Mercados de predicción y predicciones performativas

Tomás Schitter Sergio Abriola
Simposio Argentino de Inteligencia Artificial y Ciencia de Datos
Instituto de Ciencias de la Computación, Universidad de Buenos Aires
tomasschitter12@gmail.com

Abstract. En el ámbito del análisis de desempeño de pronosticadores, las llamadas proper scoring rules son una clase de regla de puntuación que incentiva al pronosticador a predecir su creencia verdadera, bajo la premisa de que las predicciones mismas no pueden afectar el desarrollo del evento en cuestión. Sin embargo, en la práctica existe el fenómeno de performative prediction, donde el pronosticador puede tratar de manipular el mundo mediante su propia predicción de manera que aumente su puntaje esperado pero sin reflejar su creencia verdadera inicial. En este trabajo estudiamos este escenario de predicciones que influyen en el evento que predicen. Nos enfocamos en el contexto de mercados de predicción, los cuales pueden considerarse como una forma de agregación de predicciones individuales para dar una única predicción final. Para esto, modelamos un mercado de predicción como un juego, y usamos herramientas de teoría de juegos para obtener propiedades sobre qué tanto pueden diferir las predicciones óptimas de los jugadores de la probabilidad real del evento, entre otras propiedades teóricas de estos modelos.

Keywords: Scoring rules, Predicciones performativas, Mercados de predicción

Prediction markets and performative predictions

Abstract. In the field of performance analysis of forecasters, proper scoring rules are a class of scoring rules that incentivize truthful forecasting, assuming that the predictions themselves cannot influence the outcome of the event in question. However, in practice, there exists the phenomenon of performative predictions, where the forecaster may attempt to manipulate the world through their own prediction in a way that increases their expected score, without reflecting their true initial belief.
In this work, we study this scenario where predictions influence the event they are trying to predict. We focus on the context of prediction markets, which can be considered a way of aggregating individual forecasts into a single final prediction. To do this, we model a prediction market as a game and use tools from game theory to derive properties about how much players' optimal predictions can differ from the true probability of the event, among other theoretical properties of these models.

Keywords: Scoring rules, Performative predictions, Prediction markets

1 Definiciones y conceptos previos

1.1 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 (o minimizar según el caso) el puntaje obtenido. Esto da lugar a las siguientes definiciones:

Definición 1. *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. Además, trabajamos siempre con scoring rules **regulares**:

Definición 2. *Una scoring rule S orientada positivamente se dice* ***regular*** *si $S(p, q)$ es un valor finito para todo $p, q \in [0, 1]$, excepto tal vez que $S(p, q) = -\infty$ si $p \neq q$.*

1.2 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, como incentivar a los pronosticadores a predecir su verdadera creencia sobre la distribución de probabilidad asociada al evento sobre el cual están prediciendo. 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%, lo cual es una característica sumamente no deseable.

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 hacen esto es asignándole puntaje esperado máximo a una predicción dada cuando ésta coincide con la distribución real asociada al evento predicho (Carvalho, 2016; Waghmare and Ziegel, 2025).

Definición 3. *Para el caso de predicciones binarias, que es en el que nos concentramos en este trabajo, el* ***puntaje esperado*** *de una scoring rule se define como:*

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

A partir de este punto, siempre que nos refiramos a una *scoring rule* nos referimos al puntaje esperado, y la llamamos simplemente S en lugar de S_e.

Definición 4. *Sea S una scoring rule orientada positivamente, decimos que S es una* ***proper scoring rule*** *si $S(q,q) \geq S(p,q)$ para todo $p, q \in [0,1]$. Decimos que S es una* ***strictly proper scoring rule*** *si además la igualdad vale si y solo si $p = q$.*

