

Capítulo 11:

Objetivos de un estudio clínico

Introducción

Este capítulo aporta claves para identificar las preguntas que un determinado estudio permite contestar. Al terminar este capítulo, aquel lector que haya realizado los ejercicios podrá especificar los objetivos de un estudio concreto así como interpretar la utilidad de los originales de investigación más relevantes.

Objetivos

Al terminar este capítulo, un lector que haya realizado los ejercicios:

- Distinguirá entre predicción e intervención.
- Percibirá las ventajas de anticipar el valor de una variable.
- Distinguirá entre una pregunta sobre causas y una pregunta sobre efectos.
- Clasificará como experimental un estudio si el investigador asigna los valores de la causa.
- Exigirá una hipótesis y un protocolo para poder considerar confirmatorio un estudio.
- Valorará las posibilidades de un análisis exploratorio.
- Sabrá que la confusión es una dificultad para la inferencia causal.
- Sabrá que para que exista confusión es necesario que la tercera variable prediga la respuesta.
- Sabrá que para que exista confusión es necesaria colinealidad.
- Identificará interacción con modificación del efecto.
- Pedirá a un diseño ausencia de sesgo.
- Pedirá a un diseño eficiencia.
- Sabrá que la interacción puede aparecer sin colinealidad.
- Sabrá que la interacción puede depender de la escala.
- Identificará el efecto causal con la diferencia en la respuesta entre dos acciones posibles.
- Identificará el efecto causal con la pregunta prospectiva, sobre efectos.
- Percibirá que el efecto causal requiere asignación de la causa.
- Percibirá que el efecto causal requiere una causa alternativa.
- Sabrá que las dos respuestas necesarias para la inferencia causal no son observables.
- Percibirá que la inferencia causal precisa premisas sobre la distribución del efecto.
- Conocerá cómo la inferencia estadística permite la inferencia causal.
- Interpretará el efecto causal ajustado como la pregunta natural de un investigador.

Cómo establecer objetivos

A continuación se interpretan, desde el punto de vista de sus implicaciones sanitarias, algunos de los resultados obtenidos en los capítulos anteriores. En este primer punto se estudia cómo establecer el objetivo de un estudio de investigación y las implicaciones que tiene para la interpretación de sus resultados.

El primer tipo de preguntas que abordan los estudios científicos hacen referencia a la distribución de una sola variable y dan respuesta al quién, cómo, dónde y cuándo. Las herramientas estadísticas disponibles se engloban bajo el concepto de «estadística univariante», y abordan el estudio gráfico, descriptivo y la inferencia mediante intervalos de confianza y pruebas de hipótesis.

Dos posibles usos de una relación: predicción e intervención

El segundo tipo de preguntas que abordan los estudios científicos incluyen el estudio simultáneo de dos variables: ¿qué relación existe entre el valor de la presión arterial y el tratamiento administrado? El establecimiento de esta relación puede tener dos utilidades prácticas bien diferentes: ¿se desea **predecir** (adivinar) acontecimientos futuros (el pronóstico), o se pretende **intervenir** sobre el curso de los mismos (el tratamiento y la prevención)? Así, mediante la administración de un tratamiento se pretende modificar la futura evolución de los pacientes. Nótese que un clínico puede decidir el tratamiento para un paciente; pero, en cambio, no puede ni parar ni acelerar el paso del tiempo, por lo que el tiempo previo de evolución sólo puede ser útil para predecir.

Ejemplo 11.1



Los partes meteorológicos permiten anticipar que pasará tal cosa con tal probabilidad, pero no permiten evitar que pase.

Cuando el objetivo es realizar esta predicción, por ejemplo para clasificar o para seleccionar, lo más importante es cuantificar la capacidad predictiva de la relación estudiada. Es decir, ¿hasta qué punto el conocimiento de una variable permite reducir la incertidumbre sobre la otra?

Ejemplo 11.2



Debemos tomar la presión arterial a Jaime, un nuevo paciente. ¿Por qué presión hay que empezar? ¿Hasta qué presión conviene subir el manómetro? Quizá haya que subir hasta 200 mmHg o quizá sea suficiente con 160 mmHg. Ahora bien, si Jaime nos dice «hace unos meses estaba a 120 de máxima». Ahora la «duda» se reduce: ¿subo hasta 160 o sólo hasta 140? Conocer la presión anterior ha reducido la incertidumbre de 40 (entre 200 y 160) a 20 (entre 160 y 140) mmHg.

Ejemplo 11.3

Mi amiga Julia tendrá un descendiente. ¿Niño o niña? Mi incertidumbre es absoluta. Si apuesto por niño, tendré un 50% de posibilidades de error. Y lo mismo si apuesto por niña. Ahora bien, si el ecógrafo, con un margen de error del 1%, predice niña, la incertidumbre se ha reducido del 50 al 1%.

Recuerde

El objetivo de la predicción es anticipar acontecimientos.

Ejemplo 11.4

Un 80% de los pacientes con un determinado nivel de CD4 no desarrollan sida durante los próximos 12 meses.

Recuerde

El objetivo de la intervención es dirigir los acontecimientos futuros.

Ejemplo 11.5

La protección con un preservativo prácticamente anula las probabilidades de contraer el sida.

Cuando el ambicioso objetivo es intervenir sobre una variable para cambiar el valor de otra, es necesario que exista relación de causa-efecto.

Recuerde

Para predecir el valor de una variable a partir del valor de otra es suficiente con establecer una simple relación numérica entre ambas variables. Sin embargo, para poder intervenir se requiere una relación más íntima, llamada de causa-efecto.

El primer punto necesario para estudiar una relación de causa-efecto es dejar claro qué representa la causa y qué el efecto. Demos el símbolo X a la causa potencial (o intervención, o tratamiento, o factor en estudio) y el símbolo Y al efecto en estudio (o consecuencia, o respuesta, o *end-point*).

Ejemplo 11.6

Al modificar la posición del interruptor (X), conseguimos cambiar el estado de la luz (Y). Si el sistema funciona, a botón apretado corresponde luz encendida; y a botón alzado, luz apagada. Dos características de este fenómeno merecen ser remarcadas. La primera es que, con la premisa de que el sistema funciona, existe una relación perfecta, al 100%, entre la posición del interruptor y el estado *encendido/apagado* de la luz. La relación es tan perfecta, que es determinista: conocido el valor de la variable X, es posible «determinar» el valor de la variable Y. La segunda característica es que, modificando la primera variable, se puede provocar un cambio en la segunda. Nótese que, de estas dos características, la segunda es la que nos induce a pensar en el concepto *causa*.

Ejemplo 11.7

Al aumentar el número de horas (X) dedicadas al estudio de una cierta materia se pretende influir en la puntuación en su examen (Y). Igual que en el caso anterior, modificando las horas de estudio, se desea influir en el dominio de esa materia y, asumiendo que el examen es válido y fiable, en la puntuación. Se cree, por tanto, que existe relación causal. Pero en este ejemplo se pueden imaginar otros fenómenos que también influyan en la respuesta: conocimientos previos, capacidad retentiva, habilidad mental, autoconfianza, suerte, etc. Así, a diferencia del ejemplo anterior, la relación no será *perfecta* o determinista. Pero sí que seguirá siendo causal.

Como el universo no se reduce a estas dos variables en estudio, X e Y, se deben considerar aquellas terceras variables Z que puedan estar relacionadas tanto con la causa X como con el efecto Y.

Ejemplo 11.8

El interruptor (X) permite encender la luz (Y) siempre que el resto del sistema (Z) esté en orden. En cambio, un ejemplo de relación numérica pero no causal sería el volumen (X) de la caja de un ordenador y su capacidad de cálculo (Y). Posiblemente, y muy en general, a mayor volumen (X), mayor capacidad (Y). Lo que permite incluso cierta predicción y seleccionar un buen ordenador. Pero no se pretenderá hacer un ordenador más voluminoso (X) para incrementar sus prestaciones (Y) sin modificar sus componentes (Z).

Ejemplo 11.9

Se podría reducir la incertidumbre sobre la mortalidad por cáncer de un país (Y) conociendo el nivel de uso de ordenadores (X), ya que ambas variables van asociadas al desarrollo socioeconómico (Z). Sin embargo, no se propondrá intervenir sobre el uso de la informática (X) para disminuir la mortalidad por cáncer (Y).

