

¿Cómo se obtiene una ley científica?

Obtención de leyes científicas
con el método correlacional

Mg. David Medianero Burga

¿Como se obtiene una ley científica?

**Obtención de leyes científicas
con el método correlacional**

Mag. David Medianero Burga

ORCID N.º 0000-0002-3886-7617

Profesor de la Universidad Nacional Mayor de San Marcos

Asesor Metodológico de la Escuela Nacional de Control

Resumen

En el contexto de los estudios analíticos, el primero y más sencillo de los diseños metodológicos de investigación es el denominado *método correlacional* o *método ecológico* (termino este último utilizado frecuentemente en las ciencias biomédicas). Consiste, como su nombre lo sugiere, en relacionar la variable de causa con la variable de efecto y de este modo sencillo pero también ingenuo establecer o confirmar una relación de causalidad. Sencillo, en tanto solo exige contar con ciertos datos sobre dos únicas variables. Ingenuo, debido precisamente a su sencillez: omite el efecto de otras variables que podrían incidir simultáneamente, con menor o mayor fuerza, en el efecto que es objeto de estudio. Pese a todo, las investigaciones casi siempre se inician con este método, hecho que explica su abrumadora popularidad entre los investigadores, especialmente en el campo de las disciplinas relacionadas a la gestión pública.

Palabras clave: *correlación, método correlacional, diagrama de dispersión, coeficiente de correlación, regresión, causalidad.*

Abstract

In the context of analytical studies, the first and simplest of the research methodological designs is the so-called correlational method or ecological method (the latter term frequently used in biomedical sciences). It consists, as its name suggests, in relating the cause variable to the effect variable and in this simple but also ingenious way establishing or confirming a causal relationship. Simple, insofar as it only requires having certain data on only two variables. Naive, precisely due to its simplicity: it omits the effect of other variables that could simultaneously affect, with greater or lesser force, the effect that is the object of study. Despite everything, investigations almost always start with this method, a fact that explains its overwhelming popularity among researchers, especially in the field of disciplines related to public management.

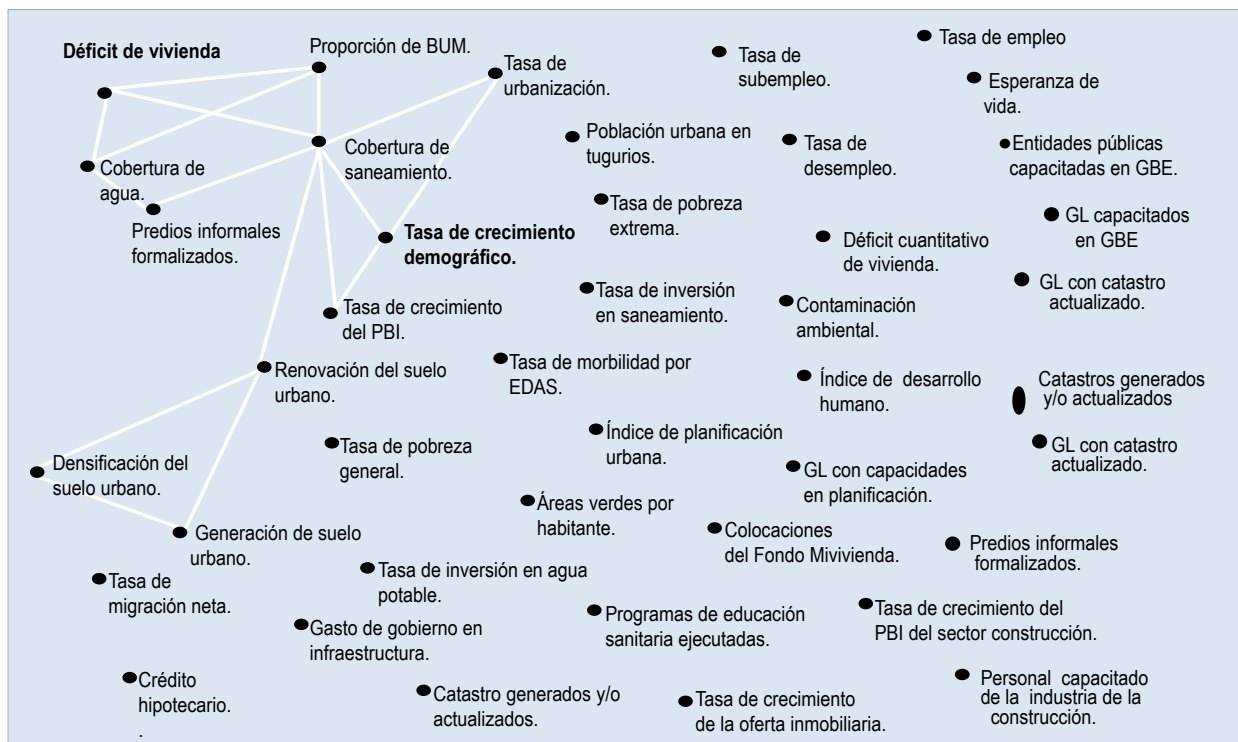
Keywords: *correlation, correlational method, scatter diagram, correlation coefficient, regression, causality.*

Introducción: ¿Qué es una ley científica?

Por definición, una ley científica es una relación necesaria, constante y general entre distintos hechos (Yuren, 2002). Una vez obtenida e integrada con otras leyes como parte de una teoría, ella sirve al mundo de un modo tan amplio y profundo que con justicia la ciencia se ha convertido en el recurso más valioso de la humanidad, inmensamente mayor que la inmensa naturaleza. El mundo se caracteriza por su variabilidad y diversidad, tanto en la esfera natural como cultural. Los hechos de la realidad varían sin cesar y las relaciones entre ellos varían incesantemente también, por razones naturales o inducidas. La turbulencia que emerge como resultado, sin embargo, empieza a escampar mediante un sencillo procedimiento gráfico por el cual se logran atisbar ciertas regularidades empíricas. Estas, tomadas como puntos de apoyo, ayudan a inferir modelos y leyes científicas que permiten resolver los problemas del mundo, grandes o pequeños, bien describiéndolos, o también explicándolos.

