

## REGRESIÓN MÚLTIPLE EN LOS TRIBUNALES<sup>1</sup>

El análisis de regresión múltiple es un instrumento para entender la relación entre dos o más variables. Implica una variable explicada – la *variable dependiente* – y variables adicionales (las *variables independientes*) que se estima producen o están asociadas con los cambios de la variable dependiente. Por ejemplo, mediante un análisis de regresión múltiple es posible estimar el efecto que tiene la cantidad de años trabajados sobre el salario. El salario sería la variable dependiente a ser explicada; los años de experiencia la variable explicativa. El análisis de regresión múltiple muchas veces es apropiado para analizar datos entre teorías competitivas entre sí, cuando hay varias explicaciones posibles de las relaciones entre cierto número de variables explicativas.<sup>2</sup> Es típico que el análisis de regresión múltiple use una única variable dependiente y diversas variables independientes a fin de evaluar los datos estadísticos pertinentes. Por ejemplo, en un caso en el que se alega discriminación sexual en el salario, un análisis de regresión múltiple examinaría no sólo el sexo, sino también otras variables de interés, como la educación y la experiencia.<sup>3</sup> El empleador-defensor podrá usar el análisis de regresión múltiple para sostener que el salario es función de la educación y experiencia del empleado – y el acusador podrá sostener que el salario también depende del sexo del individuo.

---

<sup>1</sup> Ver Daniel L. Rubinfeld, Reference Guide on Multiple Regression, in Reference Manual on Scientific Evidence, 2nd ed., Federal Judicial Center (2000), pp. 179-227; Franklin M. Fisher, Multiple Regression in Legal Proceedings, 80 Colum. L. Rev. 702, 1980; American Bar Association Section of Antitrust Law Economics Committee, Selected Readings in Antitrust Economics: Applied Econometrics (July 2008) [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=12365&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=12365&Itemid=0) contiene ensayos de econometría aplicada a la literatura anti-monopolística; Jonathan B. Baker and Daniel L. Rubinfeld, Empirical Methods in Antitrust: Review and Critique, American Law and Economics Review, Fall 1999, pp. 386-435 [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13330&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=13330&Itemid=0); Allan J. Lichtman, Passing the test - Ecological Regression Analysis in the Los Angeles County Case and Beyond, Evaluation Review (ER), Vol.15, N° 6, Dec. 1991 [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13489&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=13489&Itemid=0); Daniel L. Rubinfeld, Econometrics in the Courtroom, Columbia Law Review, June 1985, pp. 1048-1097 [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13345&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=13345&Itemid=0); Alan O. Sykes, An Introduction to Regression Analysis, Chicago Working Paper in Law & Economics [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=14283&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=14283&Itemid=0); Daniel L. Rubinfeld and Franklin M. Fisher, U.S. v. Microsoft - An Economic Analysis, The Antitrust Bulletin, Spring 2001 [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=12384&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=12384&Itemid=0)

<sup>2</sup> El análisis de regresión múltiple sólo es un tipo de análisis en el que diversas variables están involucradas. Otros son: el análisis de estratificación, el análisis de la varianza, los análisis probit y logit, el análisis discriminante y el análisis factorial.

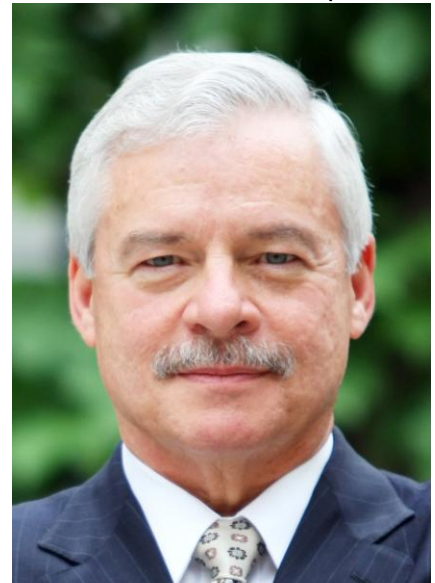
<sup>3</sup> En *Ottaviani v. State University of New York*, 875 F.2d 365, 367 (2d Cir. 1989) (citations omitted), cert. denied, 493 U.S. 1021 (1990) <http://ftp.resource.org/courts.gov/c/F3/196/196.F3d.358.98-7204.98-7202.98-7206.98-7208.98-7212.html>, la corte estableció que: “En casos dispares que involucran denuncias por discriminación de género, los demandantes usan típicamente el análisis de regresión múltiple para aislar la influencia del género sobre las decisiones de empleo relacionadas con cierta ocupación o salario de dicha ocupación. El primer paso del análisis de regresión consiste en especificar todos los factores “legítimos” (es decir, no discriminatorios) que probablemente afecten a la variable dependiente y que podrían explicar la disparidad de tratamiento de varones y mujeres. Al identificar estos criterios legítimos que afectan el proceso de decisión, los demandantes individuales pueden hacer predicciones acerca de qué trabajo o beneficios laborales recibirían empleados situados de forma similar, y entonces medir la diferencia entre el tratamiento predicho y el tratamiento real de esos empleados. Si existe una disparidad entre los resultados predichos y reales de las empleadas mujeres, las demandantes de un caso de trato dispar pueden argumentar que la diferencia neta “residual” representa el efecto ilegal de un intento discriminatorio de asignación del trabajo o de los beneficios laborales”.

El análisis de regresión múltiple es un instrumento que permite hacer estimaciones precisas y cuantitativas de los efectos de alguna variable interesante. Es una herramienta cuyos orígenes se remontan 200 años atrás al matemático Carl Friedrich Gauss, pero que se ha desarrollado sobremanera en los últimos 55 años. Este crecimiento se debió al desarrollo de las técnicas estadísticas modernas, a la creciente disponibilidad de datos “decentes”, y por encima de todo, al desarrollo de la computadora. Ha habido un uso creciente de la regresión múltiple y de técnicas vinculadas en conexión con procesos legales de distinta índole, si bien los jueces y los abogados han tenido una tendencia a ver a estos instrumentos con desconfianza general (a veces con justificación).

De interés especial para nuestro libro es el artículo escrito por los economistas Fisher y Rubinfeld, sobre el caso Microsoft.<sup>4</sup> En ese proceso anti-monopolístico, Fisher fue uno de los expertos por el gobierno, actuando en calidad de testigo, mientras que Rubinfeld actuó como economista jefe de la división anti-monopolios; su posición consistió en que: a) Microsoft alcanzó poder monopolístico en el mercado de sistemas operativos para computadoras personales compatibles con Intel; b) Microsoft anticipó la posibilidad de que su sistema operativo Windows fuera erosionado por los exploradores de Internet y por la plataforma de Java, ambos capaces de soportar aplicaciones de software independientes del sistema operativo; c) Microsoft adoptó acciones anti-competitivas (exitosas) para impedir la competencia en exploradores de Internet a fin de proteger la dominación de su sistema operativo Windows; d) también adoptó acciones anti-competitivas para restringir la disponibilidad de la plataforma tecnológica Java a fin de proteger su posición dominante de su sistema operativo Windows. Además, llevó adelante una campaña de avisos y solicitadas anti-competitivos con el objeto de convencer a otras empresas de que no compitieran con Microsoft en software de plataforma; e) Microsoft utilizó su poder de monopolio en el segmento de los sistemas operativos de las computadoras personales (PCs) para distorsionar la competencia en los exploradores de Internet; f) la conducta de Microsoft, que intentó mantener y aumentar las barreras a la entrada en el mercado de sistemas operativos de PCs, incluyó vincular a su explorador con el sistema operativo (lo que significaba que los vendedores debían adquirir el sistema del explorador de internet de Microsoft como condición necesaria para adquirir el sistema operativo Windows de Microsoft), obstaculizando así en forma severa la competencia de Netscape y limitando la amenaza de que los desarrolladores de software terminaran preparando otro explorador fuera del control de Microsoft, y no permitir a sus competidores utilizar los canales de distribución más eficientes. En resumen, se trató de una acción anti-monopolística que, de no remediarse, daría lugar a un entorno en el cual las



Franklin M. Fisher (1934- )



Richard L. Schmalensee

<sup>4</sup> Franklin M. Fisher and Daniel L. Rubinfeld, U.S. v. Microsoft - An Economic Analysis, The Antitrust Bulletin, Spring 2001. [http://www.law.berkeley.edu/faculty/rubinfeld/Profile/publications/FisherRubinfeld\\_USvsMicrosoft.pdf](http://www.law.berkeley.edu/faculty/rubinfeld/Profile/publications/FisherRubinfeld_USvsMicrosoft.pdf) Para detalles adicionales, ver [http://en.wikipedia.org/wiki/Microsoft\\_litigation](http://en.wikipedia.org/wiki/Microsoft_litigation)

plataformas que no usaran un estándar Microsoft no prosperarían, perdiéndose una ‘oportunidad crítica’ de innovación que reduzca o elimine el poder de Microsoft.

Microsoft se defendió con energía argumentando que todos sus intentos de “innovación” estaban siendo atacados por las empresas rivales celosas de su éxito, y que el litigio gubernamental era meramente un títere. Una publicidad a página plena en The Washington Post y en The New York Times del 2 de junio de 1999 del Independent Institute entregó una Carta Abierta al presidente Clinton de 240 economistas sobre el proteccionismo anti-monopolístico.<sup>5</sup> En una parte decía que *a los consumidores no les interesaban estas acciones anti-monopolísticas – sino a las empresas rivales. Los consumidores de alta tecnología habían disfrutado de precios en caída, producción en alza, y de una impresionante colección de nuevos productos e innovaciones... Pero, crecientemente, algunas empresas estaban buscando imponer trabas a sus rivales tratando que el gobierno las proteja. Muchos de estos casos están basados en daños que tendrían los consumidores en alguna fecha futura no especificada, y hay varias intervenciones gubernamentales que debilitarán a las empresas exitosas de Estados Unidos impidiendo que sean competitivas en otras partes del mundo.* El juez Jackson estableció los hechos el 5 de noviembre de 1999, afirmando que la dominación por Microsoft del mercado de sistemas operativos basados en x86 de las PCs constituía un monopolio, y que Microsoft había encarado acciones destinadas a aplastar toda amenaza a ese monopolio, incluyendo a Apple, Java, Netscape, Lotus Notes, Real Networks, Linux, y otros. El 3 de abril de 2000 emitió una sentencia en dos partes: sus conclusiones legales eran que Microsoft era un monopolio, y que Microsoft así lo había intentado, en violación de las secciones 1 y 2 de la Sherman Act, por cuyo motivo el remedio era que Microsoft debía ser separada en dos unidades independientes, una para producir el sistema operativo, y la otra para producir otros componentes de software.

El juicio también fue notorio porque tanto la procuraduría y la defensa utilizaron como testigos expertos a profesores del MIT a fin de preparar sus casos. Richard L. Schmalensee, un excelente economista y decano de la Escuela de Administración del MIT, testificó como experto a favor de Microsoft, mientras que por la otra parte, como hemos visto, Franklin Fisher, que había sido consejero de tesis doctoral de Schmalensee en el MIT, testificó a favor del Departamento de Justicia (en forma conjunta con Daniel L. Rubinfeld). Este fallo fue apelado, y por fin el DOJ anunció el 6 de setiembre de 2001 que ya no se buscaría la separación de Microsoft, y que, en su lugar, se buscaría una penalidad anti-monopolio de menor envergadura. El 2 de noviembre de 2001 el DOJ llegó a un acuerdo con Microsoft, que requería que Microsoft compartiera sus interfaces de programación aplicada con otras empresas y designara un panel de tres personas con pleno acceso a los sistemas, registros y códigos fuente de Microsoft por el término de 5 años a fin de asegurar su cumplimiento. Empero, el DOJ no le exigió a Microsoft cambiar ningún código ni le impidió que atara otro software con Windows en el futuro. El 5 de agosto de 2002 Microsoft anunció que haría ciertas concesiones en pro del acuerdo final, a la espera del veredicto del juez. El 1º de noviembre de 2002 el juez Kollar-Kotelly emitió una sentencia aceptando la mayor parte del acuerdo propuesto por el DOJ. Hubo 9 estados y el distrito de Columbia que no aceptaron el acuerdo, sosteniendo que no era lo suficientemente severo como para frenar las prácticas comerciales anti-competitivas de Microsoft. Sin embargo, el 30 de junio de 2004 la corte de apelaciones de US aprobó en forma unánime el acuerdo con el Departamento de Justicia, rechazando las objeciones. Los estados en disenso vieron al acuerdo como un mero tirón de orejas. El experto industrial Robert X. Cringely no cree que sea posible una separación, y que “ahora la única forma en que Microsoft podría morir es si se suicida”. Andrew Chin, un profesor de derecho anti-monopolístico de North Carolina en Chapel Hill, que asistió al juez Jackson en preparar el borrador de los hechos, escribió que el acuerdo le otorgaba a Microsoft “una

<sup>5</sup> [http://www.independent.org/pdf/open\\_letters/antitrust.pdf](http://www.independent.org/pdf/open_letters/antitrust.pdf)

inmunidad especial en sus licencias de Windows y otras 'plataformas de software' bajo términos contractuales que destruyen la libertad de competencia".

