Problema de calificación con revisión aleatoria

A continuación, presentaremos el **problema de calificación con revisión aleatoria** como un juego de teoría de juegos de manera **formal**. Veremos tanto la perspectiva de cada estudiante (como agente que toma decisiones para maximizar su utilidad) como la perspectiva del profesor (o diseñador del mecanismo) que busca fijar los parámetros para incentivar la honestidad y la realización de la tarea.

1. Conjunto de jugadores

- Jugadores (estudiantes): denotamos el conjunto de estudiantes por \${1, 2, \dots, N}\$. Cada estudiante \$i\$ es un jugador.
- **Profesorx** (o diseñador del mecanismo)**: aunque no siempre se modela como "jugador" en el sentido clásico (porque suele ser el "diseñador"), para completitud se puede incluir como un agente que define las reglas. Aun así, nos centraremos en el problema de decisión de los estudiantes y luego veremos cómo el profesor elige sus parámetros.

2. Espacio de estrategias de cada estudiante

Cada estudiante \$i\$ toma dos decisiones secuenciales (que podemos condensar en una sola estrategia compuesta):

- Hacer o no hacer la tarea (acción real):
 \$a_i \in {\text{Hacer},; \text{No Hacer}}\$.
- Reportar si hizo o no la tarea al inicio de la clase (declaración pública):
 \$r_i \in {\text{Reportar "Sí"},; \text{Reportar "No"}}\$.

Por lo tanto, formalmente, la **estrategia** de cada estudiante \$i\$ puede verse como un par: \$s_i ;=; (a_i,; r_i) ;\in; {\text{Hacer}, \text{No Hacer}} ,\times, {\text{"Sí"}, \text{"No"}}\$.

Notemos que, en principio, podrían existir más estrategias si consideráramos matices intermedios (por ejemplo, "hacer la tarea pero no la termino" o "partial credit"), pero aquí mantenemos la versión simplificada.

3. Elección aleatoria de \$A\$ estudiantes para revisión

El **mecanismo** definido por el profesor consiste en revisar aleatoriamente \$A\$ de las \$N\$ tareas (o estudiantes). Supongamos que la selección de \$A\$ estudiantes es **uniforme** entre todos los subconjuntos de tamaño \$A\$ (o que se elige a cada estudiante con probabilidad \$\frac{A}{N}\$, según la simplificación que usemos).

Probabilidad de ser revisado para un estudiante \$i\$: \$p = \dfrac{A}{N}\$ (o la versión exacta con combinatoria, si preferimos).

4. Sistema de recompensas y castigos (función de pagos)

Para capturar la utilidad (payoff) de cada estudiante, definimos:

- Recompensa base por la tarea completada: \$R\$ (por ejemplo, \$R = 10\$ puntos).
- Calificación de quien admite no haber hecho la tarea: \$0\$.
- **Castigo** por mentir y ser descubierto: \$C\$. Observa que \$C < 0\$ (es un número negativo, por ejemplo \$-10\$, \$-20\$, etc.).

De este modo, el pago (utilidad) de un estudiante \$i\$ depende de:

- 1. Si realmente hizo la tarea $(a_i = \text{Hacer} \setminus \{0,0\})$.
- 2. Si declara que la hizo $(r_i = Si' \text{ text} o) No'$
- 3. Si es revisado o no y, en caso de ser revisado, si se descubre una mentira.

Podemos definir la **función de utilidad** de forma más concisa. Denotemos por \$\delta_i\$ la **variable de revisión** de estudiante \$i\$, donde

 $\Delta_i =$

\begin{cases}

- 1, & \text{si el estudiante} i \text{ es revisado},\
- 0, & \text{si no es revisado}.

\end{cases}\$

El **payoff** de i, que llamaremos $u_i(a_i, r_i; delta_i)$, se puede representar (en su versión esperada) así:

- 1. Caso \$\mathbf{a_i = \text{Hacer}}\$:
 - Si reporta "Si", entonces el estudiante obtiene \$R\$, sea revisado o no (porque no miente).
 - Si reporta "No" (algo raro, pero posible), entonces recibe 0 (está renunciando a su recompensa).