Si, por ejemplo, se buscara explicar el déficit de vivienda en un país y se escudriñara en los antecedentes y el contexto del problema, se encontraría una gran variedad de otros problemas u otras variables relacionadas. De todas ellas, podríamos seleccionar las más importantes y de estas, finalmente, elegir una que se postularía como la causa principal del problema. Por ejemplo, la tasa de crecimiento demográfico. A partir de este momento el trabajo del investigador consistiría en observar la relación entre los cambios de valores en ambas variables, recopilando datos sobre los sujetos relevantes, tales como países o ciudades.

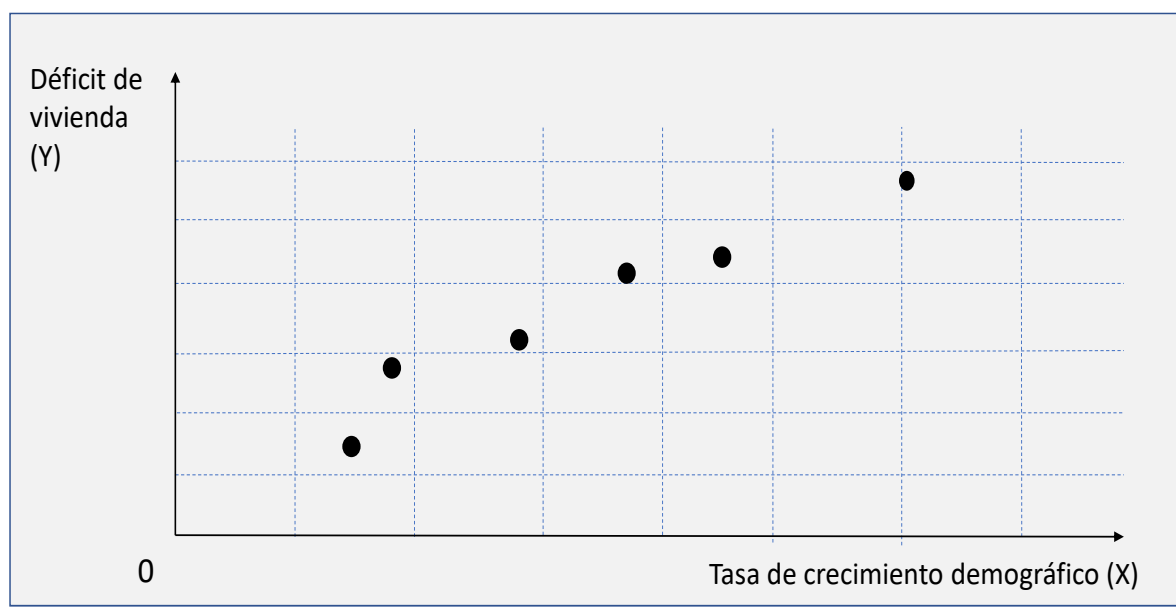
Figura 1. Representación gráfica de la complejidad real



Las leyes científicas son relaciones entre variables y estas se refieren a hechos. Los hechos son eventos, propiedades o condiciones de la realidad. Una relación, si bien no es un hecho, pues

no tiene existencia concreta, representa la conexión mental que hacemos de unos hechos con otros, y que en ocasiones puede permitirnos descubrir una relación de causalidad. Por ejemplo, podemos observar la relación entre el déficit de vivienda y la tasa de crecimiento demográfico en cinco ciudades de un país, o en cinco países distintos.