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

Teorema 5. *(Gneiting and Raftery, 2007) Una scoring rule regular S es una proper scoring rule si y solo si existe una función convexa G con subgradiente G^* tal que para todo $p, q \in [0,1]$*

$$S(p,q) = G(p) + G^*(p)(q-p) \tag{2}$$

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

En la práctica hay varias strictly proper scoring rules que se utilizan con frecuencia (Bickel, 2007). En particular, dos muy utilizadas son la *Regla de Brier* (Brier, 1950) y la regla logarítmica (Hanson, 2007), que, para el caso donde las predicciones son binarias, se definen como:

$$\begin{aligned} &logscore(p,q) = q \log p + (1-q) \log(1-p) \\ &Brier(p,q) = q(p-1)^2 + (1-q)p^2 \end{aligned}$$

1.3 Predicciones performativas

Como dijimos antes, la ventaja de las strictly proper scoring rules es que incentivan al pronosticador a predecir 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 que se conoce como *predicción performativa* (Perdomo et al., 2020) y puede llevar a resultados indeseados. Para modelar esta situación, introducimos la siguiente definición:

Definición 6. *Consideramos que existe una función $f : [0,1] \to [0,1]$, llamada* ***función de influencia****, que describe la influencia de la predicción p del pronosticador en la verdadera distribución q asociada al evento.*

La entrada de la función es la predicción realizada por el pronositcador, y la salida es la nueva probabilidad asociada al evento real *después* de realizada la predicción. 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 sii 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 (Johannes Treutlein and Oesterheld, 2022), 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, $S(p, p) = S(1 - p, 1 - p)$), el pronosticador favorecerá aquellos puntos fijos de menor entropía. 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. Finalmente, se probaron ciertas cotas sobre la distancia que puede haber entre la predicción óptima para el pronosticador y puntos fijos de f.
Además del análisis mencionado sobre predicciones performativas (Hardt et al., 2022; Johannes Treutlein and Oesterheld, 2022), se ha estudiado su efecto en el contexto de oráculos, que son herramientas de inteligencia artificial cuya única funcionalidad consiste en responder preguntas predictivas sobre eventos concretos. Además, han surgido algunas propuestas en los últimos años para solucionar o al menos paliar este fenómeno en diversos contextos (Perdomo et al., 2020), como los oráculos contrafácticos (Armstrong, 2017) o los *stop gradients* (R. J. H. Johannes Treutlein and Cooper, Enero 2023; Oesterheld et al., 2023).

1.4 Mercados de predicción

Los *mercados de predicción* son mercados abiertos donde se intercambian activos cuyo valor depende del resultado de un 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 (Wolfers and Zitzewitz, 2006). De esta forma, los mercados de predicción pueden considerarse como una forma de recoger el conocimiento de expertos o incluso el conocimiento popular de forma de generar predicciones (Watkins, 2007). 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 (Ungar et al., 2012). 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 (Surowiecki, 2004). Hay asimismo actualmente varios ejemplos conocidos de mercados y plataformas de predicción sobre diversos temas, como *predictIt, Manifold o Metaculus.*

Los mercados de predicción han sido estudiados desde diversos puntos de vista, incluyendo la economía (Wolfers and Zitzewitz, 2004), la computación (Hanson, 2007) y la matemática (Wolfers and Zitzewitz, 2006). 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 limitamos a estudiar una formalización simplificada de mercado de predicción que facilitará su estudio. En particular, nos concentramos en uno de los principales problemas que pueden tener los mercados de predicción, que es el de las *performative predictions*, similar a lo descrito en la sección 1.3. Este problema ha sido analizado en los últimos tiempos (R. J. H. Johannes Treutlein and Cooper, Enero 2023; Shi et al., 2009), y, análogamente a lo que sucedía en el caso de un solo pronosticador, consiste en que la predicción final generada por el mercado influye en el resultado del evento sobre el cual se realizan las predicciones.

