

Introducción a la Estadística

Motivación

Fiona Franco Churruarín
fionafch96@gmail.com

UTDT

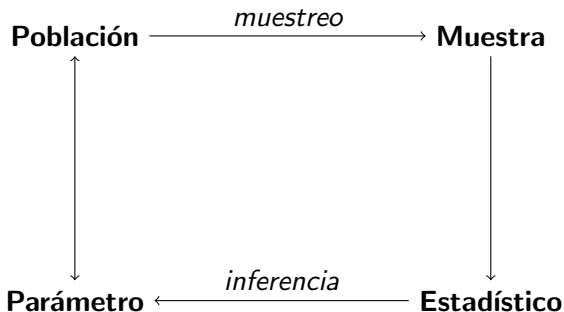
Febrero 2022

Probabilidad y Estadística

Estos dos términos suelen usarse en conjunto y confundirse, pero en realidad están bien diferenciados:

- **Teoría de Probabilidad:** provee modelos para entender (dar estructura ordenada) los experimentos o situaciones inciertas, sobre las cuales no conocemos el resultado hasta que ocurren.
- **Inferencia Estadística:** provee herramientas para, a partir de una *muestra*, hacer declaraciones precisas y fundamentadas sobre alguna característica de la *población*.

Población y Muestra



Población y Muestra

Una diferenciación esencial a la hora de hacer estadística:

- A nivel **poblacional** podemos pensar que existe una caja negra generadora de datos, sobre la cual imponemos cierta estructura utilizando teoría de probabilidad. A las cantidades que caracterizan esta caja negra se les llama **parámetros**.
- A nivel **muestral** sólo vemos una porción limitada de datos, vemos algunos tickets que fueron impresos por la caja negra, que está caracterizada por parámetros. El objetivo de la inferencia estadística es proveer herramientas para, a partir de algunos tickets, dar *la mejor adivinanza posible* sobre el valor de los parámetros en la caja negra.

Ejemplo

Mucho tecnicismo... vamos con un ejemplo bien sencillo. Supongamos que tiro una moneda 5 veces, y los resultados son [cara, cruz, cruz, cara, cara].

La pregunta del millón: ¿que tan verosímil o plausible es que la moneda esté cargada hacia cara?

Nuestro objetivo:

- 1 entender cómo funciona el evento incierto 'tirar una moneda 5 veces';
- 2 a partir de una lista particular de 5 resultados de tirar una moneda, realizar la mejor adivinanza posible sobre, por ejemplo, la probabilidad de que salga cara en un tiro; último y tal vez más importante
- 3 cuantificar que tan (in)seguros estamos de nuestra adivinanza.

Ejemplo, más sexy

Si logramos entender la situación de la filmina anterior, podremos analizar críticamente estudios como publicaciones de encuestas para elecciones.

ELECCIONES 2011 Encuesta de Poliarquía para LA NACION

Macri le lleva a Filmus 17,6 puntos

Continuación de la Pág. 1, Col. 2

que alcanza el 46 por ciento.

Pero a que no se considere necesariamente posibles de balotaje, la encuesta muestra que el panorama se complica para el kirchnerismo. El 46 por ciento de los consultados afirmó que seguramente se votará por Filmus, mientras que en el caso de Macri el "sí" oscila el 42 por ciento.

El gran desafío para Filmus será aprovechar las tres semanas que quedan de campaña para cambiar esta línea. "Lo que pasó con el caso Schekelender, si bien no hay dudas que permitan medir su impacto real, le impidió a Filmus arrancar a tiempo su campaña. No le da tiempo ahora para acercar la brecha con la que arrancó", opina Pablo Penederhitz, uno de los directores de Poliarquía.

En los próximos días, el candidato del Frente para la Victoria deberá esforzarse para dar la pelea, más no la agenda política quedará seguramente monopolizada por la decisión reelección de Cristina Kirchner sobre su candidatura presidencial y su eventual compañero de fórmula.

En un mismo momento vendrá que moviera Solanas si aspira a desplazar a Filmus del balotaje previsto para el 21 de julio.

En las elecciones de 2007, Macri obtuvo 45,8% en la primera vuelta, contra 32,7 de Filmus. El tercero fue Jorge Telerman, con 20,8%, que ahora vuelve a presentarse con nuevas posibilidades de recuperar la cuarta un 2,6%. Ya en el balotaje el resultado fue 60 a 39 para el actual jefe por sí solo.

