están definidas sobre el compacto $[0, 1]$, por lo que alcanzan máximo y mínimo. Sean

$$C_i = \max_{p \in [0,1]} |G'_i(p)|$$

$$D = \min_{p \in [0,1]} |G''_j(p)|$$

y defino:

$$C = \max_{1 \leq i \leq n} C_i$$

Acotando $f'(\hat{p})$ por L , cada $G'_i(p_i)$ del numerador por C , w_j por $\delta > 0$ (que voy a elegir después convenientemente), y $G''_j(p_j)$ por D , obtengo que

$$\left| \frac{w_j \sum_{i=1}^n G'_i(p_i) f'(\hat{p})}{G''_j(p_j)} \right| < \frac{\delta \sum_{i=1}^n LC}{D} = \frac{n\delta LC}{D}$$

Sea entonces $\epsilon > 0$, si tomo $\delta = \frac{D\epsilon}{nLC}$ entonces

$$|f(\hat{p}) - p_j| = \left| \frac{w_j \sum_{i=1}^n G'_i(p_i) f'(\hat{p})}{G''_j(p_j)} \right| < \epsilon$$

□

Observación 31. *Notar que este Teorema vale para cualquier jugador y cualquier predicción conjunta que cumpla que $G''_j(p_j) \neq 0$, y no se restringe únicamente a equilibrios de Nash como el Teorema 14 de [15] o como el Teorema 26.*

3.3. Predicciones únicas y secuenciales

Partiendo del formalismo original, consideraremos ahora que el juego es **secuencial**, es decir, las predicciones se realizan de a una por vez en algún orden predeterminado en vez de todas simultáneamente. Asumimos que el juego es de información perfecta, es decir, cada jugador conoce en todo momento no solo la información respecto a la estructura del juego, esto es, las funciones de pago y posibles estrategias de los demás jugadores y la función f , sino que además, al momento de realizar su predicción, cada jugador conoce las predicciones hechas por los jugadores anteriores hasta ese momento.

3.3.1. Resultados analíticos

Veremos entonces que, bajo ciertas hipótesis, el resultado del juego está en realidad determinado de antemano.

Teorema 32. *Si tanto la función f como cada una de las funciones S_i son continuas, cada jugador juega racionalmente para maximizar su puntaje, y hay algún criterio determinístico y conocido por todos los jugadores para elegir una predicción cuando hay más de una que maximiza la función S_i para el jugador i , entonces existe un equilibrio de Nash perfecto en subjuegos.*

Notar que si se discretiza el espacio de posibles estrategias de cada jugador (por ejemplo, limitando la cantidad de decimales de las predicciones), entonces el resultado es una consecuencia directa del teorema 24, incluso sin las hipótesis pedidas. Sin embargo, el Teorema 32 en toda su generalidad no se deriva del Teorema 24, ya que, si bien el juego planteado es secuencial y de información perfecta, las posibles estrategias de cada jugador son infinitas.

Demostración. Para demostrarlo, aplicaremos la misma idea que para el teorema 24, es decir, utilizaremos inducción hacia atrás. Consideremos como fijas las estrategias de los primeros $n - 1$ jugadores. Como la función f y la función S_n son continuas, entonces su composición lo es, y por lo tanto la función $S_n(p_n, f(\sum_{i=1}^n w_i p_i))$ resulta continua. Como S_i está definida en el conjunto compacto $[0, 1] \times [0, 1]$, entonces alcanza máximo. La estrategia elegida por el jugador n será entonces

$$p_n^*(p_1, \dots, p_{n-1}) = \max_{p_n \in [0,1]} S_n(p_n, f(\sum_{i=1}^n w_i p_i))$$

Si hay más de una predicción máxima, entonces el jugador utilizará el criterio determinístico y conocido por todos para desempatar entre ellas. De esta forma, la estrategia elegida por el jugador n queda unívocamente determinada por las estrategias de los jugadores anteriores. Consideremos ahora el jugador $n - 1$, con el mismo razonamiento su estrategia será:

$$p_{n-1}^*(p_1, \dots, p_{n-2}) = \max_{p_{n-1} \in [0,1]} S_{n-1}(p_{n-1}, f(\sum_{i=1}^{n-1} w_i p_i + w_n p_n^*(p_1, \dots, p_{n-1})))$$

En general, la estrategia para el jugador k , con $1 \leq k \leq n$, será:

$$p_k^*(p_1, \dots, p_{k-1}) = \max_{p_k \in [0,1]} S_k(p_k, f(\sum_{i=1}^k w_i p_i + \sum_{i=k+1}^n w_i p_i^*(p_1, \dots, p_{i-1}))) \quad (3.13)$$

donde en cada paso la existencia del máximo está asegurada por ser S_k y f continuas definidas en un compacto, y en caso de haber más de un máximo se utiliza siempre un criterio determinístico y conocido por todos para desempatar entre ellos. El resultado de aplicar el método será un equilibrio de Nash perfecto en subjuegos. \square

Notar, sin embargo, que para poder implementar esta estrategia se debería primero discretizar el espacio de las predicciones permitidas para cada jugador, e incluso haciendo esto el costo temporal de computar una solución de esta forma puede resultar muy alto.

3.3.2. Resultados experimentales

A continuación presentaremos una serie de experimentos realizados sobre este modelo particular, es decir, al considerar el juego secuencial de información perfecta, de forma de estudiar propiedades sobre este modelo de forma empírica.

Simularemos un juego de 4 jugadores, donde cada uno realiza una única predicción y éstas se realizan en un orden predeterminado y conocido por todos. Asumiremos en principio que todos los jugadores son puntuados con la misma scoring rule, y que todos tienen el mismo peso en el mercado. El resto del modelo se mantiene igual al planteado en las secciones anteriores. La idea de los experimentos será simular el proceso de *inducción hacia atrás* descrito en la sección anterior, de forma que cada jugador realice la predicción que maximice su puntaje, y así observar algunas características que surgen a partir de este modelo. Evidentemente, puesto que el espacio de predicciones posibles de cada jugador es infinito, es necesario discretizarlo para llevar a cabo la simulación; tomamos entonces una granularidad de 1% en las predicciones permitidas para cada jugador.

3.3.2.1. Experimento 1.1: la importancia del orden - regla de Brier

Para este primer experimento, intentaremos evaluar si el orden en el que se realizan las predicciones afecta los puntajes obtenidos por los jugadores, siendo nuestra hipótesis inicial que los jugadores que predicen primero tienen ventaja por sobre los últimos, ya que, al ser el juego de información perfecta, pueden tomar la decisión de qué predicción dar de forma más libre que los últimos jugadores, que están condicionados por las predicciones realizadas por los jugadores anteriores.

Para hacer esto, puntuaremos a cada jugador con la misma scoring rule, la regla de Brier modificada ya presentada 2.5:

$$B'(p, q) = -(q(p - 1)^2 + (1 - q)p^2)$$

de forma que resulta orientada positivamente, y usaremos la siguiente f para modelar la influencia de las predicciones en la probabilidad de evento asociado, donde q es la probabilidad a priori del evento (antes de que se realicen las predicciones):

$$f(p, q) = c * p + (1 - c) * q \tag{3.14}$$

donde c es un factor que determina el poder de influencia de la predicción. Más precisamente, estaremos tomando un redondeo a dos decimales de la f dada de esta manera.

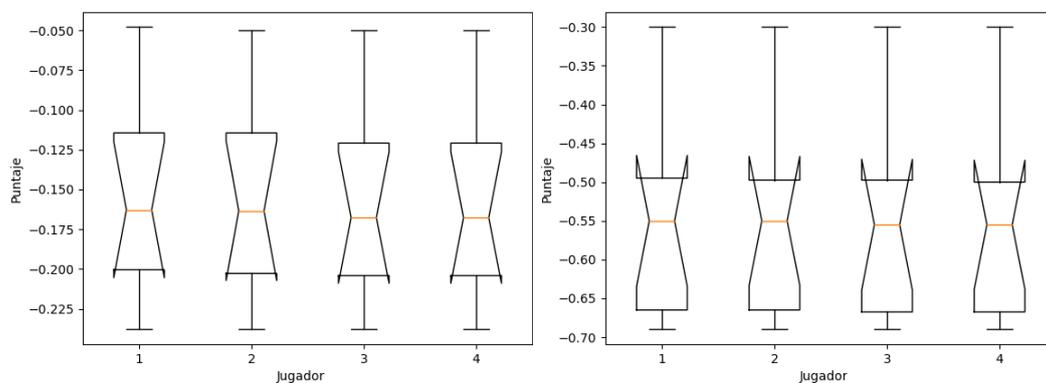
Para este experimento tomamos $c = 0,1$ (luego realizaremos otros experimentos para analizar diferencias en los resultados al variar el poder de influencia de la predicción). La idea es que la predicción mueve la probabilidad del evento más cerca de la propia

predicción, donde qué tan pronunciado es este movimiento depende del peso que le demos a la predicción. Esta f cumple con algunas propiedades esperables para la función que modela esta influencia, como ser simetría² y continuidad, salvo redondeo. En efecto, la simetría es porque:

$$\begin{aligned} 1 - f(1 - p, 1 - q) &= 1 - (c * (1 - p) + (1 - c)(1 - q)) \\ &= 1 - c + c * p - 1 + q + c - c * q \\ &= c * p + (1 - c) * q \\ &= f(p, q) \end{aligned}$$

Se realizaron 10 iteraciones del experimento, donde tomamos el q inicial, esto es, la probabilidad inicial del evento sobre el cual se predice, de manera aleatoria, tomando luego el promedio entre todos los resultados obtenidos. La elección de esta metodología se debe a que el tiempo de cómputo necesario para, por ejemplo, ejecutar todos los experimentos con todos los q posibles con granularidad 0.5 (que puede considerarse un mínimo de granularidad necesaria para que los resultados sean realmente representativos), resultaría demasiado grande. De esta forma buscamos entonces reducir la influencia de la elección de la probabilidad inicial en los resultados. Los resultados obtenidos pueden observarse en la figura 3.1a y en la tabla 5.1 en el apéndice B.

