

COLECCIÓN

CIENCIAS ECONÓMICAS

2

UMET
UNIVERSIDAD
METROPOLITANA

MODELACIÓN MATEMÁTICA **de procesos económicos.** **APLICACIONES**

Ramiro Infante Roblejo
Raúl Rodríguez Muñoz
José Enrique Martínez Serra
Pedro Manuel Ricardo Zaldívar

MODELACIÓN MATEMÁTICA **de procesos económicos.** **APLICACIONES**

Ramiro Infante Roblejo
Raúl Rodríguez Muñoz
José Enrique Martínez Serra
Pedro Manuel Ricardo Zaldívar

CIENCIAS ECONÓMICAS

Con el auspicio de la Fundación Metropolitana



FUNDACIÓN
METROPOLITANA
Fomentando la Educación Superior

MODELACIÓN MATEMÁTICA **de procesos económicos.** **APLICACIONES**

Ramiro Infante Roblejo
Raúl Rodríguez Muñoz
José Enrique Martínez Serra
Pedro Manuel Ricardo Zaldívar

Diseño de carátula: D. I. Yunisley Bruno Díaz

Composición de textos: D. I. Yunisley Bruno Díaz

Corrección: MSc. Dolores Pérez Dueñas

Edición: MSc. Isabel Gutiérrez de La Cruz

Dirección editorial: Dr. C. Jorge Luis León González

Sobre la presente edición:

© Editorial Universo Sur, 2019

© Universidad Metropolitana de Ecuador, 2019

ISBN: 978-959-257-544-8

Podrá reproducirse, de forma parcial o total, siempre que se haga de forma literal y se mencione la fuente.



Editorial: "Universo Sur".

Universidad de Cienfuegos. Carretera a Rodas, Km 3 ½.

Cuatro Caminos. Cienfuegos. Cuba.

CP: 59430

E-mail: eus@ucf.edu.cu

*“La ciencia sólo llega a la perfección
cuando consigue valerse de las matemáticas”.*

Karl Marx

Prólogo

El presente libro incursiona en temáticas relacionadas con asignaturas del área de la Matemática, y aborda contenidos esenciales que deben adquirir los profesionales de las especialidades de las Ciencias Económicas y Empresariales durante su formación curricular.

La relevancia de la obra se encuentra explícita en el tratamiento metodológico y didáctico que se da a los aspectos teóricos y prácticos relacionados con las temáticas de Programación Lineal, Programación de Meta, Pruebas de Hipótesis, Análisis de Regresión y Aplicaciones de las Derivadas, así como en lograr una integración armónica de estos métodos y modelos en sectores estratégicos de la economía, como el transporte, el agroindustrial y el agropecuario.

Los contenidos abordados se encuentran debidamente articulados y estructurados, y responden a la lógica y a la metodología que exigen estas temáticas de la Matemática Aplicada. El lenguaje técnico utilizado se ajusta al rigor científico que requieren los contenidos presentados, utilizando en la escritura términos y simbologías universalmente empleados.

El libro se encuentra estructurado en dos capítulos, referentes a la modelación matemática de los procesos económicos en el transporte y en la ganadería. El desarrollo metodológico de los capítulos se sustenta en la aplicación de la metodología de la modelación económico-matemática.

En el desarrollo inicial de los capítulos, se hace referencia a las características productivas, económicas y tecnológicas del objeto de estudio práctico, lo que permite una mejor comprensión del proceso de modelación matemática mediante la aplicación de la programación lineal y de meta lineal con prioridades.

Un aspecto significativo del libro, que contribuye a robustecer la solución de los modelos económico-matemáticos, es la determinación y validación de sus parámetros, a través del empleo de los métodos y técnicas de la Estadística, la Econometría y la Matemática. Estos métodos y técnicas, así como los modelos económico-matemáticos, se abordan teóricamente y se realizan sus aplicaciones prácticas a situaciones económicas reales.

Se presenta una amplia utilización de paquetes de programas informáticos

para la obtención de la solución de los modelos económico-matemáticos y para la estimación e inferencia de los modelos estadísticos. Además, se emplea un asistente matemático para realizar las operaciones de Álgebra Lineal. Para que exista una mejor comprensión de los resultados alcanzados, muchos de los mismos, se obtienen realizando los cálculos de forma manual y mediante la utilización de los programas informáticos.

Se utiliza de manera ajustada la bibliografía consultada y se realiza una amplia y detallada revisión de la literatura más actualizada procedente de fuentes primarias, representadas por libros y publicaciones periódicas. Además, se tuvo en cuenta la literatura clásica, de obligatoria consulta en algunas de las temáticas abordadas.

La finalidad del libro es su utilización como material de consulta para los estudiantes de pregrado de las Ciencias Económicas y Empresariales, en temas de las asignaturas de Matemática Aplicada, Inferencia Estadística, Investigación Operativa y Econometría.

Los autores

Capítulo I

Modelación económico-matemática para la planificación óptima del transporte en procesos agroindustriales

1.1. Introducción al capítulo. Aspectos esenciales de la metodología de la modelación económico-matemática

En el presente capítulo, se exponen aspectos esenciales a tener en cuenta en las fases de la metodología de la modelación económico-matemática, precisando en la formulación matemática, supuestos y elementos básicos para la construcción de un modelo de programación lineal, como un caso particular de los problemas de optimización.

A partir de las generalidades sobre la formulación y modelación de un problema de transporte, y de un estudio detallado de las características tecnológicas y económicas del proceso de transportación en la agroindustria azucarera, se realiza la modelación económico-matemática para la optimización del transporte, mediante la construcción de dos modelos de programación lineal.

Como relevante, se presenta la modelización matemática de los parámetros: capacidad industrial de los centros de recepción y costo de transportación, correspondientes a los modelos de programación lineal construidos.

La determinación y validación de la capacidad industrial se realiza mediante la estimativa e inferencia del análisis de una función de regresión lineal múltiple, representada por un modelo de regresión polinomial, donde se tienen en cuenta: la prueba de hipótesis de F de Fisher y las pruebas de los supuestos básicos de no autocorrelación, homocedasticidad y normalidad, tales como, Durbin-Watson, Breusch-Godfrey, Goldfeld-Quandt, general de la heterocedasticidad de White, Jarque-Bera y Shapiro-Wilks.

Se muestra la utilización de las derivadas a la resolución de problemas de optimización en la determinación de aprovechamiento máximo de la capacidad industrial y en el análisis de algunas de las propiedades que cumplen las funciones reales de una variable real.

En la modelización económica de la función del costo se determina una función racional de tipo lineal, y en su validación, se utiliza la prueba de hipótesis entre medias, mediante una prueba t de Student para dos muestras

independientes, y se corroboran los supuestos de normalidad y de igualdad de varianzas, a través de las pruebas de hipótesis de Shapiro-Wilks y de Levene, respectivamente.

A los contenidos tratados de programación lineal, estadística matemática, econometría, matemática y álgebra lineal inicialmente, se hace una exposición teórica y, posteriormente, en las aplicaciones prácticas, se presentan sus resultados calculados manualmente y obtenidos mediante programas informáticos correspondientes a dichas materias. En la parte final del capítulo, se hace un esbozo teórico sobre el análisis económico de la solución óptima de un modelo de programación lineal.

La planificación del transporte como actividad de la agroindustria azucarera juega un papel decisivo, tanto económico como productivo, ya que constituye el vínculo indispensable entre la agricultura y la industria. La planificación del proceso de transportación presenta sus especificidades, que radican en: la dispersión de sus objetos de trabajo, sus características heterogéneas y sus múltiples conexiones, diferentes niveles de eficiencia de los recursos productivos en dependencia de las condiciones concretas de trabajo, influencia de las condiciones climáticas y naturales, etc.

La complejidad del proceso económico-productivo del transporte en la agroindustria dificulta grandemente no solo el análisis y comprensión de dicho objeto de estudio, sino también la planificación de su desarrollo, por lo que es necesario aplicar métodos de planeación económicamente fundamentados que permitan detectar oportunamente cualquier incumplimiento y tomar decisiones para su corrección.

Una vía efectiva para lograr el perfeccionamiento de la planificación del proceso de transportación y de las actividades vinculadas a este, es mediante la aplicación de técnicas de avanzada, por medio de los métodos y modelos económico-matemáticos. Por ello, constituye una necesidad de la economía agroindustrial azucarera el empleo de técnicas modernas de gestión empresarial, adecuadas a cada escenario y basadas en las mejores y más avanzadas prácticas contemporáneas, con el propósito de garantizar la mayor eficiencia en la gestión de los procesos económicos y productivos.

La utilización de estos modernos métodos de planificación en la economía agroindustrial azucarera posibilita: seleccionar las variantes de transportación óptima, dar un uso racional a los recursos técnicos y económicos que participan en el corte mecanizado, lograr un mejor balance

de las capacidades industriales y obtener niveles máximos de producción. Dichos resultados se logran a partir de la minimización del costo total de transportación, que implícitamente incide en una disminución significativa de los costos de producción de azúcar.

Lo expresado evidencia que la utilización de los métodos y modelos económico-matemáticos permite obtener resultados superiores a los que se pueden alcanzar mediante el empleo de los métodos tradicionales de planificación.

La planificación óptima del proceso de transportación se sustenta en la utilización de la Metodología de la Modelación Económico-Matemática, la cual se fundamenta en el método de la modelación como método teórico de investigación científica. Este método de investigación juega un papel relevante en el estudio del proceso de transportación, ya que permite profundizar en la naturaleza de los fenómenos que ocurren en dicho objeto de estudio.

La metodología de la modelación económico-matemática consta de una serie de fases y etapas que deben seguirse para resolver el problema objeto de estudio. Se utilizaron como base metodológica los criterios de Moskowitz & Wright (1992); Eppen, Gould, Schmidt, Moore & Watherford (2000); Hillier & Lieberman (2010); Taha (2012); y Felipe, Ortiz, Martínez & Arias (2013), así como la experiencia de los autores del libro. Todo ello permitió definir las siguientes fases de la Metodología de la Modelación Económico Matemática:

- I. Planteamiento del problema.
- II. Construcción del modelo.
- III. Prueba del modelo.
- IV. Solución del modelo.
- V. Prueba de la solución y presentación de variantes.
- VI. Implementación y establecimiento de control y ajuste sobre los resultados.

El término problema designa una dificultad que no puede resolverse instintivamente, por lo que se requiere de una investigación conceptual o empírica. Por tanto, un problema es el primer eslabón de la cadena: problema-investigación-solución.

En la primera fase se define, desde el punto de vista cualitativo, el problema y las características que lo hacen modelable. En esta etapa del conocimiento se abarca, en un principio, al objeto como un todo, así como los nexos y relaciones exteriores que son asequibles al conocimiento empírico.

Es imprescindible en esta fase encauzar el trabajo en correspondencia con los intereses del decisor y considerar todos los factores necesarios que conformarán el modelo, lo que permitiría realizar un enfoque integral del sistema modelado. Además, se determina la técnica cuantitativa y el método de modelación matemática que efectivamente puede representar la situación problemática.

En dicha fase se requiere: realizar una descripción precisa de las metas u objetivos del estudio, identificar las variables de decisión o esenciales y los parámetros del sistema de decisión, y determinar las limitaciones o restricciones de las variables del sistema.

Definidos cualitativamente los parámetros fundamentales del modelo, se transita la segunda fase, que se caracteriza por la representación matemática del problema planteado.

El planteamiento del problema y la representación matemática de este se encuentran en unidad dialéctica, ya que la posibilidad del conocimiento de un fenómeno económico en todos sus aspectos y propiedades, comienza con la percepción sensorial de la realidad objetiva, es decir, por lo concreto sensible; pasando a lo abstracto, etapa en la cual se arriba a determinadas abstracciones que reflejan la esencia de los fenómenos económicos objeto de modelación. Después de hechas las abstracciones, el pensamiento se eleva a lo concreto, fase cualitativamente superior a la anterior, que es lo concreto pensado y se arriba así al planteamiento del modelo económico-matemático deseado.

El planteamiento del modelo es la fase más detallada de la metodología. Una vez definido el modelo que mejor se ajusta para representar el problema planteado, se especifican las relaciones cuantitativas presentes en el objeto y las restricciones del problema en términos de las variables de decisión.

La prueba del modelo consiste en comprobar si, cualitativa y cuantitativamente, sus componentes se corresponden adecuadamente con los propósitos de la investigación realizada, y persigue el objetivo de depurar al modelo para

hacerlo más representativo de la realidad, cuyos aspectos fundamentales se propone reflejar. Una característica distintiva de esta fase es su carácter de retroalimentación, ya que puede modificar el problema planteado y modelado.

Si la prueba del modelo es exitosa, se procede a obtener la solución del mismo. En caso contrario, se vuelve a la fase de planteamiento del problema o a la de construcción del modelo.

La solución del modelo planteado, cuarta fase, puede lograrse mediante la aplicación de algún procedimiento de cálculo o método de solución, tales como: el analítico, el numérico o de simulación.

Entre las fases de construcción y de solución del modelo, existe un determinado nivel de retroalimentación, pues el planteamiento matemático del problema debe satisfacer todas las condiciones que exigen el algoritmo del método de solución.

La comprobación de la solución permite incorporar mayor nivel a los resultados obtenidos y constituye un paso decisivo para demostrar su grado de aplicabilidad práctica. Por tanto, la fase de prueba de la solución y presentación de variantes es necesaria, ya que la solución obtenida no puede ser aplicada inmediatamente sin antes probar sus beneficios, pues el modelo representa a la realidad solo parcialmente y la solución no puede, consecuentemente, reflejar todos los aspectos del problema objeto de análisis.

Un momento importante es la presentación y análisis de la solución con los decisores, pues es el paso precedente de la aplicación práctica de los resultados de la investigación. Si la solución no satisface las necesidades y expectativas de los decisores, se introducen las modificaciones y ajustes correspondientes.

En la última fase, con la implantación de la solución del modelo seleccionado, se comienza a verificar la relación existente entre el objeto de la investigación y su representación, y se comienza a resolver la situación problémica. Por tanto, la investigación económica donde aplican las técnicas cuantitativas de gestión se considera como concluida cuando se ha introducido y demostrado su utilidad económica y práctica.

Si el modelo económico-matemático se utiliza más de una vez en el análisis

de los problemas de decisión, este debe revisarse cada vez que se emplee, para tener en cuenta los cambios que puedan ocurrir en las condiciones y datos del modelo; además, se debe tener en cuenta si satisface los requerimientos del decisor.

En este primer capítulo se exhibe una metodología para la planificación del proceso de transportación en la agroindustria azucarera, mediante la utilización de los métodos y modelos económico-matemáticos de optimización. Para la elaboración y desarrollo de la metodología se hace necesario procesar la información de partida y comprobar los resultados. Aquí se utiliza el Análisis Estadístico Matemático, el Análisis Econométrico y Matemático. La metodología expuesta es factible de ser aplicarla a otros procesos de transportación.

1.2. Características tecnológicas y económicas del proceso de transportación

En la planificación del proceso de transportación de la caña de azúcar es necesario tener en cuenta las características que determinan la estructura tecnológica del alza y corte mecanizado y de la recepción industrial, como actividades que se complementan en el proceso de transportación e interactúan como sistema, en el llamado *sistema de recolección de la cosecha*. Su mecanización comienza por el alza de caña cortada manualmente. La operación tecnológica del alza consiste en levantar los bultos de cañas, previamente formados por los cortadores de caña y ubicarlos sobre el medio de transporte encargado de su traslado.

Una de las operaciones tecnológicas de mayor complejidad en el proceso de recolección de la cosecha de caña lo constituye el corte mecanizado. En la realización de esta labor se utilizan máquinas cosechadoras capaces de cortar la caña, trocearla, eliminarle un buen porcentaje de cogollos y pajas, y situarlas en el medio de transporte.

Paralelamente a la introducción y difusión progresiva de la mecanización del corte, hay sistemas de recolección que utilizan un procedimiento de recepción industrial y beneficio de la caña de azúcar, por medio de los llamados centros de acopio y estaciones de limpieza, conocidos como centros de recepción industrial, que constituyen instalaciones industriales de recepción, troceado, limpieza y trasbordo de la caña.

En la transportación de la caña desde el campo hasta las instalaciones industriales o fábrica de azúcar, existe una gran variedad de métodos que van desde los más tradicionales, pasando por modernos y pesados camiones, hasta las múltiples variantes ferroviarias.

La tendencia de los sistemas de transporte de caña evoluciona en correspondencia con las particularidades locales, geográficas, ambientales, económicas y tecnológicas. En tal sentido, el desarrollo del transporte de caña en el mundo ha pasado por diferentes etapas, buscando adecuarse a las condiciones existentes en cada región en particular.

Por tales razones, el sistema de transporte a emplear está en correspondencia con las características del grado de desarrollo de la industria azucarera y con los niveles de mecanización de la agricultura cañera.

En correspondencia con lo anterior, se considera que una fábrica azucarera es moderna si emplea el transporte automotor para el traslado de la materia prima del campo a la industria. La tendencia de los países mayores productores de azúcar es mantener y desarrollar el sistema de transporte automotor por las características técnicas y económicas que este posee, al asimilar el auge acelerado de la mecanización en la cosecha cañera.

El transporte constituye el nexo decisivo para lograr el buen desarrollo y desenvolvimiento de las labores de la industria azucarera y se encuentra condicionado a los requerimientos del corte de la caña. Es imprescindible el empleo de equipos idóneos para su traslado y recepción industrial.

La selección del tipo de transporte tiene carácter económico, por lo general se recomienda que las distancias largas sean recorridas por ferrocarril, las medias y cortas por camión, y las más cortas por tractor con remolque. Además, existe la transportación con tracción animal, que constituye un tipo de transporte alternativo utilizado en distancias pequeñas, donde las condiciones concretas lo exijan.

Desde el punto de vista económico, para la elección de un medio de transporte se le tiene que atribuir un rasgo distintivo al costo de transportación y a la distancia de traslado. Otro factor relevante que incide en la selección del tipo de transporte lo constituyen las condiciones de la infraestructura vial, las cuales determinan el rendimiento de los medios de transporte.

Además de los factores anteriores, para elegir correctamente uno u otro tipo de transporte, se hace necesario examinar sus indicadores técnico- económicos más importantes, como son la capacidad de transportación, la velocidad de envío de la carga y la rentabilidad de la transportación, entre otros.

En la actualidad, cuando los países mayores productores de azúcar han acumulado un potencial económico significativo, la tarea del desarrollo económico actual y prospectivo de la agroindustria azucarera, exige del aseguramiento de la máxima eficiencia en la utilización de los recursos y entre ellos se encuentran los que participan en el proceso de transportación.

En correspondencia con lo expresado, una de las tareas de mayor relevancia en el desarrollo de la agroindustria azucarera lo constituye la elevación de su eficiencia. El aspecto central de esta tarea está representado por la reducción de los costos totales, conjuntamente con la utilización al máximo de las capacidades disponibles y la mayor producción de azúcar.

El transporte es un eslabón tan importante como cualquier otro en el sistema de fabricación de azúcar. El mismo tiene gran relevancia económica, porque los gastos generales asociados al transporte de la caña del campo a la fábrica representan un porcentaje considerable dentro del costo total de producción de azúcar, dichos gastos oscilan entre el 20 y 35 %, dependiendo de cómo se realicen los cálculos. Además, el proceso de transporte es el de mayor intensidad de fuerza de trabajo. Por tanto, el transporte constituye una condición esencial para lograr niveles de eficiencia económica en la agroindustria azucarera.

La importancia del transporte es notable, al poseer funciones indispensables en el traslado de la materia prima e insertarse como eslabón y factor de continuidad del proceso productivo, lo que determina que requiera de un alto grado de eficiencia en el cumplimiento de las tareas a él asignadas. En oposición a la producción agrícola e industrial, que se desarrollan en lugares específicos donde están ubicadas, el transporte produce mediante su tráfico en el proceso de traslado de la caña entre la agricultura y la industria.

El transporte de caña, como actividad productiva, presenta características económicas específicas al participar en el proceso productivo, ya que a diferencia de la agricultura o de la industria, no crea un nuevo producto. El mismo constituye una prolongación necesaria del proceso productivo de

la agricultura. En consecuencia, el efecto útil del transporte no se puede separar de dicho proceso productivo.

La producción de la agricultura cañera termina cuando el resultado de su proceso productivo, la caña de azúcar, es trasladada por los medios de transporte desde las áreas de producción de caña hasta las instalaciones de la industria azucarera, por lo que resulta conveniente utilizar eficientemente el transporte en esta relación agroindustrial.

En la relación agroindustrial, donde el transporte constituye el eslabón principal y desempeña un papel preponderante dentro de la producción azucarera, el costo de transportación como indicador económico para medir su efectividad, tiene una alta significación por su peso dentro de los costos totales de producción del azúcar.

El costo de transportación como criterio de eficiencia es un indicador que permite generalizar la calidad del trabajo de la transportación, pues muestra cuánto cuesta el traslado de la materia prima. En el mismo se refleja: el nivel de productividad alcanzado, el grado de eficiencia con que se usan los medios de transporte, la disciplina tecnológica, la eficiencia de la organización del trabajo, etc.

De aquí se infiere que todas las medidas encaminadas a la reducción del costo de transportación, constituyen fundamentos importantes para la elevación de la eficiencia de la producción del transporte, y en última instancia de la producción de azúcar, como producto final.

La obtención de una alta eficiencia económica en la agroindustria azucarera está indisolublemente vinculada a la más eficaz utilización de las capacidades de su equipamiento tecnológico; por lo que debe constituir un objetivo priorizado, dentro de la política económica de cualquier industria azucarera: aprovechar las capacidades tecnológicas hasta el límite máximo que permitan los recursos disponibles, con mayor eficiencia y menores costos posibles.

1.3. Formulación matemática, supuestos y elementos básicos para la construcción de un modelo de programación lineal

En el presente epígrafe y previo a la modelación económico-matemática del proceso de transportación en la agroindustria azucarera, se introducen sucintamente un conjunto de elementos y conceptos básicos relacionados con los métodos y modelos de programación lineal.

Formulación matemática y supuestos de los modelos de programación lineal

Los problemas de optimización son aquellos en los cuales se busca el máximo o mínimo de una función con un número determinado de variables, estando sus valores sujetos a ciertas limitaciones. El problema de programación lineal es un caso particular de los problemas de optimización, y puede describirse de la siguiente forma: dada una función lineal de varias variables, se quieren determinar valores no negativos para dichas variables que maximicen o minimicen el valor de la función lineal, sujeta a ciertas condiciones que asumen la forma de un sistema de ecuaciones y/o inecuaciones lineales (Pilar et al., 2013).

Considerando que n es el número de variables, y que el sistema de ecuaciones y/o inecuaciones consta de m elementos, con $m < n$, el problema de programación lineal se formula matemáticamente, como sigue:

$$\text{Min o Max } Z = C_1X_1 + C_2X_2 + \dots + C_nX_n$$

sujeito a

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n &\{ \leq, =, \geq \} b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n &\{ \leq, =, \geq \} b_2 \\ &\vdots \\ a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n &\{ \leq, =, \geq \} b_m \end{aligned}$$

$$X_j \geq 0, \quad j = 1, 2, \dots, n$$

La formulación matemática general de un problema de programación lineal se representa de la manera siguiente:

$$\text{Min o Max } Z = \sum_{j=1}^n C_j X_j \quad (I)$$

sujeito a

$$\sum_{j=1}^n a_{ij} X_j \{ \geq, =, \leq \} b_i \quad (II)$$

$$X_j \geq 0 \quad (\text{III})$$

donde:

$$i = 1, 2, \dots, m$$

$$j = 1, 2, \dots, n$$

C_j , b_i y a_{ij} son valores conocidos y constantes.

X_j son las variables esenciales o de decisión.

La función (I) se denomina función objetivo. Expresa el criterio de optimización que se utiliza en el problema, es decir, minimizar o maximizar. Los valores C_j son los coeficientes de la función objetivo o coeficientes económicos.

Al sistema de ecuaciones e inecuaciones lineales (II) se le denomina sistema de restricciones lineales, donde: b_i son los términos independientes, y a_{ij} son los coeficientes tecnológicos de las restricciones, coeficiente de las variables en las restricciones o coeficientes de conversión.

La expresión (III) se denomina condición de no negatividad de las variables e indica el requerimiento de que todas las variables esenciales o decisión sean no negativas.

Desde el punto de vista matemático no existe diferencia entre sistema de ecuaciones e inecuaciones lineales o sistema de restricciones lineales, y la condición de no negatividad de las variables, en el sentido de que ambas constituyen limitaciones a los posibles valores de X_j . La diferencia entre ambos radica en el tratamiento diferente que se le da a la condición de no negatividad de las variables cuando se procede a resolver el problema de programación lineal.

Los modelos de programación lineal cumplen los supuestos de proporcionalidad, aditividad y divisibilidad.

La *proporcionalidad* significa que los insumos requeridos para la realización de una actividad y el efecto económico que esta produce, siempre son proporcionales al nivel de actividad.

Lo anterior expresa que: si para producir una unidad de producto j se requieren de 5 horas/máquinas, y genera un efecto económico de \$ 2,00,

entonces para producir 8 unidades de producto j se requieren de 40 horas/máquinas y genera un efecto económico de \$ 16,00.

La *aditividad* expresa el efecto económico total que se obtiene, así como el total de cada recurso utilizado como resultado de la ejecución conjunta de las actividades, debe ser igual a las sumas respectivas de las cantidades resultantes de la ejecución de cada actividad.

Ilustrando este supuesto, con respecto a la utilización de los recursos, si para producir una unidad del producto uno, X_1 , se necesitan 2 horas/hombre y para producir una unidad del producto dos, X_2 , se emplean 5 horas hombre, entonces la utilización total de las horas/hombre está dada por la ejecución conjunta de las producciones, donde se suma la cantidad total de horas/hombre que son necesarias para producir una unidad del producto uno, X_1 , más la cantidad total de horas/hombres que son necesarias para producir una unidad del producto dos X_2 , resultando la operación siguiente: $2X_1 + 5X_2$.

Con respecto al efecto económico, si para producir una unidad del producto uno, X_1 , se genera un efecto económico de \$ 6,00 y para producir una unidad del producto dos, X_2 , se genera un efecto económico de \$ 11,00, entonces el efecto económico total está dado por la ejecución conjunta de las producciones, donde se suman el efecto económico total que genera producir una unidad del producto uno, X_1 , más el efecto económico total que genera producir una unidad del producto dos, X_2 , resultando la operación siguiente: $6X_1 + 11X_2$.

La divisibilidad se refiere a los valores permitidos para las variables de decisión. Las variables de decisión en un modelo de programación lineal pueden tomar valores cualesquiera, incluyendo valores no enteros, que satisfagan las restricciones funcionales y de no negatividad. Dichas variables no solo están restringidas a valores enteros, por tanto, las actividades se pueden realizar a niveles fraccionales (Moskowitz & Wright, 1992). Las variables de decisión, son variables continuas no negativas, pertenecen al conjunto numérico de los reales, y pueden tomar valores ceros y reales positivos.

Además de los supuestos de proporcionalidad, aditividad y divisibilidad, los modelos de programación lineal cumplen con la condición de ser modelos determinísticos, ya que los C_j , b_i y a_{ij} son valores conocidos y constantes.

Elementos básicos para la construcción de un modelo de programación lineal

Cuando se hayan definidos cualitativamente los parámetros fundamentales del modelo y verificado el cumplimiento de los supuestos del problema de programación lineal, se procede a la construcción del modelo que representa al planteamiento del problema objeto de estudio.

Una primera situación a tener en cuenta en la construcción del modelo es cerciorarse que la situación problemática planteada efectivamente suscita la necesidad de un modelo económico-matemático, pues existen dos tipos de errores:

- Tratar de resolver una situación compleja de múltiples alternativas por métodos empíricos e intuitivos; como consecuencia, se ofrece una solución poco fundamentada a un problema que requiere la aplicación de modelos económico-matemáticos, que permitan tomar en cuenta todos los factores condicionantes y ofrecer una solución científicamente fundamentada.
- Tratar de aplicar modelos económico-matemáticos a la solución de problemas poco complejos y con pocas alternativas, para los cuales no son requeridos.

La formulación de un problema da lugar a los planteamientos matemáticos de trabajo y general:

- El *planteamiento de trabajo* es el utilizado en la etapa experimental de la investigación y debe reflejar explícitamente la situación planteada en la formulación del problema, como por ejemplo: el problema modelado en el epígrafe 2.7.
- El *planteamiento general* expresa de forma general el problema modelado, sin expresar numéricamente los parámetros del problema, por ejemplo: los problemas modelados en los epígrafes 1.5 y 2.2.

La fase de construcción del modelo se sustenta en las etapas de definición de las variables, construcción del sistema de restricciones y planteamiento de la función objetivo.

Definición de las variables

Las variables se clasifican en variables de decisión o esenciales, y en variables de consecuencia.

Cada variable de decisión o esencial se identifica con cada una de las actividades en que se descompone el problema objeto de estudio y mediante las cuales se puede resolver en forma directa el problema que se trate. Son las variables en términos de las cuales se expresa la solución óptima del modelo matemático. Constituye el medio disponible para resolver la situación problemática.

Para definir las variables de decisión y de consecuencia se debe tener en cuenta su definición conceptual, dimensional y temporal.

Definición conceptual:

Expresa el contenido de las variables desde el punto de vista cualitativo en el contexto del problema. Para facilitar la definición resulta útil emplear el principio de unicidad: de origen, de destino, de estructura tecnológica, y de coeficiente económico en la función objetivo.

En consecuencia, una variable se puede definir cuando existen origen, destino, estructura tecnológica y coeficientes económicos, únicos, de acuerdo con el contexto del problema y con el objetivo que se persiga en el trabajo de investigación. La correcta definición de la variable permitirá identificar en forma exacta los valores numéricos de las variables de decisión en la solución óptima del problema.

A las variables de decisión se adicionan subíndices y superíndices en función de la diversidad de origen, destino, estructura tecnológica o coeficiente económico, o algún otro criterio modificador que se considere conveniente instrumentar. Por ejemplo, X_{ijk} y X_{ijk}^l .

Definición dimensional:

No basta con definir las variables por su cualidad, es necesario definir las también en términos cuantitativos y físicos, es decir, en las unidades que van a ser utilizadas para operar con estas variables.

Cuando existan varias alternativas de formas de medición, se deberá seleccionar la más representativa del problema objeto de estudio. Se deberá escoger una sola dimensión y ser consecuente con ella en todo el modelo.

Definición temporal:

Permite enmarcar la variable en un período de tiempo determinado. Se deberá escoger una misma unidad de tiempo. En problemas sencillos, es

posible obviar este tipo de definición.

Construcción del sistema de restricciones

El sistema de restricciones, formado por un sistema de ecuaciones o inecuaciones lineales, constituye el conjunto de limitaciones que restringen el objetivo primario a lograr, ya que el problema a resolver se optimiza dentro de un determinado contexto.

Procedimiento para la construcción de las restricciones:

- Establecer el valor numérico del término independiente, teniendo en cuenta su dimensión física y temporal y su significado cualitativo. El término independiente puede ser resultado de cálculos exactos, estimaciones o proyecciones.
- Establecer el sentido (signo: \geq , $=$, \leq) de la restricción.
- Analizar e incluir el conjunto de las variables esenciales o de decisión que conformarán parte de la restricción.
- Finalmente, se establecen los coeficientes tecnológicos o coeficientes de las variables en la restricción, a_{ij} , que permitan adaptar las dimensiones de las variables de decisión o esenciales a la de la restricción, para mantener la homogeneidad dimensional en la restricción; de ahí su nombre de coeficiente de conversión.

Forma de empleo de los coeficientes a_{ij} :

- El coeficiente a_{ij} puede venir dado directamente en la dimensión requerida. También puede darse indirectamente, o sea, en forma inversa, en este caso, se utilizará dividiendo a las variables de decisión por a_{ij} .
- Si el coeficiente a_{ij} no se plantea en forma directa o inversa, se debe construir basado en una determinada información. Por ejemplo, en la restricción de la capacidad unitaria de almacenamiento.
- Las restricciones de demanda, se caracterizan por tener:
 - Los coeficientes a_{ij} iguales a la unidad.
 - Los coeficientes a_{ij} que satisfagan una condición impuesta.
- En algunos problemas es necesario utilizar más de un coeficiente a_{ij} para homogenizar las dimensiones de las variables con la dimensión de la restricción.

Planteamiento de la función objetivo

La función objetivo es una función lineal que depende de todas las variables de decisión o esenciales definidas en el problema. Para su construcción, se deben tener en cuenta los supuestos de proporcionalidad, aditividad y divisibilidad de los modelos de programación lineal y se debe comprobar el análisis dimensional. La función objetivo expresa el propósito central o primario que se persigue, y es el medio a través del cual se mide la efectividad de la solución planteada.

Métodos para el planteamiento de las restricciones y función objetivo

Las interrelaciones existentes entre las variables en las restricciones y en la función objetivo son construidas mediante el método analítico-deductivo. También, puede ser necesaria la utilización de métodos estadísticos y econométricos, para la estimación y validación de los términos independientes y coeficientes del modelo.

El método analítico-deductivo consiste en establecer relaciones entre las variables. Estas relaciones son construidas tomando como punto de partida el cumplimiento de los supuestos de proporcionalidad y aditividad correspondientes a los modelos de programación lineal.

1.4. Generalidades sobre la modelación y formulación de un problema de transporte

En la práctica de la planificación de los procesos económicos, los métodos y modelos económico-matemáticos más utilizados y elaborados para la solución de problemas económicos, son los de optimización lineal. Esto se debe a que con respecto a otros modelos, estos presentan una serie de ventajas, tales como: el planteamiento matemático no resulta complejo, existe flexibilidad en la descripción matemática de los procesos económicos, se dispone de algoritmos de solución y programas computacionales confiables.

Los métodos y modelos de optimización lineal son ampliamente conocidos y difundidos en la literatura científica de investigación operacional y de los métodos cuantitativos, se dedican total o parcialmente a su estudio numerosas obras. Los contenidos más tratados son los aspectos matemáticos de los modelos, la utilización de algoritmos de solución, el modelo dual, el análisis de post-optimalidad y los problemas de transporte.

En los últimos años ha existido cierto desarrollo, a nivel empresarial, de los modelos económico-matemáticos y su aplicación en la teoría y la práctica de la planificación económica. Este desarrollo se basa en el conocimiento de la teoría de la planeación en la economía y en las nuevas posibilidades de la matemática y la computación.

Para la aplicación de la modelación económico-matemática es necesaria la práctica como fuente y base del conocimiento, pues se quieren resolver los problemas presentes en la realidad económica. Por eso es indispensable conocer esa realidad y tener conocimientos de la práctica económica para poder modelarla.

Dentro de las aplicaciones prácticas de los métodos de optimización lineal, en la modelización económica, ocupa un lugar relevante la utilización de estos métodos y modelos a los problemas de transporte, con el objetivo de que sus resultados sirvan de base para la toma de decisiones en la elaboración y fundamentación de los planes de transportación.

El problema de transporte se define como aquel en el que se requiere encontrar el método para asignar productos que están disponibles en cantidades determinadas en ciertos orígenes, a un número de destinos que tienen demanda conocida, de modo tal que el costo de transportación sea mínimo. Formalizando lo expresado, la formulación general de un problema de transporte, se refiere a cuando se desea transportar un producto homogéneo desde m orígenes hasta n destinos. En cada origen se cuenta con las cantidades a transportar O_i , mientras en cada destino se necesitan D_j cantidades de productos. Es conocido el costo de transportar una cantidad de producto de los orígenes i a los destinos j .

Este problema fue dado a conocer en 1941 por Hitchcock, en un estudio relacionado con *la distribución de un producto desde diversos orígenes a numerosas localidades*; el problema se conoce frecuentemente en la literatura como el problema de Hitchcock. A partir de entonces, se han desarrollado numerosos procedimientos para resolver el problema de transportación, entre los cuales se destacan, desde un inicio, las versiones simplificadas del Método Simplex de Dantzig, el Método de Saltamonte de Cranes y Cooper, el Método de Distribución Modificada de Ferguson y el Método de los Potenciales o de Ford-Fulkerson.

Existen otros autores, que por medio de sus obras han contribuido al desarrollo metodológico del problema de optimización del transporte, destacándose: Moskowitz & Wright (1992); Eppen, et al. (2000); Hillier & Lieberman (2010); Taha (2012); entre otros.