Ejemplo 11.10

En niños de 5 a 15 años existe una fuerte correlación entre su estatura y el número de preguntas correctamente contestadas a una prueba de inteligencia. Dado que ambas variables reflejan la maduración física y psíquica que los niños van desarrollando a esa edad, cabe esperar que a mayor estatura, mayor número de respuestas. Por lo tanto, se puede utilizar esta correlación para, conocida la estatura, determinar o, por lo menos, reducir la incertidumbre sobre el número de preguntas acertadas. Pero ¿se puede intervenir?, ¿se consigue aumentar el número de respuestas correctas (Y) si se aumenta la estatura de un niño (X)? El importante papel del paso del tiempo que representa la edad (Z) obliga a abandonar la idea de «manipular» la inteligencia mediante variaciones en la estatura.

Recuerde

La relación entre dos variables puede ser explicada por la relación de ambas con una tercera variable.

Más adelante se estudia con más detalle el papel que pueden desempeñar las terceras variables.

Perspectiva temporal: preguntas prospectivas y retrospectivas

En la relación de causa-efecto hay dos preguntas muy diferentes en función de su perspectiva temporal: la *prospectiva*, que mira hacia delante, y la *retrospectiva*, que mira hacia atrás.

Recuerde

La pregunta prospectiva es: «¿cuál es el efecto de esta causa?».

Ejemplo 11.11

Si me tomo esta medicina, ¿se me pasará el dolor?

Recuerde

La pregunta retrospectiva es: «¿cuál o cuáles son las causas de este efecto?».

Ejemplo 11.12

Se me ha pasado el dolor de cabeza, ¿es porque me tomé una medicina?

Los dos planteamientos difieren en la perspectiva desde la que se mira: el punto de partida de la pregunta puede ser la causa (la medicina) o bien el efecto (el dolor). Pero tienen en común que el efecto (Y) siempre sucede con posterioridad a la causa (X), lo que es un principio inmutable de las leyes físicas de causalidad.

Ejemplo 11.13

Primero, el buey tira, y luego, el carro se mueve.

Contraejemplo 11.14

Imagine que, para estudiar el efecto de la oración en la evolución de los pacientes, en el año 2000 se asignan a dos grupos distintos 400 historias clínicas de pacientes que enfermaron en el año 1990, uno en el que un grupo de personas rezan por ellos (¡en el año 2000!) y otro en el que no. Posteriormente se compara su nivel de salud en el año 1995. Nótese que el efecto antecede en, por lo menos, 5 años a la hipotética causa. Para justificar este diseño es necesario creer que el efecto de la oración no obedece a las leyes de la física y puede, por tanto, modificar actos pasados.

A lo largo de una investigación médica, las preguntas retrospectivas suelen anteceder a las preguntas prospectivas.

Ejemplo 11.15



En su investigación sobre las epidemias de asma en Barcelona (Antó et al., 58), el primer planteamiento (retrospectivo) de los doctores Antó y Sunyer fue: «Hay una epidemia de asma (Y), ¿cuál o cuáles pueden ser sus causas (X)?». Y observaron que las epidemias ocurrían cuando había descarga de soja en el puerto y cierto viento favorable. Entonces se hicieron la pregunta (prospectiva): «Si cubrimos con una lona (X) la descarga de soja, ¿se acabará la epidemia (Y) incluso en los días de viento (Z)?».

Obsérvese la necesaria interacción entre viento y soja. En la pregunta retrospectiva, ambas aparecen como causas (X); mientras que en la pregunta prospectiva, el viento se trata como una condición (Z) que no depende del investigador y se interviene sobre las condiciones de descarga de soja (X), que sí dependen de él.

Recuerde



Las preguntas retrospectivas suelen preceder, en la evolución del conocimiento, a las preguntas prospectivas.

Nota técnica



En el caso de fármacos, este orden se invierte. La farmacología clínica aborda preguntas prospectivas, más concretas («¿Cuál es el efecto de...?»); mientras que la farmacoepidemiología se concentra en preguntas retrospectivas, cuya formulación es más tentativa (¿cuáles son las causas de...?).

La tabla 11-1 resume los diferentes tipos de preguntas clínicas y sanitarias.

Preguntas típicas		Ejemplos	Pregunta prototipo
Descripción		¿Cuál es la proporción de hipertensos?	¿Cuál es la distribución de la variable Y?
		¿Cuánto tarda?	¿Cuáles son sus valores centrales?
Predicción		Relación entre filtrado glomerular y tiempo de evolución	¿Qué grado de dispersión tiene?
Intervención	Causa de...	Se me ha pasado el dolor de cabeza, ¿es porque me he tomado una aspirina?	¿Qué filtrado glomerular (Y) cabe esperar en un caso de tal tiempo de evolución (X)?
	Efecto de...	Si me tomo la aspirina, ¿se me pasará el dolor de cabeza?	Retrospectiva: ¿Cuál o cuáles son las causas (X) de este efecto (Y)? Prospectiva: ¿Cuál es el efecto (Y) de esta causa (X)?

Tabla 11-1 Resumen de las respuestas que ofrece la inferencia estadística

Estudios experimentales y estudios observacionales

La diferencia primordial entre los estudios observacionales y los experimentales es que en los primeros el investigador «observa» el valor de la causa en estudio, mientras que en los segundos, «asigna» su valor a las unidades.

Por ello, en las investigaciones con seres vivos, la primera consideración para ir hacia un diseño observacional o hacia uno experimental es si la causa es asignable.

Ejemplo 11.16



La esperanza de vida es mucho mayor en las mujeres, pero ningún clínico aconsejará a un paciente masculino que cambie de género. Ni puede plantearse un diseño experimental en que el género se asigne al azar.

Contraejemplo 11.17



En un estudio sobre la discriminación laboral es posible asignar al azar nombres masculinos y femeninos a una serie de currículums y preguntar a un empresario qué salario cree que merecen.

Conviene distinguir si el efecto de esta causa X podría ser beneficioso o no. Si la expectativa es de un efecto positivo, es ético plantear un experimento en seres humanos.

Ejemplo 11.18



Un ensayo clínico sobre un nuevo fármaco. Existen indicios previos tanto de un posible efecto beneficioso como de niveles bajos de efectos tóxicos. La consideración simultánea de ambos efectos, positiva y negativa, permite equilibrar la expectativa en los dos grupos de tratamiento.

En cambio, en aquellas causas potenciales cuyos efectos cabe esperar negativos, no sería ético asignar un voluntario a este tóxico, por lo que sólo se pueden plantear estudios observacionales.

Ejemplo 11.19



Un estudio epidemiológico sobre el efecto del tabaco. No se puede asignar el tabaco a un individuo y pedirle que fume o no fume durante los próximos 20 años para ver si le aparece...

Recuerde



La **asignación** de la posible causa X en estudio divide a los estudios en experimentales y observacionales.

Definición

Ensayo clínico: toda investigación efectuada en seres humanos para determinar o confirmar los efectos clínicos, farmacológicos y/o demás efectos farmacodinámicos, y/o de detectar las reacciones adversas, y/o de estudiar la absorción, distribución, metabolismo y excreción de uno o varios medicamentos en investigación con el fin de determinar su seguridad y/o su eficacia (6).

Definición

Estudio observacional: estudio en el que los medicamentos se prescriben de la manera habitual, de acuerdo con las condiciones normales de la práctica clínica (aquellas establecidas en la autorización de comercialización). La asignación de un paciente a una estrategia terapéutica concreta no estará decidida de antemano por un protocolo de ensayo, sino que estará determinada por la práctica habitual de la medicina, y la decisión de prescribir un medicamento determinado estará claramente dissociada de la decisión de incluir al paciente en el estudio. No se aplicará a los pacientes ninguna intervención, ya sea diagnóstica o de seguimiento, que no sea la habitual de la práctica clínica, y se utilizarán métodos epidemiológicos para el análisis de los datos recogidos (6).

Además de la **causa X** y la **respuesta Y**, hay las terceras variables Z. Usualmente, son condiciones o atributos de los casos, como por ejemplo, el centro o la edad. Dado que su valor no es determinable o asignable por el investigador, sino que las unidades ya se presentan con dicho valor, su única utilidad práctica es predictiva.

Ejemplo 11.20

Ya se comentó que aunque la esperanza de vida es mayor en las mujeres, no tiene sentido aconsejar a un paciente varón que cambie de género.

Recuerde

*Las **terceras variables Z** representan las condiciones o atributos con las que se muestran los casos.*

Análisis exploratorio y análisis confirmatorio

El análisis de unos datos se puede realizar con una finalidad exploratoria o con una finalidad confirmatoria.

Definición

El análisis exploratorio pretende generar ideas e hipótesis.