Los resultados se condicen así con la hipótesis planteada. Se puede ver que el promedio del puntaje obtenido decrece a medida que los jugadores realizan sus predicciones más tarde en la ronda, siendo el último jugador el de menor puntaje en promedio.



(a) Boxplot de los puntajes obtenido por cada jugador usando la regla de Brier (b) Boxplot de los puntajes obtenido por cada jugador usando la regla logarítmica

Fig. 3.1: Resultados con $c = 0, 1$

3.3.2.2. Experimento 1.2: la importancia del orden - regla logarítmica

A continuación repetimos el experimento descrito en la sección anterior, pero puntuando a cada jugador con la regla logarítmica ya presentada 2.3:

$$\text{logscore}(p, q) = q \log p + (1 - q) \log (1 - p)$$

² En este contexto, simétrica significa que $f(p, q) = 1 - f(1 - p, 1 - q)$.

en lugar de la regla de Brier, de forma de observar cómo varían los resultados al cambiar la scoring rule utilizada, sin modificar ninguna otra condición de las presentadas en la sección anterior. Para la implementación de la regla logarítmica, consideramos que si la predicción realizada fue 1 ó 0, y la verdadera probabilidad del evento no es ni 1 ni 0, entonces el puntaje es $-\infty$. Los resultados obtenidos con esta regla fueron muy similares a los obtenidos con la regla de Brier, mostrando una pequeña ventaja para los primeros jugadores, como se observa en la figura 3.1b y en la tabla 5.2 en el apéndice B.

3.3.2.3. Experimento 2.1: la importancia del orden al aumentar la influencia de las predicciones

A continuación repetimos los experimentos presentados en las secciones anteriores, pero variando el valor dado a la constante c para la función f . Probamos con los valores $c = 0,15$ y $c = 0,2$, tanto para la regla de Brier como para la logarítmica. Los resultados pueden observarse en las figuras 3.2 y 3.3, y en la sección C del apéndice. Exceptuando los resultados obtenidos para la regla de Brier con $c = 0,2$, no se observaron grandes diferencias con los resultados presentados en las secciones anteriores. El caso particular de la regla de Brier con $c = 0,2$ fue el único donde se observó un leve aumento en la diferencia entre los jugadores, siendo mayor la ventaja obtenida especialmente por el primer jugador respecto a los demás. Conjeturamos que estos resultados pueden deberse por un lado, a que el aumento de la constante c es pequeño, con lo cual la diferencia con respecto a los casos anteriores también lo es, y por otro lado porque la elección aleatoria de la probabilidad a priori del evento (a la que antes llamamos q) puede influir en los resultados obtenidos.

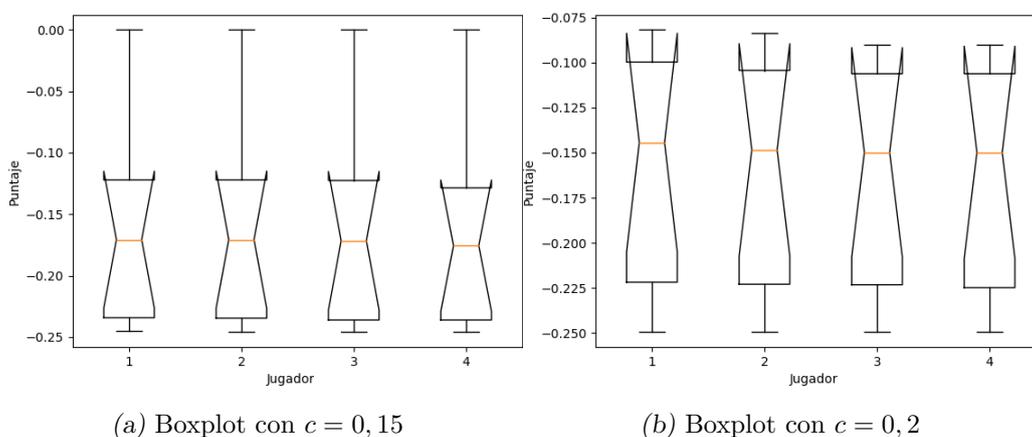


Fig. 3.2: Resultados usando la regla de Brier

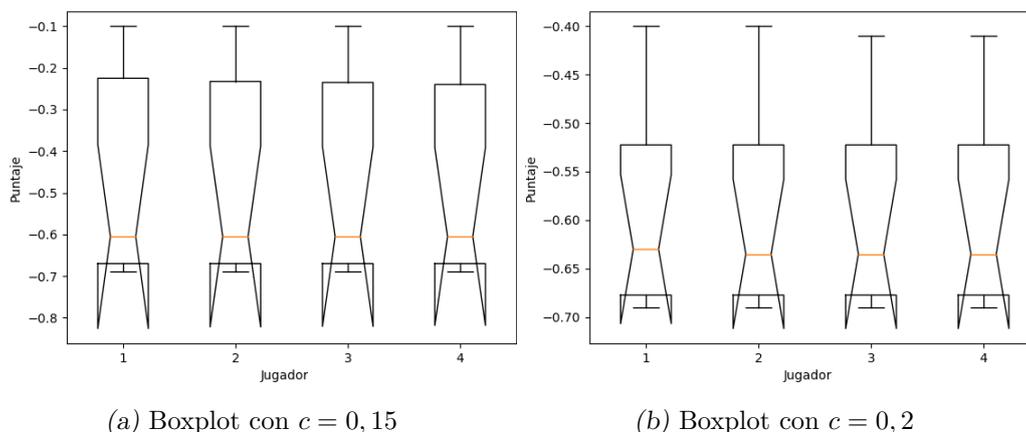


Fig. 3.3: Resultados usando la regla logarítmica

Para intentar dilucidar los resultados obtenidos y analizar la influencia del q original elegido, realizamos un nuevo experimento para que nos brinde información más completa sobre lo que está sucediendo.

3.3.2.4. Experimento 2.2: resultados completos usando la regla de Brier

Como dijimos antes, para realizar un análisis más completo corrimos la simulación variando q entre 0 y 1 con una granularidad de 0,5, de forma de observar tanto los puntajes obtenidos por cada jugador para cada q como la predicción realizada por cada jugador para cada q . Para esto, tomamos la misma f que en los experimentos anteriores 3.14, variando la constante c (con los valores 0,1, 0,15 y 0,2), asumiendo que todos los jugadores tienen el mismo peso en el mercado, y usando en principio la regla de Brier, aunque luego repetimos el experimento con la regla logarítmica para realizar una comparación entre los resultados con ambas reglas. Los datos completos para la regla de Brier obtenidos pueden encontrarse en el apéndice E.

Lo primero que podemos remarcar de los resultados es que los puntajes de todos los jugadores aumentan consistentemente al aumentar el valor la constante c . Este fenómeno resulta esperable, teniendo en cuenta que al aumentar el valor de c aumenta también la influencia de las predicciones en la probabilidad final del evento, dándole más posibilidades a los jugadores de manipularla para obtener un mayor puntaje. Otra evidencia de esto es que la diferencia entre la probabilidad final del evento y la probabilidad original, si bien es pequeña, aumenta al aumentar c , como puede verse en las tablas en el apéndice E y en las figuras 3.4 y del apéndice 5.1, donde para ilustrar mejor la situación vemos, para cada valor para c entre 0 y 1 con granularidad 0,5, la diferencia entre la probabilidad original (q) y la probabilidad final tras realizar las predicciones.

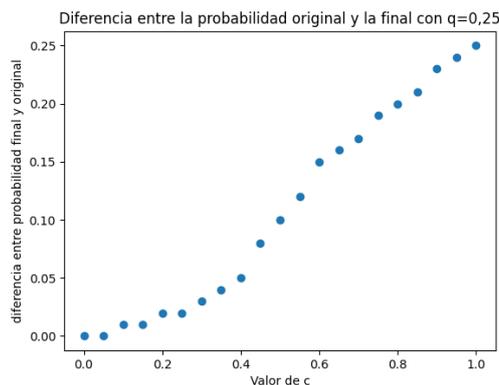


Fig. 3.4: Diferencia entre la probabilidad original y la probabilidad final al variar c con $q = 0,25$.

