

El papel lógico de las magnitudes de las relaciones en el análisis multivariado de la asociación con parciales y marginales

Alfredo Errandonea (h.)

i. Introducción

1. En las últimas décadas se han incorporado o desarrollado diversos procedimientos de análisis multivariado en la práctica de la investigación social, con variables grados de sofisticación estadística. Curiosamente, y en forma paralela, también se ha dado un crecimiento del cuestionamiento de la legitimidad en cuanto a la aplicación de las técnicas cuantitativas para la interpretación de fenómenos sociales; por lo menos —o de manera más intensa— para ciertos niveles de interpretación.¹

Es posible hipotetizar cierto tipo de relaciones entre ambos incrementos tan ostensiblemente contrapuestos, pero ello me llevaría lejos de mi interés aquí. Sólo permítaseme suponer que cierta desmesura en la tendencia a la fácil aplicabilidad de nuevos procedimientos a que la accesibilidad informática invita, puede conducir al descuido de las limitaciones y restricciones que algunos vistosos instrumentos implican. De allí a la inconsistencia y vulnerabilidad de ciertos resultados que se pretenden “probatorios” pueden quedar distancias muy cortas; sorteadas con demasiada facilidad y frecuencia para variables y relaciones de naturalezas fuertemente acotadas de restricciones.

2. Más allá de las modas, me sigue pareciendo altamente reivindicable para la mayor parte de los análisis que hacen los científicos sociales cuando se manejan con un número limitado de variables, el

clásico procedimiento multivariado de control por parciales, tan utilizado desde los tiempos de Durkheim, que Lazarsfeld formalizara a partir de 1946 (Lazarsfeld en Lazarsfeld y Rosenberg, 1955; Lazarsfeld en Lerner, 1959). Porque se trata de un procedimiento que permite discriminar los diferentes papeles lógicos de cada variable, hasta donde ello es posible hacerlo estadísticamente, cualquiera sea la naturaleza de las variables y las relaciones. Y porque la artesanía de la elaboración que requiere e implica, lejos de constituir un inconveniente, le otorga la ventaja de consistir en una lectura mucho más analítica de los datos, que entrega observaciones desagregadas, con mayor elocuencia descriptiva de las relaciones involucradas. Y éstas, a partir de ciertas asunciones teóricas, posibilitan las inferencias explicativas que procura el investigador.

3. Hace ya bastante tiempo de que llegué a la conclusión sobre la importancia prioritaria de tratar de avanzar en el instrumental de análisis a partir de este enfoque, en procura del desarrollo de esta técnica de control mediante apertura matriciales de cuadros bivariados.² Aunque ella aparezca como

1 No me ocuparé aquí de la falsa oposición que se ha pretendido establecer entre técnicas cuantitativas y técnicas cualitativas, asunto sobre el cual me ocupé en otros trabajos (Errandonea, 1985; Errandonea y Supervielle, 1986).

2 Durante varios años de docencia de Metodología y Técnicas de Investigación en la Argentina (Carreras de C. Política y de Sociología de la Universidad del Salvador —1976/1984—, y de Sociología de la Universidad de Buenos Aires —desde 1985—, principalmente), en la organización y dictado de las clases, fui trabajando un desarrollo interpretativo de este instrumental de análisis multivariado. Al punto de que circulan versiones de apuntes de clases y con fines docentes en una Guía de Clase (Errandonea, 1982; última versión: 1987). Este artículo pretende ser el comienzo de una sistematización más formalizada de estas reflexiones.

demasiado simple y –por ello mismo– no susceptible de mayor avance, estoy persuadido de que se adecúa mucho mejor al tipo de variables que manejamos en la mayor parte de la investigación social y a la tosquedad del dato “real” que a partir de ellas podemos elaborar.¹

Con singular retraso en relación a la edad que en mi tienen estas preocupaciones, con este trabajo me propongo iniciar la concreción de las reflexiones hechas.