Durante un análisis exploratorio se busca el modelo, más simple o parsimonioso, que mejor se ajuste a los datos. Para ello, todo vale y se torturan los datos en busca de hipótesis.

Recuerde



Un estudio exploratorio debe terminar con expresiones del tipo: «nuestros resultados sugieren...».

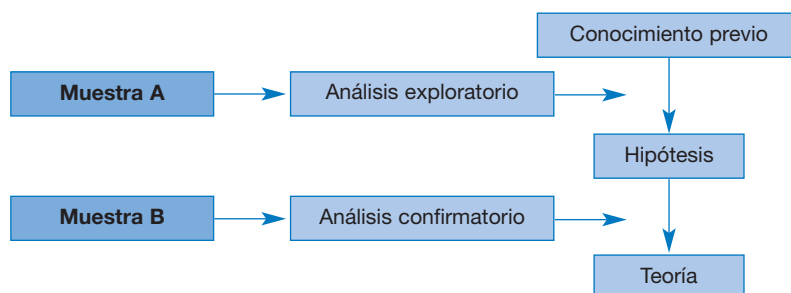


Figura 11-1 Momento del conocimiento científico en el que se emplea cada perspectiva de análisis.

En cambio, las técnicas estudiadas en los capítulos anteriores permitían poner a prueba una hipótesis previamente especificada.

Definición



El análisis confirmatorio pretende elevar a teoría una hipótesis previa.

Aquí, en el análisis confirmatorio ya no se busca la hipótesis que mejor se ajuste a los datos. Ahora se pregunta a los datos si permiten ratificar o rechazar el modelo previo. Dado que el contraste de hipótesis asume unos ciertos riesgos (α y β), este contraste sólo puede realizarse una vez. De lo contrario, estos riesgos no se mantendrían en los niveles deseados.

Ejemplo 11.21



Si se juega un número a la lotería, la probabilidad de que toque es pequeña, pero esta probabilidad sube cuantos más números se compran.

Recuerde



En un estudio confirmatorio sólo se puede hacer un contraste: el de aquella hipótesis especificada en el protocolo.

Historieta

Se dice que un bioestadístico es aquel señor que rechaza que Cristóbal Colón descubriera América porque no estaba en el protocolo de su viaje. Pero es más simple: un bioestadístico le pediría un segundo viaje a Colón para confirmar su hipótesis, tal y como hicieron los Reyes Católicos.

Contraejemplo 11.22

Se dispone de una base de datos que ha ido recogiendo a lo largo de cierto tiempo toda la información sobre partos de un hospital. El investigador ha ido observando que aquellos que hacían una determinada preparación al parto tenían mejor evolución, por lo que decide poner a prueba formalmente esta hipótesis. Si se basa en los mismos datos que la han generado, será un estudio exploratorio; pero si decide recoger nuevos datos (futuros, otros centros, etc.), será confirmatorio.

Definición

Un **estudio prolectivo** tiene una **hipótesis documentada previa** a la existencia de los datos; mientras que en el **retrolectivo** es posterior.

Lectura

Feinstein (59), punto 12.2.1. «Las palabras “prospectivo” y “retrospectivo” se usan a menudo para marcar la diferencia entre planificar la obtención de datos y desarrollar análisis a posteriori, pero su uso es ambiguo debido a que los mismos términos se utilizan para un tipo diferente de distinción temporal, en cuanto a la dirección de las observaciones. Por lo tanto, “prospectivo” también puede referirse a la dirección, hacia delante y longitudinal, en que se siguen las personas en un estudio de cohorte; y “retrospectivo” también puede referirse a la dirección hacia atrás en el seguimiento de un estudio etiológico de casos y controles [...] Para eliminar la ambigüedad, los términos “prolectivo” y “retrolectivo” se pueden utilizar para indicar la existencia de una hipótesis documentada previa a la recolección de datos.»

Lectura

«Las revistas pertenecientes al ICMJE exigen, como condición para la publicación, la inscripción previa en un registro público de ensayos clínicos. [...] Con este objetivo, el ICMJ define un ensayo clínico como cualquier proyecto de investigación que asigne de manera prospectiva a una serie de personas a un grupo de intervención o a un grupo de comparación, con objeto de evaluar la relación causa-efecto entre una intervención médica y un parámetro de salud.» (60)

En resumen, la información y las pruebas (la *evidencia*) aportados por unos datos concretos pueden servir o bien para construir una teoría o bien para ratificarla. Pero unos mismos datos no pueden dar soporte, a la vez, a la construcción y a la ratificación de una teoría.

Recuerde



Para poder finalizar un estudio diciendo «hemos demostrado que...» es necesario un protocolo y una hipótesis previos.

Ejercicio 11.1



Llano et al. (45) terminan su artículo con la siguiente frase: «En su conjunto, nuestros resultados sugieren un papel para IL-7 en el mantenimiento de la regeneración de las células T y la depleción por VIH en individuos infectados, y una posible relación entre IL-7 y la emergencia de variantes SI».

Estas afirmaciones, ¿son teorías o hipótesis? En otras palabras, ¿cree que se trata de resultados de un estudio exploratorio o de uno confirmatorio?

Ejercicio 11.2

¿Qué tipo de utilidad (predicción o modificación) cree que estos autores sugerirán en el texto para la relación entre IL-7 y SI?

¿Y entre IL-7 y la regeneración de células T?

Ejercicio 11.3

La probabilidad de tener un hijo con el síndrome de Down es aproximadamente de uno cada 600 embarazos en la población general, pero asciende hasta uno cada 60 en las gestantes mayores de 36 años.

¿Cuál es el objetivo de establecer relación entre la edad de cada progenitor y el riesgo de síndrome de Down: intervenir o predecir?

Digamos, para terminar este apartado, que las preguntas retrospectivas suelen ser más preliminares, abiertas y exploratorias, mientras que las preguntas prospectivas acostumbran a ser más concretas y confirmatorias.

Ejemplo 11.23

Recuerde que Barcelona estuvo a punto de quedarse sin sus Juegos Olímpicos por las epidemias de asma que aparecían por zonas. Obsérvese el carácter abierto de la primera pregunta retrospectiva que se formuló: «Hay una epidemia de asma, ¿cuál o cuáles pueden ser las causas?». Una vez se atribuyó el asma a la acción conjunta de la descarga de soja en el puerto y el viento que transportaba el alérgeno, se formuló la pregunta prospectiva, que fue concreta y cerrada: «Si reparamos los silos y protegemos con una lona la descarga de soja, ¿se acabará la epidemia?».

Ejercicio cooperativo

A partir de un artículo de investigación, proponga preguntas para que sus compañeros analicen las características del estudio siguiendo todos los puntos vistos en este primer apartado. Idealmente, convendría seleccionar frases en las que los autores sean ambiguos. Preste especial atención a la concordancia entre el objetivo del original y las conclusiones.

Papel de las terceras variables

Se ha dicho que en el estudio entre la causa (X) y el efecto (Y) deben considerarse las terceras variables (Z). En este punto se exponen, con la ayuda de ejemplos numéricos, las formas en las que estas terceras variables pueden interferir en el estudio de la relación entre el efecto y la causa.

El reto de la confusión o cómo evitar sesgos

Si una causa en estudio X está relacionada con una tercera variable Z, los posibles efectos de ambas no podrán diferenciarse y se dice que están confundidos. Por lo tanto, para que exista confusión es preciso que las variables X y Z estén relacionadas, lo que recibe el nombre de colinealidad.

Ejemplo 11.24

Dos tratamientos se utilizan en dos centros diferentes, de atención primaria y terciaria, si bien en el primero (CA1.^a) se tiene preferencia por el tratamiento A, de forma que A se administra 5 veces por cada vez que se ofrece B. En cambio, en el segundo centro (CA3.^a) es justo al revés, ya que A sólo se administra 1 vez por cada 5 de B. Todo ello conduce a un desequilibrio global que la *odds ratio* cuantifica en 25 (tabla 11-2). Existe, por tanto, una fuerte relación entre el tratamiento en estudio (X) y la tercera variable, género (Z).

	x = A	x = B	Razón
CA1. ^a	150	30	5 a 1
CA3. ^a	30	150	1 a 5

odds ratio = OR = 25

Tabla 11-2 El CA1.^a tiene tendencia a aconsejar la intervención A, mientras que el CA3.^a prefiere la B. Centro e intervención están relacionados (OR = 25)

Se dice, en esta situación, que X y Z son colineales y que sus posibles efectos estarán confundidos, como se muestra en la continuación del ejemplo.