Las obligaciones de Microsoft dentro del acuerdo como fue escrito originalmente, expiraron el 12 de noviembre de 2007. Empero, más tarde Microsoft admitió una extensión adicional de dos años en los Juicios Finales referidos a las licencias de protocolos de comunicaciones, y que si los demandantes quisieran extender estos aspectos del acuerdo hasta 2012, no plantearía objeciones. Los demandantes dejaron en claro que la extensión pretendía servir sólo para otorgar a la parte relevante del acuerdo "la oportunidad de tener éxito durante el tiempo que cubría ese período", antes que dar lugar a un "patrón de violaciones intencionadas y sistemáticas". La crítica a las tácticas de Microsoft, descritas como *abrazar*, *extender* y *ahogar*, implican que Microsoft se apodera de un estándar o producto competitivo, a continuación lo extiende a fin de producir su propia versión que resulta incompatible con el estándar, lo que termina extinguiendo con el correr del tiempo la competencia que no usa o no puede usar la nueva versión de Microsoft. Son estas tácticas las que dieron lugar a juicios llevados adelante por empresas y gobiernos, y miles de millones de dólares en fallos en su contra. David Meyer apuntó que "Microsoft tiene una larga historia de pedir y que se le concedan patentes por inventos que, según mucha gente que a veces puede demostrarlo, estaban basados en ideas comunes y evidentes." Esto lo dijo en respuesta al pedido de 2008 de patentar una aplicación que permite subir o bajar un incremento de página oprimiendo una sola tecla – que ha sido un método incisivo durante décadas."<sup>6</sup>

*Críticas* Milton Friedman creía que el caso anti-monopolístico en contra de Microsoft establecía un peligroso precedente que presagiaba una creciente regulación gubernamental de una industria que otrora estuvo libre de intrusiones del gobierno y que, en consecuencia, el progreso tecnológico de la industria podría verse trabado.<sup>7</sup> Jean-Louis Gassée, CEO de Be Inc., afirmó que en realidad Microsoft no ganaba dinero con el Internet Explorer, y que su inclusión en el sistema operativo se debió a las expectativas de los consumidores de tener un navegador presentado dentro de su sistema operativo. Por ejemplo, BeOs viene presentado con su propio navegador, NetPositive, y Mac OS X con Safari. Argumentó que, en cambio, la verdadera influencia anti-competitiva residía en los descuentos que ofrecía a los fabricantes de equipo original impidiendo a otros sistemas operativos que pudieran afianzarse en el mercado.<sup>8</sup>

La regresión múltiple también puede ser útil (1) en determinar si está presente un efecto determinado; (2) en medir la magnitud de un efecto determinado; y (3) en proyectar cuál sería el efecto particular ante una intervención dada. En un caso de infracción de patentes, por ejemplo, un análisis de regresión múltiple podría ser utilizado para determinar (1) si la conducta del presunto infractor ha afectado el precio del producto patentado, (2) la cuantía del efecto, y (3) cuál hubiera sido el precio del producto si la supuesta infracción no hubiese ocurrido.

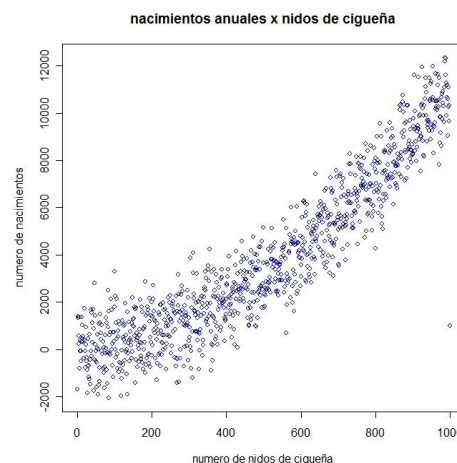
Durante las últimas décadas el uso del análisis de regresión múltiple en los tribunales de US aumentó en forma considerable. Aunque el análisis de regresión se usó con mayor frecuencia en casos de discriminación de sexo y raza y violación de defensa de la competencia, otros casos han sido el estudio de las subestimaciones censales, los derechos de voto, el estudio de la disuasión impuesta por la pena de muerte, la regulación de tarifas, y la propiedad intelectual. El análisis de

<sup>6</sup> David Meyer, Microsoft has been granted a patent on 'Page Up' and 'Page Down' keystrokes, Aug, 2008. <http://www.zdnet.com/news/microsoft-patents-page-up-and-page-down/218626>

<sup>7</sup> Milton Friedman, The Business Community's Suicidal Impulse, Cato Policy Report, March/April 1999, Vol. 21, No. 2. [http://www.cato.org/pubs/policy\\_report/v21n2/friedman.html](http://www.cato.org/pubs/policy_report/v21n2/friedman.html)

<sup>8</sup> Jean-Louis Gassée on why PC manufacturers don't sell non MS products, 26 Feb 1999. <http://lists.essential.org/info-policy-notes/msg00005.html>

regresión múltiple puede originar un valiosísimo testimonio científico en el ámbito judicial. Pero hay que tener cuidado, ya que también puede confundir cuestiones importantes que carezcan de valor probatorio. En el caso *EEOC v. Sears, Roebuck & Co.*,<sup>9</sup> en el que Sears era acusado de discriminación en contra de las mujeres en sus prácticas de contratación, el 7º Circuito reconoció que “los análisis de regresión múltiple, están diseñados a fin de determinar el efecto de diversas variables independientes sobre una variable dependiente, como sucede en un contrato, constituyen un método común y aceptado de tratar denuncias dispares”. Pero en este caso, lo que encontró el tribunal fue que “los análisis de regresión de E.E.O.C. no reflejaban el complejo proceso de toma de decisión no discriminatorio de Sears” y que sus análisis eran “tan defectuosos que no eran convincentes”.<sup>10</sup> También el análisis de regresión múltiple ha sido cuestionado en cuestiones censales y en casos de pena de muerte. Como se ha dicho antes, al interpretar los resultados de un análisis de regresión múltiple, es importante hacer una distinción entre correlación y causalidad. Cuando dos variables están correlacionadas, los eventos asociados a esas variables suceden más frecuentemente conjuntamente que lo que se puede esperar si la asociación fuera al azar. Pongamos que salarios más altos estén asociados con más años de experiencia, y salarios más bajos con menos años de experiencia. Pero si los salarios más altos están asociados con menos experiencia – y a la inversa, menores salarios lo están con una mayor experiencia – existirá una correlación *negativa* entre ambas variables.



Para decirlo una vez más, que dos variables guarden correlación entre sí no implica que una de ellas sea causa de la otra. Por consiguiente, al hacer una inferencia causal, es importante evitar las correlaciones *espurias*. Éstas surgen cuando hay dos variables estrechamente vinculadas pero no existe ninguna relación causal entre ambas, porque las dos están causadas por una tercera variable que no ha sido examinada. Un ejemplo popular es el de las estadísticas holandesas que muestran una correlación positiva entre el número de nidos de cigüeñas en distintas primaveras y la cantidad de niños que nacieron en ese momento. Por supuesto, no se trata de una conexión causal; sólo existe correlación entre ambas series porque estaban correlacionadas con la temperatura de los nueve meses previos a las observaciones.<sup>11</sup> Otro ejemplo: podría haber una correlación negativa entre la edad de algunos empleados cualificados de una empresa de computación y sus salarios. De esta correlación no debería extraerse la conclusión de que el empleador discriminó necesariamente en contra de sus empleados en base a su edad. Existe una tercera variable no examinada, que es la capacidad tecnológica de sus empleados, que podría estar explicando las diferencias de productividad y, por lo tanto, las diferencias de salarios.<sup>12</sup> O consideremos un caso de infracción a la ley de patentes en donde las ventas acrecentadas del producto que supuestamente se encuentra en infracción están asociadas a precios más bajos del producto patentado. Esta correlación sería espuria si ambos productos tuvieran nichos de

<sup>9</sup> <http://iilt.ilstu.edu/teeimer/court%20cases/eeocvs.htm>

<sup>10</sup> El tribunal del distrito comentó específicamente “los severos límites del análisis de regresión en evaluar los procesos de decisiones complejas”.

<sup>11</sup> Roger Sapsford, Victor Jupp, ed (2006). *Data Collection and Analysis*. Sage.

<sup>12</sup> Ver p.ej. *Sheehan v. Daily Racing Form Inc.*, 104 F.3d 940, 942 (7th Cir.) <http://caselaw.lp.findlaw.com/cgi-bin/getcase.pl?court=7th&navby=case&no=962123> (se rechazó la demanda porque la denuncia de discriminación usó un estudio estadístico que correlacionó la edad con la retención e ignoró la “remota posibilidad de que la edad esté correlacionada con una calificación legítima del cargo”), denegado, 521 U.S. 1104 (1997).

mercado no-competitivos y el menor precio se debiera a una baja del costo de producción del producto patentado. La falacia ha dado lugar a una conocida locución latina, *cum hoc ergo propter hoc* (“con esto, por lo tanto a causa de esto”).

Por otro lado, puntualizar la posibilidad de una correlación espuria no es suficiente para dejar de lado un argumento estadístico. No hay que prestar demasiada atención a un argumento que exhibe una correlación espuria cualitativa o cuantitativamente sustancial. Por ejemplo, puede ser necesario estimar una relación entre la aptitud tecnológica y la productividad laboral requerida en el ejemplo de discriminación por edad de más arriba.<sup>13</sup> Un punto importante: *la causalidad no sólo puede ser inferida a partir de un análisis de los datos; en realidad, se necesita una teoría que explique la relación entre ambas variables para inferir que existe una relación causal*. Recíprocamente, que dos variables estén correlacionadas entre sí no garantiza que exista una relación; podría ser que el modelo – que representa la teoría causal subyacente – no refleje en forma apropiada el juego conjunto de las variables explicativas. De hecho, recuerden que el hecho de que no exista correlación no garantiza que no haya una relación causal. La falta de correlación podría ocurrir si (1) hay datos insuficientes; (2) los datos están mal medidos; (3) los datos no permiten extraer múltiples relaciones causales; o (4) el modelo fue mal especificado porque se omitió una o más variables vinculadas con la variable de interés.

Existe tensión entre cualquier intento de llegar a conclusiones de gran nitidez y la naturaleza inherentemente probabilística de un análisis de regresión múltiple. En general, el análisis estadístico implica la expresión formal de la incertidumbre en términos de probabilidades. La realidad de que el análisis estadístico genera probabilidades que son en sí relaciones, no debería considerarse como un argumento en contra de la evidencia estadística. La única alternativa que quedaría a mano sería la menos confiable evidencia anecdótica.

En este capítulo, luego de revisar el estado del saber matemático-estadístico en el ámbito de los tribunales, introduciremos el problema estadístico del ajuste de funciones a observaciones dadas, y analizaremos ciertos temas de procedimiento y cuestiones metodológicas que tienen que ver con la admisibilidad y la ponderación que debería acordarse a los resultados obtenidos mediante los análisis de regresión múltiple. También se revisarán algunos estándares de información y análisis que un experto en esa materia debería seguir, y se pasará revista al diseño de investigación – esto es, de qué manera el análisis de regresión múltiple puede ser utilizado para elegir entre teorías alternativas de un caso dado. Luego nos concentraremos en la cuestión de cómo interpretar los resultados, tanto desde un punto de vista estadístico como práctico. Se discutirá brevemente la calificación de los expertos. Luego nos dedicaremos a aspectos vinculados con el uso de los datos. En el Apéndice, se entrará más en detalle en el problema de la regresión múltiple, y allí veremos una serie de ejemplos con aplicaciones de la técnica. Un libro de base: les recomiendo el estudio de los capítulos 1 a 8 de la obra de Robert S. Pindyck y Daniel L. Rubinfeld, *Econometric Models and Economic Forecasts*, McGraw-Hill International Edition, 4<sup>th</sup> edition, 1998. De esta obra vamos a desarrollar en clase algunos ejercicios ilustrativos.

---

<sup>13</sup> Ver, p.ej., *Allen v. Seidman*, 881 F.2d 375 (7th Cir. 1989) (Hubo escepticismo jurídico cuando el abogado defensor no incorporó una regresión logística que detallaba una variable omitida – tener o no un grado superior de educación especial; la estrategia del abogado defensor frente a las comparaciones estadísticas también debe incluir un análisis que demuestre que las comparaciones son engañosas). Los requisitos apropiados para que el abogado defensor exhiba una correlación espuria podrían depender, en general, del proceso de descubrimiento. P.ej., ver *Boykin v. Georgia Pac. Co.*, 706 F.2d 1384 (1983) <http://caselaw.lp.findlaw.com/scripts/getcase.pl?navby=search&case=/data2/circs/8th/942305p.html> (crítica del análisis del demandante por no incluir factores omitidos, cuando el demandante consideraba que toda la información de un formulario era inadecuada).

## 1. El analfabetismo matemático-estadístico en la justicia

Pero antes vamos a tratar un problema desarrollado por Meyerson.<sup>14</sup> Ustedes se habrán preguntado cómo es posible que un abogado de la Maestría tenga que entrar en vericuetos matemático-estadísticos, que son aparentemente parte de una disciplina completamente separada del derecho. Cabe preguntarse: ¿cuál es la causa de que un juez permita a un fiscal poner de relieve la raza de un acusado criminal si no hay ninguna indicación de que la raza tenga algo que ver con el caso? ¿Por qué motivo un tribunal podría pedirle a un jurado que considere a un sospechoso probablemente tan culpable como inocente, aún cuando no hay otra evidencia más que la de acusarlo? ¿Y qué puede inducirlo a un juez a que una mujer reciba una menor compensación por daños y perjuicios que un hombre idénticamente situado, bajo la expectativa tácita de que el sexismo continuará en el futuro?

Todo esto tiene que ver con un defecto muy grande del derecho, y es que aún no constituye una ciencia. No sorprende que muchos jueces sufran de distanciamiento, resistencia, e incapacidad en cuestiones matemáticas. Éstas son perversiones de la justicia, resultantes del analfabetismo matemático. Esto afecta a la mayoría de los abogados, como también afecta a la mayoría de los ciudadanos.<sup>15</sup> El analfabetismo matemático es especialmente preocupante porque el análisis numérico se ha convertido en una parte importante del sistema legal, en particular en la utilización de testimonios estadísticos en el juicio, “que ha crecido dramáticamente”.<sup>16</sup> En nuestros días la evidencia estadística es un elemento esencial que abarca a todo el universo legal. Las estadísticas son usadas regularmente para probar o rechazar cuestiones tan disímiles como los perjuicios causados en juicios por uso de tóxicos, violación de contratos, discriminación en el trabajo y de carácter político, identificación por medio de ADN en casos criminales o familiares, violaciones a las leyes de marcas y patentes, daño al medio ambiente, fraude financiero y pérdidas de ingresos futuros.<sup>17</sup> La Corte Suprema de US ha declarado que en tales casos, los tribunales federales cumplen con una “obligación general de ‘guardianes’”. Comenzando con su decisión de 1993 en el caso *Daubert v. Merrell Dow Pharmaceuticals, Inc.* el Tribunal Supremo declaró que las Reglas Federales de Evidencia asignan al juez del juzgado la tarea de asegurar que el testimonio de un experto cumpla tanto con fundamentos fiables como que sea relevante para la tarea a mano.

En lugar de ello, a los expertos no legales se les ha ido permitiendo ampliar su rol hasta el punto de que en la actualidad están adoptando decisiones políticas normativas. A veces esto ocurre porque los tribunales confunden validez científica con relevancia lógica; puede que no esté en disputa la ciencia y la matemática, pero la cuestión real es hasta qué punto hay una conexión lógica y legal de los números expuestos con el punto legal que se busca demostrar. Como observó el informe de la Comisión Carnegie sobre Ciencia y Tecnología en Decisiones Judiciales, “en los tribunales, el conocimiento científico debe informar sobre qué elección cabe hacer, pero abdicar a favor del hombre de ciencia es incompatible con la responsabilidad del juez de decidir

<sup>14</sup> Michael I. Meyerson, Significant Statistics: The Unwitting Policy Making of Mathematically Ignorant Judges, Pepperdine Law Review SSRN, 2010. [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13802&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=13802&Itemid=0)

<sup>15</sup> Así lo señala para el norteamericano promedio Marilyn Burns, *Math: Facing an American Phobia* (1998).

