UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

FACULTAD DE CIENCIAS

ACCIDENTES PROFESIONALES: UNA ESTIMACION PARA EL CASO IMSS (1960 - 2000)

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I A

P R E S E N T A :

MONICA ILIANA SANCHEZ ZARAGOZA

DIRECTOR DE TESIS:
ACT. VICTOR MANUEL SOLIS NAJERA

2005

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales de la Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:
"Accidentes profesionales: Una estimación para el caso IMSS (1960-2000)."

realizado por Sánchez Zaragoza Mónica Iliana

con número de cuenta 08821585-1, pasante de la carrera de Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	Act. Víctor Manuel Solís Nájera	
Propietario	Act. María Aurora Valdés Michell	
Propietario	Act. Laura Miriam Querol González	
Suplente	Act. Marina Castillo Garduño	
Suplente	Mat. César Eduardo Sousa Mondragón	

Consejo Departamental de Matemáticas

Act. Jaime Vázquez Alamilla

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMATICAS

A la UNAM,
por formarme

A la Facultad de Ciencias
porque es mi casa

A todos mis sinodales y revisores
por regalarme tantas cosas
importantes y sobre todo:
su talento

A mi madre, Juana Zaragoza
por darme su amor, apoyo
y por creer y confiar siempre
en mi.

A mi padre (†), Roberto Sánchez
por hacerme sentir tan especial.

A mis hermanos,
Norma, Carlos, Miguel y Pedro
porque son parte
fundamental en mi vida.

A mis hermosas niñas
Jenny y Citlali
por iluminar mi vida

A todos mis amig@s,
por creer que podía lograrlo
por estar a mi lado en todos
los momentos importantes
y por ser eso simplemente:
mis amig@s.

Indice

Introducción

Los accidentes y enfermedades derivados del trabajo causan un enorme daño y sufrimiento humano para las víctimas, familiares y personas de su entorno que ninguna compensación económica es capaz de reparar. Además de sufrimiento, estos accidentes y enfermedades suponen un gran costo económico para la sociedad y para las empresas; costo que, al igual de los accidentes, podría reducirse drásticamente. No es fácil cuantificar ese costo ya que intevienen muchas variables que en ciertos casos no son conmensurables. No obstante, buena parte de estos son medibles.

Se estima que, cada año, un total de 250 millones de trabajadores sufren accidentes de trabajo, y que de estos 335 mil son mortales. Adicionalmente, se estima que cada año 160 millones de trabajadores sufren algún tipo de enfermedad profesional, y un número todavía mayor sufren o ven amenazado su bienestar físico y mental. Se trata de estimaciones conservadoras, pues las estadísticas de que se dispone no son completas y muchos accidentes no se registran. Además, los costos sociales y humanos de los accidentes del trabajo son enormes: los que sufren los trabajadores involucrados y sus familias no se pueden medir; y el costo económico total equivaldría al 4 por ciento del Producto Interno Bruto. Se trata de una situación crítica, que afecta a cada país, a cada sector de actividad económica, y a cada comunidad. De no adoptarse medidas de prevención y de control eficaces, es de prever que la situación continuará deteriorándose, como resultado del crecimiento de la población y de la expansión del mercado del trabajo; pues al aumentar la población activa aumenta también el número de trabajadores expuestos a los riesgos profesionales.

La tendencia y la evolución que se observa en algunas áreas de las condiciones de trabajo, como resultado de la globalización de la economía también es preocupante. La aparición de nuevas formas de producción que no van acompañadas de las necesarias medidas de prevención, están introduciendo nuevos riesgos. La presión del mercado hacia la liberalización, y la privatización y la debilitación de las medidas de protección -incluyendo el traspaso de las responsabilidades en materia de salud y seguridad laboral del gobierno a las empresas- dejan a los trabajadores en una situación muy vulnerable.

Por grande que sea el reto al que nos enfrentamos, debemos tener siempre presente que todos los accidentes y las enfermedades profesionales son evitables. Pero, ¿cómo cumplir con una tarea tan importante en un mundo del trabajo que sigue cambiando a ritmo cada vez más acelerado?. Es necesario, en primer lugar, que la seguridad y la salud en el trabajo ocupe el lugar que le corresponde en la agenda política de cada país. Esto requiere un compromiso de la clase política y de la clase empresarial para asegurar que la seguridad y salud en el trabajo se asuma como un elemento clave en los procesos de análisis, y de toma de decisiones, en materia de inversión y de organización de la producción, y que las organizaciones sindicales, los trabajadores mismos y las organizaciones no gubernamentales, expresión de los grupos sociales mas vulnerables, estén involucrados en estos procesos. En todo caso, y cara al futuro, interesa no olvidarse de que no es el ser humano el que debe estar al servicio de la economía, sino que por el contrario es el desarrollo económico el que debe ponerse al servicio del ser humano; y la seguridad y salud en el trabajo es un instrumento crucial para alcanzar ese objetivo.

En virtud de lo antes expuesto se ha llevado a cabo este trabajo de investigación, en el cual se trata de explicar la relación entre inversión monetaria y el número de accidentes de trabajo, no ha sido fácil llevar a cabo este ejercicio ya que se ha tenido que trabajar con diferentes estadísticas, en particular las publicadas por el IMSS, la inversión en el ramo de accidentes de trabajo publicados por el propio Instituto y por el INEGI, así como algunas otras fuentes secundarias necesarias para hacer cruces y comprobaciones de información. El trabajo se estructura de la manera siguiente, el capítulo 1 proporciona una breve descripción de la Seguridad Social, así como los inicios del IMSS, del seguro de accidentes de trabajo y las características de este último; se define también de manera breve lo que se entenderá por población, la metodología a seguir así como la información que será utilizada. El capítulo 2 desarrolla los conceptos del modelo de regresión lineal y se presentan los características de los estimadores. El capítulo 3 trata sobre el Modelo de Regresión Lineal Múltiple que será utilizado, en el capítulo 4 se presentan los resultados de la estimación de los accidentes de trabajo con base en la población y las aportaciones por parte de empleados y empleadores. Finalmente se presentan las conclusiones y un anexo con los datos utilizados en esta pequeña investigación.

Capítulo 1

La seguridad social en México

Hacia el año de 1942 existian circunstancias favorables para que finalmente pudiera implantarse en México el Seguro Social. Auspiciado por el interés del Presidente Ávila Camacho por las cuestiones laborales se creó la Secretaría de Trabajo y Previsión Social y la encomendó a quien fuera Secretario de Gobernación del régimen anterior, el licenciado Ignacio García Téllez; la función inicial de la naciente dependencia fue limar asperezas y procurar la conciliación obrero-patronal.

En diciembre del mismo año se envió a las Cámaras la iniciativa de Ley, para "proteger a los trabajadores y asegurar su existencia, su salario, su capacidad productiva y la tranquilidad de la familia; contribuir al cumplimiento de un deber legal, de compromisos exteriores y de promesas gubernamentales". El Congreso aprobó la Iniciativa y el 19 de enero de 1943 se publicó en el Diario Oficial la Ley del Seguro Social.

Ahí se determina, desde los artículos iniciales, que la finalidad de la seguridad social es garantizar el derecho humano a la salud, la asistencia médica, la protección de los medios de subsistencia y los servicios sociales necesarios para el bienestar individual y colectivo. Como instrumento básico de la seguridad social se establece el Seguro Social y para administrarlo y organizarlo, se decreta la creación de un organismo público descentralizado, con personalidad y patrimonio propios, denominado Instituto Mexicano del Seguro Social.

El Instituto Mexicano del Seguro Social nace en 1943 en respuesta a las aspiraciones de la clase trabajadora. Actualmente, la Ley señala que la seguridad social tiene como finalidades el garantizar el derecho humano a la salud, la asistencia médica, la protección de los medios de subsistencia y los servicios sociales necesarios para el bienestar individual y colectivo, así como el otorgamiento de una pensión que, en su caso y previo cumplimiento de los requisitos legales, será garantizada por el Estado.

La Misión del Instituto Mexicano del Seguro Social es otorgar a los trabajadores mexicanos y a sus familias la protección suficiente y oportuna ante contingencias tales como la enfermedad, la invalidez, la vejez o la muerte.

El principal instrumento de la Seguridad Social es el Seguro Social, cuya organización y administración se encarga precisamente a la Institución llamada IMSS.

La protección se extiende no sólo a la salud, sino también a los medios de subsistencia, cuando la enfermedad impide que el trabajador continúe ejerciendo su actividad productiva, ya sea de forma temporal o permanente.

El propósito de los servicios sociales de beneficio colectivo y de las prestaciones fundamentales se orientan a incrementar el ingreso familiar, aprender formas de mejorar los niveles de bienestar, cultivar aficiones artísticas y culturales y hasta propiciar una mejor utilización del tiempo libre.

La Ley del Seguro Social expresa así todo lo anterior: "la Seguridad Social tiene por finalidad, garantizar el derecho humano a la salud, la asistencia médica, la protección de los medios de subsistencia y los servicios sociales necesarios para el bienestar individual y colectivo"[1].

La misión implica una decidida toma de postura en favor de la clase trabajadora y sus familiares; misión tutelar que va mucho más allá de la simple asistencia pública y tiende a hacer realidad cotidiana el principio de la solidaridad entre los sectores de la sociedad y del Estado hacia sus miembros más vulnerables.

Simultáneamente, por la misma índole de su encargo, el Instituto actúa como uno de los mecanismos más eficaces para redistribuir la riqueza social y contribuye así a la consecución de la justicia social en el país. Entre otras funciones, la labor institucional ayuda a amortiguar presiones sociales y políticas. Los trabajadores mexicanos consideran al IMSS como una de las conquistas definitivas después de muchos años de luchas sociales y como un patrimonio al que no están dispuestos a renunciar.

Este capítulo se estructura de la manera siguiente, en la primera parte se da una breve

[1] Ley del Seguro Social, Artículo 2 del Título Primero, disposiciones generales, capítulo único, 1995.

descripción de la evolución del tratamiento de los accidentes de trabajo, posteriormente se habla del caso del IMSS, y se hace mención a cuales son las diferencias entre los sistemas de Seguridad Social y las prestaciones otorgadas por el Instituto así como la cobertura en cuanto a ramas de seguridad, la tercera sección nos habla de cuál es el costo de la seguridad social en función de los gastos y la caracterización del seguro de accidentes de trabajo; se aborda la seguridad social y las características de los accidentes de trabajo, la cobertura que ofrece el IMSS en las diferentes ramas. Se define el concepto de población y las características de la misma con relación al trabajo presente, la metodología a seguir y la información a utilizar.

1.1 Evolución histórica de los accidentes de trabajo

Desde siempre ha existido la preocupación de los trabajadores ante los riesgos laborales, enfermedades, despido o retiro ya que se deben tener recursos monetarios que permitan brindar atenciones al trabajador que así lo requiera. Cuando se inicia la industrialización, el trabajo era realizado en ambientes complejos, donde los trabajadores estaban expuestos a numerosos accidentes ya sea por descuido propio, de los compañeros, del capataz o supervisor al trabajar con la maquinaria. Ante estos imprevistos, el trabajador se encontraba desprotegido y la posibilidad de obtener una indemnización era muy pequeña.

Por lo anterior, diferentes gobiernos evaluaron las consecuencias de la falta de atención a los trabajadores y con el surgimiento del concepto de seguridad social, se sentaron las bases para establecer que los accidentes de trabajo y las consecuencias que dichos accidentes llevaban consigo debian ser reparados de forma más justa. Esto se justifica con el argumento de que cualquier persona que se dedique a alguna actividad económica en donde se utilize ya sea maquinaria o el trabajo realizado por otros y por lo cual sea susceptible de sufrir lesiones accidentales, en consecuencia debe ser responsable de proporcionar a quién resulte afectado una indemnización sin que medie la pregunta acerca de quién debe imputarse la falta (ya sea al patrón, al trabajador o a un tercero).

En cuanto a las enfermedades profesionales, la protección se dió más tarde y para legislar en relación a dicha materia no fue sencillo identificar que enfermedades podían considerarse como tales, esto es, deberían excluirse enfermedades que son comunes para una población en general.

El Seguro de Accidentes de Trabajo tuvo diversas modificaciones con relación a sus origenes; entre las más importantes tenemos: el incremento a la cuantía en las pensiones, un aumento en las prestaciones monetarias y ampliación de las prestaciones en especie, la expansión de la cobertura hacia los trabajadores agrícolas asalariados quienes resultaron beneficiados con este ramo del seguro, además de la reclasificación de algunas empresas con base al riesgo registrado ante el IMSS

La reducción de las enfermedades y los accidentes de trabajo es una necesidad inaplazable dentro de los objetivos del IMSS, ya que ambos se encuentran entre las principales causas de mortalidad y morbilidad, muchos de los accidentes y enfermedades de trabajo se derivan de la falta de información, educación o cuidado.

1.2 Los accidentes de trabajo en México

Si entendemos el desarrollo económico como un medio para lograr el desarrollo social, entonces el crecimiento económico dependerá fundamentalmente de los recursos humanos con los que cuenta y de está manera la seguridad social juega un papel muy importante. La seguridad social debe cumplir con la misión de conservar y mejorar el índice de salud de los trabajadores ya que estos elevarán la productividad y por tanto el desarrollo.

Los únicos antecedentes verdaderos de la legislación moderna sobre aseguramiento de los trabajadores y de sus familiares, se encuentran a principios de este siglo en los últimos años de la época porfiriana, en dos disposiciones de rango estatal: la Ley de Accidentes de Trabajo del Estado de México, expedida el 30 de abril de 1904, y la Ley sobre Accidentes de Trabajo, del Estado de Nuevo León, expedida en Monterrey el 9 de abril de 1906. En estos dos ordenamientos legales se reconocía, por primera vez en el país, la obligación para los empresarios de atender a sus empleados en caso de enfermedad, accidente o muerte, derivados del cumplimiento de sus labores. Para 1915 se formuló un proyecto de Ley de Accidentes que establecía las pensiones e indemnizaciones a cargo del empleador, en el caso de incapacidad o muerte del trabajador por causa de un riesgo profesional.

La base constitucional del seguro social en México se encuentra en el artículo 123 de la Carta Magna promulgada el 5 de febrero de 1917. Ahí se declara "de utilidad social el establecimiento de cajas de seguros populares como los de invalidez, de vida, de cesación involuntaria en el trabajo, de accidentes y de otros con fines similares".

A finales de 1925 se presentó una iniciativa de Ley sobre Accidentes de Trabajo y Enfermedades Profesionales. En ella se disponía la creación de un Instituto Nacional de Seguros Sociales, de administración tripartita pero cuya integración económica habría de corresponder exclusivamente al sector patronal. También se definía con precisión la responsabilidad de los empresarios en los accidentes de trabajo y se determinaba el monto y la forma de pago de las indemnizaciones correspondientes. La iniciativa de seguro obrero suscitó la inconformidad de los empleadores que no estaban de acuerdo en ser los únicos contribuyentes a su sostenimiento y consideraban que también otros sectores deberían aportar. En 1929 el Congreso de la Unión modificó la fracción XXIX del artículo 123 constitucional para establecer que "se considera de utilidad pública la expedición de la Ley del Seguro Social y ella comprenderá seguros de Invalidez, de Vida, de Cesación Involuntaria del Trabajo, de Enfermedades y Accidentes y otros con fines análogos. Con todo, habrían de pasar todavía casi quince años para que la Ley se hiciera realidad.

La Seguridad Social es un conjunto de servicios y prestaciones que se otorgan a los individuos, estén o no afiliados al sistema de seguridad social del país -sin embargo es cierto que estos servicios no llegan a la totalidad del mismo de manera equitativa-; aunque en el campo de la seguridad social la palabra riesgo no indica necesariamente un daño -como en algún otro tipo de seguros- sino la necesidad de que alguna persona necesite los servicios de la seguridad social.

Si se hace referencia a cada uno de los ramos inscritos dentro de la seguridad social, se tiene:

1) enfermedades generales y maternidad,

2) vejez, invalidez y muerte,

3) accidentes y enfermedades profesionales,

4) desempleo,

5) asignaciones familiares.

En México, y en el caso muy particular del Instituto Mexicano del Seguro Social (IMSS), se tienen cinco ramos de seguro:

1) Enfermedades y maternidad; el cual otorga prestación de servicios médicos en especie y en dinero.

2) Invalidez y vida, siempre y cuando no sean ocasionados por un riesgo de trabajo; se otorgan prestaciones en dinero.

3) Retiro, Cesantía en edad avanzada y vejez; la prestación se da mediante servicio médico, en especie y en dinero, ya sea mediante una pensión al derechohabiente directo o en caso de muerte una pensión por viudez.

4) Riesgos de trabajo, se cubre al trabajador por cualquier percance que pueda ocurrirle por motivo de su trabajo. Este seguro cubre el servicio médico, pago de pensión si el trabajador quedara inhabilitado o pensión por viudez en caso de muerte.

5) Guarderias y prestaciones sociales, este ramo se encarga de la prestación de servicios sociales con el fin de fomentar la salud, prevención de enfermedades y accidentes, así como elevar el nivel de vida de la población mediante programas y servicios de asistencia, también se otorga el servicio de guarderia para los hijos de los derechohabientes.

Si se hace una comparación entre las ramas que integran la seguridad social (SS) y las ramas otorgadas por el IMSS se observa que no contempla las siguientes ramas:

1) Asignaciones familiares, donde puede considerarse el alimento y la vivienda como ejemplo de sus prestaciones.

2) Desempleo, el cual proveé una pequeña cantidad de dinero al trabajador que ha sido desempleado hasta encontrar un nuevo trabajo.

Ahora bien, la Seguridad Social tiene los siguientes tres campos de aplicación

1) Servicio público: prestación de servicio a cualquier persona que lo requiera, sin importar que se tenga o no relación de trabajo.

2) Asistencia social: prestación de servicios a personas de bajos ingresos, previa comprobación de ellos.

3) El seguro social: prestación de servicios sólo a aquellas personas que están calificadas para recibirlos ya sea por el tipo de empleo o por la zona de residencia.

En el caso del IMSS, existe un régimen voluntario, i. e., no importa que el trabajador no tenga derecho a ser asegurado ya que el interesado puede inscribirse dentro del seguro social (IMSS) para recibir los beneficio de éste.

Sectores ocupacionales cubiertos por el IMSS en México

	Comercial e industrial	Independiente(1)	Doméstico(1)	Rural(1),(2)	Empleos públicos
México	X	X	X	X	X

(1) Afiliación voluntaria

(2) Afiliación limitada a ciertas áreas, grupos de trabajadores o riesgos

El seguro social como campo de aplicación se inicia principalmente con una cobertura hacia grupos que son fáciles de administrar y financiar; y a medida que se desarrolla y estructura de una mejor manera, esta cobertura se va ampliando de manera gradual a otros grupos -esto puede observarse de manera muy clara en el IMSS-. Primeramente cubre a los trabajadores de industrias y servicios; luego a los asalariados (incluso a los empleados de la agricultura) y a los trabajadores por cuenta propia, y en la mayoría de los casos al asegurado y a sus dependientes.

Respecto al régimen del SS, conforme al artículo 4° de la Constitución que a la letra dice en el párrafo correspondiente: "...Toda persona tiene derecho a la protección de salud. La ley definirá las bases y modalidades para el acceso a los servicios de salud y establecerá la concurrencia de la Federación y las entidades federativas en materia de salubridad general"[2]

1.3 Costo de la seguridad social

El costo de la seguridad social se puede observar desde los siguientes puntos

1. Nivel de prestaciones de SS; este nivel puede examinarse desde
    1.1 Costo de prestaciones por habitante
    1.2 En relación al PIB.
2. Costo de los egresos; estos se clasifican en:
    2.1 Prestaciones en especie: servicios médicos
    2.2 Prestaciones en dinero: pensiones

[2] Constitución Política de los Estados Unidos Mexicanos, pág. 8.

2.3 Gastos en administración: salario del personal del organismo, gastos de matenimiento físico de las funciones administrativas.

3. Fuentes de financiamiento: el mecanismo más aplicado para el financiamiento de la SS es la política tripartita

3.1 Cotizaciones de obreros

3.2 Cotizaciones de patrones

3.3 Contribución del Estado.

La cotización que aporta el asegurado es como el pago por una mutualidad o un seguro, y en el caso de México particularmente está aportación tiene una característica similar al pago de un impuesto de manera individual.

En el caso de la cotización por parte de los patrones, esta tiene carácter de impuesto indirecto, ya que en el caso de los trabajadores, reciben a cambio de los servicios por parte de la SS y ayudan al desarrollo de la industria, ya que mantienen la capacidad productiva de los trabajadores. La contribución del Estado tiene como finalidad la distribución más equitativa de la riqueza, sin embargo en México, la SS no cubre al total de la población, por tanto esta finalidad se ve sesgada.

