

# ANÁLISIS DE RUTA: CONCEPTOS BÁSICOS

MSC Luisa Angelucci B.

## RESUMEN

El creciente interés en el establecimiento de leyes causales en las ciencias sociales y del comportamiento ha incrementado el número de investigadores que se abocan al uso de técnicas como el análisis de rutas o los modelos de ecuaciones estructurales. Sin embargo, no siempre estas técnicas han sido correctamente aplicadas, sobre todo por el desconocimiento o comprensión poco clara de los supuestos sobre los cuales se fundamentan dichas técnicas.

El presente artículo se brinda a manera de introducción sobre los conceptos básicos referidos al análisis de ruta. Específicamente el objetivo de este trabajo es ofrecer una aproximación a la técnica, indicando las nociones implicadas en la misma, conceptos relacionados, procedimiento básico de resolución y el establecimiento de algunas diferencias con el planteamiento de los modelos de ecuaciones estructurales.

Se propone al análisis de ruta como una estrategia que permite la comprensión y estudio de diferentes comportamientos bajo la perspectiva de la complejidad y causalidad de los fenómenos, representando un medio para construir rutas causales a partir de supuestos teóricos y para verificar, a su vez, estos supuestos. De esta manera se resalta como una técnica alternativa destinada a la investigación explicativa en las ciencias del comportamiento, específicamente la Psicología.

**Palabras claves:** análisis de ruta, modelo, causalidad.

## GENERALIDADES

El análisis de ruta fue originalmente desarrollado por Sewall Wright en 1934 como una técnica estadística para el establecimiento de las relaciones entre variables de un sistema biológico (Ortiz-Pulido, 2000). El análisis de ruta (*path analysis*) se ha constituido en una herramienta clave en la formulación de hipótesis y evaluación de relaciones causales en las investigaciones de corte no experimental y específicamente en el campo de la Psicología.

Para autores como Klem (1995) el análisis de ruta puede ser visto como una extensión de la regresión múltiple, diferenciándose del mismo en que no sólo trabaja con una variable dependiente, sino con más de una; además de poderse plantear relaciones del tipo X causa Y y Y causa Z. Por otro lado, también puede ser visto como un caso especial de un análisis estructural de covarianza, al plantear la posibilidad de verificar relaciones directas y mediadas entre variables bajo la perspectiva de un modelo.

Como se planteó anteriormente, el análisis de ruta tiene como objetivo la verificación de las relaciones entre las variables planteadas en un modelo. De esta manera se parte de dos conceptos o aspectos cruciales en la ciencia, y específicamente en las ciencias sociales, humanas o del comportamiento, como la Psicología. Uno de los aspectos es la noción de causalidad y el otro la noción de modelo; aspectos que se tratarán a continuación teniendo presente la posible superficialidad del abordaje de estos dos temas en el presente artículo.

## CAUSALIDAD

El punto de partida del análisis de rutas es una teoría acerca de las relaciones causales entre un grupo de variables (Klem, 1995). Sin embargo, cuando se habla de causalidad se hace referencia a una asunción, más que a una propiedad del resultado obtenido o de la técnica. Es decir, se asume que algunas variables están causalmente relacionadas y se utiliza la técnica para verificar la proposición planteada. En la literatura sobre el tema, no se evidencia una explicación única del término causalidad. No obstante, se puede decir que aparecen dos posiciones centrales: la ontológica y la gnoseológica (Wartosky, 1976).

En la ontológica, se supone que existe una relación unívoca en la causalidad; una relación necesaria y suficiente entre causa y efecto. La causa proviene del propio objeto, de su esencia. En la gnoseológica, la causa es una entidad inventada para explicar las relaciones. La causación concierne a la experiencia y al conocimiento, sin ser un rasgo de las cosas mismas. En esta posición se habla de relación funcional, donde son básicos tres requisitos: relación de contigüidad entre antecedentes y consecuentes,

relación temporal (precedencia temporal de la causa) y conjunción constante entre los elementos (constancia en la relación observada).

En esta línea, según Critto (1982), Jhon Stuart Mill adiciona el requerimiento de la eliminación de otras posibles explicaciones de la relación (relación no espuria). Así, en un esquema de relación causal entre dos variables, puede existir una tercera que provoque una acción sobre ellas. En este sentido se debe considerar la influencia de cualquier variable "extraña" en la relación causal.

Esta noción gnoseológica permite la inclusión de un determinismo condicional, probabilístico, entre las variables, ya sean observadas directamente o estimadas a partir de indicadores, lo que amplía la noción de causalidad, ya que puede considerarse no sólo la lógica del diseño experimental para el análisis causal de las relaciones en ciencias del comportamiento, tal como se plantea en la posición heredada de la psicología y relacionada a su vez con el positivismo tradicional (Bunge, 1981), sino otros métodos, como el análisis multivariado, estudios de panel y tendencia, estudios evaluativos, inventarios de causas y efectos, entre otros (Critto, 1982; Blalock, 1984; Mulaik, 1987).

Al considerar la causalidad como una relación funcional, donde las propiedades de una variable determinan las propiedades de la otra (cuando un elemento de una primera distribución corresponde con otro elemento de una segunda distribución), se da cabida a la noción de contingencia y correlación entre los fenómenos. Así, la metodología estadística y específicamente la asociación entre una determinada posición del caso observado en una de las variables (X) y la posición de otra (Y), permite organizar y entender las relaciones entre las variables, prediciendo lo que ocurrirá en una dimensión al conocer lo que ocurre en la otra (Critto, 1982) y por tanto favoreciendo el descubrimiento de posibles relaciones causales. De esta forma, el análisis de ruta se constituye en una modalidad de investigación que se corresponde con las estrategias basadas en la covariación entre variables, donde la relación causa-efecto se refleja por medio de una función o conjuntos de ecuaciones (de la Garza, 1995).

Cabe resaltar, tal cual lo hacen Robles (1992), Peña (1999) y Keith (2005), entre otros, que el problema de fondo para lograr inferencias causales válidas no es la posibilidad o no de manipular las variables, de usar o no la correlación. Lo esencial es la validez interna del estudio, es decir, la coherencia metodológica entre el objetivo, el planteamiento de la teoría, las hipótesis causales, la pertinencia del diseño de investigación y las operaciones de análisis a la hipótesis a contrastar.

## MODELO

La teoría de modelos en sistemas formales y distintas aproximaciones a la epistemología permiten discutir con mucha extensión acerca del significado de modelo y, cualquier definición de modelo que pretenda darse puede ser simple (Robles, 2000). Sin embargo, se acepta que en todo modelo existe un componente teórico, que se concibe como una representación de descripción de hechos observados o conjunto de datos empíricos, que “si bien son sistemas o analogías que nos describen la realidad en términos simplificados, no dejan por ello de participar del esquema teórico dentro del cual son utilizados, ni de su función explicativa” (Arnau, 1982, p. 32).

El modelo se encuentra así siempre relacionado con la teoría. Para Anguera (1989) el modelo adapta la teoría a la realidad empírica y permite su acercamiento a ella, haciéndola más comprensible y evitando los problemas propios de una rígida formalización. Para Sierra Bravo (1981) los modelos se pueden conceptualizar como construcciones teóricas hipotéticas, susceptibles de matematización, con las que se pretende representar una parte de la realidad, para estudiarla y verificar la teoría.

Los modelos son, entonces, enunciados teóricos, elaborados *a priori*, acerca de la relación entre variables y que pueden ser transformados a modelos simbólicos o matemáticos que permitan verificar su poder predictivo y el ajuste del modelo a los datos empíricos, más específicamente a un modelo que se corresponda con el manejo de los datos e incluya la probabilidad de errores (Bross, 1951; Baltes, Reese y Nesselroade, 1980; Sierra Bravo, 1981).

De esta manera, los modelos causales se evalúan determinando su valor teórico, consistencia matemática y adecuación empírica; esto deriva en la posibilidad de reconfirmación o modificación del modelo. Si el modelo se ajusta a los datos, se sigue sometiendo a prueba y evaluación, si no, se descarta como posible explicación de la estructura causal de las variables en estudio.

El análisis de ruta permite expresar formal y explícitamente una teoría por medio de un modelo que se puede representar en palabras, por ecuaciones matemáticas y generalmente por un diagrama de ruta (Klem, 1995). De hecho, para Mulaik (1993) un modelo de rutas es un sistema de hipótesis, derivado de una teoría, en el cual las hipótesis versan sobre los coeficientes de ruta. Conceptualización de análisis de ruta cónsona con la noción de modelo descrita anteriormente.

En síntesis, el modelo representado en el diagrama de rutas es un todo integrado (un sistema), con cohesión conceptual derivada de la teoría y unidad operacional dada por las rutas que conectan las variables e integran los efectos de unas sobre otras, efectos

que pueden evaluarse a la luz de su ajuste a los datos, previa matematización (Baltes y cols., 1980; Sierra Bravo, 1981; Klem, 1995; Robles, 2000).