1.5 Nuestro formalismo

Tomamos como base un modelo similar al presentado en (R. J. H. Johannes Treutlein and Cooper, Enero 2023), que formaliza un mercado de predicción como un juego. Partimos 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 (potencialmente distinta 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, consideramos que la predicción final del mercado es un promedio de las predicciones individuales de cada participante teniendo en cuenta su peso, es decir, dadas las predicciones $(p_1, ..., p_n)$, la predicción final $\hat{p}$ es:

$$\hat{p} = \sum_{i=1}^{n} w_i p_i \tag{3}$$

lo cual está bien definido ya que una combinación convexa de probabilidades da siempre como resultado otra probabilidad. Además, como en el modelo presentado en la sección 1.3, asumimos que existe una función $f : [0, 1] \to [0, 1]$ 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 $S_i(p_i, f(\sum_{i=1}^{n} w_i p_i))$. Asumimos 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 (R. J. H. Johannes Treutlein and Cooper, Enero 2023) se demostró que, si las predicciones de los jugadores resultan ser un equilibrio de Nash (Roughgarden, 2010), los jugadores con un peso suficientemente bajo 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 este trabajo realizamos diversos análisis sobre este formalismo, considerando pequeñas variaciones sobre sus reglas. 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 estudiamos en este trabajo.

2 Resultados

A continuación estudiamos distintas variaciones sobre el formalismo para *mercados de predicción* presentado en la sección 1.5. Uno de los objetivos principales es entonces generalizar el Corolario 15 presentado en (R. J. H. Johannes Treutlein and Cooper, Enero 2023) para otras variantes de este formalismo. Este resultado dice que:

Teorema 7 (R. J. H. Johannes Treutlein and Cooper, Enero 2023, Corollary 15). *Sea G la función convexa asociada a la proper scoring rule*[1] *S, sea la estrategia conjunta $\boldsymbol{p} = (p_1, ..., p_n)$ 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:*

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

2.1 Predicciones simultáneas iteradas

Comenzamos tomando un modelo como el planteado en la sección 1.5, donde todas las predicciones de los jugadores se realizan de manera simultánea, solo que ahora consideramos 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, modificamos también la forma en que puntuamos las predicciones de cada jugador. Sea S_i la strictly proper scoring rule asociada a cada jugador

[1] En el trabajo original se asumía que todos los jugadores eran puntuados con la misma scoring rule.

i (asumimos por simplicidad que todas las scoring rules son siempre positivas), ahora el puntaje de cada jugador i en la ronda k es

$$T_i(p_{i_k}, f(\hat{p}_k)) = S_i(p_{i_k}, f(\hat{p}_k)) - S_i(m(p_{-i_k}), f(\hat{p}_k)) \tag{4}$$

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).
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.
Presentamos a continuación un resultado que extiende el Teorema 7 para esta generalización del formalismo básico planteado en 1.5.

Teorema 8. *Sea G_i la función convexa asociada a la scoring rule del jugador i, S_i, proveniente del Teorema de caracterización 5, 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 $\boldsymbol{p_k} = (p_{1_k}, ..., p_{n_k})$ 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. Empezamos por notar que la derivada de la scoring rule S_i del i-ésimo jugador, utilizando el teorema de caracterización 5 y asumiendo que G_i es doblemente derivable, es:

$$\frac{\partial}{\partial p_i} S_i(p_i, f(\hat{p})) = w_i G_i'(p_i) f'(\hat{p}) + G_i''(p_i)(f(\hat{p}) - p_i) \tag{5}$$

La demostración luego es muy similar a la del Teorema 14 de (R. J. H. Johannes Treutlein and Cooper, Enero 2023). Si $\boldsymbol{p_k} = (p_{1_k}, ..., p_{n_k})$ 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 $p_{i_k} \in (0,1)$:

$$\frac{\partial}{\partial p_{i_k}} T_i(p_{i_k}, f(\hat{p}_k)) = G_i''(p_{i_k})(f(\hat{p}_k) - p_{i_k}) + w_i f'(\hat{p}_k)(G_i'(p_{i_k}) - G_i'(m(p_{-i_k}))) = 0 \tag{6}$$

donde usamos la fórmula para la derivada de S_i presentada en 5. Por lo tanto:

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

notar por hipótesis $G_i''(p_{i_k}) \neq 0$ y 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. 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:

$$C = \underset{p_1,p_2 \in [0,1]}{max} |G_i'(p_1) - G_i'(p_2)|$$
$$D = \underset{p \in [0,1]}{min} G_i''(p)$$

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

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

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 el Teorema se sigue con la misma demostración que en el primer caso. □

