

¿ES POSIBLE AUTOMATIZAR LA INFERENCIA CAUSAL?

Mauricio Suárez

Universidad Complutense de Madrid

Encuentro Ciencia-Filosofía, RAC, Madrid, 13 Mayo 2019

¿Qué es la Inferencia Causal? (I)

El tipo de inferencia que permite establecer causas a partir de los datos asociativos, probabilidades, o correlaciones observables entre sus efectos.

No es trivial: En general ***correlación no es causalidad***

¿Qué es la Inferencia Causal? (II)

- Casos rudimentarios de correlaciones cotidianas que no son causa - efecto:
 - 1) La edad de los trabajadores de una empresa, y su sueldo
 - 2) La altura media de los niños de una clase, y su grado o curso
 - 3) La subida del precio de un alimento básico, y de otros alimentos básicos
- Otros casos más complejos - pero similares - de correlaciones científicas:
 - 4) El campo gravitacional creado por un objeto, y su resistencia al movimiento
 - 5) Los rasgos evolutivos de una especie, y los homólogos en otras especies
 - 6) La baja reactividad de dos gases nobles, p.ej.: Helio ($Z = 2$), y Argón ($Z = 18$)

¿Qué es la Inferencia Causal? (III)

- En general, en estos casos, dos variables (A, B) están correlacionadas, puesto que:

$$P(A / B) \neq P(A / \neg B), \text{ o sencillamente: } P(A / B) \neq P(A)$$

lo que es otra forma de decir que: $P(A \& B) \neq P(A) P(B)$

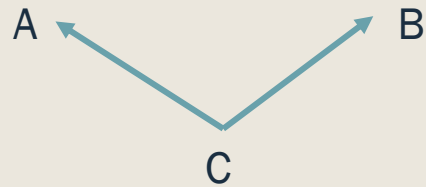
- Muchas variables correlacionadas (A, B) no son causa-efecto:

Ni es cierto “ $A \rightarrow B$ ”, ni tampoco “ $B \rightarrow A$ ”, donde “ \rightarrow ” representa *causalidad*

El Principio de Causa Común (I)

Reichenbach (*The Direction of Time*, 1956): El principio de la causa común (PCC):

“Si acaece una coincidencia improbable [correlación] debe existir una causa común”



C es causa común de A, B si:

- 1) (A, B) están correlacionados: $P(A / B) > P(A) P(B)$
- 2) A no es causa de B, y B no es causa de A
- 3) C apantalla A de B (“C screens off A from B”):

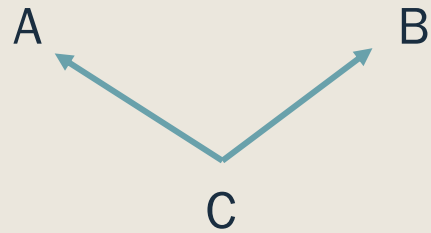
$$P(A / B \& C) = P(A / C)$$

o sea:

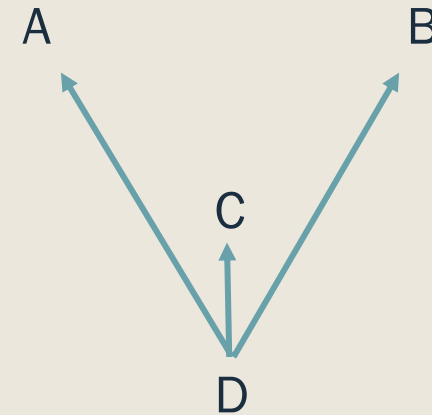
$$P(A \& B / C) = P(A / C) P(B / C)$$

El Principio de Causa Común (II)

La metodología del descubrimiento causal (Suppes, 1972; Spirtes, Glymour & Scheines, 1993; Pearl, 2000)) es *falible*:



C es causa común de (A,B)

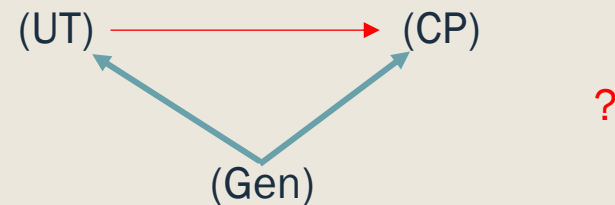


D es causa común de (A,B,C)

El Principio de Causa Común (III)

¿Cómo podemos descartar que $A \rightarrow B$, o que $B \rightarrow A$?