2. Caso \$\mathbf{a_i = \text{No Hacer}}\$:

- Si reporta "No", obtiene 0.
- Si **reporta "Sí"** y no es revisado \$(\delta_i = 0)\$, obtiene \$R\$.
- Si **reporta "Sí"** y es revisado \$(\delta_i = 1)\$, es descubierto y recibe \$C\$ (castigo negativo).

Resumiendo en una tabla de resultados (para un solo estudiante), donde asumimos \$\delta_i\$ puede ser 0 o 1:

Hizo	Reporte	Revisado?	Utilidad
Hacer	"Sí"	\$\delta_i \in {0,1}\$	\$R\$
Hacer	"No"	\$\delta_i \in {0,1}\$	\$0\$

Hizo	Reporte	Revisado?	Utilidad
No Hacer	"No"	\$\delta_i \in {0,1}\$	\$0\$
No Hacer	"Sí"	\$\delta_i=0\$	\$R\$
No Hacer	"Sí"	\$\delta_i=1\$	\$C\$ \$(< 0)\$

Para el **cálculo esperado**, consideramos que \$\delta_i = 1\$ con probabilidad \$p\$ y \$\delta_i = 0\$ con probabilidad \$1-p\$. Así, por ejemplo, la **utilidad esperada** de un estudiante que "No Hace" y "Reporta Sí" es:

 $U_{\text{miente}} := p \times C :+ (1-p)\times R$

5. Problema de decisión de cada estudiante (condiciones de equilibrio)

Dado que cada estudiante busca maximizar su utilidad, definimos el problema de decisión individual:

 $\sum_{(a_i,r_i)} \quad \mathcal{E}[\text{payoff} \in a_i, r_i, p]$ sujeto a las reglas del juego (revisión aleatoria, castigo por mentir).

En concreto, el estudiante compara las utilidades de cada una de las **cuatro estrategias** (hacer/no hacer la tarea \$\times\$ reportar sí/no):

- 1. \$(a_i=\text{Hacer},,r_i=\text{"Si"})\$:
 \$U_1 = R \quad (\text{pues siempre recibe } R)\$.
- 2. \$(a_i=\text{Hacer},,r_i=\text{"No"})\$:
 \$U_2 = 0 \quad (\text{pierde el beneficio de } R)\$.
- 3. \$(a_i=\text{No Hacer},,r_i=\text{"No"})\$:
 \$U_3 = 0\$.
- 4. \$(a_i=\text{No Hacer},,r_i=\text{"Sí"})\$: \$U_4 = (1-p),R + p,C\$.

El **estudiante racional** elegirá la estrategia que le dé mayor **utilidad esperada**.

6. El problema del profesor (diseñador del mecanismo)

El profesor (o la profesora) tiene como objetivo que **no exista** incentivo a mentir (y, además, idealmente, que los estudiantes sí hagan la tarea). Desde el punto de vista de **diseño de mecanismos**, el profesor controla:

- 1. \$A\$: el número de estudiantes que revisa.
- 2. **\$C\$**: la magnitud del castigo (un número negativo, \$-|C|\$).
- 3. (Potencialmente, también \$R\$, el valor de la tarea, pero a menudo se asume fijo, por ejemplo en 10 puntos.)

6.1. Incentivo a no mentir (incentivo a la verdad)

Para que un estudiante que **no hace la tarea** prefiera reportar "No" en lugar de mentir (reportar "Sí"), necesitamos:

\$U_3;\ge; U_4\$,

es decir,

\$0;;\ge;; (1-p),R + p,C\$.

Ello nos lleva a la desigualdad clave para desincentivar la mentira:

\$(1-p),R + p,C;\le; 0\$.

6.2. Incentivo a hacer la tarea

Si queremos que los estudiantes, en equilibrio, decidan **hacer la tarea** (y reportar "Sí"), necesitamos además que:

\$U_1;\ge; U_3\$,

o sea,

\$R;\ge; 0\$.