Ejemplo 11.25



Pongamos que, por la razón que sea, los pacientes del CA1.^a que evolucionan favorablemente doblan a los que no lo hacen: la razón u *odd* está 2 a 1 (tabla 11-3). Mientras que en el CA3.^a, los que evolucionan favorablemente son la mitad de los que no lo hacen: la *odd* está 1 a 2 (es decir, $\frac{1}{2}$). El cociente de estas dos razones (u *odds ratio*) vale 4 ($2/\frac{1}{2} = 4$), lo que indica que la razón entre evolución favorable y evolución no favorable es 4 veces superior en el centro 1 que en el centro 2. Dicho de otra manera, en el CA1.^a se curan, por el motivo que sea, 4 veces más. No es preciso ni se pretende afirmar que la relación entre el centro y la evolución de los pacientes sea causal: por ejemplo, los del CA1.^a podrían llegar antes o en mejores condiciones, con mayores posibilidades de mejorar. Imaginemos que los dos tratamientos A y B de este ejemplo no tienen ninguna diferencia en su respuesta, tal y como muestran las tablas 11-4, una para cada centro. Se observa en cada centro una *odds ratio* de 1, indicativo de que no existe relación entre tratamiento y evolución. Pero si, por brevedad o descuido, este investigador no tiene en cuenta el centro y estudia únicamente la tabla 11-5, obtendrá una *odds ratio* de 2,5 con un intervalo de confianza de 1,6 a 3,8, que indica asociación entre tratamiento y evolución. Al estudiar ambos centros en conjunto, el tratamiento parece tener un efecto (OR = 2,5) que en realidad no tiene. ¿Qué está sucediendo? Se están confundiendo los efectos de la tercera variable «centro» con los del tratamiento en estudio. Aquí no engaña la intuición: si el tratamiento A se estudia más en el CA1.^a, que tiene mejores resultados, y el tratamiento B en el CA3.^a, que tiene peores resultados, esta comparación estará sesgada a favor del tratamiento A. (Nótese que, en la tabla 11-2, 150 de los 180 pacientes tratados con A provienen del CA1.^a, y viceversa.) La comparación entre A y B en la tabla 11-5 no es «justa» porque los pacientes no tienen la misma procedencia.

	Y+	Y-	Razón
CA1. ^a	120	60	2 a 1
CA3. ^a	60	120	1 a 2

$$OR = 4$$

Tabla 11-3 La evolución es 4 veces (OR) mejor en el CA1.^a

CA1. ^a	Y+	Y-	Razón
X = A	100	50	2 a 1
X = B	20	10	2 a 1

$$OR = 1 \quad IC_{95\%} = 0,4 \div 2,3$$

CA3. ^a	Y+	Y-	Razón
X = A	10	20	1 a 2
X = B	50	100	1 a 2

$$OR = 1 \quad IC_{95\%} = 0,4 \div 2,3$$

Tabla 11-4 En ninguno de los dos centros, la intervención cambia la evolución

	Y+	Y-	Razón
X = A	110	70	11 a 7
X = B	70	110	7 a 11

$$OR \approx 2,5 \quad IC_{95\%} = 1,6 \div 3,8$$

Tabla 11-5 Estudiando ambos centros a la vez, existe relación entre el tipo de intervención y evolución

Recuerde



La relación entre los tratamientos (X) en comparación y la tercera variable (Z) provoca que se confundan sus efectos: un investigador poco atento podría atribuir al tratamiento una diferencia en el efecto que es explicable por la variable Z.

Un premio al buen diseño: ausencia de sesgo

Por supuesto, un buen investigador no cometería el error anterior, ya que habría considerado la variable Z, o bien en el diseño del estudio o bien en el análisis, y habría así evitado el fuerte desequilibrio del tratamiento en ambos centros. Veamos a continuación las ventajas de un buen diseño.

Ejemplo 11.26



Un buen diseño implicaría idéntica razón entre tratados y controles ambos centros, lo que se traduce por una $OR = 1$ (tabla 11-6).

Es decir, ahora el tratamiento está balanceado entre centros, ya que la razón tratados con A frente a tratados con B es idéntica en ambos.

Ahora, si igual que antes no existieran diferencias entre los tratamientos, se podría observar, en cada centro, una situación idéntica a la anterior, como la que muestran las tablas 11-7.

En los dos centros, estudiados por separado, el tratamiento no tiene efecto: en el CA1.^a, ambos tratamientos ($X = A$ y $X = B$) tienen la misma razón «2 a 1» entre evolución positiva ($Y+$) y negativa ($Y-$), lo que conduce a una *odds ratio* de 1, indicador de que el tratamiento no tiene efecto. Y lo mismo en el CA3.^a: como la tendencia hacia una evolución negativa se observa por igual en ambos grupos (1 a 2), tampoco existe relación ($OR = 1$) entre tratamiento y evolución. Si, por parsimonia o por claridad, este investigador quisiera presentar los resultados comparativos de los tratamientos sin distinguir por centro, obtendría los siguientes datos (tabla 11-8), por lo que concluiría también que no existe relación entre tratamiento y evolución.

Es decir, estudiando ambos centros a la vez, el tratamiento tampoco muestra efecto alguno, como era deseable.

Bueno	X = A	X = B	Razón
CA1. ^a	90	90	1 a 1
CA3. ^a	90	90	1 a 1

$$OR = 1$$

Tabla 11-6 Diseño balanceado: intervención y centro son independientes ($OR = 1$)

CA1. ^a	Y+	Y-	Razón
X = A	60	30	2 a 1
X = B	60	30	2 a 1

$$OR = 1 \quad IC_{95\%} = 0,5 \div 1,9$$

CA3. ^a	Y+	Y-	Razón
X = A	30	60	1 a 2
X = B	30	60	1 a 2

$$OR = 1 \quad IC_{95\%} = 0,5 \div 1,9$$

Tabla 11-7 En cada centro por separado, no hay relación entre intervención y evaluación ($OR = 1$)

	Y+	Y-	Razón
X = A	90	90	1 a 1
X = B	90	90	1 a 1

$$OR = 1 \quad IC_{95\%} = 0,7 \div 1,5$$

Tabla 11-8 Estudiando ambos centros a la vez, intervención y evolución también son independientes (OR = 1)

En resumen, un investigador cuidadoso obtiene los mismos resultados cuando realiza el análisis teniendo en cuenta la tercera variable Z que cuando no la considera. Y ello a pesar de que la evolución continúa siendo mejor en el CA1.^a que en el cen-CA3.^a, exactamente igual que antes.

Recuerde



Aunque una tercera variable Z esté relacionada con la respuesta Y, si se mantiene controlada mediante un buen diseño, no podrá sesgar la estimación del efecto del tratamiento.

Comentario



La tercera variable Z se puede relacionar con la respuesta Y por cualquier razón. Se podrían poner muchos ejemplos para explicar por qué los pacientes del CA1.^a se curan más que los del CA3.^a. Pero lo importante es que esta relación, sea del tipo que sea, causal o no, está ahí y no puede anularse. Ni tampoco ignorarse. La solución no pasa por «romper» la relación entre Z y Y que no puede depender del investigador: son «las cosas de la vida» que hay que considerar. La solución consiste en anular la relación entre Z y X, como se vio en el ejemplo anterior, mediante un diseño balanceado.

Ejercicio 11.4



En la tabla 11-9 figura un ejemplo de Bishop et al. (40) —analizado también por Freeman et al. (61)— sobre la evolución de un recién nacido (vive, muere) en función de la duración de la preparación materna al parto (cuidados > < 1 mes) y del hospital (A, B). ¿Qué opina sobre la influencia de la variable clínica en la evolución de recién nacido?

Ejercicio 11.5

Supongamos que en la situación anterior, los resultados fuesen los presentados en la tabla 11-10. ¿Qué opina en este caso sobre la influencia de la variable clínica en la evolución de recién nacido?