<sup>16</sup> Jeremy A. Blumenthal, Law and Social Science in the Twenty-First Century, Southern California Interdisciplinary Law Journal, Vol. 12; SSRN, December, 2004. [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13804&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=13804&Itemid=0)

<sup>17</sup> Jonathan J. Koehler, The Probity-Policy Distinction in the Statistical Evidence Debate, SSRN and Tulane Law Review, Vol. 66, p. 141, 1991. [http://www.ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13805&Itemid=0](http://www.ebour.com.ar/index.php?option=com_weblinks&task=view&id=13805&Itemid=0)

sobre la ley aplicable”.<sup>18</sup> El déficit judicial crítico no es, empero, respecto de las ciencias en general, sino de las matemáticas. Hay demasiados jueces que no “hablan en términos matemáticos” ni entienden lo que esos números comunican. También fallan en apreciar el significado de que los resultados matemáticos no son una necesidad científica, sino consecuencia de juicios de valor específicos. Al ignorar que estos juicios son inherentes a elecciones matemáticas, los jueces terminan aceptando valores que pueden hallarse en las antípodas de un sistema judicial.

Naturalmente, siempre se producen errores. No existe una “máquina de la verdad” que diga con precisión férrea el resultado correcto de una prueba médica, de una predicción económica, o de un juicio. De lo que se trata es de reducir la frecuencia y el grado de los resultados incorrectos, pero la condición humana es inevitablemente imperfecta. Los estadísticos, que se manejan con el “arte de lo probable”, han imaginado un ingenioso sistema para conducirse con este error inevitable.

Supongan, p.ej. que hay una prueba médica para determinar si los pacientes tienen una enfermedad determinada, y que, en general, un puntaje más alto está correlacionado con una mayor probabilidad de tener esa enfermedad. Los pacientes están dentro de un rango de esta prueba, y se requiere un punto de corte de manera de poder decir que, por encima de ese puntaje, el paciente será diagnosticado como enfermo. Hay dos situaciones posibles donde la prueba podría estar errada. En primer término, con un “falso positivo”, habrá pacientes sanos que serán diagnosticados como enfermos. Alternativamente, con un “falso negativo”, habrá pacientes enfermos que erróneamente serán considerados sanos. Como vimos, a estos errores los llamamos de “Tipo I” y de “Tipo II” respectivamente.

Luego, el punto de corte habrá que elegirlo basándose en qué tipo de error es el peor. Si subimos el punto de corte, tendremos más falsos negativos (más pacientes afectados declarados sanos) pero menos falsos positivos (menos pacientes sanos declarados enfermos). Si se reduce el punto de corte se tendrá el efecto inverso (con menos falsos negativos pero falsos positivos adicionales). De cualquier forma que ustedes elijan cometerán errores; no existe un “punto perfecto” para elegir. Como ambos errores siempre tendrán lugar<sup>19</sup> el punto de corte para determinar la presencia de una enfermedad reflejará el juicio de valor sobre qué tipo de error es más serio. Puede que prefiramos tener menos falsos positivos – un error de Tipo I más bajo – para emplear una prueba de medicación, así reducimos el número de empleados erróneamente acusados. Para enfermedades con graves consecuencias que podrían evitarse sólo si hay una acción inmediata (p.ej. si se trata de un cambio de la dieta que podría evitar un retraso mental durante el desarrollo del feto) podríamos desear tener un menor número de falsos negativos, un error de Tipo II más bajo, a fin de minimizar la posibilidad de que alguien enfermo no sea diagnosticado. La elección de un estándar de prueba legal refleja un cálculo similar. Como sucede con una prueba médica inevitablemente imperfecta, siempre existe la posibilidad de que el veredicto de un juicio no sea adecuado a los hechos reales. En un contexto criminal, si se declara culpable a un inocente, hemos cometido un error de Tipo I. Si exoneramos a un culpable, tendremos un error de Tipo II. Lo mismo sucede en el contexto civil. Y bien sabemos que serán cometidos errores.

¿Cuánto riesgo estamos dispuestos a aceptar basándonos en los valores del sistema legal? A diferencia de una investigación científica, “el derecho está orientado a la resolución ‘justa’ de

---

<sup>18</sup> Carnegie Commission on Science, Technology and Government, *Science and Technology in Judicial Decision Making: Creating Opportunities and Meeting Challenges*, 24 (1993). <http://www.ccstg.org/>

<sup>19</sup> Sólo se podrá evitar cometer estos errores si en el 100% de las pruebas fracasan o en el 100% pasan.



casos en lugar de la búsqueda de la verdad”.<sup>20</sup> Pero mientras que admitir la “verdad” es un valor profundo de todo sistema legal, a menudo no se alcanza porque los veredictos deben ser emitidos en condiciones de información incompleta. Por ejemplo, en muchos casos los tribunales han hecho recaer el peso de la prueba de causalidad sobre un demandante inocente ante un demandado negligente, al requerir que el demandante pruebe la causa, lo que “sería injusto y destructivo de los propósitos disuasivos involucrados en el concepto de deber ser diligente”.

A guisa de conclusiones de esta sección, cabe recordar que hace más de 40 años el Tribunal Supremo de US advirtió lo siguiente: “La matemática, un verdadero brujo de nuestra sociedad computadorizada, al tiempo que es de ayuda para el juez o el jurado responsable de decidir cuáles son las cuestiones de hecho en el juicio en la búsqueda de la verdad, no debe proyectar un hechizo sobre ellos”.<sup>21</sup> En nombre del rigor matemático incorrectamente aplicado, los tribunales de US permitieron que cuestiones raciales entren en los juicios criminales, y que los prejuicios reduzcan los subsidios por los delitos cometidos. Si se la acepta judicialmente, información que se reconoce universalmente como incorrecta y carente de base puede ser utilizada para calcular, p.ej., las probabilidades de paternidad. Y algunos malhechores lograron rehuir a sus obligaciones con jueces que renunciaron a su responsabilidad de balancear riesgos de error que forman parte inevitable de todo juicio. Al fin de cuentas, las matemáticas mal usadas no son meramente un hechicero sino un matón que se apropia del poder de emitir juicios de valor que pertenece a los tribunales. No es necesario que jueces y abogados se vuelvan “aficionados a las matemáticas” para que reclamen desempeñar un papel apropiado.<sup>22</sup> Pero deben ser conscientes de que la aparente objetividad de las matemáticas a veces enmascara un juicio subjetivo, y no deben dejarse llevar cuando los números “duros” están basados en poco más que intuición y conjeturas.<sup>23</sup> *Los números pueden comunicar información importante. Es necesario que sepamos comprender lo que los números tratan de decir.*

## 2. Ajuste de curvas

Los datos, provenientes de medir variables, pueden venir dados por una gran cantidad de fuentes y presentados de distintas maneras. Los datos que describen el movimiento de una variable a través del tiempo son llamados datos de *series de tiempo*, y pueden ser diarios, semanales, mensuales, trimestrales, o anuales. Los datos que describen las actividades de personas individuales, empresas, u otras unidades en cierto momento del tiempo son llamados datos de

<sup>20</sup> Lucinda Finley, Guarding the Gate to the Courthouse: How Trial Judges Are Using Their Evidentiary Screening Role to Remake Tort Causation Rules, SSRN; DePaul Law Review, Vol. 49, P. 335, 1999 (afirma que “el sistema de agravios no debería atarse a los valores de la epidemiología porque sus propósitos y funciones sociales siempre tuvieron en cuenta un rol ‘justiciero’ de alcance más amplio que cuando los científicos llegan a una conclusión, o lo que sucede con el consenso verdadero ‘científico’ en un momento”).

<sup>21</sup> V. People v. Collins, 438 P.2d 33, 33 (1968).  
<http://www.law.berkeley.edu/faculty/sklansky/evidence/evidence/cases/Cases%20for%20TOA/People%20v.%20Collins.htm>

<sup>22</sup> En Daubert, el por entonces Jefe de Justicia Rehnquist se quejó de que la necesidad de que los jueces “evalúen la validez científica” les imponía la obligación de convertirse en “científicos aficionados”. Daubert, 509 U.S. at 600-01 (Rehnquist, C.J., concurring in part and dissenting in part). Ver la nota (183) de Ann C. Hodges and Porcher L. Taylor, III, The Business Fallout from the rapid Obsolescence and planned Obsolescence of High-Tech Products: Downsizing of Noncompetition Agreements.  
<http://www.stlr.org/html/volume6/hodges.txt>

<sup>23</sup> Recomiendo nuevamente la lectura de David H. Kaye, Is Proof of Statistical Significance Relevant?, SSRN y Washington Law Review, Vol. 61, 1986.  
[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1411859](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1411859)

*sección cruzada o transversal*. Si los datos combinan una dimensión temporal con otra transversal, son llamados de *panel*.

Es probable que un estudio de mercado realizado sobre presupuestos de las familias en un momento del tiempo utilice datos de sección cruzada. Datos de sección cruzada también son usados para examinar un grupo de estados contables de un grupo comercial a fin de estimar el patrón de conducta entre firmas individuales de la industria. Supongan que estoy interesado en la relación entre dos variables  $X$  e  $Y$ . A efectos de describirla estadísticamente, necesito un conjunto de observaciones para cada variable y una hipótesis que plantea una forma matemática explícita de esa relación. El conjunto de observaciones es llamado una *muestra*.<sup>24</sup> En primer término, analizaremos el caso en que la relación entre  $X$  e  $Y$  es lineal, o sea que viene representada por una línea recta. Dado el carácter lineal, mi objetivo es especificar una regla mediante la cual pueda determinar la “mejor” relación lineal que vincula a  $X$  con  $Y$ .

Edad (Años)	Diámetro
97	12.5
93	12.5
88	8
81	9.5
75	16.5
57	11
52	10.5
45	9
28	6
15	1.5
12	1
11	1

Los pares de observaciones serán indicados como  $\{y_i, x_i\}_{i=1}^n$ . Estos datos están constituidos por una respuesta *escalar*  $y$  y un conjunto de *variables explicativas*  $x$ , que se reducen en nuestro caso a una única variable explicativa. El modelo *lineal* será escrito:

$$[1] \quad y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

el cual representa un modelo estadístico si agregamos los supuestos siguientes: (a) los coeficientes  $\beta_0$  y  $\beta_1$  son constantes que el análisis debe determinar (denominadas *coeficientes de regresión*); (b) la variable independiente  $x_i$  es una variable observada sin error (hay modelos más complejos que levantan este supuesto); (c) se debe incorporar un *término aleatorio* no observable  $\varepsilon_i$  que constituye una variable aleatoria cuya esperanza matemática es nula.<sup>25</sup>

Vayamos a un ejemplo. En la tabla adjunta están los datos de diámetro (medido en pulgadas) de árboles al S.E. de Wisconsin y de edad de esos árboles (en años). Ustedes saben probablemente que la edad de un árbol puede computarse contando el número de aros anuales concéntricos en el tronco del árbol. (No es necesario convertir en un muñón al árbol para efectuar este cálculo, ya que se puede sacar una muestra que vaya de la corteza hasta el núcleo muerto usando un taladro). Hay una medida usual de tamaño, conocida como “diámetro a la altura del pecho” (DAP). “Altura del pecho” se define como 4 veces y media por arriba del lado cuesta arriba del árbol. Como un tronco no es un círculo perfecto, el “diámetro” es definido como la longitud de la circunferencia dividida por  $\pi$ . Nuestros datos de edad y diámetro son los de la tabla anterior.

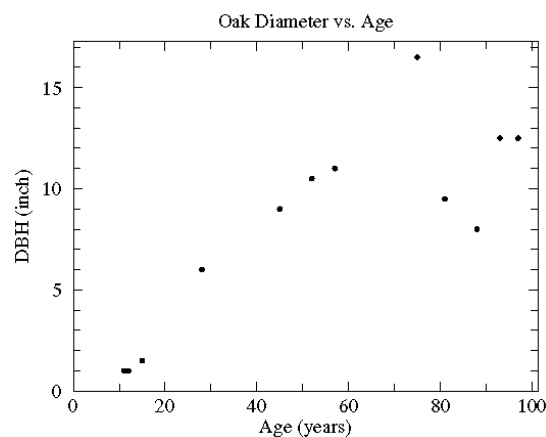


Figura 1

<sup>24</sup> Los datos muestrales son observaciones elegidas de una *población* subyacente que representa la verdadera relación estudiada.

<sup>25</sup> Luego, si  $E(\cdot)$  es la esperanza matemática de la expresión,  $E(y_i | x_i) = \beta_0 + \beta_1 x_i$ . Esta expresión se lee “esperanza matemática de  $y_i$  dada  $x_i$ ”.

La decisión siguiente es graficar estos datos en un diagrama de dispersión, en el cual debemos decidir cuál será la variable  $X$  (independiente) y la variable  $Y$  (dependiente). Una regla sugiere que conviene colocar a la variable medida con menor precisión en el eje de las  $Y$ . Cuantificar la cantidad de anillos anuales debería dar un resultado preciso; naturalmente, los árboles serán un poco más viejos que la cuenta porque les llevó algunos años crecer hasta la altura en que fue tomado el núcleo. El DAP podría medirse con bastante precisión. Los errores de edad (una subestimación sistemática de algunos años) y de DAP (error de redondeo, si lo hay) son difíciles de comparar ya que están expresados en unidades diferentes y tienen distinta naturaleza. Otra regla sería: colocar la variable de “control” en el eje de las  $X$  y la variable dependiente en el eje de las  $Y$ . Si pienso que la edad es la causa del crecimiento, más que a la inversa, pondré la edad en el eje de las  $X$ . Por otra parte, si el objetivo es terminar con una fórmula que permita predecir la edad basándose en el DAP, entonces pondré la edad en el eje de las  $Y$ . La elección que haga no es sólo una cuestión de gráfico: habrá distintas “líneas de tendencia” haciendo uso de distintas opciones. Pero el coeficiente de correlación  $r$  y el  $p$ -valor asociado no dependerán de la elección que se haga. La figura 1 muestra la elección que hice. Está claro que hay una tendencia a que los robles más viejos sean también los más grandes. Por otro lado, hay bastante variabilidad, por ejemplo el roble más grande *no* es el más viejo. Queda claro que la relación entre edad y DAP no viene dada mediante una mera “conexión entre los puntos” (Anótenlo porque es básico: ¡ está mal obtener un gráfico “conectando a los puntos entre sí”!).