Ejemplo famoso (Doll & Hill, 1950): Correlación entre el uso de tabaco (UT), y el cáncer de pulmón (CP)



(Gen) es causa común de (UT), (CP), si y solo si:

1) (CP, UT) están correlacionados: $P(\text{CP} / \text{UT}) > P(\text{CP})$

2) (CP) no es causa de (UT), y (UT) no es causa de (CP)

3) (Gen) apantalla (UT) de (CP): $P(\text{CP} / \text{UT} \ \& \ \text{Gen}) = P(\text{CP} / \text{Gen})$

o sea: $P(\text{CP} \ \& \ \text{UT} / \text{Gen}) = P(\text{CP} / \text{Gen}) P(\text{UT} / \text{Gen})$

Causalidad y Correlación (I)

BRITISH MEDICAL JOURNAL

LONDON SATURDAY SEPTEMBER 30 1950

SMOKING AND CARCINOMA OF THE LUNG

PRELIMINARY REPORT

BY

RICHARD DOLL, M.D., M.R.C.P.

Member of the Statistical Research Unit of the Medical Research Council

AND

A. BRADFORD HILL, Ph.D., D.Sc.

Professor of Medical Statistics, London School of Hygiene and Tropical Medicine; Honorary Director of the Statistical Research Unit of the Medical Research Council

In England and Wales the phenomenal increase in the number of deaths attributed to cancer of the lung provides one of the most striking changes in the pattern of mortality recorded by the Registrar-General. For example, in the quarter of a century between 1922 and 1947 the annual number of deaths recorded increased from 612 to 9,287, or roughly fifteenfold. This remarkable increase is,

whole explanation, although no one would deny that it may well have been contributory. As a corollary, it is right and proper to seek for other causes.

Possible Causes of the Increase

Two main causes have from time to time been put forward: (1) a general atmospheric pollution from the exhaust

Causalidad y Correlación (II)

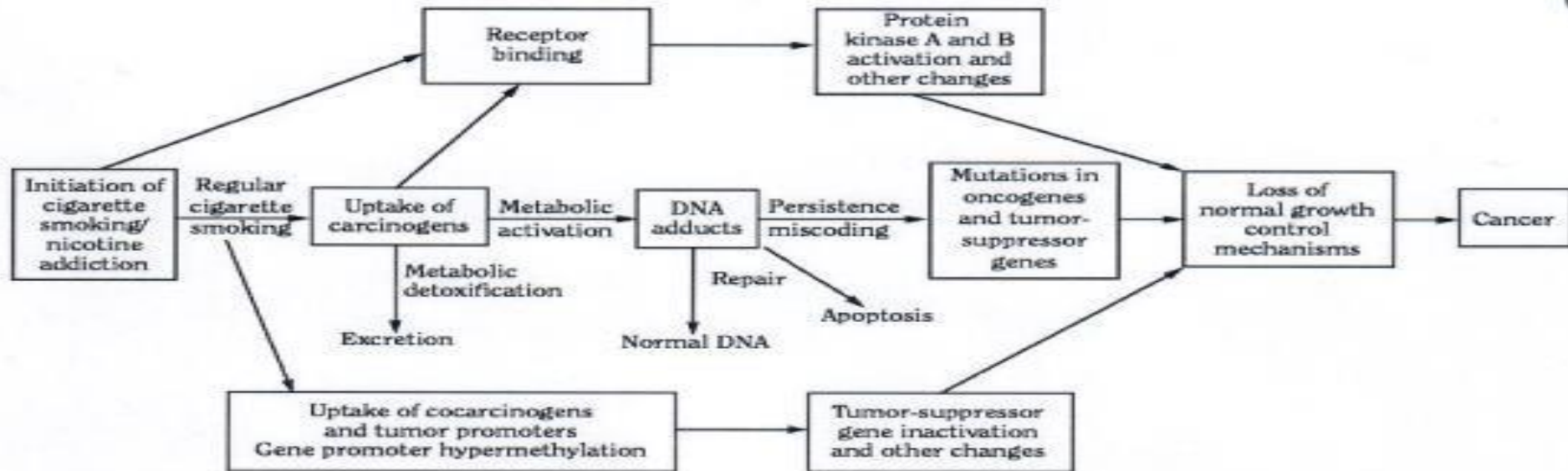


Figure 5.1 Link between cigarette smoking and cancer through carcinogens in tobacco smoke

From: 5, Cancer



How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Centers for Disease Control and Prevention (US); National Center for Chronic Disease Prevention and Health Promotion (US); Office on Smoking and Health (US). Atlanta (GA): Centers for Disease Control and Prevention (US); 2010.

Copyright Notice

NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health.

Causalidad y Correlación (III)

“No Causes In, No Causes Out”

Es imprescindible conocer el mecanismo causal para poder deducir una causa directa, o una común, como explicación correcta de cualquier correlación entre dos variables



Conclusiones

:: El *patrón* de la inferencia causal en física cuántica, química y química cuántica, biología evolutiva, ecología y otras ciencias naturales

:: La inferencia causal *no puede reducirse* a un algoritmo que extraiga la estructura causal a partir del mero análisis de correlaciones entre variables

:: El juicio experto con respecto a la plausibilidad de los mecanismos causales subyacentes a las correlaciones es *falible* pero *imprescindible*

:: *No es posible automatizar la inferencia causal*