## **PRESUPUESTOS**

Un modelo debe cumplir ciertas condiciones para que le pueda ser aplicado el análisis de ruta. Sierra-Bravo (1995) expone los siguientes:

1. El modelo debe representar relaciones de causa-efecto entre las variables.
2. Las relaciones planteadas en el modelo son lineales y aditivas. Las relaciones curvilíneas, multiplicativas o de interacción son excluidas.
3. El modelo debe ser un sistema cerrado o completo. Así, cada variable dependiente debe ser considerada como completamente determinada por alguna combinación de las variables del sistema. Si la determinación no es completa por las variables medidas, se introduce una variable residual que no presenta correlación con otras variables del modelo.
4. Se asume que las variables residuales del modelo, también llamadas errores, no están correlacionadas entre sí y ejercen una influencia aleatoria sólo sobre una variable del modelo (una endógena) y no sobre varias.
5. El modelo debe ser recursivo: es decir, las relaciones entre las variables que forman el modelo son asimétricas. No se debe asumir relaciones bidireccionales entre las variables, dos variables no pueden ser causa y efecto una de otra.
6. El modelo debe ser lineal: se supone que las relaciones que unen las variables que lo forman se representan por ecuaciones lineales.
7. El nivel de medida de las variables debe ser de intervalo o razón. Se pueden emplear variables cualitativas a condición de que sean dicotómicas o se puedan dicotomizar. Así, variables como sexo, tener o no trabajo, entre otras, que indiquen dos niveles de la variable pueden ser incluidas en el modelo. Las variables categóricas policotómicas deben convertirse en dicotómicas. Una forma de hacerlo es codificarlas como variables *dummy*. Según Kerlinger y Lee (2002) una variable *dummy*, también llamada prototipo, es una variable producto de una solución que consiste en crear tantas variables como categorías menos 1 tiene la variable original. Por ejemplo, si se tiene la variable estado civil con tres categorías: casado, divorciado y soltero, la solución por codificación *dummy* sería:

Tabla 1. Ejemplo de codificación *dummy*

	$X_1$	$X_2$
Casado	1	0
Divorciado	0	1
Soltero	0	0

De esta forma se tiene un vector 1 y 0: un 1 indica que un sujeto es miembro del grupo (en este caso: casado), y un 0 indica que el sujeto no es miembro del grupo, así  $X_1$  incluye a los casados y  $X_2$  a los divorciados, los solteros, por consecuencia, no pertenecen a los dos grupos anteriores. Con este tipo de codificación la variable puede utilizarse en modelos lineales, como el análisis de rutas.

8. Las variables utilizadas en el modelo son medidas sin error (el error de medida es mínimo / las medidas de las variables presentan una alta confiabilidad).

Otros autores plantean otros requerimientos que se adicionan a los anteriores; éstos son:

9. No deben existir errores de especificación. Los errores de especificación ocurren cuando una variable causal no es considerada en el modelo, o se incluyen variables que no presentan una relación causal relevante. La especificación del modelo tiene que ver con el establecimiento de las relaciones de correspondencia entre los conocimientos teóricos y el planteamiento de las ecuaciones matemáticas relativas a los efectos causales entre las variables (Sierra-Bravo, 1981). Para Casas (2002), además, tiene que ver con la especificación de los supuestos estadísticos sobre las fuentes de variación y el comportamiento de las variables no incluidas en el modelo. La claridad del modelo viene determinada por el grado de conocimiento del investigador; si éste es poco o impreciso, la asignación de parámetros, consecuentemente, es más susceptible de error.
10. No debe existir subidentificación o indeterminación en el modelo. La identificación del modelo se relaciona con el número de ecuaciones y el número de incógnitas existentes en un modelo. Cuando el número de ecuaciones es insuficiente para proporcionar soluciones al número de parámetros incógnitos, se dice que el modelo es infra o sub-identificado. Cuando los modelos contienen el número justo de ecuaciones, son llamados identificados, y aquellos para los que hay más ecuaciones que incógnitas se denominan sobreidentificados. Para Casas (2002) una de las reglas generales para identificar un modelo “es la de los grados de libertad, obtenidos como la diferencia entre el número de varianzas y covarianzas (ecuaciones) y el número de parámetros a estimar. Es una condición necesaria pero no suficiente. Cuando  $g < 0$ , serán *modelos*

*infraidentificados*; cuando  $g=0$ , los modelos son *posiblemente identificados*, y cuando  $g>0$  el modelo está *sobreidentificado*" (p. 5). Los modelos recursivos, tal como los propuestos como modelos de ruta, *nunca* son subidentificados y generalmente son identificados (Baltés y cols., 1980).

11. Uso apropiado del coeficiente de correlación. Cuando se usa una matriz de correlaciones como resultado, se debe hacer uso adecuado del coeficiente según el nivel de medida de las variables, por ejemplo, el uso de la correlación de Pearson para medir la asociación entre dos variables de intervalo.
12. Tamaño adecuado de la muestra para evaluar la significancia. Kline (1998) recomienda 20 casos por cada uno de los parámetros para evaluar los efectos del modelo (esto incluye todas las posibles dimensiones o subescalas de las variables medidas). Para Klem (1995) el número de casos necesarios depende de la complejidad del modelo, pero mínimo se requiere entre 200 ó 300 casos.
13. Uso de la misma muestra. Todas las relaciones planteadas en el modelo deben verificarse con la misma muestra, por lo que debe tomarse en cuenta la posibilidad de sobremuestrear y de manejar los datos perdidos (Hair y cols., 2000).
14. Ortiz-Pulido (2002) señalan que es necesario estandarizar los datos obtenidos para evitar problemas de magnitud entre las variables consideradas, proponiendo la tipificación por Z.

Igualmente, al utilizar la técnica de regresión para solucionar el análisis de rutas se presentan las mismas limitantes que se aplican a los análisis de regresión (supuestos del análisis). Algunos no contemplados en líneas anteriores son (Hair y cols., 2000; Kerlinger y Lee, 2002; Keith, 2005):

15. Normalidad de los datos de las variables individuales.
16. Debe existir nula o baja multicolinealidad, es decir; nula o baja correlación entre las variables predictoras
17. Los errores están normalmente distribuidos y son independientes entre sí.
18. La distribución de los valores de los errores alrededor de la línea de regresión debería ser constante para todos los valores de X. Este supuesto se refiere a la homocedasticidad.

En caso de no cumplir los supuestos relacionados con el análisis de regresión, se pueden llevar a cabo técnicas para intentar lograrlo (Hair y cols., 2000).

## COMPONENTES DEL MODELO DE RUTA

Un modelo de ruta está compuesto por las *rutas* que conectan en forma directa e indirecta *variables* implicadas en una teoría de forma simbólica y matemática por los *coeficientes de ruta*.

## RUTAS

Una ruta es la imagen de una relación asimétrica entre dos variables. Las rutas se representan en los diagramas usando flechas de una punta, para señalar su recorrido. Las rutas generalmente se refieren a relaciones entre dos variables, que pueden ser directas o indirectas (Wright, 1934; Wolffe, 1980). Una ruta directa representa la relación entre dos variables en conexión inmediata  $X \rightarrow Y$ , mientras que una ruta indirecta representa una relación entre dos variables, pasando a través de otras como  $X \rightarrow Z \rightarrow Y$ . Las rutas permiten la representación bidimensional de relaciones multivariadas, cuya presentación multidimensional puede ser bastante compleja (Robles, 2000).

Klem (1995) define dos tipos de modelos de rutas en función de las rutas definidas entre las variables, el modelo totalmente recursivo y el modelo no totalmente recursivo. En el modelo totalmente recursivo, cada variable tiene un efecto directo sobre todas las otras variables de la cadena causal. Un modelo no totalmente recursivo de una o más de las rutas o enlaces causales entre las variables no es propuesto (ver más adelante figuras 1 y 2, respectivamente).

## VARIABLE

Una variable es un símbolo al que se le asignan valores o números (Kerlinger y Lee, 2002) y es identificada como la operacionalización de un constructo (Robles, 2000). Como se indicó, las variables en el análisis de ruta son atributos derivados de la teoría en que se basa la investigación. Al representarlas por medio de las ecuaciones y los diagramas no hay que perder de vista sus propiedades teóricas.

En un análisis de ruta, la principal distinción que debe hacerse es entre las *variables exógenas*, que son aquellas que no reciben influencia de ninguna otra variable en el modelo, más allá del error de medida; y las *variables endógenas*, las cuales reciben influencia de otras variables, es decir, son explicadas por una o más variables del modelo (Klem, 1995; Sierra-Bravo, 1995). La clasificación metodológica entre variables independientes y dependientes, o entre variables predictoras y predichas no es estrictamente



aplicable en este contexto, dado que algunas variables reciben flechas, y a su vez, las dan en el diagrama (son influenciadas e influyen) (Klem, 1995). Las variables dependientes sólo serían aquellas que reciben flechas, las variables independientes aquellas de donde parte la flecha, y aquellas que reciben y dan flecha serían mediadoras (otros autores las llaman intervinientes) (Klem, 1995; Casas, 2002).

Igualmente se distinguen las variables residuales, que representan a los factores no observados que pueden influir en cada una de las variables explicadas en el modelo o a los errores de medida. Vienen a constituir la varianza de los datos no explicada por el modelo. Según Klem (1995) la magnitud del efecto de error para una variable endógena es igual a  $1 - R^2$  (varianza de error) donde  $R^2$  es el coeficiente de determinación múltiple entre la variable endógena y las variables que la influyen. Asimismo, se usa la fórmula del coeficiente de alienación:  $\sqrt{1 - R^2}$ .