La relación entre la edad y el DAP podría dar lugar a una recta que yerra a los datos individuales (algunos quedan arriba y otros abajo), pero en cambio se obtiene una especie de promedio entre los puntos como en la Figura 2. Hay diversos programas de computación que permiten estimar la recta de mejor ajuste, por ejemplo el EVIEWS, del cual transcribo una salida simplificada:

Tabla 1. Resultados de regresión

Dependent Variable: DIAMETRO				
Method: Least Squares				
Included observations: 12				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.285354	1.702259	0.755087	0.4676
EDAD	0.127792	0.027131	4.710220	0.0008
R-squared	0.689308	Mean dependent var		8.250000
Adjusted R-squared	0.658238	S.D. dependent var		4.997727
p value of F	0.000829	Durbin-Watson stat		1.1787630

La ecuación de regresión de la Tabla 1 se escribe (redondeando al decimal más próximo):

$$[2] \quad DIAMETRO = 1.28 + 0.13.EDAD$$

Esta información es rutinariamente provista por los programas de cómputo: el coeficiente  $\beta_0 = 1,28$  (la “ordenada al origen”) y el coeficiente  $\beta_1 = 0,13$  (la “pendiente de la línea” o “coeficiente angular”). Hay que prestar atención, además, a los errores estándar de ambos coeficientes:  $\sigma_{\beta_0} = 1,70$ , y  $\sigma_{\beta_1} = 0,03$ .<sup>26</sup> Los estadísticos  $t$  ensayan si el verdadero valor de cada coeficiente podría ser 0. Son calculados como el cociente entre el estimador de  $\beta_i$  y el correspondiente estimador de su error estándar  $\sigma_i$ . Un valor elevado de  $t$  indica que la hipótesis debe rechazarse y que el coeficiente correspondiente  $\neq 0$ . La última columna expresa el  $p$ -valor. Por convención, un  $p$ -valor inferior a 0,05 se interpreta como evidencia de que el correspondiente valor poblacional es distinto de cero. El estadístico  $t$  de EDAD es muy elevado para los estándares usuales: 4,71 que tiene un  $p$ -valor de prácticamente cero. Lo cual indica que al ser tan reducido (al nivel de aproximación decimal usado) es muy improbable que la relación aparente entre la edad del árbol y el diámetro del mismo sea un mero producto del azar. El coeficiente de correlación es  $r = 0,83$ . Este coeficiente de correlación elevado al cuadrado,  $R^2 = 0,69$ , es llamado el *coeficiente de determinación* o *de explicación* (o de Pearson). Cuando a este coeficiente se lo “penaliza” por la incorporación de  $p$  variables explicativas, se obtiene el coeficiente  $R^2_{aj} = 0,66$  (en cuyo caso hablamos del coeficiente de determinación ajustado).<sup>27</sup>

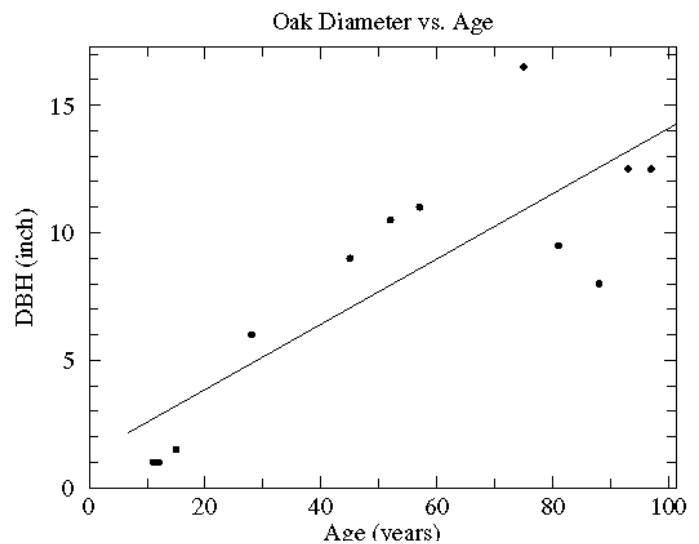


Figura 2

Hay dos estadísticos descriptivos muy simples: uno es la *media de la variable dependiente* y otro es el *error estándar de la variable*. Nos quedan finalmente dos: el  $p$ -valor del estadístico F y el contraste de Durbin y Watson. Un F-test es un estadístico que tiene una distribución F bajo la hipótesis nula (es otro tipo de distribución distinta de la normal). Es utilizado para identificar al modelo que mejor ajusta los datos muestrales de una población. El término fue inventado por George W. Snedecor, en honor a Sir Ronald A. Fisher, quien fue el que desarrolló en los 1920s a este estadístico como una relación entre dos tipos de varianzas. Este estadístico busca contrastar la hipótesis de que todos los coeficientes (excepto la ordenada al origen, con lo cual estamos probando a  $\beta_1$ ) son iguales a cero. Este estadístico tiene que calcularse teniendo en



Ronald A. Fisher (1890-1962)

<sup>26</sup> Con estos errores estándar se pueden obtener intervalos de confianza de los coeficientes de regresión. Por ejemplo, para  $\beta_0$  un intervalo de confianza al 95% vendrá dado por  $1,29 \pm t_c \cdot 0,76$  donde  $t_c$  es un cierto valor tabulado de la distribución  $t$ -Student, que resulta igual a 2,12 para una dódima bilateral al 95% (Ver Apéndice, tablas estadísticas) con 11 grados de libertad. Esto resulta en un intervalo de confianza muy impreciso, comprendido entre -0,32 y 2,90. En cambio,  $\beta_1$  tiene una estimación muy precisa ( $p = 0,0008$ ).

<sup>27</sup> Este estadístico siempre es inferior a  $R^2$ , puede disminuir al agregarse nuevas variables independientes, y hasta puede llegar a ser negativo para ecuaciones con un ajuste muy pobre:  $R^2_{aj} = 1 - [(n-1)/(n-p)] (1 - R^2)$ .

cuenta el número de observaciones ( $n$ ) y de variables explicativas<sup>28</sup> en número igual a  $p$ . Su  $p$ -valor indica la probabilidad de que la hipótesis sea corroborada; su reducido valor indica que es muy improbable que la relación aparente entre edad y DAP sea resultado del mero azar.

A pesar de su nombre, modelo clásico de regresión lineal, éste no se limita a relaciones lineales entre la variable dependiente y (una o más) variables independientes. En realidad es *potencialmente no lineal* en las variables independientes, pero *siempre es lineal en los parámetros  $\beta$*  y el término aleatorio o de error entra en forma aditiva y separable de los restantes.

Por ejemplo, la figura 3 representa funciones logarítmicas expresadas en bases distintas. Recuerden que el logaritmo de un número en una base determinada es el exponente al cual hay que elevar la base para obtener el número. Es la función matemática inversa de la función exponencial. Luego,

$$[3] \quad \log_b N = x \text{ siempre que } N = b^x.$$

No me detendré más en distintas formas funcionales que pueden aparecer en el análisis. Sugiero que lean el artículo de Schmidheiny a tal efecto.<sup>29</sup>

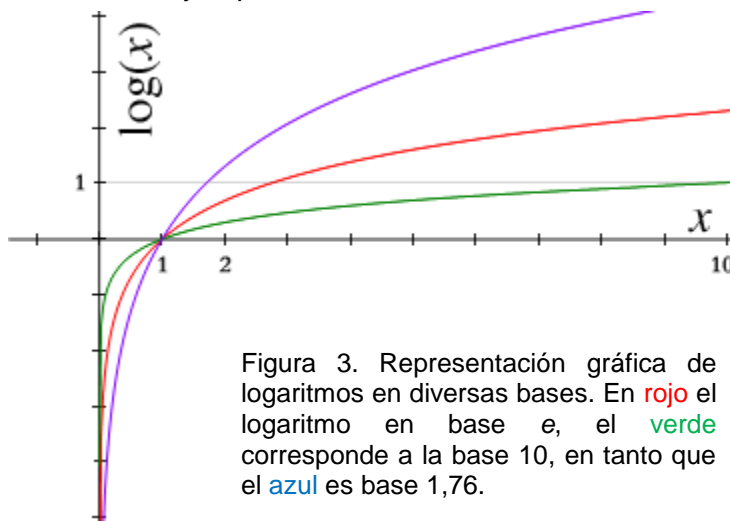


Figura 3. Representación gráfica de logaritmos en diversas bases. En rojo el logaritmo en base  $e$ , el verde corresponde a la base 10, en tanto que el azul es base 1,76.

A veces se dice que  $R^2$  es igual a la fracción de la varianza de  $Y$  explicada por su relación con  $X$ . Lo que esto significa es que el error estándar de los *residuos* de la línea de tendencia de la Figura 2 (las líneas de color azul de la Figura 4) dividido por el error estándar de  $Y$  es igual a  $1-R^2$ . Obviamente, si  $R$  está próximo a  $+1$  o  $-1$ , los desvíos con relación a la línea de tendencia deben ser “pequeños”. Pero tomen en cuenta que muchas veces se presentan situaciones de correlaciones “estadísticamente significativas” (o sea, reducidos  $p$ -valores) que explican una reducida variación de los datos (o sea, con bajos coeficientes  $R^2_{aj}$ ).

El nombre de “mínimos cuadrados” proviene del proceso de definir una línea de tendencia. Se ajusta una línea hasta que la suma de los cuadrados de los desvíos de las  $Y$  respecto de la línea (en color azul en la Figura 4) sea la más reducida posible. Pero noten que hay otras formas de obtener líneas de tendencia.

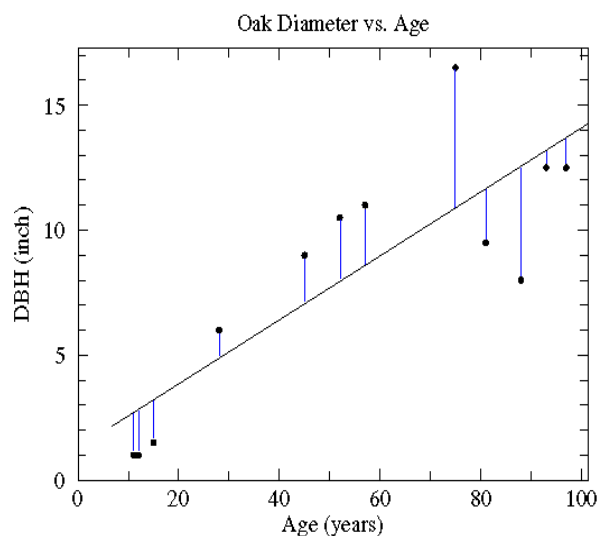


Figura 4

<sup>28</sup> Más precisamente, este estadístico tiene una distribución  $F(p-1, n-p)$  bajo la hipótesis nula y el supuesto de normalidad.

<sup>29</sup> Kurt Schmidheiny, Functional Form in the Linear Model, Universitat Pompeu Fabra, Fall 2009 <http://kurt.schmidheiny.name/teaching/functionalfom2up.pdf>

En la Figura 5. la línea sólida cumple con un buen desempeño pero deja 4 puntos bien lejos de la línea, y se basa en minimizar la distancia de los desvíos *horizontales* a partir de esa línea (graficados en color rojo). Tiene una mayor pendiente que la línea de mínimos cuadrados de la Figura 4. La curva punteada, que yerra apenas 3 puntos, es una parábola elegida para minimizar el cuadrado de los desvíos de las Y. No hay una sola manera de elegir la “mejor” línea de tendencia entre todas las líneas de tendencia. Ustedes se pueden guiar a) por lo que sugiere una teoría conocida, b) por lo que requiere un docente en particular, c) por la práctica usual (que es habitualmente una línea de mínimos cuadrados), d) conociendo qué puntos es probable que sean anómalos, o d) (desafortunadamente) por la necesidad de conseguir un resultado determinado.

Las líneas de tendencia son utilizadas a menudo para “guiar el ojo”, desplegando una tendencia promedio. También pueden servir para realizar predicciones cuantitativas. Por ejemplo, “¿cuán grande será mi roble dentro de 20 años?” o “¿cuán viejo puede ser este árbol que tiene 10 pulgadas de diámetro?”. Siempre es más seguro hacer predicciones dentro del rango de datos que definen la línea de tendencia (lo cual sería básicamente una *interpolación*). Cuando son hechas fuera del rango probado (*extrapolación*) las líneas de tendencia pueden dar lugar a respuestas erróneas, incluso disparatadas. Por ejemplo, la línea sólida de la Figura 5 (con desvíos horizontales minimizados) sugiere que un roble más joven que 1 año tiene un diámetro negativo. La línea de tendencia parabólica sugiere que los árboles comienzan a contraerse a partir de los 80 años y que tendrán un diámetro negativo si tienen más de 150 años.