Por lo general, el costo de todos los riesgos es compartido por los tres sectores, sin embargo, el ramo de accidentes de trabajo es financiado sólo por el patrón del trabajador[3], en la nueva Ley del IMSS se ha eliminado la clasificación basada en grupos y grados de riesgo creando un nuevo sistema en el que cada empresa es evaluada por su propio historial de seguridad y casos de accidentes; así, cada industria, comercio o servicio determinará sus cuotas por sus propias condiciones de seguridad y el número de accidentes que en ella ocurran, de esta manera, cada empresa puede reducir sus cuotas mediante acciones encaminadas a la prevención de riegos laborales.

4. Gastos del gobierno central y nivel de bienestar social.

El bienestar social es una de las obligaciones aceptadas por el gobierno, que brinda protección a la población mediante los servicios de SS o algunos otros servicios sociales.

Los gastos por parte del gobierno pueden clasificarse de dos maneras:

[3] Nueva Ley del IMSS

4.1 Económica: agrupación según el carácter económico.

4.2 Funcional: agrupación según destino.

Sin embargo puede hacerse una combinación de las dos anteriores, esto es, su carácter económico y su función, esto permite ver los gastos como un objetivo particular. Las transacciones realizadas por el gobierno se dividen en tres grupos.

Transacciones del gobierno	-**Gastos corrientes**: gastos en bienes y servicios para consumo. -**Gastos en Capital**: gastos que afectan al patrimonio y tienen permanencia de larga duración. -**Gastos de transferencia**: gastos que realiza el gobierno como una manera de redistribuir el ingreso, puede considerarse como ejemplo los diferentes tipos de subsidio.

Los gastos funcionales son todos aquellos gastos que se realizan con una determinada finalidad.

Gastos funcionales	Fomento económico	
	Inversión y protección social	-servicios educativos y culturales -servicios asistenciales y hospitales -bienestar y seguridad social
	Ejército y armada	
	Administración general	
	Deuda Pública	

En las partidas de servicios asistenciales y hospitalarios así como la de bienestar y seguridad social se incluyen gastos como:

-salubridad

-asistencia médica y servicios hospitalarios

-construcciones hospitalarias

-maternidad y asistencia infantil

-asistencia social y otros servicios complementarios

-aportaciones a las instituciones de seguridad social

-etc.

muchos de estos gastos son erogaciones corrientes, principalmente de transferencias y servicios personales.

1.3.1 Características del seguro de accidentes de trabajo

La parte de accidentes de trabajo está contemplada en el seguro de riesgos de trabajo; los riesgos de trabajo son todos aquellos percances que puede sufrir un trabajador como resultado de una actividad propia de su actividad laboral; ahora, se considera accidente de trabajo a toda aquella lesión orgánica o funcional ya sea inmediata o posterior así como la muerte como resultado de un acto de trabajo. También son considerados accidente de trabajo, todas aquellas lesiones que sean causadas mientras el trabajador se dirige a su centro laboral o cuando se dirige a su domicilio desde su lugar de trabajo. Estas últimas lesiones no son consideradas al calcular el índice de siniestralidad de la empresa.

1.4 Conceptos demográficos

El estudio de los fenómenos sociales necesita de conceptos que permitan remitirse a estos cuando se desea estudiar algún fenómeno en particular, en nuestro caso algunas definiciones como población y trabajo serán de utilidad para el estudio de los accidentes de trabajo sufridos por la población trabajadora de empresas incorporadas al regimen de seguridad social IMSS. Estos conceptos serán definidos en las secciones siguientes. Se introduce el concepto de población a través del concepto manejado en economía, así como las características de la fuerza de trabajo, se define como se manejarán los grupos de la población a estudiar y finalmente se hace una enumeración rápida de las fuentes de información a consultar.

1.4.1 Población

En el estudio de la población se puede encontrar las dos siguientes ramas de estudio

$$\text{Población} \begin{cases} \text{PRFT: proceso de reproducción de la fuerza de trabajo} \\ \text{Proceso de reproducción de la población} \end{cases}$$

Además agregando el factor conocido como el proceso directo de producción, definido como la reproducción del conjunto de procesos sociales que constituyen el modo de producción mismo; éste proceso se divide en

Proceso directo de producción (PDP)	-Procesos de transformación -Aplicación de energía humana por diversos medios de trabajo (FT) -Relaciones de producción

de aquí, se puede decir que existen efectos socialmente útiles (esto es, para la propia supervivencia), y procesos de intercambio que son llamados también procesos de circulación (PC). Los PDP y PC constituyen el proceso económico De esta manera si se establece el concepto de relaciones de trabajo entonces se puede determinar lo que se considera trabajo y lo que no, así como lo que son accidentes y enfermedades de trabajo y lo que no es considerado como tal.

Ahora, con la pregunta siguiente ¿agota la reproducción poblacional el proceso de reproducción de la FT? entonces se define primeramente el concepto de FT como aquella capacidad de gastar la energía humana ya sea mental o física, en cualquier proceso de trabajo. El proceso de trabajo es el proceso de transformación de un objeto en un producto mediante una actividad utilizando instrumentos de trabajo.

Se debe aclarar que en un principio la capacidad de la producción depende de manera fundamental del adiestramiento, esto es, de lo que se llama enseñanza,denominado en la actualidad Capital Humano, ya sea de manera formal e informal y que constituye de alguna forma la estructura de la actividad. Ahora, en términos demográficos, el proceso de reproducción poblacional (esto es, procesos que afectan la existencia física) comprende los siguientes fenómenos.

1. Mantenimiento.
2. Descanso (reposición de fuerzas físicas).
3. Aprendizaje básico (socialización).
4. Morbilidad.
5. Mortalidad.

Además existen dos procesos útiles en la producción de la fuerza de trabajo (FT), a saber:

1. Trabajos directamente destinados a la producción de fuerza de trabajo
2. Trabajos no directamente destinados a la producción de fuerza de trabajo.

El Proceso de Reproducción de la Fuerza de Trabajo (PRFT) puede dividirse en tres procesos básicos, los cuales a su vez comprenden fenómenos sociales.

- Proceso de reproducción de la FT
  - Proceso de reproducción poblacional
    - -Nupcialidad
    - -Fecundidad
    - -Mantenimiento
    - -Descanso
    - -Aprendizaje básico
    - -Morbilidad
    - -Mortalidad
  - Proceso de reproducción compleja
    - -Educación formal
    - -Educación no formal
  - Proceso de ajuste
    - -Migraciones de todo tipo
    - -Adscripciones dentro de la división social del trabajo (empleadores, asalariados, etc.; diversos procesos ideológico, político, etc.).
    - -Adscripciones dentro de la división del trabajo social (diversos sectores, paro, etc.)

Fuente: Leguina, Joaquin; Fundamentos de Demografía, pág. 12

Una parte de los procesos sociales que intervienen ya sea directa o indirectamente en el PRFT son protagonizados por el individuo-soporte, la otra parte la constituye lo que se denomina entorno social. A continuación se tiene un cuadro resumen para clasificar los fenómenos mencionados y su implicación como consecuencia o no del trabajo.

Procesos	Agentes: Individuo soporte de la fuerza de trabajo	Agentes: Entorno social
Proceso de trabajo	Educación formal Educación no formal Migración ADST[1] ADTS[2]	Mantenimiento Aprendizaje básico Morbilidad Mortalidad Educación formal ADTS[2]
Proceso de no trabajo	Mantenimiento Descanso Aprendizaje básico Morbilidad Mortalidad	Nupcialidad Fecundidad

[1] Adscripción dentro de la división del trabajo
[2] Adscripción dentro de la división del trabajo social

Fuente: Leguina, Joaquin; Fundamentos de Demografía., pág. 13.

Si observamos el cuadro anterior, la columna de enmedio toma en cuenta los fenómenos que forman parte del propio PRFT, la última columna clasifica eventos ajenos al individuo.

1.4.2 Estrategia metodológica

Si entendemos por población el concepto utilizado por Marx, en donde dicho concepto no es más que una abstracción si se dejan de lado las **clases** por las cuales se compone, entonces se considera como población al conjunto de relaciones diversas con una totalidad de determinantes. De esta manera, se puede decir que no es posible trabajar con teorías demográficas aisladas, sino que, muy por el contrario, deben insertarse todas esas teorías en un marco mucho más general.

Si esto es verdad, el PRFT, tiene como finalidad la supervivencia del grupo al cual pertenece, entonces la teoría demográfica necesita de dos trabajos teóricos y previos.

a) La delimitación de grupos sociales significativos.

b) Determinación de las necesidades **sociales** de la fuerza de trabajo y de qué forma se **estimula** su producción dentro de cada grupo significativo.

De esta manera se puede realizar el análisis de respuestas de dichos grupos ante diferentes estímulos.

La importancia de delimitar es, sin duda, que dada la condición de la **situación social** o dicho de otra forma, dadas las prácticas sociales de los elementos del grupo, estos tienden a tener comportamientos homógeneos. Sin embargo un problema grave es la delimitación de dichos grupos, lo que en ocasiones lleva a realizar la enumeración de todos los miembros de la población. Es aquí donde pueden surgir las diversas estrategias metodológicas a seguir.

Un método para elegir los grupos más homógeneos es aquel en donde el grupo tiene la menor varianza entre las variables elegidas (estadísticamente) y a partir de estos se escogería el mejor método de análisis (en estos casos puede elegirse el análisis estadístico multivariado: correlación múltiple, análisis de varianzas, análisis factorial, etc.); sin embargo, se debe aplicar primeramente el enfoque **grupal**, entendido por el proceso de probar si el individuo **x** pertenece o no al grupo **A**, y esto puede realizarse con una dicotomía:

1) se tiene o no un comportamiento homógeneo frente al fenómeno de estudio del grupo,

2) se detectan o no diferencias significativas entre el comportamiento del grupo y el resto de la población.

De esta manera podemos encontrar 4 tipos de resultados que se resumen en el cuadro siguiente:

	Diferencias significativas	
Homogeneidad interna	Si	No
Si	1	2
No	3	4

Para el estudio que nos ocupará, la población queda determinada por los trabajadores que cotizan al IMSS y que presentaron algún accidente de trabajo con incapacidad permanente. La población presenta homogeneidad interna y no tiene diferencias significativas ya que se cubren de la misma manera los accidentes de trabajo en las diferentes empresas.

1.4.3 Sistema de información

En Demografía como en la mayoría de las Ciencias Sociales, los datos obtenidos a través de la observación experimental son casi nulos y muchas veces se debe conformar con el contraste de hipótesis aunque también muchas de éstas podrían ser fácilmente comprobables sin utilizar métodos estadísticos propiamente dichos. Al utilizar los datos de las variables a introducir al fenómeno en estudio, se tiene que, en ocasiones es casi imposible (o en algunos casos imposible) separar estadísticamente los efectos de dichas variables entre sí, y a lo más que se puede aspirar es a una descripción correcta de los procesos sociales que junto a unas hipótesis razonables pueden explicar coherentemente las interrelaciones del fenómeno en estudio.

Sin embargo, con todo y dichas limitaciones, se debe crear un sistema de información capaz de cubrir las necesidades de dicho estudio aún cuando las fuentes de información sean inagotables pero poco aplicadas al fenómeno. En este sentido, los datos a ocupar serán, series estadísticas entre ellas las publicadas por el INEGI, por el IMSS, Informes de Gobierno de diferentes años, entre otros; que proporcionan información relacionada y relevante para el fenómeno en estudio.

Consideraciones

Los temas tratados anteriormente son sólo con la finalidad de obtener un marco general de referencia para tratar el tema de la seguridad social en México; el marco conceptual de este estudio es la demografía y como tal, se ha considerado necesario hacer el presente capítulo como una introducción a todos los puntos de vista desde donde puede estudiarse a los accidentes profesionales.

Capítulo 2

Metodología

La econometría es la aplicación de métodos matemáticos y estadísticos al análisis de datos y hechos ya sea económicos o sociales. Los métodos estadísticos ayudan a resumir datos, estimar parámetros de algún modelo que sea de nuestro interés e intrepretar la fuerza de la evidencia de las hipótesis que se quieren probar. La econometría se centra, por tanto en la estimación de algún comportamiento en particular validándolo mediante métodos estadísticos.

El capítulo se estructura de la manera siguiente: se describen las características del Modelo de Regresión Lineal, los supuestos y características de los estimadores. Se hace una deducción de cada estimador y sus propidades así como las pruebas de hipótesis a trabajar en cada supuesto.

2.1 Antecedentes del Modelo de Regresión Lineal Clásico

El objetivo primordial de la econometría consiste en especificar y estimar un modelo de relación entre variables ya sea económicas o de algún otro tipo. Ya sea de manera teórica o de manera empírica se específica un conjunto de relaciones entre las variables, además de hacer intentos para poder validar estas proposiciones con relación al comportamiento de los datos observados. La estimación de dichas relaciones se hace a partir de la información muestral acerca de las diferentes variables que se desean introducir en la estimación y se trata entonces de cuantificar la magnitud de la dependencia entre ellas

2.1.1 Especificación del Modelo de Regresión

Como se mencionó anteriormente la teoría sugiere el conjunto de relaciones entre las variables, es decir, el punto de partida para especificar una regresión. El análisis de regresión conduce a la estimación de los parámetros de un modelo, utilizando datos sobre sus variables. Generalmente una regresión contiene tres elementos básicos:

- Un Modelo Lineal
- Un conjunto de datos (generalmente una muestra y no una población)
- Un método de estimación, esto es, un método para transformar los datos y obtener estimaciones de la relación entre las diversas variables.

2.1.2 Modelo de Regresión Lineal

En el caso de la regresión lineal simple se tiene un modelo de la siguiente forma

$$\mathbf{Y}_t = \beta_0 + \beta_1 \mathbf{X}_t$$

donde **Y** es una variable denominada variable dependiente, explicada o endógena; **X** es la variable independiente, explicativa o exogéna y se supone no estocástica (fija), en las aplicaciones prácticas se dispone de una muestra y por el modelo se sugiere que esta relación se cumple para cada par de observaciones y al realizar las estimaciones se obtienen los valores de β_0 y β_1 los cuales permitirán realizar contrastación de hipótesis y donde β_1 indicará la variación que experimenta la variable endógena al registrar cambios a lo largo del ciclo, la variable exógena. Los coeficientes β_0 y β_1 reciben el nombre de intercepto y pendiente del modelo de regresión lineal simple respectivamente. El subíndice t indica que será un modelo a estimar con datos de series temporales.

Componentes del modelo de regresión

No es fácil la obtención de los coeficientes del modelo lineal simple a partir de una muestra de tamaño T. Esto se debe a que no se pueden estimar los verdaderos parámetros de la ecuación debido a que se trata sólo de una muestra y no del total de la población, es por ésta razón que el modelo queda determinado de la siguiente manera

$$\mathbf{Y}_t = \beta_0 + \beta_1 \mathbf{X}_t + u_t \tag{1}$$

donde el último componente $\mathbf{u}_t$ se denomina perturbación estructural o término de error del modelo de regresión no observable, este componente permite explicar las diferencias entre los dos miembros de la igualdad. Para la perturbación aleatoria no se tienen datos disponibles, esto es, no puede medirse por lo cual se supone que no esta correlacionada con la variable X_t. La interpretación de dicho componente es variada, entre estos se tienen:

a) No se pueden encontrar otras variables que expliquen a la variable endógena, o no pueden, o no es fácil encontrar tales variables.

b) Aún cuando se pueda especificar las variables o se conozcan cuáles son las variables que tienen una alta influencia en la variable endógena no se dispone de datos.

c) El término de error puede reflejar errores de medida en la variable endógena, que suelen surgir porque las variables que se utilizan en la estimación reflejan de manera aproximada pero no exacta los conceptos que se quieren incorporar en el modelo.

Una vez que han sido estimados los coeficientes β_0 y β_1 en (1) se tiene una ecuación lineal, una recta que se denomina recta de regresión. Si despejamos el componente de error al realizar la estimación se tiene

$$\widehat{u}_t = \mathbf{Y}_t - \widehat{\beta}_0 - \widehat{\beta}_1 \mathbf{X}_t$$

cuyo valor puede ser positivo o negativo; este valor será positivo si el valor de la variable endógena excede del estimado por la recta; cuando el resultado es negativo el valor de la variable endógena es inferior al estimado. La $\widehat{u}_t$ se conoce como residuo de la observación muestral.

El término de error es una variable aleatoria, diferente para cada observación muestral, y su realización no es observable; el residuo es observable ya que se construye a partir de las estimaciones y de los datos de las variables exógena y endógena. En consecuencia, el término de error y el residuo son elementos diferentes.

Desde el punto de vista estadístico, el modelo de regresión no tiene una connotación de causalidad en la relación entre variables; esto es, puede hacerse una estimación de $\mathbf{Y}$ con respecto a $\mathbf{X}$ o de $\mathbf{X}$ con respecto de $\mathbf{Y}$, sin embargo el análisis no será el mismo ya que una variable exógena se considera determinística, mientras que la variable endógena se considera aleatoria. El hecho de estimar una variable en términos de la otra dependerá del tipo de estudio que se quiera hacer y la relación que exista entre ambas variables.

2.1.3 Supuestos del Modelo Lineal Clásico

En el Modelo Lineal Clásico se tienen 8 supuestos que deben observarse.

1. Linealidad de las variables. Supone que el modelo lineal es el que más se ajusta a la realidad.

2. Linealidad en los parámetros.

3. Esperanza matemática igual a cero; esto es, se supone que la esperanza del término de error u_i es $E(u_i) = 0$, $\forall i = 1, 2, 3, \ldots, n$. Para el caso donde $E(u_i) = c$ con $c \neq 0$ se tendría un efecto constante sobre $\mathbf{Y}_i$ que debería ser reflejado por el término β_0.

Con relación a este supuesto, debe considerarse que se tendrán problemas de este tipo cuando se omiten variables relevantes, esto es, si el modelo que mejor se ajusta es de la forma

$$\mathbf{Y} = \beta_0 + \beta_1 \mathbf{X}_1 + \beta_2 \mathbf{X}_2 + u$$

y se estima un modelo de la forma

$$\mathbf{Y} = \beta_0 + \beta_1 \mathbf{X}_1 + v$$

en donde se ha omitido la variable $\mathbf{X}_2$, entonces el término de error tendrá la forma $v = u + \beta_2 \mathbf{X}_2$, y al obtener la esperanza se tiene entonces $E(u + \beta_2 \mathbf{X}_2) = E(u) + E(\beta_2 \mathbf{X}_2) = 0 + \beta_2 E(\mathbf{X}_2)$, donde $E(\mathbf{X}_2)$ indica la esperanza que toman los valores de la variable omitida y que se supone constante a través del tiempo; esto implica que $E(v) \neq 0$ en general.

4. Homoscedasticidad o varianza constante, se supone que la varianza del término de error, que se denota como $var(u_i) = \sigma_u^2$ $\forall i = 1, 2, 3, \ldots, n.$, es constante a lo largo de toda la muestra.

5. Ausencia de autocorrelación, supone que los términos de error correspondiente a dos observaciones cualquiera de la muestra, son dos variables aleatorias diferentes y que no están correlacionadas estadísticamente.

6. Estabilidad temporal, esto es, los coeficientes β_0 y β_1 son constantes en el tiempo, así

como que el modelo propuesto es válido para todo el periodo de estudio.

7. Causalidad unidireccional, se refiere a que la relación causal de **X** sobre **Y** se cumple sólo en ese sentido y no al revés. Esto se basa en la naturaleza del análisis y la relación teórica que existe entre dichas variables.

8. Variables exógenas deterministas. Todas las variables exógenas deben ser deterministas, ya que la variable endógena se verá afectada por el término de error u.

2.2 Estimador de Mínimos Cuadrados Ordinarios

En este apartado se tratará el uso del estimador por diferentes métodos y sus características.

Sea el modelo

$$y_i = \beta_0 + \beta_1 x_1 + u_i \tag{1}$$

donde se estiman los coeficientes y se tiene entonces

$$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 \tag{2}$$

donde $\widehat{\beta}_0$ y $\widehat{\beta}_1$ han sustituido a los parámetros poblacionales que eran desconocidos; (2) representa la estimación de la variable y al realizar el modelo econométrico. Como existe una variación entre el valor real observado y el valor obtenido por la estimación, se define entonces

$$\widehat{u}_i = y_i - \widehat{y}_i$$

como el residuo correspondiente a la observación muestral.

Lo que se desea entonces es minimizar el valor de $\widehat{u}_i$, pero como este valor no es único sino un conjunto de residuos, se tiene entonces que se debe minimizar el conjunto de tales residuos.

2.2.1 Método de Momentos

Los supuestos sobre el residuo (u) fueron

$$E(u) = 0 \qquad cov(x - u) = 0$$

en este método se usará la contraparte muestral.