Figura 2. Simplificación de la realidad

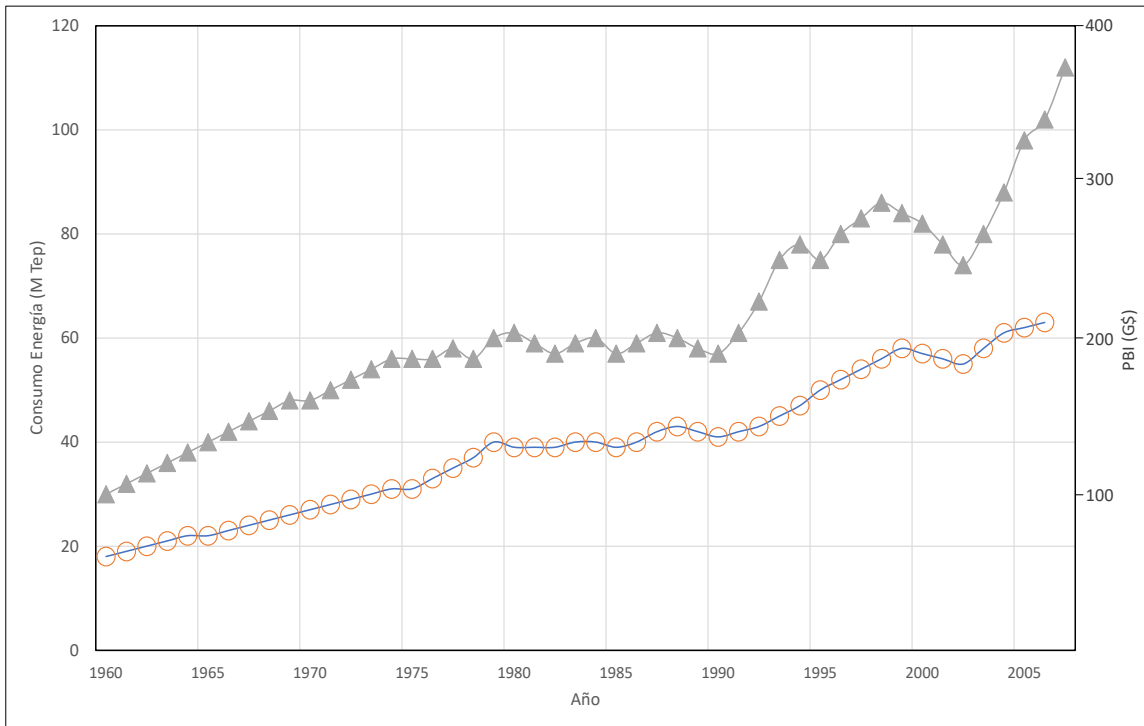


Sin embargo, no todas las relaciones son igualmente importantes o relevantes. Desde la perspectiva de la ciencia interesan aquellas relaciones que son, al mismo tiempo, necesarias, constantes y generales. Dado que el edificio de la ciencia se construye uniendo de forma coherente todos los conocimientos que emergen de la investigación, solo pueden ser aceptadas relaciones de carácter permanente y no circunstanciales, necesarias y no contingentes, generales y no particulares (Yuren, 2002). En un mundo naturalmente cambiante y caótico, el mundo de la ciencia intenta en compensación dotarse de un orden, y de cierta estabilidad y uniformidad, aun cuando es consciente de la inevitable caducidad de sus pocas certezas, las que a la larga probablemente no podrán evitar ser absorbidas en el sumidero de viejas y nuevas incertidumbres.

Razonamiento cartesiano

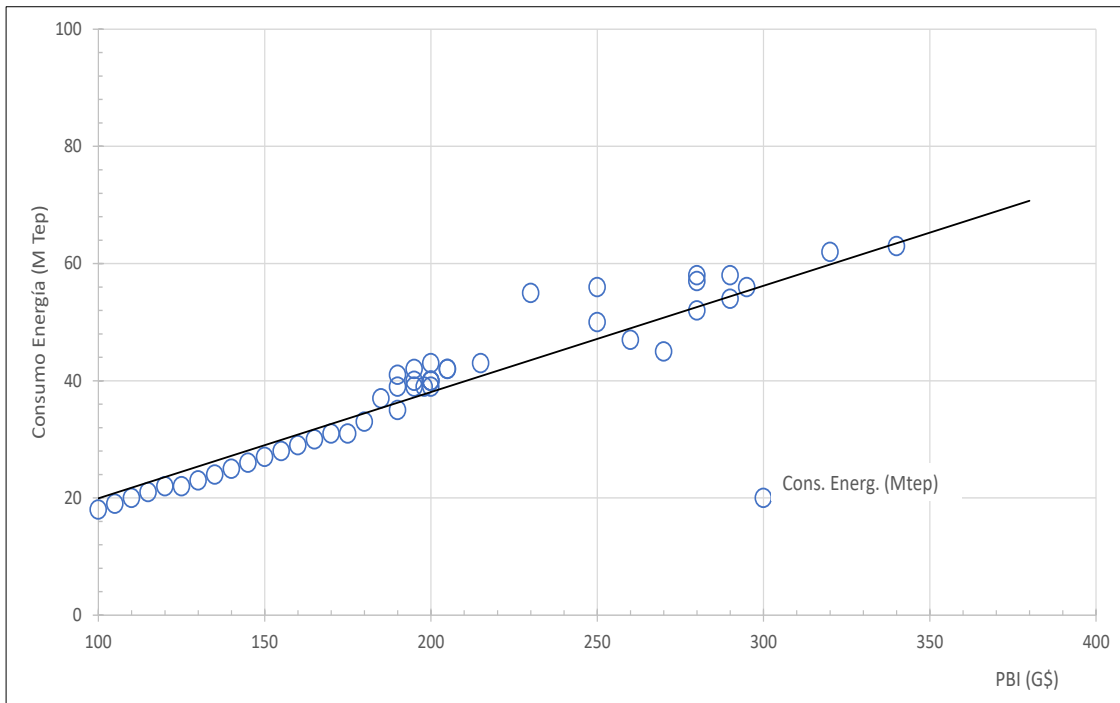
La metodología más conocida y popular para la obtención de leyes científicas es un gráfico concebido poco después que René Descartes inventara el sistema de ejes coordenados. Se conoce, como ha sido señalado, con el nombre de diagrama de dispersión o nube de puntos, y permite visualizar la relación entre dos hechos, simbólicamente denotados por las variables X e Y, la primera en calidad de causa o variable independiente y la segunda como efecto o variable dependiente. En virtud de esta simple consideración, el diagrama de dispersión se erige como la primera y más simple herramienta de la investigación analítica o explicativa, y al conjunto de procedimientos que se levantan sobre sus hombros, incluyendo el coeficiente de correlación y la ecuación de regresión, se denomina método correlacional.

Figura 3. Representación de las variables como función del tiempo



Fuente: Adaptado de Gil, S. (2014).

Figura 4. Representación del consumo de energía como función del PBI



Fuente: Adaptado de Gil, S. (2014).

La representación gráfica de los datos de las variables sobre las que se presume una relación potencialmente causal es una gran fuente de inspiración para desarrollar modelos, leyes y teorías

científicas. Por ejemplo, el crecimiento del PBI y el consumo de energía se encuentran íntimamente relacionados: varían en forma conjunta en el tiempo, por lo que se pueden comparar tanto entre sí, como cada uno de ellos por separado con la variable tiempo, tal como se puede ver en las dos figuras siguientes (Gil, 2014).