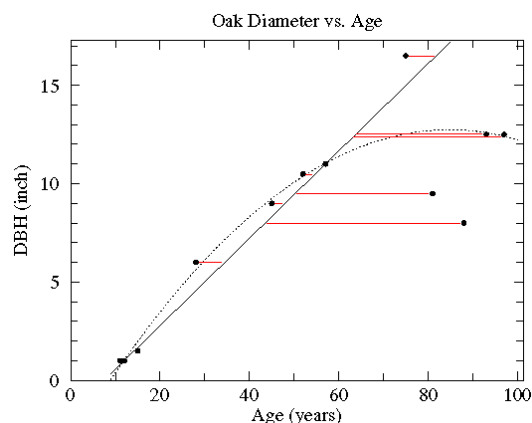


Figura 5

Otra idea básica fundamental del modelo de mínimos cuadrados es que toda la estructura está incorporada en la expresión  $\beta_0 + \beta_1 x_i$  y que lo que queda es comportamiento puramente aleatorio (esta hipótesis es levantada en modelos más sofisticados). En particular, no debe ser posible poder predecir el residuo de un año conociendo los residuos de años anteriores. ¿Cómo diagnosticar si los residuos tienen esta propiedad? En la Figura 6 se incluyen los residuos de la ecuación de regresión de la Figura 2. Obsérvese que en la zona central existe una tendencia a que residuos positivos sigan a residuos positivos, lo que no ocurre en las dos zonas de los extremos. Lo que esto significa es que, en la zona central p.ej., hay una similitud entre observaciones como función del tiempo que las separa. A esta propiedad se la denomina *auto-correlación*. La idea de auto-correlación de una serie de tiempo es que se puede encontrar un patrón repetitivo, como ser la presencia de una señal periódica que, de existir, violaría el supuesto del modelo de mínimos cuadrados. Pero los econométricos están preparados para enfrentar este tipo de situaciones. Antes que nada, distinguen entre auto-correlación positiva (algo así como la que se verifica en la región central y en ambos extremos de la Figura 6) de los casos de auto-correlación negativa, que se presenta cuando es posible predecir a partir de un residuo *positivo* que el residuo siguiente será *negativo*, y que uno negativo le seguirá otro positivo. Para ello, se construye un estadístico que en la Tabla 1 es llamado de Durbin-Watson (ya que fue descubierto por James Durbin y Geoffrey Watson), que es igual aproximadamente a  $d=2.(1-r)$  en

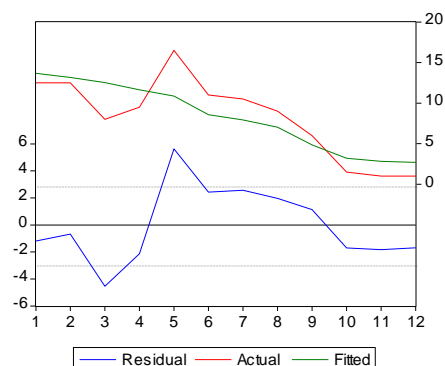


Figura 6

cuya fórmula  $r$  es el *coeficiente de auto-correlación muestral de los residuos* (es decir, que si  $r > 0$  tendremos una tendencia a que residuos positivos sean seguidos por residuos positivos, y residuos negativos por residuos negativos; mientras que si  $r < 0$  la tendencia será a que residuos con un signo sean seguidos por residuos de otro signo). Este estadístico  $d$  siempre está comprendido entre 0 (caso  $r = +1$ ) y 4 (caso  $r = -1$ ). Luego, valores reducidos de  $d$  están indicando que es probable que exista auto-correlación positiva, y valores elevados de  $d$  que sea negativa. Para *probar la auto-correlación positiva* al nivel de significación del 95%, el valor  $d$  debe ser comparado con dos valores críticos, uno inferior  $d_L$  y otro superior  $d_U$ ; se compara  $d$  con estos dos valores tabulados (ver Apéndice) de tal forma que, si  $d < d_L$  existe evidencia estadística de que los errores están positivamente auto-correlacionados; si  $d > d_U$  la evidencia estadística indica que los errores *no* están positivamente auto-correlacionados; pero si  $d_L < d < d_U$ , el contraste es inconclusivo.<sup>30</sup> En nuestro ejemplo, el estadístico  $d = 1,18$ . Con 12 observaciones y 2 variables (incluyendo la ordenada al origen como otra variable), al 95% de confianza, el estadístico queda comprendido entre los dos valores límites que proporciona la tabla (0.97076 y 1.33137) de lo cual obtenemos que el contraste es inconclusivo.

### 3. Especificación del Modelo

Volviendo al derecho, el análisis de regresión múltiple del economista que brinda testimonio le permite elegir entre teorías o hipótesis alternativas y lo ayuda a distinguir aquellas correlaciones que son claramente espurias de las que reflejan relaciones válidas. En este punto conviene tener un diseño de investigación bien armado.

*Cuestión específica investigada* Una investigación comienza formulando una pregunta clara. Los datos a ser recolectados deben estar directamente vinculados con la cuestión, ya que en caso contrario no podrán sacarse las inferencias adecuadas a partir del análisis estadístico. Por ejemplo, si la cuestión que se presenta en un caso de infracción a una patente es qué precio habría alcanzado el producto del demandante de no ser por la venta del producto en infracción, debe haber suficientes datos como para permitir que el experto tome nota en forma estadística de los factores que determinaron el precio del producto.

*Modelo a elegir* La especificación de un modelo involucra distintas etapas, cada una fundamental para que tenga éxito el esfuerzo de investigación. Idealmente, un análisis de regresión múltiple se elabora en base a una teoría que describe las variables incluir en el estudio. Por ejemplo, la teoría de los mercados laborales puede conducirnos a esperar que los salarios de una industria estén vinculados con la experiencia y la productividad de los trabajadores. Si se cree que hay discriminación laboral deberían agregarse una o más variables que reflejen semejante discriminación. Usualmente a los modelos se los caracteriza mediante parámetros, es decir características numéricas del modelo. En el ejemplo laboral, un parámetro podría reflejar el incremento salarial asociado a cada año de experiencia laboral. La regresión múltiple utiliza una muestra, o selección de datos, de la población para obtener estimadores de los valores de los parámetros del modelo. Un estimador asociado con una variable explicativa determinada es el coeficiente de regresión estimado. ¿Por qué es importante disponer de una teoría apropiada, elegir bien las variables apropiadas, o elegir bien la forma correcta del modelo? Porque si uno no

<sup>30</sup> La prueba de auto-correlación negativa sigue un enfoque similar, pero en lugar de  $d$  se usa  $(4-d)$ . Si  $(4-d) < d_L$ , la evidencia indica que los errores están negativamente auto-correlacionados, mientras que si  $(4-d) > d_U$  esa evidencia está indicando que no están negativamente auto-correlacionados. El test es inconclusivo si  $(4-d)$  está comprendido entre ambos valores. Cabe tener en cuenta que si la distribución de los errores no sigue una distribución normal, o hay otras alteraciones del modelo canónico [1], los econométricos usan en forma complementaria el test llamado de Breusch-Godfrey.

hace todo esto, los resultados estadísticos pueden estar *sesgados*, lo que significa que se genera una tendencia sistemática a que el parámetro del modelo sea demasiado alto o bajo.

*Elección de la variable dependiente* La variable elegida como “dependiente” debe ser la que corresponde.<sup>31</sup> Por ejemplo, supongan que existe una discriminación de pagos a trabajadores por hora. Una elección de variable dependiente es la tasa salarial por hora de los empleados; otra es el salario anual. La diferencia es importante, porque las diferencias de salarios anuales pueden deberse en parte a diferencias en horas trabajadas. Si la cantidad de horas trabajadas resulta de las preferencias del trabajador y no de la discriminación, el salario por hora es una elección correcta. Si la cantidad de horas está vinculada a la discriminación alegada, en ese caso será más adecuado elegir al salario anual.<sup>32</sup>

*Variable explicativa relevante* Aquí debe elegirse en forma adecuada a la variable explicativa que permita la evaluación de hipótesis alternativas. Por ejemplo, en un caso de discriminación, la variable de interés podría ser el sexo o el origen geográfico del individuo. En un caso anti-monopolístico, puede tratarse de una variable que tome valor 1 si refleja la presencia de la conducta anti-competitiva alegada y 0 en los demás casos.

*Variabes explicativas adicionales* Debería intentarse identificar variables explicativas adicionales, conocidas o hipotéticas, algunas medibles y que puedan sustentar hipótesis sustanciales tomadas en cuenta por el análisis de regresión. Así, en un caso de discriminación, una medida de la aptitud de los trabajadores puede brindar una explicación alternativa – los salarios más reducidos pueden deberse a capacidades inadecuadas.<sup>33</sup>

<sup>31</sup> En el análisis de regresión múltiple, usualmente la variable dependiente es una variable continua que adopta valores dentro de un cierto rango. Si la variable es *categorica*, es decir que adopta sólo dos o tres valores, hay formas modificadas del análisis de regresión múltiple que son apropiadas (como los análisis probit y tobit). Para ver un ejemplo de uso de este último, ver EEOC v. Sears, Roebuck & Co., 839 F.2d 302, 325 (7th Cir.1988) (EEOC utilizó el análisis logit para medir el impacto de variables tales como la edad, la educación, la experiencia en el trabajo y en las líneas de producción, sobre el porcentaje de comisiones cobradas por las mujeres). <http://openjurist.org/839/f2d/302>

<sup>32</sup> En los sistemas laborales con salarios anuales atados al grado del escalafón, podría resultar más apropiado utilizar el salario anual correspondiente a la posición laboral.

<sup>33</sup> En *Ottaviani v. State University of New York*, 679 F. Supp. 288, 306–08 (S.D.N.Y. 1988), *aff'd*, 875 F.2d 365 (2d Cir. 1989), cert. denied, 493 U.S. 1021 (1990), [http://www.dhr.state.ny.us/pdf/Commissioner's%20Orders/rossi\\_v\\_iona\\_college\\_2007.pdf](http://www.dhr.state.ny.us/pdf/Commissioner's%20Orders/rossi_v_iona_college_2007.pdf) el tribunal dictaminó en la fase de responsabilidad civil, que no existió discriminación ni en la ubicación en rangos ni en la promoción entre rangos, de modo que el rango era la variable adecuada en un análisis de regresión múltiple para determinar si las profesoras eran tratadas en forma distinta que los profesores. Sin embargo, en *Trout v. Garrett*, 780 F. Supp. 1396, 1414 (D.D.C. 1991) <http://www.ll.georgetown.edu/federal/judicial/dc/opinions/01opinions/01-5325a.html> el tribunal dictaminó, en la etapa de daños del juicio, que la duración de la experiencia laboral previa a su contratación no constituía una variable adecuada en un análisis de regresión a fin de computar los salarios atrasados en materia de discriminación de empleo. Según el tribunal, incluir los sueldos atrasados resultaría en que no se hallaría discriminación sexual, pese a las conclusiones en contrario de la etapa de daños del juicio. *Id.* ver también *Stuart v. Roache*, 951 F.2d 446 (1st Cir. 1991) <http://bulk.resource.org/courts.gov/c/F2/951/951.F2d.446.91-1483.html> (que admitía sólo 3 años de antigüedad a ser considerados como resultado de la discriminación previa, cert. denied, 504 U.S. 913 (1992)). Que haya o no variables determinadas que reflejan consideraciones “legítimas” o que incorporan sesgos ilegítimos es un tema recurrente de la discriminación. Ver p.ej. *Smith v. Virginia Commonwealth Univ.*, 84 F.3d 672, 677 (4th Cir. 1996) <http://www.ftc.gov/os/2000/10/heinzbrief.pdf> (en banc); *id.* at 681–82 (Luttig, J., concurr. in part) (sugirió que el análisis de regresión no incluyó “factores de rendimiento” al punto de tornarlo incompleto e inadmisibles); *id.* at 690–91 (Michael, J., dissenting) (sugiere que el análisis de regresión excluyó en forma correcta los



No siempre es posible incluir todas las variables que pueden ejercer influencia sobre la variable dependiente para que el análisis sea exitoso; algunas pueden ser medidas, y otras pueden ser de poca importancia.<sup>34</sup> Si un análisis preliminar muestra que la porción no explicada de la regresión múltiple es inaceptablemente alta, el experto debe tratar de descubrir si hay alguna variable que no detectó y omitió en el análisis.<sup>35</sup> Si se comete el error de excluir una variable explicativa importante correlacionada con la variable que resulta de interés en un modelo de regresión puede ocasionarse que a una variable incluida se le adjudique un efecto que en realidad está causado por la variable omitida. En otros términos, esta omisión genera sesgos de los parámetros estimados. Por regla general, las variables omitidas correlacionadas con la variable dependiente reducen el valor probatorio del análisis de regresión. (Este efecto puede llegar a ser importante si hay una fuerte relación entre la variable omitida y la variable dependiente, y una fuerte correlación entre la primera variable y las variables explicativas interesantes.) Esto puede conducir a inferencias hechas con el análisis de regresión que no resulten útiles para el juez de los hechos.<sup>36</sup>

La omisión de variables no correlacionadas con la variable interesante, en general, no representa tanto un problema, porque al parámetro que mide el efecto de la variable interesante sobre la variable dependiente se lo estima sin sesgo. Por ejemplo, supongan que el efecto de una política introducida por los tribunales para alentar a los maridos a que costeen el sostén de sus hijos ha sido probada eligiendo en forma aleatoria algunos casos manejados según políticas actuales de

---

“factores de rendimiento”); ver también *Diehl v. Xerox Corp.*, 933 F. Supp. 1157, 1168 (W.D.N.Y. 1996). <http://caselaw.findlaw.com/us-2nd-circuit/1293020.html>

<sup>34</sup> El efecto combinado de las variables excluidas aparece dentro del término aleatorio en el modelo de regresión, así como cualquier error de modelización. Ver el Apéndice del capítulo para más detalles. Pero David W. Peterson, en su ensayo crítico sobre la Reference Guide on Multiple Regression, 36 *Jurimetrics J.* 213, 214 n.2 (1996) sostiene que “el supuesto de que el efecto combinado de las variables explicativas omitidas del modelo no guardan correlación con las variables explicativas incluidas” constituye una condición de “filo de la navaja... que resulta poco verosímil”.

<sup>35</sup> Un  $R^2$  demasiado bajo es indicativo de que la porción no explicada del modelo de regresión múltiple es demasiado elevada. Pero esta conclusión dependerá del contexto y de los datos, pues como veremos en el Apéndice, un  $R^2$  bajo no necesariamente implica un modelo pobre (y *viceversa*).

<sup>36</sup> *Bazemore v. Friday*, 751 F.2d 662, 671–72 (4th Cir. 1984) <http://openjurist.org/751/f2d/662/pe-bazemore-v-c-friday-pe-bazemore>. Este análisis ratificó el rechazo de la corte distrital a aceptar un análisis de regresión múltiple como prueba de discriminación por preponderancia de la evidencia; la corte de apelaciones afirmó que, si bien la regresión utilizó cuatro variables (raza, educación, ejercicio de la ocupación y título del empleo), la falla en usar otros factores (entre ellos los incrementos de paga que variaban según el condado, impedían su consideración como evidencia), *aff'd in part, vacated in part*, 478 U.S. 385 (1986). Observen, no obstante que en *Sobel v. Yeshiva University*, 839 F.2d 18, 33, 34 (2d Cir. 1988), cert. denied, 490 U.S. 1105 (1989) <http://openjurist.org/839/f2d/18>, el tribunal dejó en claro que “un demandado [por Título VII] que cuestiona la validez de un análisis de regresión múltiple tiene que exhibir que los factores que sostiene deberían haber sido incluidos debilitaría exhibir una disparidad salarial producida en el análisis”, realizando un ataque específico y “exhibiendo la relevancia de todas las variables particulares que sostiene ... que deberían haber sido incluidas” en el análisis, en lugar de cuestionar simplemente los resultados de la demostración del demandante por la carencia de determinada variable. Ver también *Smith v. Virginia Commonwealth Univ.*, 84 F.3d 672 (4th Cir. 1996) (en banc) <http://openjurist.org/84/f3d/672> (halló que si ciertas variables deberían haber sido incluidas en un análisis de regresión es una cuestión de hecho que impide un juicio sumario). También en *Bazemore v. Friday* <http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=478&invol=385>, la Corte, al declarar que el punto de vista del Cuarto Circuito del valor de la evidencia de los análisis de regresión era claramente incorrecto, afirmó que “normalmente, la falla de incluir variables afecta el valor probatorio del análisis, no su admisibilidad. Es importante destacar que un análisis de regresión que incluya algo menos que ‘todas las variables medibles’ puede ser de utilidad para demostrar el caso de un demandante”. 478 U.S.385, 400 (1986) (footnote omitted). <http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=us&vol=478&invol=385>

los tribunales y otros manejados con arreglo a una nueva política más estricta. El efecto de la nueva política podría medirse mediante una regresión múltiple que use el éxito del pago como variable dependiente y 0 o 1 como variable explicativa (1 si se aplicó el nuevo programa; 0 si no lo fue). Si no se puede incluir una variable explicativa que refleje la edad del marido involucrado en el programa no afectaría la evaluación del tribunal de la nueva política, porque es probable que gente de todas las edades esté afectada tanto por la antigua política como por la nueva. La elección al azar de la política de los tribunales a ser aplicada en cada caso asegura que la variable edad omitida no está correlacionada con la variable de política.