II. La Ecuación de Lazarsfeld

4. Obviamente, la expresión ecuacional de la suma ponderada en que Lazarsfeld descompone la relación originaria cuando ella es controlada por “t”, puede escribirse:

$$XY = xy;t \frac{n_t}{n} + xy;t' n_t' + Xt \cdot ty$$

Lazarsfeld se interesa exclusivamente por la consideración analítica de cada término y no por la fórmula en general.²

Pero aquí me interesa la ecuación entera en la medida de que utilizaremos las magnitudes de las fuerzas de las relaciones. Como ellas se miden mediante coeficientes que constituyen normalizaciones, los parciales matriciales deben ser ponderados por la proporción del “n” que corresponde a cada uno en relación al “n” total.

Podemos entonces razonar a partir de la propia idea global de la ecuación.

5. Si la igualdad que expresa la ecuación no se cumple, debe presumirse distorsión en alguno o algunos de sus términos. Ella puede obedecer al carácter espureo, por ejemplo, de alguna de las relaciones que constituyen términos de la ecuación. Si así fuera, existiría alguna variable fuera del modelo que en realidad debería haber sido incluida en él, y que por alguna razón no visible con su exclusión, sólo se hace presente a través de su espuriedad. En ese caso (no cumplimiento de la igualdad de la ecuación) puede afirmarse que el conjunto de variables consideradas –y por lo tanto, el modelo puesto a prueba es inadecuado.

Pero esta prueba de adecuación es asimétrica: tan sólo descarta modelos inadecuados, pero no permite presumir adecuación en modelos no descartados. Por lo tanto, su eventual utilidad práctica es demasiado relativa.³

6. A partir del cumplimiento de la ecuación –la relación originaria como equivalente a la suma de relaciones parciales y producto de relaciones marginales–, está claro que cada uno de los componentes adicionados del segundo término es inferior o igual a la relación originaria o primer término de ella.

Desde que no todos los componentes del segundo término pueden ser simultáneamente iguales al primero, tenemos un primer tipo de situaciones en que alguno de ellos es efectivamente igual al primero, y por lo tanto desaparece el otro componente. Aquí deben distinguirse primariamente dos clases de casos: aquellos en que es la suma ponderada de los parciales la que se iguala con la relación originaria y los marginales los que desaparecen o se anulan; y aquellos otros en que es el producto de los marginales el que se equipara con la relación originaria y los parciales los que desaparecen. Pero además hay algunas variantes para la primera de las dos clases de casos vistos: cuando existiendo un marginal, este igual desaparece anulado por el producto de otro inexistente (o casi), en cuyo caso los parciales tienden a sumar también equiparadamente con la relación originaria.

El otro tipo de situaciones está constituido por los casos en que hay presencia de relaciones adicionales tanto en los parciales como en los marginales. Por lo tanto, los componentes a adicionar para componer el segundo término, en sí, son inferiores a la relación originaria, aunque puedan a su vez integrarse con alguna relación que la exceda. Aquí también deben distinguirse primariamente dos subclases de casos: aquellos en los cuales los parciales son similares entre sí y aquellos otros en que difieren de forma neta (incluido el caso de la ausencia de alguno). Además, también se puede discriminar entre las magnitudes de parciales y totales.

7. Como se ve, pueden distinguirse tipos, clases y variantes de composición de la ecuación según las magnitudes de las relaciones y sus referencias, y correspondencias mutuas. A partir de ellas, puede constituirse una tipología de resultados que avance en desarrollo de capacidad interpretativa sobre la célebre formulada por Lazarsfeld. Porque las diferentes composiciones referidas suponen una lógica global de relacionamiento de las variables incluidas

1 Naturalmente de que el dato se elabora por medio de una combinación de procesos de abstracción, para los cuales la realidad no constituye más que el referente empírico. Pero tanto los indicadores mediante los cuales efectivamente se realiza esta referencia fáctica, como la propia conceptualización y naturaleza “cualitativa” de las variables manejadas, justifican sobradamente esta afirmación. Mas adelante, trato de fundarlo

2 Lazarsfeld expresó su ecuación “ponderada”, escribiéndola así (Lazarsfeld, 1955):

$$(XY) = (xy;t) + (xy;t') + (Xt) \cdot (tY)$$

3 Por otra parte, puede haber otros motivos de exacto incumplimiento de la igualdad de la ecuación (Véase en Numeral 23).

en el modelo, que se expresa en las magnitudes de las relaciones involucradas. Es en esta línea en la que he trabajado para desarrollar una tipología de interpretación de resultados; cuya formalización me propongo. Pero ella debe comenzar por ciertas proposiciones tanto conceptuales metodológicas como de significación lógica de determinados resultados estadísticos.