Igualmente, en el análisis de ruta se trabaja con las variables manifiestas u observadas, a diferencia de otras técnicas de análisis causal como el modelo estructural que trabaja con la noción de variable latente. Para Bentler (1980) una variable latente es un constructo hipotético que el investigador no mide y que de hecho no puede ser medida directamente, dado que no existen métodos operacionales para medirlos. Según Kerlinger y Lee (2002) "una variable latente es una entidad no observada, que se presume subyace a las variables observadas" (p. 49); entendiéndose, por tanto, que en el análisis de ruta se trabajaría con la operacionalización de constructos y de variables latentes (no explícitas en el modelo), es decir, con variables o indicadores observables.

## **COEFICIENTES DE RUTA**

Cada relación de influencia de una variable sobre otra (una ruta) puede evaluarse a partir de un coeficiente que expresa la magnitud y dirección del efecto. Se han utilizado, en este sentido, los coeficientes Wright o los coeficientes Beta (de regresión estandarizados). Como se indicó anteriormente, el autor a quien puede considerarse creador de la técnica de los coeficientes de ruta es el biólogo Sewall Wright; específicamente, uno de los aportes *más* importantes de Wright fue el desarrollo de una técnica para calcular los coeficientes de ruta a partir de los coeficientes de correlación.

Según Sierra-Bravo (1995) estos coeficientes Wright constituyen las incógnitas cuyo valor se halla mediante la resolución del sistema de ecuaciones del modelo. La notación de estos coeficientes suele ser  $p_{ij}$ , donde  $i$  es la variable predicha y  $j$  la variable predictora (la variable efecto va primero). Es un coeficiente que muestra el efecto directo de una variable independiente sobre una dependiente en el diagrama de ruta. Actualmente, por lo general, no se utilizan los coeficientes de Wright sino los coeficientes

Beta resultantes del análisis de regresión. Según Sierra-Bravo (1995) ambos coeficientes coinciden, por lo que se pueden utilizar indistintamente.

Es útil indicar que para una regresión bivariada, el coeficiente beta –de rutas– es el mismo que el coeficiente de correlación simple o de orden cero, es el caso de un modelo de rutas con una variable endógena influenciada por una sola variable (y el término de error). Para el caso de la regresión múltiple (modelo de rutas donde la variable endógena es influenciada por dos o más variables y el término de error) el coeficiente beta –de rutas– es un coeficiente de regresión parcial, y representan la relativa influencia de la variable predictora sobre la predicha sin contaminación de las varianzas producidas por las otras variables predictoras, es decir, controlando estadísticamente el efecto de las otras variables del modelo. En otras palabras, según Cohen, Cohen, Wert y Aiken (2003) “el término coeficiente de regresión parcial se usa para aclarar que es el peso aplicable a una VI cuando no especifica más de una VI en el modelo” (pp. 66-67).

## **REPRESENTACIÓN DEL MODELO DE RUTA**

El modelo causal propuesto bajo la perspectiva del análisis de ruta se puede representar por medio de un diagrama, por medio de ecuaciones o por enunciados verbales. El uso más extendido es la representación gráfica vía diagrama de rutas.

### **DIAGRAMA DE RUTA**

Es una representación gráfica de las relaciones de causalidad que se presumen entre las variables, que se construye a manera de diagrama de flujo. Tal como se ve en la figura 1, las variables van dentro de cajas (rectángulos) (a diferencia de las variables latentes en los modelos estructurales que van en óvalos). Las relaciones entre las variables se expresan mediante flechas rectas unidireccionales que empiezan en la variable predictora (que influye) y cuya punta termina en la variable predicha o influida (Sierra-Bravo, 1995). Es preferible que las cabezas de las flechas terminen en el lado izquierdo de la caja de la variable influida, no abajo o arriba de la misma. El diagrama debe ser ordenado temporalmente, así existen variables “tempranas” (más hacia la izquierda del diagrama), como por ejemplo variable 2 y 3 en la figura 1 que influyen sobre otras “tardías”, como por ejemplo la variable 1 en la misma figura (Ortiz-Pulido, 2000). Más hacia la izquierda del diagrama se encuentran las variables exógenas.

Las relaciones entre variables exógenas (no reciben influencia de otra variable) se representan por líneas curvas, que indican correlación entre las variables, sin causación. Por esta razón se dibujan bidireccionales, con cabeza de flecha a ambos lados (ver figura 2). Algunos autores obvian la representación gráfica de esta correlación (Sierra-Bravo, 1995; Ortiz-Pulido, 2000).

En el diagrama propuesto se coloca sobre cada línea recta (rutas) el símbolo del coeficiente beta ( $\beta$ ) con los números índice de las variables, comenzando con la variable predicha, y luego el signo de la correlación esperada, tal como lo estableció Wright (Sierra-Bravo, 1995; Casa, 2002); por ejemplo, en la figura 1, el coeficiente de ruta entre la variable 1 y la 2 se expresa como  $\beta_{12+}$ . Cabe destacar que el uso de esta nomenclatura no es universal. Algunos autores representan las relaciones colocando sólo el signo esperado en la relación, otros no colocan ni betas ni signos, sino que enuncian en forma verbal cada una de las relaciones esperadas en cada ruta.

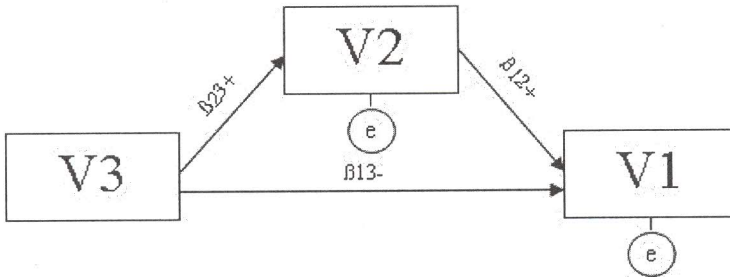


Figura 1. Ejemplo de diagramas de rutas propuesto

Las variables residuales (desconocidas, no medibles, pero que afectan al sistema) se representan como "U" o "e", según el estilo, y dirigen una flecha hacia la variable que es influida por otra (endógena). Algunos autores encierran en un círculo este tipo de variable, como se puede ver en las figuras 1 y 2.

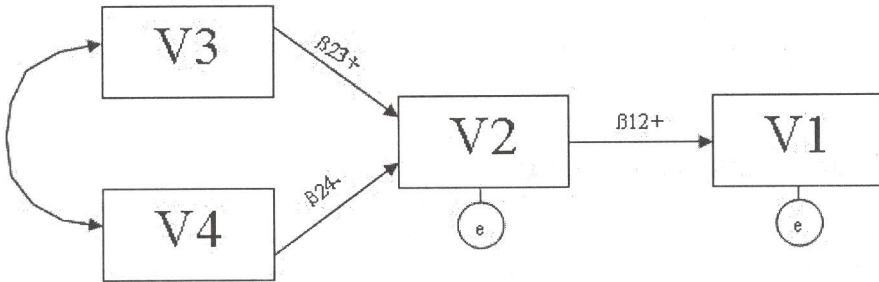


Figura 2. Diagrama de rutas propuesto con dos variables exógenas

Luego de resolver el análisis de rutas, el valor del coeficiente de ruta es colocado arriba de la línea que conecta la variable predictorora y la predicha; y el error dentro del círculo (ver figura 3). Algunos autores colocan, además, el coeficiente de determinación del modelo ( $R^2$ ) arriba o adentro de la caja de cada variable endógena.

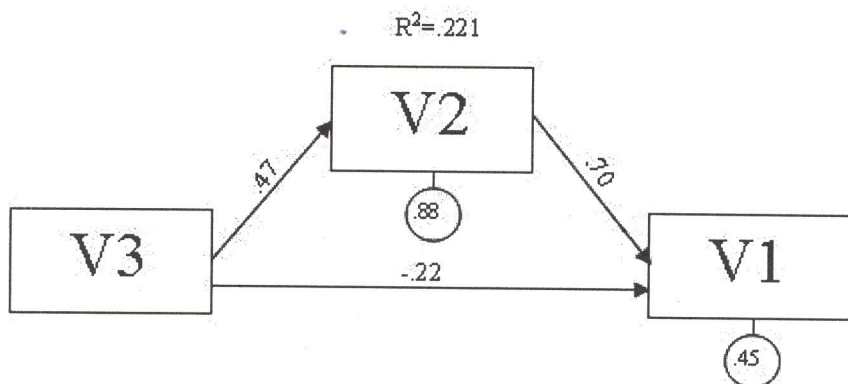


Figura 3. Ejemplo de diagrama de rutas resultante

Cuando la variable es una *dummy*, se suele indicar sus códigos en la caja, a fin de que se puedan interpretar los resultados (la dirección de los coeficientes beta) (ver figura 4).