Si $\widehat{\beta}_0$ y $\widehat{\beta}_1$ son los estimadores de β_0 y β_1 respectivamente, la contraparte muestral del residuo u_i es $\widehat{u}_i$, entonces

$$\widehat{u}_i = y_i - \widehat{\beta}_0 + \widehat{\beta}_1 x_1$$

Ahora, las ecuaciones que se necesitan para estimar $\widehat{\beta}_0$ y $\widehat{\beta}_1$ se obtienen de sustituir los supuestos poblacionales muestrales

	Supuesto poblacional	Supuesto muestral		
(1)	$E(u) = 0$	$\frac{1}{n}\sum \widehat{u}_i = 0$	ó	$\sum \widehat{u}_i = 0$
(2)	$cov(x - u) = 0$	$\frac{1}{n}\sum x_i\widehat{u}_i = 0$	ó	$\sum x_i\widehat{u}_i = 0$

de (1) y (2) se tiene

$$\sum_{i=1}^{n} \widehat{u}_i = 0 \text{ entonces } \sum_{i=1}^{n}\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) = 0 \qquad (1.1)$$

$$\sum_{i=1}^{n} x_i\widehat{u}_i = 0 \text{ entonces } \sum_{i=1}^{n} x_i\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) = 0 \qquad (2.1)$$

De (1.1) se tiene

$$\sum_{i=1}^{n}\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) = 0$$

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\widehat{\beta}_0 - \sum_{i=1}^{n}\widehat{\beta}_1 x_i = 0$$

$$\sum_{i=1}^{n} y_i - n\widehat{\beta}_0 - \widehat{\beta}_1\sum_{i=1}^{n} x_i = 0$$

$$\sum_{i=1}^{n} y_i = n\widehat{\beta}_0 + \widehat{\beta}_1\sum_{i=1}^{n} x_i \qquad (3)$$

De (2.1) se tiene

$$\sum_{i=1}^{n} x_i \left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) = 0$$
$$\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \widehat{\beta}_0 - \sum_{i=1}^{n} x_i^2 \widehat{\beta}_1 = 0$$
$$\sum_{i=1}^{n} x_i y_i - \widehat{\beta}_0 \sum_{i=1}^{n} x_i - \widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\sum_{i=1}^{n} x_i y_i = \widehat{\beta}_0 \sum_{i=1}^{n} x_i + \widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 \tag{4}$$

se tiene de (3) y (4) un sistema de la forma

$$\begin{aligned} \sum_{i=1}^{n} y_i &= n\widehat{\beta}_0 + \widehat{\beta}_1 \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i &= \widehat{\beta}_0 \sum_{i=1}^{n} x_i + \widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 \end{aligned} \tag{5}$$

Resolviendo el sistema (5), se obtienen los valores para $\widehat{\beta}_0$ y $\widehat{\beta}_1$. Los valores minimizan la suma de los residuos.

2.3 Método de Mínimos Cuadrados Ordinarios

Se utiliza como criterio la minimización de la suma de los cuadrados de los residuos, llamada también suma residual (SR), entonces se deben encontrar los valores β_0 y β_1 que resuelvan el problema

$$\min_{\beta_0,\beta_1} SR = \sum \widehat{u}_i^2$$

recordemos que $\widehat{u}_i = y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i$, entonces

$$\min_{\beta_0,\beta_1} SR = \sum \left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right)^2$$

se obtienen las derivadas parciales con respecto a β_0 y β_1 y se obtiene

$$\begin{aligned}
\frac{\partial SR}{\partial \beta_0} &= 2\sum_{i=1}^{n}(-1)\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) \\
&= -2\sum_{i=1}^{n}\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) \\
\frac{\partial SR}{\partial \beta_1} &= 2\sum_{i=1}^{n}\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right)(-x_i) \\
&= -2\sum_{i=1}^{n}\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right)x_i
\end{aligned}$$

la matriz de segundas derivadas (Hessiano) se obtiene de la siguiente forma

$$\begin{aligned}
\frac{\partial^2 SR}{\partial \beta_0^2} &= -2\sum_{i=1}^{n}(-1) = 2n \\
\frac{\partial^2 SR}{\partial \beta_0 \beta_1} &= -2\sum_{i=1}^{n} -x_i = 2\sum_{i=1}^{n} x_i \\
\frac{\partial^2 SR}{\partial \beta_1 \beta_0} &= -2\sum_{i=1}^{n}(-1)(x_i) = 2\sum_{i=1}^{n} x_i \\
\frac{\partial^2 SR}{\partial \beta_1^2} &= -2\sum_{i=1}^{n}(-1)x_i^2 = 2\sum_{i=1}^{n} x_i^2
\end{aligned}$$

por tanto, la matriz queda determinada por

$$\begin{pmatrix} 2n & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2\sum_{i=1}^{n} x_i^2 \end{pmatrix}$$

Utilizando los conceptos acerca de convexidad para funciones de varias variables se tiene que debe cumplirse

(i) $\dfrac{\partial^2 f(x_1x_2)}{\partial x_1^2}\dfrac{\partial^2 f(x_1x_2)}{\partial x_2^2} - \left[\dfrac{\partial^2 f(x_1x_2)}{\partial x_1 \partial x_2}\right]^2 \geq 0$

(ii) $\dfrac{\partial^2 f(x_1x_2)}{\partial x_1^2} \geq 0$

(iii) $\dfrac{\partial^2 f(x_1x_2)}{\partial x_2^2} \geq 0$

Obsérvese que

$(i)\ (2n)\left(2\sum_{i=1}^{n}x_i^2\right)-4\left(\sum_{i=1}^{n}x_i\right)^2=4n\sum_{i=1}^{n}x_i^2-4\left(\sum_{i=1}^{n}x_i\right)^2$

de donde

$\sum_{i=1}^{n}x_i^2>\left(\sum_{i=1}^{n}x_i\right)^2$

como $4n$ es positivo y $4n>4$ entonces

$4n\sum_{i=1}^{n}x_i^2>4\left(\sum_{i=1}^{n}x_i\right)^2$

por tanto

$4n\sum_{i=1}^{n}x_i^2-4\left(\sum_{i=1}^{n}x_i\right)^2\geq 0$

$(ii)\ \dfrac{\partial^2 SR}{\partial\beta_0^2}=2n\quad \text{y}\quad 2n\geq 0$

$(iii)\ \dfrac{\partial^2 SR}{\partial\beta_1^2}=2\sum_{i=1}^{n}x_i^2\quad \text{y}\quad \sum_{i=1}^{n}x_i^2\geq 0\Rightarrow 2\sum_{i=1}^{n}x_i^2\geq 0.$

Este resultado puede generalizarse para una función de varias variables. La función será convexa si y sólo si la matriz hessiana de $n\times n$ es semidefinida positiva para todos los valores posibles de $(x_1,x_2,\ldots,x_n)$, entonces puede concluirse que la solución al sistema

$$\begin{aligned}-2\sum_{i=1}^{n}\left(y_i-\widehat{\beta}_0-\widehat{\beta}_1x_i\right)&=0\\ -2\sum_{i=1}^{n}\left(y_i-\widehat{\beta}_0-\widehat{\beta}_1x_i\right)x_i&=0\end{aligned}\tag{6}$$

son los valores β_0 y β_1, además de que se alcanza un mínimo en la función Suma Residual. Al resolver el sistema se tiene

$\sum_{i=1}^{n}\left(y_i-\widehat{\beta}_0-\widehat{\beta}_1x_i\right)=0$

$\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}\widehat{\beta}_0-\sum_{i=1}^{n}\widehat{\beta}_1x_i=0$

$\sum_{i=1}^{n}y_i-n\widehat{\beta}_0-\widehat{\beta}_1\sum_{i=1}^{n}x_i=0$

$$\sum_{i=1}^{n}y_i=n\widehat{\beta}_0+\widehat{\beta}_1\sum_{i=1}^{n}x_i\tag{7}$$

$\sum_{i=1}^{n}\left(y_i-\widehat{\beta}_0-\widehat{\beta}_1x_i\right)x_i=0$

$\sum_{i=1}^{n}y_ix_i-\sum_{i=1}^{n}\widehat{\beta}_0x_i-\sum_{i=1}^{n}\widehat{\beta}_1x_i^2=0$

$$\sum_{i=1}^{n} y_i x_i - \widehat{\beta}_0 \sum_{i=1}^{n} x_i - \widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\sum_{i=1}^{n} y_i x_i = \widehat{\beta}_0 \sum_{i=1}^{n} x_i + \widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 \tag{8}$$

(7) y (8) constituyen un sistema de ecuaciones simúltaneas en las incógnitas $\widehat{\beta}_0$ y $\widehat{\beta}_1$.

$$\sum_{i=1}^{n} y_i = n\widehat{\beta}_0 + \widehat{\beta}_1 \sum_{i=1}^{n} x_i \tag{91}$$

$$(9)$$

$$\sum_{i=1}^{n} y_i x_i = \widehat{\beta}_0 \sum_{i=1}^{n} x_i + \widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 \tag{9.2}$$

El sistema (9) es conocido como Sistema de Ecuaciones Normales.

Si se despeja $\widehat{\beta}_0$ de (7) se tiene

$$\widehat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i - \widehat{\beta}_1 \sum_{i=1}^{n} x_i}{n} = \overline{y} - \widehat{\beta}_1 \overline{x} \tag{10}$$

que puede utilizarse para obtener el estimador de MCO de β_0 cuando se tenga al estimador β_1.

Despejando $\widehat{\beta}_1$ de (9.2) se tiene

$$\widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i - \widehat{\beta}_0 \sum_{i=1}^{n} x_i$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \widehat{\beta}_0 \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2}$$

sustituyendo el valor de $\widehat{\beta}_0$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \frac{1}{n}\left(\sum_{i=1}^{n} y_i - \widehat{\beta}_1 \sum_{i=1}^{n} x_i\right) \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \frac{1}{n} \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i - \frac{\widehat{\beta}_1}{n} \left(\sum_{i=1}^{n} x_i\right)^2}{\sum_{i=1}^{n} x_i^2}$$

$$\widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i - \frac{1}{n}\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i - \frac{\widehat{\beta}_1}{n}\left(\sum_{i=1}^{n} x_i\right)^2$$

$$\widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 + \frac{\widehat{\beta}_1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 = \sum_{i=1}^{n} y_i x_i - \frac{1}{n}\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i$$

$$\widehat{\beta}_1 \left[\sum_{i=1}^{n} x_i^2 + \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right] = \sum_{i=1}^{n} y_i x_i - \frac{1}{n}\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \frac{1}{n}\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 + \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \overline{y}\overline{x}}{\sum_{i=1}^{n} x_i^2 + \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

$$\widehat{\beta}_1 = \frac{S_{xy}}{S_x^2} = \rho_{xy}\frac{S_y}{S_x}$$

donde S_{xy} es la covarianza, S_x^2 es la varianza de x, y S_x, S_y son las desviaciones de x e y.

(10) y (11) son las estimaciones por MCO como función de los estadísticos muestrales, sin tener que resolver el sistema (9). Se calcula $\widehat{\beta}_1$ y posteriormente $\widehat{\beta}_0$. Con esto se puede comprobar que una propiedad de los MCO se cumple; la recta estimada pasa por $(\overline{y}, \overline{x})$.

Las ecuaciones anteriores, recordando las igualdades con las que se ha trabajado pueden escribirse como

$$\sum_{i=1}^{n} \widehat{u}_i = 0 \tag{12}$$

$$\sum_{i=1}^{n} x_i \widehat{u}_i = 0 \tag{13}$$

que son las dos propiedades del estimador de los MCO. Recordando se tiene

1) la suma de los residuos que son generados por la estimación de mínimos cuadrados es igual a cero (esta propiedad no puede garantizarse con otros métodos)

2) Los residuos de mínimos cuadrados no están relacionados con la variable exógena del modelo.

Cuando es considerado un modelo con K variables exógenas, esto es, un modelo de regresión lineal general o múltiple, los residuos de mínimos cuadrados no están relacionadas con ninguna de las variables exógenas.

2.3.1 Esperanza Matemática

La expresión del estimador de la pendiente del modelo de regresión lineal simple por MCO puede ser escrita como

$$\begin{aligned} \widehat{\beta}_1 &= \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2} = \frac{\sum_{i=1}^{n}(x_i-\overline{x})\,y_i}{\sum_{i=1}^{n}(x_i-\overline{x})^2} = \sum_{i=1}^{n}\left[\frac{\sum_{i=1}^{n}(x_i-\overline{x})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right] y_i \qquad (14)\\ \widehat{\beta}_1 &= \sum_{i=1}^{n} a_i y_i \end{aligned}$$

que es una combinación lineal ponderada de las observaciones de la variable endógena, en donde la ponderación está definida por

$$a_i = \frac{x_i-\overline{x}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

donde se ha utilizado el hecho de que la suma de las desviaciones de una variable con relación a la media muestral es siempre igual a cero. Entonces la suma de las ponderaciones es igual a cero.

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n}\left[\frac{x_i-\overline{x}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right] = \frac{\sum_{i=1}^{n}(x_i-\overline{x})}{\sum_{i=1}^{n}(x_i-\overline{x})^2} = 0 \quad \text{y además}$$

$$\begin{aligned} \sum_{i=1}^{n} a_i x_i &= \sum_{i=1}^{n}\left[\frac{(x_i-\overline{x})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right] x_i \\ &= \frac{\sum_{i=1}^{n}(x_i-\overline{x})\,x_i}{\sum_{i=1}^{n}(x_i-\overline{x})^2} = \frac{\sum_{i=1}^{n}x_i^2-\overline{x}\sum_{i=1}^{n}x_i}{\sum_{i=1}^{n}(x_i-\overline{x})^2} \end{aligned}$$

$$
\begin{aligned}
&= \frac{\sum_{i=1}^{n} x_i^2 - \overline{x}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n}\left(x_i^2 - 2x_i\overline{x} + \overline{x}^2\right)} = \frac{\sum_{i=1}^{n} x_i^2 - \overline{x}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - 2\overline{x}\sum_{i=1}^{n} x_i + \overline{x}^2\sum_{i=1}^{n} 1} \\
&= \frac{\sum_{i=1}^{n} x_i^2 - \overline{x}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - 2\overline{x}\sum_{i=1}^{n} x_i + n\overline{x}^2} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2} = 1
\end{aligned}
$$

Debe recordarse el supuesto de que los valores x_i son fijos. Sólo las observaciones de la variables endógena Y serán diferentes de las disponibles actualmente, esto debido al componente de perturbación aleatoria u_i.

Si se sustituye el valor de y_i según la expresión del modelo de regresión en (14) se tiene

$$
\begin{aligned}
\widehat{\beta}_1 &= \sum_{i=1}^{n} a_i\left(\beta_0 + \beta_1 x_i + u_i\right) = \sum_{i=1}^{n} a_i\beta_0 + \sum_{i=1}^{n} a_i\beta_1 x_i + \sum_{i=1}^{n} a_i u_i \\
&= \beta_0\sum_{i=1}^{n} a_i + \beta_1\sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} a_i u_i \\
&= \beta_0(0) + \beta_1(1) + \sum_{i=1}^{n} a_i u_i \\
\widehat{\beta}_1 &= \beta_1 + \sum_{i=1}^{n} a_i u_i \qquad (15)
\end{aligned}
$$

Aplicando esperanzas a (15)

$$
\begin{aligned}
E\left(\widehat{\beta}_1\right) &= E\left(\beta_1\right) + E\left(\sum_{i=1}^{n} a_i u_i\right) = \beta_1 + \sum_{i=1}^{n} a_i E\left(u_i\right) \\
&= \beta_1 + \sum_{i=1}^{n} a_i(0) = \beta_1 \quad \text{entonces} \\
E\left(\widehat{\beta}_1\right) &= \beta_1
\end{aligned}
$$

por tanto el estimador de MCO para β_1 es insesgado.

Si obtenemos el promedio del modelo de regresión y se obtienen las esperanzas

$$
\overline{y} = \beta_0 + \beta_1\overline{x} + \overline{u}
$$

$$
\begin{aligned}
E(\overline{y}) &= E(\beta_0) + E(\beta_1\overline{x}) + E(\overline{u}) \\
E(\overline{y}) &= \beta_0 + \beta_1 E(\overline{x})
\end{aligned}
$$

Si se recuerda la expresión del estimador β_0 por MCO se tiene

$$
\begin{aligned}
\widehat{\beta}_0 &= \overline{y} - \beta_1\overline{x} \\
E\left(\widehat{\beta}_0\right) &= E(\overline{y}) - E(\beta_1\overline{x}) \\
&= \beta_0 + \beta_1 E(\overline{x}) - \beta_1 E(\overline{x}) \quad \text{entonces} \\
E\left(\widehat{\beta}_0\right) &= \beta_0
\end{aligned}
$$

por tanto, el estimador de β_0 por MCO es insesgado.

2.3.2 Matriz de covarianzas para el estimador

Un estimador puntual tiene asociado una medida de dispersión del mismo y en general la medida está dada por la varianza, que indica el grado en que se aproxima al verdadero valor del parámetro que se pretende estimar.

Otra ventaja de la medida de dispersión es el hecho de realizar pruebas sobre algunos o ambos parámetros de la estimación; esto es, si algún coeficiente toma determinados valores teóricos.

Se parte entonces de la expresión del estimador β_1

$$
\widehat{\beta}_1 = \beta_1 + \sum_{i=1}^{n} a_i u_i \tag{16}
$$

se debe observar que

$$
\begin{aligned}
E(a_i u_i) &= a_i E(u_i) = a_i(0) = 0 \\
Var(a_i u_i) &= a_i^2 var(u_i) = a_i^2 \sigma_u^2 \qquad \forall i = 1, 2, \ldots, n
\end{aligned}
$$

y dado que la $cov(u_i, u_j) = 0$ entonces

$$\begin{aligned} E\left(\sum_{i=1}^{n} a_i u_i\right) &= \sum_{i=1}^{n} a_i E(u_i) = \sum_{i=1}^{n} a_i (0) = 0 \\ Var\left(\sum_{i=1}^{n} a_i u_i\right) &= \sum_{i=1}^{n} var(a_i u_i) = \sigma_u^2 \sum_{i=1}^{n} a_i^2 \end{aligned}$$

al calcular la suma de los cuadrados de las ponderaciones se tiene

$$\begin{aligned} \sum_{i=1}^{n} a_i^2 &= \sum_{i=1}^{n} \left[\frac{x_i - \overline{x}}{\sum_{i=1}^{n} (x_i - \overline{x})^2}\right]^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{\left(\sum_{i=1}^{n} (x_i - \overline{x})^2\right)^2} \\ &= \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2 \sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{1}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \\ &= \frac{1}{nS_x^2} \end{aligned}$$

el estimador $\widehat{\beta}_1$ es la suma de una constante (β_1) y una ponderación de las perturbaciones $\left(\sum_{i=1}^{n} a_i u_i\right)$, de la varianza de $\widehat{\beta}_1$ es entonces

$$\begin{aligned} var\left(\widehat{\beta}_1\right) &= var(\beta_1) + var\left(\sum_{i=1}^{n} a_i u_i\right) \\ var\left(\widehat{\beta}_1\right) &= var\left(\sum_{i=1}^{n} a_i u_i\right) = \sum_{i=1}^{n} a_i^2 var(u_i) \\ &= \frac{\sigma_u^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{\sigma_u^2}{nS_x^2} \qquad \text{entonces} \\ var\left(\widehat{\beta}_1\right) &= \frac{\sigma_u^2}{nS_x^2} \end{aligned}$$

Para obtener la varianza de $\widehat{\beta}_0$ se tiene

$$var\left(\widehat{\beta}_0\right) = var\left(\overline{y} - \widehat{\beta}_1 \overline{x}\right)$$

$$= var(\overline{y}) + var\left(\widehat{\beta}_1\overline{x}\right) - 2cov\left(\overline{y}, \widehat{\beta}_1\overline{x}\right) \tag{17}$$

De la expresión (14) del modelo lineal simple se tiene

$$\begin{aligned}
\sum_{i=1}^{n} y_i &= \sum_{i=1}^{n} (\beta_0 + \beta_1 x_i) + \sum_{i=1}^{n} u_i \\
&= \sum_{i=1}^{n} \beta_0 + \sum_{i=1}^{n} \beta_1 x_i + \sum_{i=1}^{n} u_i \\
&= \beta_0 \sum_{i=1}^{n} 1 + \beta_1 \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} u_i \\
&= n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} u_i
\end{aligned}$$

si se multiplican ambos lados por $\frac{1}{n}$

$$\begin{aligned}
\frac{\sum_{i=1}^{n} y_i}{n} &= \frac{n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} u_i}{n} \\
\overline{y} &= \beta_0 + \beta_1\overline{x} + \overline{u}
\end{aligned}$$

entonces

$$\begin{aligned}
var(\overline{y}) &= var(\beta_0) + var(\beta_1\overline{x}) + var(\overline{u}) \\
&= 0 + \beta_1^2 var(\overline{x}) + \frac{\sigma_u^2}{n} = \frac{\sigma_u^2}{n}
\end{aligned}$$

como β_0, β_1 y $\overline{x}$ son constantes, entonces las varianzas son iguales a cero, e indican que no están correlacionadas con $\overline{u}$, donde $\overline{u}$ es el promedio de la muestra de tamaño n de una variable aleatoria con esperanza igual a cero y varianza σ_u^2, entonces

$$var(\overline{y}) = var(\overline{u}) = \frac{\sigma_u^2}{n}$$

sustituyendo los valores obtenidos en (17) se tiene

$$var\left(\widehat{\beta}_0\right) = \frac{\sigma_u^2}{n} + \overline{x}^2 var(\beta_1) - 2cov\left(\overline{y}, \widehat{\beta}_1\overline{x}\right)$$

$$
\begin{aligned}
&= \frac{\sigma_u^2}{n} + \overline{x}^2 \frac{\sigma_u^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2} - 2\overline{x}cov\left(\overline{y}, \widehat{\beta}_1\right) \\
&= \sigma_u^2 \left(\frac{1}{n} + \frac{\overline{x}^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \right) - 2\overline{x}cov\left(\overline{y}, \widehat{\beta}_1\right)
\end{aligned}
$$

donde se ha tomado en cuenta que $\overline{x}$ es determínistica.