Simplificación del proceso

Un gráfico de este tipo marca el inicio modesto de lo que más adelante tomará la forma de modelo y que concluirá, finalmente, en la formulación de una nueva ley o una nueva teoría científica. En su relato más simple, para configurar una ley o modelo basta con tener dos puntos en el plano: dos observaciones de las variables de interés en dos sujetos de estudio; luego unirlos con una línea recta, calcular sus parámetros y deducir finalmente la ecuación o modelo matemático. De este modo, una cierta regularidad empírica da un salto de gigante en la escalera de la abstracción científica y se convierte en ley, asumiendo que el patrón de comportamiento de las variables es de carácter permanente, pues tiende a ocurrir siempre o casi siempre, y es general, en tanto y en cuanto afecta a todos los sujetos de estudio en un marco temporal y espacial razonables.

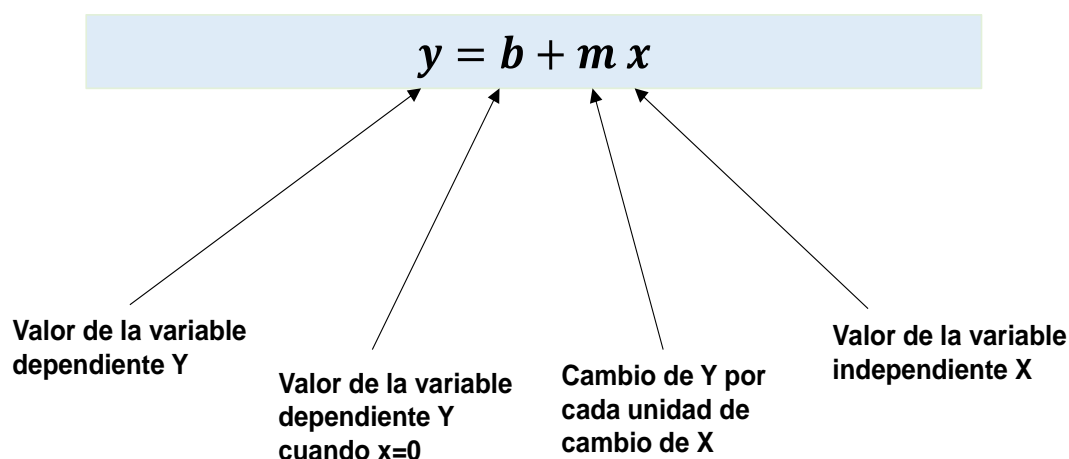
La intuitiva y cotidiana relación entre las variables de ingreso y consumo, tal como se muestran en la tabla y figura siguientes, ofrece una ilustración especialmente didáctica del proceso de obtención de una ley científica.

Datos de ingreso y consumo

Puntos del diagrama (Sujetos de estudio)	Variable X: Ingreso (S/.)	Variable Y: Consumo (S/.)
A	100	90
B	200	170

Para deducir, a partir de estos datos, la ecuación que demuestra la relación entre ingreso y consumo en forma cuantitativa, es necesario calcular los parámetros; es decir, los valores constantes que permiten expresar el valor que tendrá el consumo ante determinados valores de ingreso. Esto se logra, en su versión más sencilla, utilizando la denominada ecuación explícita de la recta:

Figura 5. Gráfica de la relación entre ingreso y consumo



Esta ecuación es una expresión compacta que captura toda la información necesaria para probar una hipótesis o modelo causal, cuando la relación se presume simple (solo dos variables) y lineal (razón de cambio constante). En matemáticas se conoce como ecuación lineal simple y en mundo específico de la investigación se denomina modelo lineal simple. Ahora bien, dado que en la tabla de datos mostrada tenemos los valores de x e y , la cuestión se reduce a encontrar los valores del intercepto (b) y la pendiente (m). Estos son los parámetros de la ecuación, aquellos que explican como determinados valores de X se transforman en ciertos valores de Y , y por lo tanto logran explicar en forma cuantitativa la relación entre estas variables.

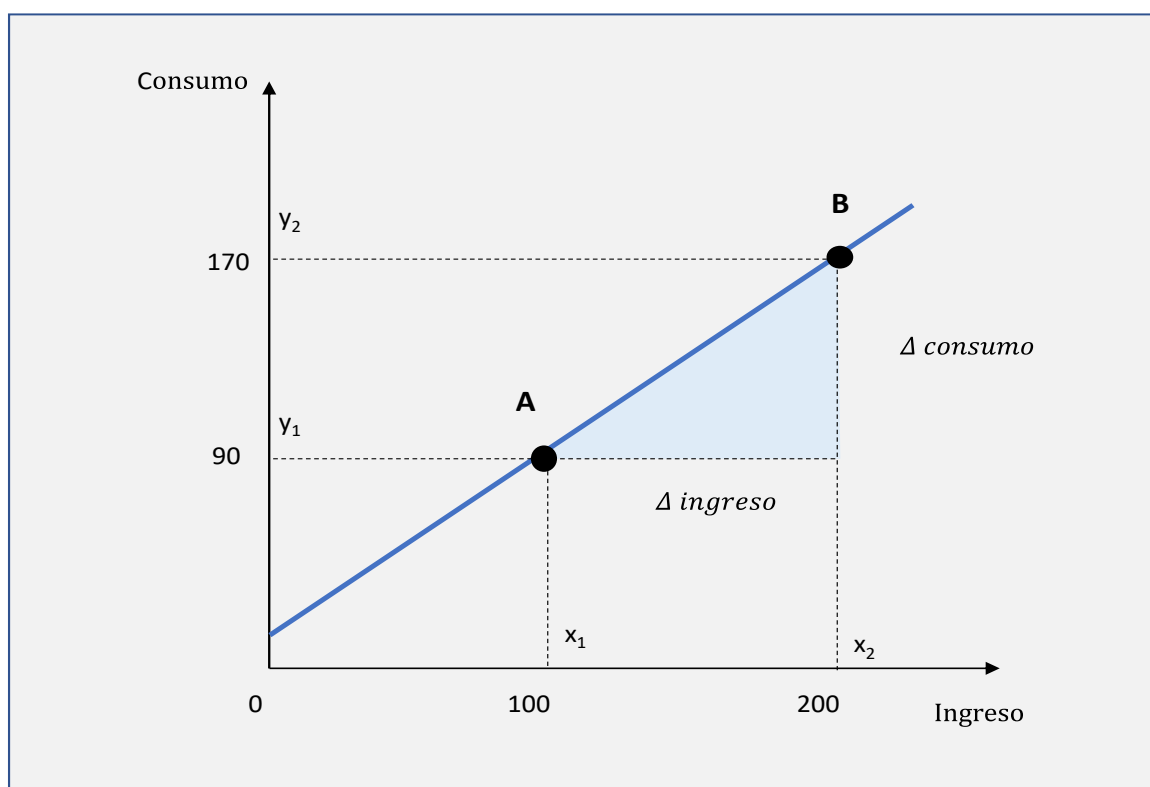
En primer lugar, se calcula la pendiente (m), que expresa la proporción del ingreso que permanentemente las personas tienden a destinar al consumo, lo cual se realiza mediante la comparación del incremento que se observa en Y con el incremento de X .