Esto se cumple naturalmente si \$R>0\$. Aun así, para que de verdad les convenga hacer la tarea en vez de no hacerla y mentir, comparamos \$U_1\$ y \$U_4\$:

\$R;\ge; (1-p),R + p,C\$.

Reordenando,

\$R - (1-p),R ;\ge; p,C \quad\Longrightarrow\quad p,R ;\ge; p,C \quad\Longrightarrow\quad R ;\ge; C\$.

Dado que \$C\$ es negativo, es muy probable que \$R \ge C\$ sea cierto de manera trivial (por ejemplo, \$10 \ge -10\$). Pero en cualquier caso, esta desigualdad refuerza la idea de que el castigo sea suficiente para que mentir no sea apetecible.

6.3. Elección de \$(A, C)\$ para inducir verdad y tarea

La decisión del profesor se formula como:

```
$\begin{aligned}
```

&\text{Dado } N, R, \

 ${\rm A,,C} \quad {\rm A,,C} \quad {\rm C,cos}, \$

&\text{sujeto a:}\

&\quad (1-p),R + p,C ;\le; 0 \quad \text{(incentivo a la verdad)},\

 $\alpha p = \frac{A}{N}, \quad 0 \le A \le N,$

 $\alpha C < 0.$

\end{aligned}\$

Donde \$\Pi(\text{profesor})\$ podría ser un criterio que el profesor optimiza (por ejemplo, "maximizar la probabilidad de que todos hagan la tarea sin mentir" o "minimizar el costo de revisar" sujeto al incentivo de honestidad).

Típicamente, el profesor podría querer **minimizar** |C| (no castigar demasiado) y **minimizar** A (no revisar demasiado) **sujeto** a la condición de incentivos. Luego, se deriva que:

 $(1-p),R + p,C \le 0 \quad 0 \quad c \$;\le; -(1-p),R \quad\Longrightarrow\quad C ;\le; -\frac{(1-p)}{p},R\$.

Con $p = \frac{A}{N}$, la relación

\$C;\le; -,\frac{N-A}{A},R\$

aparece como la "frontera" de castigo mínimo necesario para desalentar la mentira.

En la práctica, el profesor elige **un par \$(A, C)\$** sobre esta frontera de modo que los estudiantes no mientan. Por ejemplo:

- Si \$A\$ es bajo (reviso pocas tareas), debo imponer un castigo muy severo (alto en valor absoluto negativo).
- Si \$A\$ es alto (reviso más tareas), el castigo puede ser menor.

7. Resumen del juego formal

- 1. **Jugadores**: \$N\$ estudiantes (y el profesor como diseñador).
- 2. Estrategias de cada estudiante:

 $s_i = (a_i, r_i) \quad \text{hore} \$ $a_i \in \{\text{Hacer}, \text{No Hacer}\}, \$ $r_i \in \{\text{Si''}, \text{No''}\}\$.

- 3. **Resultado aleatorio**: se elige aleatoriamente a A estudiantes para revisar. Cada estudiante es revisado con probabilidad $p = \frac{A}{N}$.
- 4. Pagos:
 - Si un estudiante **hace** la tarea y declara "Sí": obtiene \$R\$.
 - Si un estudiante hace la tarea y declara "No": obtiene 0.
 - Si un estudiante **no hace** la tarea y declara "No": obtiene 0.
 - Si un estudiante no hace la tarea y declara "Sí", no revisado: obtiene \$R\$.
 - Si un estudiante no hace la tarea y declara "Sí", revisado: obtiene \$C\$ (<0).
- 5. **Condición de equilibrio deseado**: los estudiantes no quieren mentir y prefieren hacer la tarea a no hacerla.
- 6. **Objetivo del profesor**: diseñar \$(A, C)\$ de tal forma que esa estrategia (hacer la tarea y declarar correctamente) sea un **Equilibrio de Nash** que satisfaga los incentivos.

Así queda planteado **formalmente** como un **juego de teoría de juegos** con un mecanismo de verificación (revisión aleatoria) y un sistema de recompensas/castigos. La "magia" está en encontrar los parámetros adecuados para que la única estrategia óptima sea "hacer la tarea y reportar verazmente".