De acuerdo con la expresión para la estimación de $\widehat{\beta}_1$ y sustituyendo para obtener la covarianza de $\overline{y}$ y $\widehat{\beta}_1$, se tiene[1]

$$
\begin{aligned}
Cov\left(\overline{y}, \widehat{\beta}_1\right) &= Cov\left(\beta_1, \overline{y}\right) + Cov\left(\overline{y}, \sum_{i=1}^{n} a_i u_i\right) \\
&= 0 + \sum_{i=1}^{n} Cov\left(\overline{y}, a_i u_i\right) \\
&= \sum_{i=1}^{n} a_i Cov\left(\overline{y}, u_i\right) = \sum_{i=1}^{n} a_i \frac{\sigma_u^2}{n} = \frac{\sigma_u^2}{n} \sum_{i=1}^{n} a_i \\
&= \frac{\sigma_u^2}{n}(0) = 0
\end{aligned}
$$

en donde se ha utilizado que

$$
\begin{aligned}
Cov\left(\overline{y}, u_i\right) &= Cov\left(\widehat{\beta}_0 + \widehat{\beta}_1 \overline{x} + \frac{1}{n} \sum_{i=1}^{n} u_j u_i\right) \\
&= Cov\left(\widehat{\beta}_0, u_i\right) + Cov\left(\widehat{\beta}_1 \overline{x}, \overline{u}_i\right) + Cov\left(\frac{1}{n} \sum_{i=1}^{n} u_j u_i\right) \\
&= 0 + 0 + \frac{1}{n} \sum_{i=1}^{n} Cov\left(u_j, u_i\right) \\
&= \frac{1}{n} var\left({}_j u_i\right)
\end{aligned}
$$

[1] Se utilizarán propiedades de la covarianza

(1) $Cov(x + a, y) = Cov(x, y)$

(2) $Cov(x + a, y + a) = Cov(x, y)$

(3) $Cov\ (x - a, y - b) = Cov(x, y)$

(4) $Cov(ax, y) = aCov(x)$

(5) $Cov(ax, by) = abCov(x, y)$

(6) $Cov(ax + b, cy + d) = acCov(x, y)$

$$= \frac{\sigma_u^2}{n}$$

entonces

$$\begin{aligned}
Var\left(\widehat{\beta}_0\right) &= \sigma_u^2 \left(\frac{1}{n} + \frac{\overline{x}^2}{\sum\limits_{i=1}^{n} (x_i - \overline{x})^2} \right) \\
&= \sigma_u^2 \left(\frac{\sum\limits_{i=1}^{n} (x_i - \overline{x})^2 + n\overline{x}^2}{n \sum\limits_{i=1}^{n} (x_i - \overline{x})^2} \right) \\
&= \sigma_u^2 \left(\frac{\sum\limits_{i=1}^{n} (x_i^2 - 2x_i\overline{x} + \overline{x}^2) + n\overline{x}^2}{n \sum\limits_{i=1}^{n} (x_i - \overline{x})^2} \right) \\
&= \sigma_u^2 \left(\frac{\sum\limits_{i=1}^{n} x_i^2 - 2\overline{x} \sum\limits_{i=1}^{n} x_i + \overline{x}^2 \sum\limits_{i=1}^{n} 1 + n\overline{x}^2}{n \sum\limits_{i=1}^{n} (x_i - \overline{x})^2} \right) \\
&= \sigma_u^2 \left(\frac{\sum\limits_{i=1}^{n} x_i^2 - 2\overline{x}\,(n\overline{x}) + 2n\overline{x}^2}{n \sum\limits_{i=1}^{n} (x_i - \overline{x})^2} \right) \\
&= \sigma_u^2 \left(\frac{\sum\limits_{i=1}^{n} x_i^2}{n \sum\limits_{i=1}^{n} (x_i - \overline{x})^2} \right)
\end{aligned}$$

Para obtener la covarianza entre los dos estimadores se tiene

$$\begin{aligned}
Cov\left(\widehat{\beta}_0, \widehat{\beta}_1\right) &= Cov\left(\overline{y} - \widehat{\beta}_1\overline{x}, \widehat{\beta}_1\right) \\
&= Cov\left(\overline{y}, \widehat{\beta}_1\right) - Cov\left(\widehat{\beta}_1\overline{x}, \widehat{\beta}_1\right) \\
&= 0 - \overline{x}Var\left(\widehat{\beta}_1\right)
\end{aligned}$$

$$= -\overline{x}\frac{\sigma_u^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

$$= \frac{-\sigma_u^2\overline{x}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

esta última expresión indica que la covarianza entre $\widehat{\beta}_0$ y $\widehat{\beta}_1$ es contrario al signo de la media muestral de X. Se analizan los casos siguientes

(i) media muestral positiva, error de la estimación β_1 positivo (valor de $\widehat{\beta}_1$ mayor al teórico), entonces el producto por la media de X generaría en promedio una contribución positiva del error de estimación cuando se explica a la variable Y.

$$\overline{y} = \beta_0 + \beta_1\overline{x} = \underset{(1)}{\left[\widehat{\beta}_0 + \widehat{\beta}_1\overline{x}\right]} + \left[\underset{(2)}{\left(\beta_0 - \widehat{\beta}_0\right)} + \underset{(3)}{\left(\beta_1 - \widehat{\beta}_1\right)\overline{x}}\right]$$

en donde (3) tiene una contribución negativa $\left(\widehat{\beta}_1 > \beta_1\right)$. Para compensar, la estimación por MCO de β_0 quedaría por debajo de su valor verdadero;

(ii) media muestral positiva, error de estimación β_1 negativo (valor de $\widehat{\beta}_1$ menor al teórico), entonces de la expresión anterior de $\overline{y}$, se tendría que nuevamente se hace una contribución positiva y que para compensar el valor de β_0 quedaría por arriba del valor verdadero.

(iii) media muestral negativa, el caso sería semejante a (i).

Como se conocen las expresiones de las varianzas y covarianzas de los estimadores, entonces pueden realizarse pruebas estadísticas para contrastar hipótesis acerca de sus valores teóricos ya sea de manera individual o en conjunto. Sin embargo, en las expresiones anteriores se tiene la varianza del término de error σ_u^2 que es desconocido; por tanto debe estimarse tal parámetro y utilizar dicho parámetro en lugar del verdadero valor desconocido. Se utiliza entonces la varianza muestral como un estimador. Como los residuos de mínimos cuadrados tienen media igual a cero entonces su varianza muestral es

$$\widehat{\sigma}_u^2 = \frac{S_u^2}{n-2} = \sum_{i=1}^{n}\frac{\widehat{u}_i}{n-2}$$

se toma $n-2$ en el denominador para que el estimador sea insesgado. Al obtener la estimación de la varianza, entonces pueden ser utilizadas las expresiones anteriores de la varianza de los estimadores de los coeficientes de tal forma que se tienen las "estimaciones de las varianzas de los coeficientes estimados"

2.3.3 Eficiencia del estimador

En el modelo de regresión, el efecto aleatorio proviene del término de error, del cual se ha supuesto que tiene esperanza nula y varianza σ_u^2. Dicho efecto se transmite a la variable y_i, la cual tiene esperanza $E(y_i) = \beta_0 + \beta_1 x_i$ y varianza σ_u^2, al igual que u_i. De la expresión (14) se tiene que el estimador MCO β_1 depende linealmente de las observaciones de la variable aleatoria Y. $\widehat{\beta}_0$ es al igual que $\widehat{\beta}_1$ una combinación lineal de las observaciones de y_i :

$$
\begin{aligned}
\widehat{\beta}_0 &= \overline{y} - \widehat{\beta}_1 \overline{x} = \overline{y} - \left(\beta_1 + \sum_{i=1}^{n} a_i u_i \right) \overline{x} \\
&= \overline{y} - \overline{x}\beta_1 - \overline{x} \sum_{i=1}^{n} a_i u_i \\
&= \overline{y} - \overline{x}\beta_1 - \overline{x} \sum_{i=1}^{n} a_i \left(y_i - \beta_0 - \beta_1 x_i \right) \\
&= \overline{y} - \overline{x}\beta_1 - \overline{x} \sum_{i=1}^{n} a_i y_i + \overline{x}\beta_0 \sum_{i=1}^{n} a_i + \overline{x}\beta_1 \sum_{i=1}^{n} a_i x_i \\
&= \overline{y} - \overline{x}\beta_1 - \overline{x} \sum_{i=1}^{n} a_i y_i + \overline{x}\beta_0 (0) + \overline{x}\beta_1 (1) \\
&= \overline{y} - \overline{x}\beta_1 - \overline{x} \sum_{i=1}^{n} a_i y_i + \overline{x}\beta_1 \\
&= \overline{y} - \overline{x} \sum_{i=1}^{n} a_i y_i \\
&= \frac{\sum_{i=1}^{n} y_i}{n} - \overline{x} \sum_{i=1}^{n} a_i y_i = \sum_{i=1}^{n} \left(\frac{1}{n} - \overline{x} a_i \right) y_i
\end{aligned}
$$

El estimador de MCO es de mínima varianza de entre los estimadores lineales. Se utilizan los dos siguientes Teoremas para justificar esta propiedad.

Teorema de Gauss-Markov. *Bajo los supuestos del Modelo de Regresión Lineal Clásico, el estimador de MCO es el estimador lineal insesgado de mínima varianza de los coeficientes del modelo de regresión.*

Este teorema afirma que la matriz de covarianzas del estimador de MCO es inferior a la de cualquier otro estimador lineal e insesgado. Esto es, la diferencia entre ambas matrices es semidefinida negativa. Una implicación de este resultado es que la varianza del estimador MCO de β_0 es inferior a la de cualquier otro estimador lineal e insesgado de dicho coeficiente, y lo mismo ocurre con la varianza del estimador MCO de β_1.

Cuando el término de error del modelo tiene una distribución Normal se afirma que el estimador MCO es eficiente, esto es, tiene la menor varianza posible (la menor matriz de covarianzas), dentro de la clase de los estimadores insesgados, ya sean lineales o no.

Teorema de Rao. *Si se cumplen los supuestos del Modelo de Regresión Lineal Clásico, y además el término de error del modelo tiene una distribución Normal, entonces el estimador MCO es el estimador insesgado de mínima varianza de los coeficientes del modelo de regresión.*

Demostración.

Considérese el modelo de regresión con término de error Normal

$$y_i = \beta_0 + \beta_1 x_i + u_i \qquad\qquad u_i \sim N\left(0, \sigma_u^2\right)$$

el cual tiene la siguiente función de verosimilitud

$$L\left(\beta_0, \beta_1, \sigma_u^2 \mid y_1, x_1, \ldots, y_n, x_n\right) = \prod_{i=1}^{n} \frac{1}{\sigma_u \sqrt{2\pi}} e^{-u_i^2/2\sigma_u^2}$$

y su logaritmo:

$$\begin{aligned}\ln L\left(\beta_0, \beta_1, \sigma_u^2 \mid y_1, x_1, \ldots, y_n, x_n\right) &= -\frac{n}{2}\ln\sigma_u^2 - \frac{n}{2}\ln 2\pi - \sum_{i=1}^{n}\frac{u_i^2}{2\sigma_u^2} \\ &= -\frac{n}{2}\ln\sigma_u^2 - \frac{n}{2}\ln 2\pi - \sum_{i=1}^{n}\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma_u^2}\end{aligned}$$

al derivar se tiene

$$\begin{aligned}\frac{\partial \ln L}{\partial \beta_0} &= \frac{1}{\sigma_u^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) \\ \frac{\partial \ln L}{\partial \beta_1} &= \frac{1}{\sigma_u^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)\, x_i\end{aligned}$$

$$\frac{\partial \ln L}{\partial \sigma_u^2} = -\frac{n}{2}\frac{1}{\sigma_u^2} - \frac{1}{2\sigma_u^4}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Se puede obtener el estimador de máxima verosimilitud, bajo el supuesto de que los términos de error tienen distribución Normal. Se igualan a cero las dos primeras ecuaciones del sistema anterior; sin embargo estas ecuaciones son precisamente las ecuaciones normales de MCO, de tal modo se tiene que *el estimador de máxima verosimilitud de los parámetros β_0 y β_1 coincide con el estimador de MCO*. Al igualar a cero la tercera ecuación, se obtiene el estimador de la varianza del término de error:

$$\widehat{\sigma}_{MV}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\widehat{u}_i^2$$

que es parecido al estimador de MCO para dicho parámetro. De hecho, el estimador de MCO σ_u^2 es insesgado. Sin embargo, el estimador de máxima verosimilitud es sesgado

$$E\left(\widehat{\sigma}_{MV}^2\right) = E\left(\frac{n-2}{n}\widehat{\sigma}_{MCO}^2\right) = \frac{n-2}{n}E\left(\widehat{\sigma}_{MCO}^2\right) = \frac{n-2}{n}\widehat{\sigma}_u^2$$

pero el sesgo desaparece al aumentar el tamaño muestral ya que el factor $\frac{n-2}{n}$ tiende a uno. El estimador MV de la varianza es, asintóticamente insesgado.

Las segundas derivadas son:

$$\begin{aligned}
\frac{\partial^2 \ln L}{\partial \beta_0^2} &= -\frac{1}{\sigma_u^2}n \\
\frac{\partial^2 \ln L}{\partial \beta_1^2} &= -\frac{1}{\sigma_u^2}\sum_{i=1}^{n}x_i^2 \\
\frac{\partial^2 \ln L}{\partial \left(\sigma_u^2\right)^2} &= \frac{n}{2}\frac{1}{\left(\sigma_u^2\right)^2} - \frac{1}{\left(\sigma_u^2\right)^3}\sum_{i=1}^{n}\widehat{u}_i^2 \\
\frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_1} &= -\frac{1}{\sigma_u^2}\sum_{i=1}^{n}x_i \\
\frac{\partial^2 \ln L}{\partial \beta_0 \partial \sigma_u^2} &= -\frac{1}{\sigma_u^2}\sum_{i=1}^{n}\widehat{u}_i^2 \\
\frac{\partial^2 \ln L}{\partial \beta_1 \partial \sigma_u^2} &= -\frac{1}{\sigma_u^2}\sum_{i=1}^{n}\widehat{u}_i^2 x_i
\end{aligned}$$

donde se ha hecho explícito que por la coincidencia de los estimadores de máxima verosimilitud y de MCO, las diferencias $y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i$ son los residuos de MCO. Se sabe que la suma de los residuos como la de su producto por x_i son cero, al aplicar el operador esperanza en el resto de las segundas derivadas, se tiene la matriz de información:

$$I(\theta) = -E\left(\frac{\partial^2 \ln L(\beta_0, \beta_1, \sigma_u^2 \mid y_1, x_1, \ldots, y_n, x_n)}{\partial\theta\partial\theta'}\right) = \begin{bmatrix} \frac{n}{\sigma_u^2} & \frac{1}{\sigma_u^2}\sum_{i=1}^{n} x_i^2 & 0 \\ \frac{1}{\sigma_u^2}\sum_{i=1}^{n} x_i & \frac{1}{\sigma_u^2}\sum_{i=1}^{n} x_i^2 & 0 \\ 0 & 0 & \frac{n}{2}\frac{1}{\sigma_u^2} \end{bmatrix}$$

donde θ denota el vector $\theta = (\beta_0, \beta_1, \sigma_u^2)$. La inversa de esta matriz es la matriz de covarianzas del estimador de máxima verosimilitud del vector θ. Si se observa la matriz $I(\theta)$ tiene una estructura por bloques, un bloque de 2×2 y otro de 1×1, su inversa conserva esta misma estructura. Se prueba entonces que:

1) los estimadores de máxima verosimilitud de β_0 y β_1 no están correlacionados con el estimador de σ_u^2, debido a la estructura de bloques de la matriz de información. Esto implica que las covarianzas entre los estimadores de máxima verosimilitud de los coeficientes y la varianza del término de error son igual a cero.

2) Ocurre lo mismo con los estimadores de MC. Los estimadores de los coeficientes son los mismos, y el de σ_u^2 se diferencia del de máxima verosimilitud en un factor constante, por lo que se mantiene la no correlación.

3) La inversa del bloque de 2×2, que es la matriz de covariazas del estimador de máxima verosimilitud de β_0 y β_1 genera las varianzas y covarianzas que se obtuvieron para el estimador de mínimos cuadrados.

La tercera condición garantiza que el estimador de MCO de los coeficientes β_0 y β_1 es eficiente, ya que es insesgado y su matriz de covarianzas coincide con el bloque correspondiente de la inversa de la matriz de información. Este resultado es válido siempre y cuando el término de error tenga una distribución Normal.

2.3.4 Medidas de Bondad de ajuste

El procedimiento de MCO arroja una recta de regresión que proporciona la menor Suma de Cuadrados de Residuos. Sin embargo, el ajuste de la recta puede ser medido mediante un indicador del grado de ajuste de la regresión por MCO con respecto a la nube de puntos del cual se ha partido.

Recuérdese que

$$\widehat{u}_i = y_i - \widehat{y}_i = y_i - \left(\widehat{\beta}_0 + \widehat{\beta}_1 x_i\right)$$

Si u_i se distribuye con una función de distribución de probabilidad Normal, entonces $y_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i + u_i$ sigue también una distribución Normal ya que se tiene una constante y una variable que se distribuye Normal; además de que

$$\begin{aligned} E(y_i) &= E\left(\widehat{\beta}_0 + \widehat{\beta}_1 x_i + u_i\right) \\ &= E\left(\widehat{\beta}_0\right) + E\left(\widehat{\beta}_1 x_i\right) + E\left(u_i\right) \\ &= \beta_0 + \beta_1 x_i \end{aligned}$$

$$\begin{aligned} var\left(y_i\right) &= var\left(\widehat{\beta}_0 + \widehat{\beta}_1 x_i + u_i\right) \\ &= var\left(u_i\right) = \sigma_u^2 \end{aligned}$$

así, de acuerdo con el modelo, las observaciones de la variable endógena tienen la misma varianza pero diferente esperanza ya que esta depende del valor de X y éste cambia a lo largo de la serie.

El residuo correspondiente a cada observación es una combinación lineal de todos los términos de error del modelo y en consecuencia, si la perturbación aleatoria del modelo es Normal, el residuo presenta la misma distribución Normal[2].

La esperanza para el residuo es

$$\begin{aligned} E\left(\widehat{u}_i\right) &= E\left(y_i - \widehat{y}_i\right) = E(y_i) - E\left(\widehat{\beta}_0 + \widehat{\beta}_1 x_i\right) \\ &= \left(\beta_0 + \beta_1 x_i\right) - E\left(\widehat{\beta}_0\right) - E\left(\widehat{\beta}_1 x_i\right) \end{aligned}$$

[2]Sean $X_1, X_2, \ldots X_n$ v.a.i. con distribuciones respectivas $N\left(\mu, \sigma_1^2\right)$, $N\left(\mu, \sigma_2^2\right)$, $\cdots, N\left(\mu, \sigma_n^2\right)$ entonces la combinación lineal

$$X = \sum_{i=1}^{n} c_i X_i$$

tiene una distribución Normal $\sum_{i=1}^{n} c_i X_i \sim N\left(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2\right)$

$$= \quad (\beta_0 + \beta_1 x_i) - \beta_0 + \beta_1 x_i = 0$$

de β_0, $\widehat{\beta}_0$, β_1, $\widehat{\beta}_1$, únicamente $\widehat{\beta}_0$ y $\widehat{\beta}_1$ son aleatorios.