$$m = \frac{\Delta \text{consumo}}{\Delta \text{ingreso}} = \frac{Y_2 - Y_1}{X_2 - X_1}$$

$$m = \frac{170 - 90}{200 - 100} = \frac{80}{100} = 0.80$$

Si el valor de la pendiente es 0.80, como ocurre en este ejemplo, ello indica que, del total de su ingreso, las personas dedican el 80 % al gasto de consumo. En otras palabras, por cada un sol de incremento del ingreso, el consumo se incrementa en 80 céntimos.

Figura 6. Gráfica de la relación entre ingreso y consumo



Acto seguido se calcula el intercepto u ordenada al origen, que refleja el gasto de consumo que las personas realizan con independencia de su ingreso y que podría referirse a un subsidio que reciben de un pariente o del estado, o a la valorización del autoconsumo de producción propia, hecha en casa o en chacra. Para ello, se parte de la ecuación de la recta, se deduce b y se reemplazan los valores que ya se tienen: ingreso (x), consumo (y) y pendiente (m).

$$y = b + m x$$

$$b = y - m x$$

$$b = 90 - 0.80(100) = 90 - 80 = 10$$

El valor obtenido de 10 nos indica que las personas gastan 10 soles, independientemente del monto de su ingreso. El intercepto se ubica en el lugar geométrico donde la recta corta al eje de las ordenadas, eje Y, pues, en ese punto el ingreso, representado en el eje X, es cero.

Como ya se cuentan con ambos parámetros, pendiente e intercepto, reemplazamos estos valores en la ecuación general y obtenemos, finalmente, la expresión cuantitativa de la relación entre estas dos variables. Los parámetros modelan el comportamiento de las variables, razón por la cual para la ciencia el mundo, o es un mundo paramétrico, o no es mundo.

$$Y = b + m x$$

$$Y = 10 + 0.80x$$

Y así, de este modo intuitivo y lógico se logra obtener una ley que relaciona las variables de ingreso y consumo. O esto es lo que dijo Keynes, su legendario autor, que la llamó Ley psicológica fundamental, que "...consiste en que los hombres (y las mujeres), como regla general y en promedio, están dispuestos a incrementar su consumo a medida que aumenta su ingreso, pero no en la misma cuantía del aumento en su ingreso" (Keynes, 1936; citado por Gujarati, 2010).

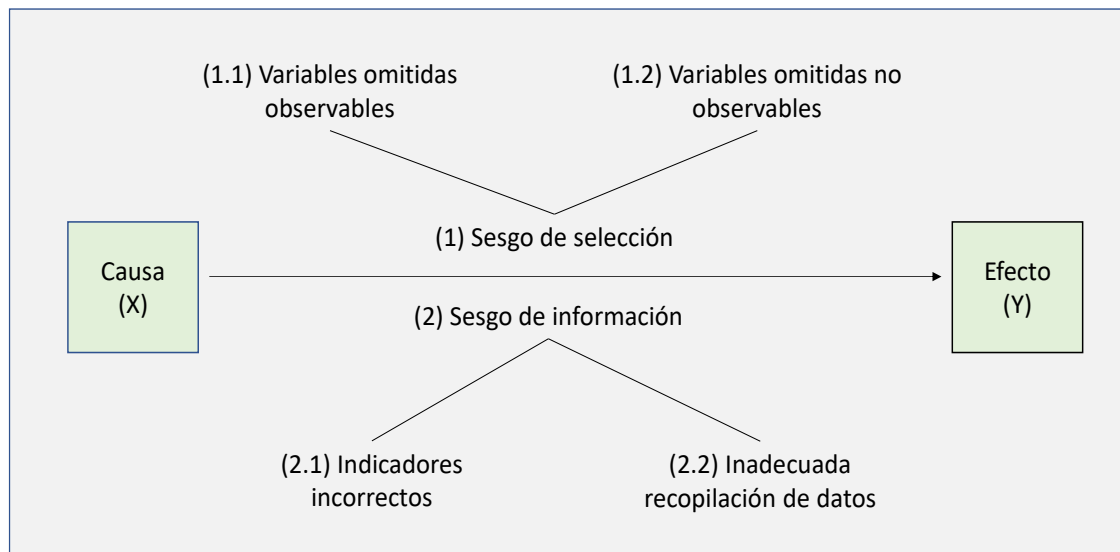
Ahora X podría tomar cualquier valor y ya no solo los que están en la tabla de datos, que apenas es una muestra, y en todos los casos se podría predecir automáticamente el valor de Y. La virtud de este procedimiento radica en su capacidad de generalización del conocimiento: la ecuación permite pasar de lo particular (los datos obtenidos mediante una encuesta o tomados de una serie estadística) a lo general, de modo que en cualquier situación similar se podrían determinar los valores de la variable dependiente a partir de los valores de la variable independiente. La ecuación no tiene un significado real particular, sino un significado formal general. Es una abstracción de la realidad, válida para cualquier situación que involucre a las variables estudiadas.

