

EL MODELO DE LAZARSFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS (Guía de clase)

FICHA N° 14

A. En forma muy sintética, el modelo de análisis que Paul Lazarsfeld desarrolló hacia mediados de los años 40 y que presentara en la reunión anual de la Asociación Americana de Sociología en 1946 se propone un método para responder a la siguiente pregunta general:



Una vez que se ha observado una relación entre dos variables categóricas (nominales u ordinales), ¿cómo se puede establecer estadísticamente si esa relación es válida o si por el contrario, es espúrea?



Se comprende que el límite del análisis bivariado en todas sus formas más elementales es la incapacidad de proponer una respuesta a esta pregunta. El análisis multivariado, por lo tanto, no sólo constituye una forma de abordar una teoría multidimensional sino también una forma estadística de establecer si es válida cualquier relación bivariada de interés que haya sido observada empíricamente.



La noción de espureidad como opuesta a la noción de validez, es por lo tanto central. En principio, si se halla mediante cualquier método que una relación entre dos variables no se mantiene luego de un adecuado control estadístico, se puede afirmar que una razón de esta invalidez es la espureidad.

❖ El ejemplo clásico de “cátedra” es la relación entre *nacimientos* (Y) y *cigüeñas* (X). La fuerte relación originalmente (en adelante, RO) observada desaparece cuando se controla la relación mediante la variable *área geográfica* (T). En las áreas rurales hay tanto más cigüeñas como nacimientos; lo contrario sucede en las áreas urbanas. De aquí que se observe la RO resulte espúrea por ser causada por una variable que antecede a las otras dos.

❖ En el campo de la investigación actual, este problema está presente en prácticamente todos los análisis que persiguen algún interés sustantivo (teórico o político). Generalmente, se utiliza el término de **sesgo** y de relación **sesgada** para referirse a una RO que desaparece luego de un control apropiado .








En el análisis las nociones de relación espúrea, válida o sesgada adquieren su mayor importancia cuando las proposiciones teóricas encierran una pretensión causal.

❖ En los términos en que se está manejando el autor, la noción conecta con la escuela empirista inglesa de Hume y de Stuart Mill.

EL MODELO DE LAZARSFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS (Guía de clase)

- La causalidad para estos autores es una noción filosófica, ideal, no empírica que se establece cuando se han verificado observacionalmente tres condiciones:
 - i) Contigüidad espacial entre ambos eventos observados.
 - ii) Precedencia y sucesión temporal del evento causa ("C") respecto del evento efecto ("E").
 - ii) Regularidad en la ocurrencia: siempre que ocurre "C", "E" ocurre, y siempre que "C" no se verifica, "E" no se verifica.
 - iv) C es causa de E siempre que se hayan descartado todas las otras posibles alternativas que puedan causar a E.

B. La lógica del modelo puede ser expuesta en forma relativamente simple a partir del análisis simultáneo de tres variables.

-  Cuando se ha constado la presencia o la ausencia de una relación entre dos variables ("X" e "Y"), el siguiente paso lógico consiste en analizar el papel que puedan estar jugando otras variables en esta **relación originaria**. Estas reciben el nombre de variables - **test** (T).
-  La lógica del modelo consiste en indagar qué sucede con la relación originaria si esta es analizada en cada uno de los sub-espacios diferentes definidos por las categorías de aquella tercera variable.
-  Los resultados obtenidos en el análisis de la relación originaria (X,Y) dentro de esos sub-espacios se denominan **relaciones parciales**.. Su notación es en el caso de una variable T dicotómica: (X,Y;T) y (X,Y;T').
-  La lógica general del procedimiento se completa con la construcción de cuadros adicionales para analizar cuál es la relación **marginal** entre cada una de las variables originales con la variable - test. La notación aquí es (T,X) y (T,Y).
-  El modelo Lazarsfield se basa en asignar a las variables papeles conceptualmente asimétricos . La variable objeto de explicación es denominada como **dependiente (Y)**. La variable originalmente propuesta como explicativa se le denomina **independiente (X)** y la tercera variable tiene, para ese análisis, un papel de **control (T)**. Detrás de esta nomenclatura está resonando la clásica clasificación de diseños en experimentales, pre-experimentales y cuasi-experimentales.