La varianza para $\widehat{u}_i$ es

$$\begin{aligned}
var\left(\widehat{u}_i\right) &= var\left[(\beta_0 + \beta_1 x_i) - \left(\widehat{\beta}_0 + \widehat{\beta}_1 x_i\right)\right] \\
&= var\left(\widehat{\beta}_0 + \widehat{\beta}_1 x_i\right) \\
&= var\left(\widehat{\beta}_0\right) + x_i^2 var\left(\widehat{\beta}_1\right) + 2x_i Cov\left(\beta_0, \beta_1\right) \\
&= \sigma_u^2 \frac{\sum_{j=1}^{n} x_j^2}{n\sum_{j=1}^{n}(x_j - \overline{x})^2} + x_i^2 \frac{\sigma_u^2}{\sum_{j=1}^{n}(x_j - \overline{x})^2} - 2\frac{\sigma_u^2 x_i \overline{x}}{\sum_{j=1}^{n}(x_j - \overline{x})^2} \\
&= \frac{\sigma_u^2}{\sum_{j=1}^{n}(x_j - \overline{x})^2}\left[\frac{\sum_{j=1}^{n} x_j^2}{n} + x_i^2 - 2x_i\overline{x}\right] \\
&= \frac{\sigma_u^2}{n\sum_{j=1}^{n}(x_j - \overline{x})^2}\left[\sum_{j=1}^{n} x_j^2 + nx_i^2 - 2x_i\sum_{j=1}^{n} x_j\right] \\
&= \frac{\sigma_u^2}{n\sum_{j=1}^{n}(x_j - \overline{x})^2}\left[\sum_{j=1}^{n}\left(x_j^2 + x_i^2 - 2x_i x_j\right)\right] \\
&= \frac{\sigma_u^2}{n}\left[\frac{\sum_{j=1}^{n}(x_j - x_i)^2}{\sum_{j=1}^{n}(x_j - \overline{x})^2}\right]
\end{aligned}$$

como se tiene esperanza cero, la varianza del residuo es un indicador adecuado del tamaño. La varianza será mayor (lo cual no es deseable), cuanto mayor es σ_u^2, pero será mayor si el tamaño muestral aumenta. También es menor, cuando la varianza muestral de la variable exógena aumenta (lo que es deseable, ya que el grado de fluctuación en X no es negativo sino positivo).

Ahora, si observamos la expresión

$$\frac{\sum_{j=1}^{n}(x_j - x_i)^2}{\sum_{j=1}^{n}(x_j - \overline{x})^2}$$

el x_i correspondiente al $\widehat{u}_i$ está en el numerador, esto es, cuanto más se separe ésta de la media de todas las x_i, la varianza del residuo aumentará.

2.3.5 Error estándar de la regresión

Como ya se vio anteriormente, la esperanza matemática de la distribución de probabilidad de cada uno de los residuos MCO es cero, además de que también se cumple que la media muestral es cero.

Si para toda la muestra, los residuos tienen media cero, entonces la desviación muestral será un indicador del tamaño promedio de ellos. Esto es importante, ya que si la recta estimada tiene un buen ajuste en la nube de puntos, los residuos deberían ser pequeños. Utilizar la desviación muestral de los residuos es un criterio de ajuste; además de que si se utiliza en el denominador $n-2$, el cuadrado es un estimador insesgado de σ_u^2. El insesgamiento puede demostrarse sin obtener previamente los residuos de la regresión.

Al recordar la expresión

$$\begin{aligned}
\widehat{\sigma}_u^2 &= \frac{S_u^2}{n-2} = \sum_{i=1}^{n}\frac{\widehat{u}_i^2}{n-2} = \frac{1}{n-2}\sum_{i=1}^{n}\widehat{u}_i^2 \\
&= \frac{1}{n-2}\sum_{i=1}^{n}\widehat{u}_i\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) \\
&= \frac{1}{n-2}\sum_{i=1}^{n}\widehat{u}_i y_i = \frac{1}{n-2}\sum_{i=1}^{n}\left(y_i - \underset{\to 0}{\widehat{\beta}_0} - \underset{\to 0}{\widehat{\beta}_1 x_i}\right) y_i \\
&= \frac{1}{n-2}\sum_{i=1}^{n}\left(y_i^2 - y_i\widehat{\beta}_0 - y_i\widehat{\beta}_1 x_i\right) \\
&= \frac{1}{n-2}\left[\sum_{i=1}^{n} y_i^2 - \widehat{\beta}_0\sum_{i=1}^{n} y_i - \widehat{\beta}_1\sum_{i=1}^{n} y_i x_i\right]
\end{aligned}$$

y al obtener la raíz cuadrada, se obtiene la desviación estimada que recibe el nombre de error estándar de la regresión (EER).

$$EER = \sqrt{\frac{\sum_{i=1}^{n} u_i^2}{n-2}} = \sqrt{\widehat{\sigma}_u^2} = \widehat{\sigma}_u^2$$

Cuando se minimiza la varianza residual; se minimiza también el EER. Se debe recordar que la desviación se encuentra medida en las mismas unidades que la variable a la que se refiere, en este caso, las mismas unidades que el residuo que a su vez tienen las mismas unidades de la variable endógena y_i. Posteriormente, para conocer que tan bueno es el ajuste, se hace una comparación con la media muestral de la variable endógena

2.3.6 Coeficiente de determinación

Cuando se tienen dos modelos que tienen variable dependiente diferente, el criterio del EER no es adecuado ya que no se pueden hacer afirmaciones sobre que modelo es mejor, es por está razón que se utiliza el criterio del coeficiente de determinación, denotado por R^2; tal coeficiente es un indicador sin unidades, que no se relaciona con ninguna de las variables del modelo.

Se tiene entonces

$$y_i - \overline{y} = (\widehat{y}_i - \overline{y}) + (y_i - \widehat{y}_i) = (\widehat{y}_i - \overline{y}) + \widehat{u}_i$$

que indica la distancia entre una observación y_i y su media $\overline{y}$ puede ser escrita como la distancia entre el valor ajustado $\widehat{y}_i$ y la $\overline{y}$ más el residuo correspondiente. La regresión estimada por MCO proporciona el valor numérico de $\widehat{y}_i - \overline{y}$ que es una aproximación de $y_i - \overline{y}$, el resto es la parte no explicada o residuo. "La desviación total respecto a la media puede escribirse como la suma de la desviación explicada y el residuo"[3].

Trabajando con la expresión anterior se tiene que si elevamos al cuadrado

$$\begin{aligned}(y_i - \overline{y})^2 &= [(\widehat{y}_i - \overline{y}) + \widehat{u}_i]^2 \\ &= (\widehat{y}_i - \overline{y})^2 + 2(\widehat{y}_i - \overline{y})\,\widehat{u}_i + \widehat{u}_i^2\end{aligned}$$

[3]Novales, pág. 501.

si se hace la suma para todas las observaciones

$$\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2+2\sum_{i=1}^{n}\widehat{u}_i(\widehat{y}_i-\overline{y})+\sum_{i=1}^{n}\widehat{u}_i^2$$

de la expresión

$$\begin{aligned}
\sum_{i=1}^{n}\widehat{u}_i(\widehat{y}_i-\overline{y}) &= \sum_{i=1}^{n}\widehat{u}_i\widehat{y}_i-\sum_{i=1}^{n}\widehat{u}_i\overline{y}\\
&= \sum_{i=1}^{n}\widehat{u}_i\widehat{y}_i-\overline{y}\sum_{i=1}^{n}\widehat{u}_i=\sum_{i=1}^{n}\widehat{u}_i\widehat{y}_i\\
&= \sum_{i=1}^{n}\widehat{u}_i\left(\widehat{\beta}_0+\widehat{\beta}_1x_i\right)=\sum_{i=1}^{n}\widehat{u}_i\widehat{\beta}_0+\sum_{i=1}^{n}\widehat{u}_i\widehat{\beta}_1x_i\\
&= \widehat{\beta}_0\sum_{i=1}^{n}\widehat{u}_i+\widehat{\beta}_1\sum_{i=1}^{n}\widehat{u}_ix_i=0 \quad \text{entonces}\\
\sum_{i=1}^{n}(y_i-\overline{y})^2 &= \sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2+\sum_{i=1}^{n}\widehat{u}_i^2 \qquad \text{esto es}\\
nS_y^2 &= nS_{\widehat{y}}^2+nS_{\widehat{u}}^2
\end{aligned}$$

donde $nS_{\widehat{y}}^2$ es la suma explicada por la regresión

$nS_{\widehat{u}}^2$ es la suma no explicada

nS_y^2 es la variación total

ahora si se divide la suma explicada por la variación total, se tiene entonces

$$\begin{aligned}
R^2 &= \frac{nS_{\widehat{y}}^2}{nS_y^2}=\frac{S_{\widehat{y}}^2}{S_y^2}=\frac{S_y^2-S_{\widehat{u}}^2}{S_y^2}=1-\frac{\sum_{i=1}^{n}u_i^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}\\
&= 1-\frac{\sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2} \qquad \text{entonces}\\
R^2 &= 1-\frac{\text{Variación no explicada en } Y}{\text{Variación total}}\\
R^2 &= \frac{\text{Variación explicada en } Y}{\text{Variación total}}
\end{aligned}$$

Teorema *El coeficiente de determinación del modelo de regresión toma valores entre cero y uno.*

Justificación

$$R^2 = \frac{\text{Variación explicada en } Y}{\text{Variación total en } Y}$$

el lado derecho es el cociente de dos números positivos, entonces dicho cociente es positivo.

El coeficiente de determinación, R^2 indica el porcentaje de la variación total en la variable Y que la regresión estimada es capaz de explicar. Si la regresión tiene un ajuste bueno, esto se debe a que la variable X explica buena parte de la variación que Y experimenta a lo largo de la muestra, y los residuos serán generalmente pequeños. La variación explicada en Y será un porcentaje elevado de su variación muestral total, y el coeficiente de determinación será próximo a la unidad. Ocurrirá lo contrario cuando el ajuste de MCO a la nube de puntos no sea el adecuado en cuyo caso el R^2 será cercano a cero.

Un modelo que proporciona un buen ajuste a los datos, puede ser utilizado para efectuar evaluaciones e inferencias acerca de la cuestión conceptual que lo originó. Cuando un R^2 es próximo a cero indica que las estimaciones obtenidas explican pobremente las variaciones que tiene la variable endógena, con lo cual no se pueden realizar evaluaciones o inferencias confiables.

Sin embargo, al evaluar el R^2 se debe tener cuidado ya que cuando se cuentan con pocas observaciones, uno o dos residuos elevados pueden generar un R^2 bajo lo que puede generar una mala evaluación de la regresión. Por el contrario, si ningún residuo es especialmente alto, el R^2 será elevado, en este caso el coeficiente debe tomarse como un indicador de información muestral escasa.

Definición En ocasiones los supuestos que justifican el uso de MCO no se cumplen, entre estos se tienen los que hacen acerca de la propiedad del término de error

a) la varianza de las perturbaciones no es la misma a lo largo de toda la muestra, lo que se denomina heteroscedasticidad[4];

b) cuando las series son temporales, los términos de error correspondientes a dos períodos consecutivos no son independientes, esto es, existe autocorrelación.

El modelo se estima con el supuesto de que cumple con las hipótesis básicas de MCO. Una

[4]Esto es, ausencia de homoscedasticidad.

vez que ha sido estimado se procede a hacer contraste de hipótesis sobre el comportamiento de los residuos que se han obtenido de la estimación para detectar heteroscedasticidad o autocorrelación.

2.4 Heteroscedasticidad

Supóngase el modelo de regresión

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

donde u_i es una variable aleatoria con distribución Normal con esperanza cero y varianza σ_i^2, diferente para las diferentes observaciones.

La heteroscedasticidad se puede presentar en datos de series temporales aunque es más común que se presente en datos de corte transaversal o de sección cruzada.

Cuando hay presencia de heteroscedasticidad los estimadores por MCO siguen siendo insesgados, pero se tienen las consecuencias siguientes:

1) No se cumple el Teorema de Gauss Markov, y las estimaciones de MCO ya no son de mínima varianza; esto es, es posible obtener estimadores lineales de menor varianza que los obtenidos por MCO.

2) Las expresiones que se han utilizado para el cálculo de varianzas y covarianzas de los estimadores de MCO son incorrectas. Las expresiones tienen dificultades cuando hay presencia de heteroscedasticidad, ya que el parámetro σ_u^2 que aparece en dichas expresiones no tiene significado cuando la variamza del término de error es diferente para cada x_i de la muestra.

3) La construcción de intervalos de confianza y la contrastación de hipótesis es inadecuada.

4) El estadístico R^2 carece de interpretación.

2.4.1 Detección

Cuando existe heteroscedasticidad, no es posible proceder sin realizar hipótesis acerca del comportamiento de la varianza de los términos de error. En ausencia de hipótesis, habría que estimar la varianza para cada observación muestral, pero, además se desea estimar los coeficientes. Dada está situación se debe efectuar alguna hipótesis sobre la variación de σ_i^2, lo común es hacer depender a σ_i^2 de los valores que toma alguna variable observable. La detección de la estructura de variación de σ_i^2 puede deducirse como sigue:

1) se hace una estimación inicial del modelo por MCO, ignorando la posible variación de σ_i^2,

2) se hacen dos tipos de análisis

a) examinar la gráfica de los residuos;

b) realizar regresiones auxiliares de los residuos o sus cuadrados.

Cada residuo es una variable aleatoria, su magnitud ya sea en valor absoluto o cuadrado es una aproximación a su desviación. La varianza de los residuos no coincide con la varianza del término de error, pero es una aproximación de la misma. Si se encuentra gráficamente alguna dependencia entre la magnitud del residuo y alguna de las variables exógenas, por ejemplo x_i, puede definirse un modelo de heteroscedasticidad como

$$\sigma_i^2 = var(u_i) = kx_i \qquad (1)$$

ó

$$\sigma_i^2 = var(u_i) = kx_i^2 \qquad (2)$$

(1) Supone una relación directa entre la variable x y un indicador del tamaño de los residuos; ésta expresión sólo tiene sentido cuando X toma valores positivos; pues la estructura de la σ^2 no tiene sentido en el caso de X negativa. En el caso de (2) la varianza del término de error y en consecuencia el valor absoluto de los residuos, aumenta con el cuadrado de la variable exógena; la ventaja es que es compatible con valores positivos y negativos.

La varianza se hace depender del término de error de la variable exógena del modelo, porque si dependiera de otra variable indicaría que la segunda variable tiene capacidad de explicar a y_i, por lo que debería incluirse como variable exógena y en consecuencia el modelo estaría mal especificado.

Un enfoque formal propuesto por White, consiste en estimar una regresión del cuadrado de los residuos sobre las variables exógenas y sus cuadrados, que en el modelo de regresión simple se específica como

$$\widehat{u}_i^2 = \delta_0 + \delta_1 x_i + \delta_2 x_i^2 + \varepsilon_i$$

y se utiliza el estadístico nR^2, donde R^2 denota el coeficiente de determinación de esta regresión auxiliar. Si no hay heteroscedasticidad dependiente de las variables exógenas, el R^2 tenderá a cero; si existe heteroscedasticidad este valor no es cercano a cero y al multiplicarlo por n tenderá a elevar ese valor. Este contraste trata de observar que efecto que efecto es predominante, si es el primero entonces se concluye que no hay heteroscedasticidad. Para determinar si el estadístico es lo suficientemente grande, en cuyo caso se rechaza la hipótesis nula (H_0 : hay homoscedasticidad) se utilizan las tablas χ^2_{n-2}, donde n es el número de variables exógenas de la regresión auxiliar.

2.4.2 Estimación

La estimación por MCO trata de minimizar la Suma de Cuadrados de los residuos pero si la varianza correspondiente a cada observación muestral es diferente el tratamiento por MCO no es adecuado: cuanto mayor sea la varianza, el componente no explicable de la variable endógena será mayor y en consecuencia la observación será menos confiable. Interesa entonces minimizar la suma de cuadrados ponderada, de manera tal que se dé menos importancia a aquellos residuos que corresponden a las observaciones muestrales de mayor varianza σ_i^2.

El proceso de estimación es el siguiente: Supóngase que las varianzas σ_i^2 son conocidas, entonces se divide el modelo por σ_i

$$\frac{y_i}{\sigma_i} = \beta_0 \frac{1}{\sigma_i} + \beta_1 \frac{x_i}{\sigma_i} + \frac{u_i}{\sigma_i}$$

que es un modelo con variables diferentes, pero el término de error satisface

$$var\left(\frac{u_i}{\sigma_i}\right) = \frac{var\left(u_i\right)}{\sigma_i^2} = \frac{\sigma_i^2}{\sigma_i^2} = 1$$

que es un modelo homoscedastico. No es necesario que la varianza sea uno, lo importante es que la varianza sea constante; esto es, que sea igual para todas las observaciones.

El procedimiento anterior no es directamente aplicable ya que la σ_i^2 es desconocida, pero pueden utilizarse hipótesis sobre las varianzas. Se describe a continuación el caso en que la varianza se supone dependiente del cuadrado de la variable exógena.

$$var\left(u_i\right) = \sigma_i^2 = kx_i^2$$

donde k es desconocida.

Cuando se tiene esta situación, se deben transformar las variables de tal manera que el modelo que se obtenga satisfaga las hipótesis de MCO. Supóngase entonces, en este caso, que se divide cada variable del modelo por x_i

$$y_i^* = \frac{y_i}{x_i}, \qquad x_i^* = \frac{x_i}{x_i} = 1, \qquad u_i^* = \frac{u_i}{x_i}$$

el modelo tiene la forma

$$\begin{aligned} \frac{y_i}{x_i} &= \beta_0 \frac{1}{x_i} + \beta_1 \frac{x_i}{x_i} + \frac{u_i}{x_i} \\ y_i^* &= \beta_0 \frac{1}{x_i} + \beta_1 + u_i^* \end{aligned}$$

reordenando

$$y_i^* = \beta_1 + \beta_0 \frac{1}{x_i} + u_i^*$$

donde β_1 es el término independiente, y la variable exógena es la inversa de X, $\dfrac{1}{x_i}$. El término de error u_i^* es el cociente entre el u_i original y una variable no aleatoria x_i, por lo cual u_i^* es estocástico. En particular u_i tiene una distribución Normal con esperanza matemática igual a cero, y varianza igual a

$$\begin{aligned} \sigma_i^{*2} &= var\left(u_i^*\right) = var\left(\frac{u_i}{x_i}\right) = \frac{1}{x_i^2} var\left(u_i\right) \\ &= \frac{1}{x_i^2}\left(kx_i^2\right) = k \end{aligned}$$

esto es, es constante para todas las observaciones. Por tanto, el método de MCO está justificado para la estimación del modelo transformado, y se obtiene

$$
\begin{aligned}
\widehat{\beta}_0 &= \frac{\sum_{i=1}^{n}\left(y_i^{\star}-\overline{y}^{\star}\right)\left(x_i^{\star}-\overline{x}^{\star}\right)}{\sum_{i=1}^{n}\left(x_i^{\star}-\overline{x}^{\star}\right)} \\
&= \frac{\sum_{i=1}^{n}\left(\frac{y_i}{x_i}-\frac{1}{n}\sum_{j=1}^{n}\frac{y_j}{xj}\right)\left(\frac{1}{x_i}-\frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j}\right)}{\sum_{i=1}^{n}\left(\frac{1}{x_i}-\frac{1}{n}\sum_{j=1}^{n}\frac{1}{xj}\right)^2} \\
\widehat{\beta}_1 &= \overline{y}^{\star}-\widehat{\beta}_0\overline{x}^{\star}=\frac{1}{n}\left(\sum_{i=1}^{n}\frac{y_i}{x_i}\right)-\widehat{\beta}_0\left(\frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j}\right)
\end{aligned}
$$

La matriz de varianzas-covarianzas del modelo transformado esta determinado por

$$
\begin{aligned}
var\left(\widehat{\beta}_0\right) &= \frac{\sigma^{\star 2}}{\sum_{i=1}^{n}\left(x_i^{\star}-\overline{x}^{\star}\right)^2}=\frac{k}{\sum_{i=1}^{n}\left(\frac{1}{x_i}-\frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j}\right)} \\
var\left(\widehat{\beta}_1\right) &= \sigma^{\star 2}\frac{\sum_{i=1}^{n}x_i^{\star 2}}{n\sum_{i=1}^{n}\left(x_i^{\star}-\overline{x}^{\star}\right)^2}=\frac{k\sum_{i=1}^{n}\frac{1}{x_i}}{n\sum_{i=1}^{n}\left(\frac{1}{x_i}-\frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j}\right)} \\
cov\left(\widehat{\beta}_0,\widehat{\beta}_1\right) &= -\frac{\overline{x}^{\star}\sigma^{\star 2}}{\sum_{i=1}^{n}\left(x_i^{\star}-\overline{x}^{\star}\right)^2}=\frac{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}\right)k}{\sum_{i=1}^{n}\left(\frac{1}{x_i}-\frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j}\right)}
\end{aligned}
$$

donde k es la varianza de los residuos del modelo transformado.