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 9 *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 $C = \underset{p_1,p_2 \in [0,1]}{sup} \left| \frac{G'_i(p_1) - G'_i(p_2)}{G''_i(p_1)} \right| < \infty$.*

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:

$$F_i(p_{i_1}, ..., p_{i_k}) = \sum_{j=1}^{k} T_i(p_{i_j}, f(\hat{p}_j)) \tag{7}$$

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 10. *Sea G_i la función convexa asociada a la scoring rule del jugador i, S_i, y sea k el 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 $p_i = (p_{i_1}, ..., p_{i_k})$, 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 6 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 $G_i''(i_{i_j}) \neq 0$. Luego la demostración prosigue de forma muy similar al Teorema anterior. □

Lo que nos dice el Teorema es que, si el i-ésimo jugador predice de 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 del evento en cada ronda es arbitrariamente chica.

2.2 Mercado de suma constante con predicciones únicas y simultáneas

Estudiamos ahora el mismo formalismo, pero considerando nuevamente que hay una única ronda de predicciones y que el mercado es de suma constante, es decir, que cumple la ecuación 8. En (Ungar et al., 2012) 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

$$\sum_{i=1}^{n} S_i(p_i, f(\hat{p})) = c \tag{8}$$

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

Teorema 11. *Si cada una de las funciones G_i asociadas a la scoring rule S_i del i-ésimo jugador 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 8, 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

$$\begin{aligned} \frac{\partial}{\partial p_j} S_i(p_i, f(\hat{p})) &= w_j G'_j(p_j) f'(\hat{p}) + G''_j(p_j)(f(\hat{p}) - p_j) && \textit{si } i = j \\ \frac{\partial}{\partial p_j} S_i(p_i, f(\hat{p})) &= w_j G'_i(p_i) f'(\hat{p}) && \textit{si } i \neq j \end{aligned}$$

Por lo tanto

$$\begin{aligned} 0 = \frac{\partial}{\partial p_j} \sum_{i=1}^{n} S_i(p_i, f(\hat{p})) &= \sum_{i=1}^{n} \frac{\partial}{\partial p_j} S_i(p_i, f(\hat{p})) \\ &= \sum_{i=1}^{n} \frac{\partial}{\partial p_j} (G_i(p_i) + G'_i(p_i)(f(\hat{p}) - p_i)) \\ &= G''_j(p_j)(f(\hat{p}) - p_j) + w_j \sum_{i=1}^{n} G'_i(p_i) f'(\hat{p}) \end{aligned}$$

usando que, como G es C^2 entonces $G^* = G'$ y la fórmula anterior. 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 8: 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

$$\begin{aligned} C &= \max_{1 \leq i \leq n} \{ \max_{p \in [0,1]} |G'_i(p)| \} \\ D &= \min_{p \in [0,1]} |G''_j(p)| \end{aligned}$$

Y acotando de forma muy similar a los Teoremas anteriores se deduce la propiedad. □

Observación 12 *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 (R. J. H. Johannes Treutlein and Cooper, Enero 2023) o como el Teorema 8.*

2.3 Predicciones únicas y secuenciales

Partiendo del formalismo original, consideramos ahora que el juego es **secuencial**, es decir, que 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, que 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.

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

Teorema 13. *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(Roughgarden, 2010).*

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 de un Teorema conocido en Teoría de juegos que dice que todo juego secuencial, finito y de información perfecta tiene un equilibrio de Nash perfecto en subjuegos (Roughgarden, 2010), incluso sin las hipótesis pedidas. Sin embargo, el Teorema 13 en toda su generalidad no se deriva de dicho Teorema, ya que, si bien el juego planteado es secuencial y de información perfecta, las posibles estrategias de cada jugador son infinitas.

Demostración. Para demostrarlo, utilizamos una técnica conocida como *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

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

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.

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

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

El resultado de aplicar este método para cada jugador será un equilibrio de Nash perfecto en subjuegos. □

3 Conclusiones y trabajo futuro

En este trabajo estudiamos el problema de las *predicciones performativas* en el contexto de *mercados de predicción*. Para esto, modelamos un mercado como un juego de n jugadores, y utilizamos herramientas de *teoría de juegos* para obtener ciertas propiedades respecto a las predicciones realizadas por los jugadores en dicho modelo, así como en los modelos que surgen de introducir pequeñas modificaciones en él. Estudiamos 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. 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. 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.