	Clínica A		Clínica B		Todos	
	Muere	Vive	Muere	Vive	Muere	Vive
Cuidados < 1	3	176	17	197	20	373
Cuidados > 1	4	293	2	23	6	316
OR IC _{95%} OR	1,25 0,28, 5,64		0,99 0,22, 4,57		2,82 1,12, 7,12	

Tabla 11-9 Evolución Y del recién nacido (vive/muere) según el nivel de cuidados prenatales (X) y la clínica (Z). Resultados discordantes

	Clínica A		Clínica B		Todos	
	Muere	Vive	Muere	Vive	Muere	Vive
Cuidados < 1	7	137	10	204	17	341
Cuidados > 1	5	92	7	137	12	229
OR IC _{95%} OR	0,94 0,29, 3,05		0,96 0,36, 2,58		0,95 0,45, 2,03	

Tabla 11-10 Evolución Y del recién nacido (vive/muere) según el nivel de cuidados prenatales (X) y la clínica (Z). Mismos resultados a nivel de clínica y global

Ejercicio 11.6



En el ejercicio anterior no había confusión al no predecir Z la respuesta Y: en la clínica A mueren 12 de 241 (5%) y en la B, 17 de 358 (5%). Por qué otra razón puede no haber confusión?

Se puede recurrir a la distribución marginal cuando no exista confusión, es decir, cuando Z sea independiente, bien de Y o bien de X.

Recuerde



Si una tercera variable Z está relacionada con las dos variables en estudio (es colineal con X y es predictora de la respuesta Y), deben estudiarse los resultados basados en las tablas definidas por la variable Z, lo que equivale a controlar o ajustar por la variable Z.

Ejercicio 11.7



Si quisiera inventarse unos datos en los que haya confusión, ¿qué tendría que tener en cuenta que pasara?

Recuerde

Para que Z sea confusora debe estar relacionada con las dos variables en estudio.

Nota técnica

La afirmación anterior sólo es estrictamente cierta en el modelo lineal (t-test, regresión, ANOVA...), que considera como respuesta Y a una variable cuantitativa. Lamentablemente, el modelo lineal generalizado (que engloba las tablas de frecuencia) carece de esta propiedad (Greenland, 62), por lo que, en raras situaciones, puede conducir a resultados paradójicos.

Otro premio al buen diseño: mayor eficiencia de las estimaciones

Además de variar el valor de la estimación puntual de la relación entre X e Y, el diseño y el ajuste también repercuten en el valor del error típico de estimación y en la consiguiente amplitud del intervalo de confianza.

Ejemplo 11.27

En los datos de los dos centros de los **ejemplos 11.25 y 11.26**, el intervalo más estrecho correspondía al buen diseño en la tabla global, sin ajustar [0,7, 1,5]; le seguía el intervalo ajustado obtenido en las dos subtablas, también del buen diseño [0,5, 1,9] y el más amplio era el ajustado en las subtablas del mal diseño [0,4, 2,3].

Comentario

No tiene secreto que el intervalo obtenido en la tabla global, sin ajustar, sea más estrecho que el obtenido en las tablas parciales, ya que dispone de mayor número de casos. Tampoco debe sorprender que el intervalo obtenido al comparar 90 casos de un tratamiento con 90 casos del otro sea más estrecho que el obtenido al comparar 150 casos de un tratamiento con 30 del otro, pues ya se vio que la comparación será más eficiente cuando los dos grupos estén más equilibrados.

Un ajuste planificado en el momento del diseño permite optimizar el rendimiento estadístico de las comparaciones, por lo que un mismo número de casos aporta mayor información y se traduce en un intervalo de confianza más estrecho.

Recuerde

*Un buen diseño tiene dos objetivos: aportar **validez** (o controlar sesgos) y aumentar la **eficiencia**, con lo que permite estimaciones más precisas.*

El reto de la modificación del efecto

La situación que se comenta a continuación no debe suponer una sorpresa para el lector, ya habituado a que ciertas condiciones (variables Z) modifiquen el efecto del tratamiento X en la respuesta Y.

Ejemplo 11.28



La ingesta de cierta cantidad de alcohol (X) tiene un efecto negativo en las habilidades motoras (Y), que es mayor en las mujeres que en los hombres (Z).

Esta modificación del efecto se conoce, en términos estadísticos, como interacción entre las variables Z y X en la respuesta Y. Esta situación es muy diferente de la provocada por la confusión. La primera diferencia es que un diseño equilibrado no puede proteger contra esta interacción.

Ejemplo 11.29



Supóngase que, para comparar dos tratamientos A y B, tanto en hombres como en mujeres, se ha procedido a realizar la siguiente asignación equilibrada (misma razón entre tratados con A y con B en ambos géneros), tabla 11-11. Si la variable género es modificadora del efecto, podrían observarse, por ejemplo, los resultados de las tablas 11-12, en las que se observa que mientras en los hombres es el tratamiento B el que obtiene mejores resultados, en las mujeres, los mejores resultados corresponden al tratamiento A. La tabla 11-13 muestra la información agregada de todos los casos. Nótese que no tiene ningún sentido, ya que no informa ni del efecto en los hombres ni del efecto en las mujeres. Una vez más, si un investigador no considera la tercera variable género, los resultados carecen de validez, ya que hablarían de dos tratamientos que tienen la misma eficacia, cuando se ha visto que su eficacia depende de la tercera variable.

	X = A	X = B	Razón
Hombres	60	60	1 a 1
Mujeres	60	60	1 a 1

$$OR = 1$$

Tabla 11-11 Asignación de la intervención X balanceada por género

Hombres	Y+	Y-	Razón	Mujeres	Y+	Y-	Razón
X = A	20	40	1 a 2	X = A	40	20	2 a 1
X = B	40	20	2 a 1	X = B	20	40	1 a 2

OR = 1/4

OR = 4

Tabla 11-12 En los hombres, B obtiene mejores resultados (OR = 1/4), pero en las mujeres lo hace A (OR = 4)

	Y+	Y-	Razón
X = A	60	60	1 a 1
X = B	60	60	1 a 1

OR = 1

Tabla 11-13 Al estudiar conjuntamente a hombres y mujeres, las dos intervenciones obtienen idénticos resultados

Recuerde



La *interacción* puede aparecer aunque no exista colinealidad.

Si hay interacción, para explicar el efecto del tratamiento se debe distinguir por el nivel de la tercera variable. En el caso de la confusión, era necesario considerar la tercera variable en el análisis y analizar los resultados condicionando o ajustando por ella, pero se podía presentar un efecto único del tratamiento, ya que era el mismo en los diferentes niveles de la variable Z.

Ejemplo 11.30



En este último ejemplo, la relación entre el tratamiento y la respuesta es diferente para cada género, y para explicar el efecto es necesario aclarar de qué género se está hablando. En el ejemplo 11.25, el efecto (nulo) era el mismo en ambos centros y se podía hablar de un único efecto (nulo) del tratamiento: todas las OR, tanto las parciales como la global, valían 1.

En el ejemplo 11.29 de modificación del efecto, la OR para los hombres vale 1/4, mientras que la de las mujeres es de 4.

Recuerde

Si Z es una variable confusora, la relación parcial (condicionando por Z) difiere de la relación marginal (sin ajustar por Z).

Si Z es una variable modificadora del efecto, las relaciones parciales (condicionando por los diferentes niveles de Z) difieren entre sí.

Recuerde

La interacción, a diferencia de la confusión, requiere aportar resultados que varían según la tercera variable.

Ejercicio 11.8 (Cont. 6.20)

En un Centro de Atención Primaria han realizado un experimento para poner a prueba un nuevo tratamiento contra las molestias gástricas pospandriales. Para ello han asignado al azar o bien el tratamiento en estudio, o bien un placebo idéntico. El paciente, cuando le aparecía el dolor, en su domicilio, tomaba el tratamiento asignado y anotaba en un cuadernillo si había curado o no.

A pesar de que la asignación había sido realizada al azar, la «mala suerte» (o una selección diferente de los casos facilitada por un diseño no enmascarado) quiso que los pacientes quedaran desequilibrados en cuanto a una variable muy importante: el nivel de dolor que sentían al inicio en el momento en que decidían tomar la medicación. Los pacientes con dolor suave fueron mayoritariamente al grupo placebo, mientras que los pacientes con dolor fuerte pertenecían al grupo tratado. En la tabla 11-14 se exponen los resultados obtenidos. ¿Qué papel desempeña el dolor basal?

	Dolor suave		Dolor fuerte		Todos	
	Cura	No cura	Cura	No cura	Cura	No cura
Tratados	100	5	650	650	750	655
Placebos	550	80	10	200	560	280
OR	2,91		20		0,57	
IC _{95%} OR	1,15, 7,36		10,5, 38,1		0,48, 0,68	

Tabla 11-14 Todas las OR son diferentes (y sus IC_{95%} no se solapan). ¿Qué está pasando?

Comentario

Que exista interacción complica el análisis. Por tanto, la vida será más fácil si no existe. Pero demostrar o negar su existencia no es un objetivo tan importante como estudiar la relación entre X e Y. Si no puede asumirse que no haya interacción, posiblemente el primer paso puede ser establecer el efecto en un subgrupo homogéneo.