Los autores citados abordan en líneas generales aspectos teóricos relacionados con la formulación del problema de transporte, descripción matemática, teoremas que enuncian las propiedades del modelo de transporte, métodos para obtener la solución básica inicial, algoritmo de transporte, dualidad del problema de transporte, etc., constituyendo estas temáticas las más importantes y coincidentes. Las situaciones prácticas referidas en sus obras son sencillas e ilustrativas, aunque dan una idea de la problemática del transporte.

En las obras y publicaciones de los autores de referencia predominan trabajos de corte teórico, y las aplicaciones de los modelos económico-matemáticos son exiguas; no obstante, son trabajos de alto nivel didáctico.

Con respecto al planteamiento matemático del problema de transporte, la generalidad de los autores no rebasan los límites de la llamada forma clásica, que consiste en determinar las cantidades de productos homogéneos que pueden ser transportados desde un conjunto de puntos llamados orígenes, a otro conjunto de puntos llamados destinos, con el fin de minimizar el costo total de transportación, suponiendo conocidos los coeficientes económicos de los costos entre origen y destino, así como las existencias u ofertas y las necesidades o demandas, en los orígenes y los destinos, respectivamente.

En la representación matemática del problema clásico de los modelos económico-matemáticos aplicados a la actividad del transporte, se puede suponer la igualdad estricta entre la oferta y la demanda. En ese caso, se dice que el modelo es cerrado. En los llamados modelos abiertos, esta igualdad no es obligada, o sea, puede existir un exceso o déficit en la oferta.

Además de la situación clásica del problema de transporte, existe la posibilidad de transportar un producto por más de un medio de transporte (automotor, ferrocarril, etc.) y que la capacidad disponible total de cada uno de estos sea limitado. También es común el caso de transportar varios tipos de productos, es decir, productos heterogéneos. Teniendo en cuenta las situaciones expresadas, de manera general se puede plantear un modelo

matemático que dé como respuesta los vínculos óptimos entre los orígenes y los destinos para cada producto, así como la vía por la cual debe realizarse la transportación.

En el caso de las condiciones expuestas donde existe más de un tipo de transporte, no se puede resolver con el planteamiento clásico del modelo matemático, por lo que se hace necesaria la elaboración de modelos matemáticos que representen esta situación.

Los planteamientos del problema clásico a las diversas situaciones generales, así como los planteamientos realizados de los autores citados, proporcionaron los elementos básicos para el estudio del planteamiento del problema de los modelos exhibidos en el presente libro.

En el libro se enfocan los aspectos relacionados con la transportación del producto caña de azúcar, mediante una metodología donde se formulan y plantean modelos económico-matemáticos de programación lineal.

1.5. Formulación general del problema y modelación económico-matemática para la optimización del transporte en la agroindustria azucarera

En el epígrafe, se define de forma detallada el sistema de transportación a estudiar, de manera que se adquiere una visión global del mismo, lo que permite realizar una descripción precisa del problema, identificándose las variables de decisión y estableciéndose las limitaciones o restricciones existentes en las variables del problema. Considerando como premisa lo anterior, se construyen los modelos económico-matemáticos para la optimización del transporte en la agroindustria azucarera.

Formulación general del problema de transporte

Para garantizar el cumplimiento de los planes de producción de azúcar, como producto final del proceso productivo, se hace indispensable la transportación de la caña de azúcar, materia prima fundamental, desde un conjunto de orígenes i , llamados campos cañeros, pertenecientes a productores de caña l , hasta un conjunto de destinos j , llamados centros de recepción industrial, correspondientes a una *única agroindustria* (en caso de transportarse dentro de la misma), o a q *agroindustrias* (de existir transportaciones entre agroindustrias colindantes).

Las transportaciones entre los campos cañeros i a los centros de recepción industria j es posible realizarlas por más de un medio de transporte de tipo k , correspondientes a los transportistas p ; además, se cuenta con medios de cosecha y alza mecanizada, cuyas capacidades disponibles son limitadas.

Conociendo la producción estimada de caña de cada campo, la capacidad industrial estimada de los centros de recepción, de los tipos de transporte, de las alzadoras y de las cosechadoras, así como el costo de transportación por cada tonelada de caña de azúcar desde los campos cañeros i correspondientes a la unidad de producción l , al centro de recepción industrial j , por el tipo de transporte k , perteneciente a la unidad de transporte p ; el problema a resolver consiste en determinar los volúmenes de caña a transportar desde los campos cañeros a los centros de recepción industrial, así como la vía por la cual debe realizarse la transportación, con el propósito de minimizar el costo total de transportación.

Como una opción para lograr la integración territorial de la producción azucarera, y alcanzar niveles altos de eficiencia económica en las agroindustrias q de una región cañera, se hace indispensable conocer, cuál sería la mejor forma de vincular y transportar la caña disponible desde los campos cañeros i de las agroindustrias colindantes, a los centros de recepción industrial j de las agroindustrias azucareras q de mayor eficiencia industrial, con el fin de obtener volúmenes de producción conjunta de azúcar iguales o superiores a los niveles previstos, con el objetivo primario de minimizar el costo total de transportación, siendo conocida la cantidad de azúcar planificada a producir en conjunto por las fábricas colindantes.

A la problemática planteada se une la importancia que tiene para los productores de caña, transportistas y fábricas azucareras, establecer un plan de transportación científicamente argumentado mediante una correcta política de transportación que permita la utilización racional de los medios de transporte y de cosecha y alza mecanizada, y de las capacidades industriales instaladas, así como obtener las transportaciones más eficientes entre orígenes y destinos.

Para lograr lo antes expuesto y dar solución a la problemática planteada, se utilizan los métodos modernos de planificación, mediante el empleo de técnicas cuantitativas de gestión, como la modelación económico-matemática que se ajusta a los requerimientos anteriores a través de un modelo de programación matemática.

En este capítulo, se analizan y desarrollan modelos de programación matemática enfocados como modelos de programación lineal, relacionados con las transportaciones cañeras dentro y entre agroindustrias azucareras.

Modelación económico-matemática para la optimización del transporte

Al concluir la fase de planteamiento del problema, una vez definidos cualitativamente sus parámetros y en presencia de una cantidad de variantes de solución, que hace poco práctica e imposible la solución del problema por métodos tradicionales o convencionales, se pasa a la construcción del modelo que representa ese planteamiento, o sea, a la representación matemática del problema objeto de estudio.

En la modelación matemática de los procesos económicos y productivos, se deben considerar la mayor cantidad de factores que influyen en el problema planteado, para lograr obtener una solución adecuada.

En la determinación de los vínculos óptimos para la transportación de caña entre los campos cañeros y los centros de recepción industrial, con el fin de lograr que el proceso económico modelado sea lo más real posible, se tuvieron en cuenta los factores más importantes, tratando de no obviar aquellos que tienen mayor incidencia en la problemática, a fin de evitar decisiones erróneas, es decir, se realizaron abstracciones de los aspectos menos esenciales y relevantes.

Los factores que deben tenerse en cuenta son los siguientes: la cantidad estimada de caña de azúcar molible de los campos cañeros, la capacidad de procesamiento industrial de los centros de recepción, los rendimientos y capacidades disponibles de los tipos de transporte, de las cosechadoras y de las alzadoras, los volúmenes de producción de azúcar planificados, el rendimiento industrial, las transportaciones previstas entre los campos cañeros y centros de recepción industrial, teniendo en cuenta el tipo de transporte, la distancia media entre los campos cañeros y los centros de recepción industrial, la distancia media recorrida por unidad de carga, los gastos fijos y variables incurridos en la transportación de la caña, el costo de transportación por cada tonelada de caña, la norma operacional y potencial, y los días de zafra operacionales y potenciales.

La posibilidad de conjugar estos factores, sobre una base científicamente argumentada, se logra mediante la elaboración de modelos económico-

matemáticos de programación lineal, cuyo objetivo primario estará encaminado a minimizar el costo total de transportación como indicador económico fundamental en estos tipos de modelos.

En la elaboración y aplicación de la metodología expuesta, se utiliza el análisis sistémico, que permite realizar un enfoque integral y completo, en el estudio de los factores que influyen en las diferentes fases del proceso de transportación y, además, posibilita concretar los elementos del sistema que van a ser representados en los modelos económico-matemáticos. Mediante el análisis sistémico es factible la interrelación de la definición cualitativa de los parámetros y las relaciones matemáticas de los modelos económico-matemáticos aplicados.

Considerando el análisis realizado sobre el objeto de estudio en la formulación del problema, se construyen dos tipos de modelos económico-matemáticos: uno relacionado con la optimización del transporte en una agroindustria azucarera y otro relacionado con la optimización del transporte entre agroindustrias azucareras colindantes.

Para el planteamiento matemático de los modelos de optimización del transporte, Moskowitz & Wright (1992); Eppen, et al. (2000), Bazaraa, Jarvis & Sherali (2009); Hillier & Lieberman (2010); Taha (2012); y Felipe, et al. (2013), contribuyeron a su planteamiento teórico; y García, et al. (2013); Mussone & Wolfler (2013); Steadie, Dellaert, Nuijten, Van Woensel & Raoufi (2014); y Ribeiro & Morabito (2017), aportaron los elementos prácticos relacionados con la concepción del sistema de restricción, planificación óptima del transporte, cálculo de capacidad de los medios de transporte y optimización de la cosecha.

Modelo económico-matemático para la optimización del transporte en una agroindustria azucarera (Modelo I)

Índices:

$i = 1, 2, \dots, n$ Campos cañeros.

$j = 1, 2, \dots, m$ Centros de recepción industrial.

$k = 1, 2, \dots, f$ Tipos de transporte.

$l = 1, 2, \dots, g$ Productores de caña.

$p = 1, 2, \dots, h$ Transportistas.

Parámetros del modelo:

O_{il} : Producción total estimada molible a cortar en el campo cañero i del productor l , durante la zafra azucarera, en toneladas de caña.

D_j : Capacidad disponible total estimada del centro de recepción industrial j , durante la zafra azucarera, en toneladas de caña.

T_{kp} : Capacidad disponible total estimada del tipo de transporte k perteneciente al transportista p , durante la zafra azucarera, en toneladas de caña.

A_l : Capacidad disponible total estimada de las alzadoras correspondientes al productor l , durante la zafra azucarera, en toneladas de caña.

C_l : Capacidad disponible total estimada de corte de las cosechadoras correspondientes al productor l , durante la zafra azucarera, en toneladas de caña.

C_{ijk}^{lp} : Costo de transportación por toneladas de caña de azúcar a transportar desde el campo cañero i , correspondiente al productor l , al centro de recepción industrial j , por el tipo de transporte k perteneciente al transportista p .

Variables de decisión:

X_{ijk}^{lp} : Toneladas de caña de azúcar a transportar desde el campo cañero i correspondiente al productor l , al centro de recepción industria j , por el tipo de transporte k , del transportista p , durante la zafra azucarera.

Sistema de restricciones:

Restricción de tipo I: Cumplimiento de la tarea de corte por campo cañero.

$$\sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lp} = O_{il}; \quad i = 1, 2, \dots, n \quad l = 1, 2, \dots, g$$

Restricción de tipo II: Disponibilidad de la capacidad estimada de los centros de recepción.

$$\sum_{i=1}^n \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h X_{ijk}^{lp} \leq D_j; \quad j = 1, 2, \dots, m$$

Restricción de tipo III: Disponibilidad de la capacidad estimada de los medios de transporte.

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{l=1}^g X_{ijk}^{lp} \leq T_{kp}; \quad k = 1, 2, \dots, f \quad p = 1, 2, \dots, h$$

Restricción de tipo IV: Disponibilidad de la capacidad estimada de los equipos de alza.

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lp} \leq A_l; \quad l = 1, 2, \dots, g$$

Restricción de tipo V: Disponibilidad de la capacidad estimada de las cosechadoras.

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lp} \leq C_l; \quad l = 1, 2, \dots, g$$

Restricción de tipo VI: Condición de no negatividad de las variables.

$$X_{ijk}^{lp} \geq 0; \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m, \quad k = 1, 2, \dots, f, \quad l = 1, 2, \dots, g$$

$$p = 1, 2, \dots, h.$$

Función objetivo

Se desarrolló una función objetivo que tiene como propósito fundamental, optimizar o minimizar el costo total de transportación entre los campos cañeros y los centros de recepción industrial. La función lineal a optimizar es la siguiente:

$$\text{Min } Z = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h C_{ijk}^{lp} X_{ijk}^{lp}$$

Las restricciones de tipo I: Indican que el volumen de caña de azúcar a transportar por los diferentes medios de transporte hacia los centros de recepción industrial, tiene que coincidir con el volumen de caña de azúcar a cosechar por los campos cañeros durante la zafra. Se establece una restricción para cada campo cañero.

Las restricciones de tipo II: Establecen el balance de capacidad industrial entre el volumen de caña estimada de todos los campos cañeros con la capacidad industrial disponible estimada de los centros de recepción durante la zafra. Se considera una restricción por cada centro de recepción.

Las restricciones de tipo III: Plantean el acotamiento de la capacidad disponible estimada por tipo de transporte, es decir, el volumen máximo posible de caña de azúcar a transportar durante la zafra entre los campos cañeros y los centros de recepción por el medio de transporte considerado.

Las restricciones de tipo IV: Establecen el acotamiento de la capacidad disponible estimada de los equipos de alza mecanizada durante la zafra azucarera, o sea, el volumen de caña de azúcar máximo posible a alzar.

Las restricciones de tipo V: Limitan la capacidad disponible estimada de las cosechadoras, indicando el volumen máximo posible de caña de azúcar a cosechar mecanizadamente durante el período de zafra.

Presentación resumida del planteamiento general del modelo económico-matemático:

$$\text{Min } Z = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h C_{ijk}^{lp} X_{ijk}^{lp}$$

suje to a

$$\sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lp} = O_{il}; \quad i = 1, 2, \dots, n, \quad l = 1, 2, \dots, g.$$

$$\sum_{i=1}^n \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h X_{ijk}^{lp} \leq D_j; \quad j = 1, 2, \dots, m.$$

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{l=1}^g X_{ijk}^{lp} \leq T_{kp}; \quad k = 1, 2, \dots, f, \quad p = 1, 2, \dots, h.$$

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lp} \leq A_l; \quad l = 1, 2, \dots, g.$$

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lp} \leq C_l; \quad l = 1, 2, \dots, g.$$

$$X_{ijk}^{lp} \geq 0; \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m, \quad k = 1, 2, \dots, f, \quad l = 1, 2, \dots, g, \\ p = 1, 2, \dots, h.$$

Modelo económico-matemático para la optimización del transporte entre agroindustrias azucareras (Modelo II)

Índices:

$i = 1, 2, \dots, n$ Campos cañeros.

$j = 1, 2, \dots, m$ Centros de recepción industrial.

$k = 1, 2, \dots, f$ Tipos de transporte.

$l = 1, 2, \dots, g$ Productores de caña.

$p = 1, 2, \dots, h$ Transportistas.

$q = 1, 2, \dots, r$ Agroindustrias azucareras.

Parámetros del modelo:

O_{ilq} : Producción total estimada molible a cortar en el campo cañero i del productor l correspondiente a la agroindustria q , durante la zafra azucarera, en toneladas de caña.

D_{jq} : Capacidad disponible total estimada del centro de recepción industrial j de la agroindustria q , durante la zafra azucarera, en toneladas de caña.

T_{kpq} : Capacidad disponible total estimada del tipo de transporte k perteneciente al transportista p de la agroindustria q , durante la zafra azucarera, en toneladas de caña.

A_{lq} : Capacidad disponible total estimada de las alzadoras correspondientes al productor l de la agroindustria q , durante la zafra azucarera, en toneladas de caña.

C_{lq} : Capacidad disponible total estimada de corte de las cosechadoras correspondientes al productor l de la agroindustria q , durante la zafra azucarera, en toneladas de caña.

S : Azúcar prevista a producir en conjunto por las agroindustrias durante la zafra azucarera, en toneladas de azúcar.

R_q : Rendimiento industrial promedio en la agroindustria q .

C_{ijk}^{lpq} : Costo de transportación por toneladas de caña de azúcar a transportar desde el campo cañero i , correspondiente al productor l , al centro de recepción industrial j correspondiente a la agroindustria q , por el tipo de transporte k perteneciente al transportista p .

Variables de decisión:

X_{ijk}^{lpq} : Toneladas de caña de azúcar a transportar desde el campo cañero i correspondiente al productor l , al centro de recepción industrial j correspondiente a la agroindustria q , por el tipo de transporte k del transportista p , durante la zafra azucarera.

Sistema de restricciones:

Restricción de tipo I: Cumplimiento de la tarea de corte por campo cañero.

$$\sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lpq} = O_{ilq}; \quad i = 1, 2, \dots, n, \quad l = 1, 2, \dots, g, \quad q = 1, 2, \dots, r.$$

Restricción de tipo II: Disponibilidad de la capacidad estimada de los centros de recepción industrial.

$$\sum_{i=1}^n \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h X_{ijk}^{lpq} \leq D_{jq}; \quad j = 1, 2, \dots, m, \quad q = 1, 2, \dots, r.$$

Restricción de tipo III: Disponibilidad de la capacidad estimada de los medios de transporte.

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{l=1}^g X_{ijk}^{lpq} \leq T_{kpq}; \quad k = 1, 2, \dots, f, \quad p = 1, 2, \dots, h, \quad q = 1, 2, \dots, r$$

Restricción de tipo IV: Disponibilidad de la capacidad estimada de los equipos de alza.

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lpq} \leq A_{lq}; \quad l = 1, 2, \dots, g, \quad q = 1, 2, \dots, r.$$

Restricción de tipo V: Disponibilidad de la capacidad estimada de las cosechadoras.

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lpq} \leq C_{lq}; \quad l = 1, 2, \dots, g, \quad q = 1, 2, \dots, r.$$

Restricción de tipo VI: Producción conjunta mínima de azúcar entre las agroindustrias.

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h \sum_{q=1}^r R_q \cdot X_{ijk}^{lpq} \geq S$$

Restricción de tipo VII: Condición de no negatividad de las variables.

$$X_{ijk}^{lpq} \geq 0; \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m, \quad k = 1, 2, \dots, f, \quad l = 1, 2, \dots, g, \\ p = 1, 2, \dots, h, \quad q = 1, 2, \dots, r.$$

Función objetivo

Se desarrolló una función objetivo que tiene como propósito fundamental optimizar o minimizar el costo total de transportación entre los campos cañeros y los centros de recepción. La función lineal a optimizar es la siguiente:

$$\text{Min } Z = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h \sum_{q=1}^r C_{ijk}^{lpq} X_{ijk}^{lpq}$$

Este modelo posibilita la vinculación entre agroindustrias colindantes y se incluye un nuevo tipo de restricción con respecto al modelo anterior: La restricción de tipo VI, que expresa la meta mínima de producción conjunta de azúcar entre las agroindustrias colindantes y garantiza que la producción de azúcar resultante de las agroindustrias sea igual o superior a la cantidad prevista.

Presentación resumida del planteamiento general del modelo económico-matemático:

$$\text{Min } Z = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h \sum_{q=1}^r C_{ijk}^{lpq} X_{ijk}^{lpq}$$

sujeto a

$$\sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lpq} = O_{ilq}; \quad i = 1, 2, \dots, n, \quad l = 1, 2, \dots, g, \quad q = 1, 2, \dots, r.$$

$$\sum_{i=1}^n \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h X_{ijk}^{lpq} \leq D_{jq}; \quad j = 1, 2, \dots, m, \quad q = 1, 2, \dots, r.$$

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{l=1}^g X_{ijk}^{lpq} \leq T_{kpq}; \quad k = 1, 2, \dots, f, \quad p = 1, 2, \dots, h, \quad q = 1, 2, \dots, r$$

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lpq} \leq A_{lq}; \quad l = 1, 2, \dots, g, \quad q = 1, 2, \dots, r.$$

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{p=1}^h X_{ijk}^{lpq} \leq C_{lq}; \quad l = 1, 2, \dots, g, \quad q = 1, 2, \dots, r.$$

$$\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^f \sum_{l=1}^g \sum_{p=1}^h \sum_{q=1}^r R_q \cdot X_{ijk}^{lpq} \geq S.$$

$$X_{ijk}^{lpq} \geq 0; \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m, \quad k = 1, 2, \dots, f, \quad l = 1, 2, \dots, g, \\ p = 1, 2, \dots, h, \quad q = 1, 2, \dots, r.$$

Los modelos económico-matemáticos expuestos para la planificación del transporte proporcionan cambios cuantitativos que permiten un análisis económico más depurado y completo, y facilitan una amplia y clara valoración de los resultados, sobre la base de una información más exacta; conduciendo a cambios cualitativos en la metodología de la planificación del transporte.

La metodología utilizada en la construcción de los modelos económico-matemáticos puede ser utilizada en sistemas de transportación de actividades económicas correspondientes a otras ramas de la economía.

Después de haber detallado el planteamiento matemático de los modelos económico-matemáticos para la optimización del transporte, y previo a la conformación del modelo numérico, se presentan de forma exhaustiva los aspectos teóricos y prácticos de la metodología utilizados para la determinación y validación de dos parámetros relevantes seleccionados de los modelos:

1 - Capacidad industrial disponible estimada de los centros de recepción

(D_j y D_{jq}).

2 - Costo de transportación (C_{ijk}^{lp} y C_{ijk}^{lpq}).

La determinación y validación del término independiente de las restricciones de tipo II y del coeficiente económico de la función objetivo, descritos anteriormente, se realizan aplicando métodos y modelos de Inferencia Estadística, de Análisis Econométrico, del Cálculo Diferencial y del Álgebra Lineal, auxiliados por software estadísticos y otros asistentes matemáticos, como se exhibe en los epígrafes 1.6 y 1.7.

1.6. Modelización de la capacidad industrial de los centros de recepción. Determinación y validación

1.6.1. Generalidades, estimativa e inferencia de la función de regresión

La utilización de las capacidades industriales de procesamiento instaladas de los centros receptores de caña es un indicador técnico-económico importante, pues estas trabajan ininterrumpidamente durante todo el período de zafra.

Como resultado del análisis de las series de tiempo sobre la utilización de las capacidades industriales, se determinó que generalmente, durante los primeros días de zafra, la capacidad industrial total disponible no se aprovecha a un alto nivel debido a los ajustes de inicio de la zafra y preparación de la molienda. A partir de aquí, el aprovechamiento de las capacidades tiende a incrementarse hasta alcanzar valores máximos.

Sin embargo, durante las semanas finales de la zafra, el aprovechamiento de las capacidades industriales comienza a disminuir, al incrementarse el tiempo perdido por condiciones técnicas de los equipos y de la maquinaria y por la influencia de las condiciones climatológicas de algunas regiones, entre otras razones.

Teniendo en cuenta los factores mencionados y por la complejidad de los procesos agroindustriales azucareros, se hace imprescindible el estudio de cada centro de recepción para determinar qué parte de la capacidad industrial potencial puede ser utilizada óptimamente durante la zafra.

La utilización óptima y eficiente de las capacidades instaladas tiene gran importancia económica y productiva, pues además del beneficio que estos centros de recepción le generan a la materia prima, facilitan una organización territorial más racional de la agricultura cañera, ya que alrededor de estos se realiza el programa de plantación y cosecha; garantizando una mejor organización del sistema de cosecha, como complemento a la labor de las cosechadoras y de los medios de transporte, y posibilita un incremento de la productividad del corte manual, simplificando el mismo.

Además de lo referido anteriormente, y debido al elevado costo de operación y mantenimiento que generan los centros de recepción industrial, se impone a la agroindustria azucarera su utilización racional y eficiente. Para lograr esto, es necesario hacer un estudio de dichos centros aplicando métodos y modelos matemáticos.

El análisis estadístico matemático a utilizar en el estudio de las capacidades industriales de los centros de recepción es la estimativa e inferencia del análisis de los modelos de Regresión Lineal Múltiple. Para determinar el modelo matemático de mejor ajuste, se tienen en cuenta: el coeficiente de determinación máximo, la prueba de hipótesis F de Fisher, así como las pruebas Durbin-Watson, Breusch-Godfrey, Goldfeld-Quandt, general de la heterocedasticidad de White, Jarque-Bera y Shapiro-Wilks. Este análisis estadístico matemático se apoya en la aplicación del paquete de programa estadístico y econométrico, EViews versión 3.1, y del paquete de programa estadístico SPSS (Statistical Package for the Social Sciences) versión 23.0, que posibilitan obtener los resultados estadísticos de los modelos matemáticos a verificar.

Para la aplicación de los paquete de programa estadístico y econométrico mencionado, se utiliza como consulta del EViews a Carrascal, González & Rodríguez (2001); y como consulta del SPSS a Gil (2017); y la Guía del usuario de IBM SPSS (2017).

Para determinar la capacidad industrial estimada de los centros de recepción con respecto a su utilización histórica en función del tiempo, se calcula previamente el volumen de caña promedio semanal procesada industrialmente, en al menos cinco zafras, ajustando el comportamiento del promedio semanal mediante la aplicación del Método de los Mínimos Cuadrados Ordinarios, por medio de los programas de computación

mencionados, que posibilitan obtener los resultados estadísticos de los modelos matemáticos a probar.

Los modelos matemáticos comprobados con la aplicación de los procedimientos metodológicos son funciones de tipo: lineal, cuadrática, cúbica, exponencial, logarítmica, entre las más representativas. Los modelos de las funciones cuadráticas (polinomios de segundo grado) y cúbicas (polinomios de tercer grado) examinadas, exhiben los coeficientes de determinación R^2 más altos, con valores superiores a 0,81. Los modelos de las funciones restantes presentan coeficientes de determinación inferiores a 0,49.

El comportamiento de la curva asociada al aprovechamiento de la capacidad industrial de los centros de recepción se puede ajustar a un modelo cuadrático o cúbico en función de los factores inicialmente mencionados, o sea, a partir del promedio de la caña procesada semanalmente, es posible ajustar curvas continuas representadas por las funciones de regresión polinomiales estimadas, tales como: $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t - \hat{\beta}_3 X_t^2$ o $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t + \hat{\beta}_3 X_t^2 - \hat{\beta}_4 X_t^3$,

donde, \hat{Y}_t es la variable regresada o respuesta, capacidad industrial estimada del centro de recepción, (en toneladas de caña), X_t es la variable regresora, explicativa o predictora, tiempo (en semanas), y $\hat{\beta}_j$, con $j = 2,3,4$ son los coeficientes de regresión parcial. Los signos de la función dependen del comportamiento de la serie de los datos u observaciones ajustadas.

Las funciones de regresión polinomiales en la situación objeto de estudio:

- están definidas solo para los números reales positivos, o sea, el $\text{Dom } f = \{x \in \mathfrak{R} : x > 0\}$, y
- toman solamente valores positivos, es decir, la $\text{Im } f = \{y \in \mathfrak{R} : y > 0\}$.

Los modelos de regresión polinomiales $E_{(Y_t/X_t)} = Y_t = \beta_1 + \beta_2 X_t - \beta_3 X_t^2 + U_t$ o $E_{(Y_t/X_t)} = Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_t^2 - \beta_4 X_t^3 + U_t$, reseñados, son modelos donde existe una sola variable explicativa, X_t , en el miembro derecho de la función, la cual se encuentra elevada a distintas potencias X_t , X_t^2 y X_t^3

lo cual convierte a dichas funciones en modelos de regresión lineal múltiples.

Estos modelos de regresión múltiple, modelos de regresión polinomial, polinomios de segundo y tercer grado, son lineales con respecto a los parámetros, por consiguiente las β_j pueden ser estimadas mediante la aplicación de las metodologías usuales del Método de los Mínimos Cuadrados Ordinarios o por el Método de Máxima Verosimilitud sin que se presenten problemas especiales de estimación (Gujarati, 2003).

Seleccionar una u otra función de regresión polinomial estimada para la proyección de la capacidad industrial de los centros de recepción es válido. No obstante, es factible seleccionar la función de regresión polinomial cuadrática, $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t - \hat{\beta}_3 X_t^2$ por los resultados satisfactorios obtenidos con la verificación de las pruebas estadísticas, y porque, como ventaja adicional, poseer una expresión más sencilla, lo que facilita los cálculos.

En el estudio teórico y práctico de la estimativa, inferencia y validación de la función de regresión correspondiente al presente capítulo, se toma como fundamento los procedimientos metodológicos abordados en los temas de econometría relacionados con la regresión lineal múltiple y con los supuestos básicos, desarrollados por Gujarati (2003); Wooldridge (2008); Gujarati & Porter (2010); y Espallarga & Solís (2012).

Para el desarrollo del Método de los Mínimos Cuadrados, de las pruebas de hipótesis que validan a los modelos de regresión y de las aplicaciones de la estadística a los procesos económicos, se consideran los procedimientos expuestos por Anderson, Sweeney & Williams (2008); Wackerly, Mendenhall & Scheaffer (2010); y Díaz (2012).

Estimadores mínimos cuadráticos de la función de regresión. Determinación por la vía matricial

La regresión por la vía matricial es válida para cualquier modelo que presenta linealidad con respecto a los parámetros. Dado el modelo general de regresión polinomial de segundo grado $E_{(Y_t/X_t)} = Y_t = \beta_1 + \beta_2 X_t - \beta_3 X_t^2 + U_t$ se obtiene para cada par ordenado (x_i, y_i) , donde $i = 1, 2, 3, \dots, n$, un conjunto

de ecuaciones simultáneas, representado por el siguiente sistema de sistema de n ecuaciones:

$$\begin{aligned}
 Y_1 &= \beta_1 + \beta_2 X_1 - \beta_3 X_1^2 + U_1 \\
 Y_2 &= \beta_1 + \beta_2 X_2 - \beta_3 X_2^2 + U_2 \\
 Y_3 &= \beta_1 + \beta_2 X_3 - \beta_3 X_3^2 + U_3 \\
 \cdot &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
 Y_n &= \beta_1 + \beta_2 X_n - \beta_3 X_n^2 + U_n
 \end{aligned}$$

Representación en notación matricial detallado del sistema de ecuaciones:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ 1 & X_3 & X_3^2 \\ \cdot & \cdot & \cdot \\ 1 & X_n & X_n^2 \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ \dots \\ U_n \end{bmatrix}$$

El sistema de ecuaciones que representa el modelo de regresión lineal

$Y = \beta_1 + \beta_2 X + \beta_3 X^2 + U_i$ se puede representar en notación matricial condensa mediante la ecuación matricial siguiente: $Y = X \cdot \beta + U$

donde:

Y es una matriz columna de orden $n \times 1$, que representa las observaciones de la variable dependiente.

X es una matriz de orden $n \times p$, que representa las observaciones de la variable explicativa, siendo p el número de parámetros del modelo.

β es una matriz columna de orden $p \times 1$, que representa los p parámetros del modelo.

U es una matriz columna de orden $n \times 1$, que representa los errores aleatorios.

La ecuación matricial $Y = X \cdot \beta + U$ formula cualquier modelo de regresión lineal. Entonces $\hat{Y} = X \cdot \hat{\beta}$ representa la función de regresión estimada, donde la matriz columna β constituye la matriz de los estimadores mínimos cuadrados de los parámetros. Esta matriz se obtiene mediante el principio de los mínimos cuadrados, donde se persigue minimizar la suma de los cuadrados de los errores de estimación o residuos, es decir, se minimiza la $\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$, y seleccionar los estimadores representados en

β de tal forma que para un conjunto muestral de información la $\sum_{i=1}^n e_i^2$ es la más pequeña posible.

Mediante la aplicación del cálculo diferencial se obtiene la siguiente derivada parcial, $\frac{\partial e^T e}{\partial \underline{\hat{\beta}}} = -2\underline{X}^T \underline{Y} + 2\underline{X}^T \underline{X} \underline{\hat{\beta}}$ (lo subrayado indica que se

encuentra en notación matricial). Igualando la derivada a cero se obtiene la ecuación matricial $\underline{X}^T \underline{X} \underline{\hat{\beta}} = \underline{X}^T \underline{Y}$, la cual permite la determinación de los estimadores mínimos cuadrados.

El desarrollo de $\underline{X}^T \underline{X} \underline{\hat{\beta}} = \underline{X}^T \underline{Y}$ conduce a un sistema de ecuaciones, denominado ecuaciones normales, formado por igual cantidad de ecuaciones que de incógnitas tenga el modelo. Para los modelos de regresión polinomiales de segundo grado, el sistema de ecuaciones tiene tres incógnitas, siendo estas los estimadores mínimos cuadrados de los parámetros.

Considerando la representación de la ecuación matricial $\underline{X}^T \underline{X} \underline{\hat{\beta}} = \underline{X}^T \underline{Y}$, para la obtención de los estimadores mínimos cuadrados del modelo de regresión lineal múltiple $Y = \beta_1 + \beta_2 X + \beta_3 X^2 + U_i$, tenemos que:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ 1 & X_3 & X_3^2 \\ \dots & \dots & \dots \\ 1 & X_n & X_n^2 \end{bmatrix} \quad X^T = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ X_1 & X_2 & X_3 & \dots & X_n \\ X_1^2 & X_2^2 & X_3^2 & \dots & X_n^2 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$$

Sustituyendo las matrices Y , X , X^T y β en la ecuación matricial

$X^T X \beta = X^T Y$, se obtiene la igualdad:

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ X_1 & X_2 & X_3 & \dots & X_n \\ X_1^2 & X_2^2 & X_3^2 & \dots & X_n^2 \end{bmatrix} \begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ 1 & X_3 & X_3^2 \\ \dots & \dots & \dots \\ 1 & X_n & X_n^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ X_1 & X_2 & X_3 & \dots & X_n \\ X_1^2 & X_2^2 & X_3^2 & \dots & X_n^2 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{bmatrix}$$

Realizando en la ecuación matricial las multiplicaciones de las matrices $X^T X$ y $X^T Y$, se obtiene el siguiente resultado:

$$\begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i} X_{3i} \\ \sum X_{3i} & \sum X_{3i} X_{2i} & \sum X_{3i}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \sum X_{3i} Y_i \end{bmatrix}$$

La matriz $(X^T X)$ es una matriz cuadrada y simétrica.

Calculando el producto de las matrices $(X^T X)$ y β , se obtiene:

$$\begin{bmatrix} n\hat{\beta}_1 & + & \hat{\beta}_2 \sum X_{2i} & + & \hat{\beta}_3 \sum X_{3i} \\ \hat{\beta}_1 \sum X_{2i} & + & \hat{\beta}_2 \sum X_{2i}^2 & + & \hat{\beta}_3 \sum X_{2i} X_{3i} \\ \hat{\beta}_1 \sum X_{3i} & + & \hat{\beta}_2 \sum X_{3i} X_{2i} & + & \hat{\beta}_3 \sum X_{3i}^2 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \sum X_{3i} Y_i \end{bmatrix}$$

Aplicando la igualdad de matrices lo anterior conduce a un sistema de ecuaciones normales, tal como:

$$\begin{aligned} n\hat{\beta}_1 &+ \hat{\beta}_2 \sum X_{2i} + \hat{\beta}_3 \sum X_{3i} = \sum Y_i \\ \hat{\beta}_1 \sum X_{2i} &+ \hat{\beta}_2 \sum X_{2i}^2 + \hat{\beta}_3 \sum X_{2i} X_{3i} = \sum X_{2i} Y_i \\ \hat{\beta}_1 \sum X_{3i} &+ \hat{\beta}_2 \sum X_{3i} X_{2i} + \hat{\beta}_3 \sum X_{3i}^2 = \sum X_{3i} Y_i \end{aligned}$$

Igualando cada ecuación a cero, o sea:

$$\begin{aligned} \sum Y_i - n\hat{\beta}_1 - \hat{\beta}_2 \sum X_{2i} - \hat{\beta}_3 \sum X_{3i} &= 0 \\ \sum X_{2i} Y_i - \hat{\beta}_1 \sum X_{2i} - \hat{\beta}_2 \sum X_{2i}^2 - \hat{\beta}_3 \sum X_{2i} X_{3i} &= 0 \\ \sum X_{3i} Y_i - \hat{\beta}_1 \sum X_{3i} - \hat{\beta}_2 \sum X_{3i} X_{2i} - \hat{\beta}_3 \sum X_{3i}^2 &= 0 \end{aligned}$$

Resolviendo el sistema de ecuaciones normales se obtienen los estimadores

mínimos cuadrados $\hat{\beta}_1$, $\hat{\beta}_2$ y $\hat{\beta}_3$. La resolución se puede efectuar aplicando cualquiera de los métodos estudiados en Álgebra Lineal, siempre que estén dadas las condiciones para ello.