Otro fenómeno interesante que podemos observar es que tanto los valores más extremos de q (0 y 1) como los valores intermedios (cerca de 0,5) resultan los menos manipulables (notar que para los valores 0, 1 y 0,5 todos los jugadores predicen siempre la probabilidad real del evento). Al alejarnos de estos valores, las predicciones empiezan a diferenciarse de la probabilidad real q , intentando siempre volver más extrema la probabilidad del evento. Finalmente, notamos un fenómeno interesante y quizás no muy intuitivo, que es que en la mayoría de los casos, las predicciones de los últimos jugadores son más lejanas a la probabilidad original que la de los primeros. Además, también sucede que, en todos los casos, el jugador cuya predicción está más cercana a la probabilidad final es el primero. Incluso sucede en la mayoría de los casos que la predicción del primer jugador es más cercana a la probabilidad final que la predicción final del mercado, siendo así el puntaje del primer jugador mayor al obtenido por la predicción del mercado, aunque el obtenido por la predicción del mercado es mayor al puntaje de los últimos jugadores en la mayoría de los casos. Creemos que esto puede deberse a que, si el primero realizara una predicción más extrema, más alejada de la probabilidad original, sería más susceptible a que los siguientes jugadores manipulen la probabilidad del evento en la dirección contraria, haciendo que la predicción del primero se aleje de la probabilidad final y por ende baje su puntaje. Además, las predicciones más extremas de los últimos jugadores hacen que la predicción final del mercado se aleje también de la probabilidad final. Esto se condice con el hecho de que, a mayor valor de c , más extremas son las predicciones de los últimos jugadores, y peor es el puntaje de la predicción del mercado. Para visualizar mejor la situación, realizamos una simulación para algunos valores particulares de q fijando las predicciones de los primeros tres jugadores en las obtenidas en la simulación original, de forma de observar el puntaje para cada posible predicción del último jugador. En la figura 3.5 pueden verse los resultados obtenidos para los distintos valores de c con $q = 0,25$, en el apéndice E pueden verse resultados para otros valores de q , que resultaron muy similares. Vemos que al último jugador le conviene extremizar la probabilidad final del evento, es decir, intentar llevar a un valor más cercano a 0 en este caso, pero sin que su predicción quede demasiado alejada de la probabilidad real. Es por eso que su puntaje crece hasta llegar a un máximo en un punto menor a la probabilidad original (de forma de llevarla hacia 0), y luego comienza a decrecer a medida que se aleja de dicho punto. Además, como ya marcamos antes, podemos ver como este punto máximo se aleja de la probabilidad

real del evento a medida que aumentamos la influencia de las predicciones (es decir, al aumentar el valor de c).

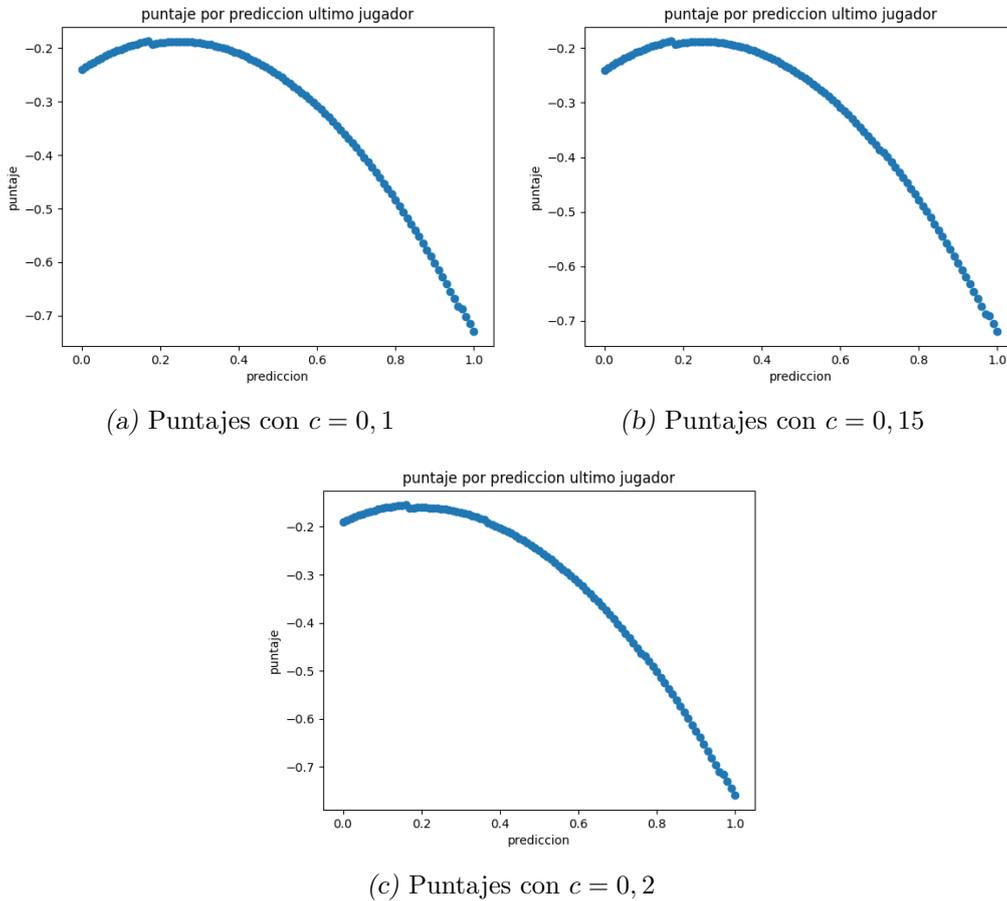


Fig. 3.5: Puntajes del último jugador con $q = 0,25$

Por último, realizamos la misma simulación para obtener resultados completos, pero utilizando la regla logarítmica en lugar de la regla de Brier para puntuar a los jugadores, de forma de ver el efecto de la elección de la scoring rule en los resultados. Los resultados pueden verse en las tablas del apéndice F. Puede verse que éstos son muy similares a los obtenidos para la regla de Brier. Nuevamente vemos como, a mayor valor de la constante c , los puntajes de todos los jugadores también aumentan. De la misma forma, a mayor valor de c aumenta también la distancia entre la probabilidad final del evento y la probabilidad original para algunos valores de q . También vemos que, como antes, en los valores extremos (0 y 1) e intermedio (0,5) las predicciones coinciden con la probabilidad del evento, mientras que al alejarse de dichos valores empiezan a aparecer las diferencias. Como última similitud podemos observar que, en la mayoría de los casos, las predicciones de los últimos jugadores se alejan más tanto de la probabilidad original del evento predicho como de la probabilidad final, y que el comportamiento de la predicción y el puntaje del mercado es muy similar al observado para la regla de Brier.

La principal diferencia que encontramos en los resultados, y que es resultado de una de las diferencias entre ambas reglas, es que, mientras que al utilizar la regla de Brier había algunos valores de q distintos de 0 y 1 para los cuales los últimos jugadores eligen predecir 0 ó 1, en la regla logarítmica esto no sucede. Con la regla logarítmica el único caso donde los jugadores eligen predecir 0 es cuando ésta es la probabilidad real del evento, y lo mismo sucede para 1. La razón para esto es que la regla logarítmica penaliza muy fuertemente (da menos infinito) si se da una predicción extrema (0 ó 1) que no coincide con la probabilidad del evento, mientras que la regla de Brier no. Esta es una ventaja que presenta utilizar la regla logarítmica, ya que es altamente deseable no incentivar predicciones tan extremas a menos que se tenga mucha seguridad de ellas.

3.3.2.5. Experimento 3: la importancia del peso de los jugadores

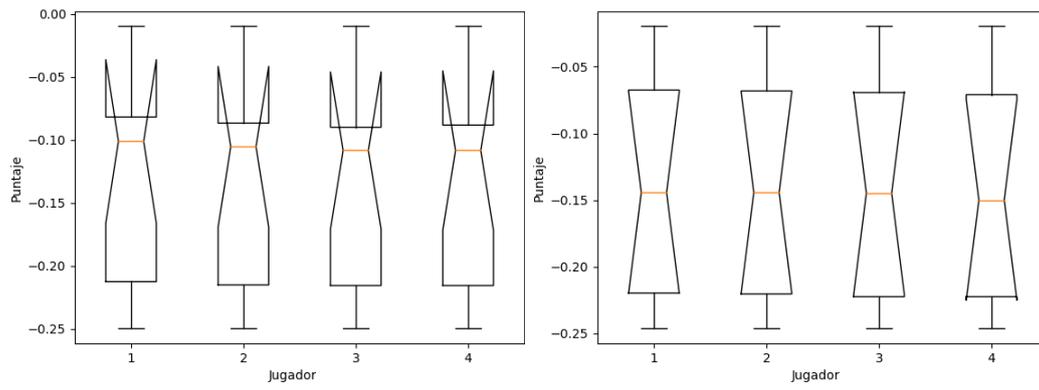
En todos los experimentos realizados hasta el momento asumíamos que todos los jugadores tenían el mismo peso en el mercado. A continuación analizaremos la influencia de variar los pesos, tanto al darle más peso a los primeros jugadores como a los últimos. La función f utilizada es la misma que en los experimentos anteriores 3.14:

$$f(p) = c * p + (1 - c) * q$$

con constante $c = 0, 1$, considerando que todos los jugadores son puntuados con la regla de Brier. Los resultados numéricos se encuentran en el apéndice D.