Cabe destacar que existen varios estilos para presentar un diagrama de rutas resultante: presentando sólo las rutas cuyos coeficientes resultaron significativos (según el nivel de significancia establecido *a priori*); presentando todas las rutas independientemente de su significancia estadística; presentando sólo las rutas cuyos coeficientes superan un valor previamente establecido, por ejemplo 0.30 y, presentando todas las rutas pero diferenciando las líneas de aquellas rutas cuyos coeficientes son significativos de las líneas de las rutas no significativas (las significativas se colocan más gruesas, grandes o llenas) (Ortiz-Pulido, 2000).

## ECUACIONES MATEMÁTICAS

Igualmente, el modelo de rutas se puede plantear por medio de ecuaciones matemáticas (Wolfe, 1980). Se establecen tantas ecuaciones lineales como variables endógenas tenga el modelo. Por ejemplo, para la figura 1 se plantean dos ecuaciones matemáticas, donde **a** es el intercepto, **b** son los coeficientes de regresión (en Sierra-Bravo (1981) los coeficientes p) y **e** el error.

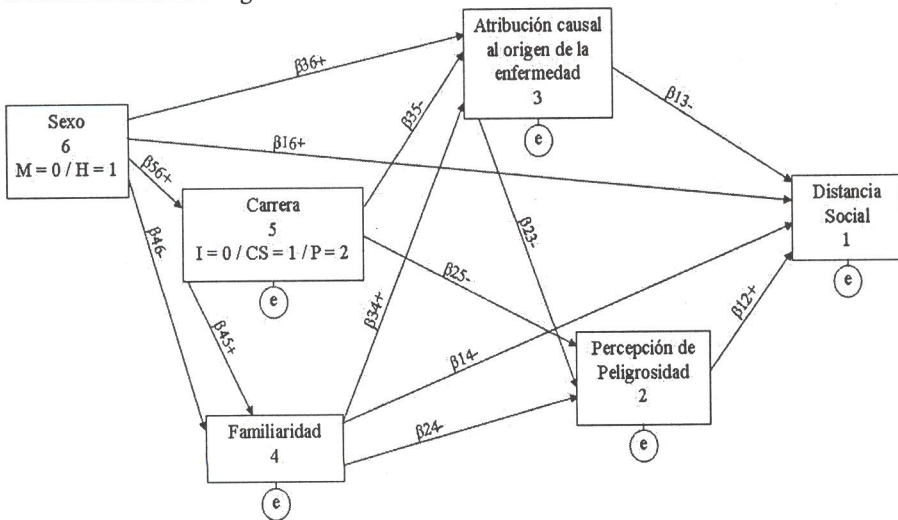
$$(1) \quad \begin{aligned} X_1 &= a + b_2 X_2 + b_3 X_3 + e \\ X_2 &= a + b_3 X_3 + e \end{aligned}$$

## ENUNCIADOS VERBALES

Algunos autores exponen verbalmente (escriben) cada una de las hipótesis o relaciones causales planteadas entre cada variable endógena y su predictora. Generalmente esta forma de presentación no se hace aislada, sino que acompaña al diagrama de rutas.

De seguido se muestra la representación de un modelo de rutas por medio de su diagrama de ruta. Se toma como ejemplo el estudio de Angelucci, Lazo, Loaiza, Montiel, Mora, Páez, Perales, Rojas y Rojas (2006), quienes con una muestra de 398 estudiantes universitarios realizaron una investigación para estudiar la influencia de factores psicosociales, sexo y carrera sobre la distancia social hacia las personas con esquizofrenia, y relacionar estas variables entre sí.

Las hipótesis fueron planteadas en un diagrama de rutas tal como se presenta a continuación en la figura 4.

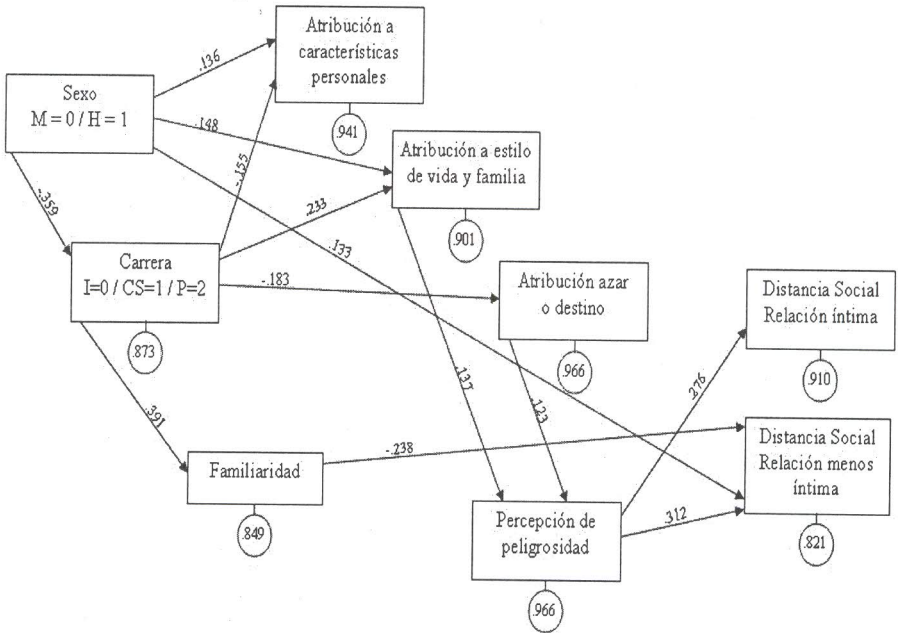


**Figura 4.** Diagrama propuesto para la distancia social hacia la esquizofrenia. Angelucci y cols. (2006).

Se puede observar que en este diagrama se enumeran las variables del modelo de derecha a izquierda. Se coloca el símbolo de Beta y su signo en cada ruta, expresando la relación esperada y los errores en círculo y debajo de cada caja de las variables endógenas. Se destaca que hay dos variables categóricas, el sexo y la carrera.

Como ya se dijo, es recomendable colocar dentro de las cajas de este tipo de variables (a veces se coloca fuera del diagrama) el código asignado a cada modalidad de la variable. La variable carrera no se convirtió en *dummy*, dado que se asumió como ordinal en cuanto a la cantidad de materias relacionadas con esquizofrenia en cada *pensum* (Ingeniería [I] menor contenido, Psicología [P] mayor contenido).

Luego de realizar los análisis estadísticos pertinentes, se obtuvo el diagrama de rutas definitivo (ver figura 5).



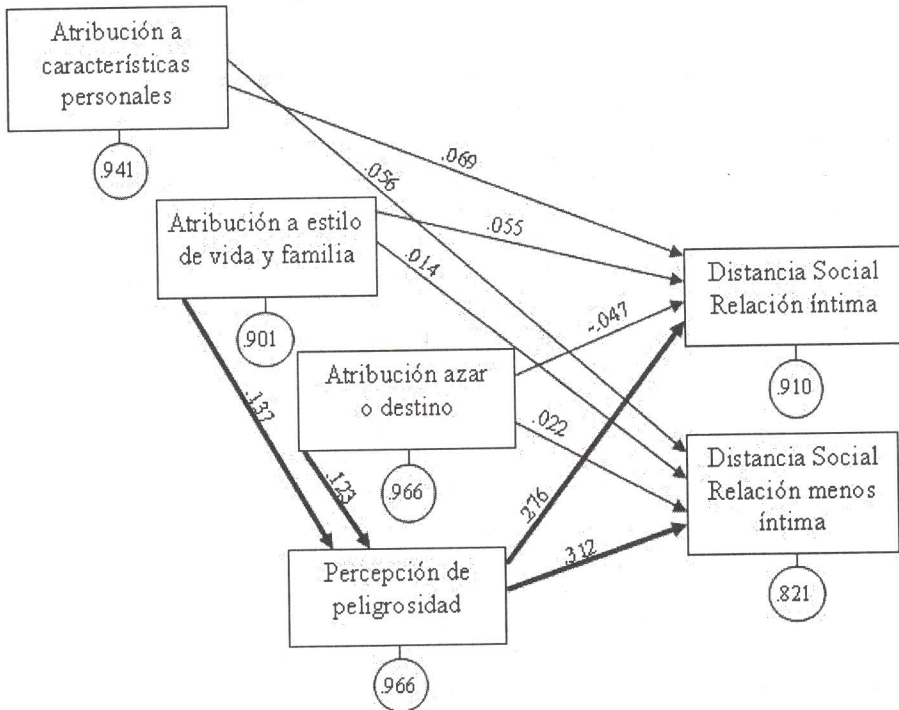
**Figura 5.** Diagrama resultante en la explicación de la distancia social hacia la esquizofrenia. Angelucci y cols. (2006).

Como se puede observar, este diagrama resulta diferente al propuesto. Una diferencia que viene dada por la cantidad de variables representadas (cajas) en ambos diagramas. Por ejemplo, en el diagrama propuesto, la variable atribución es representada en una sola caja, sin embargo, se realizó posteriormente un análisis de componentes principales donde resultaron tres dimensiones, las cuales se representan diferenciadas en el diagrama de rutas resultante; lo mismo ocurrió con distancia social.