La interacción o modificación del efecto es un concepto íntimamente ligado a la medida del efecto empleada. Puede aparecer con una medida y no aparecer con otra.

Ejemplo 11.31

En la tabla 11-15 puede verse que mientras el riesgo relativo es idéntico, el riesgo atribuible es marcadamente diferente. Usando el riesgo relativo podrá decirse que el tratamiento multiplica por 1,5 la probabilidad de curarse, mientras que si se usa el riesgo atribuible habrá que especificar que, cuando el dolor es suave, la proporción de curados aumenta un 20% mientras que cuando es fuerte, sólo lo hace un 5%.

	Dolor suave		Dolor fuerte	
	Cura	No cura	Cura	No cura
Tratados	60	40	15	85
Placebos	40	60	10	90
OR	2,25		1,59	
RR	1,50		1,50	
RA	0,20		0,05	

Tabla 11-15 Al presentar interacción, un único valor del RA no sirve para resumir la relación en ambos grupos de pacientes

Dado que el objetivo científico de parsimonia aconseja utilizar el número mínimo de parámetros para explicar un fenómeno, siempre será preferible utilizar aquella medida que no precisa interacción, ya que entonces puede aplicarse a todos los grupos la misma medida del efecto.

Comentario

El objetivo principal de un estudio suele ser si interviniendo sobre X se consigue modificar la respuesta Y en las condiciones Z. La cuestión sobre si Z es o no confusora (o modificadora) es, en el fondo, secundaria y sirve para decidir el tipo de análisis, de presentación de resultados y la población objetivo. Por ello, en un estudio confirmatorio debe ser algo ya conocido previamente, que formaría parte de las premisas del estudio actual.

Ejercicio cooperativo

Busque ejemplos de la literatura en los que las terceras variables puedan actuar como confusoras o modificadoras del efecto. Proponga preguntas concretas que permitan comprobar que su compañero distingue ambos roles.

Definición de efecto causal

Llevamos un rato hablando de efectos. Vamos ya a definir el efecto causal, primero en una unidad y luego en la población de unidades. Posteriormente, abordaremos el reto de considerar a las terceras variables Z .

Definición del efecto causal en una unidad

Para definir el efecto causal necesitamos establecer un marco. Escojamos primero a la respuesta Y en la que se desea estudiar dicho tratamiento. En segundo lugar, por simplicidad, estudiemos el efecto de cambiar el control c por el tratamiento t . Representemos ahora por $Y_{t,i}$ el valor de la respuesta que se observaría en el individuo i si se le asignara t . Sea $Y_{c,i}$ la respuesta que se observaría si se le asignara c .

Ejemplo 11.32

En el control de la presión arterial, si se desea cambiar en el paciente José López el tratamiento «dieta baja en sal» (c) por el tratamiento «dieta baja en sal más bloqueadores β » (t) y se decide analizar como respuesta Y a la presión arterial diastólica; $Y_{t,i}$ representa la PAD que tendrá el Sr. José López cuando se le administren bloqueadores β más dieta, mientras que $Y_{c,i}$ representa la PAD del mismo paciente si se le hubiera aconsejado dieta sola.

Definición

El efecto causal (Rubin, 63; Holland, 64) es la diferencia entre la respuesta al asignar t y la respuesta al asignar c : $Y_{t,i} - Y_{c,i}$

Ejemplo 11.33

El efecto, en la PAD del Sr. López, de añadir bloqueadores β a la dieta es la PAD de este paciente cuando se le administran bloqueadores β y dieta menos la PAD del mismo señor cuando sólo se le aconseja dieta. Esta definición, aunque pretende ser simple, tiene una serie de connotaciones nada triviales.

La definición de efecto causal es prospectiva

En primer lugar, nótese que se ha definido el *efecto de una causa* y no al revés. Es decir, se está contestando la pregunta prospectiva.

Ejemplo 11.34



Al buscar la causa del tétanos (pregunta retrospectiva) podría encontrarse cierta competencia entre causas que se engloban (el *Clostridium* engloba la toxina), de forma que diversos investigadores, en diferentes momentos del conocimiento científico, podrían atribuir la causa a diferentes componentes del *Clostridium*. En cambio, no habría ambigüedad a la hora de estimar el efecto de cada una de estas posibles diferentes causas (pregunta prospectiva).

Comentario



Esta definición «prospectiva» es la que responde al objetivo final de la sanidad, que es, mediante intervenciones, mejorar el estado de salud de la población.

Nota técnica



Aickin (65) ha propuesto un planteamiento alternativo que pretende modelar el entorno multicausal que suele aparecer en las preguntas retrospectivas. En cambio, la aproximación prospectiva que se presenta aquí considera a las concausas como condiciones en las que se estima el efecto de la causa que se cree que se está en condiciones de intervenir.

La definición de efecto causal habla de asignación

Una segunda consideración en la definición realizada es que se ha hablado de asignación. Esta definición sólo estudia el efecto de aquellos fenómenos en los que el investigador puede o podrá decidir el valor que toman. En una palabra, sólo sirve para aspectos **intervenibles**.

Ejemplo 11.35



Ya se ha dicho que aunque la esperanza de vida de las mujeres es superior a la de los hombres, un clínico, al ver entrar a un paciente masculino no le pide que abandone ese mal hábito.

De acuerdo con esta definición, es fundamental distinguir entre tratamientos —causas susceptibles de una intervención— y los atributos de las unidades, que son condiciones o características poseídas por los individuos. La capacidad de intervención se reduce a los primeros.

Ejemplo 11.36

Mientras no se pueda asignar la edad a un ser vivo, esta definición no permitirá estudiar si la relación entre edad y riesgo de morir es causal. Pero ¿es una limitación? A fin de cuentas, si no se puede decidir el valor de la variable *edad*, ¿para qué importa saber si es una relación causal? Desde una perspectiva práctica, es irrelevante si la relación es causal o no: dado que la edad no depende del investigador, se ha perdido toda posibilidad de influir en la esperanza de vida por medio de la edad. Tanto si la relación es causal como si no lo es, se podrán utilizar las propiedades predictivas de una relación y así predecir el tiempo de vida en función de la edad. También se pueden definir subgrupos homogéneos para un estudio determinado. Pero no tiene utilidad práctica, mientras no se pueda *asignar* la edad, preguntarse si esta relación es o no es causal.

Contraejemplo 11.37

Ciertos hábitos de vida pueden acelerar el envejecimiento celular. Tiene pleno sentido preguntarse si cambiarlos tiene un efecto causal sobre la cantidad y la calidad de vida.

Lectura

El concepto de lo que es intervenible no es inmutable. Como apunta Cox (66), aunque no se puede cambiar la edad de una persona, quizá sí se pueden ralentizar ciertos cambios bioquímicos o moleculares ligados al envejecimiento. Así, la edad no sería asignable, pero sí que lo sería la decisión de ralentizar o no esos cambios.

Ejemplo 11.38

Cabe esperar que un paciente hospitalizado cumpla al 100% con el tratamiento establecido, pero eso no es así para un paciente ambulatorio. El profesional sanitario puede tomar la decisión de asignar un tratamiento u otro a un paciente, pero el grado de cumplimiento posterior depende de este último. Si se quiere valorar el impacto de una intervención sanitaria, se debe estudiar aquello que se desea cambiar en el futuro: la decisión del clínico.

Contraejemplo 11.39

Por supuesto, pueden diseñarse intervenciones para mejorar el grado de cumplimiento. Por ejemplo, podría compararse el nivel de cumplimiento de la medicación entre un blister clásico y un expendedor electrónico con voz que recordara el momento de la toma de la medicación. En ese caso, el clínico podría decidir entre asignar el paciente a un método u otro para dispensar el tratamiento.

Ejercicio 11.9

Cuando se habla de que cierta variable es un factor de riesgo, ¿quiere decir que la relación es causal y por tanto, que mediante una intervención sobre esa variable se lograría disminuir el riesgo?

La definición de efecto causal es relativa a otra causa

Una tercera consideración relevante de la definición presentada es que se ha definido el efecto de una causa de forma relativa a otra causa alternativa. Ello no tiene por qué ser una limitación, ya que emula el proceso de decisión de la vida real, en la que, por lo menos, aparece la alternativa de no hacer nada. ¿Para qué se podría desear conocer las consecuencias o efectos de nuestros actos, si no hay alternativa posible?