Cálculo de la matriz columna $\beta = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$

Partiendo de:

$$(X^T X)\beta = X^T Y$$

$$\begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i} X_{3i} \\ \sum X_{3i} & \sum X_{3i} X_{2i} & \sum X_{3i}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \sum X_{3i} Y_i \end{bmatrix}$$

Multiplicando cada miembro de $(X^T X)\beta = (X^T Y)$ por la matriz inversa de $(X^T X)$; es decir $(X^T X)^{-1}$, y aplicando las propiedades de las matrices, se obtiene:

$$\begin{aligned} (X^T X)^{-1}(X^T X)\beta &= (X^T X)^{-1}(X^T Y) \\ I_n \cdot \beta &= (X^T X)^{-1}(X^T Y) \\ \beta &= (X^T X)^{-1}(X^T Y) \end{aligned}$$

donde:

I_n es la matriz idéntica de orden tres.

A partir de la aplicación de $\beta = (X^T X)^{-1}(X^T Y)$ se obtienen los estimadores mínimos cuadrados $\hat{\beta}_1$, $\hat{\beta}_2$ y $\hat{\beta}_3$, de la función de regresión polinomial estimada $\hat{E}_{(Y_i/X_i)} = \hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i - \hat{\beta}_3 X_i^2$ correspondiente a la

situación objeto de estudio.

Validación de la función de regresión estimada. Prueba F de Fisher

Dada la necesidad de tomar decisiones sobre la población a partir de información de la muestra, se plantean hipótesis que representan un conjunto de aseveraciones sobre la población que pueden ser verificadas, constituyendo su objetivo fundamental validar los resultados obtenidos en la muestra a partir de la función matemática seleccionada con respecto a la población de procedencia.

Una vez determinada la función de regresión polinomial cuadrática

$\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t - \hat{\beta}_3 X_t^2$, correspondiente al problema analizado, es necesario el análisis de su precisión mediante la significación de la regresión. La validación del modelo se plantea por la vía matricial, es decir, se verificará si la variable independiente es capaz de explicar significativamente el comportamiento de la variable dependiente.

Con el objetivo de determinar si el modelo explica o no el fenómeno objeto de estudio, se realiza el análisis de varianza en la regresión mediante la descomposición de la variación total, y se realiza el planteamiento de una prueba de hipótesis, donde la hipótesis nula que se somete a prueba representa que el modelo no explica el fenómeno en estudio, versus la hipótesis alternativa a contrastar que representa que el modelo sí explica el fenómeno en estudio.

Estudiar el análisis de regresión desde el punto de vista del análisis de varianza, representa un medio esclarecedor y complementario de analizar el problema de la inferencia estadística.

El análisis de varianza, desde el punto de vista de la regresión, se basa en la descomposición de la suma de cuadrados total (SC_T) en la suma de cuadrados explicada (SC_E), y en la suma de cuadrados residual o del error (SC_R), expresado por la ecuación $SC_T = SC_E + SC_R$.

Asociados a la suma de cuadrados se encuentran sus grados de libertad (gl), o número de observaciones independientes sobre las cuales está basada, como se muestra a continuación:

- SC_T tiene $n-1$ grados de libertad, pues se pierde un grado de libertad al calcular la media muestral de \bar{Y} .
 - SC_R tiene $n-p$ grados de libertad.
 - SC_E tiene $p-1$ grados de libertad.
- donde: n es el número de observaciones, p es el número de parámetros que se estiman.

Observe que la suma en la ecuación $SC_T = SC_E + SC_R$ tiene asociada los siguientes grados de libertad $n-1 = (n-p) + (p-1)$.

En la tabla de análisis de varianza (ANOVA) se presenta la suma de cuadrados en forma matricial, resultado de la aplicación del análisis de varianza por la vía matricial en la regresión.

Tabla 1.1. Análisis de Varianza (ANOVA) por la vía matricial.

Fuente de variación	Suma de cuadrados SC	Grados de libertad gl	Suma de cuadrados medios SCM
Debido a la regresión (SC_E)	$SC_E = \underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2$	$(p-1)$	$SCM \text{ de } SC_E = \frac{SC_E}{p-1} = S_E^2$
Debido a los residuos (SC_R)	$SC_R = \underline{Y}^T \underline{Y} - \underline{\beta}^T \underline{X}^T \underline{Y}$ $SC_R = SC_T - SC_E$	$(n-p)$	$SCM \text{ de } SC_R = \frac{SC_R}{n-p} = S_R^2$
Total (SC_T)	$SC_T = \underline{Y}^T \underline{Y} - n\bar{Y}^2$	$(n-1)$	////

Fuente: elaborada por los autores.

Lo subrayado indica que se encuentra en notación matricial.

donde:

SC_E es la suma de cuadrados debido a la regresión o explicada por la regresión.

SC_R es la suma de cuadrados debido a los residuos o errores de estimación.

$p-1$ son los grados de libertad del numerador.

$n-p$ son los grados de libertad del denominador.

\underline{Y}^T , $\underline{\beta}^T$ y \underline{X}^T son las matrices transpuestas de \underline{Y} , $\underline{\beta}$ y \underline{X} , respectivamente.

\underline{Y} es la matriz columna de la variable dependiente.

S_E^2 es la varianza explicada por la regresión.

S_R^2 es la varianza debido a los residuos.

A partir de los resultados del análisis de varianza de la regresión, se determina el estadístico de prueba o estadígrafo F . El valor de F calculado, estimado, obtenido u observado, se obtiene mediante el cociente de la varianza explicada por la regresión y la varianza debido a los residuos, es decir,

$$F = \frac{S_E^2}{S_R^2} \quad \text{o} \quad F = \frac{\frac{SC_E}{p-1}}{\frac{SC_R}{n-p}} = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R}$$

La región crítica para el estadígrafo F , dado el nivel de significación α , se encuentra dada por:

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0 : F > F_{1-\alpha}(p-1, n-p) \right\}$$

El valor crítico, tabulado o teórico del estadístico F , está definido por:

$$F_{tab} = F_{1-\alpha}(p-1, n-p)$$

Como regla de decisión:

- Si el valor de obtenido de F excede al valor crítico $F_{1-\alpha}(p-1, n-p)$ o sea, $F > F_{1-\alpha}(p-1, n-p)$, se rechaza la hipótesis nula. Por tanto, la función de regresión estimada es de buen ajuste y el valor de F es altamente significativo para el nivel de significación fijado. Se acepta que la variable regresora tiempo, X , en semana, explica el comportamiento de la

capacidad industrial del centro de recepción, Y . Se admite la función seleccionada como buena estimativa de $E_{(Y_i/X_i)}$ y el modelo es de buen ajuste.

- Si el valor de obtenido de F no excede al valor crítico $F_{1-\alpha}(p-1, n-p)$ o sea, $F \leq F_{1-\alpha}(p-1, n-p)$, no se rechaza la hipótesis nula. Por tanto, la función de regresión estimada no es de buen ajuste.

Alternativamente, si el valor p o p -valor de F calculado es pequeño; es decir, cuando la Prob (F -statistic) o valor p es menor que la probabilidad de rechazar la hipótesis nula siendo cierta o nivel de significación fijado, se rechaza la hipótesis nula.

Resumen de la prueba de hipótesis.

Formulación de las hipótesis nula y alternativa:

$$H_0: \beta_2 = \beta_3 = 0$$

$$H_1: \text{Existe algún } \beta_j \neq 0, \text{ para } j = 2, 3$$

Región crítica para el estadígrafo F , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0: F > F_{1-\alpha}(p-1, n-p) \right\}$$

donde:

$$F = \frac{S_E^2}{S_R^2} \quad \text{o} \quad F = \frac{\frac{SC_E}{p-1}}{\frac{SC_R}{n-p}} = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R}$$

Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha}(p-1, n-p)$.
- No rechazar H_0 , si $F \leq F_{1-\alpha}(p-1, n-p)$

Alternativamente:

- Rechazar H_0 , si la Prob (F -statistic) o valor $p > \alpha$.
- No rechazar H_0 , si la Prob (F -statistic) o valor $p > \alpha$.

La decisión se puede tomar indistintamente, considerando el valor estadígrafo F , o la probabilidad a partir del valor p .

Si de aplicar la prueba de significación a la función de regresión cuadrática estimada de los centros de recepción, se obtiene como resultado rechazar la hipótesis nula, con un nivel de significación de 0,05, entonces se infiere que la función cuadrática estimada es de buen ajuste.

Para verificar la validez de la prueba estadística de F de Fisher, así como la eficiencia de la estimación obtenida a partir de la función de regresión cuadrática, se procede a verificar los supuestos básicos del modelo, mediante la comprobación de los procedimientos de varias pruebas para detectar la presencia de no autocorrelación, homocedasticidad y normalidad. De ser necesario se realizan las correcciones necesarias con el objetivo de lograr su cumplimiento y de esta forma seleccionar la mejor función de regresión que sirva de base para el pronóstico.

Coefficiente de determinación múltiple

A partir de la descomposición de la suma de cuadrados, $SC_T = SC_E + SC_R$, y

dividiendo cada término entre la SC_T se obtiene: $1 = \frac{SC_E}{SC_T} + \frac{SC_R}{SC_T}$

La relación $\frac{SC_R}{SC_T}$ expresa qué parte de la variabilidad total se debe al error o residuos, y la relación $\frac{SC_E}{SC_T}$ qué parte de la variabilidad total queda explicada por la ecuación de regresión. Esta última relación se denomina *coeficiente de determinación* y se denota por R^2 , o sea, el coeficiente de determinación es el cociente de la suma de cuadrados explicado por la regresión entre la suma de cuadrados total y se cumple que:

$$R^2 = \frac{SC_E}{SC_T} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SC_R}{SC_T}$$

Las fórmulas abreviadas para la determinación de las sumas de cuadrados

son, $SC_E = \underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2$, $SC_T = \underline{Y}^T \underline{Y} - n\bar{Y}^2$ y $SC_R = \underline{Y}^T \underline{Y} - \underline{\beta}^T \underline{X}^T \underline{Y}$

donde se deduce que:

$$R^2 = \frac{\underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2}{\underline{Y}^T \underline{Y} - n\bar{Y}^2}$$

El coeficiente de determinación toma valores entre cero y uno. Cuando R^2 es igual a uno, la SC_E coincide con el valor numérico de la SC_T ($SC_E = SC_T$), por tanto, la SC_R toma valor cero ($SC_R = 0$), en este caso el ajuste es perfecto. Cuando R^2 es igual a cero la SC_E también es cero. En la medida que se acerque a uno el ajuste será mejor, ya que se disminuyen los errores de la estimación, ofreciendo una buena medida de la bondad del ajuste realizado.

El coeficiente de determinación múltiple mide la bondad del ajuste de la función de regresión. Para modelos con más de una variable explicativa, expresa la proporción de la variación en Y , que es explicada conjuntamente por las variables explicativas, es decir, expresa qué porcentaje de la variabilidad total de los valores de la variable dependiente Y , en la muestra, es explicada o controlada por la regresión.

Coefficiente de determinación ajustado

Es un coeficiente de determinación alternativo que se define y se denota por:

$$\bar{R}^2 = 1 - \frac{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p}}{\frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{n-1}} = 1 - \frac{\frac{SC_R}{n-p}}{\frac{SC_T}{n-1}}$$

El término ajustado significa ajustado por los grados de libertad asociados con las sumas de los cuadrados de los residuos y total.

1.6.2. Pruebas de los supuestos del modelo de regresión

Pruebas para detectar la presencia de no autocorrelación, homocedasticidad y normalidad, y supuesto no multicolinealidad:

- I. Pruebas para detectar la presencia de no autocorrelación: Prueba de Durbin-Watson y Prueba Breusch-Godfrey.
- II. Pruebas para detectar la presencia de homocedasticidad: Prueba de Goldfed-Quandt y Prueba de White.
- III. Pruebas para detectar la presencia de normalidad: Prueba de Jarque-Bera y Prueba Shapiro-Wilks.
- IV. Supuesto de no multicolinealidad.

A continuación, se hace referencia de forma sucinta a las pruebas.

Pruebas para detectar la presencia de no autocorrelación: Prueba de Durbin-Watson y Prueba Breusch-Godfrey

Prueba de Durbin-Watson

Se utiliza para detectar si existe autocorrelación de primer orden, es decir, entre cada perturbación o error aleatorio U_i y la que precede. La no autocorrelación existe cuando hay independencia entre las perturbaciones o errores, (Espallargas & Solís, 2012); es decir, la $Cov(U_i, U_j) = E(U_i, U_j) = 0$, para todos los i, j con $i \neq j$.

El planteamiento de las pruebas de hipótesis se realiza teniendo en cuenta el tipo de autocorrelación, y sobre el parámetro ρ ; contrastando las siguientes hipótesis:

Caso I:

$$H_0 : \rho_{U_i, U_{-1}} \leq 0 \Leftrightarrow Cov(U_i, U_{-1}) \leq 0 \text{ No hay autocorrelación positiva.}$$

$$H_0 : \rho_{U_i, U_{-1}} > 0 \Leftrightarrow Cov(U_i, U_{-1}) > 0 \text{ Hay autocorrelación Positiva.}$$

Caso II:

$H_0 : \rho_{U_i, U_{-1}} \geq 0 \Leftrightarrow \text{Cov}(U_i, U_{-1}) \geq 0$ No hay autocorrelación negativa.

$H_0 : \rho_{U_i, U_{-1}} < 0 \Leftrightarrow \text{Cov}(U_i, U_{-1}) < 0$ Hay autocorrelación negativa.

Caso III:

$H_0 : \rho_{U_i, U_{-1}} = 0 \Leftrightarrow \text{Cov}(U_i, U_{-1}) = 0$ No hay autocorrelación.

$H_0 : \rho_{U_i, U_{-1}} \neq 0 \Leftrightarrow \text{Cov}(U_i, U_{-1}) \neq 0$ Hay autocorrelación.

La última situación corresponde a una prueba bilateral o de dos colas, por lo que el nivel de significación se divide por dos.

El estadístico de prueba d de Durbin-Watson, se calcula como sigue:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

donde:

$e_i = \hat{Y}_i - \hat{Y}$ es el error de estimación o residuos.

Y_i es el valor de cada observación.

\hat{Y}_i es el valor estimado, obtenido a partir de la función de regresión estimada.

El estadístico de prueba d es el cociente de la suma de las diferencias cuadradas de los residuos sucesivos entre la suma de cuadrados de los residuos. Como el coeficiente muestral de autocorrelación de primer orden es un estimador de ρ , el estadístico de prueba d se puede expresar como $d \approx 2(1 - \rho)$. Puesto que $-1 \leq \rho \leq 1$, entonces $0 \leq d \leq 4$, y por tanto, los valores de d se encuentran entre 0 y 4. De tomar un valor cercano a 0, indica que hay autocorrelación positiva de primer orden, y cuando toma valores cercanos a 4, indica que existe autocorrelación negativa de primer orden (Gujarati, 2003).

Los valores críticos, teóricos o tabulados se encuentran representados por cotas superiores e inferiores, lo que está dado por los valores críticos d_u (d_{Upper} valor máximo) y d_L (d_{Lower} valor mínimo), respectivamente, tabulados en dependencia del número de observaciones (n), del número de variables explicatorias (k') del modelo y en función del nivel de significación.

A continuación, se presenta una representación esquemática para la toma de decisión de la prueba de Durbin-Watson.

↓	Rechazar H_0 hay autocorrelación positiva	↓	Zona de indeterminación o de duda	↓	No rechazar H_0	↓	Zona de indeterminación o de duda	↓	Rechazar H_0 hay autocorrelación negativa	↓
0		d_L		d_u	2	$4-d_u$		$4-d_L$		4

La regla de decisión se realiza en función del estadístico de prueba Durbin-Watson y de sus valores tabulados, como se muestra a continuación según Gujarati (2003):

Tabla 1.2. Resumen de la toma de decisión para la prueba de Durbin-Watson.

Hipótesis Nula	Decisión	Si
No existe autocorrelación positiva No existe autocorrelación positiva No existe autocorrelación positiva	No se rechaza Se rechaza No hay decisión	$d > d_u$ $0 < d < d_L$ $d_L \leq d \leq d_u$
No existe autocorrelación negativa No existe autocorrelación negativa No existe autocorrelación negativa	No se rechaza Se rechaza No hay decisión	$d < 4 - d_u$ $4 - d_L < d < 4$ $4 - d_u \leq d \leq 4 - d_L$
No existe autocorrelación positiva o negativa No existe autocorrelación positiva o negativa	No se rechaza Se rechaza	$d_u < d < 4 - d_u$ $0 < d < d_L$ $4 - d_L < d < 4$

Fuente: Gujarati (2003).

Cuando se toma la decisión de no rechazar la hipótesis nula, significa que no se detecta autocorrelación de primer orden en los residuos.

La prueba Durbin-Watson presenta las siguientes limitaciones: tiene una zona o región de indeterminación, no se puede aplicar cuando se tienen funciones en las que aparezcan valores retardados de Y , y la prueba solo se puede

aplicar a esquemas de autocorrelación de primer orden.

Prueba general de autocorrelación de Breusch-Godfrey

Para evitar algunos elementos de la rigidez presentes en la prueba Durbin-Watson de autocorrelación, se puede aplicar la prueba Breusch-Godfrey para la autocorrelación de orden superior, la que se basa en el principio del Multiplicador de Lagrange.

La hipótesis nula que se somete a prueba es $H_0: \rho_1 = \rho_2 = \dots = \rho_p = 0$.

Si el tamaño de la muestra es grande, entonces el valor obtenido $(n-p)R^2$ sigue una distribución ji-cuadrado $\chi^2_{1-\alpha, p}$, donde $(n-p)R^2 \sim \chi^2_{1-\alpha, p}$, con p grados de libertad. Si el valor obtenido $(n-p)R^2$ es superior al valor crítico ji-cuadrado para el nivel de significación seleccionado, se puede rechazar la hipótesis nula, lo que implica que al menos existe un ρ significativamente diferente de cero.

Pruebas para detectar la presencia de homocedasticidad: Prueba de Goldfeld-Quandt y Prueba general de heterocedasticidad de White.

Las varianzas del error aleatorio o perturbación deben ser constantes e iguales para todas las observaciones, para que en ellas exista homocedasticidad, donde las varianzas de las perturbaciones no dependen de la observación correspondiente.

Prueba de Goldfeld-Quandt

Para la aplicación de la prueba, se supone que la varianza heterocedástica σ_i^2 está relacionada positivamente con una de las variables explicatorias en el modelo de regresión (Gujarati, 2003), es decir, se aplica cuando se supone que alguna variable provoca heterocedasticidad y se conoce la forma de la relación. Por ejemplo: $V(U_i) = \sigma^2 X_i^3$, donde la varianza crece cuando X_i crece en esa proporción (Espallargas & Solís, 2012).

El estadístico de prueba utilizado para verificar el supuesto de homocedasticidad es el desarrollado por Goldfeld y Quandt en su método paramétrico. Para conocer los pasos a seguir para determinar el estadístico de prueba λ , el lector puede consultar los procedimientos expuestos por Gujarati (2003); y Espallargas & Solís (2012). No obstante, a continuación, se presenta un resumen.

Pasos para realizar prueba de Goldfeld-Quandt:

- Se ordenan los datos de mayor a menor tomando como referencia la variable explicativa X .
- Se determinan los valores centrales c , normalmente deben de cumplir la condición $c < n/3$.
- Se dividen las observaciones en dos muestras: la primera recoge las observaciones correspondiente a los valores mayores de X , y la segunda recoge las observaciones correspondiente a los valores menores de X .
- Se realiza la estimación de cada una de las muestras por separado.
- Se determina la suma de cuadrados residuales.
- Se realiza la prueba de igualdad de varianzas.

La prueba parte del crecimiento o decrecimiento de las varianzas de las perturbaciones.

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$$

$$H_1: \sigma_1^2 > \sigma_2^2 > \dots > \sigma_n^2 \quad \text{o} \quad H_1: \sigma_1^2 < \sigma_2^2 < \dots < \sigma_n^2, \text{ o ambas.}$$

También, para verificar la hipótesis de homocedasticidad, se utiliza una prueba que permita determinar si las varianzas de los dos grupos de muestras son iguales, es decir:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

$$\text{El estadístico de prueba es: } \lambda = F = \frac{\frac{SC_{R2}}{(n-c-2p)/2}}{\frac{SC_{R1}}{(n-c-2p)/2}}.$$

donde: SC_{R2} es la suma de cuadrados de los residuos de la regresión de las primeras observaciones correspondiente a los valores mayores de X ; SC_{R1} es la suma de cuadrados de los residuos de la regresión de las últimas observaciones correspondiente a los valores menores de X ; n es el número total de observaciones; c es el número de observaciones centrales; p es el número de parámetros de la función.

Este estadístico se distribuye según una distribución F de Fisher, con

$(n-c-2p)/2$ grados de libertad para el numerador y con $(n-c-2p)/2$ grados de libertad para el denominador, siendo el valor crítico o teórico $F_{1-\alpha}[(n-c-2p)/2, (n-c-2p)/2]$.

Regla de decisión: si el estadístico de prueba obtenido $\lambda = F$ es superior al valor crítico, o sea, $\lambda > F_{1-\alpha}[(n-c-2p)/2, (n-c-2p)/2]$, se rechaza la hipótesis nula $H_0: \sigma_1^2 = \sigma_2^2$. Por tanto, existe heterocedasticidad y no se cumple con el supuesto de homocedasticidad.

Prueba general de heterocedasticidad de White

La prueba general de heterocedasticidad de White no requiere del ordenamiento de las observaciones con respecto a la variable X , que supuestamente provoca la heterocedasticidad como en la prueba de Goldfeld-Quandt. Mediante la prueba de White se hace más factible llevar a cabo la detección de la heterocedasticidad y no se apoya en el supuesto de normalidad.

Para realizar la prueba de White, inicialmente se estima la función de regresión lineal simple o múltiple y se obtienen los residuos e_i , mediante el Método de los Mínimos Cuadrados Ordinarios, y se comprueba que la función obtenida sea de buen ajuste y buena predictora. Teniendo en cuenta estos resultados se realiza la prueba general de heterocedasticidad de White, obteniéndose la ecuación de los errores, dada por la regresión auxiliar asociada a la prueba.

En la formulación de la hipótesis, como hipótesis nula que se somete a prueba, se plantea el cumplimiento del supuesto de homocedasticidad, es decir la no existencia de heterocedasticidad, y como hipótesis alternativa a contrastar se plantea la existencia de heterocedasticidad.

El estadístico de prueba es el producto del tamaño de la muestra, n , por el coeficiente de determinación de la regresión auxiliar, R^2 , o sea, $n \cdot R^2$ obtenido de la regresión auxiliar, que sigue una distribución ji-cuadrado con grados de libertad igual al número de regresoras en la regresión auxiliar, sin incluir el término constante, es decir, $n \cdot R^2 \sim \chi_{(1-\alpha)gl}^2 = \chi_{(1-\alpha)k}^2$. Siendo k el número de parámetros de la ecuación de los errores menos la constante, o sea, los grados de libertad.

Si el valor calculado u obtenido del estadístico de prueba $n \cdot R^2$ es superior que el valor ji-cuadrado crítico o teórico, al nivel de significación considerado, se

rechaza la hipótesis nula, por lo tanto, hay heterocedasticidad y no se cumple con el supuesto de homocedasticidad.

Si el valor calculado u obtenido del estadístico de prueba $n.R^2$ es menor que el valor ji-cuadrado crítico, al nivel de significación fijado, no se rechaza la hipótesis nula y se cumple con el supuesto de homocedasticidad, lo que implica que en la regresión auxiliar $\alpha_2 = \alpha_3 = \alpha_4 \dots$, son iguales a cero.

Considerando la probabilidad, si valor p de $n.R^2$ es mayor que el nivel de significación α , no se rechaza la hipótesis nula y se cumple con el supuesto de homocedasticidad.

Considerando el valor obtenido de F , si es menor que el valor crítico de $F_{1-\alpha}(p-1, n-p)$, entonces no se rechaza la hipótesis nula y se cumple con el supuesto de homocedasticidad.

Pruebas para detectar la presencia de Normalidad

La normalidad se puede probar empleando los estadísticos Ji-Cuadrado, Kolmogorov-Smirnov, Shapiro-Wilks y Jarque-Bera.

Prueba de Jarque-Bera

Esta prueba utiliza el estadístico Jarque-Bera (*JB*) para probar la normalidad de un conjunto de datos. Los datos pueden ser los relacionados con la variable dependiente (regresada o respuesta) o con los residuos o errores. Por tanto, con el estadístico de prueba Jarque-Bera, se puede verificar el supuesto de normalidad en el análisis de regresión, es decir, si la variable o los residuos siguen una distribución normal o no (Espallarga & Solís, 2012). Los fundamentos teóricos y metodológicos que se tienen en cuenta para el desarrollo de esta prueba se basan en los criterios de Gujarati (2003); y de Espallarga & Solís (2012).

La prueba de Jarque-Bera calcula la asimetría y la curtosis y aparece en el paquete de programa estadístico y econométrico EViews.

Prueba de Jarque-Bera para verificar el supuesto de normalidad de la variable dependiente (regresada o respuesta) en el análisis de regresión.

El estadístico de Prueba a utilizar es:

$$JB = n \left[\frac{A^2}{6} + \frac{(K-3)^2}{24} \right]$$

donde:

n es el número de observaciones.

A es la simetría.

K es la curtosis.

También se puede utilizar la expresión siguiente para el estadístico de prueba:

$$JB = \frac{n-k}{6} \left[A^2 + \frac{1}{4}(K-3)^2 \right]$$

Asintóticamente, el estadístico de prueba Jarque-Bera (JB) sigue una distribución ji-cuadrado, con dos grados de libertad (gl).

Formulación de las hipótesis nula y alternativa:

La hipótesis nula que se somete a prueba, expresa que la variable dependiente (regresada o respuesta) está normalmente distribuida.

$H_0: Y_i \sim N$.

$H_1: Y_i$ no sigue una distribución normal.

Región crítica para el estadístico de prueba JB , dado el nivel de significación α , es

$W = \left\{ JB \in \mathfrak{R} \text{ con } JB \geq 0 : JB > \chi_{1-\alpha}^2 (gl) \right\}$ o sea,

$W = \left\{ JB \in \mathfrak{R} \text{ con } JB \geq 0 : JB > \chi_{1-\alpha}^2 (2) \right\}$

Regla de decisión:

- Rechazar H_0 , si $JB > \chi_{1-\alpha}^2 (2)$.

- No rechazar H_0 , si $JB \leq \chi_{1-\alpha}^2 (2)$.

Alternativamente:

- Rechazar H_0 , si el valor de la probabilidad o valor $p < \alpha$.
- No rechazar H_0 , si el valor de la probabilidad o valor $p > \alpha$.

Si el valor probabilístico del estadístico de prueba JB , valor p , es muy bajo en una aplicación, lo cual sucederá si el valor estadístico de prueba difiere en gran medida de cero, o sea, si $JB > X_{1-\alpha}^2(2)$, entonces se rechaza la hipótesis nula, de que la variable dependiente (regresada o respuesta) está normalmente distribuida. Por el contrario, si el valor probabilístico del estadístico de prueba JB , valor p , es muy alto en una aplicación, lo cual ocurrirá cuando el valor del estadístico de prueba se encuentra cercano a cero, o sea, si $JB \leq X_{1-\alpha}^2(2)$, entonces no se rechaza la suposición de normalidad, no existen elementos para rechazar la hipótesis nula, o sea, la variable dependiente sigue una distribución normal.

El valor del estadístico de prueba JB depende de la asimetría y de la curtosis.

La asimetría mide la deformación de la distribución y se calcula con las fórmulas

$$A = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^3, \text{ o } A = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3}{\hat{\sigma}^3}$$

La curtosis mide el apuntamiento de la distribución y se calcula como

$$K = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^4, \text{ o por } K = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^4}{\hat{\sigma}^4}$$

En las expresiones de A y K , $\hat{\sigma} = S$ representa el estimador sesgado de la desviación estándar poblacional, y se calcula por:

$$S = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}$$

Prueba de Jarque-Bera para verificar el supuesto de normalidad en los residuos (errores), en el análisis de regresión.

Estadístico de Prueba a utilizar es:

$$JB = n \left[\frac{A^2}{6} + \frac{(K-3)^2}{24} \right]$$

donde:

n es el número de observaciones.

A es la asimetría.

K es la curtosis.

También se puede utilizar la expresión siguiente para el estadístico de prueba:

$$JB = \frac{n-k}{6} \left[A^2 + \frac{1}{4}(K-3)^2 \right]$$

Asintóticamente, el estadístico Jarque-Bera sigue una distribución ji-cuadrado, con dos grados de libertad.

Formulación de las hipótesis nula y alternativa:

La hipótesis nula a ser probada, afirma que los residuos están normalmente distribuidos.

$$H_0: U_i \sim N.$$

$H_1: U_i$ no siguen una distribución normal.

La región crítica para el estadístico de prueba JB , dado el nivel de significación α , es

$$W = \left\{ JB \in \mathfrak{R} \text{ con } JB \geq 0 : JB > \chi_{1-\alpha}^2(2) \right\}$$

Regla de decisión:

- Rechazar H_0 , si $JB > \chi_{1-\alpha}^2(2)$.

- No rechazar H_0 , si $JB \leq \chi^2_{1-\alpha} (2)$.

Alternativamente:

- Rechazar H_0 , si el valor de la probabilidad o valor $p < \alpha$.

- No rechazar H_0 , si el valor de la probabilidad o valor $p > \alpha$.

Si el valor probabilístico del estadístico de prueba JB , valor p , es muy bajo en una aplicación, lo cual sucederá si el valor estadístico de prueba difiere en gran medida de cero, o sea, si $JB > \chi^2_{1-\alpha} (2)$, entonces se rechaza la hipótesis de que los residuos o errores de estimación están normalmente distribuidos. Por el contrario, si el valor probabilístico del estadístico de prueba JB , valor p , es muy alto en una aplicación, lo cual ocurrirá cuando el valor del estadístico de prueba se encuentra cercano a cero, o sea, si $JB \leq \chi^2_{1-\alpha} (2)$, entonces no se rechaza la suposición de normalidad, no existen elementos para rechazar la hipótesis nula, o sea, los residuos o errores de estimación siguen una distribución normal.

La asimetría y la curtosis se calculan de forma similar a lo expresado anteriormente.

Al analizar la normalidad es importante tener en cuenta las características de la asimetría y de la curtosis en la distribución.

La asimetría mide la deformación de una distribución, donde se pueden presentar las situaciones siguientes:

- Si el coeficiente de asimetría es igual a cero, la serie de observaciones es simétrica, cuyo principal patrón es la distribución normal, siendo las medidas de tendencia central media aritmética, mediana y moda, iguales
- Si el coeficiente de asimetría es mayor que cero, es decir, si es positivo, la serie de observaciones es asimétrica positiva, la distribución es deformada a la derecha, siendo la media aritmética mayor que la moda.
- Si el coeficiente de asimetría es menor que cero, es decir, si es negativo, la serie de observaciones es asimétrica negativa, la distribución es deformada a la izquierda, siendo la moda mayor que la media aritmética.

En una distribución simétrica, los datos de la serie se encuentran dispuestos simétricamente en relación con sus promedios. Además, la media aritmética,

la mediana y la moda son iguales; a medida que estas medidas se alejan, se acentúa la asimetría de la serie.

Como se aprecia en las figuras siguientes, la deformación corresponde al lado en que la distribución tiene la cola más larga.

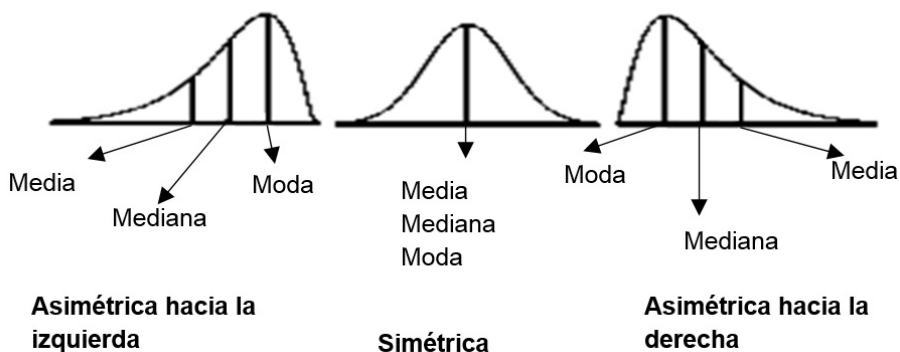


Figura 1.1. Representación de la asimetría de una distribución.

La curtosis mide el grado de apuntamiento o achatamiento de una distribución, donde se pueden presentar las situaciones siguientes:

- Si el valor de la curtosis es igual a tres, la distribución es conocida como mesocúrtica, cuya principal muestra es la distribución normal, existiendo una concentración normal de los valores en la región central de la distribución.
- Si el valor de la curtosis es mayor que tres, la distribución es más apuntada que la normal y recibe el nombre de leptocúrtica, existiendo una alta concentración de los valores en la región central de la distribución.
- Si el valor de la curtosis es menor que tres, la distribución es menos apuntada que la normal, es más achatada, y recibe el nombre de platicúrtica, existiendo una baja concentración de los valores en la región central de la distribución.

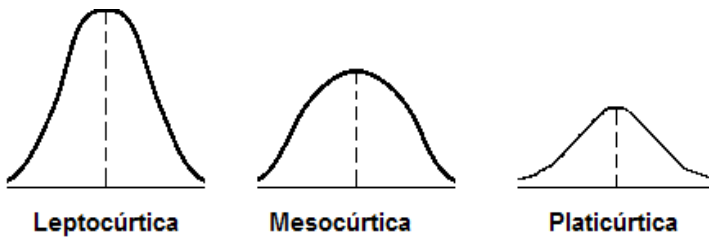


Figura 1.2. Representación de la curtosis de una distribución.

Cuando el coeficiente de asimetría se aleja de cero y la distribución es deformada, o cuando el coeficiente de curtosis difiere de tres y la distribución presenta apuntamiento, y el valor del estadístico de prueba Jarque-Bera siga siendo positivo, aumenta la probabilidad de rechazar la hipótesis nula, de que los datos provienen de una distribución normal.

Otras pruebas de normalidad

Dada la importancia que tiene la presencia de normalidad en el contexto de la evaluación económica que se presenta, se considera útil tener en cuenta otras pruebas que generalmente son utilizadas para probar la normalidad de un conjunto de datos como la de Kolmogoro-Smirnov y Shapiro-Wilks. La prueba Kolmogorov-Smirnov es apropiada, cuando el tamaño de la muestra es mayor que 50, y la prueba Shapiro-Wilk, cuando el tamaño de la muestra es como máximo 50. Las mismas se encuentran disponibles en el paquete de programa estadístico SPSS.

Las pruebas permiten probar si la variable en estudio de una muestra tiene distribución normal. Como hipótesis a contrastar se tiene que la hipótesis nula plantea que las observaciones siguen una distribución normal, y la hipótesis alternativa que las observaciones no siguen una distribución normal. Para tomar la decisión, si el valor p asociado a la prueba seleccionada, es mayor que el nivel de significación, no existen evidencias para rechazar el supuesto de normalidad de la distribución.

Supuesto de no multicolinealidad

A la multicolinealidad se hace referencia, desde el punto de vista práctico, en el epígrafe 1.6.5, cuando se estudia el modelo de regresión polinomial $Y_t = \beta_1 + \beta_2 X_t - \beta_3 X_t^2 + U_t$, como modelo que mejor ajusta al aprovechamiento de la capacidad industrial de los centros de recepción.

Si como resultado de aplicar, a la capacidad industrial disponible estimada de los centros de recepción, las pruebas de Durbin-Watson; Breusch-Godfrey; Goldfeld-Quandt; White; Jarque-Bera y Shapiro-Wilks, se comprueba que no existe autocorrelación, que existe homoscedasticidad en las varianzas del error, es decir, son perturbaciones homoscedásticas, que existe normalidad para el nivel de significación seleccionado, y que no se viola el supuesto de no multicolinealidad, entonces se infiere que las funciones de regresión polinomiales estimadas cuadráticas constituyen los modelos matemáticos que permitirán realizar las proyecciones adecuadas y eficientes de la capacidad industrial de los centros de recepción.

1.6.3. Esbozo sobre la aplicación de las derivadas para el análisis de las propiedades de funciones

Las funciones reales de una variable real cumplen algunas propiedades que no se cumplen para cualquier tipo de función, tales como: *simetría, paridad, monotonía, existencia de extremos, concavidad y existencia de puntos de inflexión*. Las propiedades se utilizan para describir el comportamiento de las funciones, para trazar sus gráficos y para describir las características fundamentales de las funciones más utilizadas en las Ciencias Administrativas y Económicas.