EL MODELO DE LAZARFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS
(Guía de clase)

 **Relación original bi-variada (X,Y)**

Variable Y	Variable X		Total
	Categoría X	Categoría X'	
Categoría Y			(n _y)
Categoría Y'			(n _{y'})
Total	100% (n _x)	100% (n _{x'})	100% (N)

 **Relaciones parciales (P) (X,Y; T) y (X,Y;T')**

Variable "Y"	Variable "T"				Total
	Categoría T		Categoría T'		
	Variable "X"		Variable "X"		
	Categoría X	Categoría X'	Categoría X	Categoría X'	
Categoría Y					
Categoría Y'					
Total	100% (n _a)	100% (n _b)	100% (n _a)	100% (n _b)	100% (N)

 **Relaciones marginales (M)**

(X,T)

(Y,T)

Variable X	Variable T		Total	Variable Y	Variable T		Total
	T	T'			T	T'	
X				Y			
X'				Y'			
Total	100% (n _x)	100% (n _{x'})	100% (N (x,t))	Total	100% (n _t)	100% (n _{t'})	100% (N (y,t))

EL MODELO DE LAZARSFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS (Guía de clase)

C. La ecuación de covarianzas representa la formalización de la lógica expuesta en el anterior apartado y constituye en el instrumento principal que luego se utilizará para realizar una interpretación sistemática de los hallazgos.

- El modelo Lazarsfeld se representa matemáticamente en la siguiente ecuación para el caso de tres variables dicotómicas, donde la relación original se descompone en la suma ponderada de dos relaciones parciales (P) y el producto de las relaciones marginales (M).

$$(X,Y) = (X,Y; T) \oplus (X,Y; T') \oplus [(T,X) \otimes (T,Y)]$$

- Cada uno de los términos de la anterior ecuación es el coeficiente de asociación calculado en cada una de las 5 tablas.
- Es importante recordar que se debe usar **siempre el mismo coeficiente** de asociación seleccionado para la relación originaria.
- Los signos de “+” encerrados entre círculos indican que se trata de una suma ponderada. Los ponderadores de esta suma son la proporción de casos sobre el total en cada tabla parcial:

$$n_t / N$$

$$n_{t'} / N$$

- Tanto Lazarsfeld como Kendall en sus trabajos originales de los años 50, descartaron aplicar los ponderadores para interpretar los resultados, dado que estaban interesados en los “casos puros” de elaboración.
- En la práctica usual se dice entonces que se trata de una “suma cualitativa” donde en el fondo no interesa demasiado corroborar la igualdad, salvo para chequear de que no se hayan cometido errores de cálculo, etc.
- La ecuación implica una abstracción donde se están comparando los resultados de las cinco tablas construidas (para el caso de tres variables dicotómicas). Es necesario haber comprendido bien qué sucede en estas para luego pasar a analizar aquella.

EL MODELO DE LAZARFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS
(Guía de clase)

D. En la interpretación de los resultados es conveniente comenzar por entender la tipología de casos “casos polares” propuesta originalmente.



Los casos polares se proponen como posibilidades lógicas resultantes de la ecuación de covarianzas examinada más arriba.

• En términos generales, pueden observarse dos grandes tipos de elaboraciones:

[Por parciales] $(X,Y) = (X,Y; T) \oplus (X,Y;T') \oplus 0$

[Por marginales] $(X,Y) = 0 \oplus 0 \oplus [(T,X) \otimes (T,Y)]$

• Una elaboración por parciales implica que no se ha observado relación entre las variables T y X en las tablas marginales con la variable Y.

• Una elaboración por marginales resulta de la anulación de **ambos** parciales. Es decir, ha desaparecido la relación original una vez que se ha controlado por T.



“Sostenemos que hay esencialmente cuatro operaciones que pueden ser practicadas con dos variables originales y una de prueba:”

	Forma estadística (casos polares)	
	Elaboración por Parciales (P) $(X,Y;T) \neq 0; (X;Y;T') \neq 0$ $(X,T) = 0$	Elaboración por Marginales (M) $(X,T) \neq 0$ $(X,Y;T) = 0; (X,Y;T') = 0$
Variable “T” es anterior a la variable “X”	PA	MA
Variable “T” es intermedia a “X” e “Y”	PI	MI

PA - Se refiere a la relación general dada por una condición o disposición (T), un estímulo (X) y una respuesta (Y).

PI - Representa una situación típica en la que en la relación original interviene

EL MODELO DE LAZARSFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS (Guía de clase)

una variable (T). El factor T es denominado *contingencia*.

MA - Representa un caso de relaciones equívocas o espúreas en las que toda la relación originaria se debía a las relaciones marginales entre las variables.

MI - Se trata de una variable intermedia entre X e Y que da cuenta (en el caso típico) de toda la relación originaria.