Ejemplo 11.40

Supóngase que acude a urgencias un paciente con una fractura de hueso largo para la que el mínimo tratamiento éticamente aceptable es la reducción e inmovilización. Por tanto, no puede observarse la evolución de una alternativa que consista en no hacer nada, en no aplicar el tratamiento.

Existe una situación en la que es posible sustituir la observación de la evolución por el valor hipotético de lo que habría pasado si no se hubiera hecho nada. Se trata de la invarianza o no variabilidad de los datos. Exige poder predecir sin error los valores futuros en caso de que no se interviniera.

Ejemplo 11.41

¿Aceptaría un oftalmólogo la premisa de invarianza en un paciente con cataratas? Si los miembros más críticos de este colectivo científico estuvieran dispuestos a asumir que, en ausencia de intervención, la evolución de las cataratas es perfectamente conocida, creerían en esta invarianza.

Ejemplo 11.42

En los primeros años de la aparición del sida, ningún paciente de cierto estadio evolutivo lograba sobrevivir más de, pongamos, seis meses desde el diagnóstico: invariablemente morían antes. Si un científico aplica un tratamiento nuevo a un paciente que, sin duda alguna, tiene el sida en dicho estadio evolutivo, y este caso único permanece vivo a los doce meses, el resto de la comunidad científica estará muy interesada en dicho tratamiento. A pesar de disponer de una sola observación.

La premisa de invarianza es raramente sostenible. La situación habitual es justamente la contraria: existe una variación notable entre e intra individuos. Si la evolución de los pacientes es ciertamente impredecible, la obtención de datos sobre el tratamiento control es obligada.

Además de esta incertidumbre sobre la evolución de los pacientes, si existe variabilidad intrasujeto y si se seleccionan a los individuos que estén en peores condiciones, aparecerá un fenómeno denominado regresión a la media que originará que, en una nueva observación, estos casos estén en mejores condiciones.

Ejemplo 11.43

La artritis o «reuma» es una afectación con grandes variabilidades dentro de un mismo paciente. Si un paciente decide acudir al médico cuando «ya no aguanta más», posiblemente esté acudiendo en el peor momento de su enfermedad, por lo que tiene más posibilidades de encontrarse mejor en los próximos días, aun en ausencia de tratamiento.

Ejemplo 11.44

En el rendimiento académico suele existir una cierta variabilidad, ya que puede influir la suerte o la inspiración del momento, por ejemplo. El alumno que ha sacado el mejor puesto ya sabe lo difícil que es mantenerse ahí. Es frecuente dar un estímulo negativo (castigo o bronca) después de los peores rendimientos, y uno positivo (premio), después de los mejores. Pero si existe variabilidad intraindividual en el rendimiento académico, cabe esperar que después de la peor nota (que se acompaña de bronca) vengan mejores resultados, y después de la mejor nota (que se acompaña de premio), peores resultados. Dado que los estímulos positivos y negativos se dan en condiciones diferentes, un observador poco atento podría confundir los efectos de la regresión a la media con los hipotéticos efectos de los estímulos.

Recuerde

La presencia de variabilidad obliga a recoger información sobre el control.

Comentario

Puede argumentarse que si existe un único tratamiento para una determinada enfermedad, no tiene ningún interés estimar su efecto ya que no hay alternativa y dicho tratamiento deberá aplicarse forzosamente. Sin embargo, el concepto de éticamente aceptable depende del entorno y de los recursos disponibles: el responsable de la administración de los recursos sanitarios decide el patrón de mínimo tratamiento éticamente aceptable en función del beneficio y el coste de las diferentes intervenciones sanitarias. Ahora, la pregunta no es asignar un tratamiento u otro a un mismo paciente, sino asignar el recurso sanitario a una u otra intervención. A diferencia del profesional asistencial, que cuida personas, el administrador debe velar por la salud de toda una población. Su paciente es la comunidad, y debe decidir la prioridad de las diferentes intervenciones, exactamente de la misma forma en que debe hacerlo un clínico delante de un paciente politraumatizado.

Problema fundamental de la inferencia causal

Para estimar el efecto causal en una unidad determinada, se ha propuesto hacer la resta entre los dos valores de la respuesta bajo los dos tratamientos alternativos, $Y_{t,i}$ y $Y_{c,i}$. Ahora bien, nótese que estas dos respuestas no son observables a la vez y en las mismas condiciones.

Ejemplo 11.45

O bien se puede observar la respuesta al tratamiento bloqueadores β más dieta (PAD_{BB+D}) en unas unidades y la respuesta a la dieta aislada (PAD_D) en otras; o bien, en el mismo individuo, se puede observar la respuesta $PAD_{BB+D, \text{López}}$ en unas condiciones en un cierto momento y la respuesta $PAD_{D, \text{López}}$ en otras condiciones de otro momento. Pero nunca en idénticas circunstancias.

Al exponer la confusión, se vio que estas diferencias —sea en unidades, sea en condiciones, sea en ambas— abren la posibilidad de que terceras variables Z fueran la auténtica explicación de las diferencias observadas.

Ejemplo 11.46

Podría ser que se estudiaran pacientes de diferente edad (Z_1) en cada grupo. O que ambos grupos sean de diferentes centros (Z_2). O que en un segundo período ya no exista alérgeno en el ambiente. O que el primer tratamiento haya conseguido curar la enfermedad antes de dar el segundo.

¿Puede una tercera variable Z ser una explicación de la relación observada entre X e Y ? O, dicho de otra manera, ¿cómo se podrá afirmar que las diferencias en la respuesta Y son debidas a X y no a diferencias entre unidades, o a diferentes condiciones o variables Z ?

Recuerde

Se ha definido el efecto causal en una unidad, pero no es directamente observable. Este problema es tan delicado que recibe el nombre de problema fundamental de la inferencia causal, y constituye el reto de la inferencia estadística y el diseño de experimentos.

Definición de efecto causal en la población

El primer paso consiste en extender al conjunto de todas las unidades el efecto definido en una única unidad. Para generalizar el efecto causal es preciso realizar alguna premisa sobre su distribución a lo largo de las unidades. Al estudiar la comparación de dos medias mediante la t de Student se ha usado la premisa de un efecto constante y aditivo, es decir: idéntico en cada individuo, y que simplemente se suma o resta a su tendencia natural.

Comentario

Este modelo se conoce técnicamente con el nombre de «efectos fijos». En él, los diferentes resultados en las distintas unidades se interpretan por divergencias iniciales entre los individuos, no porque el efecto del tratamiento vaya cambiando de una unidad a otra.

Nota técnica

Ya se vio en el capítulo 8 que esta premisa tan simple está detrás del análisis estadístico habitual: el efecto del tratamiento se manifiesta en las medias, pero no en las dispersiones.

Recuerde

El efecto promedio, o diferencia entre medias, se corresponde con el efecto del tratamiento en cada uno de los casos.

Bajo la premisa de efecto constante, todas las unidades proporcionan información del mismo efecto y pueden combinarse para estimarlo. Ello permite extender a toda la población aquella definición que sólo servía para el Sr. José López.

Definición



El efecto causal es la diferencia entre la respuesta media al administrar t y la respuesta media al asignar c : $E[Y_t] - E[Y_c]$.

Nota técnica



$E[Y_t]$ se lee como «esperanza de Y_t », y es la media poblacional, μ_t , que representa el centro de todos los posibles casos tratados con t .

Establecer una premisa sobre el comportamiento del efecto a lo largo de las unidades permite, no sólo observar qué ha sucedido en el pasado en la unidad estudiada, el Sr. José López, sino también anticipar qué sucederá en las próximas unidades en las que se administre el tratamiento.

Nota técnica



Si la premisa de efecto constante se considera simplista, puede relajarse de varias formas. Una de ellas, conocida como modelo de efectos aleatorios, asume que el efecto del tratamiento tiene una distribución de valores que puede resumirse mediante su media y su desviación típica. Ahora se deben estimar, pues, dos parámetros: el del efecto promedio y el de la distancia usual del efecto en cada unidad a este efecto promedio. Además de requerir diseños sofisticados, que implican determinaciones repetidas en cada individuo, este modelo tiene el inconveniente de su aplicabilidad posterior, ya que, cuanto mayor sea esta dispersión, mayor será la incertidumbre de cuál será el efecto en una unidad determinada, lo que puede llegar a requerir estudios « n igual a 1» para determinar el efecto en cada caso particular.