Para complementar el estudio sobre la determinación del aprovechamiento de las capacidades industriales, a partir de la función de regresión polinomial estimada cuadrática, en el presente epígrafe se expone un análisis conceptual de las propiedades de las funciones reales de una variable real, mediante las aplicaciones de las derivadas. En los epígrafes subsiguientes (1.6.4, 1.6.5 y 1.6.6), se realizan aplicaciones prácticas de las propiedades.

El desarrollo del epígrafe actual y los subsiguientes, se realizan considerando los aspectos teórico y práctico abordados por Arya, Lardner & Ibarra (2009); Zill & Wright (2011); y Haeussler, Paul & Wood (2014), referentes a las temáticas relacionadas con funciones de una variable, funciones racionales enteras, aplicaciones económicas de las funciones y aplicaciones de la derivada a los problemas de optimización.

A continuación, se hace referencia al análisis conceptual de las propiedades: monotonía, existencia de extremos locales, concavidad y existencia de puntos de inflexión.

Monotonía

Se dice que una función f es monótona creciente [decreciente] en un intervalo, si para cualquier par de valores x_1, x_2 de ese intervalo, tales que $x_1 < x_2$, se cumple que $f(x_1) < f(x_2)$ [$f(x_1) > f(x_2)$]. En otras palabras, la función es creciente cuando a medida que la x aumenta, la y aumenta también, y es decreciente, si cuando la x aumenta, la y disminuye.

Precisando el análisis, se muestra el siguiente teorema que ofrece una condición necesaria y suficiente para analizar la monotonía de una función en un intervalo donde dicha función es derivable.

Teorema: Una función $y = f(x)$, derivable en el intervalo $]a, b[$, es creciente [decreciente] en ese intervalo, si y solo si $f'(x) \geq 0$ [$f'(x) \leq 0$] para todo $x \in]a, b[$.

La aplicación de este teorema al análisis de la monotonía se reduce al análisis del signo de la derivada.

Para analizar la monotonía de una función se procede como sigue: se calcula la primera derivada de la función; se determinan los ceros de la primera derivada; se determinan los puntos que indefinen la primera derivada; se determina el signo de la primera derivada en los intervalos determinados por los puntos críticos (obtenidos por los ceros de la primera derivada o por los puntos que la indefinen), y por último, se interpreta el signo de la primera derivada, y se concluye si la monotonía es decreciente o creciente, para el intervalo donde se efectúa el análisis.

Existencia de extremos

Se dice que una función f alcanza su valor máximo [mínimo] en el punto x_0 si para todo $x \in \text{Dom } f$ se cumple que $f(x_0) \geq f(x)$ [$f(x_0) \leq f(x)$].

A los valores máximo y mínimo de una función se les llama extremos absolutos o extremos globales de la función.

Cuando la condición de ser el mayor o menor valor de la función se cumple solamente en una vecindad del punto x_0 , entonces se dice que el extremo es local y no global. En correspondencia, se puede decir que un extremo global o absoluto, es también un extremo local; sin embargo, lo contrario, por lo

general, no se cumple.

A continuación, se define formalmente el concepto extremo local.

Una función f tiene un *máximo [mínimo] local* en el punto x_0 si existe una vecindad de x_0 , tal que, $f(x_0) \geq f(x)$ [$f(x_0) \leq f(x)$] para todo x de dicha vecindad. El número $y_0 = f(x_0)$ es un *valor máximo [mínimo] local* de la función y el punto (x_0, y_0) es un *máximo [mínimo] local* del gráfico de f .

Antes de dar la condición necesaria y la condición suficiente para la existencia de extremos locales, se presentan algunos detalles sobre estos extremos:

- En cada punto máximo local, la curva presenta una cumbre.
- En cada punto mínimo local, la curva presenta una hondonada.
- Al valor máximo (mínimo) local se le llama *extremo local*, porque es el mayor (el menor) valor de la función en una vecindad de un punto. Eso no tiene por qué ocurrir en otra parte del dominio de la función. Incluso, un máximo local puede ser menor que un mínimo local.
- En los puntos donde la curva tiene un extremo local, si existe la tangente, es horizontal. Por tanto, su pendiente es cero. Esto lo confirma el siguiente teorema.

Teorema: Si la función $y = f(x)$ es derivable en un punto x_0 donde tiene un extremo local, entonces $f'(x_0) = 0$.

Este teorema da una condición necesaria para la existencia de extremos.

Eso significa que una función no puede tener extremos en puntos donde $f'(x) \neq 0$.

La condición dada por el teorema anterior no llega a ser suficiente, porque hay funciones que no tienen extremos en puntos donde $f'(x) = 0$.

Hay funciones que tienen extremos en puntos donde no son derivables.

De acuerdo con el teorema anterior y el análisis hecho sobre el mismo, los únicos puntos donde puede haber extremos son aquellos donde $f'(x) = 0$, o donde $f'(x)$ no está definida. Estos puntos se denominan puntos críticos de primera especie. Ahora bien, ¿cómo se analiza si en un punto crítico hay un extremo? En el caso

de que haya extremo, ¿cómo se sabe si es un máximo o un mínimo?

En cada punto donde una función tiene un extremo, se produce un cambio en la monotonía de la misma. Relacionando esto con el signo de la derivada se obtiene la siguiente condición necesaria y suficiente para la existencia de extremos.

Teorema: Sea f una función continua en x_0 y derivable en una vecindad reducida de x_0 , entonces f tiene un extremo en x_0 , si y solo si $f'(x)$ cambia de signo cuando x pasa por x_0 de izquierda a derecha. El extremo es:

- máximo, si $f'(x)$ es positiva para $x < x_0$ y negativa para $x > x_0$; y
- mínimo, si $f'(x)$ es negativa para $x < x_0$ y positiva para $x > x_0$.

Los pasos lógicos para analizar la existencia de extremos, además del procedimiento abordado para determinar la monotonía de una función, expuestos anteriormente, incluyen el análisis de la existencia de extremos locales, a partir del teorema anterior.

Concavidad y punto de inflexión

Para trazar exitosamente el gráfico de una función, es importante tener en cuenta hacia dónde está dirigida la concavidad (hacia arriba o hacia abajo) y en qué puntos se produce el cambio de la dirección de la concavidad.

Se dice que una función f es *cóncava hacia arriba* [*hacia abajo*] en un intervalo, si para cualquier punto de ese intervalo el gráfico queda por encima [por debajo] de la tangente. El punto del gráfico donde cambia la dirección de la concavidad se llama *punto de inflexión*.

En otras palabras, la función es cóncava hacia arriba en el intervalo donde la curva *abre hacia arriba*, y cóncava hacia abajo en el intervalo donde la curva *abre hacia abajo*.

Precisando el análisis, se presenta el siguiente teorema: Sea f una función dos veces derivable en el intervalo $[a, b]$. Entonces el gráfico de f tiene la concavidad dirigida hacia arriba [hacia abajo] en $]a, b[$, si y solo si, para todo $x \in]a, b[$ se cumple que $f''(x) \geq 0$ [$f''(x) \leq 0$].

Según este teorema, en los puntos donde $f''(x) \neq 0$ no puede haber puntos de inflexión. Por tanto, solo puede haber en aquellos en los que $f''(x) = 0$ o

$f''(x)$ no está definida, denominados *puntos críticos de segunda especie*.

Teorema: Si la función f es continua en el punto crítico de segunda especie x_0 y dos veces derivable en una vecindad reducida del mismo, entonces $(x_0, f(x_0))$ es un punto de inflexión del gráfico de f , si $f''(x)$ cambia de signo en x_0 .

La aplicación de los dos teoremas anteriores permite analizar la concavidad y la existencia de puntos de inflexión, usando la segunda derivada, de la misma forma que se analiza la monotonía y extremos locales con la primera derivada.

Para analizar la concavidad y la existencia de puntos de inflexión del gráfico de una función se procede como sigue: se calcula la segunda derivada de la función; se determinan los ceros de la segunda derivada; se determinan los puntos que indefinen la segunda derivada; se determina el signo de la segunda derivada en los intervalos determinados por los puntos críticos; y por último se interpreta el signo de la segunda derivada, y se concluye si la concavidad se encuentra dirigida hacia arriba o hacia abajo para el intervalo donde se efectúa el análisis. Además, se dan las conclusiones sobre la existencia de puntos de inflexión.

1.6.4. Determinación del aprovechamiento máximo de la capacidad industrial

Uno de los problemas relevantes en los procesos económicos radica en cómo lograr la eficiencia. Para ello, es necesario desde la perspectiva de la matemática, la minimización o maximización de determinados indicadores económicos. Esos problemas en los que es necesario hallar el valor máximo o el valor mínimo de una función se denominan problemas de optimización. Los valores máximos o mínimos obtenidos como solución de un problema de optimización, que coinciden con los extremos absolutos de la función, se denominan valores óptimos.

En correspondencia con lo expresado, una vez determinada la función matemática que mejor ajusta a la serie de datos, representada por la función de regresión polinomial estimada cuadrática, $\hat{E}_{(Y_t|X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t - \hat{\beta}_3 X_t^2$ siendo \hat{Y}_t la capacidad industrial estimada del centro de recepción (en toneladas de caña) y X_t el tiempo (en semanas), se procede a determinar

la semana de máximo u óptimo aprovechamiento de las capacidades industriales, mediante el empleo de los métodos clásicos de optimización. Estos métodos son aquellos en los que se utiliza la derivada para la resolver problemas de optimización. Uno de ellos fue descrito en el epígrafe anterior.

Como en este caso se trata de una función cuadrática, la cual es derivable en todo su dominio y cuya segunda derivada es constante, se puede proceder en la forma siguiente:

- Se calcula la primera deriva la función estimada polinomial de grado dos.
- Se iguala la primera derivada a cero y se resuelve la ecuación así obtenida para hallar el punto crítico X_0 , correspondiente a la posible semana de máximo u óptimo aprovechamiento de las capacidades industriales de los centros de recepción.
- Se calcula la segunda derivada de la función estimada ($\frac{d^2\hat{Y}}{dX^2} = \hat{Y}''$). Si la segunda derivada de la función estimada es menor que cero, para cualquier valor de la variable explicativa X en el dominio de la función, se puede afirmar que para el valor X_0 la función alcanza su valor máximo u óptimo.
- Se evalúa la función de regresión polinomial estimadas cuadráticas para el valor crítico X_0 y se obtiene el valor máximo u óptimo de la función, o sea, $\hat{Y}_{\max} = \hat{\beta}_1 + \hat{\beta}_2 X_0 - \hat{\beta}_3 X_0^2$, que representa la capacidad industrial máxima estimada del centro de recepción.

A partir de \hat{Y}_{\max} , correspondiente a la semana X_0 , se determina el aprovechamiento máximo u óptimo de las capacidades productivas de los centros de recepción, que representa la producción óptima por unidad de tiempo que puede alcanzar el centro de recepción, si se utiliza eficiente y racionalmente, sin pérdida de tiempo y sin sobre explotación.

Las capacidades disponibles calculadas de los centros de recepción representan los términos independientes de *las restricciones de tipo II* de los modelos planteados, es decir, los D_j o D_{jq} .

El procedimiento expuesto para determinar la capacidad máxima estimada del centro de recepción, se puede generalizar de la siguiente forma:

- Calculando la primera derivada de la función general de regresión polinomial estimada $\hat{E}_{(Y/X)} = \hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X + \hat{\beta}_3 X^2$, se obtiene que

$$\frac{d\hat{Y}}{dX} = \hat{\beta}_2 + 2\hat{\beta}_3 X$$

- Igualando el resultado de la primera derivada a cero $\frac{d\hat{Y}}{dX} = 0$, se forma la ecuación $\hat{\beta}_2 + 2\hat{\beta}_3 X = 0$.

- Despejando a X en la ecuación anterior, se obtiene $X = -\frac{\hat{\beta}_2}{2\hat{\beta}_3}$.

Este valor de X es la abscisa del vértice de la parábola que representa gráficamente a la función de regresión polinomial estimada y será

denotado por X_0 , o sea, $X_0 = -\frac{\hat{\beta}_2}{2\hat{\beta}_3}$.

- Calculando la segunda derivada de esta función se obtiene, $\frac{d^2\hat{Y}}{dX^2} = 2\hat{\beta}_3$.

- Si $\hat{\beta}_3 > 0$, la segunda derivada es positiva, el gráfico de la función tiene la concavidad dirigida hacia arriba y la función tiene su valor mínimo en X_0 .

- Si $\hat{\beta}_3 < 0$, la segunda derivada es negativa, el gráfico de la función tiene la concavidad dirigida hacia abajo y la función tiene su valor máximo en X_0 .

- Sustituyendo el valor numérico de X_0 en \hat{Y} , y realizando los cálculos correspondientes, se obtiene el valor máximo o mínimo de la función de regresión polinomial estimada.

- El valor máximo o mínimo también puede calcularse en función de los estimadores mínimos cuadráticos, sustituyendo la expresión de X_0 en \hat{Y} , y realizando las operaciones algebraicas correspondientes, se obtiene la expresión que permite calcular el valor máximo o mínimo de la función de regresión estimada, o sea, $\hat{Y}_{\max \text{ o } \min} = \hat{\beta}_1 - \frac{\hat{\beta}_2^2}{4\hat{\beta}_3}$. Con solo aplicar las

expresiones de X_0 y $\hat{Y}_{\max \text{ o } \min}$ se obtiene el resultado deseado.

Desarrollando el procedimiento expuesto:

$$\text{Sustituyendo } X = X_0 = -\frac{\hat{\beta}_2}{2\hat{\beta}_3} \text{ en } \hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X + \hat{\beta}_3 X^2$$

$$\begin{aligned} \hat{Y}_{\max \text{ o min}} = \hat{Y}_0 &= \hat{\beta}_1 + \hat{\beta}_2 X_0 + \hat{\beta}_3 X_0^2 \\ &= \hat{\beta}_1 + \hat{\beta}_2 \left(-\frac{\hat{\beta}_2}{2\hat{\beta}_3} \right) + \hat{\beta}_3 \left(-\frac{\hat{\beta}_2}{2\hat{\beta}_3} \right)^2 \\ &= \hat{\beta}_1 - \frac{\hat{\beta}_2^2}{2\hat{\beta}_3} + \hat{\beta}_3 \left(\frac{\hat{\beta}_2^2}{4\hat{\beta}_3^2} \right) = \hat{\beta}_1 - \frac{2\hat{\beta}_2^2}{4\hat{\beta}_3} + \frac{\hat{\beta}_2^2}{4\hat{\beta}_3} = \hat{\beta}_1 - \frac{\hat{\beta}_2^2}{4\hat{\beta}_3} \end{aligned}$$

$Y = \hat{Y}_0$ es el valor extremo, máximo o el mínimo, de la función de regresión polinomial estimada y, geoméricamente, es la ordenada del vértice de la parábola que representa gráficamente a dicha función.

Se ha demostrado que tanto el punto X_0 , donde la función alcanza su valor extremo, como el valor extremo de la función Y_0 pueden calcularse en función de los estimadores mínimos cuadráticos $\hat{\beta}_1$, $\hat{\beta}_2$ y $\hat{\beta}_3$.

En resumen, la función de regresión estimada $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X + \hat{\beta}_3 X^2$ alcanza el valor extremo $\hat{Y}_0 = \hat{\beta}_1 - \frac{\hat{\beta}_2^2}{4\hat{\beta}_3}$ en el punto $X_0 = -\frac{\hat{\beta}_2}{2\hat{\beta}_3}$. El extremo es máximo, si $\hat{\beta}_3 < 0$, y es mínimo, si $\hat{\beta}_3 > 0$.

El procedimiento expuesto para la determinación del aprovechamiento óptimo de la capacidad industrial se fundamenta teóricamente en el estudio de las propiedades de las funciones a partir de la aplicación de las derivadas expuestas en el epígrafe 1.6.3

1.6.5. Estudio de un caso real sobre la utilización de la capacidad industrial y validación de resultados

Para el estudio de la capacidad industrial de un centro de recepción se tiene en cuenta la:

- I. Estimativa de la capacidad industrial.
- II. Determinación del aprovechamiento máximo u óptimo de la capacidad industrial.
- III. Inferencia realizada al modelo de regresión lineal múltiple asociado a la capacidad industrial.

I- Estimativa de la capacidad industrial

Para mostrar la determinación de la capacidad estimada de los centros de recepción, se presenta una situación práctica real. Como paso inicial se calcula el promedio de la caña procesada semanalmente durante las últimas cinco zafras azucareras, lo que representa la capacidad industrial real promedio semanal, tabla 1.3.

Tabla 1.3. Capacidad industrial real promedio de caña procesada semanal (toneladas).

Semana	Capacidad industrial real promedio semanal (t)	Semana	Capacidad industrial real promedio semanal (t)
1	2729,43	9	4764,75
2	2804,27	10	5172,02
3	3458,96	11	4352,56
4	3069,45	12	3646,38
5	4205,78	13	2401,87
6	5207,66	14	789,366
7	5221,93	15	873,937
8	5165,48		

Fuente: elaborada por los autores.

Se propone investigar si la variable dependiente capacidad industrial real promedio semanal, se encuentra relacionada con la variable independiente, tiempo (en semanas), donde se considera una muestra de quince años, como se exhibe en la tabla anterior.

La información expuesta se representa gráficamente por la nube de puntos que se muestra en el siguiente diagrama de dispersión, formado por 15 pares ordenados (x_i, y_i) , donde i toma valores desde 1 hasta 15.

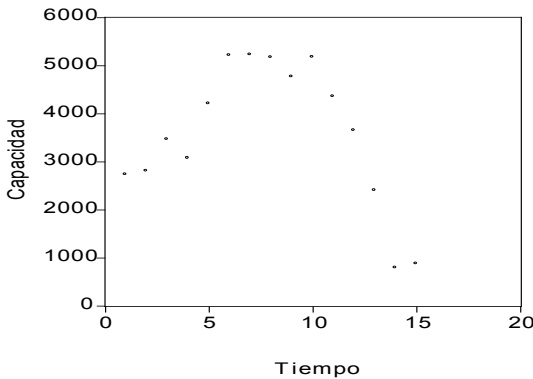


Figura 1.3. Diagrama de dispersión representado por los puntos (x_i, y_i) (EViews).

El diagrama de dispersión da una idea de la forma gráfica de la función, el mismo se obtiene representando en un sistema de coordenadas todos los puntos (x_i, y_i) .

Teniendo en cuenta la representación de los pares ordenados, se puede representar una función racional entera o polinómica de grado dos, función cuadrática, que se puede expresar en la forma:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2}, \text{ con } n = 2 \text{ y } a_n < 0, \text{ o sea, } f(x) = ax^2 + bx + c, \text{ donde } a, b \text{ y } c \text{ son constantes y } a < 0.$$

Suponiendo que el comportamiento de la capacidad industrial promedio se encuentra en relación estrecha con el tiempo (en semanas), entonces, con el propósito de hacer estimaciones, se determina que la función $f(x) = -a x^2 + bx + c$, es la que mejor expresa esta relación.

Conocida la relación entre la capacidad industrial promedio con el tiempo en semanas, se investiga qué modelo matemático, ecuación matemática o función matemática expresa adecuadamente dicha relación. El modelo, ecuación o función buscada será una función de regresión representada por

$$\hat{E}_{(Y_t|X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t - \hat{\beta}_3 X_t^2.$$

Esta función está definida para todos los números reales, en particular para la situación objeto de estudio, está definida para los números reales positivos.

Su gráfico es una parábola con las siguientes propiedades: el vértice es el punto para el cual $X = -\frac{\hat{\beta}_2}{2\hat{\beta}_3}$, como el estimador mínimo cuadrático o

coeficiente de regresión $\hat{\beta}_3 < 0$, el gráfico de la función tiene la concavidad dirigida hacia abajo y el vértice tiene su valor máximo, donde la abscisa es $X = -\frac{\hat{\beta}_2}{2\hat{\beta}_3}$.

Estimadores mínimos cuadráticos de la función de regresión. Determinación de $\hat{\beta}_1$, $\hat{\beta}_2$ y $\hat{\beta}_3$ por la vía matricial.

Dado el modelo de regresión lineal $Y = \beta_1 + \beta_2 X - \beta_3 X^2 + U_i$, se tendrá para cada par ordenado (x_i, y_i) , donde $i = 1, 2, 3, \dots, 15$ un conjunto de ecuaciones simultáneas, representadas por el siguiente sistema de 15 ecuaciones:

$$Y_1 = \beta_1 + \beta_2 X_1 - \beta_3 X_1^2 + U_1$$

$$Y_2 = \beta_1 + \beta_2 X_2 - \beta_3 X_2^2 + U_2$$

$$Y_3 = \beta_1 + \beta_2 X_3 - \beta_3 X_3^2 + U_3$$

$$Y_4 = \beta_1 + \beta_2 X_4 - \beta_3 X_4^2 + U_4$$

$$Y_5 = \beta_1 + \beta_2 X_5 - \beta_3 X_5^2 + U_5$$

$$Y_6 = \beta_1 + \beta_2 X_6 - \beta_3 X_6^2 + U_6$$

$$Y_7 = \beta_1 + \beta_2 X_7 - \beta_3 X_7^2 + U_7$$

$$Y_8 = \beta_1 + \beta_2 X_8 - \beta_3 X_8^2 + U_8$$

$$Y_9 = \beta_1 + \beta_2 X_9 - \beta_3 X_9^2 + U_9$$

$$Y_{10} = \beta_1 + \beta_2 X_{10} - \beta_3 X_{10}^2 + U_{10}$$

$$Y_{11} = \beta_1 + \beta_2 X_{11} - \beta_3 X_{11}^2 + U_{11}$$

$$Y_{12} = \beta_1 + \beta_2 X_{12} - \beta_3 X_{12}^2 + U_{12}$$

$$Y_{13} = \beta_1 + \beta_2 X_{13} - \beta_3 X_{13}^2 + U_{13}$$

$$Y_{14} = \beta_1 + \beta_2 X_{14} - \beta_3 X_{14}^2 + U_{14}$$

$$Y_{15} = \beta_1 + \beta_2 X_{15} - \beta_3 X_{15}^2 + U_{15}$$

Representación del sistema de ecuaciones en notación matricial:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \\ Y_8 \\ Y_9 \\ Y_{10} \\ Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{15} \end{bmatrix} = \begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ 1 & X_3 & X_3^2 \\ 1 & X_4 & X_4^2 \\ 1 & X_5 & X_5^2 \\ 1 & X_6 & X_6^2 \\ 1 & X_7 & X_7^2 \\ 1 & X_8 & X_8^2 \\ 1 & X_9 & X_9^2 \\ 1 & X_{10} & X_{10}^2 \\ 1 & X_{11} & X_{11}^2 \\ 1 & X_{12} & X_{12}^2 \\ 1 & X_{13} & X_{13}^2 \\ 1 & X_{14} & X_{14}^2 \\ 1 & X_{15} & X_{15}^2 \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ -\beta_3 \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \\ U_7 \\ U_8 \\ U_9 \\ U_{10} \\ U_{11} \\ U_{12} \\ U_{13} \\ U_{14} \\ U_{15} \end{bmatrix}$$

A partir de la aplicación de la expresión $\beta = (X^T X)^{-1} (X^T Y)$ o

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = (X^T X)^{-1} (X^T Y)$$

obtenida en el epígrafe 1.6.1, se determinan los estimadores mínimos cuadráticos $\hat{\beta}_1$, $\hat{\beta}_2$ y $\hat{\beta}_3$. Para su determinación se utiliza el asistente matemático Derive versión 6.0, como se exhibe a continuación.

Representación de la matriz columna Y de orden o tamaño de la matriz 15×1 , correspondiente a la variable dependiente capacidad industrial, y de la matriz rectangular X de orden o tamaño de la matriz 15×3 , correspondiente a la variable explicativa tiempo.

$$Y = \begin{bmatrix} 2729,43 \\ 2804,27 \\ 3458,96 \\ 3069,45 \\ 4205,78 \\ 5207,66 \\ 5221,93 \\ 5165,48 \\ 4764,75 \\ 5172,02 \\ 4352,56 \\ 3646,38 \\ 2401,87 \\ 789,366 \\ 837,937 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \\ 1 & 8 & 64 \\ 1 & 9 & 81 \\ 1 & 10 & 100 \\ 1 & 11 & 121 \\ 1 & 12 & 144 \\ 1 & 13 & 169 \\ 1 & 14 & 196 \\ 1 & 15 & 225 \end{bmatrix}$$

Matriz transpuesta X^T .

La matriz transpuesta de la matriz X , se obtiene a partir de X intercambiando las filas por las columnas.

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & 81 & 100 & 121 & 144 & 169 & 196 & 225 \end{bmatrix}$$

Multiplicación de la matriz transpuesta X^T por la matriz X .

Como X^T y X son dos matrices de órdenes 3×15 y 15×3 , respectivamente, el producto de las matrices X^T y X , es la matriz $X^T X$ de orden 3×3 o de tercer orden, cuyos elementos para la situación objeto de estudio son números reales positivos, resultantes de la suma de los productos de los elementos de la fila i de la matriz X^T por los elementos correspondientes de la columna j de

la matriz X , resultando:

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ X_1 & X_2 & X_3 & \dots & X_5 \\ X_1^2 & X_2^2 & X_3^2 & \dots & X_5^2 \end{bmatrix} \begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ 1 & X_3 & X_3^2 \\ \cdot & \cdot & \cdot \\ 1 & X_5 & X_5^2 \end{bmatrix} = \begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_2 X_{3i} \\ \sum X_{3i} & \sum X_{3i} X_{2i} & \sum X_{3i}^2 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & 81 & 100 & 121 & 144 & 169 & 196 & 225 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \\ 1 & 8 & 64 \\ 1 & 9 & 81 \\ 1 & 10 & 100 \\ 1 & 11 & 121 \\ 1 & 12 & 144 \\ 1 & 13 & 169 \\ 1 & 14 & 196 \\ 1 & 15 & 225 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 15 & 120 & 1240 \\ 120 & 1240 & 14400 \\ 1240 & 14400 & 178321 \end{bmatrix}$$

La matriz cuadrada $(X^T X)$ de orden 3 es simétrica, porque se cumple que la matriz $(X^T X)$ es igual a la transpuesta de la matriz $(X^T X)$, es decir, $(X^T X) = (X^T X)^T$.

El determinante de la matriz cuadrada $(X^T X)$ de orden 3 o de tercer orden es el número real que se define como:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32})$$

Un procedimiento práctico desarrollado para obtener cada uno de los sumandos anteriores es la Regla de Sarrus, que consiste en agregar las dos primeras filas de la matriz original a la propia matriz; las flecha dirigidas hacia abajo indican el producto de esos tres elementos con el mismo signo y las flechas dirigidas hacia arriba indican el producto de esos tres elementos, y el resultado del producto se multiplica por menos uno (-1), es decir, al producto de los tres elementos se le cambia el signo.

$$\begin{vmatrix}
 a_{11} & a_{12} & a_{13} \\
 a_{21} & a_{22} & a_{23} \\
 a_{31} & a_{32} & a_{33} \\
 a_{11} & a_{12} & a_{13} \\
 a_{21} & a_{22} & a_{23}
 \end{vmatrix} = (a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}) - (a_{31}a_{22}a_{13} + a_{11}a_{32}a_{23} + a_{21}a_{12}a_{33})$$

Se obtiene como resultado del determinante de la matriz $(X^T X)$, el siguiente:

$$\det (X^T X) = \begin{vmatrix} 15 & 120 & 1240 \\ 120 & 1240 & 14400 \\ 1240 & 14400 & 178321 \end{vmatrix} = 17326400$$

Como el determinante de la matriz cuadrada $(X^T X)$ igual a 17326400, es diferente de cero, la matriz cuadrada es no singular o regular; entonces se puede calcular la inversa por el método de la matriz adjunta.

Cálculo de la matriz inversa $(X^T X)^{-1}$ mediante el método de la matriz adjunta.

$$(X^T X)^{-1} = \frac{1}{\det (X^T X)} (X^T X)^+ = \frac{1}{|X^T X|} (X^T X)^+$$

La matriz adjunta de la matriz cuadrada $(X^T X)$ es:

$$(X^T X)^+ = \begin{bmatrix} 13746880 & -3541440 & 190400 \\ -3541440 & 1137080 & -67200 \\ 190400 & -67200 & 4200 \end{bmatrix}$$

Resultando como matriz inversa:

$$(X^T X)^{-1} = \frac{1}{17326400} \begin{bmatrix} 13746880 & -3541440 & 190400 \\ -3541440 & 1137080 & -67200 \\ 190400 & -67200 & 4200 \end{bmatrix} = \begin{bmatrix} \frac{361}{455} & -\frac{93}{455} & \frac{1}{91} \\ \frac{93}{455} & \frac{4061}{61880} & -\frac{6}{1547} \\ \frac{1}{91} & -\frac{6}{1547} & \frac{3}{12376} \end{bmatrix}$$

Multiplicación de la matriz transpuesta de $X(X^T)$ por la matriz columna Y , se obtiene la matriz $(X^T Y)$.

Como X^T y Y son dos matrices de órdenes 3×15 y 15×1 , respectivamente, el producto de las matrices X^T y Y , es la matriz $(X^T Y)$ de orden 3×1 , resultando:

$$X^T Y = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ X_1 & X_2 & X_3 & \dots & X_{15} \\ X_1^2 & X_2^2 & X_3^2 & \dots & X_{15}^2 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_{15} \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \sum X_{3i} Y_i \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & 81 & 100 & 121 & 144 & 169 & 196 & 225 \end{bmatrix} \begin{bmatrix} 2729,43 \\ 2804,27 \\ 3458,96 \\ 3069,45 \\ 4205,78 \\ 5207,66 \\ 5221,93 \\ 5165,48 \\ 4764,75 \\ 5172,02 \\ 4352,56 \\ 3646,38 \\ 2401,87 \\ 789,366 \\ 837,937 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 53863,843 \\ 402767,019 \\ 3685426,721 \end{bmatrix}$$

Determinando a $\hat{\beta}_1$, $\hat{\beta}_2$ y $\hat{\beta}_3$ como resultado de la multiplicación de las matrices $(X^T X)^{-1}$ y $(X^T Y)$.

$$(X^T X)^{-1} (X^T Y) = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \frac{361}{455} & -\frac{93}{455} & \frac{1}{91} \\ -\frac{93}{455} & \frac{4061}{61880} & -\frac{6}{1547} \\ \frac{1}{91} & -\frac{6}{1547} & \frac{3}{12376} \end{bmatrix} \begin{bmatrix} 53863,843 \\ 402767,019 \\ 3685426,721 \end{bmatrix} = \begin{bmatrix} \frac{414648161}{455000} \\ \frac{13972912971}{12376000} \\ -\frac{951054101}{12376000} \end{bmatrix}$$

$$(X^T X)^{-1} (X^T Y) = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \frac{414648161}{455000} \\ \frac{13972912971}{12376000} \\ -\frac{951054101}{12376000} \end{bmatrix} = \begin{bmatrix} 911,3146 \\ 1129,0330 \\ -76,8466 \end{bmatrix}$$

Valores de los estimadores mínimos cuadráticos obtenidos por la vía matricial:

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 911,3146 \\ 1129,0330 \\ -76,8466 \end{bmatrix}$$

Por tanto, la función de regresión mínimo cuadrática o ecuación de predicción estimada es:

$$\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2.$$

Representación matricial de la ecuación de predicción estimada \hat{Y} .

$$\hat{Y} = \begin{bmatrix} 1 & X_t & X_t^2 \end{bmatrix} \begin{bmatrix} 911,3146 \\ 1129,0330 \\ -76,8466 \end{bmatrix}$$

Las matrices $(X^T X)$ y $(X^T Y)$ también se pueden obtener calculando sus elementos como se exhibe a continuación.

$$(X^T X) = \begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i} X_{3i} \\ \sum X_{3i} & \sum X_{3i} X_{2i} & \sum X_{3i}^2 \end{bmatrix}$$

$$(X^T Y) = \begin{bmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \sum X_{3i} Y_i \end{bmatrix}$$

Siendo $(X^T X)$ una matriz cuadrada de tercer orden y simétrica.

Se tabula la información de las variables X y Y para calcular los valores correspondientes a los elementos de las matrices $(X^T X)$ y $(X^T Y)$.

Tabla 1.4. Información para determinar las matrices $(X^T X)$ y $(X^T Y)$.

Y	X_2	X_3	X_2^2	X_3^2	$X_2 X_3$	YX_2	YX_3
2729,430	1	1	1	1	1	2729,430	2729,430
2804,270	2	4	4	16	8	5608,540	11217,080
3458,960	3	9	9	81	27	10376,880	31130,640
3069,450	4	16	16	256	64	12277,800	49111,200
4205,780	5	25	25	625	125	21028,900	105144,500
5207,660	6	36	36	1296	216	31245,960	187475,760
5221,930	7	49	49	2401	343	36553,510	255874,570
5165,480	8	64	64	4096	512	41323,840	330590,720
4764,750	9	81	81	6561	729	42882,750	385944,750
5172,020	10	100	100	10000	1000	51720,200	517202,000
4352,560	11	121	121	14641	1331	47878,160	526659,760
3646,380	12	144	144	20736	1728	43756,560	525078,720
2401,870	13	169	169	28561	2197	31224,310	405916,030
789,366	14	196	196	38416	2744	11051,124	154715,736
873,930	15	225	225	50625	3375	13108,950	196634,250
53863,836	120	1240	1240	178312	14400	402766,914	3685425,150
$\sum Y_i$	$\sum X_{2i}$	$\sum X_{3i}$	$\sum X_{2i}^2$	$\sum X_{3i}^2$	$\sum X_{2i} X_{3i}$	$\sum X_{2i} Y_i$	$\sum X_{3i} Y_i$

Fuente: elaborada por los autores.

$$\left(X^T X \right) = \begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i} X_{3i} \\ \sum X_{3i} & \sum X_{3i} X_{2i} & \sum X_{3i}^2 \end{bmatrix} = \begin{bmatrix} 15 & 120 & 1240 \\ 120 & 1240 & 14400 \\ 1240 & 14400 & 178312 \end{bmatrix}$$

$$\left(X^T Y \right) = \begin{bmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \sum X_{3i} Y_i \end{bmatrix} = \begin{bmatrix} 53863,836 \\ 402766,914 \\ 3685425,150 \end{bmatrix}$$

Por las razones expuestas en el análisis estadístico, a partir del promedio de la caña procesada semanalmente en cinco zafras, se ajusta la curva continua, representada por la función de regresión estimada $\hat{E}_{(Y_i/X_i)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$, con coeficiente de determinación y coeficiente de determinación ajustado asociados a valores altos, lo que demuestra que la variable predictora, explica en buena medida la variación de la variable dependiente.

Obtención de los estimadores mínimos cuadráticos de la función de regresión capacidad industrial, utilizando el paquete estadístico EViews.

Tabla 1.5. Estimadores mínimos cuadráticos $\hat{\beta}_1, \hat{\beta}_2$ y $\hat{\beta}_3$ (EViews).

Dependent Variable: Y Method: Least Squares Sample: 1 15 Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	$\hat{\beta}_1 = 911.3146$	494.3730	1.843374	0.0901
X	$\hat{\beta}_2 = 1129.033$	142.1832	7.940693	0.0000
X^2	$\hat{\beta}_3 = -76.84665$	8.641266	-8.892985	0.0000
R-squared	0.880321	Mean dependent var		3590.923
Adjusted R-squared	0.860375	S.D. dependent var		1485.336
S.E. of regression	555.0178	Akaike info criterion		15.65273
Sum squared resid	3696537.	Schwarz criterion		15.79434
Log likelihood	-114.3955	F-statistic		44.13416
Durbin-Watson stat	1.468850	Prob(F-statistic)		0.000003

II- Determinación del aprovechamiento máximo u óptimo de la capacidad industrial

Para determinar la semana de máximo u óptimo aprovechamiento de las capacidades instaladas, correspondiente a la función estimada, se procede a derivar la función anterior con respecto a X , resultando: $\frac{d\hat{Y}}{dX} = \hat{Y}'_t = 1129,033 - 153,692X_t$, igualando a cero para hallar el punto

crítico, se obtiene que $X_0 = 7,35$ semana.