La transformación en el modelo se justifica por el hecho de cumplir con las hipótesis del modelo de MCO, sin embargo la interpretación de los coeficientes no debe realizarse en el contexto del modelo transformado sino en el modelo original; esto es, mediante la transformación se han obtenido buenos estimadores del modelo original cuando se calculen los residuos del modelo, el R^2, σ_u^2, ó ERR. Ningún estadístico de los anteriormente mencionados se calcula sobre el modelo transformado.

El procedimiento anterior se denomina Mínimos Cuadrados Ponderados, y esto se debe a que cada observación se pondera de manera diferente, en este caso en particular, las ponderaciones han sido las inversas de x_i, aunque cada problema donde se presenta heteroscedasticidad tiene una ponderación diferente. El método de Mínimos Generalizados Ponderados es un caso particular del método de Mínimos Cuadrados Generalizados en presencia de heteroscedasticidad.

Por tanto, los residuos del modelo son

$$\widehat{u}_i = y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i$$

donde aparecen las variables del modelo original.

La varianza residual es kx_i^2 y el R^2 puede calcularse como $R^2 = 1 - \dfrac{SCR}{SCT}$

2.5 Autocorrelación

La autocorrelación se produce cuando los términos de error correspondientes a distintas observaciones muestrales están correlacionadas. Este hecho contradice una de las hipótesis del modelo de regresión. Al igual que la heteroscedasticidad, la autocorrelación puede deberse a la omisión de alguna variable exógena relevante y que tiene una estructura correlacionada en el tiempo.

Pueden existir situaciones de autocorrelación complejas, sin embargo, sólo se tratará la autocorrelación de primer orden

$$u_t = \rho u_{t-1} + \varepsilon_t \qquad |\rho| < 1 \tag{1}$$

Debe aclararse que (1) no implica que cada término de error dependa únicamente del anterior. Esta estructura indica que cada término de error depende en mayor medida del anterior y de forma decreciente de los anteriores. La autocorrelación también se denomina positiva o negativa, dependiendo del signo del parámetro ρ.

Las consecuencias de la autocorrelación son

1) El estimador de MCO es lineal e insesgado, pero ya no es el de mínima varianza, y por

tanto existe un estimador más eficiente.

2) Las expresiones utilizadas para calcular varianzas y covarianzas de los estimadores de MCO son sesgados.

3) Los intervalos de confianza y los estadísticos para el contraste de hipótesis no son adecuados.

4) El cálculo de la varianza residual σ_u^2 es un estimador sesgado.

5) El R^2 es sesgado.

El análisis de los residuos mediante gráficas es un buen procedimiento para detectar autocorrelación. Los residuos de MCO siguen manteniendo una media muestral igual a cero. Si los residuos no están autocorrelacionados se presentaran oscilaciones alrededor del cero de forma aleatoria; por el contrario, si dichos residuos están correlacionados se presentarán valores consecutivos con igual signo que, oscilarán alrededor de cero, pero no tendrán cambios bruscos de signo. Así, es conveniente examinar el gráfico de los residuos en función del tiempo para observar si existen los patrones antes descritos. Otra forma de observar el comportamiento de los residuos es observar a $\widehat{u}_t$. como función del residuo del periodo anterior $\widehat{u}_{t-1}$. Si la gráfica presenta una nube de puntos con un mayor número de estos en el primer y tercer cuadrantes, existe entonces autocorrelación positiva, si por el contrario, los puntos están en el segundo y cuarto cuadrantes, existe entonces autocorrelación negativa.

Otra forma de detectar autocorrelación de primer orden es mediante el estadístico Durbin-Watson (DW)

$$DW = \frac{\sum_{t=2}^{n} (\widehat{u}_t - \widehat{u}_{t-1})^2}{\sum_{t=2}^{n} \widehat{u}_t^2}$$

$$DW = \frac{\sum_{t=2}^{n} (\widehat{u}_t - \widehat{u}_{t-1})^2}{\sum_{t=2}^{n} \widehat{u}_t^2} = \frac{\sum_{t=2}^{n} (\widehat{u}_t^2 - 2\widehat{u}_t\widehat{u}_{t-1} + \widehat{u}_{t-1})}{\sum_{t=2}^{n} \widehat{u}_t^2} = \frac{\sum_{t=2}^{n} \widehat{u}_t^2 - 2\sum_{t=2}^{n} \widehat{u}_t\widehat{u}_{t-1} + \sum_{t=2}^{n} \widehat{u}_{t-1}}{\sum_{t=2}^{n} \widehat{u}_t^2}$$

$$
\begin{aligned}
&\simeq 2\frac{\sum_{t=2}^{n}\widehat{u}_t^2-\sum_{t=2}^{n}\widehat{u}_t\widehat{u}_{t-1}}{\sum_{t=2}^{n}\widehat{u}_t^2}=2\left[\frac{\sum_{t=2}^{n}\widehat{u}_t^2}{\sum_{t=2}^{n}\widehat{u}_t^2}-\frac{\sum_{t=2}^{n}\widehat{u}_t\widehat{u}_{t-1}}{\sum_{t=2}^{n}\widehat{u}_t^2}\right]=2\left[1-\frac{\sum_{t=2}^{n}\widehat{u}_{t-1}}{\sum_{t=2}^{n}\widehat{u}_t}\right]\\
&= 2\left[1-\widehat{\rho}\right]
\end{aligned}
$$

donde $\widehat{\rho}$ es el coeficiente de correlación entre $\widehat{u}_t$ y $\widehat{u}_{t-1}$. Obsérvese que $\widehat{\rho}$ es el estimador de MCO del parámetro ρ de (1). Así para efectuar el constraste se estima una regresión el modelo por MCO original y posteriormente se estima una regresión de los residuos sobre un periodo anterior. La estimación MCO de la pendiente de ésta última regresión permite el cálculo del estadístico DW por la relación DW y $\widehat{\rho}$.

Bajo la hipótesis nula H_0 : no existe autocorrelación por lo que $\rho = 0$, y el estadístico DW es igual a 2. Entonces, los valores del estadístico alrededor de dos indica que no existe autocorrelación. En otros casos se tiene que si $\rho \to +1$ entonces el estadístico DW se acerca a cero por tanto hay autocorrelación positiva. Si $\rho \to -1$ el estadístico DW$\simeq 4$, por tanto existe autocorrelación negativa. Se puede concluir entonces que existe autocorrelación si los valores se alejan de dos.

La tabla del estadístico (incluida al final de este trabajo) proporciona dos valores críticos para cada nivel de significancia, ambos valores se encuentran entre cero y dos; y que están denotados por d_i y d_s, donde d_i es el límite inferior y d_s el límite superior. Las hipótesis que se manejan son $H_0 : \rho = 0$ vs $H_1 : \rho > 0$, si se tienen valores del estadístico DW entre $[0, d_i]$ se tiene evidencia de que $\rho > 0$, esto es se debe rechazar H_0. Si por el contrario, los valores del estadístico están entre $[d_s, 2]$ y entonces no se debe rechazar H_0. Valores de DW entre d_i y d_s no proporcionan evidencia a favor de ninguna de las dos hipótesis.

No debe olvidarse que el estadístico DW sólo sirve para detectar autocorrelación de primer orden, esto es, la que corresponde al modelo

$$\widehat{u}_t = \rho\widehat{u}_{t-1}$$

Sin embargo, el problema de autocorrelación puede tener una estructura aún más compleja. En general debe examinarse la función de autocorrelación que queda definida como la sucesión de estadísticos r_k :

$$r_k = \frac{\sum_{t=k+1}^{T} \widehat{u}_t \widehat{u}_{t-k}}{\sum_{t=k+1}^{T} \widehat{u}_t^2} \qquad k = 1, 2, 3, \ldots$$

donde r_0 es

$$r_0 = \frac{\sum_{t=1}^{T} \widehat{u}_t \widehat{u}_t}{\sum_{t=1}^{T} \widehat{u}_t^2} = 1$$

y r_1 es

$$r_1 = \frac{\sum_{t=2}^{T} \widehat{u}_t \widehat{u}_{t-1}}{\sum_{t=2}^{T} \widehat{u}_t^2} = \widehat{\rho}$$

2.5.1 ¿Qué hacer en presencia de autocorrelación?

Supóngase que el término de error del modelo de regresión presenta autocorrelación, esto es,

$$u_t = \rho u_{t-1} + \varepsilon_t \qquad |\rho| < 1 \tag{1}$$

$$y_{t-1} = \beta_0 + \beta_1 x_{t-1} + u_{t-1}$$

multiplicando por ρ

$$\rho y_{t-1} = \rho \beta_0 + \rho \beta_1 x_{t-1} + \rho u_{t-1} \tag{2}$$

recuérdese que el modelo en el tiempo t es

$$y_t = \beta_0 + \beta_1 x_t + u_t \tag{3}$$

restando (2) de (3) se tiene

$$y_t - y_{t-1} = \beta_0 (1 - \rho) + \beta_1 (x_t - \rho x_{t-1}) + (u_t \rho - \rho u_{t-1})$$

$$= \delta_0 + \delta_1 (x_t - \rho x_{t-1}) + \varepsilon_t$$

el término de error mantiene la estructura de (1) por lo que puede escribirse como

$$y_t^* = \beta_0^* + \beta_1 x_t^* + \varepsilon_t \tag{4}$$

donde las variables de la ecuación (4) son transformaciones de las variables originales

$$\begin{aligned} y_t^* &= y_t - \rho y_{t-1} \\ x_t^* &= x_t - \rho x_{t-1} \end{aligned}$$

el término de error del modelo (4) ya no presenta autocorrelación, por lo que es válido el uso de MCO. No debe olvidarse que la transformación de las variables se hace para obtener estimadores con buenas propiedades.

Al recuperar las estimaciones de los parámetros del modelo original a partir de los obtenidas en el modelo transformado se debe tener en cuenta

$$\widehat{\sigma}_u = \frac{\widehat{\sigma}_\varepsilon}{1-\rho^2}; \qquad \widehat{\beta}_0 (1-\widehat{\rho}) = \widehat{\delta}_0; \qquad \widehat{\beta}_0 = \widehat{\delta}_1.$$

Por tanto, si el modelo presenta autocorrelación se tiene entonces

1) se estima el modelo de MCO; no se toma en cuenta si existe autocorrelación;

2) se hace uso de las estimaciones para generar los residuos MCO: $\widehat{u}_t = y_t - \widehat{\beta}_0 - \widehat{\beta}_1 x_1$;

3) con los residuos, se estima el modelo de autocorrelación: $\widehat{u}_t = \rho \widehat{u}_{t-1} + \varepsilon_t$;

4) con el residuo $\widehat{\rho}$ estimado, se transforman las variables.

5) se estima $y_t^* = \widehat{\beta}_0^* + \widehat{\beta}_1 x_1^* + \varepsilon_t$;

6) los coeficientes y las desviaciones típicas estimadas del modelo transformado son los que se asocian al modelo original. Los residuos se obtienen con relación a las variables del modelo original, aunque los coeficientes se estiman con el modelo transformado.

El procedimiento donde se utiliza MCO en dos ocasiones se conoce como Estimación por Mínimos Cuadrados Generalizados en presencia de autocorrelación.

Capítulo 3

El modelo de regresión múltiple

De forma general una variable dependiente puede tener más de un sola variable explicativa, por lo que el modelo de regresión simple puede ser demasiado débil cuando se quiere explicar algún evento de interés.

Cuando se tiene más de una variable exógena determinando el comportamiento de una variable endógena, deben considerarse de manera conjunta y por tanto hay que utilizar los procedimientos de regresión y correlación múltiple que se encuentran dentro del análisis de regresión múltiple. Consi-
dérese el siguiente modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

en el que se tienen dos variables exógenas, X_1 y X_2 y una variable endógena Y. En este modelo se observa que se tienen dos componentes: uno determínista $\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}$ y otro aleatorio u_i.

El coeficiente β_0 sigue llamandóse término independiente de la regresión, mientras que β_1 y β_2 son las pendientes del modelo, también denominadas coeficientes de regresión parcial ya que miden el efecto que tendría sobre la variable endógena una variación determinada de una de las variables exógenas si la otra permaneciera constante. Los coeficientes estimados miden la variación esperada en Y por unidad de variación en X_1 y X_2, respectivamente. Al aplicar la estimación de β_1 al recorrido muestral de X_1 (valor mínimo-valor máximo), se tendrá el rango de efectos de esta variable sobre Y. Se tiene además que la variación media que una variable ha experimentado a lo largo de la muestra puede ser aproximada por su desviación típica muestral, si se multiplica ésta por la estimación de su coeficiente, se puede obtener una medida de la variación media ocasionada en Y por cada variable exógena.

Debe mencionarse que si existe alguna variable relevante para la determinación de la variable endógena Y que quede omitida en el modelo de regresión, entonces las estimaciones de mínimos cuadrados serán sesgadas, por lo que es recomendable utilizar el modelo de regresión múltiple en lugar de un modelo de regresión simple.

El análisis que se hace de un modelo de regresión múltiple mantiene las mismas hipótesis que el modelo de regresión simple, considerando las variables exógenas X_1 y X_2 son deterministas, además de suponer homoscedasticidad y ausencia de autocorrelación y agregando la hipótesis de **ausencia de colinealidad** entre las variables exógenas X_1 y X_2, lo que significa que no guardan una proporción exacta; esto es, el valor absoluto de su coeficiente de correlación es menor a 1 en valor absoluto.

Este capítulo se estructura de la manera siguiente: se habla de los fundamentos del modelo de regresión múltiple, el contraste de hipótesis en estos modelos, y finalmente se describen dos supuestos adicionales, la prueba de cambio estructural y el coeficiente de correlación ajustado.

3.1 Estimación por MCO del Modelo de Regresión Múltiple

Se trata de minimizar la Suma de los Cuadrados de los Residuos, se tiene entonces que

$$\underset{\beta_0,\beta_1,\beta_2}{Min} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i})^2$$

la solución al sistema de ecuaciones es el estimador MCO. Una vez que se ha resuelto, la diferencia: $\widehat{u}_i = y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_{1i} - \widehat{\beta}_2 x_{2i}$ será el residuo que corresponde a la observación $i-ésima$. Antes de resolver el sistema se puede escribir de forma reducida como:

$$\sum_{i=1}^{n} \widehat{u}_i = 0; \qquad \sum_{i=1}^{n} \widehat{u}_i x_{1i} = 0; \qquad \sum_{i=1}^{n} \widehat{u}_i x_{2i} = 0;$$

que tiene propiedades similares a las que se obtuvieron en el modelo de regresión lineal simple:

a) la suma de los residuos de mínimos cuadrados es igual a cero

b) los residuos de mínimos cuadrados no están relacionados con las variables exógenas del modelo.

Los incisos (a) y (b) no se cumplen con otro proceso de estimación.

Para resolver el sistema y encontrar el estimador de MCO, primero se despeja β_0 de la primera ecuación:

$$\widehat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} y_i - \widehat{\beta}_1\frac{1}{n}\sum_{i=1}^{n} x_{1i} - \widehat{\beta}_2\frac{1}{n}\sum_{i=1}^{n} x_{2i}$$

y sustituyendo en las otras dos se tiene

$$\sum_{i=1}^{n} y_i x_{1i} = \left(\frac{1}{n}\sum_{i=1}^{n} y_i - \widehat{\beta}_1\frac{1}{n}\sum_{i=1}^{n} x_{1i} - \widehat{\beta}_2\frac{1}{n}\sum_{i=1}^{n} x_{2i}\right)\sum_{i=1}^{n} x_{1i} + \widehat{\beta}_1\sum_{i=1}^{n} x_{1i}^2 + \widehat{\beta}_2\sum_{i=1}^{n} x_{1i}x_{2i}$$

$$\sum_{i=1}^{n} y_i x_{2i} = \left(\frac{1}{n}\sum_{i=1}^{n} y_i - \widehat{\beta}_1\frac{1}{n}\sum_{i=1}^{n} x_{1i} - \widehat{\beta}_2\frac{1}{n}\sum_{i=1}^{n} x_{2i}\right)\sum_{i=1}^{n} x_{2i} + \widehat{\beta}_1\sum_{i=1}^{n} x_{1i}x_{2i} + \widehat{\beta}_2\sum_{i=1}^{n} x_{2i}^2$$

que pueden ser escritas como

$$\begin{aligned} S_{x_1y} &= \widehat{\beta}_1 S_{x_1}^2 + \widehat{\beta}_2 S_{x_1x_2} \\ S_{x_2y} &= \widehat{\beta}_1 S_{x_1x_2} + \widehat{\beta}_2 S_{x_2}^2 \end{aligned} \tag{1}$$

donde

$$S_{x_1y} = \frac{1}{n}\sum_{i=1}^{n}(x_{1i} - \overline{x}_1)(y_i - \overline{y})$$

$$S_{x_2y} = \frac{1}{n}\sum_{i=1}^{n}(x_{2i} - \overline{x}_2)(y_i - \overline{y})$$

$$S_{x_1}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{1i} - \overline{x}_1)^2$$

$$S_{x_2}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{2i} - \overline{x}_2)^2$$

$$S_{x_1x_2} = \frac{1}{n}\sum_{i=1}^{n}(x_{1i} - \overline{x}_1)(x_{2i} - \overline{x}_2)$$

el sistema (1) es lineal con dos ecuaciones y dos incógnitas que tiene por solución

$$\widehat{\beta}_1 = \frac{S_{x_1y}S_{x_2}^2 - S_{x_2y}S_{x_1x_2}}{S_{x_1}^2 S_{x_2}^2 - (S_{x_1x_2})^2}$$

$$\widehat{\beta}_2 = \frac{S_{x_2y}S_{x_1}^2 - S_{x_1y}S_{x_1x_2}}{S_{x_1}^2 S_{x_2}^2 - (S_{x_1x_2})^2}$$

a partir de estas dos expresiones, se puede demostrar el siguiente teorema:

Si las dos variables exógenas, X_1 y X_2 no están correlacionadas, entonces el estimador MCO del modelo que incluye a ambas simúltaneamente:

$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$

coincide con los estimadores MCO que se obtendrían a partir de modelos que las incluyesen por separado como exógenas:

$y_i = \gamma_0 + \gamma_1 x_{1i} + \varepsilon_{1i}$ *e* $y_i = \delta_0 + \delta_1 x_{2i} + \varepsilon_{2i}$, *esto es* $\widehat{\beta}_1 = \widehat{\gamma}_1$; $\widehat{\beta}_2 = \widehat{\delta}_1$.

Demostración

Para realizar los contrastes de hipótesis acerca de los valores numéricos de los coeficientes del modelo, es necesario disponer de las varianzas de los estimadores, de la misma manera en que se hacia en el modelo de regresión simple. La deducción del cálculo de las varianzas no se realiza y sólo se presenta la fórmula para calcular éstas.

$$Var(\widehat{\beta}_0) = \sigma_u^2 \frac{1}{n} + \left(\frac{\overline{x}_1^2 S_{x_2}^2 + \overline{x}_2^2 S_{x_1}^2 - 2\overline{x}_1\overline{x}_2 S_{x_1x_2}}{S_{x_1}^2 S_{x_2}^2 - (S_{x_1x_2})^2}\right)$$

$$Var(\widehat{\beta}_1) = \sigma_u^2 \frac{S_{x_2}^2}{S_{x_1}^2 S_{x_2}^2 - (S_{x_1x_2})^2}$$

$$Var(\widehat{\beta}_2) = \sigma_u^2 \frac{S_{x_1}^2}{S_{x_1}^2 S_{x_2}^2 - (S_{x_1x_2})^2}$$

$$Cov\left(\widehat{\beta}_1,\widehat{\beta}_2\right)=-\frac{\sigma_u^2 S_{x_1x_2}}{S_{x_1}^2 S_{x_2}^2-\left(S_{x_1x_2}\right)^2}$$

Si se quiere contrastar una hipótesis nula que hace referencia a que se cumplan simúltaneamente los diferentes valores numéricos de los coeficientes cuando estos han sido estimados, será preciso entonces utilizar en la construcción del estadístico la covarianza de los estimadores.

El estimador de MCO del modelo de regresión múltiple es insesgado, además el estimador lineal es de mínima varianza, ya que el Teorema de Gauss-Markov sigue siendo válido para los modelos multivariados. De la misma manera el Teorema de Rao también se cumple, ya que, el estimador de mínimos cuadrados es eficiente cuando el término de error sigue una distribución Normal, esto es, tiene la mínima varianza posible de entre todos los estimadores insesgados, ya sean lineales o no.