Al ser un tema con múltiples parámetros que pueden variar, existen muchos caminos posibles para trabajos futuros, y cambios en el modelo pueden llevar a cambios en algunas de sus propiedades, ya que podrían modelar distintos aspectos de la realidad. Por un lado, resultaría interesante realizar simulaciones sobre variaciones del modelo que relajen algunas de las hipótesis, principalmente la información perfecta o completa, que es una hipótesis muy fuerte, de forma de observar resultados experimentales de modelos más realistas que resulten difíciles de estudiar analíticamente. Podrían relajarse asimismo las condiciones de diferenciabilidad y continuidad de las distintas funciones involucradas. Sería también interesante estudiar otras formas de agregación para obtener la predicción final del mercado, como la mediana, y ver cómo esta elección impacta en las predicciones de cada jugador. Finalmente, otro posible camino para un trabajo futuro sería estudiar el caso donde las predicciones no sean binarias, sino predicciones sobre espacios muestrales discretos, o incluso espacios muestrales continuos.

References

Armstrong, S. (2017). Good and safe uses of AI oracles. *ArXiv*, *abs/1711.05541*.
Bickel, J. (2007). Some comparisons among quadratic, spherical, and logarithmic scoring rules. *Decision Analysis*, *4*, 49–65.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *78*, 1–3.

Carvalho, A. (2016). An overview of applications of proper scoring rules. *Decision Analysis*, *13*(4), 223–242.

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*, 359–378.

Hanson, R. (2007). Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation. *Journal of Prediction Markets*, *1*(1), 3–15.

Hardt, M., Jagadeesan, M., & Mendler-Dünner, C. (2022). Performative power. *ArXiv*, *abs/2203.17232*.

Johannes Treutlein, R. J. H., Caspar Oesterheld, & Cooper, E. (Enero 2023). Stop-gradients lead to fixed point predictions. https://www.alignmentforum.org/posts/i3v7WeCXyWiYfhihF/stop-gradients-lead-to-fixed-point-predictions

Johannes Treutlein, R. J. H., & Oesterheld, C. (2022). Proper scoring rules don't guarantee predicting fixed points. https://www.alignmentforum.org/posts/Aufg88v7mQ2RuEXkS/proper-scoring-rules-don-t-guarantee-predicting-fixed-points

Oesterheld, C., Treutlein, J., Cooper, E., & Hudson, R. (2023). Incentivizing honest performative predictions with proper scoring rules. *ArXiv*, *abs/2305.17601*.

Perdomo, J. C., Zrnic, T., Mendler-Dünner, C., & Hardt, M. (2020). Performative prediction. *ArXiv*, *abs/2002.06673*.

Roughgarden, T. (2010). Algorithmic game theory. *Communications of the ACM*, *53*, 78–86.

Shi, P., Conitzer, V., & Guo, M. (2009). Prediction mechanisms that do not incentivize undesirable actions. *Workshop on Internet and Network Economics*.

Surowiecki, J. (2004). The wisdom of crowds: Why the many are smarter than the few. london. *Abacus*, 295.

Ungar, L. H., Mellers, B. A., Satopaa, V. A., Tetlock, P. E., & Baron, J. (2012). The Good Judgment Project: A large scale test of different methods of combining expert predictions. *AAAI Fall Symposium: Machine Aggregation of Human Judgment*.

Waghmare, K., & Ziegel, J. (2025). Proper scoring rules for estimation and forecast evaluation. *arXiv preprint arXiv:2504.01781*.

Watkins, J. H. (2007). Prediction markets as an aggregation mechanism for collective intelligence.

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Stanford Graduate School of Business Research Paper Series*.

Wolfers, J., & Zitzewitz, E. (2006). Interpreting prediction market prices as probabilities. *Federal Reserve Bank of San Francisco, Working Paper Series*.