El procedimiento para calcular a X_0 , semana de máximo u óptimo aprovechamiento de las capacidades, se puede simplificar aplicando de la expresión $X_0 = -\frac{\hat{\beta}_2}{2\hat{\beta}_3}$, donde, $X_0 = -\frac{1129,033}{2(-76,846)} = 7,35$.

La segunda derivada de la función es $\frac{d^2\hat{Y}}{dX^2} = \hat{Y}'' = -153,692$; como $\hat{Y}'' < 0$, para cualquier valor del dominio de la función; se puede afirmar que para $X_0 = 7,35$ semana, la función de regresión estimada alcanza su valor máximo u óptimo.

Los resultados exhibidos se consolidan cuando se analizan las propiedades de la función regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$, en el epígrafe 1.6.6.

Evaluando la función de regresión para el valor de X_0 , $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033(7,35)_t - 76,846(7,35)_t^2 = 5058,29$ se obtiene la capacidad industrial máxima, \hat{Y}_{\max} , igual a 5058,29 toneladas. Otra forma para calcular la capacidad máxima es a partir de los estimadores mínimos cuadráticos, como

$$\hat{Y}_{\max} = \hat{\beta}_1 - (\hat{\beta}_2^2 / 4\hat{\beta}_3) = 911,314 - \left[(1129,033)^2 / 4(-76,846) \right] = 5058,29$$

Por tanto, la capacidad industrial operacional óptima semanal del centro de recepción es de 5058,29 toneladas, que representa el 85,75 % de la capacidad industrial disponible total o capacidad potencial.

Multiplcado este aprovechamiento óptimo de las capacidades del centro de recepción por las 12 semanas de zafra operacionales planificados, se obtiene la capacidad industrial óptima del centro para la zafra, que es de 60339,48 toneladas.

Los resultados satisfactorios obtenidos de aplicar de la función de regresión polinomial cuadráticas, demuestran que la función seleccionada es buena predictora para proyectar la capacidad industrial óptima de los centros de recepción.

III- Inferencia realizada al modelo de regresión lineal múltiple asociado a la capacidad industrial

Después de haber estudiado el modelo de regresión lineal múltiple asociado a la capacidad industrial, se pueden hacer inferencias de la población objeto de estudio.

La validez de la función de regresión ajustada se comprueba a partir del análisis de varianza en la regresión y la aplicación de la prueba de F de Fisher.

Prueba de F de Fisher

1- Formulación de las hipótesis nula y alternativa:

$$H_0: \beta_2 = \beta_3 = 0$$

$$H_1: \text{Existe algún } \beta_j \neq 0, \text{ para } j = 2, 3.$$

La hipótesis nula indica que el modelo no explica la situación objeto de estudio, en contraposición la hipótesis alternativa a contrastar expresa que el modelo sí explica la situación objeto de estudio, es decir, la función de regresión estimada, es buena predictora y puede ser utilizada para predecir los valores de la variable aleatoria Y a partir de los valores dados de X .

2- Región crítica para el estadígrafo F , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0: F > F_{1-\alpha}(p-1, n-p) \right\}.$$

$$\text{donde: } F = \frac{S_E^2}{S_R^2} \text{ o } F = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R}$$

3- Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha}(p-1, n-p)$.

- No rechazar H_0 , si $F \leq F_{1-\alpha}(p-1, n-p)$.

Alternativamente:

- Rechazar H_0 , si la Prob (F -statistic) o valor $p < \alpha$.

- No rechazar H_0 , si la Prob (F -statistic) o valor $p > \alpha$.

3.1- Cálculo del estadístico de prueba F : $F = \frac{S_E^2}{S_R^2}$ o $F = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R}$.

A partir del análisis de varianza en la regresión se calcula el estadístico de prueba F , como se muestra en la tabla de análisis de varianza.

Tabla 1.6. Tabla de Análisis de Varianza (ANOVA) por la vía matricial.

Fuente de variación	Suma de cuadrados SC	Grados de libertad gl	Suma de cuadrados medios SCM
Debido a la regresión (SC_E)	$SC_E = \underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2$	$(p-1)$	$SCM \text{ de } SC_E = \frac{SC_E}{p-1} = S_E^2$
Debido a los residuos (SC_R)	$SC_R = \underline{Y}^T \underline{Y} - \underline{\beta}^T \underline{X}^T \underline{Y}$ $SC_R = SC_T - SC_E$	$(n-p)$	$SCM \text{ de } SC_R = \frac{SC_R}{n-p} = S_R^2$
Total (SC_T)	$SC_T = \underline{Y}^T \underline{Y} - n\bar{Y}^2$	$(n-1)$	////

Fuente: elaborada por los autores.

Para calcular a F hay que determinar las sumas de cuadrados: SC_E , SC_T y SC_R , y las varianzas: S_E^2 y S_R^2 . A continuación, se determinan las sumas de cuadrados por la vía matricial, mediante el asistente matemático Derive.

- Suma de cuadrados debido a la regresión o explicada por la regresión.

Información necesaria para calcular la $SC_E = \underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2$:

$$\beta = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 911,3146 \\ 1129,0330 \\ -76,8466 \end{bmatrix} \quad \beta^T = [\hat{\beta}_1 \quad \hat{\beta}_2 \quad \hat{\beta}_3] = [911,31 \quad 1129,03 \quad -76,84]$$

$$Y = \begin{bmatrix} 2729,43 \\ 2804,27 \\ 3458,96 \\ 3069,45 \\ 4205,78 \\ 5207,66 \\ 5221,93 \\ 5165,48 \\ 4764,75 \\ 5172,02 \\ 4352,56 \\ 3646,38 \\ 2401,87 \\ 789,366 \\ 837,937 \end{bmatrix} \quad X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & 81 & 100 & 121 & 144 & 169 & 196 & 225 \end{bmatrix}$$

$$\sum_{i=1}^{15} Y_i = 53863,836$$

Cálculo de la $SC_E = \underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2$

$$SC_E = [911,31 \quad 1129,03 \quad -76,84] \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & 81 & 100 & 121 & 144 & 169 & 196 & 225 \end{bmatrix} \begin{bmatrix} 2729,43 \\ 2804,27 \\ 3458,96 \\ 3069,45 \\ 4205,78 \\ 5207,66 \\ 5221,93 \\ 5165,48 \\ 4764,75 \\ 5172,02 \\ 4352,56 \\ 3646,38 \\ 2401,87 \\ 789,366 \\ 837,937 \end{bmatrix} - 15 \cdot \left(\frac{53863,836}{15} \right)^2$$

$$SC_E = 220611649,2426 - 193420855,2423$$

$$SC_E = 27190794,00$$

- *Suma de cuadrados total:*

Información necesaria para calcular la $SC_T = \underline{Y}^T \underline{Y} - n\bar{Y}^2$:

$$\underline{Y}^T = \left[\begin{array}{cccccccccccccccccccc} 2729,43 & 2804,27 & 3458,96 & 3069,45 & 4205,78 & 5207,66 & 5221,93 & 5165,48 & 4764,75 & 5172,02 & 4352,56 & 3646,38 & 2401,87 & 789,366 & 837,937 \end{array} \right]$$

$$\underline{Y}' \underline{Y} = \left[\begin{array}{cccccccccccccccccccc} 2729,43 & 2804,27 & 3458,96 & 3069,45 & 4205,78 & 5207,66 & 5221,93 & 5165,48 & 4764,75 & 5172,02 & 4352,56 & 3646,38 & 2401,87 & 789,366 & 837,937 \end{array} \right]$$

$$\underline{Y}^T \underline{Y} = 224308033,7409$$

$$SC_T = \underline{Y}^T \underline{Y} - n\bar{Y}^2$$

$$SC_T = 224308033,7409 - 193420855,2423 \quad SC_T = 30887178,4986$$

- *Suma de cuadrados debido a los residuos:*

$$SC_R = \underline{Y}^T \underline{Y} - \underline{\beta}^T \underline{X}^T \underline{Y}$$

$$SC_R = 224308033,7409 - 220611649,2426 \quad SC_R = 3696384,49$$

Además, SC_R se puede calcular como:

$$SC_R = SC_T - SC_E$$

$$SC_R = 30887178,49 - 27190794 \quad SC_R = 3696384,49$$

Varianza asociada a la regresión y a los residuos:

$$SCM \text{ de } SC_E = \frac{SC_E}{p-1} = \frac{27190794}{3-1} = 13595397 = S_E^2$$

$$SCM \text{ de } SC_R = \frac{SC_R}{n-p} = \frac{3696384,49}{15-3} = 308032,04 = S_R^2$$

Estadístico de prueba o estadígrafo F :

$$F = \frac{S_E^2}{S_R^2} = \frac{13595397}{308032,04} = 44,13 \text{ o}$$

$$F = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R} = \frac{27190794}{3-1} \cdot \frac{15-3}{3696384,49} = 44,13$$

Tabla 1.7. Resultado del Análisis de Varianza (ANOVA) por la vía matricial.

Fuente de Variación	Suma de Cuadrado SC	Grados de libertad	Suma de Cuadrados Medios SCM	F
Debido a la regresión (SC_E)	$SC_E = \underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2$ 27190794	$(p-1)$ 3-1=2	$SCM \text{ de } SC_E = \frac{SC_E}{p-1} = S_E^2$ 13595397	44,13
Debido a los residuos (SC_R)	$SC_R = \underline{Y}^T \underline{Y} - \underline{\beta}^T \underline{X}^T \underline{Y}$ 3696384,49	$(n-p)$ 15-3=12	$SCM \text{ de } SC_R = \frac{SC_R}{n-p} = S_R^2$ 308032,04	
Total (SC_T)	$SC_T = \underline{Y}^T \underline{Y} - n\bar{Y}^2$ 30887178,49	$(n-1)$ 15-1=14	////	

Fuente: elaborada por los autores.

3.2- Determinar el valor crítico o teórico:

$$F_{1-\alpha}(p-1, n-p) = F_{1-0,05}(3-1, 15-3) = F_{0,95}(2, 12) = 3,98.$$

Como:

- $F > F_{1-\alpha}(p-1, n-p)$, o sea $44,134 > 3,98$, se rechaza H_0 .

- La Prob (F -statistic) $< \alpha$, o sea $0,000003 < 0,05$, se rechaza H_0 .

A partir del análisis de varianza en la regresión y con la aplicación de la prueba F de Fisher, tomando los resultados expresados anteriormente y en el informe del programa EViews de la tabla 1.8, se puede comprobar que el valor del estadístico de prueba F igual a 44,13 es superior al valor crítico F de 3,89, y se ubica dentro de la región crítica; por lo que se rechaza la hipótesis nula $\beta_2 = \beta_3 = 0$, para un nivel de significación de 0,05. Otra forma de corroborar lo analizado viene dado porque la probabilidad de F -statistic o valor p presenta un valor extremadamente bajo de 0,000003 menor que el nivel de significación de 0,05.

Por tanto, se llega a la conclusión que, basado en los resultados de la muestra y con un nivel de significación de 0,05, se puede afirmar que la función de regresión cuadrática estimada, $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$ es de buen ajuste y puede ser utilizada para predecir el valor medio estimado de la capacidad industrial, a partir de valores dados del tiempo.

Cálculo del coeficiente de determinación múltiple R^2 :

$$R^2 = \frac{SC_E}{SC_T} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SC_R}{SC_T}$$

$$R^2 = \frac{27190794}{30887178,49} = 1 - \frac{3696384,49}{30887178,49} = 0,8803$$

A partir del resultado de la relación anterior y del informe de la tabla 1.8, el coeficiente de determinación múltiple R^2 es de 0,8803.

La función de regresión $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$

es una función de regresión polinomial con una sola variable explicativa, pero elevada a distintas potencias, por lo que se considera como un modelo de regresión múltiple. No obstante, la variable dependiente capacidad industrial del centro de recepción Y , depende de una sola variable explicativa tiempo, X .

Entonces, el coeficiente de determinación significa que el 88,03 % de la variación de la capacidad industrial del centro de recepción es explicado por el tiempo, o sea, el 88,03 % de la variación total queda explicado por la ecuación de regresión y se considera que el ajuste es bueno. Existe una cuantía o porcentaje considerablemente alto del poder explicativo.

Por el resultado obtenido, se considera que el tiempo es una buena variable predictora para determinar la capacidad de explotación industrial de los centros de recepción, y que la función de regresión cuadrática estimada de la muestra se ajusta convenientemente a los datos.

Cálculo del coeficiente de determinación ajustado \bar{R}^2 :

$$\bar{R}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / n - p}{\sum_{i=1}^n (y_i - \bar{y})^2 / n - 1} = 1 - \frac{SC_R / n - p}{SC_T / n - 1}$$

$$\bar{R}^2 = 1 - \frac{3696384 / 12}{30887178,49 / 14} = 1 - \frac{308032,04}{2206227,035} = 1 - 0,1396 = 0,8603$$

Expresa que después de tener en cuenta los grados de libertad, el tiempo X , explica en un 86,03% la variación de la capacidad industrial del centro de recepción Y .

Tabla 1.8. Resultados de $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, estadígrafo F , R^2 y \bar{R}^2 (EViews).

Dependent Variable: Y				
Method: Least Squares				
Sample: 1 15				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	$\hat{\beta}_1 = 911.3146$	494.3730	1.843374	0.0901
X	$\hat{\beta}_2 = 1129.033$	142.1832	7.940693	0.0000
X^2	$\hat{\beta}_3 = -76.84665$	8.641266	-8.892985	0.0000
R-squared	0.880321	Mean dependent var		3590.923
Adjusted R-squared	0.860375	S.D. dependent var		1485.336
S.E. of regression	555.0178	Akaike info criterion		15.65273
Sum squared resid	3696537.	Schwarz criterion		15.79434
Log likelihood	-114.3955	F-statistic	$F = \frac{S_E^2}{S_R^2} = 44.13416$	
Durbin-Watson stat	1.468850	Prob(F-statistic)	Valor p = 0.00000	

$$\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$$

En la tabla 1.9 y en la figura 1.4, se representa la estimación de la utilización de la capacidad industrial del centro de recepción correspondiente al período objeto de estudio.

Tabla 1.9. Valores de la capacidad industrial real y estimada (EViews).

Semana	Capacidad industrial real	Capacidad industrial estimada	Semana	Capacidad industrial real	Capacidad industrial estimada
1	2729,43	1963,50	9	4764,75	4848,03
2	2804,27	2861,99	10	5172,02	4516,98
3	3458,96	3606,79	11	4352,56	4032,23
4	3069,45	4197,90	12	3646,38	3393,79
5	4205,78	4635,31	13	2401,87	2601,66
6	5207,66	4919,03	14	789,36	1655,83
7	5221,93	5049,06	15	873,93	556,31
8	5165,48	5025,39	-	-	-

Fuente: elaborada por los autores.

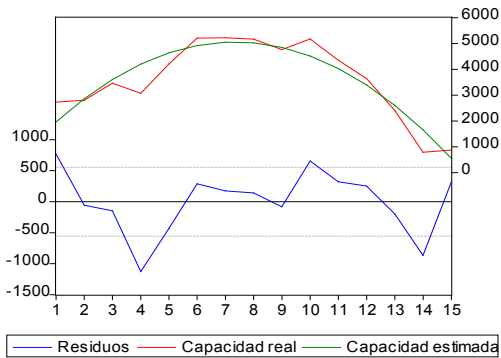


Figura 1.4. Comportamiento de la capacidad industrial real y estimada (EViews).

Para verificar la confiabilidad de la función de regresión ajustada y comprobar la validez de las proyecciones realizadas, se aplican las pruebas de hipótesis relacionadas con los supuestos básicos del modelo.

Verificación del supuesto de no autocorrelación. Prueba de Durbin-Watson

Para comprobar la existencia de autocorrelación en los residuos, se sigue el siguiente procedimiento:

- Estimar el modelo por el Método de los Mínimos Cuadrados Ordinarios bajo el supuesto de no autocorrelación.
- Determinar el estadístico de prueba d de Durbin-Watson a partir de los residuos del modelo e_t .
- Contrastar las hipótesis.

La estimación de la función de regresión cuadrática se realizó de forma manual y mediante el programa estadístico EViews, bajo el supuesto de no autocorrelación, se obtuvo: $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$.

Determinación del estadístico de prueba d de Durbin-Watson a partir de los residuos del modelo, mediante su expresión:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Seguidamente, se muestran los residuos obtenidos mediante el programa estadístico EViews.

Tabla 1.10. Error de estimación o residuos y su ploteo (EViews).

Y_i	\hat{Y}_i	$(Y_i - \hat{Y}_i) = e_i$	Residual Plot
Actual	Fitted	Residual	
2729,43	1963,50	765,929	. . *
2804,27	2861,99	-57,724	. * .
3458,96	3606,79	-147,834	. * .
3069,45	4197,90	-1128,450	* . .
4205,78	4635,31	-429,534	. * .
5207,66	4919,03	288,626	. * .
5221,93	5049,06	172,870	. * .
5165,48	5025,39	140,086	. * .
4764,75	4848,03	-83,283	. * .
5172,02	4516,98	655,040	. * .
4352,56	4032,23	320,326	. * .
3646,38	3393,79	252,586	. * .
2401,87	2601,66	-199,791	. * .
789,36	1655,83	-866,469	* . .
873,93	556,31	317,622	. * .

Fuente: elaborada por los autores.

Tabulando la información necesaria para determinar el estadístico de prueba d de Durbin-Watson.

Tabla 1.11. Información necesaria para calcular el estadístico de Durbin-Watson.

Y_i	\hat{Y}_i	$(Y_i - \hat{Y}_i) = e_i$	$e_i - e_{i-1}$	$(e_i - e_{i-1})^2$	e_i^2
2729,43	1963,50	765,93	-	-	586648,76
2804,27	2861,99	-57,72	-823,65	678399,32	3331,59
3458,96	3606,79	-147,83	-90,11	8119,81	21853,70
3069,45	4197,90	-1128,45	-980,62	961615,58	1273399,40
4205,78	4635,31	-429,53	698,92	488489,16	184496,02
5207,66	4919,03	288,63	718,16	515753,78	83307,27
5221,93	5049,06	172,87	-115,76	13400,37	29884,03
5165,48	5025,39	140,09	-32,78	1074,52	19625,20
4764,75	4848,03	-83,28	-223,37	49894,15	6935,55
5172,02	4516,98	655,04	738,32	545116,42	429077,40
4352,56	4032,23	320,33	-334,71	112030,78	102611,30
3646,38	3393,79	252,59	-67,74	4588,70	63801,70
2401,87	2601,66	-199,79	-452,38	204647,66	39916,04
789,366	1655,83	-866,46	-666,67	444454,22	750759,86
873,937	556,314	317,62	1184,08	1402062,02	100884,37
Total				$5429646,55$	$3696532,27$
				$\sum_{i=2}^n (e_i - e_{i-1})^2$	$\sum_{i=1}^n e_i^2$

Fuente: elaborada por los autores.

Donde:

$$\sum_{i=2}^n (e_i - e_{i-1})^2 = 5429646,5 \qquad \sum_{i=1}^n e_i^2 = 3696532,27 = SC_R$$

Entonces:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} = \frac{5429646,55}{3696532,27} = 1,4688$$

Determinación del estadístico de prueba d de Durbin-Watson, a partir del reporte de salida del programa estadístico EViews.

Tabla 1.12. Estadístico de prueba Durbin-Watson (EViews).

Dependent Variable: Y				
Method: Least Squares				
Sample: 1 15				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	911.3146	494.3730	1.843374	0.0901
X	1129.033	142.1832	7.940693	0.0000
X^2	-76.84665	8.641266	-8.892985	0.0000
R-squared	0.880321	Mean dependent var		3590.923
Adjusted R-squared	0.860375	S.D. dependent var		1485.336
S.E. of regression	555.0178	Akaike info criterion		15.65273
Sum squared resid	3696537.	Schwarz criterion		15.79434
Log likelihood	-114.3955	F-statistic		44.13416
Durbin-Watson stat	d = 1.468850	Prob(F-statistic)		0.000003

De ambas formas, se obtiene el valor del estadístico de prueba de Durbin-Watson $d = 1,468$.

Formulación de la hipótesis nula y alternativa, consideradas:

$$H_0 : \rho_{U,U-1} = 0 \Leftrightarrow Cov(u, u-1) = 0 \text{ No hay autocorrelación.}$$

$$H_1 : \rho_{U,U-1} \neq 0 \Leftrightarrow Cov(u, u-1) \neq 0 \text{ Hay autocorrelación.}$$

Determinando los valores tabulados que se encuentran representados por los valores críticos d_u valor máximo y d_L valor mínimo, tabulados en dependencia del número de observaciones $n = 15$ y del número de variables explicatorias $k' = 2$, y en función del nivel de significación, como la prueba de hipótesis es bilateral o de dos colas, es decir, es una alternativa bilateral, entonces para un $\alpha = 0,05$ se tendría que buscar en $\alpha / 2 = 0,025$ o 2,5%, obteniéndose en la tabla estadística como resultado, los valores críticos máximo y mínimo siguientes: $d_u = 1,40$ y $d_L = 0,83$.

Regla de decisión:

↓	Rechazar H_0 hay autocorrelación positiva	↓	Zona de indeterminación o de duda	↓	No rechazar H_0	↓	Zona de indeterminación o de duda	↓	Rechazar H_0 hay autocorrelación negativa	↓
0		d_L		d_u	2	$4 - d_u$		$4 - d_L$		4
		0,83		1,40	Valor de	2,60		3,17		
					$d = 1,468$					

Mediante la Prueba de Durbin-Watson se determinó que el valor del estadístico de prueba d es 1,468, cumple la condición $d_u < d < 4 - d_u$, o sea, $1,40 < 1,468 < 2,64$. Por tanto, no se rechaza la hipótesis nula, $H_0 : \rho_{U, U-1} = 0 \Leftrightarrow Cov(U, U-1) = 0$, por lo que se puede afirmar que no existe autocorrelación serial de primer orden en los residuos.

No hay necesidad de aplicar la prueba Breusch-Godfrey para determinar la no existencia de autocorrelación de orden superior.

Verificación del supuesto de homoscedasticidad. Pruebas de Goldfeld-Quandt y de White

Prueba de Goldfeld-Quandt

Para realizar la prueba, se tiene en cuenta el procedimiento descrito en el epígrafe 1.6.2.

1- Formulación de las hipótesis nula y alternativa:

Se utiliza una prueba para determinar si las varianzas de los dos grupos de muestras son iguales, es decir:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

2- Región crítica para el estadístico de prueba $\lambda = F$, dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0: F > F_{1-\alpha} \left[(n-c-2p)/2 ; (n-c-2p)/2 \right] \right\}$$

$$W = \left\{ \lambda \in \mathfrak{R} \text{ con } \lambda \geq 0: \lambda > F_{1-\alpha} \left[(n-c-2p)/2 ; (n-c-2p)/2 \right] \right\}$$

Donde el estadístico de prueba es $\lambda = F = \frac{SC_{R2}}{\frac{(n-c-2p)/2}{SC_{R1}}}$.

3- Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha} \left[(n-c-2p)/2 ; (n-c-2p)/2 \right]$.

- No rechazar H_0 , si $F \leq F_{1-\alpha} \left[(n-c-2p)/2 ; (n-c-2p)/2 \right]$.

3.1- Cálculo del estadístico de prueba $\lambda = F$:

$$\lambda = F = \frac{SC_{R2}}{\frac{(n-c-2p)/2}{SC_{R1}}}$$

Para calcular el estadístico de prueba $\lambda = F$, se determina primeramente la suma de cuadrados de los residuos: SC_{R1} y SC_{R2} .

En la tabla 1.13, se ordena la información de los datos de mayor a menor, tomando como referencia la variable explicativa X , y se dividen las observaciones en dos muestras: una que recoge los menores valores y otra que recoge los mayores valores de la muestra, sin considerar los valores centrales.

Tabla 1.13. Submuestras de los valores mayores y menores prueba de Goldfeld-Quandt.

Submuestra de las primeras observaciones correspondiente a los valores mayores de X	873,937	15	Se le determina la suma de cuadrados de los residuos a las primeras observaciones correspondiente a los valores mayores de X . SC_{R2}
	789,366	14	
	2401,87	13	
	3646,38	12	
	4352,56	11	
	5172,02	10	
Valores centrales C	4764,75	9	
	5165,48	8	
	5221,93	7	
Submuestra de las segundas observaciones correspondiente a los valores menores de X	5207,66	6	Se le determina la suma de cuadrados de los residuos a las segundas observaciones correspondiente a los valores menores de X . SC_{R1}
	4205,78	5	
	3069,45	4	
	3458,96	3	
	2804,27	2	
	2729,43	1	

Fuente: elaborada por los autores.

Estimación de la submuestra de las primeras observaciones, correspondiente a los valores mayores de X y determinación de la suma de cuadrados de los residuos SC_{R2} .

Tabla 1.14. Regresión para las primeras observaciones correspondiente a los valores mayores de X (EViews).

Dependent Variable: Y1				
Method: Least Squares				
Sample(adjusted): 1 6				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	17260.22	11881.82	1.452658	0.2423
X1	-1354.477	1926.694	-0.703006	0.5327
X1^2	15.97963	76.93655	0.207699	0.8488
R-squared	0.960141	Mean dependent var	2872.689	
Adjusted R-squared	0.933568	S.D. dependent var	1823.867	
S.E. of regression	470.0901	Akaike info criterion	15.45058	
Sum squared resid	$SC_{R2} = 662954.1$	Schwarz criterion	15.34646	
Log likelihood	-43.35174	F-statistic	36.13259	
Durbin-Watson stat	2.458210	Prob(F-statistic)	0.007958	

Función de regresión estimada $\hat{Y}_1 = 17260,22 - 1354,47X_1 + 15,97X_1^2$,

donde la suma de cuadrados de los residuos SC_{R2} es 662954,1.

Estimación de la submuestra de las segundas observaciones, correspondiente a los valores menores de X y determinación la suma de cuadrados de los

residuos SC_{R1} .

Tabla 1.15. Regresión para las segundas observaciones correspondiente a los valores menores de X (EViews).

Dependent Variable: Y2 Method: Least Squares Sample(adjusted): 10 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3052.268	634.0000	4.814303	0.0171
X2	-357.1866	414.7802	-0.861147	0.4525
X2^2	117.1743	58.00506	2.020070	0.1366
R-squared	0.918811	Mean dependent var		3579.258
Adjusted R-squared	0.864684	S.D. dependent var		963.4742
S.E. of regression	354.4168	Akaike info criterion		14.88568
Sum squared resid	$SC_{R1} = 376833.8$	Schwarz criterion		14.78156
Log likelihood	-41.65703	F-statistic		16.97530
Durbin-Watson stat	2.951932	Prob(F-statistic)		0.023134

Función de regresión estimada $\hat{Y}_2 = 3052,26 - 357,18X_2 + 117,17X_2^2$, donde la suma de cuadrados de los residuos SC_{R1} es 376833,8.

Calcular el cociente de la suma de cuadrados de los residuos de ambas regresiones.

Para detectar la heterocedasticidad, a partir de la prueba de Goldfeld-Quandt, se determina el valor estimado $\lambda = F$, obtenido como el cociente de la suma de cuadrados de los residuos de las primeras observaciones, correspondiente a los valores mayores de X , con respecto a la suma de cuadrados de los residuos de las segundas observaciones, correspondiente a los valores menores de X , es decir:

$$\lambda = F = \frac{\frac{SC_{R2}}{(n - c - 2p) / 2}}{\frac{SC_{R1}}{(n - c - 2p) / 2}}$$

$$\lambda = F = 662954 / 376833,8 = 1,75$$

3.2- Determinar el valor crítico o teórico de $\lambda = F$:

$$F_{1-\alpha} [(n - c - 2p) / 2 ; (n - c - 2p) / 2] = F_{0,95} [(15 - 3 - 6) / 2 , (15 - 3 - 6) / 2] = F_{0,95} (3;3) = 9,28$$

Como: $F \leq F_{1-\alpha} [(n - c - 2p) / 2 ; (n - c - 2p) / 2]$, o sea $1,20 < 6,39$, no se rechaza H_0 .

Como el valor obtenido $\lambda = F = 662954/376833,8 = 1,75$ no excede el valor crítico o teórico de $F_{0,95}(3;3)$ igual a 9,28, entonces no se rechaza la hipótesis nula que se somete a prueba, infiriéndose que existe homocedasticidad en las varianzas del error, pues las varianzas se mantienen constantes para todas las observaciones y no dependen de la observación correspondiente. Se está en presencia de un modelo con residuos homocedásticos, siendo los estimadores obtenidos insesgados y eficientes.

Pruebas de White

Una vía general para detectar la heterocedasticidad lo constituye la prueba de White. Para aplicar la prueba inicialmente, a partir de la serie de datos correspondientes a la capacidad industrial del centro de recepción durante quince años, se estima la función de regresión representada por $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t + \hat{\beta}_3 X_t^2$, cuyo resultado mostrado anteriormente es el siguiente:

Tabla 1.16. Resultado estadístico del modelo de regresión lineal (EViews).

Dependent Variable: Y				
Method: Least Squares				
Sample: 1 15				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	911.3146	494.3730	1.843374	0.0901
X	1129.033	142.1832	7.940693	0.0000
X^2	-76.84665	8.641266	-8.892985	0.0000
R-squared	0.880321	Mean dependent var		3590.923
Adjusted R-squared	0.860375	S.D. dependent var		1485.336
S.E. of regression	555.0178	Akaike info criterion		15.65273
Sum squared resid	3696537.	Schwarz criterion		15.79434
Log likelihood	-114.3955	F-statistic		44.13416
Durbin-Watson stat	1.468850	Prob(F-statistic)		0.000003

Como: $F > F_{1-\alpha}(p-1, n-p)$ y la Prob (F -statistic) o valor p menor que el nivel de significación, entonces se rechaza la hipótesis nula.

Como resultado de aplicar la prueba F de Fisher, la función de regresión lineal múltiple $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$ es de buen ajuste y significativa para un nivel de significación de 0,05.

A partir de los resultados anteriores, se realiza la prueba general de heterocedasticidad de White, obteniéndose el siguiente resultado:

Tabla 1.17. Prueba de la heterocedasticidad de White (EViews).

White Heteroskedasticity Test:				
F-statistic	$F = 0.291986$	Probability	0.830370	
Obs*R-squared	$nR^2 = 1.106384$	Probability	Valor $p = 0.775533$	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1 15				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	473141.4	487101.8	0.971340	0.3523
X	-40705.59	219812.0	-0.185184	0.8565
X^2	-562.1388	23055.84	-0.024382	0.9810
(X^2)^2	12.23215	51.50159	0.237510	0.8166
R-squared	$R^2 = 0.073759$	Mean dependent var	246435.8	
Adjusted R-squared	-0.178852	S.D. dependent var	365637.3	
S.E. of regression	396990.5	Akaike info criterion	28.84439	
Sum squared resid	1.73E+12	Schwarz criterion	29.03320	
Log likelihood	-212.3329	F-statistic	0.291986	
Durbin-Watson stat	2.506367	Prob(F-statistic)	0.830370	

La ecuación de los errores está dada por la regresión auxiliar asociada a la prueba, siendo: $e_i^2 = \alpha_1 + \alpha_2 X_{i2} + \alpha_3 X_{i3}^2 + \alpha_4 X_{i4}^4 + v_i$

1- Formulación de las hipótesis nula y alternativa:

$H_0: \alpha_2 = \alpha_3 = \alpha_4 = 0$ Existe homocedasticidad (no hay heterocedasticidad).

$H_1:$ Alguna $\alpha_j \neq 0$ Existe heterocedasticidad.

2- Región crítica para el estadístico de prueba nR^2 , dado el nivel de significación α , es

$$n \cdot R^2 \sim \chi_{1-\alpha}^2(gl) = \chi_{1-\alpha}^2(k)$$

$$W = \left\{ n \cdot R^2 \in \mathfrak{R} \text{ con } n \cdot R^2 \geq 0: n \cdot R^2 > \chi_{1-\alpha}^2(k) \right\}$$

k es el número de parámetros de los errores al cuadrado menos una constante.

3- Regla de decisión:

- Rechazar H_0 , si $n \cdot R^2 > \chi_{1-\alpha}^2(k)$.
- No rechazar H_0 , si $n \cdot R^2 \leq \chi_{1-\alpha}^2(k)$.

3.1- Cálculo del estadístico de prueba $n \cdot R^2$.

$$n \cdot R^2 = 15 \cdot 0,073759 = 1,10638$$

Información del reporte de salida del EViews: Obs*R-squared = 1,10638.

3.2- Determinando el valor crítico o teórico $\chi_{1-\alpha}^2(gl) = \chi_{1-\alpha}^2(k)$

Existen cuatro parámetros y una constante, por tanto: $p - 1 = 4 - 1 = 3$.

Grados de libertad igual a 3.

$$\text{Entonces: } \chi_{(1-\alpha)}^2 gl = \chi_{(1-\alpha)}^2 k = \chi_{(0,95)}^2 3 = 7,81.$$

Como:

- $n \cdot R^2 < \chi_{(1-\alpha)}^2 gl = \chi_{(1-\alpha)}^2 k$, o sea, $1,10638 < 7,81$, no se rechaza H_0 .
- El valor p de $n \cdot R^2 > \alpha$, o sea, $0,775 > 0,05$, no se rechaza H_0 .

Corroborando el resultado anterior, mediante la aplicación de la prueba F de Fisher.

1- Formulación de las hipótesis nula y alternativa:

$H_0: \alpha_2 = \alpha_3 = \alpha_4 = 0$ Existe homocedasticidad (no hay heterocedasticidad).

$H_1: \text{Alguna } \alpha_i \neq 0$ Existe heterocedasticidad.

2- Región crítica para el estadígrafo F , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0: F > F_{1-\alpha}(p-1, n-p) \right\}.$$

3- Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha}(p-1, n-p)$.
- No rechazar H_0 , si $F \leq F_{1-\alpha}(p-1, n-p)$.

3.1- Tomar como valor del estadístico de prueba el que proporciona la prueba de White, $F = 0,2919$.

3.2- Determinar el valor crítico o teórico de F :

$$F_{1-\alpha}(p-1, n-p) = F_{1-0,05}(4-1, 15-4) = F_{0,95}(3, 11) = 3,59.$$

Como: $F < F_{1-\alpha}(p-1, n-p)$, o sea, $0,2919 < 3,59$, no se rechaza la hipótesis nula.

La aplicación de la prueba F de Fisher complementa los resultados obtenidos con el estadístico de prueba $n.R^2$. Al no rechazar la hipótesis nula, no existe heterocedasticidad, por tanto, se cumple con el supuesto de homocedasticidad.

Verificación del supuesto de normalidad. Prueba Jarque-Bera

Prueba de Jarque-Bera para verificar el supuesto de normalidad de la variable dependiente (regresada o respuesta) en el análisis de regresión.

1- Formulación de las hipótesis nula y alternativa:

$$H_0: Y_i \sim N.$$

$H_1: Y_i$ no sigue una distribución normal.

2- Región crítica para el estadístico de prueba JB , dado el nivel de significación α , es

$$W = \left\{ JB \in \mathfrak{R} \text{ con } JB \geq 0: JB > \chi_{1-\alpha}^2(2) \right\}.$$

3- Regla de decisión:

- Rechazar H_0 , si $JB > \chi_{1-\alpha}^2(2)$.
- No rechazar H_0 , si $JB \leq \chi_{1-\alpha}^2(2)$.

3.1- Cálculo del estadístico de prueba JB :

Determinar primero el coeficiente de asimetría y el valor de la curtosis, pues después el estadístico de prueba JB depende de ambos.

En las expresiones de la asimetría y la curtosis, $\hat{\sigma} = S$ representa el estimador sesgado de la desviación estándar poblacional, y se calcula por:

$$S = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}} \quad S = \sqrt{\frac{30887128,23}{15}} = 1434,971$$

Tabular la información correspondiente para la determinación de la asimetría y la curtosis.