Este inicio es el objeto del presente trabajo.

3. Los papeles lógicos posibles de la variable de control

Puede considerarse una obviedad tautológica que cualquier análisis explicativo se propone dar cuenta explicativa de alguna variable. Por lo tanto, parto de considerar como variable necesariamente independiente en la relación originaria a la que "ubicamos" como tal. Se eliminan así, gran número de combinaciones posibles.

Y, por lo pronto, no tendré en cuenta *ese* papel entre los posibles de la variable de control (lo cual no quiere decir que no pueda llegar a resultar dependiente de la dependiente de la relación principal u originaria; pero en ese caso, obviamente, no intervendría en la explicación que es objeto del análisis).

Restan los demás papeles posibles, todos los cuales deben ser considerados. Para lo cual, los redifino aquí en función de su significación como resultado posible del análisis en su ubicación de única variable de control (tercera variable; a la que designaremos, siguiendo la convención, como "t").

9. La variable incorporada como de control ("t") podría llegar a resultar la verdadera *variable determinante o independiente* (o una de ellas, en situación de codeterminación). Se entiende por tal aquella variable (o una de aquellas) a que puede atribuirse el efecto o fenómeno estudiado; estadísticamente, aquella (o una de aquellas) a partir de la cual es previsible, anticipable.

En consecuencia, puede considerarse la eventualidad de que la variable que ocupa la posición de control ("t") sea variable independiente o determinante, cuando la magnitud de su relación con la variable dependiente ("Y") sea netamente superior a la originaria ("Y Y") y sus parciales (si no lo es tanto, podría ser codeterminante con "X"). Pero hasta aquí esta eventualidad es una hipótesis bivariada, que requerirá ser contrastada mediante la lectura con "t" girada a la posición variable independiente (de "x"), y "x" a la posición de "t"; obteniendo así una lectura de "t Y" realmente travariada, bajo el control de "X" (además de las otras que se requiera testar como hipótesis alternativas).

10. Los textos más clásicos suelen no ser excesivamente prolijos en el uso terminológico referido a variables que participan de la relación *especificán-*

dola o condicionándola. Voy a utilizar la expresión "*variable interviniente*" para identificar este papel; aunque ella es utilizada con habitualidad -y a mi juicio, incorrectamente- para referir situaciones en las cuales la variable considerada intermedia la relación entre otras dos, como se verá.

Entonces, defino al papel de "*variable interviniente*" como el de aquella variable que participa en una relación posibilitándola, condicionándola, facilitándola, obstruyéndola, mitigándola; pero que en ningún caso, por sí sola, produce el efecto en que consiste la variable dependiente. En especificación o condicionamiento puede ser temporalmente precedente a la situación constitutiva de la relación principal, o no. En el primer caso, la literatura clásica se afana por denominarla "variable antecedente" y atribuirle especificidades lógicas que lógicas que no alcanzo a ver (Salvo, claro, que se confunda prelación "temporal" con "lógica"; pero entonces debe despojarse su conceptualización de la connotación cronologista y ubicarla en el papel lógico que le corresponde). Por lo que no me sumaré a la distinción del papel de "Antecedencia", que tal como se define habitualmente, o es efectivamente interviniente sin diferencia lógica ni estadística alguna, o es variable independiente de una relación intermediada.