En cuanto a esto, no existen reglas universales, sin embargo, se recomienda representar todas las dimensiones de las variables en el diagrama propuesto, lo que implica, por supuesto, conocimiento del comportamiento de la escala de medida de cada variable.

Por otra parte, en el diagrama resultante sólo se representan las rutas con coeficientes de rutas (regresión) significativos estadísticamente. Por ejemplo, la relación entre familiaridad y percepción de peligrosidad no es dibujada por no ser significativa.

En este tipo de casos hay varias opciones para expresar los diagramas resultantes; algunos autores recomiendan no colocar las flechas cuando la relación no es significativa, como en el caso del ejemplo. Otros, colocar líneas discontinuas para las relaciones negativas (----), y otros colocan todas las rutas, pero dibujan las significativas con algún otro distintivo como color, grosor, entre otros (–, =), como se ve en la figura 6.



**Figura 6.** Parte del diagrama de ruta resultante en la explicación de distancia social hacia la esquizofrenia. Angelucci y cols. (2006). Un estilo utilizado para representar rutas no significativas.

## TÉCNICA DE ANÁLISIS DE RUTA

### ANÁLISIS DE REGRESIÓN

El análisis de rutas es una técnica similar a la regresión pero con poder explicativo, que estudia los efectos directos e indirectos en el conjunto de variables observables (Casas, 2002).

El análisis de regresión implica la estimación de coeficientes de regresión parcial (coeficiente Beta), los cuales pueden ser empleados, como se indicó anteriormente, como coeficientes de rutas en el análisis de rutas.

Para el cálculo de estos coeficientes Beta se utiliza la técnica de regresión múltiple. Se basa en lo que se conoce como modelo lineal general (Hair y cols., 2000; Kerlinger y Lee, 2002; Keith, 2005), el cual puede expresarse como la ecuación de la línea recta, más un término de error aleatorio:

$$(2) \quad Y = a + bX + e$$

En términos de análisis de regresión,  $Y$  es la variable predicha (endógena),  $X$  es la variable predictora,  $b$  es el coeficiente o pendiente de la recta,  $a$  es el intercepto y  $e$  es el término de error aleatorio, donde  $a$  y  $b$  son constantes producto del análisis de regresión. Ambas constantes pueden ser positivas o negativas. El signo del coeficiente  $b$  indica la dirección de la relación entre las variables.

El tratamiento más simple y aplicable a un rango más amplio de situaciones para resolver el análisis de ruta ha sido la estimación de la ecuación de regresión vía mínimos cuadrados (Wolfe, 1980). Por ello, el procedimiento que se describe se remite a este tipo de estimación.

El procedimiento es muy sencillo y consta de cinco pasos básicos:

1. Identificar las variables endógenas en el modelo; esto dará el total de ecuaciones de regresión a resolver. En el ejemplo anterior de distancia social, se tiene que las variables endógenas son distancia social (en sus dos dimensiones), percepción de peligrosidad, atribución causal al origen de la enfermedad (en sus tres dimensiones), familiaridad y carrera. De esta manera habría un total de ocho ecuaciones de regresión ( $\eta=8$ ).
2. Para cada una de las variables endógenas, conformar el bloque de variables con rutas directas hasta ella. A este bloque de variables que tienen efecto directo sobre  $Y\eta$ , se denominará  $X\eta$ , y constituye las variables predictoras del modelo de regresión. En el ejemplo anterior las variables predictoras de la variable peligrosidad, por ejemplo, son: familiaridad, carrera y atribución causal al origen de la enfermedad (en sus tres dimensiones).
3. Realizar el análisis de regresión para cada bloque  $X\eta \rightarrow Y\eta$ . Con esto ya se tienen los estimados de los coeficientes de ruta (coeficientes de regresión Beta) para todos los efectos directos, más los  $R^2$  para cada variable endógena y su nivel de significancia. Si es una variable predictora, se realiza una regresión simple, si existen 2 ó



más variables predictoras se usa el análisis de regresión múltiple simultáneo, donde todas las variables predictoras “entran” en la ecuación de regresión al mismo tiempo (Keith, 2005).

4. Graficar el diagrama resultante.
5. De existir rutas indirectas en el modelo, calcular los efectos indirectos para cada ruta. También se pueden estimar los efectos totales (método descrito en el siguiente apartado).

### CÁLCULO DE EFECTOS

Existen tres tipos generales de efectos en el análisis de rutas: efecto directo, efecto indirecto y efecto total, que vienen dados por la posibilidad de descomponer el efecto total del modelo.

El *efecto directo* viene dado por el coeficiente de regresión entre la variable predictora y la predicha, y responde a la pregunta de si existe relación entre las dos variables.

El *efecto indirecto* es el producto de los coeficientes en rutas indirectas entre una variable y otra variable. En un modelo  $X \rightarrow Z \rightarrow Y$ , el efecto indirecto de  $X$  sobre  $Y$  viene dado por  $bxz * bzy$ .

El *efecto total* es la suma de los efectos directos e indirectos. Si bien se pueden hacer distinciones más finas entre los tipos de efectos, ésta es la descomposición básica planteada por diferentes autores (Wolffe, 1980; Robles, 2000).

### SIGNIFICANCIA Y BONDAD DE AJUSTE DE MODELOS DE RUTAS

Este aspecto se corresponde con la etapa de evaluación del modelo, y se refiere a la exactitud de los supuestos del modelo de rutas especificado para determinar si el modelo es correcto, en el sentido de si se ajusta o aproxima al fenómeno (a los datos).

Generalmente se requiere la consideración de varios indicadores del ajuste del modelo. Para evaluar los coeficientes de rutas individuales se usan los índices de significancia estadística “t” o “F” resultantes de la ecuación de regresión. Igualmente, se evalúa la significancia de los coeficientes de determinación ( $R^2$ ) para cada variable endógena (Robles, 2000).

Una forma alternativa para evaluar el modelo con sus coeficientes de ruta es utilizar índices de ajuste del modelo a partir de la estrategia de modelos estructurales de covarianza. Si el modelo está correctamente especificado (incluye toda las variables

relevantes y excluye todas las variables irrelevantes), con flechas (rutas) correctamente indicadas, la suma de los valores de rutas desde “i” a “j” van a ser iguales a los coeficientes de regresión de “j” predicha a partir de “i”. Se puede comparar la matriz de correlación de los coeficientes de rutas estimados con la matriz de correlaciones observadas, con el fin de evaluar la bondad de ajuste del modelo de ruta.

En la práctica se puede evaluar la bondad de ajuste del modelo por medio de programas estadísticos, como LISREL O AMOS; éstos calculan una variedad de coeficientes de ajuste del modelo, bajo la noción de modelos de ecuaciones estructurales (Keith, 2005). Comúnmente se usa el índice de significancia estadística  $X^2$  (chi cuadrado), con una distribución con los mismos grados de libertad del modelo. Se plantea como hipótesis nula que el modelo que se ajusta “es bueno”, por tanto mientras mayor sea el valor del  $X^2$  mayor la posibilidad de que sea significativo y por tanto el modelo no se ajuste (Casas, 2002).

### **ANÁLISIS DE RUTA Y MODELOS ESTRUCTURALES**

Como se dijo al principio de este documento, el análisis de ruta puede verse como un caso especial de análisis estructural de covarianza, relacionándose con el modelo de ecuaciones estructurales (SEM, por sus siglas en inglés). Esta estrategia se basa en el análisis estructural de covarianza, para la estimación de parámetros de modelos que permiten evaluar el ajuste de un modelo esperado a los datos observados. Según Robles (2000) “el análisis de rutas puede considerarse como una forma general de modelos de ecuaciones estructurales, con menor cantidad de suposiciones acerca de la teoría de la medición y menores requisitos computacionales” (p. 9).

Para Robles (2000) algunos puntos separan a las dos clases de modelos: 1) el tipo de estimación de parámetros y 2) la estimación del ajuste del modelo.

En cuanto al tipo de estimación, se tiene que las técnicas de análisis de rutas tradicionalmente utilizadas, se basan en estimaciones sucesivas de los parámetros de cada ecuación; por su parte, el modelo de ecuaciones estructurales se basa en técnicas de estimación simultánea con cada paso del algoritmo computacional. Para este autor, otra distinción relacionada con el tipo de estimación es la que se puede hacer en relación a la función de optimizar en el proceso de estimación. Como ya se señaló, una de las técnicas utilizadas para resolver el análisis de ruta es la regresión múltiple. En este tipo de análisis la función a optimizar más fácilmente manejable es la de mínimos cuadrados (Kerlinger y Lee, 2002).

Existen otras formas de funciones a optimizar, o maneras de realizar la optimización de las mismas funciones, por lo que podrá encontrarse en la literatura análisis de regresión por mínimos cuadrados en dos y tres etapas, mínimos cuadrados genera-

lizados, mínimos cuadrados ponderados, y máxima verosimilitud, entre otras (Lance, Cornwell y Mulaik, 1988).