En términos generales, es de esta forma que ciertas regularidades empíricas se convierten en leyes científicas, asumiendo que los patrones de comportamiento de las variables son, en un marco temporal y espacial pertinente y razonable, de vigencia suficientemente permanente y de alcance suficientemente general. El camino de la ciencia supone algún contexto estable, alguna teoría o paradigma compartidos, dentro de los cuales las nuevas explicaciones derivadas de la investigación puedan adquirir sentido, aun cuando desde una perspectiva más amplia su vigencia talvez solo sea efímera.

Amenazas a la validez

Las ecuaciones o modelos, sin embargo, se logran siempre al precio de una drástica simplificación de la realidad: este es el papel de los supuestos, en ocasiones altamente controversial. Las amenazas a la validez de un modelo provienen de dos fuentes genéricas: las variables relevantes que no se hayan tomado en cuenta, sean o no observables; y la deficiente recolección de datos, ya sea por un inadecuado muestreo o por la selección de indicadores o constructos inadecuados, aunque esto último se da principalmente cuando se trabaja con variables cualitativas.

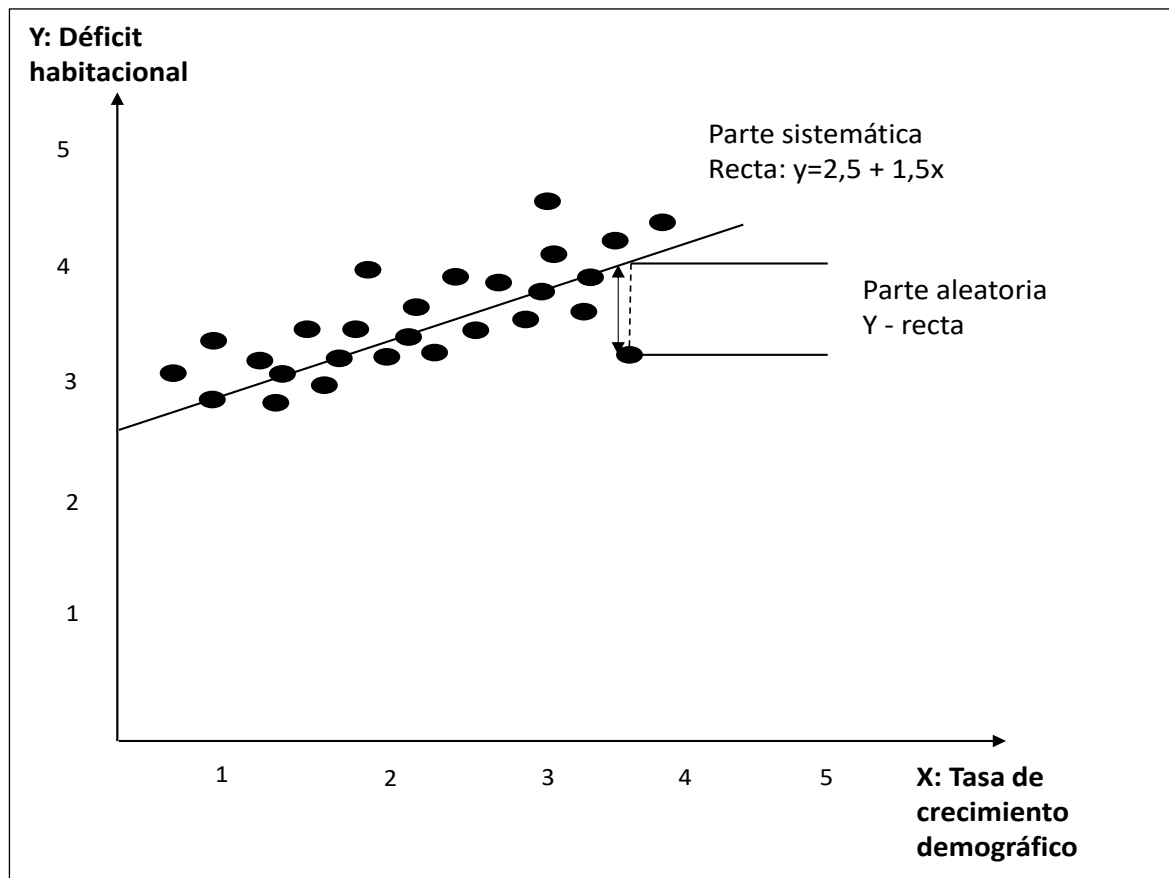
Figura 7. Amenazas a la validez de una ley científica



No se debe pensar por lo tanto que, obtenida la ecuación, los cambios de la variable dependiente Y son explicados totalmente por la variable independiente X elegida. Por ejemplo, no todo el déficit de vivienda en las ciudades es explicado o determinado por la tasa de crecimiento demográfico. Otras variables podrían estar operando “en la sombra”: tasa de empleo, tasa de inflación y tasa de interés bancario, por ejemplo. El efecto de estas otras variables es lo que gráficamente se expresa en la diferencia que existe entre los puntos de la recta de regresión (recuérdese que una recta está formada por un conjunto de puntos) y los puntos construidos con los datos reales obtenidos por el investigador.

La recta de regresión y su correspondiente ecuación solo representa la parte sistemática de la relación entre las variables, entendido el termino sistemático como tendencia o promedio. La diferencia entre el punto observado real y el punto correspondiente sobre la recta representa la parte aleatoria del déficit de vivienda, no explicada por la tasa de crecimiento demográfico, sino por las otras variables (también llamados sesgos de selección) y también por los errores derivados de los sesgos de información (muestreo, indicadores o unidades de medida inadecuados y recopilación deficiente de datos).