11. A la definición delimitativa conceptual del universo a estudio, en tanto noción lógica clasificatoria, se le suele denominar "*variable contextual*"; aunque por cierto que para la investigación concreta de que se trata se transforma en una constante (sólo se opera con un solo valor de ella). En otro "contexto" suele denominarse variable contextuales a aquellas que otorgan una característica (valor de la variable) a unidades que son colectivos ecológicos (en el llamado "análisis contextual"). Ambas conceptualizaciones aportan ingredientes a un específico papel "*contextual*" que queremos definir aquí: cuando los diferentes valores de la variable de control ("t") cambien el sentido (en el caso de relaciones lineales: el signo) de la relación original (entre "X" e "Y"), le atribuiremos el papel de cambiarle la significación contextual a la relación principal.

12. Como se ha dicho (nal. 10), suele haber confusión terminológica para identificar la situación en que la participación de una variable en una relación consiste en constituirse en trasmisor intermedio de una relación principal entre una variable independiente ("X") y una dependiente ("Y"), y *no* en condicionante o especificador de esa relación principal. A ese papel suele denominársele "interviniente", lo que me parece doblemente infeliz: a) porque lo confunde con los efectos especificante o condicionante, ya que la idea de "antecedencia" con ese sentido es incorrecta; y b) porque terminológicamente esta participación no opera como alterante que

es lo que la expresión "interviniente" sugiere (y es una adecuada sugerencia para la especificación o condicionamiento), sino como mera transmisión, pasaje o mediación. Para este papel, pues, creo que debe utilizarse la denominación de "variable intermedia-ria".

Entonces, reservo el nombre de *variable intermedia-ria* para aquella por la cual pasa toda la determinación entre una variable independiente ("X") y otra dependiente ("Y"). O sea, el papel que la ubica como dependiente de la independiente e independiente de la dependiente.

Su distinción con el papel de variable intervi-niente puede graficarse así.



13. Un caso especial de intermediación de "t" al que vale la pena consignar por separado por la significación diferente que asume en el conjunto, es cuando esa intermediación corresponde a una cadena cerrada de tres variables, cada una de las cuales desempeña ese mismo papel porque constituyen un sistema circular autosustentado en el cual la determinación se recicla. Aquí, entonces, "t" se constituye como variable dependiente de "Y" e independiente de "X" (la circulación inversa queda descartada al presuponer a "Y" siempre como dependiente en relación principal –numeral 8–).

4. La contribución de las magnitudes de las relaciones para interpretar el papel lógico de la variable de control ("t")

4. En definitiva, la interpretación del papel a atribuir a la variable de control en el juego total de las tres variables –y por lo tanto, también a las otras– constituye una tarea conceptual que implica una responsabilidad teórica. Pero el resultado, como queda dicho, debe ser consistente con el referente empírico; que aquí son los datos resultantes del análisis estadístico.

Lo que me importa subrayar aquí es que las magnitudes obtenidas en las medidas de las diferentes relaciones bivariadas simples y controladas involucradas en este tipo de análisis, consideradas en sus comparaciones mutuas, constituyen poderosos indicios instrumentales para aquella tarea interpretativa. Razón por la cual, su consideración me parece tan vital y rica como la propia lectura analítica de los

cuadros. Ambos procederes aportan fertilidad a la elaboración artesanal de la interpretación.

15. Mas que la propia formalización de una tipología interpretativa –que de cualquier manera vamos a concluir en el contexto de otros trabajos– importa explicitar las proposiciones generales a partir de las cuales pueden elaborarse tales tipos, así como orientar el análisis en cualquier caso concreto.

Si parto de que la igualdad de la ecuación de Lazarsfeld se cumple, las proposiciones generales para el análisis pueden esquematizarse según el siguiente desarrollo.

16. La relación originaria (entre "X" e "Y") es un buen punto de partida si en ella se dan las mayores magnitudes de relaciones, en comparación con las que entregan cada una de esas variables con la de control. *La relación originaria será realmente relación principal si las mayores magnitudes se mantienen a nivel de los Parciales, especialmente en relación con el producto de los marginales:*

$$xy;t_1 \frac{n_{t1}}{n} + xy;t_2 \frac{n_{t2}}{n} \geq Xt \cdot tY$$

Aún en el caso de que no sea así, pero el marginal vertical ("tY") sea superior a la suma ponderada de los parciales, es evidente que el carácter de determinante principal de "X" queda en entredicho.