Podría ser que esta variación del efecto entre las unidades siguiera cierto patrón, de forma que dependiera de una característica observable, por ejemplo, el género del paciente. Así, el tratamiento podría tener un efecto en los hombres y otro en las mujeres, que empezaríamos por considerar como constantes o fijos. Se trata de la interacción ya vista anteriormente entre género y tratamiento (o de modificación del efecto del tratamiento por el género). Ahora, se deben estimar dos parámetros, el del efecto en cada una de las poblaciones, lo que requiere duplicar el esfuerzo experimental.

Ejemplo 11.47

La ingestión de alimentos puede modificar (o interferir o interactuar) la absorción y, por tanto, el efecto de un tratamiento. Por lo general, las agencias reguladoras no exigen estudios en los que se recoja simultáneamente la información en las dos situaciones, antes y después de la ingesta, sino que aceptan estudios complementarios que aporten la información adicional no contemplada en los estudios principales.

Definición de efecto causal ajustado por terceras variables

El siguiente paso para solucionar el problema fundamental de la inferencia causal es considerar las terceras variables.

Todo estudio se realiza en ciertas condiciones. Un investigador debe explicar estas condiciones, no sólo para describir la situación en la que se realiza el estudio sino también para permitir su réplica por futuros investigadores.

Se debe, por tanto, matizar la definición de efecto causal, incluyendo estas características Z.

Definición

El efecto causal ajustado por Z es: $E[Y_t|Z] - E[Y_c|Z]$.

Nota técnica

De la misma forma que $E[Y_t]$ representa la media μ en los casos tratados con t, ahora $E[Y_t|Z]$ representa la media μ en los casos tratados con t a nivel fijo de la variable Z.

En esta expresión, utilizar el mismo símbolo Z en los dos términos de la resta significa que se está condicionando en ambos por los mismos valores de Z. Es decir, se habla de valores de Y_t «dada Z»: se ha dejado «fijo» o constante el valor de Z.

Ejemplo 11.48

Si en el estudio de la PAD, las variables Z que se desean controlar son la edad y el centro, el efecto causal ajustado será el incremento al cambiar el tratamiento c por el tratamiento t en la PAD ajustada por edad y centro, es decir, para un paciente de una edad y de un centro determinados.

Esta definición contesta la pregunta de interés y, además, convierte en constantes las variables Z que, por tanto, ya no pueden ser una explicación alternativa a la relación causal en estudio: ¡no puede existir confusión porque las variables Z son fijas!

Comentario



El problema fundamental de la inferencia causal aparecía porque el efecto causal definido en las unidades no podía observarse, ya que la misma unidad no podía ser estudiada en las mismas condiciones bajo ambos tratamientos.

En cambio, una población sí que puede ser estudiada en las mismas condiciones, bajo tratamientos diferentes. Esta estimación del parámetro central de una población es el reto habitual de la estadística, que ya se ha solucionado en los capítulos anteriores.

Si se conoce la influencia de las terceras variables y, además, éstas son observables y medibles sin error, es posible tenerlas en cuenta mediante las técnicas de ajuste estadístico.

Comentario



Un investigador serio, precavido y buen conocedor de su profesión tendrá en cuenta estas variables al diseñar el estudio y dejará constancia de ello en el protocolo del mismo. Un estudio confirmatorio que pretenda convertir una hipótesis en teoría debe especificar de forma concreta, en el protocolo, las condiciones del estudio y la metodología para tenerlas en cuenta. Las desviaciones del protocolo restarán credibilidad a la inferencia causal.

Ejercicio 11.10



(Continuación del ejercicio 11.3) Si el objetivo de conocer la relación entre la edad materna y el riesgo de Down es intervenir sobre esta edad aconsejando a las mujeres que adelanten su gestación, ¿qué papel juega la edad paterna?

(Nótese la elevada correlación entre la edad de los padres: por un lado, se sabe que a mayor edad de uno, cabe esperar mayor edad del otro; por otro, si se deja fija la edad de uno de los padres, disminuye considerablemente la variabilidad de la edad del otro. Por tanto, también es razonable que exista relación entre la edad del padre y el riesgo de síndrome de Down.)

Soluciones a los ejercicios

11.1 La expresión *sugieren* («*suggest*» en el original) es claramente indicativa de un estudio exploratorio. Estos autores están aportando al conocimiento nuevas ideas que deberán ser contrastadas posteriormente.

11.2 No queda claro. Estas expresiones son habituales en la jerga científica: podrían estar redactadas así para fomentar la lectura del texto y la discusión, donde se aclara que IL-7 puede ser utilizado como «indicador» de la evolución (utilidad predictiva). No se afirma nada sobre si, además, puede ser utilizado para intervenir.

11.3 Si lo que se pretende es cambiar la edad de la madre o del padre para disminuir el riesgo, se trata de una intervención. Pero si se desea seleccionar aquellos embarazos con elevada probabilidad de síndrome de Down para justificar una prueba diagnóstica no exenta de riesgo, entonces el objetivo es predecir.

11.4 Mientras que el estudio marginal, juntando todos los casos, muestra la existencia de relación, la asociación parcial no detecta esta relación. Por tanto, la variable clínica es confusora: condicionando por ella, no es necesario considerar la relación entre evolución y cuidados. Nótese la mayor administración de cuidados en la clínica A (relación ZX) y la mayor mortalidad en la B (relación ZY).

11.5 Se puede observar que la OR marginal y las asociaciones parciales coinciden, lo que indica que la variable clínica no es confusora y podemos recurrir a la distribución marginal, donde concluimos que la variable clínica no tiene relación con la evolución del recién nacido. Utilizar la tabla marginal cumple con el principio científico de parsimonia (si es posible, utilice el modelo más simple) y con la necesaria claridad en toda comunicación. Pero además, tiene el premio de una mayor eficiencia estadística: los intervalos de confianza son más estrechos, más precisos.

11.6 Porque las variables Z y X, por el diseño experimental, sean independientes.

11.7 Para que exista confusión, las variables Z y X deben ser colineales (estar relacionadas) y Z debe ser predictora de la respuesta Y.

11.8 Utilizando los intervalos de confianza al 95%, la interpretación sería la siguiente. En los que tienen dolor suave, el fármaco tiene un efecto positivo moderado: aumenta la disparidad «cura/no cura» entre 1,15 y 7,36 veces. En los que tienen un dolor fuerte, el efecto es más marcado: aumenta la razón «cura/no cura» entre 10,5 y 38,1 veces. En cambio, al realizar el estudio global con todos los casos, se llega a la conclusión contraria: el fármaco disminuye las posibilidades de curarse entre 0,48 y 0,68 veces.

Ajustando por dolor basal, ambos intervalos no se solapan, lo que hace pensar en un efecto diferente del tratamiento según la intensidad del dolor basal: esta variable actúa como modificadora del efecto (dolor y tratamiento interactúan). No tendría sentido ofrecer una estimación conjunta o global del efecto del tratamiento, ya que éste no es el mismo en los dos grupos de dolor basal. Pero, además, mientras que los dos intervalos ajustados se sitúan por encima de 1, lo que indica que los curados se asocian con los tratados, el intervalo sin ajustar se sitúa por debajo del 1, indicador

de la situación contraria: los curados se asocian con los placebos. Es decir, el dolor basal, además de modificador del efecto, es confusor. Esta indeseable confusión ha podido aparecer por el sesgo de selección (los más graves han ido al grupo tratado), resultado de una asignación no enmascarada.

11.9 No hay uniformidad en su uso. Por un lado, el término riesgo hace referencia tan sólo a una cierta probabilidad de un evento negativo, por lo que no especifica si detrás hay una relación causal. Por otro, algunos autores aconsejan, quizá acertadamente, reservar el término «factor de riesgo» para relaciones causales y hablar de factor pronóstico cuando no es así. Dada esta diversidad en el uso del término, cuando un autor se refiera a una variable como factor de riesgo, es muy aconsejable no dar por sentado que la relación es causal, mientras no se diga explícitamente.

11.10 Si el objetivo fuera intervenir, para poder establecer la necesaria relación causal, la mejor respuesta vendría de un diseño experimental en el que se estudiara la edad de cada uno de los padres dejando equilibrada la del otro. Por ejemplo, se aparearían tanto las madres de 20 como las de 40 años, por igual, con padres de 20 y de 40 años. Este diseño no sería ético, por supuesto, pero es que además no tendría sentido práctico en nuestro entorno sociocultural. ¿Se plantea una madre o un padre buscar una pareja «independientemente» de su propia edad? Si una madre potencial decidiera adelantar su edad de maternidad, posiblemente también estaría adelantando la edad de su pareja.