

1. En 1930 se realizó un experimento en 20,000 niños de edad escolar de Inglaterra. Los maestros fueron los responsables de asignar a los niños de manera aleatoria al grupo de tratamiento -que consistía en recibir 350 ml de leche diaria - o al grupo de control, que no recibía suplementos alimenticios. Se registraron peso y talla antes y después del experimento. El estudio descubrió que los niños que recibieron la leche ganaron más en peso en el lapso del estudio. Una investigación posterior descubrió que los niños del grupo control eran de mayor peso y talla que los del grupo de intervención, antes de iniciar el tratamiento. ¿Qué pudo haber ocurrido? ¿Podemos utilizar los resultados del estudio para inferir causalidad?

El estudio anterior se puede ubicar en el recuadro de selección **no aleatorizada** pues los niños no se seleccionaron de la población en general, por el contrario, fueron seleccionados de niños que asisten a escuelas. La asignación de unidades a grupos fue **aleatorizada** pues se indica que los maestros seleccionaron a los niños que recibían o no leche de manera aleatoria. Por lo tanto, ubicamos este estudio en el recuadro inferior izquierdo.

La investigación posterior puede haber descubierto que la selección de los maestros no fue completamente aleatoria y pudo estar sesgada, pues de manera inconsciente los maestros asignaron a los niños más delgados al grupo que recibía el tratamiento. Esto pudo haber influido en los datos observados. Debido a esta situación no se pueden utilizar los resultados para inferir causalidad pues los grupos no fueron asignados de manera verdaderamente aleatoria.

2. Supongamos que de los registros de un conjunto de doctores se selecciona una muestra aleatoria de individuos americanos caucásicos y de americanos de ascendencia china, con el objetivo de comparar la presión arterial de las dos poblaciones. Supongamos que a los seleccionados se les pregunta si quieren participar y algunos rechazan. Se compara la distribución de presión arterial entre los que accedieron a participar. ¿En qué cuadro cae este estudio? ¿Qué supuesto es necesario para permitir inferencias a las poblaciones muestreadas?

El estudio anterior se puede ubicar en el recuadro de selección **no aleatorizada** pues, aunque se selecciona aleatoriamente de registros de doctores, la selección original no es seleccionada aleatoriamente de la población general de todos los americanos caucásicos y americanos de ascendencia china. En cambio, se selecciona de un subconjunto de personas que tienen registro médico con ciertos doctores. Por otro lado, la asignación de grupos también fue **no aleatorizada** pues se les pregunta a los individuos si deseaban participar o no voluntariamente. Los que aceptan podrían tener mejores números de presión arterial que los que no aceptan, lo cual podría dar resultados erróneos en el estudio. Por lo tanto, ubicamos este estudio en el recuadro inferior derecho.

Para que sea posible hacer inferencias a las poblaciones muestreadas es necesario suponer que los registros de los doctores son representativos de toda la población en general. Esto es difícil pues las personas con registros médicos pueden estar sesgadas por problemas de salud preexistentes (por esta razón acudieron al doctor).

3. Un grupo de investigadores reportó que el consumo moderado de alcohol estaba asociado con un menor riesgo de demencia (Mukamal et al. (2003)). Su muestra consistía en 373 personas con demencia y 373 sin demencia. A los participantes se les preguntó cuánta cerveza, vino, o licor consumían. Se observó que aquellos que consumían de 1-6 bebidas por semana tenían una incidencia menor de demencia comparado a aquellos que se abstendían del alcohol. ¿se puede inferir causalidad?

No, no se puede inferir causalidad del estudio anterior. Esto ocurre debido a que no se están asignando grupos de tratamiento a la población, es decir, grupo que consuma alcohol y grupo que no consuma alcohol, sino que solo se está observando a dos poblaciones distintas ya preexistentes.

4. Un estudio descubrió que los niños que ven más de dos horas diarias de televisión tienden a tener mayores niveles de colesterol que los que ven menos de dos horas diarias. ¿Cómo se pueden utilizar estos resultados?

Dependiendo la manera que el estudio se realizó y los datos se recolectaron podríamos utilizar los resultados. Si la población del estudio se seleccionó de niños menores de 13 años de la población en general y se seleccionaron grupos aleatoriamente donde un grupo veía dos horas de televisión mientras el otro grupo menos de dos horas, manteniendo además una dieta balanceada, ejercicio constante, hábitos saludables para ambos grupos, podríamos decir que el estudio encontró que mirar más de dos horas de televisión causa mayores niveles de colesterol. Si el estudio no se realizó de esta manera **no podríamos decir que se encontró causalidad**, solo podría reportarse que se encontró una asociación entre dos variables y se podría intentar generar conciencia, pero hay muchas más variables que entran en juego como las ya mencionadas.

5. Más gente se enferma de gripe en temporada de invierno, ¿esto prueba que las temperaturas bajas ocasionan las gripas? ¿Qué otras variables podrían estar involucradas?

No, esto no prueba que las bajas temperaturas causen gripe. Encontrar una asociación no significa haber encontrado causalidad. Hay muchos otros factores a tomar en cuenta que podrían estar asociados con el invierno que favorecen la propagación del virus de la gripe, pero el invierno no es la causa directa de la gripe. Se necesitan estudiar todas esas variables en conjunto para encontrar una causalidad clara.

6. ¿Cuál es la diferencia entre un experimento aleatorizado y una muestra aleatoria?

Una muestra aleatoria es solo un subconjunto más pequeño seleccionado aleatoriamente de una población más grande con el objetivo de asegurar que la muestra sea representativa de la población en general. Por otro lado, un experimento aleatorizado es un tipo de estudio en el cual los participantes son asignados aleatoriamente a distintos grupos del estudio (grupo tratamiento o grupo control) con el objetivo de hacer inferencias causales del fenómeno estudiado.