3.1.1 Grado de ajuste del Modelo de Regresión Múltiple

Una vez que se han obtenido los valores estimados de los coeficientes del modelo por MCO, los valores explicados de la variable endógena $\widehat{y}_i$, así como los residuos $\widehat{u}_i$ se tiene entonces

$$\widehat{y}_i=\widehat{\beta}_0+\widehat{\beta}_1 x_{1i}+\widehat{\beta}_2 x_{2i} \tag{1}$$

$$\widehat{u}_i=y_i-\widehat{y}_i=y_i-\widehat{\beta}_0-\widehat{\beta}_1 x_{1i}-\widehat{\beta}_2 x_{2i}$$

y se estima la varianza del término de error mediante la *varianza residual*[1]:

$$\widehat{\sigma}_u=\frac{1}{n-3}\sum_{i=1}^{n}\widehat{u}_i^2=\frac{\widehat{S}_u^2}{n-3} \tag{2}$$

La notación para $\widehat{S}_u^2$ indica que se trata de la suma residual que se obtiene para la variable endógena a partir de las variables exógenas. El Error Estándar de la Regresión (EER) se expresa como:

$$EER=\sqrt[2]{\widehat{\sigma}_u^2}$$

[1] *El denominador muestra el número de observaciones menos el número de coeficientes a estimar, y que se denomina grados de libertad.*

e indica el grado de ajuste del modelo. El EER es la desviación típica de los residuos y constituye una medida del tamaño medio del error cometido en cada observación muestral.

Una expresión equivalente para el cálculo del ERR es:

$$\widehat{S}_u^2 = \sum_{i=1}^{n} y_i^2 - \widehat{\beta}_0 \sum_{i=1}^{n} y_i - \widehat{\beta}_1 \sum_{i=1}^{n} y_i x_{1i} - \widehat{\beta}_2 \sum_{i=1}^{n} y_i x_{2i}. \tag{3}$$

sin embargo, se debe mencionar que a pesar de ser una medida de ajuste del modelo, ésta es sólo una medida relativa, esto es, es útil para comparar modelos alternativos que tratan de explicar la misma variable endógena, pero no se puede utilizar para comparar el grado de ajuste de modelos diseñados para explicar variables endógenas diferentes.

Una medida absoluta de ajuste del modelo, es el **coeficiente de determinación múltiple**, definido como

$$R^2 = 1 - \frac{S_u^2}{S_y^2} \tag{4}$$

donde $S_y^2 = \sum_{i=1}^{n} (y_i - y)^2$, R^2 también puede calcularse sin necesidad de obtener previamente los residuos, sino utilizando los momentos muestrales

$$R^2 = \frac{\widehat{\beta}_0 \sum_{i=1}^{n} y_i + \widehat{\beta}_1 \sum_{i=1}^{n} y_i x_{1i} + \widehat{\beta}_2 \sum_{i=1}^{n} y_i x_{2i} - n\overline{y}^2}{\sum_{i=1}^{n} y_i^2 - n\overline{y}^2} = \frac{\widehat{\beta}_1 S_{x_1 y} + \widehat{\beta}_2 S_{x_2 y}}{S_y^2} \tag{5}$$

a la que se llegó utilizando la expresión (3) y la definición de suma residual.

Si se utilizan los mismos argumentos que cuando se tenia el caso simple, la Suma de Cuadrados de los residuos puede descomponerse como:

$$\begin{aligned} \sum_{i=1}^{n}(y_i - \overline{y})^2 &= \sum_{i=1}^{n}(y_i - \widehat{y}_i + \widehat{y}_i - \overline{y})^2 \qquad (6) \\ &= \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 + \sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2 + 2\sum_{i=1}^{n}(y_i - \widehat{y}_i)(\widehat{y}_i - \overline{y}) \end{aligned}$$

pero el último sumando es igual a cero. Como $y_i - \widehat{y}_i = \widehat{u}_i$, se tiene entonces

$$\begin{aligned}
\sum_{i=1}^{n}(y_i-\widehat{y}_i)(\widehat{y}_i-\overline{y}) &= \sum_{i=1}^{n}\widehat{u}_i\left(\widehat{\beta}_0+\widehat{\beta}_1x_{1i}+\widehat{\beta}_2x_{2i}-\overline{y}\right)\\
&= \widehat{\beta}_0\sum_{i=1}^{n}\widehat{u}_i+\widehat{\beta}_1\sum_{i=1}^{n}\widehat{u}_ix_{1i}+\widehat{\beta}_2\sum_{i=1}^{n}\widehat{u}_ix_{2i}-\overline{y}\sum_{i=1}^{n}\widehat{u}_i
\end{aligned}$$

donde cada uno de los sumandos es igual a cero. Entonces utilizando (6)

$$\begin{aligned}
S_y^2 &= \sum_{i=1}^{n}\widehat{u}_i^2+\sum_{i=1}^{n}(\widehat{y}_i-y_i)^2 = S_u^2+\sum_{i=1}^{n}(\widehat{y}_i-y_i)^2\\
\textit{Suma Total} &= \textit{Suma de Cuadrados de Residuos} + \textit{Suma Explicada}
\end{aligned}$$

lo que indica que la Suma de Cuadrados de Residuos o Suma Residual, S_u^2 es el porcentaje de la variación total en la variable endógena, S_y^2, que no es explicada por las variables exógenas. De esta manera, el coeficiente de determinación está en el intervalo $[0,1]$ y mientras más se acerque a 1 las variables exógenas explicarán mejor los cambios de la variable endógena Y.

3.1.2 Coeficiente de correlación parcial

El modelo de regresión múltiple permite considerar la capacidad de explicación de cada una de las variables exógenas del modelo que de manera independiente se tiene sobre la variable endógena Y. Para realizar este análisis se utiliza el coeficiente de correlación parcial que se define como:

Definición: El coeficiente de correlación parcial entre Y y X_i que se denotará ρ_{x_iy}, en el universo de variables Y, X_i es el coeficiente de correlación simple entre la variables Y y X_i, una vez que se ha eliminado la influencia de las otras variables.

Los coeficientes de correlación parcial suelen también definirse como el grado de relación que existe entre X_i e Y cuando las otras X_j variables se mantienen fijas.

Si se denota por ρ_{x_iy} o $\rho_{x_ix_j}$ a los coeficientes de correlación simple habituales entre Y y X_i ó X_i y X_j respectivamente, entonces los coeficientes de correlación parcial pueden expresarse como

$$\rho_{x_i y} = \frac{\rho_{x_i y} - \rho_{x_j y}\rho_{x_i x_j}}{\sqrt[2]{\left(1-\rho_{x_j y}^2\right)\left(1-\rho_{x_i x_j}^2\right)}}$$

3.2 Contraste de hipótesis en el Modelo de Regresión Múltiple

La contrastación de hipótesis es un ejercicio de inferencia estadística en el cual es necesario conocer la distribución de probabilidad del estadístico en el cual se basará el contraste.

3.2.1 Contraste acerca de coeficientes individuales

En el caso del modelo de regresión múltiple (sucede lo mismo en el caso simple), la distribución de los estimadores MCO proviene de la distribución de los términos de error. A los supuestos que ya se habían majeado anteriormente: esperanza igual a cero, homoscedasticidad y ausencia de autocorrelación se debe sumar el supuesto sobre la distribución de probabilidad de cada u_i es Normal, esto es $u_i \sim N(0, \sigma^2)$.

Los estimadores de mínimos cuadrados son combinaciones lineales de las u_i, por lo que tomando en cuenta que el estimador es insesgado y conociendo sus varianzas se puede concluir que

$$\frac{\widehat{\beta}_0 - \beta_0}{\sqrt{Var\left(\widehat{\beta}_0\right)}} \sim N(0,1); \qquad \frac{\widehat{\beta}_1 - \beta_1}{\sqrt{Var\left(\widehat{\beta}_1\right)}} \sim N(0,1); \qquad \frac{\widehat{\beta}_2 - \beta_2}{\sqrt{Var\left(\widehat{\beta}_2\right)}} \sim N(0,1)$$

lo que sirve para construir intervalos de confianza acerca del verdadero valor del parámetro, y de esta manera hacer contrastación de hipótesis. Sin embargo, las expresiones que se han dado anteriormente utilizan la varianza del término de error que es desconocida, pero que será sustituida ahora por su estimación $\widehat{\sigma}_u$; las distribuciones de probabilidad de los estadísticos que se obtienen una vez que se han sustituido las varianzas teóricas por las varianzas estimadas, pasan a ser t de *Student*

$$\frac{\widehat{\beta}_0 - \beta_0}{\sqrt{\widehat{Var}\left(\widehat{\beta}_0\right)}} \sim t_{n-k}; \qquad \frac{\widehat{\beta}_1 - \beta_1}{\sqrt{\widehat{Var}\left(\widehat{\beta}_1\right)}} \sim t_{n-k}; \qquad \frac{\widehat{\beta}_2 - \beta_2}{\sqrt{\widehat{Var}\left(\widehat{\beta}_2\right)}} \sim t_{n-k}$$

donde el circunflejo sobre las varianzas indica que son varianzas estimadas, esto es, se ha sustituido el verdadero valor de σ_u^2 por su estimación $\widehat{\sigma}_u^2$.

El intervalo de confianza del coeficiente β_i se forma utilizando la distribución t de *Student* del estimador MCO

$$\begin{aligned} 1-\alpha &= P\left[-t_{n-k} \leq \frac{\widehat{\beta}_i - \beta_i}{\sqrt{\widehat{Var}\left(\widehat{\beta}_i\right)}} \leq t_{n-k}\right] \\ &= P\left[\widehat{\beta}_i - t_{n-k,\frac{\alpha}{2}}\widehat{Var}\left(\widehat{\beta}_i\right) \leq \beta_i \leq \widehat{\beta}_i + t_{n-k,\frac{\alpha}{2}}\widehat{Var}\left(\widehat{\beta}_i\right)\right] \end{aligned}$$

de este modo, la probabilidad de que el intervalo fomado por la estimación de β_i y de su varianza, contenga al verdadero valor del parámetro es $1-\alpha$.

3.2.2 Contraste de significancia global

No sólo es importante conocer si los parámetros del modelo son significativos de manera individual, es importante conocer si globalmente el modelo tiene capacidad para explicar el comportamiento de la variable endógena. Es importante considerar que los coeficientes son significativamente diferentes de cero ya que esto indica que el modelo tendrá alguna capacidad explicativa, sin embargo puede presentarse el caso de que aunque por separado las variables no sean significativas, al estar conjuntamente en un modelo estos tengan alguna capacidad de explicación.

Si se recueda que

$$Suma\ Total = Suma\ de\ Cuadrados\ de\ Residuos + Suma\ Explicada$$

entonces la Suma Explicada es

$$Suma\ Explicada = Suma\ Total - Suma\ de\ Cuadrados\ de\ Residuos$$

bajo Normalidad, la Suma Total es n veces la varianza muestral de Y, que sigue una distribución χ^2_{n-1}. La Suma Residual sigue una distribución χ^2_{n-k}. La Suma Residual y la Suma Explicada son independientes. La diferencia de dos variables chi-cuadrada independientes sigue una distribución χ^2 con grados de libertad igual a la diferencia de grados de libertad entre las dos variables. En consecuencia, la Suma Explicada que es la diferencia entre dos variables independientes con distribuciones χ^2_{n-1} y χ^2_{n-k} sigue una distribución χ^2_{n-k-1} y la Suma Residual es χ^2_{n-1}, independiente de la Suma Explicada. Su cociente para k variables exógenas es

$$\frac{\frac{SE}{k-1}}{\frac{SR}{n-k}} \sim F_{k-1,n-k}$$

El contraste se basa en poder concluir si el modelo es bueno o no, la Suma Explicada será pequeña y la Suma Explicada será alta, por lo cual el estadístico tomará un valor muy bajo con relación al valor crítico de la distribución $F_{k-1,n-k}$. El estadístico también puede escribirse como función del coeficiente de determinación

$$\frac{\frac{SE}{k-1}}{\frac{SR}{n-k}} = \frac{\frac{SE}{k-1}}{\frac{ST-SE}{n-k}} = \frac{SE}{ST-SE}\frac{n-k}{k-1} = \frac{\frac{SE}{ST}}{1-\frac{SE}{ST}}\frac{n-k}{k-1} = \frac{R^2}{1-R^2}\frac{n-k}{k-1} \sim F_{k-1,n-k}$$

Está función es monótona creciente en R^2, siendo igual a cero cuando R^2 lo es, y tiene a infinito cuando R^2 se aproxima a 1.

3.3 Contraste de cambio estructural

La prueba de Chow

La prueba de Chow es una herramienta que se usa para comprobar la correcta especificación del modelo. De forma concreta se utiliza cuando se tiene sospecha de que ha ocurrido un cambio estructural en la muestra, lo suficientemente importante como para suponer que el modelo responde a dos muestras diferentes.

Supóngase que se parte de una muestra de tamaño n, entonces se divide en dos submuestras de tamaño n_1 y n_2, de tal forma que $n_1 + n_2 = n$, las muestras son independientes una de la otra, y en cada una de ellas el término de error de cada observación sigue una distribución $N(0, \sigma_u^2)$. Supónga que cada submuestra obedece al modelo

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i & i &= 1, 2, \ldots, n_1 \\ y_j &= \beta_0' + \beta_1' x_{1j} + \beta_2' x_{2j} + u_j & j &= n_1 + 1, n_1 + 2, \ldots, n \end{aligned}$$

ls hipótesis nula es $H_0 : \beta_0 = \beta'_0; \beta_1 = \beta'_1; \beta_2 = \beta'_2$ contra la hipótesis alternativa de que al menos uno de ellos es diferente. Se tomará el caso general de k variables explicativas. El contraste se efectúa primero estimando un modelo para las n observaciones; a este se le denominará modelo restringido, ya que supone que los coeficientes son los mismos para ambas partes de la muestra. Denótese a la suma de cuadrados de los residuos que se ha obtenido como SRR con $n - k$ grados de libertad. Estímese ahora cada uno de los modelos con su submuestra correspondiente, entonces denote a cada una de las sumas de residuos por SR_1 y SR_2 y a la suma de ambos por SRS. Este será el modelo sin restringir, ya que se permite a cada submuestra tener su propio valor de parámetros. Los grados de libertad para cada una de las submuestras son respectivamente $n_1 - k$ y $n_2 - k$, por lo que el número de grados de libertad para el modelo sin restringir es la suma $n_1 - k + n_2 - k = n - 2k$. Posteriormente se comparan ambas sumas residuales, formando el estadístico

$$\frac{\frac{SRR-(SR_1+SR_2)}{k}}{\frac{SR_1+SR_2}{n-2k}} \sim F_{k,n-2k}$$

que relaciona la diferencia de sumas residuales, con la del modelo restringido. El número de restricciones es k, debido a que se impone la igualdad de todos los coeficientes del modelo, por lo que k aparece en el numerador. Por último, el número de grados de libertad del modelo sin restringir aparece en el denominador. Si el estadístico de Chow toma un valor lo *suficientemente grande* , esto es, superior al de las tablas $F_{k,n-2k}$ se tendrá que admitir que las restricciones no son satisfechas, rechazando la hipótesis nula de ausencia de cambio estructural.

Debe hacerse notar que para realizar el contraste, el número de observaciones en cada submuestra debe ser "grande" y mayor que el número de variables exógenas.

3.4 Coeficiente de determinación ajustado

La Suma Explicada no puede disminuir si se agrega una variable exógena al modelo de regresión, esta suma será ahora mayor y, como la Suma Total de la variable endógena continúa siendo la misma, el coeficiente de determinación será mayor. La Suma Explicada sigue una distribución χ^2_{k-1}. Por tanto, que la suma explicada sea considerada grande o pequeña, esto es, exceda al valor crítico de las tablas de la chi-cuadrada, depende del número de variables exógenas del modelo, ya que se debe comparar este valor con el de las tablas de la distribución con los grados de libertad que dependen de k. Este efecto no se ha considerado en el coeficiente de determinación, definido como el cociente entre la Suma Explicada y la Suma Total, esta última no se ve afectada por esta cuestión, ya que el número de grados de libertad es igual a $n-1$, que es independiente del número de variables exógenas; por tanto, variaciones en k afectan el resultado del contraste, ya que no hay corrección por cambios en el número de grados de libertad.

Es por lo antes expuestos que se utiliza el *coeficiente de determinación ajustado* $\overline{R^2}$ que se define como

$$\overline{R^2} = 1 - \frac{\frac{Suma\ Residual}{n-k}}{\frac{Suma\ Explicada}{n-1}} = 1 - \frac{n-1}{n-k}\frac{Suma\ Residual}{Suma\ Explicada} = 1-(1-R^2)\frac{n-1}{n-k}$$

que es una transformación monótona del coeficiente R^2. Se tiene además que siempre se cumple que $\overline{R^2} \leq R^2$; al aumentar el número de variables exógenas k, el factor $\frac{n-1}{n-k}$ que aparece con signo contrario, hace que el coeficiente de detrminación ajustado tienda a disminuir. Sin embargo el R^2 habrá aumentado, de tal modo que el efecto final sobre el coeficiente de determinación ajustado es indeterminado, este incluso puede ser negativo. El coeficiente de determinación ajustado aumentará su valor numérico al incorporar una variable exógena si y sólo si el estadístico t de *Student* de esta nueva variable es superior a 1 en valor absoluto.

Capítulo 4

Modelo para el análisis

Visto las características del Modelo de Regresión así como sus supuestos se procedió a realizar las estimaciones previstas para el caso de los accidentes de trabajo. El capítulo presente muestra el modelo de regresión tabajado y cada una de las pruebas estadísticas aplicadas a los supuestos.

Con base en lo expuesto en los capítulos anteriores, el modelo que determina la influencia de la inversión de los trabajadores y de los empleadores así como de la población sobre el número de accidentes de trabajo de forma empírica es[1]

$$\begin{aligned} ACCIDENTES &= C(1) + C(2)POBLACIÓN \\ &\quad + C(3)EMPLEADORES + C(4)TRABAJADORES \end{aligned}$$

en donde se esperarían signos positivos para la variable población y trabajadores y signo negativo para la variable empleadores. Los signos se explican de la manera siguiente: en el caso de la población el signo será positivo ya que a mayor número de empleados el número de accidentes tiende a incrementarse, para el caso de la inversión de los empleados se espera un signo positivo ya que éstos tenderán a un menor cuidado ya que cubren sus cuotas para la seguridad social. Finalmente en el caso de los empleadores se espera un signo negativo ya que los empresarios

[1] Como ya se mencionó anteriormente los modelos pueden basarse en explicaciones teóricas o en aplicaciones pragmáticas. Este modelo tiene esta última característica, se ha supuesto este tipo de relación entre las variables estudiadas.

tratarán de mejorar los sistemas de prevención de accidentes que existen en sus instalaciones a fin de reducir el índice de siniestralidad.

Al realizar la estimación descrita anteriormente, se presentan problemas de carácter estadístico. esto es, no cumple con los supuestos del método de mínimos cuadrados. entre los problemas se tiene heteroscedasticidad y autocorrelación de primer orden.

Al buscar modelos alternativos. los cuales debían cumplir con los signos buscados y que sea estadísticamente válido se llega al modelo siguiente

$$\begin{aligned} LACCIDENTES &= C(1) + C(2)POBLACIÓN \\ &\quad +C(3)LEMPLEADORES(-1) \\ &\quad +C(4)LTRABAJADORES(-1) \end{aligned}$$

en donde se han calculado los logaritmos para las variables ACCIDENTES, EMPLEADORES y TRABAJADORES, y las últimas dos variables han sido rezagadas un periodo. Este modelo presenta los signos esperados y es válido estadísticamente en cada uno de los supuestos que pide el método de estimación: Mínimos Cuadrados Ordinarios.

Al sustituir los coeficientes obtenidos se tiene

$$\begin{aligned} LACCIDENTES &= 4.9889 + 0.0003POBLACION \\ &\quad -1.1801LEMPLEADORES(-1) \\ &\quad +1.0540LTRABAJADORES(-1) \end{aligned}$$

Donde se observa que

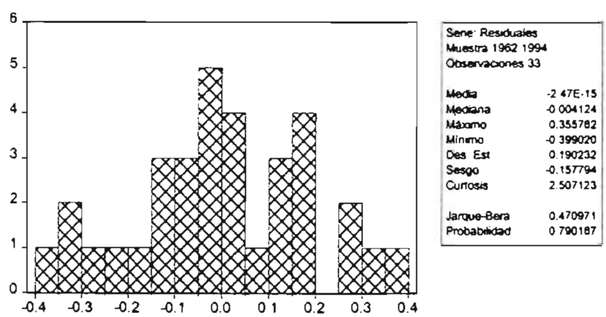
Cuando aumenta la población en una unidad los accidentes se incrementa en .03%, cuando la inversión en el rubro de accidentes de trabajo de los empleadores aumentan en 1% los accidentes disminuyen en 1.18% aunque estas cuotas influyen con un periodo de retraso y cuando las cuotas de los trabajadores se incrementan en 1% los accidentes aumentan en 1.05%. también con un periodo de retraso.

Las pruebas que se tienen que realizar son: prueba de normalidad para los errores y se observará la gráfica de los residuos para determinar si existe algún periodo que presente problemas, autocorrelación serial de segundo orden, multicolinealidad, heteroscedasticidad.