Empezamos analizando el caso donde le damos más peso a los primeros jugadores. Al tomar los pesos $[0,5, 0,2, 0,2, 0,1]$, se puede observar claramente en la figura 3.6a un aumento en la diferencia entre el puntaje obtenido por el primer y el segundo jugador, que al considerar pesos iguales se mantenían más cercanos. Este resultado se condice con lo esperado, ya que al tener más peso el primer jugador tiene más poder para mover la predicción final del mercado en una dirección particular y así manipularlo para obtener un mayor puntaje.

Por otro lado, al tomar los pesos $[0,1, 0,2, 0,2, 0,5]$ se ve en la figura 3.6b como se achica considerablemente la diferencia entre los jugadores. Es interesante observar que, si bien la diferencia entre los primeros tres jugadores se achica considerablemente respecto a los demás experimentos, el último jugador sigue teniendo una desventaja claramente observable, obteniendo un puntaje más bajo que el resto, incluso siendo el jugador de más peso. Esto puede deberse a que, a pesar de su mayor peso, el hecho de predecir último condiciona mucho su predicción, sacándole la posibilidad de usar su peso para mejorar su puntaje, como si podía hacerlo el primer jugador en el experimento anterior.



(a) Boxplot con $pesos = [0,5, 0,2, 0,2, 0,1]$ (b) Boxplot con $pesos = [0,1, 0,2, 0,2, 0,5]$

Fig. 3.6: Resultados al variar los pesos de los jugadores

4. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo estudiamos el problema de las *profecías autocumplidas* en el contexto de *mercados de predicción*, y el fenómeno de extremalización de eventos a través de las predicciones. Para esto, modelamos un mercado como un juego de n jugadores, y utilizamos herramientas de la *teoría de juegos* para obtener ciertas propiedades respecto a las predicciones realizadas por los jugadores en dicho modelo, así como en los juegos que surgen de introducir pequeñas modificaciones en el modelo original. Estudiamos así ciertas propiedades de los equilibrios de Nash y de las predicciones que llevan a los mejores puntajes en el caso donde las predicciones se realizan de manera simultánea, permitiendo varias rondas de predicciones 3.1. Por otro lado, estudiamos también qué tan buenas son las predicciones en el caso donde estas se realizan de manera simultánea, cada jugador realiza una única predicción y el mercado es de suma constante 3.2. Por último, analizamos las mejores estrategias en el caso donde cada jugador da una única predicción, pero estas se realizan de forma secuencial, asumiendo información perfecta 3.3.

Por otro lado, realizamos un análisis experimental para estudiar ciertas propiedades de forma empírica 3.3.2. Partiendo del modelo donde cada jugador realiza una única predicción y estas se realizan de forma secuencial, asumiendo información perfecta, igualdad de pesos entre todos los jugadores y una única scoring rule para todos (alternativamente Brier y logarítmica), y dimos una implementación de este modelo discretizando el espacio de predicciones posibles y tomando cuatro participantes. El primer resultado importante que obtuvimos es que el orden en el que se realizan las predicciones es relevante, ya que vimos que predecir antes da una mayor posibilidad de obtener mayor puntaje. Vimos también que al aumentar el peso del primer jugador, la diferencia en los puntajes obtenidos también aumenta, y que al aumentar el peso del último jugador y disminuir el de los primeros, la brecha se acorta considerablemente, especialmente entre los primeros tres jugadores.

Por otra parte, observamos que el aumentar la influencia de las predicciones en la probabilidad del evento tiene dos efectos importantes: por un lado todos los jugadores obtienen consistentemente un mayor puntaje, y por otro lado, que también aumenta la diferencia entre la probabilidad final del evento tras realizadas las predicciones y la predicción original.

En cuanto a la diferencia entre las predicciones de los jugadores y la probabilidad a priori del evento, notamos que los últimos jugadores son más proclives a realizar predicciones más extremas, más lejanas tanto de la probabilidad a priori del evento como de su probabilidad final, mientras que los primeros jugadores, especialmente el primer jugador, realizan predicciones más cercanas a la probabilidad real, y que en muchos casos coinciden con la probabilidad final del evento tras las predicciones. Vimos también que la predicción final del mercado suele estar más lejos de la probabilidad final del evento que la predicción del primer jugador, pero más cercana que las predicciones de los últimos jugadores. Es interesante remarcar que los resultados obtenidos fueron muy similares para la regla de Brier y la regla logarítmica. La principal diferencia que notamos es que la regla logarítmica castiga más duramente las predicciones más extremas (0 y 1) cuando son equivocadas, por lo que ningún jugador predice 0 ó 1 a menos que esa sea la probabilidad real del evento, mientras que con la regla de Brier hay casos donde la probabilidad original del evento no es

ni 0 ni 1, y, sin embargo, a los últimos jugadores les conviene realizar predicciones extremas.

Al ser un tema con múltiples parámetros que pueden variar, y donde pequeños cambios en el modelo llevan a fuertes cambios en sus propiedades, existen múltiples caminos posibles para trabajos futuros. Por el lado de las simulaciones, es importante remarcar que dentro del modelo de trabajo que tomamos, realizamos una asunción muy importante y restrictiva, y es que todos los jugadores conocen cuál será el efecto de su predicción en la probabilidad del evento predicho. Sería interesante estudiar el caso donde se relaje esta suposición, y trabajar casos donde los jugadores no sepan cuál es el efecto exacto de sus predicciones, sino que tengan una creencia a priori e intenten maximizar la esperanza de su puntaje. Resultaría también interesante estudiar otras formas de agregación para obtener la predicción final del mercado, como por ejemplo la mediana, y ver cómo esta elección impacta en las predicciones de cada jugador. Otras posibles variaciones del modelo podrían ser que los jugadores sean puntuados por scoring rules distintas, cambiar la función f que modela la influencia de las predicciones en la probabilidad del evento, o considerar el caso donde las predicciones se realizan de forma simultánea en lugar de secuencial.

Otro posible camino para un trabajo futuro es estudiar el caso donde las predicciones no sean binarias, sino predicciones sobre espacios muestrales discretos, o incluso espacios muestrales continuos. También podría ser interesante estudiar el efecto de relajar las hipótesis de información perfecta o información completa en los distintos modelos, especialmente, como mencionamos antes, la hipótesis de que todos los jugadores conocen la función f y saben de forma exacta el efecto de sus predicciones sobre la probabilidad a priori del evento, que resulta una hipótesis muy fuerte; o, por otro lado, relajar las condiciones de diferenciabilidad y continuidad de las distintas funciones involucradas en el modelo.

5. APÉNDICE

A. Implementación

Todo el código con la implementación del modelo utilizado para las simulaciones y la generación de figuras puede encontrarse en el siguiente repositorio público:
<https://gitlab.com/tschitter/prediction-markets-and-self-fulfilling-prophecies>.

El código desarrollado implementa de forma directa el proceso de inducción hacia atrás, utilizando fuerza bruta para hacerlo. No fue el objetivo de esta tesis trabajar en optimizar el código, y el rendimiento obtenido por el algoritmo de fuerza bruta resultó suficientemente bueno para llevar a cabo los experimentos.

Dentro del código, la parte más interesante y costosa en términos de complejidad es justamente el cálculo de la mejor predicción posible para cada jugador. Dicha función tiene un caso base, que es cuando estamos calculando la mejor predicción del último jugador, es decir, donde los primeros $N - 1$ jugadores ya realizaron sus predicciones, donde simplemente verifica el puntaje obtenido por cada una de las predicciones posibles, y devuelve la que mejor puntaje obtenga. En el caso recursivo, es decir, si estamos calculando la mejor predicción del jugador $i < N$, nuevamente prueba todas las predicciones posibles, pero ahora, por cada una, llama recursivamente a la función para calcular la mejor predicción posible del jugador $i + 1$, habiendo fijado ya la del jugador anterior. De esta forma, para cada posible predicción p_i del jugador i encuentra la mejor predicción para cada uno de los jugadores $i + 1, \dots, N$, y en base a dichas predicciones calcula el puntaje que obtendría la predicción p_i (esto es porque sabemos que cada jugador elige siempre la predicción que maximice su puntaje, y en caso de empate la predicción más pequeña). Finalmente, el jugador i elige la predicción que maximice su puntaje. Si comenzamos este procedimiento desde el jugador 1, obtendremos la mejor predicción posible para cada jugador.

En términos generales, podemos decir que el código es simplemente una implementación recursiva de la estrategia descrita en 3.13 al discretizar el espacio de posibles predicciones.

B. Experimento 1: la importancia del orden

Jugadores	1	2	3	4
Promedio	-0,153	-0,154	-0,157	-0,157
Desvío Estándar	0,063	0,063	0,063	0,063

Tab. 5.1: Experimento 1.1: promedio y desvío estándar del puntaje de cada jugador, usando la regla de Brier.