El sesgo ocasionado por la omisión de una variable importante vinculada con la variable interesante puede constituir un problema serio. Sin embargo, es posible que el experto tenga en cuenta este sesgo en forma cualitativa si dispone de conocimiento (aunque no sea cuantificable) sobre la relación entre la variable omitida y la variable explicativa. Por ejemplo, supongan que al experto del demandante de un caso de salario discriminado por sexo le resulta imposible hallar datos cuantificables que reflejen las aptitudes necesarias para una tarea, y además que, en promedio, las mujeres son más aptas que los hombres. Supongan también que hay un análisis de regresión del salario de los empleados (variable dependiente) con respecto a los años de experiencia y a una variable que refleja el sexo de cada empleado (variables explicativas) que sugiere que a los hombres se les paga más que a las mujeres que tienen la misma experiencia. Dado que las diferencias de aptitud no han sido tomadas en cuenta, el experto puede concluir en forma razonable que la diferencia salarial medida por la regresión es un estimador conservador de la verdadera diferencia salarial discriminatoria.

También resulta importante la precisión con que es medido el efecto de una variable de interés sobre la variable dependiente.<sup>37</sup> En general, cuanto más completa sea la relación explicada entre las variables explicativas incluidas y la variable dependiente, tanto más precisos serán los resultados. Hay que tener en cuenta, sin embargo, que la inclusión de variables explicativas irrelevantes (que no guardan correlación con la variable dependiente) *reduce la precisión de los resultados de la regresión*. Esto puede ser fuente de preocupación con muestras pequeñas, pero no es probable que lo sea en caso de muestras grandes.

*Forma Funcional* La selección del conjunto de variables a incluir en un modelo de regresión múltiple no da término al ejercicio de modelización, porque el experto también debe elegir una forma funcional apropiada. La que se elige con mayor frecuencia es el modelo de regresión lineal. En este modelo, *la magnitud en que cambia la variable dependiente asociada al cambio de cualquier variable explicativa no depende del nivel de la variable explicativa*. Supongan que 1 año más de experiencia añade \$5,000 anuales a los salarios recibidos, sin tener en cuenta la experiencia previa del empleado. Empero, en ciertos casos puede haber motivos para pensar que cambios de la variable explicativa tengan un efecto diferencial sobre la variable dependiente a medida que cambian los valores de las variables explicativas. En esos casos, el experto debería utilizar un modelo *no lineal*. Si no se toman en cuenta las no linealidades, se podría estar sobre- o sub-estimando el efecto de un cambio del valor de una variable explicativa sobre la variable dependiente.

Hay un tipo especial de no linealidad que implica la interacción entre diversas variables. Una variable de interacción es el *producto* de otras dos variables incluidas en el modelo de regresión múltiple. La variable de interacción contemplada por el experto le permite tomar en cuenta la posibilidad de que el efecto del cambio de una variable sobre la variable dependiente cambie

<sup>37</sup> Un estimador más preciso de un parámetro es un estimador que tiene un error estándar más reducido. Ver detalles en el capítulo XXIII.

cuando cambia el nivel de otra variable explicativa. Por ejemplo, en un caso de discriminación salarial, la inclusión de un término de interacción entre una variable que mide el sexo del empleado (1 si es mujer, 0 si es hombre) le permite al experto dictaminar si el diferencial por sexo varía con el nivel de experiencia. Un estimador significativo del parámetro asociado con el sexo sugiere que existe discriminación en contra de las mujeres, mientras que un estimador negativo significativo del parámetro de interacción sugiere que la magnitud de la discriminación aumenta con la experiencia.<sup>38</sup>

Hay que notar que los coeficientes no significativos de un modelo con interacciones pueden sugerir falta de discriminación, mientras que un modelo sin interacciones puede sugerir todo lo contrario. Resulta de interés especial tener en cuenta el tipo de interacción de la discriminación; si no se lo hace bien, se puede llegar a conclusiones falsas.

*Elección de la Regresión Múltiple como Método* Hay diversas técnicas estadísticas multivariadas, además de la regresión múltiple, que son útiles en el ámbito legal. Hay algunos métodos estadísticos útiles cuando se presentan no linealidades importantes.<sup>39</sup> Otros son aplicables en modelos en los que la variable dependiente es discreta, en lugar de continua (modelos de elección cualitativa, como los modelos logit y probit). Incluso hay otros que han sido aplicados especialmente a cuestiones metodológicas que surgen en el contexto de los litigios discriminatorios.<sup>40</sup>

Resulta esencial aplicar un método estadístico válido para asistir al análisis de cada caso legal. Luego, el experto debe prepararse para explicar por qué el método elegido, incluyendo el de la regresión múltiple, fue más adecuado que las alternativas.

<sup>38</sup> Para detalles adicionales respecto del efecto de interacción, ver capítulo XXV. Observen que en el caso *Ottaviani v. State University of New York*, 875 F.2d 365, 367 (2d Cir. 1989), cert. denied, 493 U.S. 1021 (1990) <http://www.employmentlawgroup.net/Opinions/MurtaghCooke-NTSB.asp>, la parte demandada se apoyó en un modelo de regresión con una variable dummy que reflejaba el género como variable explicativa. La mujer demandante, por el contrario, utilizó un enfoque alternativo con un modelo de regresión que sólo había sido desarrollado para los hombres (el supuesto grupo protegido). Los salarios femeninos predichos por esta ecuación fueron comparados luego con los salarios reales; según la demandante, una diferencia positiva implicaría una evidencia de discriminación. Para evaluar las ventajas y desventajas de este enfoque, véase Joseph L. Gastwirth, *A Clarification of Some Statistical Issues in Watson v. Fort Worth Bank and Trust*, 29 *Jurimetrics J.* 267 (1989).

<sup>39</sup> Estas técnicas incluyen, entre otras, la regresión lineal a trozos [http://en.wikipedia.org/wiki/Segmented\\_regression](http://en.wikipedia.org/wiki/Segmented_regression), la regresión polinómica [http://en.wikipedia.org/wiki/Polynomial\\_regression](http://en.wikipedia.org/wiki/Polynomial_regression), la estimación por máxima verosimilitud de modelos con relaciones funciones no lineales [http://en.wikipedia.org/wiki/Maximum\\_likelihood](http://en.wikipedia.org/wiki/Maximum_likelihood), y los modelos autorregresivos y de promedios móviles de series de tiempo [http://en.wikipedia.org/wiki/Autoregressive\\_moving\\_average\\_model](http://en.wikipedia.org/wiki/Autoregressive_moving_average_model). V. R. Pindyck and D. Rubinfeld, *Econometric Models and Economic Forecasts*, p. 117–21, 136–37, 273–84, 463–601 (4th ed. 1998).

<sup>40</sup> Los economistas debaten cuál es el modelo mejor adaptado en cuestiones de discriminación salarial. Se han comenzado a evaluar enfoques alternativos, incluyendo modelos de urna (Bruce Levin & Herbert Robbins, *Urn Models for Regression Analysis, with Applications to Employment Discrimination Studies*, *Law & Contemp. Probs.*, Autumn 1983, at 247) y, como forma de corregir los errores de medición, regresión invertida (Delores A. Conway & Harry V. Roberts, *Reverse Regression, Fairness, and Employment Discrimination*, 1 *J. Bus. & Econ. Stat.* 75 (1983)). También ver Arthur S. Goldberger, *Redirecting Reverse Regressions*, 2 *J. Bus. & Econ. Stat.* 114 (1984); Arthur S. Goldberger, *Reverse Regression and Salary Discrimination*, *The Journal of Human Resources*, Vol. 19, No. 3 (Summer, 1984), pp. 293-318 <http://isites.harvard.edu/fs/docs/icb.topic709943.files/goldberger%20reverse%20discrimination%201984.pdf>; Arlene S. Ash, *The Perverse Logic of Reverse Regression*, in *Statistical Methods in Discrimination Litigation* 85 (D.H. Kaye & Mikel Aickin eds., 1986).

#### 4. Interpretación de los Resultados de una Regresión Múltiple

Los resultados de una regresión múltiple pueden ser interpretados en términos puramente estadísticos, usando tests de significación, o en forma más práctica y menos estadística. Si bien la evaluación del significado práctico de los resultados de regresión casi siempre es relevante para un tribunal, los tests de significación estadística sólo son adecuados en determinadas circunstancias.

*Significación Práctica de los Resultados de Regresión* El significado práctico implica que la magnitud del efecto estudiado no es *de minimis* – es lo suficientemente importante, en términos sustanciales, como para que el tribunal se ocupe del mismo. V.g., si el salario horario promedio es \$10,00/hora, un diferencial salarial de \$0,10/hora entre hombres y mujeres será prácticamente insignificante porque representa sólo un 1% del salario promedio. No existe un % umbral que indique si un resultado es significativo desde el punto de vista práctico. Esa misma diferencia podría ser estadísticamente significativa si fuera estudiada una muestra grande de hombres y mujeres.<sup>41</sup> El motivo es que la significación estadística viene dada, parcialmente, por el número de observaciones en el conjunto de datos.

Manteniendo lo demás constante, la significación estadística de un coeficiente de regresión aumentará a medida que crezca el tamaño de la muestra. Luego, un diferencial de \$1 entre los salarios horarios de hombres y mujeres que sería no significativamente distinto de 0 en una muestra de 20 individuos podría ser muy significativo si la muestra se incrementara hasta 200.

A veces los resultados significativos desde el punto de vista práctico también lo son desde el estadístico.<sup>42</sup> Sin embargo, también es posible hallar coeficientes estadísticamente significativos con grandes conjuntos de datos que no lo son desde el punto de vista práctico. Y también es posible (especialmente en muestras pequeñas) tener resultados prácticamente significativos pero

<sup>41</sup> El significado práctico también puede aplicarse a la credibilidad global de los resultados de regresión. Así, en *McCleskey v. Kemp*, 481 U.S. 279 (1987) [http://en.wikipedia.org/wiki/McCleskey\\_v.\\_Kemp](http://en.wikipedia.org/wiki/McCleskey_v._Kemp), los coeficientes de las variables raciales eran estadísticamente significativos, pero la Corte se inclinó por no considerarlos legal o constitucionalmente significativos. Un resumen del caso puede hallarse en *The New Georgia Encyclopedia, McCleskey v. Kemp (1987)* <http://www.georgiaencyclopedia.org/nge/Article.jsp?path=/GovernmentPolitics/Government/LegalCases&id=h-2933>

<sup>42</sup> En *Melani v. Board of Higher Education*, 561 F. Supp. 769, 774 (S.D.N.Y. 1983) [Ver Robert C. Anderson, *An Alternative Approach to Higher Education* <http://ageconsearch.umn.edu/bitstream/17684/1/ar830208.pdf>, donde se menciona la decisión adoptada por el Juez Gagliardi] Una causa por el Título VII fue llevada contra la Universidad de la Ciudad de New York (CUNY) por haber discriminado supuestamente en contra del plantel de instrucción femenino en el pago de sus salarios. Un enfoque que usó el experto de la parte demandante fue el análisis de regresión múltiple. El coeficiente que reflejaba al sexo del empleado fue aproximadamente de \$1,800 cuando todos los años eran incluidos. Prácticamente (en términos de salarios promedio de esa época) y estadísticamente (al 5% de significación) el resultado fue significativo. Por consiguiente, el tribunal dictaminó que la demandante “había producido evidencia estadísticamente significativa de que las mujeres contratadas como plantel de instrucción a partir de 1972 recibieron salarios sustancialmente más reducidos que hombres de calificación similar”. Id. at 781. Para otro análisis vinculado que involucró comparaciones múltiples, ver *Csicseri v. Bowsher*, 862 F. Supp. 547, 572 (D.D.C. 1994) (donde se señala que el experto de la parte demandante halló “instancias estadísticamente significativas de discriminación” en 2 de 37 comparaciones, pero que sugirió que “2 de 37 representa groseramente un 5% y escasamente puede ser considerado un patrón discriminatorio”), *aff'd*, 67 F.3d 972 (D.C. Cir. 1995). [http://dc.findacase.com/research/wfrmDocViewer.aspx/xq/fac.%5CFDCT%5CDDC%5C1994%5C19940915\\_0000120.DDC.htm/qx](http://dc.findacase.com/research/wfrmDocViewer.aspx/xq/fac.%5CFDCT%5CDDC%5C1994%5C19940915_0000120.DDC.htm/qx)

no significativos estadísticamente. Por ejemplo, supongan que un experto encara un estudio por daños en un caso de infracción a la ley de patentes y predice “lo que hubieran sido las ventas si no hubiera tenido lugar la infracción” usando datos a lo largo del período de infracción. Si sólo hay datos conocidos durante 3 o 4 años del período pre-infracción, la diferencia entre “las ventas que hubieran sido” y las ventas reales durante el período de la infracción que se alega podría ser significativa desde el punto de vista práctico pero no desde el punto de vista estadístico.