Tabla 1.18. Cálculo de $\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^3$ y $\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^4$ para determinar la asimetría y la

curtosis.

y_i	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$\frac{(y_i - \bar{y})}{\hat{\sigma}}$	$\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^3$	$\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^4$
2729,43	-861,493	742170,18	-0,6003	-0,216384	0,12990757
2804,27	-786,653	618822,94	-0,5482	-0,164748	0,09031507
3458,96	-131,963	17414,23	-0,0919	-0,000777	0,00007152
3069,45	-521,473	271934,08	-0,3634	-0,047991	0,01744033
4205,78	614,857	378049,13	0,4284	0,078667	0,03370729
5207,66	1616,737	2613838,52	1,1266	1,430173	1,61133174
5221,93	1631,007	2660183,83	1,1366	1,468378	1,66897854
5165,48	1574,557	2479229,74	1,0972	1,321130	1,44964279
4764,75	1173,827	1377869,82	0,8180	0,547372	0,44775849
5172,02	1581,097	2499867,72	1,1018	1,337661	1,47387792
4352,56	761,637	580090,91	0,5307	0,149525	0,07936327
3646,38	55,457	3075,47	0,0386	0,000057	0,00000223
2401,87	-1189,053	1413847,03	-0,8286	-0,568950	0,47144638
789,36	-2801,557	7848721,62	-1,9523	-7,441645	14,5286514
873,93	-2716,986	7382012,92	-1,8934	-6,787858	12,8521876
53863,843	//////	30887128,23	//////	-8,895381	34,85468231

Fuente: elaborada por los autores.

A continuación, se exhibe el comportamiento de la variable dependiente capacidad industrial del centro de recepción durante 15 años.

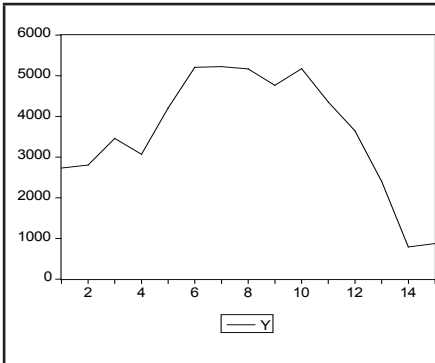


Figura 1.5. Gráfico de línea que representa el comportamiento de la variable dependiente (EViews).

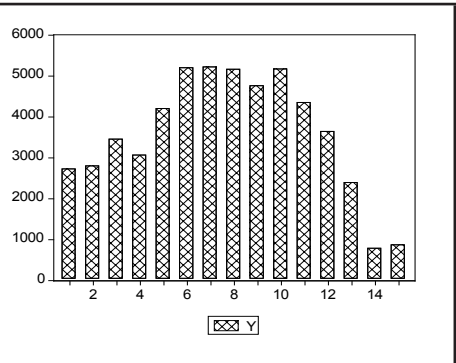


Figura 1.6. Gráfico de barra que representa el comportamiento de la variable dependiente (EViews).

$$A = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^3$$

$$A = \frac{1}{15} (-8,895389751) = -0,593026$$

$$K = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^4$$

$$K = \frac{1}{15} (34,85468231) = 2,323645$$

Cálculo del estadístico de prueba JB :

Sustituyendo los valores de la asimetría y de la curtosis en la expresión del estadístico de JB .

$$JB = n \left[\frac{A^2}{6} + \frac{(K - 3)^2}{24} \right]$$

$$JB = 15 \left[\frac{(-0,593026)^2}{6} + \frac{(2,323645 - 3)^2}{24} \right] = 15(0,058613 + 0,019061) = 1,165105$$

Los resultados obtenidos de la asimetría, la curtosis y del estadístico Jarque-Bera se corroboran con los resultados de la prueba Jarque-Bera para la variable regresada, obtenidos a través del programa EViews, que se muestra en la siguiente figura:

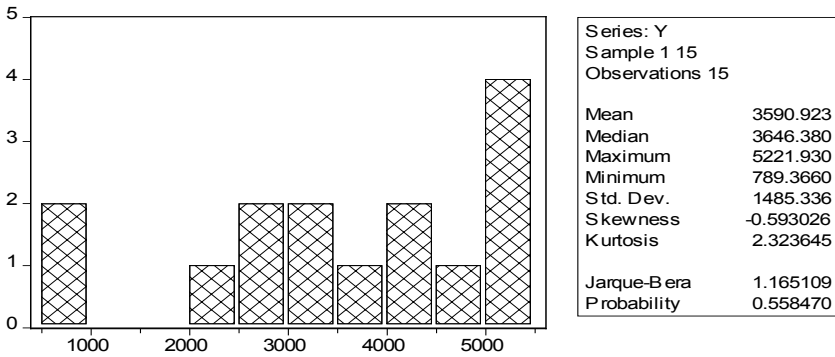


Figura 1.7. Resultado de la prueba Jarque-Bera para la variable regresada. Histograma y estadística de la prueba de normalidad (EViews).

El histograma representado es el método gráfico más utilizado para verificar el supuesto de normalidad.

3.2- Determinación del valor del percentil $\chi^2_{1-\alpha} (2)$:

Asintóticamente el estadístico Jarque-Bera sigue una distribución ji-cuadrado, con dos grados de libertad.

$$\chi^2_{1-\alpha} (2) = \chi^2_{1-0,05} (2) = \chi^2_{0,95} (2) = 5,99$$

Como: $JB \leq \chi^2_{1-\alpha} (2)$, o sea, $1,165 < 5,99$, no se rechaza la hipótesis nula.

Alternativamente:

- Rechazar H_0 , si el valor de la probabilidad o valor $p < \alpha$.
- No rechazar H_0 , si el valor de la probabilidad o valor $p > \alpha$.

Como el valor de la probabilidad o valor p es mayor que el nivel de significación, o sea, $0,558 > 0,05$, no se rechaza la hipótesis nula.

Toma de decisión:

- Como el estadístico JB igual a 1,165 es menor que el valor crítico de $\chi^2_{1-\alpha} (2)$ igual a 5,99; no existen elementos para rechazar la hipótesis nula, lo que implica que la variable regresada capacidad industrial del centro de recepción sigue una distribución normal, para un nivel de significación del 0,05.

- Alternativamente, se comprueba que la probabilidad asociada al estadístico JB o valor p de 0,558 es superior al nivel de significación del 0,05. Se arriba a la misma conclusión sobre la existencia de normalidad.
- Como el coeficiente de asimetría igual a -0,593 es menor que cero, la serie de observaciones es asimétrica negativa, presentando la distribución normal deformación a la izquierda.
- Al ser la curtosis igual a 2,323 es menor que tres; la distribución es menos apuntada que la normal, es más achatada, y se denomina platicúrtica. Existe una baja concentración de los valores de la distribución en la región central.

Prueba de Jarque-Bera para verificar el supuesto de normalidad de los residuos en el análisis de regresión.

1- Formulación de las hipótesis nula y alternativa:

$$H_0: U_i \sim N.$$

$$H_1: U_i \text{ no sigue una distribución normal.}$$

2- Región crítica para el estadístico de prueba JB , dado el nivel de significación α , es

$$W = \left\{ JB \in \mathfrak{R} \text{ con } JB \geq 0: JB > \chi_{1-\alpha}^2 (2) \right\}.$$

3- Regla de decisión:

$$\text{- Rechazar } H_0, \text{ si } JB > \chi_{1-\alpha}^2 (2).$$

$$\text{- No rechazar } H_0, \text{ si } JB \leq \chi_{1-\alpha}^2 (2).$$

3.1- Cálculo del estadístico de prueba JB :

Primero se determina el coeficiente de asimetría y el valor de la curtosis, y se obtiene que $A = -0,727684$ y $K = 3,054611$.

Luego se sustituyen los valores de la asimetría y de la curtosis en la expresión del estadístico de JB .

$$JB = n \left[\frac{A^2}{6} + \frac{(K-3)^2}{24} \right]$$

$$JB = 15 \left[\frac{(-0,727684)^2}{6} + \frac{(3,054611 - 3)^2}{24} \right]$$

$$JB = 15(0,088254 + 0,000124) = 1,325670$$

Los resultados obtenidos del estadístico Jarque-Bera se corroboran con los resultados de la prueba Jarque-Bera para los residuos, obtenidos a través del programa EViews, que se muestra en la siguiente figura:

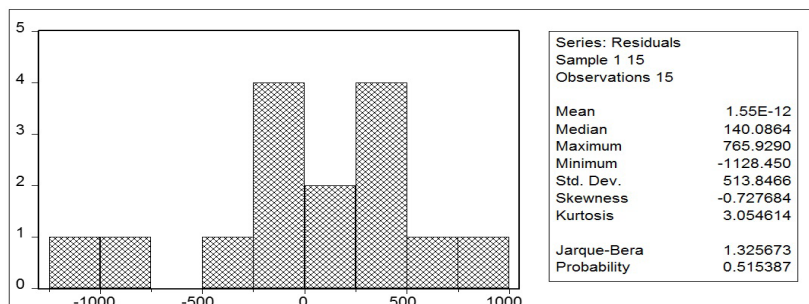


Figura 1.8. Resultado de la prueba Jarque-Bera para los residuos. Histograma y estadística de la prueba de normalidad (EViews).

El histograma de los residuos es un método gráfico usado para verificar la normalidad del error de estimación. A partir de la gráfica del histograma, se puede representar la curva de densidad de la distribución normal.

3.2- Determinar el valor del percentil $\chi^2_{1-\alpha} (2)$:

Asintóticamente el estadístico Jarque-Bera sigue una distribución ji-cuadrado, con dos grados de libertad.

$$\chi^2_{1-\alpha} (2) = \chi^2_{1-0,05} (2) = \chi^2_{0,95} (2) = 5,99$$

Como: $JB \leq \chi^2_{1-\alpha} (2)$, o sea, $1,325 < 5,99$, no se rechaza la hipótesis nula.

Alternativamente:

- Rechazar H_0 , si el valor de la probabilidad o valor $p < \alpha$.
- No rechazar H_0 , si el valor de la probabilidad o valor $p > \alpha$.

Como el valor de la probabilidad o valor p es mayor que el nivel de significación,

o sea, $0,515 > 0,05$, no se rechaza la hipótesis nula.

Toma de decisión:

- Como el estadístico JB igual a 1,325 es menor al valor crítico de $\chi^2_{1-\alpha}$ (2) igual a 5,99; no existen elementos para rechazar la hipótesis nula, lo que implica que los residuos siguen una distribución normal, para un nivel de significación del 0,05.
- Alternativamente, se comprueba que la probabilidad asociada al estadístico JB de 0,515 es superior al nivel de significación del 0,05, arribándose a la misma conclusión sobre la existencia de normalidad.
- Como el coeficiente de asimetría igual a -0,727, es menor que cero, la serie de observaciones es asimétrica negativa, presentando la distribución normal deformación a la izquierda.
- La curtosis es igual a 3,05, mayor que tres; siendo la distribución más apuntada que la normal, y se denomina leptocúrtica. Existe una alta concentración de los valores de la distribución en la región central.

Evaluando el supuesto de la normalidad mediante la Prueba de Shapiro-Wilk

Para comprobar si las observaciones de la capacidad industrial y el tiempo en años, siguen una distribución normal, existen dos pruebas que también es posible aplicar: la prueba *Kolmogorov-Smirnov*, para cuando el tamaño de la muestra es mayor que 50, y la prueba *Shapiro-Wilk*, para cuando el tamaño de la muestra es menor que 50; las mismas se encuentran disponibles en el paquete de programa estadístico SPSS.

Considerando la formulación de las hipótesis nula que la capacidad industrial sigue una distribución normal, y como hipótesis alternativa a contrastar la capacidad industrial no sigue una distribución normal, es decir difieren de la distribución normal, como el número de observaciones es menor que 50, se considera el resultado de la prueba de Shapiro-Wilk, expuesto en la tabla 1.19, donde el valor p del estadístico Shapiro-Wilk es 0,099, superior al nivel de significación de 0,05 o margen de error que se corre al momento de rechazar la hipótesis nula, siendo cierta, por tanto no se rechaza la hipótesis de normalidad, y se concluye que la variable capacidad industrial se ajusta a una distribución normal.

De igual manera, considerando las observaciones de la variable tiempo, se

toma el resultado de la prueba de Shapiro-Wilk, exhibido en la tabla 1.19, donde el valor p del estadístico Shapiro-Wilk es 0,755, superior al nivel de significación de 0,05 o margen de error que se corre al momento de rechazar la hipótesis nula, siendo cierta. Por tanto, no se rechaza la hipótesis de normalidad, concluyendo que la variable el tiempo no difiere de la distribución normal.

Tabla 1.19. Resultado de la prueba de normalidad Kolmogorov-Smirnov y Shapiro-Wilk (SPSS).

Pruebas de normalidad						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Capacidad industrial	,136	15	,200 [*]	,901	15	,099
Tiempo	,082	15	,200 [*]	,964	15	,755

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Verificación del supuesto de no multicolinealidad

Para los modelos de regresión polinomiales $Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_t^2 + U_t$ o $Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_t^2 + \beta_4 X_t^3 + U_t$, las diferentes X no están altamente correlacionadas ya que todas son potencias de X , pero términos como X^2 , X^3 , etc. son todas funciones no lineales de X y por consiguiente, de manera estricta, no violan el supuesto de no multicolinealidad (Gujarati, 2003).

La existencia de no autocorrelación, de perturbaciones homoscedásticas, de una distribución normal y la no violación del supuesto de no multicolinealidad, en el caso mostrado, corrobora la validez de la prueba estadística F de Fisher, así como la validez y eficiencia del valor estimado de la capacidad industrial máxima del centro de recepción y de cualquier proyección que se realice con la función de regresión estimada.

1.6.6. Análisis de las propiedades de la función de regresión estimada

Análisis de las propiedades de monotonía, existencia de extremos locales, concavidad y existencia de punto de inflexión de la función de regresión múltiple estimadas, $\hat{E}_{(Y_t / X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$.

Análisis de la monotonía

1- Se calcula la primera derivada de la función.

$$\hat{E}_{(Y_t / X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$$

$$\frac{d\hat{Y}}{dX} = \hat{Y}'_t = 1129,033 - 153,692X_t$$

2- Se determinan los ceros de la primera derivada:

Igualando el resultado de la primera derivada a cero y hallar el punto crítico

$$\frac{d\hat{Y}}{dX} = \hat{Y}'_t = 0$$

$$1129,033 - 153,692X_t = 0$$

$$X_t = 7,35$$

3- Se determinan los puntos que indefinen la primera derivada:

No hay puntos que indefinen la primera derivada, porque $\frac{dY}{dX} = \hat{Y}'_t = 1129,033 - 153,692X_t$, está definida en los números reales positivos; luego, el único punto crítico es 7,35.

4- Se determina el signo de la primera derivada en los intervalos definidos por el punto crítico:

- Para los valores de $X < 7,35$, en el dominio de la función dado por el conjunto numérico de los números reales positivos, la primera derivada

$$\frac{d\hat{Y}}{d\hat{X}} = \hat{Y}'_t = 1129,033 - 153,692X_t \text{ toma valores mayores que cero.}$$

- Para los valores de $X > 7,35$, en el dominio de la función dado por el conjunto numérico de los números reales positivos, la primera derivada

$$\frac{d\hat{Y}}{dX} = \hat{Y}'_t = 1129,033 - 153,692X_t \text{ toma valores menores que cero.}$$

5- Se interpreta el signo de la derivada y se concluye sobre la existencia de la monotonía:

- Como la primera derivada es mayor que cero, $\frac{d\hat{Y}}{dX} = \hat{Y}'_t > 0$, para los valores de $X < 7,35$ en el dominio de la función, la función de regresión estimada

$$\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2 \text{ es creciente.}$$

- Como la primera derivada es menor que cero, $\frac{d\hat{Y}}{dX} = \hat{Y}'_t < 0$, para los valores de $X > 7,35$ en el dominio de la función, la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$ es decreciente.

Análisis de la existencia de extremo

Como la primera derivada, $\frac{dY}{dX} = \hat{Y}'_t$, cambia de signo de positivo a negativo en el punto crítico $X=7,35$, y $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$ es continua en $X = 7,35$, por tanto la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$ tiene un máximo local, absoluto u óptimo en $X = 7,35$.

El valor máximo local, absoluto u óptimo de la función de regresión estimada es 5058,29, y se calcula como:

$$\hat{Y}_t = 911,314 + 1129,033X - 76,846X^2$$

$$\hat{Y}_t = 911,314 + 1129,033(7,35) - 76,846(7,35)^2$$

$$\hat{Y}_t = 5058,29$$

El punto máximo local y absoluto del gráfico de la función es (7,35; 5028,29).

Análisis de la concavidad

1- Se calcula la segunda derivada de la función o la derivada de la primera derivada:

$$\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$$

$$\frac{d\hat{Y}}{dX} = \hat{Y}'_t = 1129,033 - 153,692X_t$$

$$\frac{d^2\hat{Y}}{dX^2} = \hat{Y}'' = -153,692$$

2- Se interpreta el signo de la segunda derivada y se concluye sobre la existencia de la concavidad:

Como la segunda derivada es menor que cero, $\frac{d^2\hat{Y}}{dX^2} = \hat{Y}'' < 0$, para los valores de $X > 0$ en el dominio de la función, dado por el conjunto numérico de los números reales positivos el gráfico de la función, $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$, es cóncava hacia abajo en todo su dominio.

Punto de inflexión

El gráfico de la función no tiene punto de inflexión para ningún valor de X en su dominio, porque la segunda derivada no cambia de signo para ningún valor de X en su dominio, es decir no existe punto en el gráfico de la función donde cambie de dirección la concavidad.

A partir de los resultados obtenidos del análisis de la monotonía, existencia de extremos locales, concavidad y existencia de puntos de inflexión de la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 911,314 + 1129,033X_t - 76,846X_t^2$, se puede realizar el trazado del esbozo gráfico de la función. No obstante, el lector puede consultar la representación gráfica de la función en la figura 1.4 obtenido del Eviews.

1.7. Modelización económico-matemática de la función de costo. Aplicaciones

1.7.1. Determinación y validación de la función de costo

En el transporte de la caña, la producción del trabajo se mide en toneladas y toneladas-kilómetro; donde la magnitud de los gastos referidos a la producción del trabajo, representa el costo de transportación; por tanto, de forma general, el costo de transportación se considera como la cantidad de gastos que se emplean en transportar una tonelada de caña a la distancia de un kilómetro.

Para calcular el costo de los medios de transporte en la agroindustria azucarera, en función de los modelos económico-matemáticos de programación lineal, se tienen en cuenta las características que se exponen.

El proceso de transportación de carga está formado por los elementos fundamentales que determinan los gastos de explotación, para la transportación y traslado de la carga y descarga. En el caso particular de la transportación de la caña, se consideran los costos relacionados con su traslado desde los campos a los centros de recepción industrial, donde influyen los costos variables y fijos.

Los costos variables son todos aquellos gastos relacionados con el traslado de la caña, durante el movimiento del medio de transporte; se determinan sobre la base de los kilómetros recorridos. Los costos fijos son los gastos relacionados con el tiempo de existencia del medio de transporte y, generalmente, no dependen de los kilómetros recorridos.

Para el cálculo de los costos totales de transportación se han empleado numerosas expresiones matemáticas, que van desde las representaciones clásicas, hasta las más variadas funciones de costo en función de parámetros técnico-económicos, tales como: aprovechamiento de las capacidades de carga y del recorrido, tiempo de trabajo, velocidad técnica, número de viajes, entre otros.

Todas las expresiones tienen como punto de partida los supuestos básicos de las funciones clásicas, y se pueden plantear como sigue:

$$\text{Costo por tonelada (\$/t)} = \frac{\text{Costos totales}}{\text{Carga transportada}}$$

$$\text{Costo por tonelada kilómetro (\$/t \cdot km)} = \frac{\text{Costos totales}}{\text{Tráfico realizado}}$$

Los costos totales (en unidades monetarias), la carga transportada (en toneladas) y el tráfico realizado (en t.km), tienen que estar calculados para el mismo período de tiempo.

En los modelos clásicos de optimización de transporte, en general, y en los modelos económico-matemáticos destinados a la transportación de la caña de azúcar, desarrollados en el presente capítulo, se emplean como criterio de optimización para la toma de decisiones económicas el costo de transportación, es decir, los costos de transportación por tonelada de caña de azúcar a transportar desde los campos hasta los centros de recepción

industrial por el tipo de transporte considerado, constituyen los coeficientes económicos de las funciones objetivos de los modelos económico-matemáticos planteados.

Como limitante para la implementación práctica de los modelos, se presenta que los costos de los medios de transportación no se pueden obtener directamente en la generalidad de los casos, no se pueden obtener con grado de detalle por tipo de transporte, o existen de una forma global o general, lo que distorsiona el resultado esperado. Dada esta problemática, donde la información no siempre se encuentra completa, ni en la forma requerida, se presenta una metodología para determinar la función del costo, a partir de la cual se calcula el costo de transportación por tonelada-kilómetro para los medios de transporte.

Teniendo en cuenta las características de los costos y de los demás indicadores de los medios de transporte, para la determinación del costo de transportación de manera general se puede plantear que la función racional entera o polinómica de grado uno de tipo lineal, representa adecuadamente la relación funcional del costo, y se representa como $f(x) = m x + n$, donde m y n son constantes. Siendo m la pendiente de la función y expresa la tasa de variación de y con respecto a x , si m es mayor que cero, la función es creciente, y aumenta en m unidades por cada unidad que aumente x .

Para calcular los costos en que se incurre en el traslado de la caña por los medios de transporte, de forma análoga a lo expresado anteriormente, se considera la siguiente función lineal del costo:

$$Y = \alpha' + \beta X$$

donde:

Y es el costo de una unidad de carga a una distancia determinada (unidades monetarias).

X es la distancia a que se transporta la unidad de carga (kilómetros).

α' es el coeficiente de gasto fijo por unidad de carga transportada.

β es el coeficiente de gasto variable por unidad de carga-kilómetro.

El costo de una unidad de carga a una distancia dada (Y), y la distancia

a que se transporta la unidad de carga (X), representan las variables dependiente e independiente, respectivamente, por lo que el problema a resolver consiste en determinar los valores constantes de los *coeficientes de gastos fijos* (α') y los *coeficientes de gastos variables* (β). A continuación, se presenta el procedimiento desarrollado para la determinación de los coeficientes y la validación de los resultados.

Determinación de los coeficientes de gastos

Para facilitar la determinación de los coeficientes de los gastos fijos y variables, se tabula el procedimiento de cálculo en la tabla 1.20; donde, además, se refleja la dimensión de cada elemento.

Tabla 1.20. Coeficientes de gastos.

Zafras	Volumen de caña transportada (t)	Distancia media de una unidad con carga (km)	Gastos Variables (\$)	Gastos Fijos (\$)	Tráfico de carga realizado (t.km)	Coeficiente de Gasto Variable (\$/t.km)	Coeficiente de Gasto Fijo (\$/t)
.
.
.
Total						////	////
Promedio						Promedio (β)	Promedio (α')

Fuente: elaborada por los autores.

Para conformar la tabla se procede como sigue:

- Se tomarán al menos cinco zafras azucareras anteriores al período corriente que se planifica.
- El volumen de caña transportada expresa la carga realmente transportada desde los campos cañeros a los centros de recepción industrial en las zafras del período que se toma como referencia.
- La distancia media de una unidad de carga refleja la distancia promedio a recorrer por cada unidad de carga.

Se calcula dividiendo el tráfico producido por el número de unidades de carga transportadas durante el período de referencia. Si este indicador no se controla directamente o no se puede determinar por desconocer el tráfico producido, se puede calcular la distancia media ponderada de transportación de carga.

- Para los gastos variables, se requiere una definición inicial de los elementos de gastos lo más amplia posible, asegurando que el medio de transporte en estudio refleje estos gastos con bastante precisión, al imputársele el mayor número de conceptos de gastos asociados con la producción del equipo.
- El elemento de gasto fijo a tener en cuenta es la amortización.
- El tráfico de carga realizado expresa la magnitud del trabajo en el transporte de la caña de azúcar, y combina la carga y la distancia a que se transporta la caña. Este indicador se obtiene de multiplicar el volumen de caña transportada en cada zafra por la distancia media de una unidad de carga.
- El coeficiente de gasto variable se determina dividiendo los gastos variables por el tráfico de carga realizado.
- El coeficiente de gasto fijo se determina dividiendo los gastos fijos por el volumen de caña transportada.

Determinación de los coeficientes de gastos promedios:

- El coeficiente de gasto variable promedio, se calcula dividiendo el gasto total promedio variable, por el tráfico de carga total promedio realizado.
- El coeficiente de gasto fijo promedio, se obtiene dividiendo el gasto total promedio fijo, por el volumen de caña total promedio transportada.

De esta forma, se determinan los valores constantes de α' y β , de la función lineal del costo $Y = \alpha' + \beta X$. Evaluando la función lineal del costo para un valor dado de la distancia existente entre un campo cañero y un centro de recepción industrial, se obtiene el costo de transportación de una unidad de carga a la distancia dada.

La función lineal del costo, $Y = \alpha' + \beta X$, para la situación objeto de

estudio, está definida solo para los números reales positivos, o sea, el

$\text{Dom } f = \{x \in \mathfrak{R} : x > 0\}$, y la función toma solamente valores positivos,

es decir, la a $\text{Im } f = \{y \in \mathfrak{R} : y > 0\}$, representando la constante β la tasa de variación del costo de transportación con respecto a la distancia en kilómetros. Si β es mayor que cero, la función es creciente, el costo de transportación Y aumenta en β unidades monetarias, por cada kilómetro que aumente X .

Los procedimientos expuestos para determinar la función lineal del costo, y calcular por consiguiente, el costo de transportación por tonelada-kilómetro, se utilizan en la determinación de los coeficientes económicos, costo de transportación, C_{ijk}^{lp} y C_{ijk}^{lpq} de los modelos económico-matemáticos.

Análisis estadístico matemático

Para determinar el grado de validez y confiabilidad del costo total de transportación, calculado a partir de la función de costo lineal $Y = \alpha' + \beta X$ con relación al costo total de transportación real, se aplicó el análisis estadístico matemático que se expone.

A los resultados de los costos totales de transportación calculados entre los orígenes (campos cañeros) y los destinos (centros recepción) a la distancia dada, obtenidos de la aplicación de la función lineal del costo $Y = \alpha' + \beta X$, y a los resultados de los costos totales de transportación reales, se les aplica la prueba de hipótesis entre medias de dos poblaciones, con distribución normales y varianzas desconocidas, pero iguales, para comparar las medias correspondientes a ambos costos, con el propósito de verificar la no existencia de diferencia significativa entre el comportamiento de los valores promedios de las variables costo total de transportación real y el determinado de aplicar la función de costo lineal; la utilización de la prueba de hipótesis se apoya en el paquete de programa estadístico SPSS.

En la validación de la prueba de hipótesis entre las medias, se utiliza el estadístico de prueba t que se distribuye de acuerdo a una distribución probabilística t de Student con $n_x + n_y - 2$ grados de libertad, para muestras menores o iguales que treinta; las muestras proceden de poblaciones con

distribución normales y con varianzas poblacionales desconocidas e iguales.

En esta prueba de hipótesis para comparar medias, se plantea como hipótesis nula a someter a prueba $H_0: \mu_x = \mu_y$, versus la hipótesis alternativa a contrastar $H_1: \mu_x \neq \mu_y$, donde, μ_j es la media poblacional, $j = x, y$.

Con el planteamiento de la hipótesis anterior, se pretende comprobar de la población, si entre los valores promedios del costo calculado y el real, existe o no diferencias significativas; en tal sentido, se obtiene de dicha población una muestra que permite verificar la hipótesis planteada y probar que no existe diferencia significativa para ambos costos.

Para la prueba de significación, se utiliza el estadístico de prueba o estadígrafo t .

$$t = \frac{\bar{x} - \bar{y}}{S_o \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}}, \text{ donde } S_o = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$$

También se puede usar la expresión:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2} \right) \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$$

El estadígrafo t calculado, obtenido u observado, sigue una distribución t de Student con $n_x + n_y - 2$ grados de libertad.

donde:

\bar{X} es la media de la muestra obtenida a partir de los costos reales.

\bar{Y} es la media de la muestra de los costos determinada a partir de la función lineal del costo.

n_x, n_y son los tamaños de las muestras.

S_x^2, S_y^2 son las varianzas de las muestras.

Siendo S_x^2 y S_y^2 las varianzas de las dos muestras simples aleatorias

independientes de las poblaciones normales, con medias μ_x y μ_y , respectivamente, y varianzas iguales a σ^2 .

Tenemos que:

$$S_x^2 = \frac{\sum_{i=1}^{n_x} (X_i - \bar{X})^2}{n_x - 1} \quad \text{y} \quad S_y^2 = \frac{\sum_{i=1}^{n_y} (Y_i - \bar{Y})^2}{n_y - 1}$$

Regla de decisión:

- Si, $t < -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$ o $t > t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$, el valor obtenido del estadígrafo t cae en la región crítica, por tanto, se rechaza la hipótesis nula, aceptándose que existe diferencia significativa entre los valores promedios del costo real y del costo determinado por la función lineal del costo, para el nivel de significación considerado.

- Si, $t \geq -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$ o $t \leq t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$, o sea,

$-t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \leq t \leq t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$, no se rechaza la hipótesis nula, y de cumplirse el requerimiento básico de varianzas desconocidas e iguales, y de normalidad, se considera que no hay diferencia significativa entre los valores promedios de ambos costos, para el nivel de significación considerado.

Si la prueba de significación para la situación analizada tiene validez, esto confirma que las variables correspondientes al costo real y al costo determinado a partir de la función lineal del costo se distribuyen según t de Student, no existiendo diferencia significativa entre los valores promedios de los costos, para un nivel de significación fijado. Por tanto, la función matemática que mejor se ajusta al costo de transportación es la función lineal considerada $Y = \alpha' + \beta X$.

Resumiendo la prueba de hipótesis.

Formulación de la hipótesis nula y alternativa:

$$H_0: \mu_x = \mu_y$$

$$H_1: \mu_x \neq \mu_y$$

El estadístico de prueba o estadígrafo a utilizar es t :

$$t = \frac{\bar{x} - \bar{y}}{S_o \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \quad S_o = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$$

También se puede usar la expresión:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2} \right) \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$$

donde:

$$S_x^2 = \frac{\sum_{i=1}^{n_x} (X_i - \bar{X})^2}{n_x - 1} \quad y \quad S_y^2 = \frac{\sum_{i=1}^{n_y} (Y_i - \bar{Y})^2}{n_y - 1}$$

Región crítica para el estadígrafo t , dado el nivel de significación α , es

$$W = \left\{ t \in \mathfrak{R} : t < -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \quad o \quad t > t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \right\}.$$

donde:

$t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$ es el percentil de orden $1 - \frac{\alpha}{2}$ de la distribución t de Student con $(n_x + n_y - 2)$ grados de libertad.

Regla de decisión:

- Rechaza H_0 , si $t < -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$ o $t > t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$.

- No rechazar H_0 , si $t \geq -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$ o $t \leq t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$

o sea, $-t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \leq t \leq t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$.

Alternativamente:

- Rechazar H_0 , si el valor p asociado al t -statistic $< \alpha$.
- No rechazar H_0 , si el valor p asociado al t -statistic $> \alpha$.

Para la prueba de hipótesis entre media de dos poblaciones con distribuciones normales y varianzas desconocidas, pero iguales, donde se realiza la prueba de t de Student para dos muestras independiente, hay que corroborar los supuestos de normalidad y de igualdad de varianzas.

Prueba de normalidad

La prueba de normalidad se emplea para corroborar si la variable numérica aleatoria correspondiente al costo real y al costo calculado se distribuye normalmente. Puede utilizarse, en dependencia del tamaño de la muestra, la prueba Komogorov-Smirnov, o la prueba Shapiro-Wilk. Dichas pruebas se realizan mediante la aplicación del paquete de programas estadístico SPSS.

Prueba para comprobar el supuesto de varianzas poblacionales iguales

Para comprobar el supuesto de varianzas poblacionales iguales, se verificó a partir de una prueba de hipótesis de comparación de varianzas, prueba concerniente a las varianzas de dos poblaciones con distribuciones normales, como se muestra a continuación, así como, la prueba de Levene's para dos muestras independientes, mediante la aplicación del paquete de programas estadísticos SPSS.

Prueba de hipótesis para la comparación de varianzas mediante la prueba relativa a las varianzas de dos poblaciones con distribuciones normales.

Se plantea una prueba de hipótesis bilateral o de dos colas, alternativa bilateral, donde se formula como hipótesis nula, que se somete a prueba, $H_0: \sigma_x^2 = \sigma_y^2$, versus la hipótesis alternativa a contrastar $H_1: \sigma_x^2 \neq \sigma_y^2$.

Para la prueba de significación es utilizado el estadístico de prueba F de Fisher, resultado del cociente entre dos varianzas.

$$F = \frac{S_x^2}{S_y^2}$$

Regla de decisión:

- Si, $F < F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1)$ o $F > F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$, como el valor obtenido del estadígrafo F cae en la región crítica, por consiguiente, se rechaza la hipótesis nula, aceptándose que existe diferencia significativa entre las varianzas de ambos costos, para el nivel de significación fijado.

- Si, $F \geq F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1)$ o $F \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$, o sea

$$F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) \leq F \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1),$$

no se rechaza la hipótesis nula; se puede considerar que las varianzas poblacionales de ambos costos no se diferencian significativamente, para el nivel de significación fijado.

Si al aplicar esta prueba de hipótesis a la situación analizada, no se rechaza la hipótesis nula, es factible considerar que no existe diferencia significativa entre las varianzas para el nivel de significación considerado.

Resumiendo la prueba de hipótesis.

Formulación de la hipótesis nula y alternativa:

$$H_0: \sigma_x^2 = \sigma_y^2$$

$$H_1: \sigma_x^2 \neq \sigma_y^2$$

El estadístico de prueba o estadígrafo a utilizar es F de Fisher: $F = \frac{S_x^2}{S_y^2}$

Región crítica para el estadígrafo F , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} : F < F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) \text{ o } F > F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1) \right\}$$

$$\text{donde: } F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) = \frac{1}{F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)}$$

Siendo $F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1)$ el percentil de orden $\alpha/2$ de la distribución F de Fisher con $n_x - 1$ grados de libertad del numerador y $n_y - 1$ grados de libertad del denominador.

Regla de decisión:

- Rechazar H_0 , si $F < F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1)$ o $F > F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$
- Norechazar H_0 , si $F \geq F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1)$ o $F \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$

$$\text{o sea, } F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) \leq F \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1).$$

Mediante la aplicación de la prueba de hipótesis de Levene's para probar la hipótesis de la igualdad de varianzas, a través del paquete de programas estadísticos SPSS, se plantea la siguiente regla de decisión:

- Rechazar H_0 , si el valor de significancia o valor $p < \alpha$.
- No rechazar H_0 , si el valor de significancia o valor $p > \alpha$.

El procedimiento propuesto constituye una contribución importante ya que proporciona al transportista y al productor la información detallada sobre el cálculo del costo de transportación para la toma de decisiones económicamente fundamentadas. Además, dichos costos se utilizan como coeficientes económicos en el planteamiento y solución de los modelos económico-matemáticos objeto de estudio. La metodología constituye también una vía para comprobar la veracidad del costo calculado con respecto al costo real, mediante la utilización del análisis estadístico matemático.

Además del procedimiento analizado para determinar el costo de transportación a partir de la función lineal del costo, existe la posibilidad de aplicar otro método mediante la construcción de la ficha de costo por tipo de transporte.

1.7.2. Determinación de la validez del costo total de transportación calculado

Para comprobar la igualdad de las medias poblacionales, se seleccionan dos muestras simples aleatorias de 16 observaciones del costo real y del costo calculado, correspondientes a dos poblaciones. Se supone que las varianzas correspondientes a ambas poblaciones son desconocidas, pero iguales, y que las distribuciones de ambas poblaciones son normales.

La información muestral seleccionada se exhibe a continuación.

Tabla 1.21. Costo real y calculado de la muestra seleccionada.

Costo real por tonelada X_i	Costo calculado por tonelada Y_i	Costo real por tonelada X_i	Costo calculado por tonelada Y_i
121,918	110,681	103,901	95,852
117,941	106,868	106,868	96,515
115,312	117,902	115,998	112,025
114,521	115,210	110,828	113,808
109,327	108,791	117,237	116,151
100,915	107,217	119,230	104,767
103,001	96,202	109,208	108,228
104,984	95,747	125,770	112,250

Fuente: elaborada por los autores.

En correspondencia con la planteado anteriormente, para determinar el grado de validez y confiabilidad del costo total de transportación calculado a partir de la función lineal del costo $Y = \alpha' + \beta X$, con relación al costo total de transportación real, se aplica la prueba de hipótesis referentes a las medias de dos poblaciones con distribuciones normales y varianzas desconocidas, pero iguales; así como, la prueba de normalidad y la prueba de hipótesis referentes a las varianzas de dos poblaciones con distribuciones normales.

Para que la prueba de t de Student para dos muestras independientes sea válida, hay que corroborar dos supuestos: el supuesto de normalidad y el supuesto de igualdad de varianzas.