Jugadores	1	2	3	4
Promedio	-0,547	-0,549	-0,551	-0,552
Desvío Estándar	0,13	0,128	0,129	0,129

Tab. 5.2: Experimento 1.2: promedio y desvío estándar del puntaje de cada jugador, usando la regla logarítmica.

C. Experimento 2: la importancia del orden al aumentar la influencia de las predicciones

Jugadores	1	2	3	4
Promedio	-0,156	-0,1561	-0,157	-0,159
Desvío Estándar	0,09	0,09	0,091	0,09

Tab. 5.3: Promedio y desvío estándar del puntaje de cada jugador, usando la regla de Brier y $c = 0,15$.

Jugadores	1	2	3	4
Promedio	-0,1568	-0,1594	-0,161	-0,162
Desvío Estándar	0,067	0,066	0,064	0,064

Tab. 5.4: Promedio y desvío estándar del puntaje de cada jugador, usando la regla de Brier y $c = 0,2$.

Jugadores	1	2	3	4
Promedio	-0,468	-0,469	-0,47	-0,472
Desvío Estándar	0,244	0,243	0,242	0,241

Tab. 5.5: Promedio y desvío estándar del puntaje de cada jugador, usando la regla logarítmica y $c = 0,15$.

Jugadores	1	2	3	4
Promedio	-0,591	-0,593	-0,595	-0,595
Desvío Estándar	0,107	0,106	0,103	0,103

Tab. 5.6: Promedio y desvío estándar del puntaje de cada jugador, usando la regla logarítmica y $c = 0,2$.

D. Experimento 3: la importancia del peso de los jugadores

Jugadores	1	2	3	4
Promedio	-0,1349	-0,1372	-0,1388	-0,1385
Desvío Estándar	0,086	0,086	0,085	0,085

Tab. 5.7: Promedio y desvío estándar del puntaje de cada jugador con $pesos = [0,5, 0,2, 0,2, 0,1]$.

Jugadores	1	2	3	4
Promedio	-0,1435	-0,1437	-0,1445	-0,1462
Desvío Estándar	0,0879	0,0879	0,0881	0,0881

Tab. 5.8: Promedio y desvío estándar del puntaje de cada jugador con $pesos = [0,1, 0,2, 0,2, 0,5]$.

E. Datos completos regla de Brier

A continuación presentamos los puntajes obtenidos por cada jugador y las predicciones realizadas por cada jugador para cada q entre 0 y 1 con granularidad de 0,05, al usar la regla de Brier para puntarlos, manteniendo siempre el mismo peso para todos y variando el valor de la constante c de la función f , donde la f es la siguiente función (con redondeo a dos decimales):

$$f(p) = c * p + (1 - c) * q$$

Recordemos además que, en este contexto donde todos los jugadores tienen el mismo peso, la predicción final del mercado se calcula como:

$$\hat{p} = \frac{1}{N} \sum_{i=1}^N p_i$$

Donde N es la cantidad total de jugadores y p_i la predicción de cada jugador. Recordemos por último que, para esta simulación, el puntaje obtenido por cada jugador estará dado por:

$$Brier(p_i, f(\hat{p}, q))$$

Presentamos entonces los resultados obtenidos para $c = 0,1$, $c = 0,15$ y $c = 0,2$.

$c = 0,1$

q	puntaje 1	puntaje 2	puntaje 3	puntaje 4	puntaje mercado
0.0	0.0	0.0	0.0	0.0	0.0
0.05	-0.0475	-0.0475	-0.0475	-0.0475	-0.0475
0.1	-0.0819	-0.0819	-0.0883	-0.09	-0.0835
0.15	-0.1204	-0.1205	-0.1268	-0.1268	-0.122
0.2	-0.1539	-0.1548	-0.1588	-0.1588	-0.1555
0.25	-0.1824	-0.1828	-0.1873	-0.1873	-0.184
0.3	-0.2059	-0.2084	-0.2095	-0.2095	-0.2075
0.35	-0.2245	-0.2269	-0.2269	-0.2269	-0.226
0.4 ¹	-0.24	-0.24	-0.24	-0.24	-0.24
0.45	-0.2475	-0.2475	-0.2475	-0.2475	-0.2475
0.5	-0.25	-0.25	-0.25	-0.25	-0.25
0.55	-0.2475	-0.2475	-0.2475	-0.2475	-0.2475
0.6	-0.24	-0.24	-0.24	-0.24	-0.24
0.65	-0.2245	-0.2269	-0.2269	-0.2269	-0.226
0.7	-0.2059	-0.2075	-0.2095	-0.2095	-0.2075
0.75	-0.1824	-0.1828	-0.1873	-0.1873	-0.184
0.8	-0.1539	-0.1543	-0.1588	-0.1588	0.1555
0.85	-0.1204	-0.1205	-0.1268	-0.1268	-0.122
0.9	-0.0819	-0.0819	-0.0868 ²	-0.09	-0.0835
0.95	-0.0475	-0.0475	-0.0475	-0.0475	-0.0475
1.0	0.0	0.0	0.0	0.0	0.0

Tab. 5.9: Puntaje obtenido por cada jugador con $c = 0,1$.

¹ Notar que entre 0,4 y 0,6 todos los jugadores obtuvieron el mismo puntaje. En la siguiente table se ve que en dichos valores todos predijeron la verdadera probabilidad del evento. También puede verse luego que esto deja de suceder al aumentar el valor de c .

² Notar que los puntajes deberían ser simétricos, aunque en este caso (y en algunos más en las siguientes tablas) presentan pequeñas diferencias con lo esperado. Verificamos que esto se debe al redondeo realizado en las predicciones.

q	predicción 1	predicción 2	predicción 3	predicción 4	predicción mercado	probabilidad final
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.1	0.09	0.09	0.01	0 ³	0.05	0.09
0.15	0.14	0.13	0.06	0.06	0.1	0.14
0.2	0.19	0.16	0.12	0.12	0.15	0.19
0.25	0.24	0.22	0.17	0.17	0.2	0.24
0.3	0.29	0.24	0.23	0.23	0.25	0.29
0.35	0.33	0.29	0.29	0.29	0.3	0.34
0.4	0.4	0.4	0.4	0.4	0.4	0.4
0.45	0.45	0.45	0.45	0.45	0.45	0.45
0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.55	0.55	0.55	0.55	0.55	0.55	0.55
0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.65	0.67	0.71	0.71	0.71	0.7	0.66
0.7	0.71	0.75	0.77	0.77	0.75	0.71
0.75	0.76	0.78	0.83	0.83	0.8	0.76
0.8	0.81	0.83	0.88	0.88	0.85	0.81
0.85	0.86	0.87	0.94	0.94	0.9	0.86
0.9	0.91	0.91	0.98 ⁴	1.0 ⁵	0.95	0.91
0.95	0.95	0.95	0.95	0.95	0.95	0.95
1.0	1.0	1.0	1.0	1.0	1.0	1.0

Tab. 5.10: Predicción realizada por cada jugador con $c = 0, 1$.

³ Notar que estas predicciones extremas son sumamente indeseables, y, como se ve después, no suceden al utilizar la regla logarítmica

⁴ Notar que, como en la tabla anterior, las predicciones no son perfectamente simétricas, debido al redondeo.

⁵ Ver nota 3.

$c = 0,15$

q	puntaje 1	puntaje 2	puntaje 3	puntaje 4	puntaje mercado
0.0	0.0	0.0	0.0	0.0	0.0
0.05	-0.0384	-0.0388	-0.04	-0.04	-0.0393
0.1	-0.0819	-0.0819	-0.082	-0.09	-0.0823
0.15	-0.1204	-0.1204	-0.1208	-0.1268	-0.1208
0.2	-0.1539	-0.1539	-0.1548	-0.1588	-0.1543
0.25	-0.1824	-0.1824	-0.1833	-0.1873	-0.1833
0.3	-0.2059	-0.2059	-0.2075	-0.2095	-0.2063
0.35	-0.2244	-0.2244	-0.2269	-0.2269	-0.2248
0.4 ⁶	-0.2379	-0.2383	-0.2395	-0.2395	-0.2388
0.45	-0.2465	-0.2473	-0.2473	-0.2473	-0.2473
0.5	-0.25	-0.25	-0.25	-0.25	-0.25
0.55	-0.2465	-0.2473	-0.2473	-0.2473	-0.2468
0.6	-0.2379	-0.2383	-0.2395	-0.2395	-0.2388
0.65	-0.2244	-0.2244	-0.2269	-0.2269	-0.2253
0.7	-0.2059	-0.2059	-0.2075	-0.2095	-0.2063
0.75	-0.1824	-0.1824	-0.1833	-0.1873	-0.1833
0.8	-0.1539	-0.1539	-0.1548	-0.1588	-0.1543
0.85	-0.1204	-0.1204	-0.1208	-0.1268	-0.1213
0.9	-0.08	-0.08	-0.08	-0.08	-0.08
0.95	-0.0384	-0.0388	-0.04	-0.04	-0.0388
1.0	0.0	0.0	0.0	0.0	0.0

Tab. 5.11: Puntaje obtenido por cada jugador con $c = 0,15$.