4.1 Prueba de normalidad

Esta prueba se realiza para determinar si la serie de los errores esta normalmente distribuida. El estadístico mide la diferencia del sesgo y la curtosis de la serie con respecto a la distribución normal. Bajo la hipótesis nula de una distribución normal, el estadístico Jarque-Bera se distribuye como una $\chi^2_{(2)}$. El valor probabilistico reportado es la probabilidad de que la Jarque-Bera exceda (en valor absoluto) al valor observado bajo la nula y pequeña probabilidad de rechazar la hipótesis nula de una distribución normal

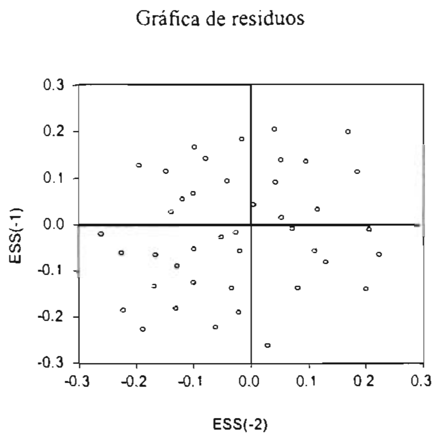
Jarque-Bera



La probabilidad de la Jarque-Bera es en este caso de $0.79 > 0.05$, por lo tanto no se rechaza la hipótesis nula, de esta manera la serie de los residuos se distribuye de manera normal.

4.2 Autocorrelación serial

Existen varias maneras para probar si existe autocorrelación serial, en este caso se observará primeramente el método gráfico que puede ser de gran ayuda para observar si existe correlación serial o no



aquí puede observarse que los errores no presentan un patrón específico, por tanto puede concluirse de la observación que no hay autocorrelación serial de orden (2), se mide autocorrelación de orden 2 por tener variables rezagadas un periodo.

Otra de las pruebas es la de Correlación serial LM de Breusch-Godfrey, la cual da como resultado

Estadístico F	0.851878	Probabilidad	0.442275
R^2	2.057304	Probabilidad	0.357489

Al observarse los datos obtenidos para esta prueba vemos que las dos probabilidades asignadas tanto al estadístico F como a la R^2 son ambos mayores a 0.05, por tanto se puede asegurar que no existe autocorrelación serial de orden 2.

4.3 Multicolinealidad

La detección de multicolinealidad se hace en este trabajo mediante la observación de la R^2 y la probabilidad del estadístico t para cada una de las variables exógenas, se tienen entonces los siguientes resultados

	Probabilidad
R^2	0.968300
$\overline{R}^2$	0.964904
t_C	0.0004
$t_{POBLACIÓN}$	0.0000
$t_{LEMPLEADORES(-1)}$	0.0020
$t_{LTRABAJADORES(-1)}$	0.0000

puede observarse que la R^2 es alta y que las variables exógenas son significativas, puede suponerse entonces que no existe multicolinealidad. Otra manera de probar si existe multicolinealidad es observar la matriz de correlación de las variables

	$LACCID$	$POBLA$	$LEMPLEA.(-1)$	$LTRAB.(-1)$
$LACCID$	1	0.916756	0.851555	0.804434
$POBLA$	0.916756	1	0.897395	0.826130
$LEMPLEA.(-1)$	0.851555	0.897395	1	0.906542
$LTRAB.(-1)$	0.804434	0.826130	0.906542	1

Se utilizarán las abreviaciones siguientes en esta matriz

$LACCID$ para $LACCIDENTES$

$POBLA$ para $POBLACIÓN$

$LEMPLEA(-1)$ para $LEMPLEADORES(-1)$

$LTRAB(-1)$ para $LTRABAJADORES(-1)$

se calcula el determinante de la matriz de correlaciones si este valor es cercano es cero entonces se asume que existe multicolinealidad, al calcular el determinante se obtiene

$$\det = 0.005262474$$

con lo que puede concluirse que no existe multicolinealidad

4.4 Heteroscedasticidad

El hecho de que exista heteroscedasticidad en el modelo tiene como consecuencia que el R^2 carezca de interpretación. La heteroscedasticidad indica que la varianza de los errores no es la misma para todas las observaciones muestrales y por tanto no puede hacerse inferencia sobre los parámetros.

Las hipótesis son H_0 : existe heteroscedasticidad vs. H_1 : No existe heteroscedasticidad

Las pruebas para detectar este problema son la ARCH y las pruebas WHITE y White con términos cruzados. La primera de estas pruebas da como resultado

Prueba ARCH

Estadístico F	1.474043	Probabilidad	0.246847
R^2	2.953196	Probabilidad	0.228413

si se observan los resultados de la prueba tanto la probabilidad del estadístico F como de la R^2 son mayores a 0.05, por lo tanto se rechaza la hipótesis nula.

Para el caso de la prueba WHITE sin términos cruzados se tiene el siguiente resultado

Prueba WHITE (sin términos cruzados)

Estadístico F	1.912505	Probabilidad	0.118198
R^2	10.06719	Probabilidad	0.121851

donde las probabilidades son mayores a 0.05, por tanto, también se rechaza la hipótesis nula.

Por último se tiene la prueba White con términos cruzados, en la que se tiene

Prueba WHITE (con términos cruzados)

Estadístico F	1.248354	Probabilidad	0.317662
R^2	10.81763	Probabilidad	0.288417

que al igual que en el caso anterior, las probabilidades son mayores a 0.05, por tanto se rechaza la hipótesis nula. Puede asegurarse entonces, con base en estas tres pruebas, que no hay presencia de heteroscedasticidad.

4.5 Cambio estructural

Se realizó la prueba Chow Breakpoint para diferentes años, principalmente el año en que entra en vigor la Reforma a la Ley del Seguro Social (1997). Esta prueba indica si el modelo es adecuado para todo el periodo de estimación o si existieron cambios fundamentales que hacen que el modelo no sea válido para todo el período y que haya entonces necesidad de estimar para diferentes muestras modelos alternativos.

Las hipótesis son H_0 : Hay cambio estructural vs. H_1 : No hay cambio estructural.

Los resultados obtenidos para algunos de los años que fueron probados

Prueba Chow (1970)

Estadístico F	2.416174	Probabilidad	0.076615

Prueba Chow (1980)

Estadístico F	1.312752	Probabilidad	0.305014

Prueba Chow (1990)

Estadístico F	1.397874	Probabilidad	0.264694

Prueba Chow (1997)

Estadístico F	1.134750	Probabilidad	0.363870

En cada una de los diferentes años, las probabilidades son mayores a 0.05, por tanto se rechaza la hipótesis nula. No existe evidencia de cambio estructural.

Coeficiente de determinación

El modelo cumple con los supuestos de Modelo de Mínimos Cuadrados. Observando los coeficientes de determinación, ajustado como no ajustado se tiene

$R^2 = 0.968300$, esto es, el modelo explica el comportamiento de los accidentes en un 96%, pero debe observarse el coeficiente de determinación ajustado, ya que muchas veces al incluir variables se hace que el valor del R^2 crezca tan sólo por adicionar variables. El coeficiente de determinación ajustado da el siguiente resultado

$\overline{R}^2 = 0.964904$; el cual indica que las variables exógenas explican el comportamiento de los accidentes en un 96%. De esta manera puede entonces asegurarse que el modelo explica el fenómeno en estudio en un 96%.

Conclusiones

Las consecuencias humanas, familiares y socio-económicas que originan los riesgos de trabajo. son lo suficientemente importantes para provocar un especial intéres en todos los sectores de la vida económica del país. El IMSS en particular está obligado a promover el grado de responsabilidad entre todos lo actores del problema, esta tarea se lleva a cabo a través no sólo del IMSS sino que están involucradas la Secretaría del Trabajo y Previsión Social, los organismos patronales y obreros para promover la prevención de los riegos de trabajo. Los patrones tienen la obligación de avisar al IMSS de todo accidente de trabajo para así tratar de identificar las causas y circunstancias bajos las cuales ocurrió el siniestro.

Un accidente se produce siempre por dos razones: una circunstancia peligrosa, una convicción peligrosa en el centro de trabajo, o la realización y el acto inseguro efectuado por un trabajador. Estas causas inmediatas de los riesgos de trabajo coinciden en un alto porcentaje. Se debe agregar que si los patrones se preocuparan porque en sus centros de trabajo no existieran condiciones peligrosas y los trabajadores tuvieran una conciencia de seguridad, entre ambos se lograria abatir los altos índices de accidentes de trabajo que se registran. Los patrones o empleadores creen que con cumplir con las cuotas están libres de invertir en Seguridad Industrial dentro de la empresa, lo mismo que los trabajadores, al cumplir con una cuota dejan de lado muchas de las condiciones que evitarian el accidentarse aunque dentro de la cuota que pagan los trabajadores no se incluye la de accidentes de trabajo, ya que esta es cubierta por el empleador). Sin embargo, los datos pueden verse afectados ya que en México el tiempo de traslado de la casa del trabajador a la empresa y de la empresa a la casa esta considerado como riesgo de trabajo "Los accidentes de tránsito como riesgos de trabajo han sido la causa del desfinanciamiento del sistema. Este tipo de accidentes es cada vez más grave en las regiones industrializadas y urbanizadas del país. como resultado de la forma en que se transportan los trabajadores y por la violencia en la vía pública"[1].

En 1994, las empresas afiliadas al IMSS eran aproximadamente 675,000 en su mayoría microempresas tenían pocos incentivos para prevenir los riegos de trabajo, ya que se gravaba de manera excesiva a la planta productiva, ya que la mayoría se encontraba clasificada en las clases más altas. Para fijar las primas de este seguro, las empresas son clasificadas y agrupadas según la actividad a la que se dediquen. La cuotas del seguro de riegos de trabajo se cubren integramente por los empleadores y se determinan con relación a la cuantía del salario base de cotización y con los riesgos inherentes a la actividad de que se trate. A partir de 1997, se establecieron cinco clases de riesgo en las que se agrupan los diversos tipos de actividades y ramas industriales, de mayor a menor peligrosidad a la que se exponen los trabajadores. Cada clase se divide en grados y cada grado presenta una siniestralidad distinta. Las empresas se ordenan conforme al Catálogo de Actividades; en esta nueva clasificación se toma en cuenta la cantidad invertida de los empleadores en Seguridad Industrial y no puede bajar su prima más allá del límite superior de la clase siguiente. El trabajo presente no puede hacer inferencias sobre este nuevo esquema ya que sólo se incluyen tres años de la nueva Ley. Sería importante realizar un nuevo estudio tomando en cuenta las nuevas características, las empresas que invierten en seguridad son reconocidas y los trabajadores se encuentran en un ambiente más seguro.

Con base en las estimaciones ya presentadas, la cantidad que invierte el empleador en cuotas, así como el número de trabajadores y las cuotas que estos pagan son un determinante en el número de accidentes que se reportan año con año al IMSS, es factible pensar que estas variables actúen con un año de retraso, ya que los datos que se presentan en los estudios actuariales sobre los accidentes, así como la determinación de cuotas se realiza cada año. Es necesario hacer patente la necesidad de una cultura de responsabilidad y seguridad industrial a todos lo niveles y en todos los ámbitos.

[1] IMSS, "Aportaciones para el Debate", pág.22

Bibliografía

Constitución Política de los Estados Unidos Mexicanos, Berbera Editores S.A. de C.V.; México. 2000.

.Leguina, Joaquín, Fundamentos de demografía, edit. Siglo XXI, España. 1973.

Ley del Seguro Social.

Halli, Shiva S. y Vaninadha Rao, K. Advanced Techniques of Population Analysis, edit. Plenum Press, New York y Londres; 1992.

IMSS. Aportaciones para el debate. La seguridad Social ante el futuro. 1995

IMSS. Lecturas en Matería de Seguridad Social. Riesgos de Trabajo. México 1979.

INEGI; Indicadores sociodemográficos de México (1930-2000). Año de publicación 2001.

Mélendez B., J. "Reflexiones sobre tendencias de los seguros de riesgo de trabajo", en *Revista de Seguridad Social*, Número 240, 2003. pp.37-47

Memoria Estadística del IMSS, varios años.

Anexo

Datos Modelo de Regresión

Año	población asegurada IMSS/1	No. Pensiones en curso de pago	Ingresos del IMSS			Egresos por riesgos de trabajo
		riesgos de trabajo (incapacidad permanente)	ingresos totales	de empleados y empleadores	por cuotas de accidentes de trabajo	
1960	1,200,708	3,246	1,800,000	1,700,000	-	-
1961	1,419,030	4,265	2,200,000	2,100,000	-	-
1962	1,594,315	5,390	2,600,000	2,500,000	200,000	200,000
1963	1,703,402	6,371	3,100,000	3,000,000	300,000	300,000
1964	2,069,480	9,025	3,600,000	3,500,000	300,000	400,000
1965	2,209,915	11,004	4,400,000	4,100,000	400,000	500,000
1966	2,315,103	13,111	5,100,000	4,900,000	600,000	500,000
1967	2,447,398	15,980	5,600,000	5,400,000	600,000	600,000
1968	2,633,054	18,327	6,400,000	6,200,000	700,000	600,000
1969	2,901,907	21,289	6,900,000	6,600,000	800,000	700,000
1970	3,120,763	25,274	8,300,000	7,000,000	1,000,000	900,000
1971	3,232,658	28,621	10,800,000	9,000,000	1,200,000	1,000,000
1972	3,581,084	31,926	12,300,000	10,200,000	1,400,000	1,200,000
1973	3,900,811	24,559	14,300,000	11,800,000	1,600,000	1,500,000
1974	4,019,884	25,922	20,200,000	17,000,000	2,200,000	2,000,000
1975	4,306,000	20,589	25,700,000	21,800,000	2,800,000	2,700,000
1976	4,338,000	24,185	33,900,000	28,800,000	3,500,000	3,600,000
1977	4,554,000	24,762	44,400,000	37,700,000	4,500,000	4,800,000
1978	5,157,000	27,931	54,400,000	46,200,000	5,700,000	6,000,000
1979	5,500,000	32,277	70,400,000	59,500,000	7,600,000	7,800,000
1980	6,369,000	41,376	99,400,000	83,500,000	11,300,000	10,900,000
1981	7,113,000	49,048	143,600,000	120,200,000	17,800,000	15,600,000
1982	7,037,000	57,986	235,800,000	192,800,000	34,100,000	27,100,000
1983	7,059,000	66,431	347,400,000	283,600,000	43,800,000	48,300,000
1984	7,615,000	73,912	538,700,000	454,600,000	61,600,000	74,100,000
1985	8,132,000	83,019	930,400,000	783,900,000	112,100,000	129,500,000
1986	7,986,000	91,811	1,592,900,000	1,377,000,000	184,900,000	219,500,000
1987	8,757,000	99,169	3,650,700,000	3,234,600,000	458,800,000	512,900,000
1988	8,917,000	106,398	7,457,800,000	6,751,600,000	800,200,000	945,700,000
1989	9,926,000	115,322	11,699,200,000	10,662,900,000	1,110,600,000	1,291,500,000
1990	10,764,000	125,020	15,781,500,000	14,377,700,000	1,517,100,000	1,424,700,000
1991	11,333,000	135,798	22,645,900,000	20,453,700,000	2,101,100,000	1,954,500,000
1992	11,369,000	150,041	28,801,800,000	26,397,600,000	2,611,300,000	2,441,200,000
1993	11,317,000	162,861	34,377,800,000	31,180,000,000	3,125,300,000	3,038,200,000
1994	11,561,000	173,657	40,880,000,000	37,180,000,000	3,663,000,000	3,558,000,000
1995	10,932,000	175,284	43,250,000,000	39,320,000,000	3,891,000,000	3,689,000,000
1996	11,895,000	183,250	45,689,700,000	41,256,000,000	4,200,000,000	3,895,600,000
1997	12,714,000	189,982	48,623,500,000	45,623,500,000	4,725,630,000	4,132,500,000
1998	13,611,000	192,922	53,563,000,000	47,856,400,000	4,632,600,000	4,156,000,000
1999	14,560,000	197,113	55,678,900,000	49,562,300,000	4,925,700,000	4,538,200,000
2000	15,433,000	200,659	62,358,000,000	56,251,000,000	5,238,600,000	4,856,700,000

Estadístico d de Durbin-Watson. Valores críticos de d_i y d_s al nivel de significación del 5 por 100 *

n	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$		$k' = 5$	
	d_i	d_s	d_i	d_s	d_i	d_s	d_i	d_s	d_i	d_s
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	1,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,66	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,62	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,63	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,63	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

n = número de observaciones

k' = número de variables explicativas, excluyendo el término constante.

* Esta tabla está tomada de *Biometrika*, vol. 41, pág. 173, 1951.

Estadístico *d* de Durbin-Watson. Valores críticos de d_i y d_s al nivel de significación del 1 por 100

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_i	d_s	d_i	d_s	d_i	d_s	d_i	d_s	d_i	d_s
15	0,81	1,07	0,70	1,25	0,59	1,46	0,49	1,70	0,39	1,96
16	0,84	1,09	0,74	1,25	0,63	1,44	0,53	1,66	0,44	1,90
17	0,87	1,10	0,77	1,25	0,67	1,43	0,57	1,63	0,48	1,85
18	0,90	1,12	0,80	1,26	0,71	1,42	0,61	1,60	1,52	1,80
19	0,93	1,13	0,83	1,26	0,74	1,41	0,65	1,58	0,56	1,77
20	0,95	1,15	0,86	1,27	0,77	1,41	0,68	1,57	0,60	1,74
21	0,97	1,16	0,89	1,27	0,80	1,41	0,72	1,55	0,63	1,71
22	1,00	1,17	0,91	1,28	0,83	1,40	0,75	1,54	0,66	1,69
23	1,02	1,19	0,94	1,29	0,86	1,40	0,77	1,53	0,70	1,67
24	1,04	1,20	0,96	1,30	0,88	1,41	0,80	1,53	0,72	1,66
25	1,05	1,21	0,98	1,30	0,90	1,41	0,83	1,52	0,75	1,65
26	1,07	1,22	1,00	1,31	9,93	1,41	0,85	1,52	0,78	1,64
27	1,09	1,23	1,02	1,32	0,95	1,41	0,88	1,51	0,81	1,63
28	1,10	1,24	1,04	1,32	0,97	1,41	0,90	1,51	0,83	1,62
29	1,12	1,25	1,05	1,33	0,99	1,42	0,92	1,51	0,85	1,61
30	1,13	1,26	1,07	1,34	1,01	1,42	0,94	1,51	0,88	1,61
31	1,15	1,27	1,08	1,34	1,02	1,42	0,96	1,51	0,90	1,60
32	1,16	1,28	1,10	1,35	1,04	1,43	0,98	1,51	0,92	1,60
33	1,17	1,29	1,11	1,36	1,05	1,43	1,00	1,51	0,94	1,59
34	1,18	1,30	1,13	1,36	1,07	1,43	1,01	1,51	0,95	1,59
35	1,19	1,31	1,14	1,37	1,08	1,44	1,03	1,51	0,97	1,59
36	1,21	1,32	1,15	1,38	1,10	1,44	1,04	1,51	0,99	1,59
37	1,22	1,32	1,16	1,38	1,11	1,45	1,06	1,51	1,00	1,59
38	1,23	1,33	1,18	1,39	1,12	1,45	1,07	1,52	1,02	1,58
39	1,24	1,34	1,19	1,39	1,14	1,45	1,09	1,52	1,03	1,58
40	1,25	1,34	1,20	1,40	1,15	1,46	1,10	1,52	1,05	1,58
45	1,29	1,38	1,24	1,42	1,20	1,48	1,16	1,53	1,11	1,58
50	1,32	1,40	1,28	1,45	1,24	1,49	1,20	1,54	1,16	1,59
55	1,36	1,43	1,32	1,47	1,28	1,51	1,25	1,55	1,21	1,59
60	1,38	1,45	1,35	1,48	1,32	1,52	1,28	1,56	1,25	1,60
65	1,41	1,47	1,38	1,50	1,35	1,53	1,31	1,57	1,28	1,61
70	1,43	1,49	1,40	1,52	1,37	1,55	1,34	1,58	1,31	1,61
75	1,45	1,50	1,42	1,53	1,39	1,56	1,37	1,59	1,34	1,62
80	1,47	1,52	1,44	1,54	1,42	1,57	1,39	1,60	1,36	1,62
85	1,48	1,53	1,46	1,55	1,43	1,58	1,41	1,60	1,39	1,63
90	1,50	1,54	1,47	1,56	1,45	1,59	1,43	1,61	1,41	1,64
95	1,51	1;55	1,49	1,57	1,47	1,60	1,45	1,62	1,42	1,64
100	1,52	1,56	1,50	1,58	1,48	1,60	1,46	1,63	1,44	1,65

n = número de observaciones

k' = número de variables explicativas, excluyendo el término constante.

* Esta tabla está tomada de *Biometrika*, vol. 41, pág. 173, 1951.