Prueba de normalidad

La prueba de normalidad se emplea para corroborar si las variables aleatorias costo real y costo calculado se distribuyen normalmente.

Formulación de la hipótesis nula y alternativa:

H_0 : Las variables costo real y costo calculado siguen una distribución normal.

H_1 : Las variables costo real y costo calculado no siguen una distribución normal.

Para ello, se pueden utilizar la prueba de Komogorov-Smirnov, cuando el tamaño de la muestra es grande, o sea, mayor a 50 observaciones, o la prueba Shapiro-Wilk cuando el tamaño de la muestra es pequeño, es decir, menor o igual a 50 observaciones.

Los resultados de las pruebas se obtienen mediante la aplicación del SPSS, como se muestra en la siguiente información:

Tabla 1.22. Prueba de normalidad correspondiente al costo real y calculado (SPSS).

Resumen de Procesamiento de Casos									
	Costo real y calculado	Casos							
		Válido		Perdidos		Total			
		N	Porcentaje	N	Porcentaje	N	Porcentaje		
Costos	Costo real	16	100,0%	0	0,0%	16	100,0%		
	Costo calculado	16	100,0%	0	0,0%	16	100,0%		
Descriptivos									
	Costo real y calculado	Estadístico			Error estándar				
Costos	Costo real	Media			112,3099			1,82126	
		95% de intervalo de confianza para la media	Límite inferior		108,4280				
			Límite superior		116,1919				
		Media recortada al 5%			112,1952				
		Mediana			112,6745				
		Varianza			53,072				
		Desviación estándar			7,28503				
		Mínimo			100,92				
		Máximo			125,77				
		Rango			24,85				
Rango intercuartil			12,31						
Asimetría			,108			,564			
Curtosis			-,920			1,091			

Costo calculado	Media	107,3884	1,90119
	95% de intervalo de confianza para la media	Límite inferior	103,3361
		Límite superior	111,4407
	Media recortada al 5%	107,4511	
	Mediana	108,5095	
	Varianza	57,832	
	Desviación estándar	7,60475	
	Mínimo	95,75	
	Máximo	117,90	
	Rango	22,16	
Rango intercuartil	14,84		
Asimetría	-,517	,564	
Curtosis	-1,014	1,091	

Pruebas de Normalidad

	Costo Real y Calculado	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Costos	Costo Real	,119	16	,200*	,969	16	,828
	Costo Calculado	,174	16	,200*	,894	16	,063

*. Esto es un límite inferior de la significación verdadera. a. Corrección de significación de Lilliefors

Como el porcentaje correspondiente a ambos costos es del 100 %, entonces el 100 % de los casos son válidos y no existen valores perdidos.

La media del costo real es de \$ 112,31 por tonelada y la media del costo calculado es de \$ 107,38 por tonelada. Numéricamente, el valor del costo real es mayor que el costo calculado, hay una diferencia de \$ 4,93, pero se demostrará con la aplicación de la prueba *t* de Student que la diferencia no es significativa.

Con la prueba de normalidad se corrobora si la variable costo se comporta normalmente en el grupo del costo real y calculado. El SPSS muestra las pruebas de normalidad Komogorov-Smirnov y Shapiro-Wilk. Como el tamaño de la muestra es menor que 50, se contrasta la normalidad con la prueba Shapiro-Wilk, se aprecia que el valor de significancia o valor *p* de la variable costo real es de 0,828 y el grado de significancia o valor *p* de la variable costo calculado es de 0,063.

Toma de decisión:

- Como el valor p de la prueba de normalidad, correspondiente a la variable costo real es de 0,828, excede el nivel de significación de 0,05, por tanto, no se rechaza la hipótesis nula, y se puede afirmar que los datos provienen de una distribución normal.
- Como el valor p de la prueba de normalidad, correspondiente a la variable costo calculado es de 0,063, excede el nivel de significación de 0,05, por tanto, no se rechaza la hipótesis nula, y se puede afirmar que los datos provienen de una distribución normal.

Se concluye que, al ser el valor p de la prueba de normalidad mayor que el nivel de significación para las variables de ambos costos y no rechazarse la hipótesis nula, se infiere que los datos provienen de una distribución normal y que la variable costo de ambos grupos se comporta normalmente. Con estos resultados se verifica el supuesto de normalidad.

Prueba de hipótesis referentes a las varianzas de dos poblaciones con distribuciones normales

1- Formulación de la hipótesis nula y alternativa:

$$H_0: \sigma_x^2 = \sigma_y^2$$

$$H_1: \sigma_x^2 \neq \sigma_y^2$$

2- Región crítica para el estadígrafo F , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} : F < F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) \text{ o } F > F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1) \right\}$$

3- Regla de decisión:

- Rechazar H_0 , si $F < F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1)$ o $F > F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$
 - No rechazar H_0 , si $F \geq F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1)$ o $F \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$
- o sea, $F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) \leq F \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$.

3.1- Cálculo del estadístico de prueba o estadígrafo F :

Cálculo de la información necesaria para la determinación de las varianzas.

Tabla 1.23. Cálculo de la información necesaria para determinación de la varianza.

Costo real por tonelada X_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	Costo calculado por tonelada Y_i	$(Y_i - \bar{Y})$	$(Y_i - \bar{Y})^2$
121,918	9,608	92,313	110,681	4,541	10,843
117,941	5,631	31,708	106,868	0,728	0,270
115,312	3,002	9,012	117,902	11,762	110,544
114,521	2,211	4,888	115,210	9,070	61,183
109,327	-2,983	8,898	108,791	1,403	1,968
100,915	-11,395	129,846	107,217	-0,171	0,029
103,001	-9,309	86,657	96,202	-9,938	125,126
104,984	-7,326	53,670	95,747	-10,393	135,512
103,901	-8,409	70,711	95,852	-10,288	133,079
106,868	-5,442	29,615	96,5159	-9,624	118,202
115,998	3,688	13,601	112,025	5,885	21,501
110,828	-1,482	2,196	113,808	7,668	41,216
117,237	4,927	24,275	116,151	10,011	76,790
119,230	6,92	47,886	104,767	-1,373	6,869
109,208	-3,102	9,622	108,228	2,088	0,705
125,770	13,46	181,171	112,250	6,110	23,639
1796,959	////	796,074	1718,214	////	867,484

Fuente: elaborada por los autores.

Primeramente determinar: \bar{X} , \bar{Y} , S_x^2 y S_y^2 .

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \qquad \bar{X} = \frac{1796,959}{16} = 112,31$$

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} \qquad \bar{Y} = \frac{1718,214}{16} = 107,38$$

$$S_x^2 = \frac{\sum_{i=1}^{n_x} (X_i - \bar{X})^2}{n_x - 1} \qquad S_x^2 = \frac{796,074}{16 - 1} = 53,07$$

$$S_y^2 = \frac{\sum_{i=1}^{n_y} (Y_i - \bar{Y})^2}{n_y - 1} \qquad S_y^2 = \frac{867,484}{15} = 57,83$$

Cálculo de F :

$$F = \frac{S_x^2}{S_y^2} \quad F = \frac{53,07}{57,83} = 0,92$$

3.2- Determinación del valor crítico de F :

$$F_{\frac{0,05}{2}}(16-1, 16-1) = \frac{1}{F_{1-\frac{0,05}{2}}(16-1, 16-1)}$$

$$F_{0,025}(15, 15) = 0,349 = \frac{1}{F_{0,975}(15, 15)} = \frac{1}{2,86} = 0,349$$

Como: $F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) \leq F \leq F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1)$, es decir, $0,349 \leq 0,92 \leq 2,86$, se cumple; por tanto, no se rechaza la hipótesis nula.

Considerando los resultados de las muestras y con un nivel de significación 0,05, se puede aseverar que se cumple el supuesto de igualdad de varianzas.

Prueba de Levene de igualdad de varianzas

El resultado anterior se corrobora con el obtenido de la prueba de Levene, de igualdad de varianzas, mediante la aplicación del paquete de programa estadístico SPSS, como se exhibe en la tabla 1.24.

La igualdad de varianzas en la prueba t de Student para dos muestras independientes en el paquete de programa estadístico SPSS, se comprueba a partir de la prueba de *Levene para la igualdad de varianzas*. La prueba de Levene toma un valor de significancia o valor p de 0,93, que excede el nivel de significación o nivel de error de 0,05, en consecuencia no se rechaza la hipótesis nula de igualdad de varianzas $H_0: \sigma_x^2 = \sigma_y^2$, por lo que no existe diferencia significativa entre las varianzas, o sea se puede asumir que las varianzas de la variable de los costos son iguales.

Tabla 1.24. Prueba de Levene para la igualdad de varianzas (SPSS).

		Prueba de muestras independientes								
		Prueba de Levene de igualdad de varianzas		Prueba <i>t</i> para la igualdad de medias						
		F	Sig.	t	Grados de libertad	Sig. (bilateral)	Diferencias de medias	95% de intervalo de confianza de la diferencia		
								Diferencia de error estándar	Inferior	Superior
Costos	Se asumen varianzas iguales	,007	,936	1,869	30	,071	4,92151	2,63277	-,45533	10,298
	No se asumen varianzas iguales			1,869	29,9	,071	4,92151	2,63277	-,45575	10,298

Se concluye que se cumplen con los supuestos de normalidad y de igualdad de varianzas.

Prueba de hipótesis referentes a las medias de dos poblaciones con distribuciones normales y varianzas desconocidas pero iguales

Comprobados los supuestos de normalidad y de igualdad de varianzas, se realiza la prueba *t* de Student como se expone a continuación.

1- Formulación de la hipótesis nula y alternativa:

$H_0: \mu_x = \mu_y$. No existe una diferencia significativa entre la media del costo real y la media del costo calculado.

$H_1: \mu_x \neq \mu_y$. Existe una diferencia significativa entre la media del costo real y la media del costo calculado.

2- Región crítica para el estadígrafo *t*, dado el nivel de significación α , es

$$W = \left\{ t \in \mathbb{R} : t < -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \quad \text{o} \quad t > t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \right\}.$$

donde:

$$t = \frac{\bar{x} - \bar{y}}{S_o \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \quad \text{o} \quad t = \frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2} \right) \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$$

$t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$ es el percentil de orden $1 - \frac{\alpha}{2}$ de la distribución t de Student con $(n_1 + n_2 - 2)$ grados de libertad.

3- Regla de decisión:

- Rechazar H_0 , si $t < -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$ o $t > t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$.

- No rechazar H_0 , si $t \geq -t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$ o $t \leq t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$

o sea, $-t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \leq t \leq t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$.

3.1- Cálculo del estadístico de prueba o estadígrafo t :

Información necesaria para calcular el estadístico de prueba t .

$$S_x^2 = \frac{\sum_{i=1}^{n_x} (X_i - \bar{X})^2}{n_x - 1} \quad S_x^2 = \frac{796,074}{16 - 1} = 53,07$$

$$S_y^2 = \frac{\sum_{i=1}^{n_y} (Y_i - \bar{Y})^2}{n_y - 1} \quad S_y^2 = \frac{867,484}{15} = 57,83$$

$$S_o = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}} \quad S_o = \sqrt{\frac{(16-1)53,07 + (16-1)57,83}{16+16-2}}$$

$$S_o = \sqrt{\frac{796,05 + 867,45}{30}}$$

$$S_o = 7,44$$

Cálculo de t :

$$t = \frac{\bar{x} - \bar{y}}{S_o \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \quad t = \frac{112,31 - 107,38}{7,44 \sqrt{\frac{1}{16} + \frac{1}{16}}}$$

$$t = \frac{4,93}{2,63} = 1,87$$

Expresión más resumida para calcular el estadístico de prueba t :

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2} \right) \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$$

$$t = \frac{112,31 - 107,388}{\sqrt{\left(\frac{(16-1)53,07 + (16-1)57,83}{16+16-2} \right) \left(\frac{1}{16} + \frac{1}{16} \right)}} = \frac{4,92}{\sqrt{55,45 \cdot 0,125}} = \frac{4,92}{2,63} = 1,87$$

3.2- Determinación del valor del percentil o valor crítico de t :

$$t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) = t_{1-\frac{0,05}{2}}(16 + 16 - 2) = t_{0,975}(30) = 2,04$$

Como: $-t_{1-\frac{\alpha}{2}}(n_x + n_y - 2) \leq t \leq t_{1-\frac{\alpha}{2}}(n_x + n_y - 2)$, es decir,

$-2,04 \leq 1,87 \leq 2,04$, se cumple; por tanto, no se rechaza la hipótesis nula.

Considerando los resultados de las muestras y con un nivel de significación 0,05, se puede aseverar que no existe diferencia significativa entre el costo promedio de transportación calculado a partir de la función lineal del costo $Y = \alpha' + \beta X$, y el costo promedio real de transportación, ya que provienen de poblaciones con igualdad en las medias.

Elementos a tener en cuenta para el caso de muestras grandes, para n_x y n_y suficientemente grandes la distribución t de Student se aproxima a la distribución normal, y por tanto, la distribución del estadígrafo t será aproximadamente igual a la normal estándar, por lo que en las regiones críticas se podrán sustituir los percentiles $t_{\beta}(n_1 + n_2 - 2)$ de la distribución t de Student por los percentiles Z_{β} de la distribución normal estándar, para todo $\beta \in]0,1[$. Donde $t_{\beta}(n_1 + n_2 - 2) \approx Z_{\beta}$.

Los resultados expuestos se confirman con los resultados obtenidos a partir de la prueba t de Student para dos muestras independientes, mediante la aplicación del SPSS, como se exhibe en la tabla 1.25.

El nivel de significancia o valor p para la prueba t de Student es 0,071. Este valor corresponde a cuando se asume la igualdad de varianzas. Contrastando el valor p de 0,071 con el nivel de significación 0,05, al ser mayor, no hay evidencia para rechazar la hipótesis nula, por tanto, no existen diferencias significativas entre la media del costo real y la media del costo calculado. Evidentemente, existe una diferencia entre los promedios de los costos, ya que el costo real es de \$ 112,31 por tonelada y el costo calculado es de \$ 107,38 por tonelada, pero la diferencia no es significativa.

Tabla 1.25. Prueba t de Student para dos muestras independientes (SPSS).

		Prueba de Levene de igualdad de varianzas		Prueba t para la igualdad de medias						
		F	Sig.	t	Grados de libertad	Sig. (bilateral)	Diferencias de medias	95% de intervalo de confianza de la diferencia		
								Diferencia de error estándar	Inferior	Superior
Costos	Se asumen varianzas iguales	,007	,936	1,869	30	,071	4,92151	2,63277	-,45533	10,298
	No se asumen varianzas iguales			1,869	29,9	,071	4,92151	2,63277	-,45575	10,298

1.7.3. Determinación de la distancia económica de los medios de transportes

En el proceso de planificación y organización de los medios de transportes, es indispensable tener en cuenta la distancia económica o crítica de transportación para la toma de decisiones científica y económicamente fundamentadas, es decir, a partir de qué distancia resulta más económico emplear un medio de transporte.

A partir de las funciones racionales enteras de tipo lineal, representadas por las funciones de los costos determinados para cada medio de transporte, se puede determinar la distancia económica o crítica entre los medios de transporte comparados, de la siguiente forma:

- Simultaneando la función lineal de costo correspondiente a cada medio de transporte.
- A partir del punto de equilibrio lineal algebraico, donde los costos

coinciden numéricamente.

- Otra forma es aplicando el análisis del punto de equilibrio lineal, mediante la representación gráfica de las funciones de los costos de cada medio de transporte en un sistema de coordenadas; comparándolas se obtiene el punto de equilibrio lineal, donde las funciones se interceptan, representado por el par ordenado (x,y) para las funciones en cuestión, donde el valor correspondiente a la abscisa (x) constituye la *distancia económica o crítica*.

Para demostrar lo expuesto, a partir de una situación real, se toman dos medios de transportación diferentes: camión y tractor con remolque, y los costos representados por las funciones lineales resultantes que se exponen a continuación:

- Camión $Y_c = 0,498 + 0,193 X$.
- Tractor con remolque $Y_t = 0,368 + 0,215 X$.

Las funciones lineales de los costos, para la situación objeto de estudio, está definida solo para los números reales positivos, o sea, el **Dom $f = \{x \in \mathfrak{R} : x > 0\}$** y la función toma solamente valores positivos, es decir, la **Im $f = \{y \in \mathfrak{R} : y > 0\}$** .

Como es de esperar, para ambos medios de transporte la tasa de variación es mayor que cero, siendo la función creciente en todo su dominio, por lo que el costo de los medios de transporte camión y tractor aumentan en 0,193 y 0,215, respectivamente, por cada kilómetro que aumente la distancia a recorrer con carga.

En la determinación de la distancia económica entre los medios de transporte, se procede de las siguientes formas:

- Simultaneando las funciones de costo, se obtiene que:

$$\begin{array}{r} Y_c = 0,498 + 0,193 \\ -Y_t = -0,368 - 0,215 X \\ \hline 0 = 0,130 - 0,022 X \quad X = 5,91 \text{ kilómetros} \end{array}$$

- Mediante el análisis del punto de equilibrio lineal, donde las funciones de los costos se igualan, se tiene algebraicamente que:

$$0,368 + 0,215 X = 0,498 + 0,193 X \quad 0,022 X = 0,130 \quad X = 5,91 \text{ kilómetros}$$

- Análisis gráfico del punto de equilibrio lineal.

Evaluando las funciones de costo para diferentes distancias, se obtienen los resultados siguientes:

Tabla 1.26. Costo de los medios de transporte por distancia recorrida.

Distancia (km)	0	1	2	3	4	5	5,91	6	7	8
Costo Tractor	0,368	0,583	0,798	1,013	1,228	1,443	1,638	1,658	1,873	2,088
Costo Camión	0,498	0,691	0,884	1,077	1,270	1,463	1,638	1,656	1,849	2,042

Fuente: elaborada por los autores.

Representándose los datos en un sistema de coordenadas, se obtiene el siguiente comportamiento gráfico:

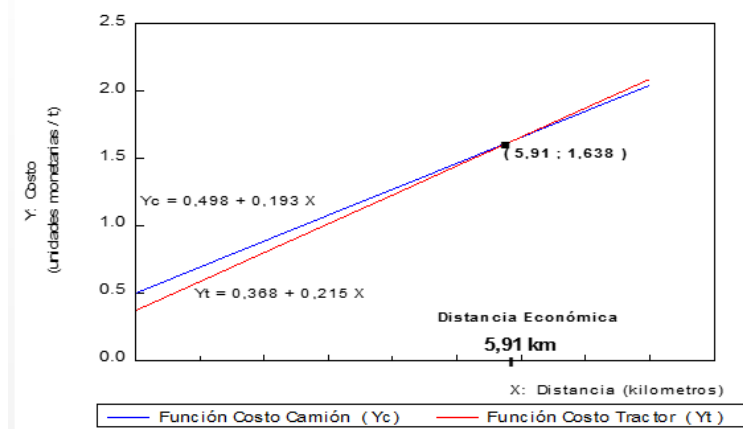


Figura 1.9. Representación de las funciones de costo y distancia económica.

Fuente: Elaboración de los autores.

Con los resultados obtenidos a partir del análisis analítico y gráfico, se aprecia que para una distancia de 5,91 kilómetros, los valores resultantes de las funciones de los costos son iguales a \$ 1,638, representando estos valores el punto de equilibrio. A partir de esta distancia, el costo correspondiente al medio de transporte camión con respecto al tractor comienza a disminuir, y el de tractor con respecto al de camión aumenta; dicha distancia constituye la distancia económica o crítica. Por tanto, para distancias superiores a 5,91

kilómetros resulta más económico emplear la transportación por camiones para el traslado de la carga.

1.7.4. Utilidad práctica de las funciones matemáticas de los costos en la modelización económico-matemática

Evaluando la función lineal de costo correspondiente a cualquier medio de transporte, para un valor dado de la distancia (X) entre el origen (campo cañero) y el destino (centro de recepción industrial), se obtiene el costo total de transportación por cada tonelada de caña a la distancia (X).

El costo calculado a la distancia dada entre los campos y los centros de recepción, constituyen los coeficientes económicos C_{ijk}^{lp} y C_{ijk}^{lpq} de las funciones objetivos de modelos económico-matemáticos planteados; constituyendo un resultado importante para la aplicación de estos modelos.

Utilizar un medio de transporte en dependencia de los resultados de la función de costo, posibilita su explotación económica a partir de la *relación costo-distancia*, lo que contribuye a disminuir el costo de transportación, ya que se utilizará el medio de transporte a partir de la distancia que económicamente se justifique, en correspondencia con lo expuesto en el epígrafe anterior.

Los costos de transportación determinados a partir de las funciones lineales obtenidas constituyen criterios de decisión económicos que proporcionan una mayor eficiencia al resultado o solución obtenida, y garantiza una mejor utilización de los recursos materiales y financieros disponibles.

La utilización adecuada, tanto económica como técnica, de cada medio de transporte de acuerdo los criterios expuestos, y en correspondencia con las condiciones concretas de que se trate, contribuye a lograr la eficiencia económica del transporte, expresado a partir de la relación favorable entre el efecto económico obtenido por la realización de transportación y los gastos indispensables para su aplicación.

La metódica desarrollada y validada, con resultados teóricos y prácticos significativos y superiores a los métodos tradicionales, permite el perfeccionamiento de los métodos convencionales de planificación.

1.8. Presentación del modelo numérico

Realizado el planteamiento matemático de los modelos económico-matemáticos teóricos objeto de estudio, procesada la información de partida y determinados los parámetros de los modelos, es indispensable buscar un procedimiento que permita consolidar e integrar los modelos teóricos con los resultados obtenidos; esto se puede lograr mediante la construcción del *modelo numérico*, expuesto en la tabla 1.27.

Como se exhibe, las dos primeras columnas reflejan los campos cañeros por productores. En el cuerpo de esta, se reflejan las vinculaciones a establecer entre los campos cañeros (orígenes) y los centros de recepción industrial (destino), mediante las variables de decisión o esenciales X_{ijk}^{lp} o X_{ijk}^{lpq} . Además, complementan a las vinculaciones la distancia media (en kilómetros) y el costo de transportación C_{ijk}^{lp} o C_{ijk}^{lpq} .

De utilizarse el modelo II de vinculación entre agroindustrias, se añade una cuarta columna a cada vinculación, donde se refleje el rendimiento industrial promedio (Rq).

Las filas corresponden a las restricciones de oferta o de producción. Al final de cada fila se encuentra la producción estimada de caña de azúcar de cada campo cañero (O_{il} o O_{ilq}). Las columnas corresponden a las restricciones de demanda o de capacidad industrial de los centros de recepción y al final de cada columna aparece la capacidad industrial estimada disponible de cada centro (D_j o D_{jq}); de tal forma, esta tabla resulta una representación adecuada del sistema de restricciones de oferta y demanda correspondiente a un modelo económico-matemático de transporte.

Como complemento importante al modelo numérico, se determinan las capacidades disponibles: de los medios de transporte (T_{kp} o T_{kpq}); de las alzadoras (A_l o A_{lq}) y de las cosechadoras (C_l o C_{lq}).

Tabla 1.27. Modelo Numérico.

Entidades de Producción Cañera	Campos Cañeros (orígenes)	Centros de Recepción (destinos)												Producción Estimada (toneladas)			
		Centro de Recepción (Destino) No. 1				Centro de Recepción (Destino) No. 2				Centro de Recepción (Destino) n							
		1	2	3	4	1	2	3	4	1	2	3	4				
Productor No. 1	001	X															
	002					X											
	005	X															
	010										X						
	015										X						
	025					X											
Productor No. 2	.																
	.																
Productor No. 3	.																
	.																
Productor No. n	.																
	.																
	.																
Capacidad Disponible (toneladas)																	////

Fuente: elaborada por los autores.

donde:

- 1- Vínculo planificado entre el origen y el destino, donde, X es la variable de decisión.
- 2- Distancia entre el origen y el destino.
- 3- Costo por tonelada transportada a la distancia dada.
- 4- Rendimiento industrial promedio.

1.9. Generalidades sobre la obtención y análisis de la solución del modelo de programación lineal

Los problemas de planificación de los procesos económicos presentan múltiples variantes, cuyos elementos y relaciones se pueden presentar cuantitativamente, por lo que es factible encontrar la solución óptima o satisfactoria de dicho problema mediante la aplicación de los métodos y modelos económico-matemáticos.

Existe una amplia variedad de métodos de solución, debido a la gran diversidad de problemas que se presentan en la práctica de la planificación y organización de los procesos económicos. Los métodos de solución pueden dividirse esencialmente en tres grupos: analítico-deductivo, numérico-inductivo y simulación.

El método analítico-deductivo consiste en el uso de la deducción matemática por medio de la aplicación de varios procedimientos de la matemática, se estudia el fenómeno pasando de lo general a lo particular. Para la solución de los modelos económico-matemáticos de programación lineal y de meta lineal que se construyen en el libro, se utiliza el procedimiento analítico-deductivo.

El método numérico-inductivo consiste en estudiar un fenómeno pasando de lo particular a lo general. En este método, se emplean procedimientos numéricos, que consisten en comprobar varios valores de las variables controlables en el modelo, comparando los resultados obtenidos y seleccionando aquel conjunto de valores de las variables controladas que produzcan la mejor solución.

Para obtener la solución óptima de los problemas de optimización, se emplea el Método Simplex desarrollado por G. Dantzig. No obstante, posterior a este, se han desarrollado la aplicación de otros algoritmos de solución, entre los que se encuentran el simplex revisado, el simplex compacto, el dual simplex, el primal dual, etc.

Variables de holgura

Para el desarrollo teórico y práctico de la Programación Lineal, resulta importante el establecimiento de una forma básica en la que se pueda escribir cualquier modelo de Programación Lineal.

Todo modelo lineal puede ser transformado a una forma estándar, donde el sistema de restricciones esté constituido por ecuaciones lineales, ya que en la práctica generalmente los sistemas de restricciones constan de inecuaciones o la combinación de inecuaciones y ecuaciones.

Las transformaciones se efectúan introduciendo variables adicionales, denominadas variables de holgura.

Características de las variables de holgura:

- Se utilizan para convertir el Sistema de Inecuaciones Lineales en Sistema de Ecuaciones Lineales.
- Desde el punto de vista del significado económico, asumen la dimensión de la restricción a la que están asociadas.
- Su coeficiente económico en la función objetivo es cero. Entonces, la conversión del sistema de restricciones originales en un sistema de ecuaciones lineales no cambia la función a ser optimizada. Optimizar el modelo utilizando las variables de holgura es equivalente a optimizar el modelo original, es decir, existe correspondencia entre las soluciones posibles y óptimas del conjunto original de las restricciones y del conjunto de ecuaciones lineales.

Puede darse el caso que las restricciones originales sean de igualdad, y donde el sistema contenga la matriz unitaria, no siendo necesario incluir las variables de holgura.

Variables artificiales

No siempre se logra de manera inmediata la solución básica inicial agregando variables de holguras, es decir, no se logra un sistema que contenga la matriz unitaria.

En tal sentido, se utiliza un procedimiento denominado adición de una base artificial, con el que se representa la matriz unitaria.

Las variables utilizadas, llamadas artificiales, se introducen solo para construir de forma inmediata una solución básica inicial, las mismas desaparecen durante el proceso de cómputo.

Las variables artificiales se utilizarán:

- En restricciones de mayor o igual, pues con las variables de holgura no se puede construir una matriz básica unitaria.
- En las restricciones de igualdad, pero no necesariamente.

Para asegurar que las variables artificiales no aparezcan en la solución óptima, se le asignan coeficientes económicos que no mejoren el valor de la función objetivo, mientras permanezcan en la base variables artificiales.

Tipos especiales de solución

Cuando se procede a obtener la solución de un problema de programación lineal utilizando el Método Simplex, se pueden presentar distintas situaciones, y arribar a tipos especiales de solución, tales como: solución óptima alterna, solución degenerada, redundancia, solución no acotada, y no existencia de solución.

Obtener soluciones óptimas alternas al resolver problemas económicos modelados mediante la programación lineal es conveniente, ya que amplía las posibilidades en el análisis de la solución al obtener diferentes variantes de solución del problema objeto de estudio y un mismo valor óptimo para la función objetivo. Por tanto, se dice que un problema de programación lineal tiene solución óptima alterna, cuando el valor óptimo para la función objetivo se obtiene en dos o más soluciones posibles básicas diferentes.

Software para obtener la solución

Los algoritmos de solución del Método Simplex han sido programados para computadoras, posibilitando que existan paquetes de programas computacionales para la solución de problemas de optimización, entre lo que se destacan: Linear Optimization (LO); Linear, Interactive, Discrete Optimizer (LINDO e HIPERLINDO); The Management Scientist (MS); Computer Models for Management Science (JOVINI); Linear Programming (LP88), QM for Windows; Quantitative System Business for Windows (WinQSB), entre los más usados.

El paquete de programa WinQSB es uno de los softwares más utilizados en la actualidad en la ciencia de la Investigación de Operaciones y en los Métodos Cuantitativos para la Toma de Decisiones en Administración, constituyendo una aplicación versátil que, debido a su facilidad y potencia de manejo, se convierte en una herramienta indispensable para la solución de problemas relacionados con los métodos cuantitativos aplicados a la administración (Quesada & Vergara, 2006). El libro de los autores de referencia, proporciona al lector un análisis de las diferentes técnicas cuantitativas de Investigación Operaciones, mediante la utilización del WinQSB, y en particular de la Programación Lineal y de la Programación de Meta Lineal.

La solución óptima de los modelos económico-matemáticos para la

optimización del transporte en la agroindustria azucarera, desarrollado en el presente capítulo, se puede obtener por medio de la aplicación del programa WinQSB, utilizando el módulo Lineal and Integer Programming (IGP).

La solución satisfactoria del modelo económico-matemático de meta con prioridades del movimiento de rebaño, desarrollado en el capítulo II, se puede obtener por medio de la aplicación del programa WinQSB, utilizando el módulo Linear goal programming (GP) and Integer linear goal programming.

Análisis económico de la solución

Obtenida la solución óptima a partir del reporte de salida de un programa de computación o de la tabla de iteración óptima del Método Simplex, se procede a realizar el análisis económico de la solución óptima, observando qué vectores forman parte de la base, y qué valor toman las variables asociadas a dichos vectores. En tal sentido, se realizará el siguiente análisis teniendo en cuenta la información que se exhibe en la tabla 1.28.

Tabla 1.28. Situaciones que se presentan en las variables esenciales y de holgura en la solución óptima.

Vector	Variables con significado económico	Signo de la restricción	Valor de la variable
Básico	Decisión	-	$X_j > 0$
No Básico	Decisión	-	$X_j = 0$
Básico	Holgura	\leq	$X_j > 0$
Básico	Holgura	\geq	$X_j > 0$
No Básico	Holgura	\leq	$X_j = 0$
No Básico	Holgura	\geq	$X_j = 0$

Fuente: elaborada por los autores.

Para interpretar económicamente el valor óptimo de las variables esenciales o de decisión hay que tener en cuenta su definición, es decir, lo que expresa desde el punto de vista cualitativo, dimensional y temporal en el contexto del problema.

Un vector básico asociado a una variable de decisión o esencial indica que el producto que se analice se produzca, se transporte, se venda, se compre, etc. Un vector no básico asociado a una variable esencial indica que no se lleve a cabo, lo que la variable asociada signifique, es decir, que no se

produzca, no se transporte, no se venda, no se compre, etc.

Para el análisis económico de las variables de holguras hay que tener en cuenta el signo, y el significado dimensional y conceptual de la restricción a la cual se encuentre asociada la variable de holgura, existiendo los siguientes contextos:

- Un vector básico asociado a una variable de holgura en una restricción con signo \leq . Si la restricción es de recurso, indica ociosidad, subutilización. Si la restricción es de demanda máxima, indica lo dejado de producir, elaborar, transportar, etc.
- Un vector básico asociado a una variable de holgura en una restricción con signo \geq . Si la restricción es de recurso, indica una utilización o aprovechamiento por encima de un mínimo establecido de un recurso. Si la restricción es de demanda mínima, indica las unidades que de dicho producto se satisfacen por encima del mínimo establecido.
- Un vector no básico asociado a una variable de holgura en una restricción con signo \leq . Si la restricción es de recurso, indica la utilización total del recurso, constituyendo un recurso crítico. Si la restricción es de demanda máxima, indica la satisfacción total de la demanda máxima.
- Un vector no básico asociado a una variable de holgura en una restricción con signo \geq . Si la restricción es de recurso, indica la utilización del mínimo establecido del recurso. Si la restricción es de demanda mínima, indica la satisfacción de la demanda mínima.

Cuando se obtiene la solución óptima de un problema en específico, debe tenerse presente que la política que aparenta ser mejor en términos del modelo, puede no ser la mejor en la realidad:

- En primer lugar, porque el modelo es una representación de los aspectos más importantes y esenciales del problema real.
- En segundo lugar, porque pueden variar las condiciones de los parámetros originales del problema.

La solución obtenida se somete a consideración de los decisores responsables del trabajo, previamente revisada y analizada económica y matemáticamente.

La aplicación a la práctica de la solución obtenida indica, en última instancia, si se ha seguido el camino acertado. Por tanto, cuando se hace referencia a la solución, se hace con respecto al modelo, y no necesariamente, al sistema real representado en el modelo.

Capítulo II

Modelación económico-matemática para la planificación de procesos económicos en la ganadería

2.1. Introducción al capítulo. Generalidades sobre la modelación matemática y aplicaciones a los procesos económicos en la ganadería

Inicialmente en este capítulo, de forma sucinta, se compara la programación de meta lineal y la programación lineal, y se esbozan las características de la programación de meta. Se exhibe una aplicación de los modelos de optimización al movimiento de rebaño, lo que constituye el antecedente más importante del modelo económico-matemático de meta desarrollado.

A partir de la formulación matemática general y característica de las restricciones y función objetivo del modelo de programación de meta lineal, y de la formulación de un problema de programación del movimiento de rebaño, sobre la base de sus particularidades tecnológicas y económicas, se realiza construcción del modelo económico-matemático de meta con prioridad. Dada la complejidad del proceso económico modelado y para lograr una mayor comprensión de este, se realiza un análisis de la construcción de las restricciones y función objetivo del modelo.

Como relevante, se presenta la modelización matemática del parámetro cantidad de cabezas del rebaño planificadas al final del año correspondiente al modelo de programación de meta lineal.

La determinación y validación de la cantidad de cabezas del rebaño planificadas al final del año se realiza mediante la estimativa e inferencia del análisis de una función de regresión lineal simple, donde se tienen en cuenta las pruebas de hipótesis de F de Fisher y t de Student, así como las pruebas de los supuestos básicos de no autocorrelación, homocedasticidad y normalidad, tales como: Durbin-Watson, Breusch-Godfrey, Goldfeld-Quandt, general de la heterocedasticidad de White, Jarque-Bera y Shapiro-Wilks. Además, se realizan inferencias a partir de la estimación del intervalo de confianza al coeficiente de regresión β_2 .

Se utiliza la prueba no paramétrica Rango de Wilcoxon para comparar la

existencia de correspondencia entre la estructura real del rebaño y la obtenida de la solución del modelo.

Al igual que en el capítulo precedente, en los contenidos tratados de programación de meta lineal, estadística matemática, econometría, matemática y álgebra lineal inicialmente, se hace una exposición teórica y, posteriormente, en las aplicaciones prácticas, se presentan sus resultados calculados manualmente y obtenidos mediante programas informáticos correspondientes a dichas materias.

En la parte final del capítulo, se hace un análisis económico de la solución del modelo matemático de meta con prioridad.

Los modelos económico-matemáticos pueden desempeñar un papel relevante como medio del conocimiento y comprobación en la planificación de la economía, en la medida que sean capaces de representar las relaciones económicas y productivas en forma matemática adecuada. La creciente complejidad de la reproducción en los procesos productivos hace que la preparación y la toma de decisiones en la planificación sea, en general, un proceso multifactorial y complejo.

La utilización de los métodos y modelos económico-matemáticos en la planificación y gestión de los procesos económicos no significa solo cambios cuantitativos, expresados en un análisis económico más complejo y exacto de la información; además de esta ventaja, dichos métodos y modelos deben conducir a cambios cualitativos en la metodología de la planificación.