E. La interpretación con la tipología de Errandonea.

- En los años ochenta, Alfredo Errandonea desarrolló lógicamente la tipología original de Lazarsfeld y Kendall para poder analizar en forma sistemática un importante conjunto de situaciones empíricas que se verifican usualmente.
- La extensión de la tipología se basa en dos tipos de criterios. El primero es interno: se profundiza el análisis de la elaboración por parciales o por marginales preguntando cuál es la relación o magnitud relativa de los parciales entre sí, y de los marginales entre sí. El segundo criterio es incorporar un nuevo tipo de elaboración: la "mixta".
- Hecha esta sofisticación, Errandonea introduce nuevos términos para las celdas creadas y discute la propiedad de algunos otros términos clásicos de Lazarsfeld y Kendall.
- El cuadro siguiente está construido en base a una adaptación de la Guía de clases n°9 usada tanto en la UDELAR como en la UBA por el mencionado Profesor. Se ha seguido el criterio de Lazarsfeld de distinguir en la tabla el papel lógico causal que la teoría se le asigna a "t".

EL MODELO DE LAZARSFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS
(Guía de clase)

- Tipos de elaboración (Errandonea "Guía 9") -

		"T" ANTECEDENTE	"T" INTERVINIENTE
Elaboración por parciales $M_1 * M_2 = 0$	$(P_1 = P_2) \text{ y } (P_1 + P_2) = (x,y)$ $M_1 = 0 \text{ y } M_2 = 0$	irrelevancia de "t"	irrelevancia de "t"
	$(P_1 < P_2) \text{ ó } (P_1 > P_2)$	especificación	especificación
	$P_1 = 0 \text{ ó } P_2 = 0$	condición (Antecedente)	condición (contingencia)
	$(x,y) = 0$ $P_1 < (x,y) < P_2 \text{ ó } P_2 < (x,y) < P_1$	variación o cambio de contexto	variación o cambio de contexto
Elaboración mixta $P_1 \neq 0 \text{ ó } P_2 \neq 0 \text{ y } M_1 \neq 0 \text{ ó } M_2 \neq 0$	$(P_1 < P_2) \text{ ó } (P_1 > P_2)$ $M_1 * M_2 = 0$ $(x,t) \neq 0 \text{ y } (t,y) = 0$	especificación	especificación
	$(P_1 = 0) \text{ ó } (P_2 = 0)$ $M_1 * M_2 = 0$ $(x,t) \neq 0 \text{ y } (t,y) = 0$	condición	condición
	$P_1 = P_2 \text{ y } M_1 * M_2 = 0$ $(x,t) \neq 0 \text{ y } (t,y) = 0$	(intermediación total de x)	efectos independientes de X
	$(P_1 = P_2) \text{ y } (x,y) M_1 = 0$ $(t,y) \neq 0$	determinantes independientes de Y	determinantes independientes de Y
	$(x,y) = 0$ $P_1 < (x,y) < P_2 \text{ ó } P_2 < (x,y) < P_1$ $(x,t) \neq 0 \text{ ó } (t,y) = 0$	variación de contexto	variación de contexto
	$(x,y) > (P_1 + P_2)$ $(P_1 = P_2) \text{ y } M_1 \neq 0 \text{ y } M_2 \neq 0$	resultantes relacionadas sistema circular autosustentado	doble vía de determinación
Elaboración por marginales $P_1 = P_2 = 0$	$M_1 \geq (x,y) \text{ y } M_2 \geq (x,y)$	MA "espureidad"	MI intermediación

EL MODELO DE LAZARSFELD (1946) PARA EL ANÁLISIS DE TRES VARIABLES DICOTÓMICAS (Guía de clase)

Bibliografía básica para este tema:

CORTES, Fernando & RUBALCABA, Rosamaría

1987 *Métodos estadísticos aplicados a la investigación en ciencias sociales*. El Colegio de México. México. DF.

ERRANDONEA, Alfredo

1987 *Guía de clases para Metodología III*. Licenciatura en Sociología. UDELAR. Montevideo. [Hay una versión de las mismas para la UBA, Buenos Aires].

KENDALL, Patricia

1971 "La introducción de variables adicionales y la elaboración del análisis". En HAYMAN, H. (1971) *Diseño y análisis de las encuestas sociales*. Editorial Amorrortu. Buenos Aires. [Original 1955 *Survey Design and Analysis. Principles, Cases and Procedures*. The Free Press. NY]

LAZARSFELD, Paul

1966 "La interpretación de las propiedades estadísticas como propiedad de investigación". En BOUDON, R. & LAZARSFELD, P. (1966) *Metodología de las ciencias sociales. Tomo II: análisis empírico de la causalidad*. Editorial Laia. Barcelona. Ponencia presentada en 1946 al Congreso de la Sociedad Americana de Sociología (ASA) en Cleveland.