Según Robles (2000), la propiedad fundamental del análisis de regresión basado en mínimos cuadrados es su simplicidad. De allí su difusión y uso como la base para técnicas elementales de análisis de rutas. En el modelo de ecuaciones estructurales, suelen usarse funciones que cumplen con propiedades cuantitativas más exigentes, como mínimos cuadrados generalizados, mínimos cuadrados ponderados, y máxima verosimilitud. Entre las implicaciones más importantes del uso de estas funciones, está el hecho de que el modelo de ecuaciones estructurales debe usarse para modelos estrictamente confirmatorios y sobreidentificados (Bentler, 1980; Kerlinger y Lee, 2002; Keith, 2005).

Esto crea grandes diferencias en la eficiencia y el significado de la estimación de parámetros, por lo que desde el punto de vista computacional, el modelo de ecuaciones estructurales es más eficiente y elegante (Robles, 2000).

Por otra parte, el análisis de rutas, en su versión más simple, implica el uso de mínimos cuadrados, con estimación sucesiva de parámetros. Esto puede ser objeto de muchos cuestionamientos desde el punto de vista computacional y estadístico. Sin embargo, desde el punto de vista práctico y didáctico el análisis de rutas por técnicas convencionales es más simple y fácil de aprender e interpretar (Bentler, 1980); además, se apunta como una primera aproximación a la construcción de modelos, resulta una opción viable y con muchos antecedentes en la literatura (Wolfe, 1980).

La otra distinción entre estas dos técnicas, según Robles (2000), viene dada por la forma de evaluar el ajuste del modelo. Como se dijo anteriormente en el análisis de rutas, basado en mínimos cuadrados, lo más común es evaluar el ajuste para cada variable endógena, usando la significancia de cada uno de los coeficientes de regresión estandarizados y la del coeficiente de determinación múltiple ( $R^2$ ). En los modelos de ecuaciones estructurales, para cada modelo existe un índice de ajuste global. Esto se relaciona con la diferencia entre estimación simultánea y sucesiva. La estimación simultánea en modelos de ecuaciones estructurales permite la especificación de un índice de ajuste global (Lance y cols., 1988; Bentler, 1980).

Por otro lado, Bentler (1980) establece que los parámetros estimados por medio de ecuaciones estructurales presentan un mayor significado teórico que los coeficientes resultantes de la ecuación de regresión. Las ecuaciones estructurales, además, no representan explícitamente los parámetros del proceso causal. Implícita en cada ecuación están los parámetros asociados con las varianzas y covarianzas de las variables predictoras.

De esta manera, con el mismo número de variables involucradas en el modelo existen más parámetros asociados con el proceso causal representados en la ecuación estructural. Por ejemplo, si se tienen tres variables predictoras ( $X_1$ ,  $X_2$ ,  $X_3$ ) y una a predecir ( $Y$ ) en una ecuación de regresión (análisis de rutas) habría tres parámetros representados por cada uno de los coeficientes de regresión ( $b_1$ ,  $b_2$  y  $b_3$ ) más un parámetro representado por el término de error ( $e$ ), en total cuatro. Mientras que en el modelo de ecuaciones estructurales habrían 10 parámetros (tres coeficientes:  $b_1$ ,  $b_2$  y  $b_3$ , el error:  $e$ , tres términos de varianza:  $\sigma_1^2$ ,  $\sigma_2^2$ ,  $\sigma_3^2$ , y tres términos de covarianzas:  $\sigma_{12}^2$ ,  $\sigma_{13}^2$ ,  $\sigma_{23}^2$ ), donde sólo tres de ellos serían mostrados en la ecuación estructural.

Otra distinción es que la resolución del modelo de ecuaciones estructurales supone previamente un modelo de medida que requiere un análisis factorial confirmatorio y, por tanto, mayor conocimiento de las medidas de cada variable (Bentler, 1980; Kerlinger y Lee, 2002; Keith, 2005).

Por último, ya no relacionada con la estimación, viene dada por el tipo de variable utilizada en cada técnica. Como ya se indicó anteriormente, el modelo de ecuaciones estructurales se propone explicar las propiedades estadísticas de variables medidas en términos de variables latentes, mientras que los modelos de rutas trabajan con variables manifiestas u observables (Kerlinger y Lee, 2002, Keith, 2005).

Este último aspecto lleva a la consideración del uso o no de la noción de variables latentes en el planteamiento de modelos, lo cual, va unido a la posición teórica asumida para la explicación de los hechos; la cual, en las ciencias del comportamiento, tiende a ser diversa. Al respecto, Gintis (2006) establece “todas las disciplinas del comportamiento incluyen un modelo de la conducta humana. Además, estos modelos no sólo son diferentes, como debe esperarse ya que sus aproximaciones explicativas son distintas, sino que son discordias incompatibles” (pp. 1-2).

En Psicología, estas discrepancias son tales que le dan a la disciplina una imagen interna más bien caleidoscópica (Peña, 2006); y, específicamente, una de las repercusiones más sonada de estas controversias está en las divergencias relativas al nivel del discurso de los datos, al punto que los enfoques conceptuales se podrían ordenar de acuerdo a la distancia que asumen entre la realidad y los datos, lo cual repercute a su vez sobre el uso o no de variables intermedias en las explicaciones teóricas (Peña, 2006, *comunicación personal*).

Más específicamente en Psicología, los enfoques conceptuales van desde los que se aferran a elementos directamente observables, la conducta explícita, a los que emplean dimensiones que sólo pueden inferirse, variables latentes. Metodológicamente, para los enfoques en el primer polo el manejo multivariable con intensidad explicativa debería

recurrir al Análisis de Ruta; y para los que se ubican en el segundo polo la herramienta sería Modelos Estructurales (Peña, 2006, *comunicación personal*).

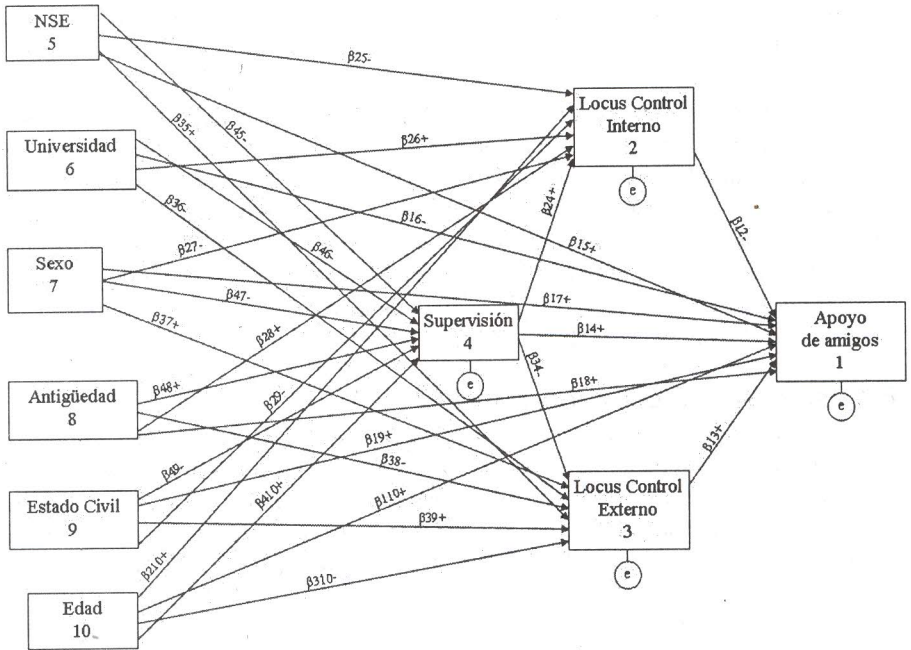
Es importante destacar que a estas diferencias subyacen, más en el fondo, modos distintos de asumir la epistemología; por tanto, la elección entre una forma u otra de resolver el cálculo de las covariaciones, contingencia o correlación, análisis de ruta o modelos estructurales, va más allá de las modas o de las competencias informáticas, adentrándose en las raíces filosóficas de cada enfoque; aun más, en Psicología el desarrollo de los modelos de cálculo siempre ha dependido de las pretensiones de los enfoques conceptuales, una regla que sigue aplicando al momento de elegir una metodología y un modo de análisis de datos (Peña, 1995).

### **EJEMPLO DEL USO DEL PROCEDIMIENTO**

A continuación se va a presentar un fragmento del diagrama de ruta propuesto por Angelucci (2001) para evaluar cómo un conjunto de variables sociodemográficas, laborales, el locus de control de salud, el apoyo social percibido y la salud psicológica, influyen la salud física autopercebida, y evaluar cómo se relacionan entre ellas dichas variables.

Cabe destacar que sólo se muestra el planteamiento del modelo, los resultados a partir del análisis de regresión, el modelo resultante y el cálculo de los efectos directos, indirectos y totales para la parte de modelo original que corresponde a la predicción del apoyo social percibido por parte de amigos.

Como se puede ver en la figura 7, se plantea un modelo que pretende explicar a la variable apoyo percibido de amigos a partir de las relaciones directas e indirectas de esta variable con el locus de control de salud, dividido en dos subescalas (interno y externo), el cumplir con un cargo de supervisor y un conjunto de variables sociodemográficas (nivel socioeconómico, sexo, estado civil y edad) y laborales (antigüedad en el trabajo y tipo de universidad donde laboral). Se representa el error para cada variable que recibe influencia de cualquier otra variable del modelo.