En los casos en que la relación originaria no presente el grado de prevalencia indicado en la expresión de las magnitudes (incluyendo la variante mencionada recién), será de orden realizar el análisis mediante la lectura en que se gira "t" a la posición de "X" y viceversa. Se tratará siempre de asumir el punto de partida de que la relación originaria sea la relación principal.

Por otra parte, el procedimiento girado de lectura siempre proporciona elementos de juicio adicionales, especialmente cuando la primera lectura arroja resultados poco convincentes lógicamente y/o teóricamente.

17. Se ha dicho –con razón– que este tipo de análisis tiene cierto grado de inspiración en el diseño experimental. Cuando en un grupo experimental no aparecen diferencias en relación con el correspondiente grupo testigo en lo que tiene que ver con un determinado efecto, luego de que el primero ha sido sometido a un cierto estímulo no expuesto al segundo, se dice que la variable que representa dicho estímulo no tiene nada que ver con el efecto indagado. O sea, que es ajeno a él.

Los parciales son asimilables al grupo experimental y al grupo testigo en la exposición al estímulo que los segrega ("t"). Entonces, cuando los parciales se mantienen iguales entre sí e iguales a la relación originaria, debe concluirse que la variable de control ("t") es totalmente ajena a la relación originaria en

sí. No la condiciona, ni la altera, ni la intermedia. Tal *ajenidad*, en referencia a la relación originaria en sí, no se modifica cualquiera que sean las relaciones que las variables de la relación originaria pudieran tener con "t".

De cualquier manera, la situación típica será:

$$XY = X_{y;t_1} \frac{n_{t_1}}{n} + X_{y;t_2} \frac{n_{t_2}}{n} + 0$$

o donde le 0 resulta de: $X_y \cdot tY = 0$ y

$$X_{y;t_1} \frac{n_{t_1}}{n} = X_{y;t_2} \frac{n_{t_2}}{n}$$

18. Por el contrario, cualquiera sea lo que ocurra a nivel de los marginales, cuando al interior de las matrices parciales resultantes del desagregado por la variable de control, entre todas o algunas de ellas se manifiesten diferentes magnitudes de relación entre las variables de la relación originaria, deberá concluirse que la variable de control esté interfiriendo a la relación originaria: posibilitándola, condicionándola, estimulándola, obstruyéndola. Por lo tanto, se trata de una *variable interviniente* (aquí con asunción de la connotación modificativa que la designación sugiere).

Como queda dicho, esta conclusión no se altera cualquiera sea la o las relaciones bivariada(s) que la variable de control evidencie con alguna o ambas de las variables de la relación originaria.

En este caso, la fórmula de Lazarsfeld en su suma ponderada tendría únicamente la particularidad de que:

$$X_{y;t'} \frac{n_{t'}}{n} > XY > X_{y;t''} \frac{n_{t''}}{n}$$

o viceversa.

El único efecto diferencial de que $X_{t;tY}$ sea igual o diferente de 0, será –respectivamente– la coincidencia o leve inferioridad de la suma ponderada de los parciales con respecto a la relación originaria bivariada. (Tal inferioridad puede llegar a no ser leve, en algunos casos límites).

En el límite, cuando alguno o algunos de los parciales llevan su diferenciación para con otro u otros al grado de desaparecer (convertirse en 0), esta "especificación" que realiza la variable "t", se convierte en "condición".

19. Aún manteniéndose igual –o muy similar– las magnitudes de los parciales entre sí, puede ocurrir que la variable de control "t" esté jugando un papel importante en la relación entre las variables originales "X" e "Y". Ello sucede cuando esa similitud de magnitud de parciales, sin embargo resulta colocarse claramente por debajo de la de la relación originaria ("XY"). Para que ello acontezca, si el modelo es adecuado, será necesario que el producto de margi-

nales sea diferente de cero: exactamente igual a la diferencia entre los equiparados parciales y la relación originaria.