*¿Cuándo cabe usar un Contraste Estadístico?* Un contraste en una contienda específica, un test de hipótesis, ayuda con frecuencia al tribunal a determinar si se produjo una violación de la ley en áreas donde la evidencia directa es inaccesible o inconcluyente. Por ejemplo, un experto puede usar un contraste de hipótesis en casos de discriminación regional o sexual para determinar la presencia de un efecto discriminatorio. La evidencia estadística sola nunca puede demostrar con certeza absoluta la validez de una teoría. Pero al proveer evidencia *contraria* al punto de vista de que no tuvo lugar una forma particular de discriminación, por ejemplo, el análisis de regresión múltiple puede ser de ayuda al juez de los hechos para evaluar la probabilidad de que se haya producido discriminación.<sup>43</sup> Los tests de hipótesis son útiles en análisis de sección cruzada, donde los datos del estudio de regresión han sido elegidos como muestra de una población en un momento dado del tiempo, y en análisis de series temporales, cuando los datos evaluados cubren un cierto número de períodos. En cualquiera de ambos análisis, el experto puede desear evaluar una hipótesis específica, vinculada habitualmente a una cuestión de responsabilidad o a determinar si existe un impacto medible de una supuesta violación. Por caso, en una cuestión de discriminación sexual, al experto le puede interesar evaluar la hipótesis nula de ausencia de discriminación contra la hipótesis alternativa de que existió discriminación en cierto sentido. Otro ejemplo es un juicio por daños anti-monopolístico, donde el experto puede desear probar la hipótesis nula de que no existió impacto legal en contra de la hipótesis alternativa de que ese impacto sí existió. En todo caso, es importante reflexionar acerca de que el rechazo de la hipótesis nula no es prueba en sí misma de responsabilidad legal. Es posible rechazar la hipótesis nula y creer que hay que tener en cuenta una explicación alternativa distinta de la responsabilidad legal.

Es frecuente que la hipótesis nula se plantee en términos de que cierto coeficiente de regresión es igual a 0. Por ejemplo, en el caso de discriminación salarial, la hipótesis nula sería que no hay discriminación salarial entre sexos. Si se observa una diferencia negativa (que implicaría que las mujeres ganan menos que los hombres, luego de controlar otras explicaciones legítimas alternativas), la diferencia debe evaluarse en cuanto a significación estadística utilizando el *t*-test.<sup>44</sup> El *t*-test utiliza el estadístico *t* para evaluar la hipótesis de que un parámetro del modelo toma un valor particular, habitualmente 0.

Por convención, en la mayoría de los trabajos científicos el nivel de significación requerido para rechazar la hipótesis nula – es decir, obtener un resultado estadísticamente significativo – es .05 o 5%. El nivel de significación mide la probabilidad de que la hipótesis nula sea incorrectamente rechazada, en caso de ser verdadera. En general, cuanto más bajo sea el porcentaje requerido de significación estadística, tanto más difícil resultará rechazar la hipótesis nula; luego a menor probabilidad menos yerros se cometerán al hacerlo. El criterio del 5% es típico, pero también dan información útil informar los tests de significación más estrictos al 1% o menos estrictos al 10%.

<sup>43</sup> Ver *International Bhd. of Teamsters v. United States*, 431 U.S. 324 (1977) (la Corte infirió discriminación a partir de una aplastante evidencia estadística por preponderancia de la evidencia).  
<http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=431&invol=324>

<sup>44</sup> El *t*-test es estrictamente válido sólo si se cumple un número de importantes condiciones. Empero, en muchos modelos de regresión, el test es aproximadamente válido si el tamaño de la muestra es lo suficientemente grande. Ver capítulo XXV para una discusión más amplia de los supuestos subyacentes a la regresión múltiple.

Al practicarse un contraste estadístico, es útil computar un nivel de significación observado, o  $p$ -valor. El  $p$ -valor asociado con la hipótesis nula de que un coeficiente de regresión es 0 es la probabilidad de que un coeficiente de tal magnitud, o mayor, haya ocurrido al azar si la hipótesis nula fuera verdadera. Si el  $p$ -valor fuera menor o igual que 5%, el experto rechazaría la hipótesis nula a favor de la hipótesis alternativa; si el  $p$ -valor fuera mayor que 5%, el experto no podría rechazar la hipótesis nula.<sup>45</sup>

Cuando el experto evalúa la hipótesis nula de que una variable de interés no guarda asociación con una variable dependiente, en contra de la hipótesis alternativa de que sí existe, lo adecuado es utilizar un test a dos colas, que permita que el efecto sea tanto positivo como negativo. Un test a una sola cola sería habitualmente usado si el experto cree, tal vez basándose en otro tipo de evidencia presentada en el juicio, que la hipótesis alternativa es positiva o negativa, pero no ambas. Por caso, un experto podría usar un contraste a una sola cola en un alegato de infracción al precio de un producto infractor si cree que la infracción fue cero o negativa. (Las ventas del producto infractor competían con las ventas del producto que fue víctima, bajando así el precio). Como al usarse un test a una cola se obtienen  $p$ -valores que son la mitad de los  $p$ -valores de un test a dos colas, elegir un test a una cola le hace más fácil al experto rechazar una hipótesis nula. Luego, al usarse un test a dos colas el rechazo de la hipótesis es menos probable. Como elegir una hipótesis alternativa es algo arbitrario, Rubinfeld sostiene que *los tribunales deberían evitar descansar únicamente en un test estadístico definido nítidamente*.<sup>46</sup> *Debería alentarse a informar el  $p$ -valor o un intervalo de confianza, ya que brinda información útil al tribunal, sea o no rechazada una hipótesis nula.*

**Robustez** La idea de robustez – saber si los resultados de regresión son sensibles a pequeñas variaciones de los supuestos (p.ej., que los datos están medidos en forma precisa) – tiene una importancia vital. Si los supuestos del modelo de regresión son válidos, los tests estándar de los modelos pueden aplicarse. Empero, si se violan los supuestos del modelo, los tests estándar pueden sobre- o subestimar la significación de los resultados. Observen que *la violación de un supuesto no necesariamente implica que el análisis de regresión pierda su validez*. Hay casos en los cuales los supuestos del análisis de regresión múltiple no están presentes, pero entonces habrá que utilizar otros métodos estadísticos. Luego, a los expertos debería alentárselos a que faciliten información adicional más allá de que los resultados de la regresión sean válidos, y si no

<sup>45</sup> Usar 1%, 5%, y a veces, 10% como niveles para determinar la significación estadística sigue siendo un tema controvertido. Se puede decir, por ejemplo, que cuando se hace análisis de regresión para fijar precios en un caso anti-monopolístico a fin de contrastar una alternativa relativamente específica de la hipótesis nula (es decir, fijación de precios) un nivel de confianza algo más reducido (un mayor nivel de significación, como 10%) podría ser adecuado. En caso contrario, cuando la alternativa de la hipótesis nula es menos específica, p.ej. la alternativa bastante vaga de “efecto” (p.ej., el aumento de precios está causado por el aumento del costo de producción, la demanda incrementada, un fuerte aumento de la publicidad, o la fijación de precios), un nivel de confianza elevado (asociado con un menor nivel de significación, como por ejemplo 1%) puede ser más apropiado. V. Ramona L. Paetzold, Problems with Statistical Significance in Employment Discrimination Litigation, New England Law Review, Winter, 1991; John Donohue and Peter Siegelman, The Changing Nature of Employment Discrimination Litigation, Yale Law School Faculty Scholarship, 1991. [http://digitalcommons.law.yale.edu/fss\\_papers/42](http://digitalcommons.law.yale.edu/fss_papers/42)

<sup>46</sup> Aparentemente, los tribunales de US han manifestado su preferencia por pruebas a dos colas. Por ejemplo, Palmer v. Shultz, 815 F.2d 84, 95–96 (D.C. Cir. 1987), en cuyo juicio el tribunal encontró que, como algunos recurrentes solicitaban una selección excesiva en determinados empleos, un contraste a dos colas era más adecuado en el Título VII. También Csicseri v. Bowsher, 862 F. Supp. 547, 565 (D.D.C. 1994 [http://www.leagle.com/xmlResult.aspx?page=19&xmlDoc=19941409862FSupp547\\_11319.xml&docbase=CS LWAR2-1986-2006&SizeDisp=7](http://www.leagle.com/xmlResult.aspx?page=19&xmlDoc=19941409862FSupp547_11319.xml&docbase=CS LWAR2-1986-2006&SizeDisp=7)) (donde se sostuvo que si bien un test a una cola “no carece de mérito”, es preferible un test a dos colas).

lo son, en qué medida son robustos. A continuación veremos algunos de los supuestos más importantes del análisis de regresión.

En un contexto de regresión múltiple, el experto supone que a menudo los cambios de las variables explicativas influyen sobre la variable dependiente, y que cambios de la variable dependiente no afectan a las variables explicativas – en otras palabras, que no hay *feedback*.<sup>47</sup> A partir de este supuesto, el experto saca la conclusión de que la correlación entre una variable explicativa y la variable dependiente se debe al efecto de la primera sobre la segunda y no a la inversa. Si el supuesto no fuera válido, la correlación espuria podría llevar al experto y al juez a una conclusión errónea.<sup>48</sup>

La Figura 7 ilustra la cuestión. En la parte (a), la variable dependiente, Precio, es explicada mediante una estructura de regresión múltiple por tres variables explicativas, Demanda, Costo y Publicidad, sin feedback. En la parte (b), hay un feedback, ya que Precio incide sobre Demanda, y la Demanda, el Costo y la Publicidad inciden sobre el Precio. Pero el Costo y la Publicidad no están afectados por el Precio. Por regla general, no existe un contraste estadístico directo para determinar la relación de causalidad; en lugar de ello, cuando se le pregunta al experto, debería estar preparado para defender su supuesto basándose en una comprensión de cuál es la conducta subyacente de las empresas o individuos involucrados.

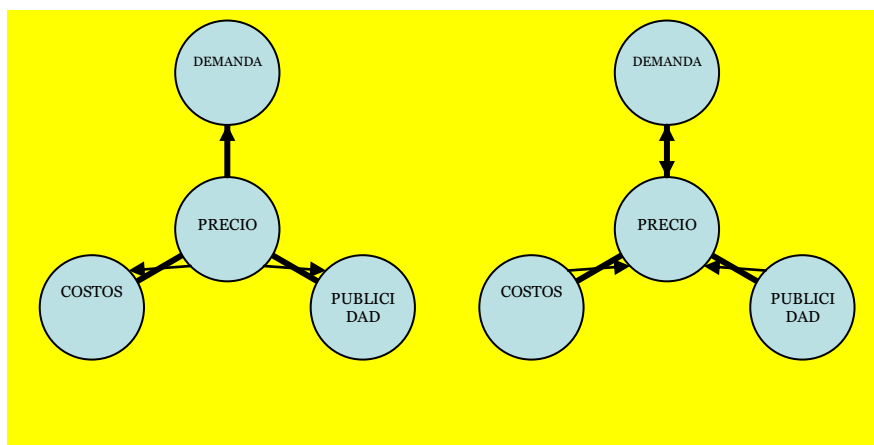


Figura 7 (a)

Figura 7(b)

Aunque no exista un único enfoque que sea completamente aplicable para estimar modelos en los cuales la variable dependiente afecta a una o más variables explicativas, una posibilidad es que el experto omita de la regresión a la variable cuestionada para apreciar si su exclusión implica alguna diferencia. Si no lo hace, la cuestión es debatible. Otro enfoque es expandir el modelo de regresión múltiple agregando una o más ecuaciones que expliquen la relación entre la variable explicativa en cuestión y la variable dependiente.

Supongan, por ejemplo, que en un juicio por discriminación sexual en materia salarial el experto del demandado considera que las valuaciones de las pruebas del empleador son una variable explicativa del salario (la variable dependiente). Si el demandante informa que el empleador ajustó las valuaciones a fin de penalizar a las mujeres, entonces lo que podría invalidarse es el supuesto

<sup>47</sup> Si ambos efectos tienen lugar al mismo tiempo se habla de "simultaneidad".

<sup>48</sup> El supuesto de que no existe un feedback es particularmente importante en los litigios, porque es posible que el defensor (si es responsable, por ejemplo, de la fijación de precios o de la discriminación) afecte las variables explicativas y termine así sesgando los contrastes usuales en regresión múltiple.

de que los salarios se determinan mediante valuaciones y no que las valuaciones fueron afectadas por los salarios. Si resulta inapropiada, la variable valuación de las pruebas debería ser excluida. En forma alternativa, la información sobre el uso que hace el empleador de las valuaciones podría traducirse en una segunda ecuación con una nueva variable dependiente, la valuación de la prueba, explicada por el sexo y el salario del trabajador. Un test de hipótesis de que el salario y el sexo afectan las valuaciones constituiría una prueba adecuada de ausencia de feedback.

Resulta esencial en el análisis de regresión múltiple que *las diversas variables explicativas de interés no estén perfectamente correlacionadas entre sí*. Si existiera correlación perfecta entre dos variables, *el experto no podría separar el efecto de la variable de interés sobre la variable dependiente del efecto de la otra variable*. Supongan por ejemplo que un juicio por discriminación sexual se requiere un cierto tipo de experiencia laboral para que sea motivo de salarios más elevados. Si todos los hombres tuvieran la experiencia laboral requerida y no es el caso de las mujeres, resultaría imposible decir si los diferenciales salariales se deben a discriminación sexual o a diferencias de experiencia. Si dos o más variables están perfectamente correlacionadas – o sea, que existe una colinealidad perfecta – la regresión puede estimarse, pero con problemas. A mayor multicolinealidad entre dos variables, tanto menos precisos serán los estimadores de los parámetros de regresión individuales (si bien puede no haber ningún problema en estimar la influencia *conjunta* de ambas variables y de los restantes parámetros de regresión).

Por fortuna, los estadísticos de regresión informados toman en cuenta si existe alguna multicolinealidad.<sup>49</sup> Como corolario, es importante tener en cuenta que no hallar una relación estrecha entre una variable de interés y una variable dependiente no implica que no exista ninguna relación.<sup>50</sup> Una muestra relativamente pequeña, o una muestra grande con un grado sustancial de multicolinealidad, pueden no facilitarle al experto suficiente información como para determinar si existe alguna relación.