⁶ Notar que, como remarcamos antes, ya para $c = 0,15$ se empiezan a ver diferencias en los puntajes obtenidos por cada jugador entre los valores 0,4 y 0,6 de q , excepto para $q = 0,5$, que es el valor de máxima incertidumbre.

q	predicción 1	predicción 2	predicción 3	predicción 4	predicción mercado	probabilidad final
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.05	0.04	0.02	0	0	0.01	0.04
0.1	0.09	0.09	0.08	0	0.07	0.09
0.15	0.14	0.14	0.12	0.06	0.12	0.14
0.2	0.19	0.19	0.16	0.12	0.17	0.19
0.25	0.24	0.24	0.21	0.17	0.21	0.24
0.3	0.29	0.29	0.25	0.23	0.27	0.29
0.35	0.34	0.34	0.29	0.29	0.32	0.34
0.4	0.39	0.37	0.35	0.35	0.36	0.39
0.45	0.43	0.41	0.41	0.41	0.41	0.44
0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.55	0.57	0.59	0.59	0.59	0.58	0.56
0.6	0.61	0.63	0.65	0.65	0.64	0.61
0.65	0.66	0.66	0.71	0.71	0.69	0.66
0.7	0.71	0.71	0.75	0.77	0.73	0.71
0.75	0.76	0.76	0.79	0.83	0.79	0.76
0.8	0.81	0.81	0.84	0.88	0.83	0.81
0.85	0.86	0.86	0.88	0.94	0.89	0.86
0.9	1.0	1.0	1.0	1.0	1.0	0.92
0.95	0.96	0.98	1.0	1.0	0.98	0.96
1.0	1.0	1.0	1.0	1.0	1.0	1.0

Tab. 5.12: Predicción realizada por cada jugador con $c = 0, 15$. $c = 0, 2$

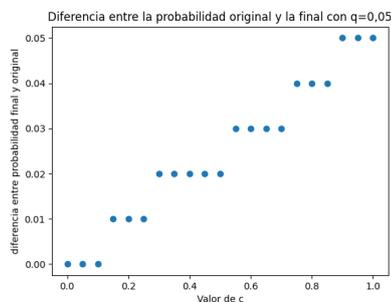
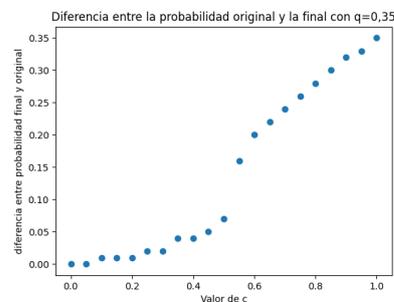
q	puntaje 1	puntaje 2	puntaje 3	puntaje 4	puntaje mercado
0.0	0.0	0.0	0.0	0.0	0.0
0.05	-0.0384	-0.0384	-0.0393	-0.04	-0.0388
0.1	-0.0736	-0.0785	-0.08	-0.08	-0.0772
0.15	-0.1131	-0.118	-0.1195	-0.1195	-0.1167
0.2	-0.148	-0.1525	-0.1525	-0.1525	-0.1512
0.25	-0.1775	-0.182	-0.182	-0.182	-0.1807
0.3	-0.2032	-0.2052	-0.2052	-0.2052	-0.2052
0.35	-0.2244	-0.2244	-0.2245	-0.2269	-0.2245
0.4	-0.2379	-0.2379	-0.2388	-0.2395	-0.2383
0.45	-0.2464	-0.2465	-0.2473	-0.2473	-0.2468
0.5	-0.25	-0.25	-0.25	-0.25	-0.25
0.55	-0.2464	-0.2464	-0.2473	-0.2473	-0.2465
0.6	-0.2379	-0.2379	-0.2388	-0.2395	-0.2383
0.65	-0.2244	-0.2244	-0.2245	-0.2269	-0.2248
0.7	-0.2041	-0.2052	-0.2052	-0.2052	-0.2052
0.75	-0.1772	-0.182	-0.182	-0.182	-0.1796
0.8	-0.1477	-0.1525	-0.1525	-0.1525	-0.1512
0.85	-0.1131	-0.1167	-0.1195	-0.1195	-0.1167
0.9	-0.0736	-0.0772	-0.08	-0.08	-0.0761
0.95	-0.0384	-0.0384	-0.0388	-0.04	-0.0385
1.0	0.0	0.0	0.0	0.0	0.0

Tab. 5.13: Puntaje obtenido por cada jugador con $c = 0, 2$.

q	predicción 1	predicción 2	predicción 3	predicción 4	predicción mercado	probabilidad final
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.05	0.04	0.04	0.01	0	0.02	0.04
0.1	0.08	0.01	0	0	0.02	0.08
0.15	0.13	0.06	0.05	0.05	0.07	0.13
0.2	0.16	0.11	0.11	0.11	0.12	0.18
0.25	0.21	0.16	0.16	0.16	0.17	0.23
0.3	0.24	0.22	0.22	0.22	0.22	0.28
0.35	0.34	0.34	0.33	0.29	0.33	0.34
0.4	0.39	0.39	0.36	0.35	0.37	0.39
0.45	0.44	0.43	0.41	0.41	0.42	0.44
0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.55	0.56	0.56	0.59	0.59	0.57	0.56
0.6	0.61	0.61	0.64	0.65	0.63	0.61
0.65	0.66	0.66	0.67	0.71	0.68	0.66
0.7	0.77	0.78	0.78	0.78	0.78	0.72
0.75	0.78	0.84	0.84	0.84	0.82	0.77
0.8	0.83	0.89	0.89	0.89	0.88	0.82
0.85	0.87	0.93	0.95	0.95	0.93	0.87
0.9	0.92	0.98	1.0	1.0	0.97	0.92
0.95	0.96	0.96	0.98	1.0	0.97	0.96
1.0	1.0	1.0	1.0	1.0	1.0	1.0

Tab. 5.14: Predicción realizada por cada jugador con $c = 0, 2$.

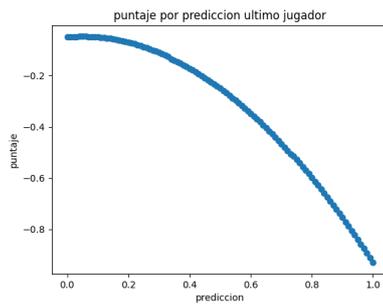
Diferencia entre la probabilidad final y la probabilidad original al variar c

(a) Diferencia con $q = 0, 05$.(b) Diferencia con $q = 0, 35$.Fig. 5.1: Diferencia entre la probabilidad original y la probabilidad final al variar c .

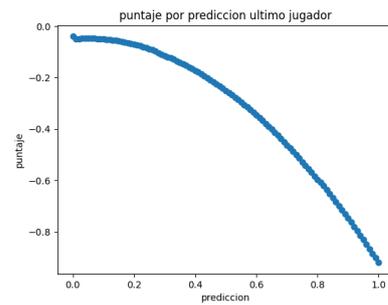
Notar que, si bien en ambos casos la distancia entre la probabilidad final y la inicial crece, lo hace de formas distintas. Esto se debe a la granularidad utilizada para las predicciones. En ambos casos los jugadores se benefician al llevar la predicción hacia el extremo más cercano, es decir, hacia 0. En el primer caso, como la probabilidad original es 0,05 y la granularidad utilizada para las predicciones es de 0,01, son menos las opciones de los jugadores para extremizar la probabilidad, y de ahí el crecimiento escalonado. En el otro caso, donde la probabilidad original es 0,35, hay más predicciones posibles que llevan la probabilidad hacia el extremo.

Gráficos de los puntajes del último jugador

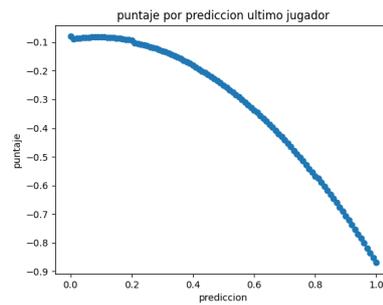
Para analizar por qué las predicciones de los últimos jugadores tienden a alejarse más de la probabilidad original del evento, realizamos un experimento donde, para ciertos valores de q , fijamos las predicciones de los primeros tres jugadores en las obtenidas en la simulación anterior (las que figuran en las tablas anteriores), de forma de observar qué puntaje obtendría el último jugador para cada posible predicción. A continuación incluimos los gráficos obtenidos para algunos valores específicos de q .