Figura 8. Divergencia entre la abstracción científica y la realidad



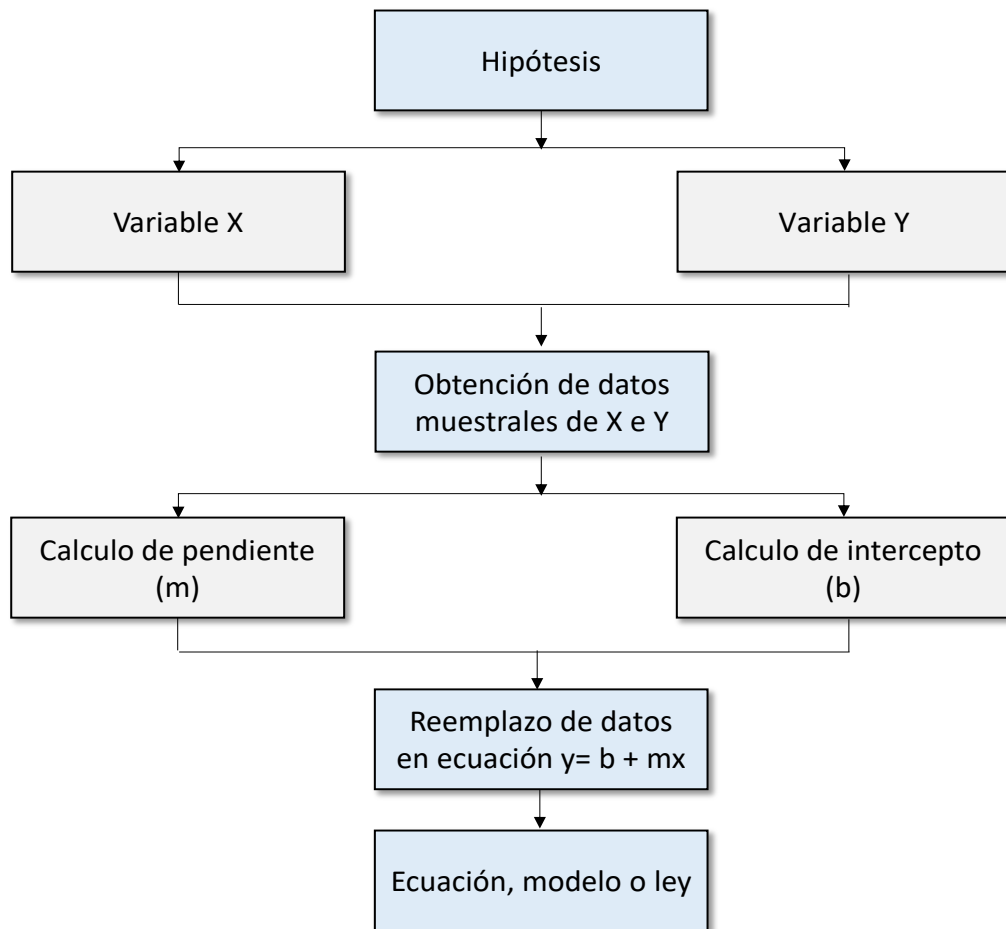
Adaptado de: Peña, D. (2014)

Vida de una ley

En conclusión, aunque el proceso de obtención de una ley científica es largo y complejo, la lógica subyacente puede ser percibida de forma sencilla mediante el soporte visual que ofrece el diagrama de dispersión en un plano cartesiano. Toda ley nace con una hipótesis. Esta es, primero operativizada, luego sometida a contrastación empírica, y finalmente convertida en una ecuación matemática.

En este proceso, tan importante como la intuición para generar hipótesis plausibles, es la investigación empírica sobre las variables. Y así como lo primero exige contar con un sólido marco teórico, lo segundo implica conocer determinadas técnicas estadísticas, desde la habilidad para transformar hipótesis en preguntas de cuestionario, hasta el diseño de la muestra, la recopilación de los datos y el procesamiento de la información. Con los datos recopilados y utilizando la ecuación de la recta, de preferencia en su forma explícita, se calculan los parámetros, tanto la pendiente como el intercepto. El proceso está casi concluido: los valores de los parámetros se reemplazan en la ecuación o, de modo más general, se usan las fórmulas del criterio denominado Mínimos Cuadrados Ordinarios, y se obtiene, en la cúspide del trabajo de investigación, la ecuación o ley científica.

Figura 9. Proceso de obtención de leyes científicas



Pero ahora el proceso nuevamente pareciera demasiado complejo, con muchas formulas y muchos datos numéricos. En realidad, la lógica sería la misma si uno deseara trabajar sin números y sin formulas. Solo habría que asumir que la relación causal que se espera demostrar debiera tener tres tipos de evidencia, bajo la forma de documentos, entrevistas u opiniones. Primero, evidencia suficiente para confirmar empíricamente que el problema que se quiere explicar varía (aumenta o disminuye) conjuntamente con los factores considerados como factores causales. Segundo, que existen otros factores no tomados en cuenta que afectan la relación, pero cuyo efecto sobre el problema no es tan fuerte como los primeros. Y tercero, que las evidencias encontradas conforman una masa de información suficiente para estar seguros que la relación causal postulada no es una excepción o un error.

De este modo, el investigador cualitativo si bien no tendría en rigor una ecuación, si mostraría un razonamiento plausible y empíricamente sustentable de la existencia de una relación causal. En la escala de la ciencia, si bien la medición y por ende el uso de variables cuantitativas es crucial, eso no significa que siempre se deba empezar con mediciones rigurosas, y que los análisis cualitativos no deban considerarse como investigaciones académicas de considerable valor, un paso necesario en la perspectiva más amplia de obtención de las leyes que explican el mundo natural o social.