La encuesta de Poliarquía la consultó en una muestra de 1000 casos telefónicos (sumados entre el 10 y el 17 del actual), con un error estadístico de +/- 3,56 por ciento y un nivel de confianza del 95 por ciento.

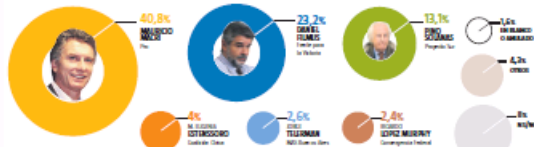
Entre los datos relevantes que arroja muestra el poco conocimiento del electorado sobre el proceso de

EL VOTO DE LOS PORTEÑOS

NIVEL DE CONOCIMIENTO DE LAS ELECCIONES A JEFE DE GOBIERNO ¿Sabe en qué fecha se realizarán?



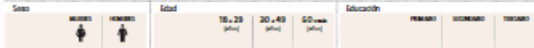
INTENCION DE VOTO A JEFE DE GOBIERNO



RAZONES DE VOTO



SEGUN VARIABLES SOCIOECONOMICAS



OPINION

Las razones que explican la diferencia

ALEJANDRO CATTENBERG PARA LA NACION

Si imagináramos un análisis interno a la maquinaria de la política porteña a quien se le pidiere una impresión sobre quién ganará las próximas elecciones, programaría la siguiente: ¿cuál es el momento electoral más propicio y similar a la presente elección?

Respondería como respuesta que en 2007 fue de los tres principales candidatos ya no me acuerdo y que Macri se impuso con el 46% de los votos contra el 24% de Filmus. "Y como se el nivel de aprobación de la gestión de gobierno después de 4 años", remarcaría. Razionalmente buena, la mayoría de los porteños (64%) la aprueba. (Y la imagen del candidato del oficialismo después de mucho en estos años? No. Macri mantiene una imagen positiva del 52% y una negativa del 36%. Y la oposición no presenta unificada en una oportunidad? No, está incluso más fragmentada que hace cuatro años. Sin pensar demasiado, la conclusión del análisis sería que el electorado porteño no vendría de muchas dificultades para renovar su mandato de gobierno.

Expandiendo el análisis, existen dos ejes que cruzan la política porteña, la aprobación y rechazo al gobierno nacional y al de Mauricio Macri. De la combinación de esos dos variables se obtienen cuatro tipos de electores. Por un lado, están quienes están satisfechos con ambos gobiernos, que son alrededor del 20% de los porteños. Entre ellos, Macri

Ejemplo, más sexy

La letra chica...

"Ficha técnica: universo: Población de 18 años y más en condiciones de votar, residentes de la Ciudad de Buenos Aires. Tipo de encuesta: Telefónica por muestreo (Sistema Cati for Windows). Características de las muestra: Aleatoria, polietápica, estratificada por zonas para la selección de las características y números telefónicos, y por cuotas de edad y sexo para la selección del entrevistado. Tamaño de la muestra: 1000 casos telefónicos. Error estadístico: $\pm 3,16\%$ para un nivel de confianza del 95%. Fecha del trabajo de campo: Entre el 10 y 17 de junio de 2011."

Población: todas las personas en condiciones de votar.

Muestra: 1000 personas.

Este curso

En resumen, si queremos aspirar a hacer declaraciones sobre la población a partir de una muestra, primero debemos entender como funciona la caja negra que genera los datos. Si no sabemos ni a qué le estamos apuntando, es difícil lograr un 'buen' disparo.

En el horizonte tenemos el campo de la inferencia estadística. Eso es lo que buscamos entender y analizar al final del día. Por ahora, dedicaremos algunas clases (tal vez algo técnicas) a revisar rápidamente conceptos de probabilidad.

Este curso se enfocará mas en los conceptos que en los desarrollos matemáticos puntuales.

Alto ahí

Uno podría preguntarse, ¿por qué no mirar a toda la población directamente? Algunos comentarios:

- Puede resultar caro en términos de encuestas, aumentar la muestra implica generalmente un mayor costo.
- Puede resultar imposible en términos prácticos por una cuestión de tiempo y decisión, dimensión de la población, factibilidad física o ética.

Pero lo más importante: si toleramos algo de incertidumbre, no hace falta. Mas vale maña que fuerza.

“La estadística es la disciplina que justifica por qué cuando cocinamos salsa no hace falta probar toda la cacerola para saber si le falta sal, sino que con una cucharadita alcanza.”