**Figura 7.** Diagrama de ruta propuesto para explicar apoyo social percibido de amigos. Fragmento de Angelucci (2001).

Luego de haber representado, y siguiendo los pasos del procedimiento basado en la técnica de análisis de regresión, se tiene que:

1. Las variables endógenas del modelo son  $X_1$  (apoyo),  $X_2$  (locus interno),  $X_3$  (locus externo) y  $X_4$  (supervisión) por lo que  $\eta = 4$
2. Las cuatro ecuaciones de regresión tendrían la siguiente forma, considerando las variables predictoras de cada variable endógena. Se plantean tomando las variables endógenas de derecha a izquierda del diagrama:

$$\begin{aligned}
 (3) \quad X_1 &= a - b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 - b_6 X_6 + b_7 X_7 + b_8 X_8 + b_9 X_9 + b_{10} X_{10} + u \\
 X_2 &= a + b_4 X_4 - b_5 X_5 + b_6 X_6 - b_7 X_7 + b_8 X_8 - b_9 X_9 + b_{10} X_{10} + u \\
 X_3 &= a - b_4 X_4 + b_5 X_5 - b_6 X_6 + b_7 X_7 - b_8 X_8 + b_9 X_9 - b_{10} X_{10} + u \\
 X_4 &= a - b_5 X_5 - b_6 X_6 - b_7 X_7 + b_8 X_8 - b_9 X_9 + b_{10} X_{10} + u
 \end{aligned}$$

Como se puede ver, las ecuaciones representan los signos de los coeficientes de regresión para cada variable predictora en función de la relación esperada con la variable predicha o endógena. Cabe destacar que un mayor puntaje de nivel socioeconómico (NSE) representa menor nivel. Supervisión fue categorizada por presencia (1) y ausencia (0). Sexo: mujer 1 y hombre 0. Universidad: pública 0 y privada 1, y estado civil: casado 0 y soltero-divorciado 1.

3. Se realizan los  $\eta$  análisis de regresión. Los resultados aparecen en las tablas siguientes, las cuales contienen los coeficientes de regresión estandarizados y no estandarizados y su nivel significancia asociado. Se toman como significativos los coeficientes con una probabilidad igual o menor a 0.05 ( $\alpha = 5\%$ ).

De esta manera se tiene, en primer lugar, para el *apoyo de amigos* y la combinación lineal de las variables predictoras una correlación moderada, explicándose significativamente el 15% de la varianza total ( $R = .441$ ,  $R^2 = .15$ ,  $gl = 9/169$ ,  $F = 4526$ ,  $p = 0.000$ ). Al evaluar cada una de las variables del modelo y su correlación con el apoyo de amigos se observa que el locus de control externo correlaciona bajo y negativo ( $B = -.256$ ,  $p = 0.01$ ), en el sentido que a menor locus de control externo mayor apoyo de amigos percibido.

Asimismo, correlacionan en forma baja el nivel socioeconómico ( $B = -.181$ ,  $p = .018$ ), la supervisión de personal ( $B = .146$ ,  $p = .048$ ) y el estado civil ( $B = .160$ ,  $p = .042$ ); así, poseer un mayor nivel socioeconómico, supervisar personal y el ser soltero y divorciado se asocia con un mayor apoyo de amigos percibido (Ver tabla 2).

Tabla 2. Coeficientes de regresión y su significancia para apoyo percibido de amigos

	b	Error estándar	Beta	t	p
(Constante)	34,425	3,514		9,796	,000
Locus de control interno	-7.102E-02	,069	-,079	-1,026	,306
Locus de control externo	-,343	,098	-,256	-3,493	,001*
NSE	-,363	,152	-,181	-2,384	,018*
Sexo	,363	,831	,033	,437	,663
Supervisión	1,670	,838	,146	1,992	,048*
Universidad	,152	,796	,014	,191	,848
Antigüedad	-3.093E-02	,067	-,045	-,464	,643*
Edad	-8.525E-02	,057	-,152	-1,490	,138
Estado civil	,932	,454	,160	2,054	,042*

\*significativo al 0.05

Para el *locus de control externo* se presenta una correlación baja y no significativa entre esta variable y el conjunto de variables predictoras, explicándose sólo el 0,7% de la varianza total ( $R = .195$ ,  $R^2 = .007$ ,  $gl = 6/183$ ,  $F = 1.211$ ,  $p = 0.302$ ). Como se puede observar en la tabla 3, ninguna variable predice el locus de control externo.

Tabla 3. Coeficientes de regresión y su significancia para locus de control externo

	b	Error estándar	Beta	t	p
(Constante)	14,394	1,742		8,262	,000
NSE	-,115	,116	-,078	-,991	,323
Sexo	,902	,623	,112	1,447	,150
Supervisión	-,699	,648	-,082	-1,080	,282
Antigüedad	-1.994E-02	,050	-,040	-,400	,689
Edad	6,404E-02	,044	,156	1,453	,148
Estado civil	-,591	,350	-,135	-1,691	,093

\*significativo al 0.05

Por otra parte, al analizar el *locus de control interno* se obtuvo una correlación múltiple moderada baja, explicándose significativamente el 10% de la varianza total ( $R = .358$ ,  $R^2 = .10$ ,  $gl = 6/178$ ,  $F = 4352$ ,  $p = 0.000$ ). La variable que predice el locus de control interno es el sexo con una correlación baja y positiva ( $B = .216$ ,  $p = .004$ ), donde el ser mujer se asocia con mayor locus de control interno en la salud (ver tabla 4).

Tabla 4. Coeficientes de regresión y su significancia para locus de control interno

	b	Error estándar	Beta	t	p
(Constante)	36,948	2,609		14,162	,000
NSE	-,326	,172	-,145	-1,896	,060
Sexo	2,636	,916	,216	2,878	,004*
Supervisión	1,626	,941	,127	1,728	,086
Antigüedad	7,339E-02	,072	,099	1,013	,312
Edad	-6.113E-02	,064	-,098	-,949	,344
Estado civil	-,800	,505	-,121	-1,585	,115

\*significativo al 0.05

Por último, la variable *supervisión del personal* correlaciona bajo pero significativamente con la combinación lineal de las variables involucradas en el modelo, las cuales explican el 5% de esta variable ( $R = .337$ ,  $R^2 = .05$ ,  $gl = 4/191$ ,  $F = 1.693$ ,  $p = 0.073$ ). De las predictoras, se tiene que sólo el nivel socioeconómico se relaciona con la supervisión de personal ( $B = -.216$ ,  $p = .004$ ), en el sentido de que a mayor nivel socioeconómico mayor posibilidad de supervisar personal (ver tabla 5).

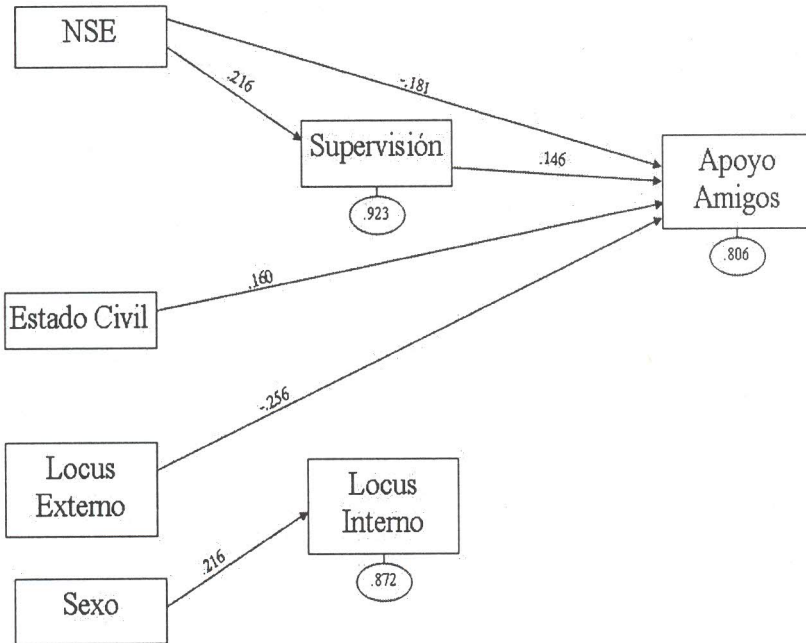
Tabla 5. Coeficientes de regresión y su significancia para supervisión de personal

	b	Error estándar	Beta	t	p
(Constante)	,517	,190		2,719	,007
NSE	-3.721E-02	,013	-,216	-2,940	,004*
Sexo	-8.682E-02	,068	-,092	-1,272	,205
Antigüedad	7,664E-03	,006	,132	1,387	,167
Edad	4,088E-03	,005	,085	,882	,379

\*significativo al 0.05

4. Se construye el diagrama de rutas resultante. En este caso con los resultados de  $p \geq 0.05$ .





**Figura 8.** Diagrama de ruta resultante en la explicación del apoyo social percibido de amigos. Fragmento de Angelucci (2001).

