

Fundamentos de Estadística

Tarea 3: Tipos de estudio y PGD

Dorely Morales Santiago

31 de agosto de 2020

Indicaciones: Para cada uno de los siguientes estudios, ubícalos en el recuadro y contesta lo que se pide. Envíen las respuestas por correo electrónico (con título fundamentos-tarea03)

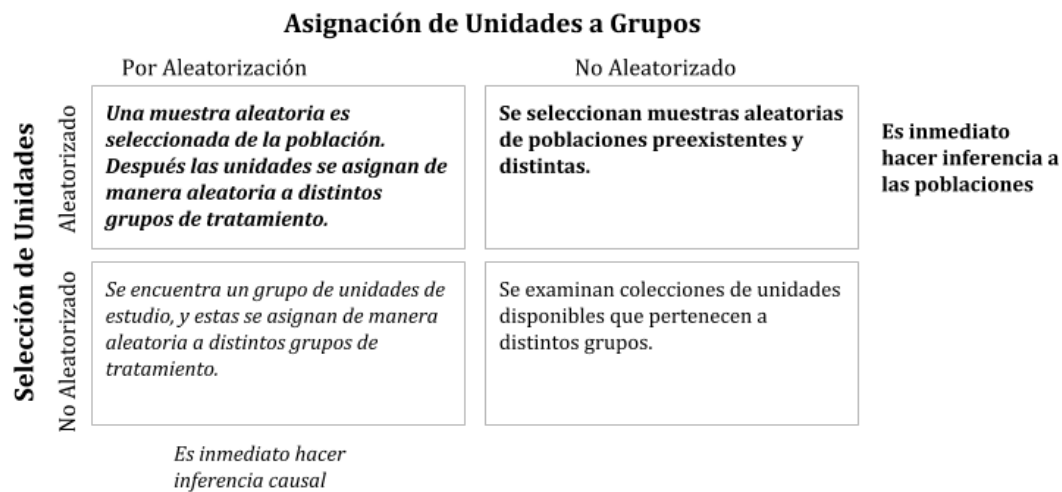


Figura 1: Inferencia estadística de acuerdo al tipo del diseño (Ramsey and Schafer 2012)

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?

Respuesta: Si bien se especifica que los alumnos fueron seleccionados al azar en el experimento, lo que pudo haber sucedido es que la asignación a los grupos de control y tratamiento no fue aleatoria dado que se descubrió que los niños de grupo control eran de mayor peso y talla que los del grupo de intervención antes de iniciar el tratamiento. Como consecuencia, debido a que los grupos no son similares o comparables pues el peso y talla los distingue entonces estrictamente no podemos inferir una causalidad directa por el consumo de leche. **Este caso se coloca en el cuadro de arriba a la derecha.**

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 que cuadro cae este estudio? ¿Qué supuesto es necesario para permitir inferencias a las poblaciones muestreadas?

Respuesta: La selección de unidades (individuos) es no aleatoria dado que se obtiene de los registros de un conjunto de doctores en lugar de hacerse un levantamiento al azar entre todos los individuos caucásicos y americanos de ascendencia china. Además, la asignación de los individuos a los grupos (americanos caucásicos y de americanos de ascendencia china) tampoco es aleatoria pues está asociada a su raza. **Por lo anterior, este caso se coloca en el cuadro de abajo a la derecha.** El supuesto necesario para permitir inferencias a las poblaciones muestreadas es que la selección de unidades (individuos) es aleatoria en el estudio.

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?

Respuesta: No, porque en el estudio sólo están tomando en cuenta la variable de consumo de alcohol pero muchas otras variables pueden estar ligadas a que las personas padezcan demencia y por tanto se deberían tomar en cuenta en el análisis. Además, para que se pueda inferir causalidad, cada grupo de personas debería tener condiciones idénticas con excepción de las variables analizadas lo cual no necesariamente sucede. **Si suponemos que la muestra de personas se generó de manera no aleatoria, este caso se coloca en el cuadro de abajo a la derecha.**

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?

Respuesta: Este ejemplo es similar al anterior. **Si de nuevo suponemos que la muestra de niños de generó sin aleatoriedad el caso también se coloca en el cuadro de abajo a la derecha.** Antes de utilizar estos resultados, se debe conocer el

proceso de generación de los datos para poder hacer inferencia con base en el universo que representa esta muestra y en caso necesario de predecir para un universo diferente se deberán realizar más supuestos, que resultan conclusiones menos sólidas.

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

Respuesta: No, porque puede haber otras variables que estén involucradas para provocar una gripe adicional a las temperaturas bajas. Por ejemplo, que las personas tengan defensas bajas y eso cause los brotes de contagio de gripe. O bien que las personas se contagien más fácilmente en invierno porque se tocan más la cara para calentarse del frío y previamente hayan tocado alguna superficie con el virus. Además, si las temperaturas bajas fueran las causantes de la enfermedad nadie se enfermaría en verano. **Si la afirmación se hace dada la evidencia que la persona tiene a su alrededor. Este caso se coloca en el cuadro de abajo a la derecha.**

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

Respuesta: Una muestra aleatoria se refiere a que las unidades son seleccionadas al azar para participar en el análisis mientras que un experimento aleatorizado es cuando de las unidades seleccionadas éstas se dividen nuevamente al azar para dividir en el grupo control y el grupo de tratamiento. **Si el experimento tiene ambas características, este caso se coloca en el cuadro de arriba a la izquierda.**