Si este es el caso, debe inferirse que "t" no está "interviniendo" para alterar (especificar o condicionar) la relación entre "X" e "Y". O sea que existen *dos vías de determinación complementarias*: una *directa* ("XY") cuyo quantum se expresa en la suma ponderada de los equiparados parciales; y una *indirecta a través de la intermediación de "t"*, de magnitud medible a través del producto de los marginales (o sea, igual a la diferencia entre las sumas ponderada de los parciales y el volumen de la relación originaria).

$$XY = X_{y;t_1} \frac{n_{t_1}}{n} + X_{y;t_2} \frac{n_{t_2}}{n} + X_{t;tY}$$

donde

$$XY > X_{y;t_1} \frac{n_{t_1}}{n} = X_{y;t_2} \frac{n_{t_2}}{n}$$

y

$$X_{t;tY} = XY - \left(X_{y;t_1} \frac{n_{t_1}}{n} + X_{y;t_2} \frac{n_{t_2}}{n} \right)$$

20. Aunque poco común, una variante curiosa de la incidencia de una variable de control en una relación original, es el caso en que esta se aparece a primera vista como encubierta o inexistente, y sólo se manifiesta en el desagregado por parciales. Ello ocurre en una forma distinta de cumplirse la igualdad entre ella y la suma ponderada de los parciales: cuando se opera algebraicamente. La relación originaria es cero o próxima (en todo caso, prácticamente es expresión de lo que ocurre con el producto de los marginales) y cada parcial tiene signos contrarios, de manera tal que se compensan.

$$XY = \left(+ X_{y;t_1} \frac{n_{t_1}}{n} \right) + \left(- X_{y;t_2} \frac{n_{t_2}}{n} \right) + X_{t;tY}$$

$$\text{donde } \left(+ X_{y;t_1} \frac{n_{t_1}}{n} \right) + \left(- X_{y;t_2} \frac{n_{t_2}}{n} \right) = 0$$

$$\text{léase } X_{y;t_1} \frac{n_{t_1}}{n} - X_{y;t_2} \frac{n_{t_2}}{n} = 0$$

En este caso, la variable de control ("t") indica un cambio de contexto (*variable contextual*), que implica inversión de efecto; es decir, cambio de sentido de la relación originaria expresado por la alteración de signo. Lo cual hace a esta como globalmente "encubierta", no visible a nivel bivariado.

Desde luego que, a estos efectos, poco importa lo que pase con los marginales; los cuales darán lugar a variantes en la tipología.

21. Finalmente tenemos el resultado al cual D en "El Suicidio" convertiría en clásico: la despari-

ción de la relación originaria ("XY") en los parciales ("xy;t'" y "xy;t'"):

$$XY = 0 + 0 + X_t \cdot tY$$

porque

$$xy;t' \cdot \frac{n_t'}{n} = 0 ; xy;t'' \cdot \frac{n_t''}{n} = 0$$

por lo tanto

$$XY = X_t \cdot tY$$

O sea que toda la relación detectada a nivel bivariado como relación originaria ("XY") se refiere en realidad a la variable introducida como de control ("t"). Y ello puede deberse a dos razones.

a) En primer lugar, puede ocurrir que la relación originaria bivariada se debía a que ambas variables principales, en realidad, son efectos de una tercera introducida como control ("t"), la cual es su verdadera determinante. Al exponerse la aparente relación bivariada al desagregado por "t", ella desaparece en sus parciales. Estamos frente a una evidente relación "espurea".

b) En segundo lugar, este mismo resultado también puede deberse a que toda la relación originaria "XY" pase a través de "t". En este caso, "t" es variable "intermediaria". El que toda la varianza de "Y" que explica "X" pase a través de "t", hace que cuando "XY" es controlada por "t", en los parciales de ésta, la relación originaria desaparezca.