Destino de una ley

Una vez obtenida, el destino de una ley, incluso el de las leyes exitosamente formuladas, es dolorosamente incierto. Podría subsistir en el mundo académico como un fragmento de conocimiento aislado, en calidad de explicación plausible y comprobada de una de las innumerables relaciones causales, aunque relativizada en la maraña de relaciones causales que forman y dan forma al mundo. O más bien, podrían integrarse de forma coherente con otras leyes, que a su vez expliquen relaciones causales adyacentes, y gracias a esta cercanía y a su carácter mutuamente complementario, integrarse y dar forma a una teoría. Es el destino ideal para una ley, el destino merecido luego del largo y esforzado proceso que implicó su obtención.

Una ilustración perfecta del rumbo positivo de una ley es el que han tenido, por un lado, la ley de la demanda y, por otro, la ley de la oferta.

La ley de demanda afirma:

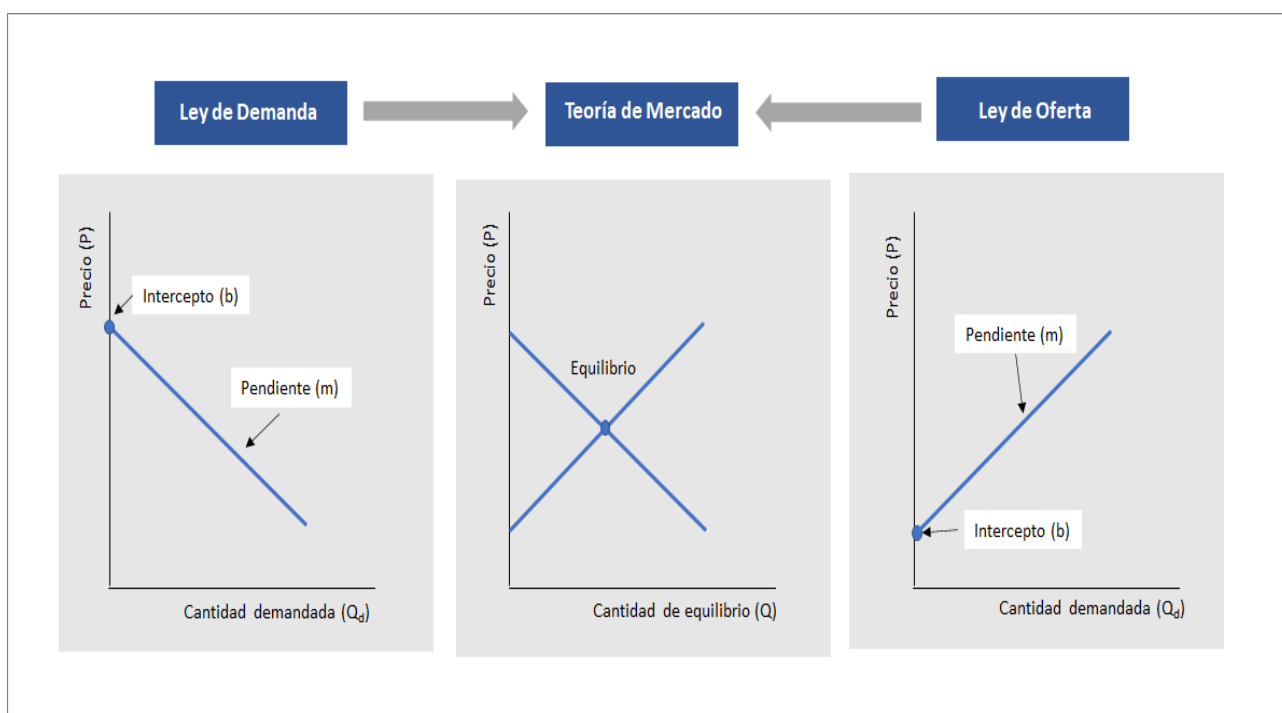
Si todo lo demás permanece constante, cuando sube el precio de un bien, la cantidad demandada disminuye; y cuando el precio de un bien baja, la cantidad demandada aumenta.

Por su parte, la ley de la oferta afirma:

Si todo lo demás permanece constante, cuando sube el precio de un bien, aumenta la cantidad ofrecida; y cuando baja el precio de un bien, disminuye la cantidad ofrecida.

Como es ampliamente conocido, ambas leyes se integraron finalmente en la teoría del equilibrio de mercado y esta teoría que en sus fundamentos originarios fuera formulada por Alfred Marshal hoy constituye la teoría más aceptada del funcionamiento del mercado; y como este es el núcleo del sistema económico, la teoría de la oferta y la demanda es la teoría más importante del funcionamiento de la economía.

Figura 10. Ley y teoría



Así, el noble investigador, al principio acompañado solo de una hipótesis y unos exiguos datos, emprende, como si tratara de un épico emprendimiento productivo, su esforzado trabajo empírico. Comprueba su modelo y formula una ley. En algún otro lugar, otros investigadores, o talvez el mismo investigador en otros momentos, formulan otras leyes, las cuales en su destino ideal logran paulatinamente integrarse hasta formar una férrea unidad conceptual. El investigador, o los investigadores, aun sin conocerse entre sí, podrán entregar finalmente al mundo un sistema explicativo completo de alguna dimensión importante de la realidad. El resultado de este proceso, tan largo como emocionante, recibe el nombre nunca tan mercedamente solemne de *teoría*.

Referencias

Gil, S. (2014). *Experimentos de Física, usando TIC y elementos de bajo costo*. Alfaomega.

Triola, M. F. (2014). *Estadística*. Pearson educación.

Wooldridge, J. (2010). *Introducción a la econometría*. Gengage Learning.

Yuren, M. (2002). *Leyes, teorías y modelos*. Trillas.

Gujarati, D (2011). *Econometría Básica*. Mc Graw Hill.