5. Como existen rutas indirectas, se calculan los efectos indirectos, por ejemplo  $X_5$  (nivel socioeconómico) sobre  $X_1$  (apoyo), siendo igual a  $\beta_{45} * \beta_{14}$ , cubriendo la ruta  $X_5 \rightarrow X_4 \rightarrow X_1$ . Es decir, el nivel socioeconómico tiene un efecto directo sobre apoyo percibido de amigos, pero además tiene un efecto directo mediado (pasando por) por la variable tener o no un cargo de supervisor. Así se multiplican los coeficientes:  $0.216 * 0.146$ , siendo el efecto indirecto igual a 0.032.

En la tabla 6 se presenta los efectos directos, indirectos y totales para la variable apoyo de amigos únicamente. Sólo se tomaron los efectos significativos estadísticamente. Como se puede evidenciar, los efectos totales más altos se presentan entre el apoyo percibido de amigos y locus externo, supervisión y NSE, en ese orden. El único efecto indirecto es con NSE.

**Tabla 6.** Efectos entre el apoyo percibido de amigos y las variables del modelo

Variable Dependiente	Variable Independiente	Efectos		
		Directos	Indirectos	Totales
Apoyo percibido de amigos	Locus interno	---	---	---
	Locus externo	-0.256	---	-0.256
	Supervisión	0.146	---	0.146
	NSE	-0.181	0.032	-0.149
	Universidad	---	---	---
	Sexo	---	---	---
	Antigüedad	---	---	---
	Estado civil	0.160	---	---
Edad	---	---	---	

Este cálculo puede hacerse para cada una de las variables endógenas del modelo.

### ALCANCES DEL ANÁLISIS DE RUTA

Según Ortiz-Pulido (2000) cuando se considera la temporalidad en los sistemas, el análisis de rutas da frecuentemente mejores resultados que otros análisis, principalmente debido a que:

- Permite construir rutas causales. Uno de las mayores ventajas de este tipo de análisis es que verifica los presupuestos de una teoría acerca de las relaciones entre sus términos más que simplemente evaluar un conjunto de datos basada en alguna relación lineal.
- Provee un medio para descomponer la correlación entre dos variables en componentes que representen contribuciones causales y no causales.
- Su estructura no es fija, en el sentido de que una misma variable puede cumplir la función de variable dependiente e independiente a la vez.
- Las aplicaciones del análisis de rutas son innumerables. En Psicología se encuentran estudios en áreas como organizacional, clínica, escolar, social, salud, básica, entre otras.

### LIMITACIONES DEL ANÁLISIS DE RUTA

Se adecua más a un análisis exploratorio que confirmatorio. En este sentido, algunos autores establecen que su utilidad es restringida para evaluar pocas hipótesis o en estadios tempranos o exploratorios de la investigación (Webley y Lea, *s/f*).

Cuando se observa la asociación o falta de asociación entre dos dimensiones, hay que tomar en consideración que la relación observada puede “depende” de influencias distorsionantes de otras variables, pudiéndose dar una relación espúrea.

Hay que tomar en cuenta que el análisis de rutas es usado para evaluar hipótesis acerca de relaciones causales entre variables, lo que supone un conocimiento previo acerca del fenómeno por parte del investigador, además del manejo de las asunciones (epistemológicas y técnicas) sobre las cuales se basa esta estrategia.

## REFERENCIAS BIBLIOGRÁFICAS

- Angelucci, L.; Lazo, V.; Loaiza, G.; Montiel, L.; Mora, N.; Páez, V.; Perales, A.; Rojas C. y Rojas, J. (2006). *Distancia social hacia las personas con esquizofrenia: Influencia de la familiaridad, percepción de peligrosidad, atribución causal, carrera y sexo*. Trabajo no publicado presentado en el V Congreso Iberoamericano de Psicología Clínica y de la Salud. San José, Costa Rica, 19- 23 de septiembre.
- Angelucci, L. (2001). *Influencia de las variables sociodemográficas, el apoyo social y locus de control de salud sobre la salud en empleados universitarios*. Trabajo de ascenso para optar a categoría de agregado. Departamento de Ciencia y Tecnología del Comportamiento. Universidad Simón Bolívar, Caracas, Venezuela.
- Anguera, M. (1989). “Hacia una representación conceptual: Teorías y modelos”. En J. Mayor y J. Pinillos (Eds.) *Tratado de Psicología General* (tomo 1). Madrid: Alambra.
- Arnau, J. (1982). *Psicología experimental: un enfoque metodológico*. México: Trillas.
- Baltes, P.; Reese, H. y Nesselroade, J. (1980). *Métodos de investigación en psicología evolutiva: Enfoque del ciclo vital*. Madrid: Morata.
- Bentler, P. (1980). “*Multivariate analysis with latent variables: Causal modeling*”. *Annual Review of Psychology*, 31, 419-456.
- Blalock, H. (1984). *Construcción de teorías en las ciencias sociales*. México: Trillas.
- Bross, I. (1958). *La decisión estadística*. Madrid: Aguilar.
- Bunge, M. (1981). *Epistemología*. Barcelona: Ariel.

- Casas, M. (2002). "Los modelos de ecuaciones estructurales y su aplicación en el Índice Europeo de Satisfacción del Cliente". En *Rect@ Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA, actas 10*, 1-11. Recuperado el 15 de febrero de 2007 de <http://www.doaj.org/doaj?func=abstract&cid=176630&toc=y>
- Cohen, J.; Cohen, P.; Wert, S. y Aiken, L. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. (3ra. edición). Nahwah: Lawrence Erlbaum.
- Critto, A. (1982). *El método científico en las ciencias sociales*. Buenos Aires: Paidós.
- De la Garza, J. (1995). *Mercadología*. México: Alambra.
- Gintis, H. (2006) *Towards a Unified Behavioral Science*. Recuperado el 15 de febrero 2007 de <http://www.umass.edu/preferen/gintis/TowardsUnity.pdf>
- Hair, J.; Anseron, R.; Tatham, R. y Black, W. (2000). *Análisis multivariante*. (5ta. edición). Madrid: Prentice Hall.
- Keith, T. (2005). *Multiple regression and beyond*. New York: Pearson.
- Kerlinger y Lee, (2002). *Investigación del comportamiento: Métodos de investigación en Ciencias Sociales*. México: Mc Graw Hill
- Klem, L. (1995). "Path analysis". En *Reading and understanding multivariate statistic*, L. Grimm y P. Yarnold (eds.) (pp. 65-98). Washington: American Psychological Association.
- Lance, C.; Cornwell, J. y Mulaik, S. (1988). "Limited information parameter estimates for latent and mixed manifest and latent variable models". *Multivariate Behavioral Research*, 32, 215-254.
- Mulaik, S. (1987). "Toward a conception of causality applicable to experimentation and causal modeling". *Child Development*, 58, 18-32.
- Ortiz-Pulido, R. (2000). "Análisis de rutas en biología: Estadística para sistemas multicausales". *Interciencia*, 25, 329-336 (versión electrónica). Recuperado el 12 de febrero 2007 de <http://redalyc.uaemex.mx/redalyc/pdf/339/33904903.pdf>

- Peña, G. (1999). *Enfoque causal-comparativo*. Material no publicado de la cátedra de Psicología experimental. Escuela de Psicología, Universidad Católica Andrés Bello, Caracas, Venezuela.
- \_\_\_\_\_ (1995). "Dos factores explicativos de la cuantofobia: la revolución de la inferencia estadística y el factor humano en la concepción de los métodos estadísticos". *Revista AVEPSO*, XVIII, 2, 65-83.
- \_\_\_\_\_ (2006). "De una psicología caleidoscópica a un cuerpo disciplinar integrado funcionalmente". En G. Peña; Y. Cañoto y Z. Santalla (eds) *Una introducción a la psicología*. Caracas: Universidad Católica Andrés Bello.
- Robles, J. (2000). *Introducción al análisis de ruta*. Guía de estudio (no publicada). Cátedra de Metodología II. Escuela de Psicología. Universidad Católica Andrés Bello, Caracas, Venezuela.
- \_\_\_\_\_ (1992). *Contrastación del modelo n: mayor efecto del tamaño muestral que el método de estimación y especificación sobre los índices de ajuste de un modelo estructural*. Tesis de Licenciatura en Psicología (no publicada), Universidad Católica Andrés Bello, Caracas, Venezuela.
- Sierra Bravo, R. (1981). *Ciencias sociales: análisis estadístico y modelos matemáticos*. Madrid: Paraninfo.
- \_\_\_\_\_ (1995). *Técnicas de investigación social: teoría y ejercicios*. (10ma. edición). Madrid: Paraninfo.
- Wartofsky, M. (1976). *Introducción a la filosofía de la ciencia*. Madrid: Alianza Editorial.
- Webley, P. y Lea, S. (s/f). *PSY6003 Advanced statistic: Multivariate analysis II: Manifest variables analyses. Topic 3: Path analysis*. Department of Psychology, University of Exeter, United Kingdom. Recuperado el 12 de febrero 2007 de <http://www.people.ex.ac.uk/SEGLea/multivar2/pathanal.html>
- Wolfe, L. (1980). "Strategies of path análisis". *American Educational Research Journal*, 17, 183-209.
- Wright, S. (1934). "The method of path coefficients". *Annals of Mathematical Statistics*, 5, 161-215.