El que sea uno u otro caso ("espureidad" o "intermediación") es una cuestión de decisión teórica. Todo consiste en saber si conceptual y lógicamente "X" puede determinar a "t"; o por el contrario, teóricamente hay que postular que "t" determina a "X". O sea que está en juego el sentido de las flechas que sólo la teoría puede decidir.

Sin embargo hay algún indicio empírico que puede concurrir a corroborar la interpretación, a partir también de las magnitudes de las relaciones.

Lógicamente, para el caso de la intermediación, se requiere que cada uno de los marginales (X_t y tY) sean individualmente, por lo menos, iguales a la relación originaria. Si es que ellos sucesivamente van a conducir la determinación que X tiene sobre Y, su magnitud no puede ser inferior a ella. Por la misma razón de que dos cables que sucesivamente tienen que transferir una determinada corriente eléctrica, cada uno de ellos debe ser capaz (tener la dimensión suficiente) de transportar el total de corriente que alternativamente conducirán, so pena de fundirse cualquier de los dos que no reúna dicha condición.

Por lo tanto, para que sea intermediación, el menor de los marginales no puede ser inferior a la relación originaria bivariada. Pero esta es una condición necesaria aunque no suficiente: su incumplimiento descarta la intermediación, pero su

cumplimiento no confirme necesariamente que este sea el pel de "t". Máxime si se tiene en cuenta que la otra alternativa para este resultado (la "espureidad"), el requerimiento de la igualdad del producto de los marginales con la relación originaria, supone suficiente fuerza en aquellos como para comprender numerosos casos en que ambos marginales individualmente también superen a la relación originaria (casi debería suponerse como la situación normal).

22. Los precedente serían los criterios básicos que permiten elaborar interpretaciones; los cuales, sin duda, serán mas complejas en la medida de que combinaran a ellos entre sí. Avanzar sobre tales combinaciones equivaldría la tarea concreta de formalizar la tipología que nos hemos propuesto postergar.

5. Defensa de este procedimiento de análisis

3. Abrimos este trabajo refiriendo la existencia y uso de modelos y procedimientos de análisis numerosos de presentación sofisticada que predominan en los análisis empíricos contemporáneos, frente a los cuales reivindicábamos este clásico y artesanal instrumento. Quiero abundar aquí en esa defensa, en favor de una manera de encarar el análisis que pone énfasis en la adecuación procesal del instrumento en relación con la naturaleza del dato, en la legitimidad epistemológica incuestionable, en la aptitud para la lectura lógica analítica de la información empírica. Que privilegia tales razones metodológicas frente a la moda.

En ese sentido, deben señalarse algunas observaciones que nos parecen trascendentes en la comparación de los procedimientos de análisis.

a. Si se utilizan los coeficientes que en cada caso correspondan, este procedimiento -en sí mismo- no está sujeto a requerimientos o condiciones cuyo incumplimiento pueda poner en juego la validez de las inferencias lógicas asumidas, como ocurre con la mayor parte de los modelos de análisis habitualmente mas utilizados en la actualidad.

El nivel de medida de las variables y las características de las relaciones, deben atenderse en la opción de coeficientes; pero el modelo en sí no requiere intervalidad, por ejemplo. Tampoco supone linealidad de las relaciones, aunque la excesiva no linealidad puede llegar a ser responsable de ciertos desajuste finales en la ecuación de Lazarsfeld. Estos también pueden resultar de la ausencia de normalidad en la distribución. Pero en ninguno de estos casos el desajuste llega a revestir importancia cuantitativa, por lo que si se acepta una cierta flexibilidad, de todas maneras impuesta por necesario redondeo de decimales, la igualdad se cumple salvo que el modelo teórico sea inadecuado. Que es lo que en definitiva requerimosb. En el mecanismo de este procedimien-

to de análisis se pone énfasis en los resultados por términos (nosotros subrayamos, además, sus magnitudes recíprocas). A diferencia de la mayor parte de los otros, el producto de este análisis es esencialmente analítico antes que sintética. La labor de síntesis queda reservada a la interpretación lógico-teórica. Por lo dicho en "a", este tipo de análisis no debe reservar parte importante de su aplicación a datos contextuales en que las unidades sean colectivos ecológicos. En todos los casos y tipos de variables y relaciones, puede ser aplicado. Y, para ello, es indistinto el nivel de las unidades de análisis. El riesgo de la falacia ecológica aquí se reduce totalmente a la adecuación de uso que haga el investigado