*Independencia de los Errores del Modelo de Regresión* Si el experto calculó los parámetros de un modelo de regresión múltiple usando datos de toda la población es posible que los estimadores midan sus parámetros con errores. Los errores pueden surgir por diversas razones, entre ellas: (1) fallas del modelo que no incorpora las variables explicativas apropiadas; (2) fallas del modelo que no refleja fallas de linealidad que podrían presentarse; (3) inclusión de variables inadecuadas en el modelo. (Naturalmente, otra fuente de error podría surgir porque se usa una muestra o subconjunto de la población para estimar los parámetros de regresión). Resulta útil visualizar el efecto acumulado de todas estas fuentes de errores de modelización como representado mediante un término adicional, el *término de error*. Hay un supuesto importante del análisis de regresión múltiple: *el término de error de cada variable explicativa es independiente de los demás*, lo que implica que no están correlacionados entre sí. Si esto es cierto, el experto puede estimar los parámetros del modelo sin incurrir en sesgos; la magnitud del término de error tendrá efecto sobre la precisión con la que es estimado el modelo, pero no será causa de que este estimador sea demasiado alto o bajo. El supuesto de independencia de los errores puede resultar inapropiado en diversas circunstancias. En algunos casos, la falla del supuesto convierte al

<sup>49</sup> V. Denny v. Westfield State College, 669 F. Supp. 1146, 1149 (D. Mass. 1987) (El tribunal aceptó el testimonio de un experto de que “la presencia de multicolinealidad llevaría a sobre-estimar el error asociado con el estimador... En otros términos, los *p*-valores serán artificialmente más elevados que lo que serían de no existir el problema”). <http://openjurist.org/880/f2d/1465/denny-v-westfield-state-college>

<sup>50</sup> Si una variable explicativa preocupante y otra variable explicativa están muy correlacionadas, eliminar a la segunda variable de la regresión puede resultar instructivo. Si el coeficiente de la variable explicativa preocupante se torna significativo, se sugiere que existe una relación entre la variable dependiente y la variable preocupante.



análisis de regresión múltiple en una técnica estadística inútil; en otros, hay modificaciones o ajustes que permiten que la estructura de regresión se acomode a la falla. El supuesto de independencia puede fallar, por ejemplo, en un estudio de la conducta de los individuos a través del tiempo, cuando un error demasiado elevado en un período es seguido por un error demasiado alto en el siguiente. Por ejemplo, si pronostico el PIB de este año muy por debajo del nivel real, lo más probable es que también lo haga el año próximo; el factor causante del error de predicción (p.ej. un supuesto incorrecto sobre el coeficiente de riesgo país) es probable que siga siendo una fuente de error en el futuro.

También es probable que el supuesto de independencia falle en un estudio de un grupo de empresas en un momento determinado del tiempo, cuando el error producido para las empresas grandes sea sistemáticamente mayor que el de las pequeñas. Por ejemplo, un análisis de rentabilidad de las empresas puede que no tome en cuenta en forma adecuada a la publicidad como fuente de incremento de las ventas y los beneficios. Si las firmas más grandes hacen más inversión en publicidad que las pequeñas, los errores de regresión serán más amplios para las grandes empresas y más pequeños para las más pequeñas.

Hay veces en que se dispone de pruebas estadísticas adecuadas para evaluar la hipótesis de independencia.<sup>51</sup> Si falla el supuesto, lo primero que debería hacer el experto es preguntarse si la fuente de la falla de independencia podría deberse a la omisión de alguna importante variable explicativa de la regresión. Si es así, cuando sea posible la variable debería ser incluida, o estimarse el efecto potencial de su omisión en caso de no incluirla. Si no hubiera ninguna variable explicativa faltante, el experto podrá aplicar un procedimiento que modifica la técnica de regresión múltiple estándar para obtener estimadores más precisos de los parámetros de regresión.<sup>52</sup>

*Sensibilidad a Datos Puntuales* Los coeficientes de regresión pueden ser muy sensibles a datos puntuales particulares. Supongan, p.ej., que un dato puntual se desvía en gran medida de su valor esperado, indicado por la ecuación de regresión, en tanto que los puntos restantes no ostentan grandes desvíos. En esta situación no sería de extrañar que los coeficientes de una regresión múltiple cambien sustancialmente si ese dato fuera excluido de la muestra.

Evaluar cuán robustos son los resultados de regresión múltiple es un cometido complejo. Por lo tanto, no existe ningún conjunto de pruebas de robustez que conciten el interés de los econométricos. En general resulta importante explorar los motivos por los cuales hay puntos atípicos. Si la fuente es un error en la captación de datos, será posible hacer correcciones adecuadas. Si todos los datos atípicos tienen ciertas características en común, p.ej. todos están asociados con un supervisor que siempre otorga un puntaje más alto en un caso de igualdad de paga) el modelo de regresión debería modificarse de modo apropiado. Una técnica generalmente útil es determinar en qué medida el parámetro estimado cambia cuando cada punto del análisis de regresión es eliminado de la muestra. Un dato puntual influyente – que causa que el parámetro estimado cambie sustancialmente – debería ser entonces estudiado ulteriormente con el fin de

---

<sup>51</sup> En los análisis de series temporales, la correlación de los errores a lo largo del tiempo, o correlación serial, puede ser contrastada (en la mayoría de los casos) usando un test de Durbin-Watson. También puede ser contrastada la hipótesis de que algunos términos de error sean altos sistemáticamente, y otros sistemáticamente bajos, problema denominado de *heterocedasticidad*.

<sup>52</sup> Si hay correlación serial, existen métodos estadísticos estrechamente vinculados que resultan apropiados, como una diferenciación generalizada (que es un *método de mínimos cuadrados generalizados*) y la estimación mediante máxima verosimilitud. Si el problema es la heterocedasticidad, los mínimos cuadrados ponderados y los estimadores de máxima verosimilitud son apropiados. Todas estas técnicas están disponibles en los programas de cómputo usuales. Los programas permiten realizar tests estadísticos de significación de los coeficientes de regresión.

determinar si se cometieron errores en el uso de los datos o si fueron omitidas variables explicativas importantes.<sup>53</sup>

*Errores de Medición* En análisis de regresión múltiple suponemos que las variables están correctamente medidas.<sup>54</sup> Si hay errores de medición de la variable dependiente, los estimadores de los parámetros de regresión serán menos precisos, aunque no estarán necesariamente sesgados. Empero, si una o más variables independientes vienen medidas con error, los estimadores paramétricos correspondientes estarán probablemente sesgados, muy probablemente hacia cero (y probablemente haya otros coeficientes estimados sesgados). Para entender por qué, supongan que la variable dependiente (salario) es medida sin margen de error, y que la variable independiente (experiencia) está sujeta a error de medición. (La antigüedad en la empresa o los años de experiencia deberían ser exactos, pero el tipo de experiencia está sujeto a error, ya que los postulantes pueden exagerar sus responsabilidades laborales previas). A medida que crece el error de medición, los parámetros estimados asociados con la variable experiencia tenderán a 0, es decir que eventualmente no existirá relación alguna entre salario y experiencia.

Resulta importante que toda fuente de error de medición sea cuidadosamente evaluada. En algunos casos, se podrá hacer poco para corregir el problema de error de medición; los resultados de regresión deben ser interpretados a la luz de este hecho. En otros casos, el experto podrá corregir el error de medición buscando nuevas fuentes de datos más confiables. Finalmente, hay técnicas de estimación alternativas (que usan variables relacionadas medidas sin error) que pueden ser aplicadas a fin de remediar el problema del error de medición en ciertas situaciones.<sup>55</sup>

## 5. El Experto

El análisis de regresión múltiple se dicta a estudiantes en campos tan diversos como la estadística, la economía, las ciencias políticas, la sociología, la psicología, la antropología, la salud pública y la historia. En consecuencia, un individuo con entrenamiento sustancial y experiencia en regresión múltiple y otros métodos estadísticos puede ser calificado de experto. Un grado doctoral en una disciplina que dicta estadística teórica o aplicada, como la economía, la historia y la psicología, significa habitualmente para otros científicos que el experto propuesto satisface las pruebas preliminares del proceso de calificación.

La decisión de calificar a un experto en análisis de regresión pertenece al tribunal. Obviamente, el experto propuesto debería ser capaz de demostrar que entiende de qué se trata. Indicadores de la experiencia profesional pueden ser las publicaciones realizadas vinculadas con análisis de regresión en revistas calificadas por sus pares, su pertenencia activa a organizaciones profesionales vinculadas, cursos dictados sobre métodos de regresión, y su experiencia práctica en análisis de regresión. Empero, la capacidad y experiencia con respecto a las cuestiones

<sup>53</sup> Para un tratamiento más completo y formal del problema del carácter robusto de una regresión ver David A. Belsley et al., *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, Wiley, 2004. 229–44 <http://staff.deuv.cl/leiva/archivos/books/diagnostic.pdf>. Una útil discusión de los valores atípicos y de la evaluación de los conjuntos de puntos con influencia puede verse en R. Dennis Cook and Sanford Weisberg, *Residuals and Influence in Regression*, in *Monographs on Statistics and Applied Probability* (1982). [http://ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13859&Itemid=0](http://ebour.com.ar/index.php?option=com_weblinks&task=view&id=13859&Itemid=0)

<sup>54</sup> La imprecisión puede darse no solamente con respecto a la variable particular medida, sino también a la precisión con que la variable medida corresponde al constructo teórico apropiado especificado por el modelo de regresión.

<sup>55</sup> V. Robert S. Pindyck and Daniel L. Rubinfeld, *Econometric Models and Economic Forecasts*, McGraw-Hill International Edition, 4th edition, 1998, pp. 178-98 (discusión de la técnica de variables instrumentales).

determinadas aplicables en cuestiones específicas e instrumentos aplicables a un caso particular también deberían ser considerados durante el proceso de calificación.

## 6. Presentación de la Evidencia Estadística

Los costos de evaluar la evidencia estadística pueden reducirse, y aumentarse la precisión de esa evidencia, si el proceso de descubrimiento se usa de manera efectiva. Cuando se evalúa cuán admisible es la evidencia estadística, un tribunal debería prestar atención a las dos cuestiones siguientes:

- ¿Acercó el experto información suficiente como para replicar el análisis de regresión múltiple?
- ¿Son razonables las decisiones metodológicas tomadas por el experto, o resultan arbitrarias e injustificadas?

*Desacuerdo sobre Datos* En general, realizar una presentación clara y comprensiva de la metodología de investigación subyacente forma parte integral del proceso de descubrimiento. Al experto debería alentársele a que revele tanto la naturaleza de los experimentos llevados a cabo como la sensibilidad de los resultados a los datos y a la metodología. Sugerencias:

- En la medida posible, a las partes se las debería alentar a que usen una base de datos común. Si hay desacuerdo sobre lo que esos datos significan, un acuerdo oportuno sobre una base de datos en común ayuda a concentrar el esfuerzo de descubrimiento sobre las cuestiones importantes del caso.
- Una parte que ofrece datos a ser utilizados en el trabajo estadístico, incluyendo tareas de regresión múltiple, debería ser alentada a proveer lo siguiente a la otra parte: (a) si está disponible y es de tamaño manejable, una *hard-copy* de los datos con sus fuentes subyacentes; (b) soportes magnéticos de computación de los datos grabados; (c) una documentación completa de los discos, cintas, etc.; (d) los programas de computación usados para generar los datos (la *hard-copy*, los discos o cinta, o ambos); y (e) la documentación de esos programas de computación.
- La parte que ofrece los datos debería poner a disposición de la otra parte al personal involucrado en la compilación de datos para responder a preguntas técnicas sobre los datos y los métodos de recopilación o de compilación.
- Una parte que ofrece una regresión de un experto en un juicio debería pedirle al experto que informe plenamente sobre: (a) la base de datos y sus fuentes;<sup>56</sup> (b) el método de recolección de datos; y (c) los métodos de analizarlos. Si es posible, esta información debería ser proporcionada con bastante antelación al juicio, de manera que la parte oponente pueda consultar a sus expertos y preparar un examen cruzado. El tribunal debe decidir caso por caso dónde debe establecerse la línea de revelación de los datos.
- A la parte oponente debe dársele la oportunidad de objetar la base de datos o el método de análisis de la base de datos ofrecida en el juicio. Las objeciones pueden ir desde errores cometidos por los oficinistas a cuestiones más complejas vinculadas con la

<sup>56</sup> Estas fuentes incluyen variables usadas en los análisis estadísticos realizados por el experto, no sólo las utilizadas en el análisis final sobre el que los expertos esperan fundamentarse.

selección de datos, la construcción de variables, y a veces la forma particular del método estadístico utilizado. Si es posible, estas cuestiones deberían ser resueltas antes del juicio.

- A las partes hay que alentarlas a que resuelvan sus diferencias en cuanto al carácter apropiado y precisión de los datos, en la medida que sea posible mediante una conversación informal. El tribunal debería hacer un esfuerzo en resolver las diferencias con anterioridad al juicio.

*Resolver Diferencias*<sup>57</sup> Las siguientes pautas son sugeridas para que los expertos presenten sus bases de datos y procedimientos analíticos. Seguir las puede ser de gran utilidad para resolver disputas en cuestiones estadísticas.

- El experto debe establecer claramente los objetivos del estudio, así como el período de tiempo al que es aplicable y la población estadística cuyos resultados son proyectados.
- El experto debe informar las unidades observadas (es decir, consumidores, empresas, o empleados).
- El experto debe medir cada variable en forma clara.
- El experto debe identificar claramente la muestra con la cual fueron estudiados los datos,<sup>58</sup> como así también el método mediante el cual dicha muestra fue extraída.
- El experto debe revelar si hay datos faltantes, ya fuera causados por su indisponibilidad (p.ej. datos comerciales) o ausencia de respuesta (p.ej., en datos censales) y el método que fue seguido para su tratamiento (p.ej., supresión de observaciones).
- El experto debe informar las investigaciones realizadas con respecto a los errores asociados a la elección de variables y de supuestos que subyacen al modelo de regresión.
- Si las muestras fueron seleccionadas al azar en una población (es decir, que se siguió un muestreo probabilístico),<sup>59</sup> el experto debe hacer un esfuerzo de buena fe para proporcionar una estimación del error muestral, y la medición de la diferencia entre el estimador muestral de un parámetro (como la media de la variable dependiente estudiada) y el parámetro poblacional (desconocido) – la media poblacional de la variable.<sup>60</sup>
- Si fueron utilizados procedimientos no probabilísticos, el experto debe informar acerca del conjunto de procedimientos utilizados para minimizar los errores muestrales.

<sup>57</sup> Una discusión más completa de estos requerimientos será hallada en Lirioka Meintjes-Van der Valt, *The Proof of the Pudding: The Presentation and Proof of Expert Evidence in South Africa*, *Journal of African Law*, 47, 1 (2003), 88–106; School of Oriental and African Studies. [http://ebour.com.ar/index.php?option=com\\_weblinks&task=view&id=13864&Itemid=0](http://ebour.com.ar/index.php?option=com_weblinks&task=view&id=13864&Itemid=0)

<sup>58</sup> Esto es importante porque le permitirá realizar inferencias sobre la población relevante.

<sup>59</sup> En un muestreo probabilístico, cada representante de la población tiene una probabilidad conocida dentro de la muestra. El muestreo probabilístico es ideal porque está muy estructurado, y en principio puede ser replicado. Los muestreos no-probabilísticos son menos deseables porque con frecuencia son subjetivos, y descansan en gran medida en el juicio del experto.

<sup>60</sup> Los errores muestrales son frecuentemente informados en términos de errores estándar o intervalos de confianza. Ver el Capítulo XXIII para detalles adicionales.