

4. Herramienta para la ciencia

La estadística interviene activamente en todas las etapas que componen el método científico.

Aunque para el método científico no exista una secuencia única, señalamos los siguientes pasos generales:

- planteo de preguntas,
- planificación y realización de los estudios,
- recolección de datos,
- análisis de la información,
- obtención de las conclusiones.

Aunque en sí misma es una ciencia que se dedica al desarrollo de nuevos métodos y modelos, la estadística es una disciplina que ofrece un conjunto de ideas y herramientas para el tratamiento de datos. En este sentido, podemos decir que se trata de una disciplina metodológica, su necesidad se deriva de la omnipresencia de la variabilidad.



¿Omnipresencia de la variabilidad?

Los estadísticos trabajan junto a expertos en diferentes disciplinas.

La estadística está involucrada en el proceso que va desde la recolección de la evidencia, el planteo de las preguntas, hasta hallar las respuestas.

Lo importante es que haya preguntas. Para hallar respuestas se pueden seguir diferentes caminos.

Toda investigación comienza con preguntas como:

- ¿Hace mal comer papas fritas?
- ¿Cuánto cuesta enviar un/a niño/a al colegio?

- ¿Quién ganará las próximas elecciones de un club de fútbol?
- ¿Está el peso del cerebro, relacionado con la inteligencia?
- ¿Es posible tomar demasiada agua?

Aunque ninguna de las preguntas anteriores se refiere directamente a números, para responderlas se requiere del uso de datos y de un procedimiento estadístico.

Veamos un ejemplo:

Supongamos que una investigadora quiere determinar quién ganará las próximas elecciones para presidente del Club Grande de Fútbol (58.210 socios), y supongamos también que ya tiene la pregunta, para responderla deberá seguir los siguientes pasos:

- **Determinar el grupo de personas a participar del estudio**

Puede utilizar la lista de socios en condiciones de votar.

- **Recolectar los datos**

Este paso es más difícil. No se puede preguntar a todos los socios si van a ir a votar, y en el caso afirmativo, ¿a quién? Supongamos que alguien dice que irá a votar y declara a quién votará: ¿realmente irá esa persona a votar el día de las elecciones?; ¿dirá esa persona a quién piensa votar realmente? ¿Y si el día de la votación cambia de opinión?

- **Organizar, resumir y analizar los datos**

Luego de recolectar los datos, la investigadora necesita organizar, resumir y analizarlos. Habitualmente, esta parte de su trabajo se reconoce como una tarea para estadísticos.

- **Obtener los resúmenes, tablas y gráficos, realizar el análisis, conclusiones tratando de responder la pregunta del investigador.**

Presentaremos los tipos de resúmenes, tablas y gráficos que podrá utilizar a partir del capítulo 7. Aunque no podamos todavía describir detalladamente el análisis, sabemos que la investigadora no podrá tener un 100% de confianza en que sus resultados sean correctos, porque no le ha podido preguntar a todas las personas y, además algunas pueden cambiar de opinión. Pero sí es posible tener una confianza cercana al 100%, digamos 95% de que la estimación es correcta. De hecho, si ha tomado una muestra representativa (sección 5.4.1) de –por ejemplo– unas 600 personas, de manera que todos los socios tienen igual chance de ser elegidos (muestra insesgada), puede tenerse un resultado preciso con un margen de error de más o menos 4%. Siempre existe la posibilidad de que la conclusión de un estudio sea errónea. Veremos más adelante (sección 10.1) que **el margen de error sólo depende del tamaño de la muestra** y no del tamaño de la población. También veremos el significado del porcentaje de confianza y cuán preciso se espera que sea un resultado.

- **Nuevas preguntas**

Cuando se concluye una investigación y se han contestado las preguntas, los resultados suelen llevar a nuevas preguntas. Podría averiguarse porqué los socios jóvenes prefieren al candidato Rolando Forzudo y los socios mayores a su oponente. A Forzudo podría interesarle estudiar qué factores hacen que los jóvenes realmente vayan a votar.

Hemos dicho que la necesidad de la estadística se deriva de la omnipresencia de la variabilidad. Pero, ¿dónde se encuentra la variabilidad en el ejemplo de la encuesta sobre la preferencia del candidato? Hay varias fuentes (o motivos) que producen variabilidad. La primera, y la más importante, es que no todos los socios piensan igual, si lo hicieran alcanzaría con saber que piensa uno de ellos. La segunda resulta porque las personas cambian de opinión, si esta fuente de variabilidad es muy grande, la validez del resultado depende de cuán cerca de las elecciones se realice la encuesta. La tercera se debe a que los encuestados pueden mentir (sección 6.3.2.).