24. Es cierto que no es adecuado ni práctico para el manejo de gran número de variables. Y que, aún en la cantidad razonablemente manipulable de variables en un análisis de este tipo, no es aconsejable proceder al cruce simultáneo de más de variables (lo cual es posible, pero complejiza el análisis al multiplicar marginales y producir marginales de marginales, y reduce el "n2 por celda-base de manera drástica).

Pero a estas limitaciones ciertas debe realizarse las siguientes puntualizaciones:

a) El manejo de un gran número de variables tampoco es demasiado adecuado al tipo de análisis sociológico más habitual. Por lo menos, el requiere un primer desglose que sí puede hacerse con otro procedimiento (en particular sugiero el cluster analysis con el único objeto de reducir el número de variables a manejar, para estos casos; o el propio "ojo" del investigador frente a una matriz de correlaciones...).

b) Quizás fuera más defendible la utilización de modelos que partan de la matriz de correlaciones (análisis factorial, por ejemplo) desde una perspectiva fuertemente inductivista; de la cual no participo.

c) La artesanía que supone la elaboración de un "grases" global de una media docena de variables,

realizado en base a varios lazarsfeld de nivel 3, con los giros que correspondan, lejos de parecerme una limitación engorrosa, me parece que aporta una verdadera riqueza fermental.

Bibliografía

- BLALOCK, H.M. "Causal Inference in Nonexperimental Research", Chapel Hill, 1964.
- BLALOCK, H.M. "Theory Building and The Statistical Concept of Random Measurement" en *American Sociological Review*, 1965.
- BOUDON, Raymond "L'analyse mathématique des faits sociaux", Plon, Paris, 1967
- COLEMAN, James S. "Introduction to mathematical sociology", Glencoe, Ill., The Free press, 1964.
- CORTES, Fernando "La insostenible levedad del dato", " estudios demográficos y urbanos-6, el colegio de México, vol.2, Nº 3, 1987.
- CORTES, Fernando y RUBALCAVA, Rosa María "Métodos estadísticos aplicados a la investigación en Ciencias Sociales.- Análisis de asociación.-", El Colegio de México, 1987
- ERRANDONEA, Alfredo (II.) "¿ Metodología cualitativa versus cuantitativa?", Cuadernos del CLAEI, Nº 35, MONTEVIDEO, 1985
- ERRANDONEA, Alfredo (h.) "Colección de Guías de Clase de Metodología de la Investigación", Nº 9, Buenos Aires, s/f (1982). (Última versión: Montevideo, 1987).
- ERRANDONEA, Alfredo (h) "Notas sobre la comparación de magnitudes de las relaciones parciales en la interpretación mediante el modelo Lazarsfeld de análisis multivariado" en "Memorial sociológico", Asociación de Graduados en Sociología, La Plata, 1983 (Nº 0).
- ERRANDONEA, Alfredo y SUPERVIELLE, Marcos "El lugar de las técnicas cualitativas" en "Revista de Ciencias Sociales", I.C.S., Montevideo, 1986, Nº 1.
- KENDALL, Patricia en HYMAN, Heber. "Diseño y análisis de las encuestas sociales", Amorrortu, Bs. Aires, 1972 (Primera edición en inglés 1955).
- LAZARSFELD, Paul F. "Interpretation of Statistical Relations as a Research Operation" en "The Language of Social Research", (Lazarsfeld y Rosenberg, eds.). Glencoe, Ill The Free Press, 1955.
- LAZARSFELD, Paul F. "The Analysis of Statistical Relations" en "Evidence and Inference" (Daniel Lerner, ed.), The Free Press, Glencoe, Ill, 1959.