(a) Puntajes con $c = 0,1$

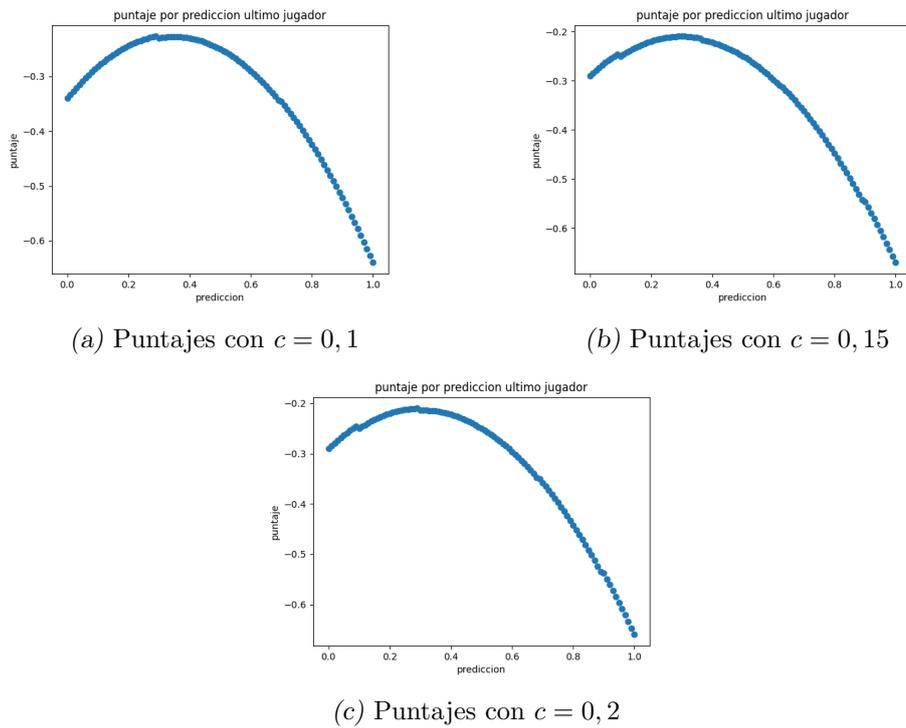


(b) Puntajes con $c = 0,15$



(c) Puntajes con $c = 0,2$

Fig. 5.2: Puntajes del último jugador con $q = 0,05$

Fig. 5.3: Puntajes del último jugador con $q = 0,35$

F. Datos completos regla logarítmica

A continuación presentamos los puntajes obtenidos por cada jugador y las predicciones realizadas por cada jugador para cada q entre 0 y 1 con granularidad de 0,05, con la regla logarítmica, manteniendo siempre el mismo peso para todos y variando el valor de la constante c de la función f , con la misma función f utilizada para la regla de Brier.

$c = 0, 1$

q	puntaje 1	puntaje 2	puntaje 3	puntaje 4	puntaje mercado
0.0	0.0	0.0	0.0	0.0	0.0
0.05	-0.1985	-0.1985	-0.1985	-0.1985	-0.1985
0.1	-0.3251	-0.3251	-0.3251	-0.3251	-0.3251
0.15	-0.4068	-0.4182	-0.4182	-0.4182	-0.413
0.2	-0.4862	-0.4957	-0.5004	-0.5004	-0.4921
0.25	-0.5514	-0.5587	-0.5587	-0.5587	-0.5559
0.3	-0.6032	-0.6087	-0.6087	-0.6087	-0.6063
0.35	-0.6413	-0.6469	-0.6469	-0.6469	-0.6448
0.4	-0.673	-0.673	-0.673	-0.673	-0.673
0.45	-0.6881	-0.6881	-0.6881	-0.6881	-0.6881
0.5	-0.6931	-0.6931	-0.6931	-0.6931	-0.6931
0.55	-0.6881	-0.6881	-0.6881	-0.6881	-0.6881
0.6	-0.673	-0.673	-0.673	-0.673	-0.673
0.65	-0.6413	-0.6469	-0.6469	-0.6469	-0.6448
0.7	-0.6024	-0.6087	-0.6087	-0.6087	-0.6063
0.75	-0.5514	-0.5587	-0.5587	-0.5587	-0.5559
0.8	-0.4866	-0.4957	-0.4957	-0.4957	-0.4921
0.85	-0.4068	-0.4182	-0.4182	-0.4182	-0.413
0.9	-0.3163	-0.3163	-0.3163	-0.3163	-0.3163
0.95	-0.1985	-0.1985	-0.1985	-0.1985	-0.1985
1.0	0.0	0.0	0.0	0.0	0.0

Tab. 5.15: Puntaje obtenido por cada jugador con $c = 0, 1$.

q	predicción 1	predicción 2	predicción 3	predicción 4	predicción mercado	probabilidad final
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.15	0.12	0.09	0.09	0.09	0.1	0.14
0.2	0.19	0.14	0.13	0.13	0.15	0.19
0.25	0.23	0.19	0.19	0.19	0.2	0.24
0.3	0.27	0.24	0.24	0.24	0.25	0.29
0.35	0.33	0.29	0.29	0.29	0.3	0.34
0.4	0.4	0.4	0.4	0.4	0.4	0.4
0.45	0.45	0.45	0.45	0.45	0.45	0.45
0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.55	0.55	0.55	0.55	0.55	0.55	0.55
0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.65	0.67	0.71	0.71	0.71	0.7	0.66
0.7	0.72	0.76	0.76	0.76	0.75	0.71
0.75	0.77	0.81	0.81	0.81	0.8	0.76
0.8	0.82	0.86	0.86	0.86	0.85	0.81
0.85	0.88	0.91	0.91	0.91	0.9	0.86
0.9	0.95	0.95	0.95	0.95	0.95	0.91
0.95	0.95	0.95	0.95	0.95	0.95	0.95
1.0	1.0	1.0	1.0	1.0	1.0	1.0

Tab. 5.16: Predicción realizada por cada jugador con $c = 0, 1$. $c = 0, 15$

q	puntaje 1	puntaje 2	puntaje 3	puntaje 4	puntaje mercado
0.0	0.0	0.0	0.0	0.0	0.0
0.05	-0.1695	-0.1939	-0.1939	-0.1939	-0.1759
0.1	-0.3025	-0.3054	-0.3163	-0.3163	-0.3054
0.15	-0.405	-0.405	-0.4182	-0.4182	-0.4092
0.2	-0.4862	-0.4862	-0.4921	-0.5004	-0.4876
0.25	-0.5511	-0.5511	-0.5587	-0.5587	-0.5537
0.3	-0.6022	-0.6022	-0.6087	-0.6087	-0.6032
0.35	-0.641	-0.641	-0.6469	-0.6469	-0.6419
0.4	-0.6687	-0.6696	-0.6722	-0.6722	-0.6707
0.45	-0.6861	-0.6878	-0.6878	-0.6878	-0.6878
0.5	-0.6931	-0.6931	-0.6931	-0.6931	-0.6931
0.55	-0.6861	-0.6878	-0.6878	-0.6878	-0.6867
0.6	-0.6687	-0.6696	-0.6722	-0.6722	-0.6707
0.65	-0.641	-0.641	-0.6469	-0.6469	-0.6431
0.7	-0.6022	-0.6022	-0.6087	-0.6087	-0.6032
0.75	-0.5511	-0.5511	-0.5587	-0.5587	-0.5537
0.8	-0.4862	-0.4862	-0.4957	-0.4957	-0.4876
0.85	-0.405	-0.405	-0.4182	-0.4182	-0.4092
0.9	-0.3025	-0.3054	-0.3163	-0.3163	-0.3095
0.95	-0.1695	-0.1939	-0.1939	-0.1939	-0.1939
1.0	0.0	0.0	0.0	0.0	0.0

Tab. 5.17: Puntaje obtenido por cada jugador con $c = 0, 15$.

q	predicción 1	predicción 2	predicción 3	predicción 4	prediccion mercado	probabilidad final
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.05	0.03	0.01	0.01	0.01	0.02	0.04
0.1	0.09	0.07	0.05	0.05	0.07	0.09
0.15	0.14	0.14	0.09	0.09	0.11	0.14
0.2	0.19	0.19	0.15	0.13	0.17	0.19
0.25	0.24	0.24	0.19	0.19	0.21	0.24
0.3	0.29	0.29	0.24	0.24	0.27	0.29
0.35	0.34	0.34	0.29	0.29	0.32	0.34
0.4	0.39	0.37	0.35	0.35	0.36	0.39
0.45	0.43	0.41	0.41	0.41	0.41	0.44
0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.55	0.57	0.59	0.59	0.59	0.58	0.56
0.6	0.61	0.63	0.65	0.65	0.64	0.61
0.65	0.66	0.66	0.71	0.71	0.69	0.66
0.7	0.71	0.71	0.76	0.76	0.73	0.71
0.75	0.76	0.76	0.81	0.81	0.79	0.76
0.8	0.81	0.81	0.86	0.86	0.83	0.81
0.85	0.86	0.86	0.91	0.91	0.89	0.86
0.9	0.91	0.93	0.95	0.95	0.94	0.91
0.95	0.97	0.99	0.99	0.99	0.99	0.96
1.0	1.0	1.0	1.0	1.0	1.0	1.0

Tab. 5.18: Predicción realizada por cada jugador con $c = 0, 15$.

$c = 0, 2$

q	puntaje 1	puntaje 2	puntaje 3	puntaje 4	puntaje mercado
0.0	0.0	0.0	0.0	0.0	0.0
0.05	-0.1679	-0.1695	-0.1939	-0.1939	-0.1759
0.1	-0.3025	-0.3025	-0.3095	-0.3163	-0.3054
0.15	-0.405	-0.405	-0.4054	-0.4182	-0.4068
0.2	-0.4862	-0.4862	-0.4866	-0.5004	-0.4876
0.25	-0.5472	-0.551	-0.551	-0.551	-0.551
0.3	-0.6022	-0.6022	-0.6024	-0.6087	-0.6024
0.35	-0.641	-0.641	-0.6413	-0.6469	-0.6413
0.4	-0.6687	-0.6687	-0.6707	-0.6722	-0.6696
0.45	-0.6859	-0.6861	-0.6878	-0.6878	-0.6867
0.5	-0.6931	-0.6931	-0.6931	-0.6931	-0.6931
0.55	-0.6859	-0.6859	-0.6878	-0.6878	-0.6861
0.6	-0.6687	-0.6687	-0.6707	-0.6722	-0.6696
0.65	-0.641	-0.641	-0.6413	-0.6469	-0.6419
0.7	-0.6022	-0.6022	-0.6032	-0.6087	-0.6032
0.75	-0.5442	-0.551	-0.551	-0.551	-0.551
0.8	-0.4862	-0.4862	-0.4866	-0.4957	-0.4866
0.85	-0.405	-0.405	-0.4054	-0.4182	-0.4068
0.9	-0.3025	-0.3025	-0.3054	-0.3163	-0.3054
0.95	-0.1679	-0.1679	-0.1939	-0.1939	-0.1759
1.0	0.0	0.0	0.0	0.0	0.0

Tab. 5.19: Puntaje obtenido por cada jugador con $c = 0, 2$.

q	predicción 1	predicción 2	predicción 3	predicción 4	predicción mercado	probabilidad final
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.05	0.04	0.03	0.01	0.01 ⁷	0.02	0.04
0.1	0.09	0.09	0.06	0.05	0.07	0.09
0.15	0.14	0.14	0.13	0.09	0.12	0.14
0.2	0.19	0.19	0.18	0.13	0.17	0.19
0.25	0.18	0.17	0.17	0.17	0.17	0.23
0.3	0.29	0.29	0.28	0.24	0.28	0.29
0.35	0.34	0.34	0.33	0.29	0.33	0.34
0.4	0.39	0.39	0.36	0.35	0.37	0.39
0.45	0.44	0.43	0.41	0.41	0.42	0.44
0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.55	0.56	0.56	0.59	0.59	0.57	0.56
0.6	0.61	0.61	0.64	0.65	0.63	0.61
0.65	0.66	0.66	0.67	0.71	0.68	0.66
0.7	0.71	0.71	0.73	0.76	0.73	0.71
0.75	0.81	0.83	0.83	0.83	0.83	0.77
0.8	0.81	0.81	0.82	0.86	0.82	0.81
0.85	0.86	0.86	0.87	0.91	0.88	0.86
0.9	0.91	0.91	0.93	0.95	0.93	0.91
0.95	0.96	0.96	0.99	0.99	0.98	0.96
1.0	1.0	1.0	1.0	1.0	1.0	1.0

Tab. 5.20: Predicción realizada por cada jugador con $c = 0, 2$.

⁷ Notar que en este caso, con la regla de Brier el último jugador predecía 0, mientras que eso no sucede con la regla logarítmica.

Bibliografía

- [1] Glenn W. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78:1–3, 1950. (Citado en 1, 9)
- [2] Tilmann Gneiting and Adrian E. Raftery. Weather forecasting with ensemble methods. *Science*, 310:248 – 249, 2005. (Citado en 1)
- [3] Laurence J. Wilson, William R. Burrows, and Andreas Lanzinger. A strategy for verification of weather element forecasts from an ensemble prediction system. *Monthly Weather Review*, 127:956–970, 1999. (Citado en 1)
- [4] J. Bickel. Some comparisons among quadratic, spherical, and logarithmic scoring rules. *Decision Analysis*, 4:49–65, 06 2007. (Citado en 1, 9)
- [5] Robin Hanson. Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation. *Journal of Prediction Markets*, 1(1):3–15, 2007. (Citado en 1, 2, 9, 11)
- [6] Stuart Armstrong. Good and safe uses of AI oracles. *ArXiv*, abs/1711.05541, 2017. (Citado en 2, 10)
- [7] Stefan Luckner. Prediction markets: Fundamentals, key design elements, and applications. In *Bled eConference*, 2008. (Citado en 2)
- [8] Justin Wolfers and Eric Zitzewitz. Interpreting prediction market prices as probabilities. *Federal Reserve Bank of San Francisco, Working Paper Series*, 02 2006. (Citado en 2, 10, 11)
- [9] Jennifer H. Watkins. Prediction markets as an aggregation mechanism for collective intelligence. 2007. (Citado en 2)
- [10] Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5:107–119, 2003. (Citado en 2, 11)
- [11] Lyle H. Ungar, Barbara A. Mellers, Ville A. Satopaa, Philip E. Tetlock, and Jonathan Baron. The Good Judgment Project: A large scale test of different methods of combining expert predictions. In *AAAI Fall Symposium: Machine Aggregation of Human Judgment*, 2012. (Citado en 2, 10, 23)
- [12] Justin Wolfers and Eric Zitzewitz. Prediction markets. *Stanford Graduate School of Business Research Paper Series*, 2004. (Citado en 2, 11)
- [13] Yiling Chen and Jennifer Wortman Vaughan. A new understanding of prediction markets via no-regret learning. *ArXiv*, abs/1003.0034, 2010. (Citado en 2, 11)
- [14] Tim Roughgarden. Algorithmic game theory. *Communications of the ACM*, 53:78 – 86, 2010. (Citado en 2, 16)

-
- [15] Rubi J. Hudson Johannes Treutlein, Caspar Oesterheld and Emery Cooper. Stop-gradients lead to fixed point predictions. https://www.alignmentforum.org/posts/i3v7WeCXyWiYfhihF/stop-gradients-lead-to-fixed-point-predictions#7_Prediction_markets, Enero 2023. (Citado en 2, 11, 12, 18, 20, 24)
- [16] Nuño Sempere and Alex Lawsen. Alignment problems with current forecasting platforms. *ArXiv*, abs/2106.11248, 2021. (Citado en 7)
- [17] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359 – 378, 2007. (Citado en 8)
- [18] Rubi J. Hudson Johannes Treutlein and Caspar Oesterheld. Proper scoring rules don't guarantee predicting fixed points. <https://www.alignmentforum.org/posts/Aufg88v7mQ2RuEXkS/proper-scoring-rules-don-t-guarantee-predicting-fixed-points>, Diciembre 2022. (Citado en 10)
- [19] Claude E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:623–656, 1948. (Citado en 10)
- [20] Moritz Hardt, Meena Jagadeesan, and Celestine Mendler-Dünner. Performative power. *ArXiv*, abs/2203.17232, 2022. (Citado en 10)
- [21] Juan C. Perdomo, Tijana Zrnica, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. *ArXiv*, abs/2002.06673, 2020. (Citado en 10, 11)
- [22] Rubi J. Hudson and Johannes Treutlein. Conditional prediction with zero-sum training solves self-fulfilling prophecies. <https://www.lesswrong.com/posts/A48amesEmqD8KNSmY/conditional-prediction-with-zero-sum-training-solves-self>, Mayo 2023. (Citado en 10)
- [23] Jakob N. Foerster, Gregory Farquhar, Maruan Al-Shedivat, Tim Rocktäschel, Eric P. Xing, and Shimon Whiteson. DiCE: The infinitely differentiable Monte-Carlo estimator. *ArXiv*, abs/1802.05098, 2018. (Citado en 10)
- [24] Caspar Oesterheld, Johannes Treutlein, Emery Cooper, and Rubi Hudson. Incentivizing honest performative predictions with proper scoring rules. *ArXiv*, abs/2305.17601, 2023. (Citado en 10)
- [25] James Surowiecki. The wisdom of crowds: Why the many are smarter than the few. london. *Abacus*, page 295, 01 2004. (Citado en 10)
- [26] <https://www.predictit.org/>. (Citado en 10)
- [27] <https://manifold.markets/>. (Citado en 10)
- [28] <https://www.metaculus.com/home/>. (Citado en 10)
- [29] Wisdom of the crowd vs. the best of the best of the best. <https://www.metaculus.com/notebooks/15760/wisdom-of-the-crowd-vs-the-best-of-the-best-of-the-best/>, Abril 2023. (Citado en 10)

- [30] More is probably more — Forecasting accuracy and number of forecasters on Metaculus. <https://www.metaculus.com/notebooks/14951/more-is-probably-more--forecasting-accuracy-and-number-of-forecasters-on-metaculus/>, Enero 2023. (Citado en 10)
- [31] Predictive performance on Metaculus vs. Manifold Markets. <https://www.metaculus.com/notebooks/15359/predictive-performance-on-metaculus-vs-manifold-markets/>, Marzo 2023. (Citado en 10)
- [32] Peng Shi, Vincent Conitzer, and M. Guo. Prediction mechanisms that do not incentivize undesirable actions. In *Workshop on Internet and Network Economics*, 2009. (Citado en 11)
- [33] Bernhard von Stengel and Theodore Turocy. Game theory. *Encyclopedia of Information Systems*, 2, 12 2003. (Citado en 13)