En el epígrafe 1.4, se hizo referencia a las ventajas de los modelos de optimización lineal con respecto a otros métodos cuantitativos; no obstante, debido a que la mayoría de las decisiones empresariales se caracterizan por metas, criterios y objetivos múltiples más que por un objetivo único, desde la década de los años sesenta y setenta del siglo pasado, paralelo a la programación lineal, se han desarrollado métodos y modelos de programación de meta (Goal Programming), técnica cuantitativa aplicada a los problemas de toma de decisiones multicriterios, donde exhiben una manera de resolver problemas de programación lineal no factibles que se originan en la interacción de restricciones meta-recursos, tales interacciones ocurren cuando las metas son conflictivas e inconmensurables.

Haciendo una comparación entre la programación de metas con respecto a la programación lineal, y teniendo en cuenta las consideraciones realizadas por Moskowitz & Wright (1992); y Eppen, et al. (2000), en la programación lineal todos los objetivos deben incluirse en la función objetivo y reducirse a un solo criterio o dimensión medible, mientras que la programación de meta permite incluir los objetivos o metas en la formulación del problema con más de una dimensión.

La programación de meta es más flexible, ya que permite que las metas conflictivas e inconmensurables sean especificadas y conduzcan a una solución satisfactoria (mejor solución obtenida con la aplicación de un modelo económico-matemático de meta), en término de las metas prioritarias de la administración. La programación lineal bajo estas circunstancias conduce, en muchos casos, a soluciones no factibles. La programación de meta no tiene muchas de las limitaciones de la programación lineal y la solución de estos modelos se puede obtener usando también el algoritmo del Método Simplex.

Dentro de las características más destacadas de la programación de meta se encuentran que:

- Es una extensión de la programación lineal y proporciona solución a problemas de decisión que tengan metas múltiples, conflictivas e inconmensurables, en correspondencia con la estructura de prioridad establecida por el decisor.
- Las metas se satisfacen en una secuencia ordinal, por lo que deben de clasificarse en orden de prioridad, de importancia, por el decisor, y son satisfechas secuencialmente por el algoritmo de solución.
- Las metas con prioridad baja se consideran solamente después del cumplimiento de las metas de prioridad alta.
- La programación de meta es un proceso de satisfacción, en el sentido que el decisor tratará de alcanzar un nivel satisfactorio de objetivos múltiples, en lugar del mejor resultado posible o solución óptima, como ocurre en programación lineal para un solo objetivo.
- En la programación de meta, en lugar de minimizar o maximizar la función objetivo directamente, se minimizan las desviaciones entre las metas y los

límites alcanzables establecidos por el conjunto dado de restricciones en los recursos.

- Las variables de desviaciones se dividen en desviaciones positivas y desviaciones negativas de cada una de las metas. El objetivo se convierte en la minimización de estas desviaciones, dentro de la estructura prioritaria asignada a estas.

Dentro de las aplicaciones prácticas de los métodos y modelos de programación lineal y programación de meta para la toma de decisiones multicriterios, ocupa un lugar importante la utilización de estos métodos y modelos en los procesos económicos del sector agropecuario, con el objetivo de que sus resultados sirvan de base para la toma de decisiones, en la elaboración y fundamentación de los planes de producción en la ganadería y la agricultura.

En la planificación de la rama ganadera, en particular, se hace imprescindible la aplicación de métodos de planificación científicamente fundamentados, ya que para el desarrollo eficiente de esta actividad se hace necesaria la integración armónica de múltiples factores estrechamente correlacionados, que demandan un riguroso estudio de sus relaciones, lo cual no es posible lograr por los métodos de planificación convencionales.

Los trabajos sobre la utilización de los métodos cuantitativos en la planificación de la ganadería vacuna se encuentran dispersos y no son muy abundantes. Sin embargo, existen aplicaciones de modelos económico-matemáticos de programación lineal y de toma de decisiones multicriterios en la planificación de los procesos agropecuarios; entre otros se encuentran: los realizados por González & Brezo (1987); Ahumada & Villalobos (2009); Mezghani, Rebai, Dammak & Loukil (2009); Stygar & Makulska (2010); Ortuño & Vitoriano (2011); Dave (2015); y Zgajnar & Kavcic (2016). Los trabajos desarrollados y aplicados de manera general han estado encaminados a la aplicación de métodos cuantitativos para la toma de decisiones en la planificación agropecuaria, en la optimización del balance alimentario, en la planificación del desarrollo perspectivo y optimización del movimiento de rebaño.

Con respecto a la optimización del movimiento de rebaño, que constituye la parte consubstancial de este capítulo, la literatura científica relacionada con la modelación matemática del movimiento de rebaño vacuno es exigua,

incluso en publicaciones de corriente principal. No obstante, el trabajo que aborda metodológicamente con mayor claridad y profundidad este problema es el desarrollado por los autores González & Brezo (1987), que plantean un modelo matemático de programación lineal donde se maximiza la producción mercantil de carne, sujeto a restricciones relativas del movimiento de rebaño, transferencias de animales a categorías superiores, restricciones que garantizan la existencia final por categoría y del rebaño global, restricciones que determinan la correspondencia entre categorías, restricciones que regulan la extracción para la venta y restricciones de correlación estructural.

Dadas las complejidades de las condiciones productivas, económicas y biológicas del proceso que se modela, se considera como relevante que los autores de referencia, para el planteamiento del modelo económico-matemático, establecen una simbología para los elementos del movimiento de rebaño en sus diferentes fases, es decir, consideran por categoría de edad y sexo la simbología referente a: existencia al inicio y final de año, nacimientos, porcentaje de ganado extraído para carne o venta, porcentaje de pérdidas, peso promedio vivo de cada cabeza de ganado, transferencias a categorías superiores y llegadas de categorías inferiores.

Entre las limitaciones más relevantes de la investigación realizada por González & Brezo (1987), sobre la construcción y aplicación de un modelo económico-matemático de optimización lineal para el movimiento de rebaño, se encuentran que:

- Los trabajos de investigación sobre la optimización del movimiento de rebaño no han tenido un uso y seguimiento sistemáticos, que permitieran validar los resultados obtenidos en la práctica y, como consecuencia, realizar aplicaciones sucesivas.
- El modelo de programación lineal desarrollado no se encuentra debidamente estructurado, pues no se definen las variables de decisión, ni los parámetros del modelo, que permitan una construcción consecuente del sistema de restricción y de la función objetivo, lo que imposibilita la comprensión adecuada del modelo presentado.
- El modelo no hace un análisis integral de las diferentes fases del movimiento de rebaño, ya que no incluye las compras a terceros.
- No se utilizan los métodos estadístico-matemáticos y procedimientos

estadísticamente fundamentados para la estimación y validación de algunos de los parámetros del modelo de programación lineal, con el propósito de flexibilizar y robustecer los resultados obtenidos.

- El análisis económico realizado a la solución óptima es breve y poco profundo, pues no se hace un análisis integral y exhaustivo de los resultados obtenidos, correspondiente a las variables, restricciones y función objetivo.
- Para la modelación del sistema complejo y multifactorial que se estudia, no se consideran las bondades y flexibilidad que proporcionan otras técnicas cuantitativas entre las que se encuentran los modelos de toma de decisiones multicriterios.

El estudio realizado a los trabajos de Ahumada & Villalobos (2009); Mezghani, et al. (2009); Stygar & Makulska (2010); Ortuño & Vitoriano (2011); Dave (2015); y Zgajnar & Kavcic (2016); y de manera relevante del estudio efectuado al problema de optimización del balance de movimiento de rebaño planteado por González & Brezo (1987), permitió realizar el análisis propuesto, desarrollar y aplicar un modelo económico-matemático de meta con prioridad; técnica cuantitativa aplicada a los problemas de toma de decisiones multicriterios, fundamentada en los criterios dados por Moskowitz & Wright (1992); y Eppen, et al. (2000), donde se exhibe una manera de resolver los problemas de programación lineal no factibles.

En el presente capítulo, se enfocan los aspectos de la planificación del movimiento de rebaño, mediante una metodología donde se formula y plantea un *modelo económico-matemático de meta lineal con prioridades*. Además, se resuelve el problema planteado y se validan los resultados obtenidos; este último aspecto tiene lugar, pues el método propuesto resuelve las expectativas de los decisores empresariales.

En el siguiente epígrafe, se presenta la formulación matemática general del modelo de programación de meta lineal y su aplicación a la programación del movimiento de rebaño.

2.2. Formulación general del problema y modelación económico-matemática de la programación del movimiento de rebaño

En el epígrafe, se presenta la formulación matemática general del modelo

programación de meta lineal, y las características de sus variables, restricciones y función objetivo. Además, se define con todos sus detalles y particularidades el sistema de planificación del movimiento de rebaño a estudiar, lo que posibilita el planteamiento del modelo económico-matemático de meta con prioridad para la programación del movimiento de rebaño.

Formulación matemática del modelo general de programación de meta lineal

Los supuestos básicos del modelo de programación lineal se aplican igualmente al modelo de programación de metas.

La diferencia en la estructura es que la programación de metas no intenta minimizar o maximizar la función objetivo, como lo hace el modelo programación lineal; lo que busca es minimizar las desviaciones entre las metas deseadas y los resultados reales, en correspondencia a las prioridades asignadas.

La función objetivo de un modelo de programación de meta es expresada en términos de las desviaciones de las metas, es decir, las variables de holguras y sobrantes de las restricciones se colocan en la función objetivo y deben minimizarse.

Formulación matemática general:

$$\text{Min } Z = \sum_{i=1}^m W_i (d_i^+ + d_i^-)$$

sujeto a

$$\sum_{j=1}^n a_{ij} X_j - d_i^+ + d_i^- = b_i$$

$$X_j, d_i^+, d_i^- \geq 0, \quad j = 1, 2, \dots, n$$

donde:

X_j es la variable de decisión o esenciales.

W_i son los pesos de ponderación (ordinal o cardinal) asignado a cada una de las metas.

d_i^- , d_i^+ son las variables de desviación no negativas.

$-d_i^+$ es el grado de sobreloro de la meta, logro por encima, y exceso. Cantidad que sobra o sobrante o cantidad en que supera la meta.

$+d_i^-$ es el grado de sublogro, logro por debajo, y deficiencia. Cantidad que falta u holgura o cantidad que no alcanza la meta.

Como no podemos tener exceso (sobrante) y deficiencia (holgura), a la vez, entonces una o ambas de estas variables deben ser igual a cero, es decir, $d_i^- \cdot d_i^+ = 0$.

Características de las restricciones y función objetivo del modelo de programación de meta lineal

El tratamiento metodológico dado a las variables desviación, restricciones y función objetivo en la construcción del modelo de programación de meta lineal, que se exhibe en el presente capítulo, se apoya en los procedimientos y elementos conceptuales expuestos por Moskowitz & Wright (1992); y por Eppen, et al. (2000), expresados en las siguientes características:

- Cada restricción de meta consta de dos miembros:
 - Primer miembro: $g(X_p, \dots, X_n)$.
 - Segundo miembro: b_i .
- Las restricciones de metas se escriben utilizando variables de desviación no negativas: d_i^- , d_i^+ .
- En la optimización, al menos una de las variables d_i^- , d_i^+ , siempre será cero.
- La variable d_i^- representa la deficiencia, y la variable d_i^+ representa el exceso, en la formulación matemática general se dan otras definiciones similares.
- Siempre que se usa d_i^- va sumada en $g(X_p, \dots, X_n) : g(X_p, \dots, X_n) + d_i^-$.
- Siempre que se usa d_i^+ va restada en $g(X_p, \dots, X_n) : g(X_p, \dots, X_n) - d_i^+$.
- Solo aparecen variables de desviación o subconjunto de ellas en la función objetivo. El objetivo siempre es minimizado.

- Las variables de decisión X_j no aparecen en la función objetivo.

Tipos de restricciones de metas

I- Blanco o propósito:

Hacer $g(X_p, \dots, X_n)$ tan próximo a b_i como sea posible.

Restricción: $g(X_p, \dots, X_n) + d_i^- - d_i^+ = b_i$.

Función objetivo, *Mínimo*: $+d_i^- + d_i^+$.

II- Minimizar deficiencias o minimizar el valor faltante:

Restricción: $g(X_p, \dots, X_n) + d_i^- - d_i^+ = b_i$.

Función objetivo, *Mínimo*: $+d_i^-$.

Restricción equivalente: $g(X_p, \dots, X_n) + d_i^- \geq b_i$.

III- Minimizar excedente:

Restricción: $g(X_p, \dots, X_n) + d_i^- - d_i^+ = b_i$.

Función objetivo, *Mínimo*: $+d_i^+$.

Restricción equivalente: $g(X_p, \dots, X_n) - d_i^+ \leq b_i$.

IV- Restricción del intervalo de la meta:

Restricción $a_i \leq g(X_p, \dots, X_n) \leq b_i$.

Expresándolo como una meta: $a_i + d_i^- \leq g(X_p, \dots, X_n) \leq b_i - d_i^+$.

Lo que equivale a las dos restricciones: $g(X_p, \dots, X_n) + d_i^- \geq a_i$.

$$g(X_p, \dots, X_n) - d_i^+ \leq b_i$$

En general, las restricciones de meta son expresadas con mayor frecuencia en forma de igualdad, usando las variables de desviación de exceso $-d_i^+$ y de holgura $+d_i^-$ que se requieran.

Formulación general del problema de programación del movimiento de rebaño

La producción ganadera y el crecimiento del rebaño se planifican con ayuda

de cálculos especiales en dependencia de las particularidades biológicas de cada una de las categorías de los animales.

En correspondencia con lo anterior, la reproducción del rebaño en las empresas ganaderas se realiza de forma planificada, mediante el balance conocido como movimiento de rebaño, constituyendo uno de los procesamientos de datos y tratamientos de la información más importantes dentro del proceso de planificación del sistema de balance de la ganadería (González & Brezo, 1987).

El balance del movimiento de rebaño sirve de base a la empresa ganadera, para realizar el cálculo de: la producción bruta y mercantil ganadera, el crecimiento del rebaño, los gastos de la formación del rebaño, los costos unitarios de los productos ganaderos, etc. Se utiliza en la determinación de los requerimientos de los alimentos para el ganado, de las instalaciones productivas, del número de trabajadores, cantidad de salario a pagar, etc.

El ganado vacuno se subdivide en una serie de categorías de edad-sexo, partiendo del desarrollo fisiológico de la masa ganadera y de las formas organizativas tomadas para su cría y explotación.

Una forma muy utilizada para dividir el ganado vacuno en varias categorías es la siguiente:

- Hembras: vacas, novillas, añojas, terneras hasta 12 meses y terneras recién nacidas.
- Machos: bueyes, sementales, toros de ceba, toretes, añojos, terneros hasta 12 meses y terneros recién nacidos.

Esta representación de división por edad-sexo está basada en plazos anuales de desarrollo del ganado vacuno y posee la característica de que, al año siguiente del período analizado, todos los animales se transfieren a la categoría superior de edad, o se utilizan para la venta con destino a carne o a otras empresas.

En la producción ganadera, aparecen condiciones alternativas para el movimiento de rebaño que se reflejan en la siguiente problemática ¿Cuántas cabezas de cada categoría edad-sexo es necesario extraer y vender, así como transferir, a la siguiente categoría en orden superior?

La transferencia o extracción del ganado depende de la existencia de cabezas al inicio del año, de la cantidad de nacimientos, y de la necesidad de tener al final del año una determinada estructura en el rebaño que garantice la reproducción normal del ganado en el futuro y el crecimiento de la producción de carne y leche según los ritmos planificados.

Además, es necesario tener en cuenta las siguientes situaciones: primero, el *porcentaje de extracción de ganado para carne o para vender por cada categoría*, el cual deberá ser calculado de forma que se obtenga la cantidad máxima de producción ganadera y se asegure, al mismo tiempo, una correlación estructural entre categorías, que permita la reproducción normal del rebaño en los siguientes ciclos de su desarrollo; segundo, *las transferencias de los animales de una categoría a la siguiente en edad- sexo*, deben realizarse manteniendo los plazos de cría normales del rebaño, en correspondencia con las características de cada categoría (González & Brezo, 1987).

Para lograr lo antes expuesto, dada su complejidad, y dar solución a la problemática planteada, se utilizan los métodos modernos de planificación, mediante el empleo de técnicas cuantitativas de gestión como la modelación económico-matemática, que se ajusta a los requerimientos anteriores, a través de un modelo de *programación matemática de meta lineal con prioridades*. Según Moskowitz & Wright (1992), este tipo de modelo matemático es más flexible que el de programación lineal, ya que permite que las metas conflictivas e inconmensurables sean especificadas y conduzcan a una solución satisfactoria, en término de las metas prioritarias de la administración.

En el capítulo se analiza y desarrolla un modelo de Programación Matemática enfocado como modelo de Programación de Meta Lineal, relacionado con la planificación ganadera.

Planteamiento del modelo económico-matemático

Formulado el problema y definidos cualitativamente sus parámetros, se procede a la construcción del modelo fase, que se caracteriza por la representación matemática del problema objeto de estudio.

En la modelación matemática de los procesos económicos y productivos, se deben considerar la mayor cantidad de factores que influyen de una forma u

otra en el problema planteado, para lograr obtener una solución adecuada. En la determinación de la planificación ganadera, con el fin de lograr que el proceso económico modelado sea lo más real posible, se tienen en cuenta los factores más relevantes de la planificación de movimiento de rebaño, tratando de no obviar aquellos que, de una manera u otra, tienen mayor incidencia en la problemática modelada, es decir, se realizaron abstracciones de los aspectos menos esenciales, a fin de evitar decisiones erróneas.

Los factores que deben tenerse en cuenta son los siguientes: existencia del rebaño al inicio del año, nacimientos de terneras y terneros programados, ganado extraído para carne o ventas, tasa de mortalidad del ganado, peso vivo promedio de cada cabeza, llegadas de las categorías inferiores, compras a terceros, transferencia de ganado a categorías inmediatas superiores, producción promedio de leche anual por vaca, correlación entre categorías por sexo, criterios de reemplazo, crecimiento del rebaño, y la producción mercantil de carne en pie del rebaño, entre los más importantes.

La posibilidad de conjugar estos factores sobre una base científicamente argumentada se logra mediante el planteamiento de modelos económico-matemáticos de programación de meta con prioridades, para determinar la estructura satisfactoria del movimiento de rebaño, cuyo objetivo primario estará encaminado a minimizar la desviación de las metas, garantizando al mismo tiempo la producción de carne y leche planificada.

Considerando el análisis realizado sobre el objeto de estudio práctico se construye un modelo económico-matemático de meta con prioridades para determinar la estructura satisfactoria del movimiento de rebaño.

Tomando como referencia a González & Brezo (1987), los elementos más relevantes del movimiento de rebaño del ganado vacuno pueden ser expresados mediante la nomenclatura de las variables de decisión y parámetros del modelo que se muestran en la tabla 2.1, donde el ganado vacuno se subdivide en una serie de categorías de edad-sexo, partiendo del desarrollo fisiológico de la masa ganadera y de las formas organizativas consideradas para su cría y explotación.

Tabla 2.1. Nomenclatura de las variables de decisión y parámetros del movimiento de rebaño (donde: $i = 1, \dots, 12$ diferentes categorías de edad y sexo).

Categorías de edad y sexo	Existencia al inicio del año (cabezas)	Nacimientos previstos (cabezas)	Porcentaje de ganado extraído para carne o ventas X_i (%)	Porcentaje de pérdidas (tasa de mortalidad) (%)	Peso vivo promedio de cada cabeza de ganado (kg)
1- Vacas	a_1	-	X_1	m_1	p_1
2- Novillas	a_2	-	X_2	m_2	p_2
3- Añojas	a_3	-	X_3	m_3	p_3
4- Terneras, hasta un año	a_4	-	X_4	m_4	p_4
5- Terneras, crías	$a_5 = 0$	c_5	X_5	m_5	p_5
6- Toros Ceba	a_6	-	X_6	m_6	p_6
7- Bueyes	a_7	-	X_7	m_7	p_7
8- Sementales	a_8	-	X_8	m_8	p_8
9- Toretos	a_9	-	X_9	m_9	p_9
10- Añojos	a_{10}	-	X_{10}	m_{10}	p_{10}
11- Terneros, hasta un año	a_{11}	-	X_{11}	m_{11}	p_{11}
12- Terneros, crías	$a_{12} = 0$	c_{12}	X_{12}	m_{12}	p_{12}
Total	a	-	-	-	-

Categorías de edad y sexo	Llegadas de las categorías inferiores X_{i+n} (cabezas)	Transferencia de ganado a categorías superiores X_{i+2n} (cabezas)	Existencia al final del año X_{i+3n} (cabezas)	Compras a terceros X_{i+4n} (cabezas)
1- Vacas	X_{13}	$X_{25} = 0$	X_{37}	X_{49}
2- Novillas	X_{14}	X_{26}	X_{38}	X_{50}
3- Añojas	X_{15}	X_{27}	X_{39}	X_{51}
4- Terneras, hasta un año	X_{16}	X_{28}	X_{40}	X_{52}
5- Terneras, crías	$X_{17} = 0$	X_{29}	$X_{41} = 0$	X_{53}
6- Toros Ceba	X_{18}	$X_{30} = 0$	X_{42}	X_{54}
7- Bueyes	X_{19}	$X_{31} = 0$	X_{43}	X_{55}
8- Sementales	X_{20}	$X_{32} = 0$	X_{44}	X_{56}
9- Toretos	X_{21}	X_{33}	X_{45}	X_{57}
10- Añojos	X_{22}	X_{34}	X_{46}	X_{58}
11- Terneros, hasta un año	X_{23}	X_{35}	X_{47}	X_{59}
12- Terneros, crías	$X_{24} = 0$	X_{36}	$X_{48} = 0$	X_{60}
Total	-	-	b	-

Fuente: Elaboración de los autores a partir de González & Brezo (1987).

En la especificación de la nomenclatura, se anulan algunas variables de decisión (X_{i+n} , X_{i+2n} , y X_{i+3n}) y parámetros del modelo matemático de metas con prioridades, dadas las siguientes condiciones:

- En las categorías superiores de vacas, toros, bueyes y sementales, no existen transferencias de animales a categorías superiores, por tanto, las variables transferencia de ganado a categorías superiores (\mathbf{X}_{i+2n}): X_{25} , X_{30} , X_{31} , y X_{32} , son iguales a cero.
- En las categorías inferiores de crías recién nacidas no existen llegadas de otras categorías inferiores, ya que no existen; en estas condiciones, las variables llegadas de categorías inferiores (\mathbf{X}_{i+n}): X_{17} y X_{24} se anulan.
- Los animales pertenecientes a las categorías terneras crías y terneros crías, al finalizar el año pasarán a la categoría siguiente: terneras y terneros hasta un año de edad. En correspondencia con lo expresado, las variables existencia al final del año (\mathbf{X}_{i+3n}): X_{41} y X_{48} son iguales a cero, y por consiguiente, la existencia a inicio de año de terneras y terneros crías son iguales a cero, por tanto los parámetros a_5 y a_{12} se anulan.

Planteamiento general del modelo económico-matemático de meta con prioridad para la programación del movimiento de rebaño

Para el desarrollo y la construcción práctica del modelo económico-matemático de meta con prioridad, para la programación del movimiento de rebaño, que se expone a continuación, se tiene como referencia el modelo de programación lineal desarrollado por González & Brezo (1987), las aplicaciones de modelos de planificación en la cadena de suministro agroalimentario realizada por Ahumada & Villalobos (2009), el modelo de programación de meta para la planificación de la producción presentado por Mezghani, et al. (2009), las aplicaciones de modelos matemáticos en el manejo de ganado vacuno realizada por Stygar & Makulska (2010), la aplicación de la programación de meta para la planificación agrícola realizada por Ortuño & Vitoriano (2011), la aplicación de la programación de meta en la gestión agrícola expuesta por Dave (2015), y la aplicación del enfoque de multicriterios para la asignación óptima de recursos realizada por Zgajnar & Kavcic (2016).

Índice:

$i = 1, 2, \dots, n$ Categorías de edad y sexo, con un valor de $n = 12$.

Donde: $i = 1$ Vacas, $i = 2$ Novillas, $i = 3$ Añejas, $i = 4$ Terneras hasta un año, $i = 5$ Terneras crías, $i = 6$ Toros de Ceba, $i = 7$ Bueyes, $i = 8$ Sementales,

$i = 9$ Toretos, $i = 10$ Añojos, $i = 11$ Terneros hasta un año, e $i = 12$ Terneros crías.

Variables de decisión:

X_i	Porcentaje de cabezas del rebaño destinadas a la venta o al sacrificio por categoría de edad y sexo i .
X_{i+n}	Cantidad (número) de cabezas del rebaño correspondientes a la categoría de edad y sexo i , llegada de la categoría inmediata inferior.
X_{i+2n}	Cantidad (número) de cabezas del rebaño correspondientes a la categoría de edad y sexo i , transferidas a la categoría inmediata superior.
X_{i+3n}	Cantidad (número) de cabezas del rebaño por categoría de edad y sexo i , al final del año.
X_{i+4n}	Cantidad (número) de cabezas de animales a comprar por categoría de edad y sexo i a terceros.

Definición general de las variables de desviación:

d_i^+ : Sobregro de la meta en la restricción i -ésima (cantidad en que se supera la meta).

d_i^- : Sublogro de la meta en la restricción i -ésima (cantidad en que no supera la meta).

Definición particular de las variables de desviación:

$d_{i_1}^+$ y $d_{i_1}^-$: Cantidad por encima y por debajo, respectivamente, respecto a la meta de la restricción del movimiento de las categorías por edad y sexo i del rebaño.

d_{2i}^+ y d_{2i}^- : Cantidad por encima y por debajo, respectivamente, respecto a la meta de la restricción que garantiza la transferencia de animales hacia las categorías i inmediatas superiores.

d_3^- : Cantidad por debajo respecto a la meta de la restricción que garantiza la existencia del rebaño planificado al final del año.

d_{4i}^+ y d_{4i}^- : Cantidad por encima y por debajo, respectivamente, respecto a la meta de la restricción que garantiza la correlación estructural necesaria entre

distintas categorías i .

d_3^+ y d_3^- : Cantidad por encima y por debajo, respectivamente, respecto a la meta de la restricción que garantiza la producción mercantil de carne en pie del rebaño.

Definición de los coeficientes y parámetros del modelo.

Notación:	Definición de parámetros y coeficientes del modelo:
a_i	Cantidad (número) de cabezas del rebaño a inicio del año, por categoría de edad y sexo i .
a	Cantidad (número) total de cabezas del rebaño a inicio del año.
c_i	Cantidad (número) de cabezas previstas a nacer de las categorías terneras crías y terneros crías i .
m_i	Porcentaje de pérdida (tasa de mortalidad), por categoría de edad y sexo i .
p_i	Kilogramos del peso vivo promedio de cada cabeza de ganado, por categoría de edad y sexo i .
g_i g'_i	Cantidad mínima y máxima de cabezas de ganado, por categoría de edad y sexo i , respectivamente.
d_i d'_i	Cantidad mínima y máxima de cabezas de ganados a comprar, por categoría de edad y sexo i , respectivamente.
V	Cantidad (número) de cabezas mínimas de vacas necesarias para garantizar la producción de leche planificada de las vacas.
R	Producción promedio de leche anual por vaca, en litros.
P	Plan de producción de leche para el año, litros.
b	Cantidad (número) de cabezas del rebaño planificadas total, al final del año.
h_i h'_i	Porcentaje mínimo y máximo de ganado extraído para carne o venta por categoría de edad y sexo i , respectivamente.

t'	Valor máximo de proporcionalidad de la correlación estructural entre la categoría de edad y sexo inferior $(i+3n)$, con respecto a la categoría de edad y sexo inmediata superior $[(i+3n)-1]$ del rebaño al final del año.
\hat{B}_2	Ritmo de crecimiento o decrecimiento del rebaño, donde: $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$
L_{superior}	Límite superior del intervalo de confianza para β_2 .
L_{inferior}	Límite inferior del intervalo de confianza para β_2 .
B	Producción mercantil de carne en pie del rebaño.
$\frac{a_i X_i}{100}$	Cantidad (número) de cabezas del rebaño destinadas a la venta o al sacrificio, por categoría de edad y sexo i , excepto terneras crías y teneros crías.
$\frac{a_i m_i}{100}$	Cantidad (número) de cabezas del rebaño que desaparezcan por las diferentes causas de pérdida, como mortalidad, por categoría de edad y sexo i , excepto terneras crías y teneros crías.
$\frac{c_i X_i}{100}$	Cantidad (número) de cabezas del rebaño destinadas a la venta o al sacrificio, por categoría de edad y sexo, de terneras crías y teneros crías.
$\frac{c_i m_i}{100}$	Cantidad (número) de cabezas del rebaño que desaparezcan por las diferentes causas de pérdida, como mortalidad, por categoría de edad y sexo, de terneras crías y teneros crías.
$\frac{a_i X_i}{100} p_i$	Término que representa la producción mercantil de carne en pie de las categorías de edad y sexo i , excepto las categorías terneras crías y teneros crías.
$\frac{c_i X_i}{100} p_i$	Término que representa la producción mercantil de carne en pie de las categorías de edad y sexo i , terneras crías y teneros crías.

Definición de las metas de preferencia:

Se definen en orden de importancia decreciente, siendo P_1 la prioridad más importante.

P_1 : Para garantizar producción mercantil de carne en pie del rebaño.

P_2 : Para garantizar el plan de cabezas del rebaño por categoría de edad y sexo al final del año.

P_3 : Para garantizar la adecuada correlación estructural entre distintas categorías.

P_4 : Para garantizar el adecuado movimiento de las categorías por edad y sexo del rebaño.

P_5 : Para garantizar la transferencia de animales hacia las categorías inmediatas superiores.

Sistema de restricciones:

Restricción de tipo I: Para el movimiento de las categorías por edad y sexo del rebaño.

$$\frac{a_i X_i}{100} - X_{i+n} - X_{i+4n} + X_{i+2n} + X_{i+3n} + d_{1i}^- - d_{1i}^+ = a_i - \frac{a_i m_i}{100}$$

$$i = 1, 2, 3, 4, 6, 7, 8, 9, 10, 11$$

Restricción de tipo II: Para el movimiento de las categorías terneras crías y terneros crías del rebaño.

$$\frac{c_i X_i}{100} + X_{i+2n} = c_i - \frac{c_i m_i}{100}; \quad i = 5, 12$$

Restricción de tipo III: Para garantizar la transferencia de animales hacia las categorías inmediatas superiores.

$$\frac{a_i X_i}{100} + X_{i+2n} + d_{2i}^- - d_{2i}^+ = a_i - \frac{a_i m_i}{100}; \quad i = 2, 3, 4, 9, 10, 11$$

Restricción de tipo IV: Para garantizar un determinado volumen o rango de animales en algunas de las categorías.

$$X_{i+3n} \geq g_i; \quad i = 1, 2, \dots, n$$

$$X_{i+3n} \leq g'_i ; \quad i = 1, 2, \dots, n$$

Una aplicación de la restricción a las vacas consiste en garantizar un rebaño final de vacas con el que se podrá cumplir el plan de producción de leche; se expresa por:

$$X_{i+3n} \geq V ; \quad i = 1$$

La cantidad correspondiente a V se determina como:

$$V = \frac{2P}{R} - a_1$$

Que se obtiene a partir de:

$$P = \left(\frac{a_i + X_{i+3n}}{2} \right) R$$

$$P = \left(\frac{a_i + V}{2} \right) R$$

Restricción de tipo V: Para garantizar el rebaño planificado al final del período de un año.

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} + d_3^- = b$$

donde: $b = a + L_{\text{superior}}$

Restricción de tipo VI: Para garantizar la correspondencia entre la transferencia de animales de una categoría a otra.

$$X_{i+n} - X_{(i+2n)+1} = 0 ; \quad i = 1, 2, 3, 4, 9, 10, 11$$

$$\sum_{i=6,7,8} X_{i+n} - X_3 = 0 ; \text{ transferencia de toretes a las categorías}$$

superiores.

Restricción de tipo VII: Para regular la extracción de animales para carne o venta dentro de cada una de las categorías.

$$X_i \geq h_i ; \quad i = 1, 2, \dots, n$$

$$X_i \leq h'_i \quad ; \quad i = 1, 2, \dots, n$$

Restricción de tipo VIII: Para garantizar la correlación estructural necesaria entre distintas categorías.

$$X_{i+3n} - t'X_{(i+3n)-1} + d_{4i}^- - d_{4i}^+ = 0; \quad i = 1, 2, \dots, n$$

Restricción de tipo IX: Para garantizar las compras a terceros.

$$X_{i+4n} \geq d_i; \quad i = 1, 2, \dots, n$$

$$X_{i+4n} \leq d'_i; \quad i = 1, 2, \dots, n$$

Restricción de tipo X: Para garantizar la producción mercantil de carne en pie del rebaño.

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} \frac{a_i X_i}{100} p_i^+ + \sum_{i=5,12} \frac{c_i X_i}{100} p_i^+ + d_5^- - d_5^+ = B$$

Restricción de tipo XI: Condición de no negatividad de las variables.

$$X_i \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+n} \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+2n} \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+3n} \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+4n} \geq 0 \quad i = 1, 2, \dots, n$$

$$d_{1i}^- \geq 0 \quad d_{1i}^+ \geq 0 \quad i = 1, 2, 3, 4, 6, 7, 8, 9, 10, 11$$

$$d_{2i}^- \geq 0 \quad d_{2i}^+ \geq 0 \quad i = 2, 3, 4, 9, 10, 11$$

$$d_3^- \geq 0$$

$$d_{4i}^- \geq 0 \quad d_{4i}^+ \geq 0 \quad i = 1, 2, \dots, n$$

$$d_5^- \geq 0 \quad d_5^+ \geq 0$$

Función objetivo

$$\text{Mínimo } Z = P_1(d_5^- + d_5^+) + P_2 d_3^- + P_3(d_{4i}^- + d_{4i}^+) + P_4(d_{1i}^- + d_{1i}^+) + P_5(d_{2i}^- + d_{2i}^+)$$

Presentación resumida del planteamiento general del modelo económico-matemático de meta con prioridad para la programación del movimiento de rebaño.

$$\text{Mínimo } Z = P_1(d_5^- + d_5^+) + P_2 d_3^- + P_3(d_{4i}^- + d_{4i}^+) + P_4(d_{li}^- + d_{li}^+) + P_5(d_{2i}^- + d_{2i}^+)$$

sujeito a

$$\frac{a_i X_i}{100} - X_{i+n} - X_{i+4n} + X_{i+2n} + X_{i+3n} + d_{li}^- - d_{li}^+ = a_i - \frac{a_i m_i}{100}$$

$$i = 1, 2, 3, 4, 6, 7, 8, 9, 10, 11$$

$$\frac{c_i X_i}{100} + X_{i+2n} = c_i - \frac{c_i m_i}{100} \quad i = 5, 12$$

$$\frac{a_i X_i}{100} + X_{i+2n} + d_{2i}^- - d_{2i}^+ = a_i - \frac{a_i m_i}{100} \quad i = 2, 3, 4, 9, 10, 11$$

$$X_{i+3n} \geq g_i \quad i = 1, 2, \dots, n$$

$$X_{i+3n} \leq g'_i \quad i = 1, 2, \dots, n$$

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} + d_3^- = b + L_{\text{superior}}$$

$$X_{i+n} - X_{(i+2n)+1} = 0 \quad i = 1, 2, 3, 4, 9, 10, 11$$

$$\sum_{i=6,7,8} X_{i+n} - X_3 = 0$$

$$X_i \geq h_i \quad i = 1, 2, \dots, n$$

$$X_i \leq h'_i \quad i = 1, 2, \dots, n$$

$$X_{i+3n} - t' X_{(i+3n)-1} + d_{4i}^- - d_{4i}^+ = 0 \quad i = 1, \dots, n$$

$$X_{i+4n} \geq d_i \quad i = 1, 2, \dots, n$$

$$X_{i+4n} \leq d'_i \quad i = 1, 2, \dots, n$$

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} \frac{a_i X_i}{100} p_i + \sum_{i=5,12} \frac{c_i X_i}{100} p_i + d_5^- - d_5^+ = B$$

$$X_i \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+n} \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+2n} \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+3n} \geq 0 \quad i = 1, 2, \dots, n$$

$$X_{i+4n} \geq 0 \quad i = 1, 2, \dots, n$$

Todos:

$$d_{1i}^- \geq 0; d_{1i}^+ \geq 0; d_{2i}^- \geq 0; d_{2i}^+ \geq 0; d_3^- \geq 0; d_{4i}^- \geq 0; d_{4i}^+ \geq 0; d_5^- \geq 0; d_5^+ \geq 0$$

2.3. Análisis de la construcción del modelo económico-matemático de meta con prioridad para la programación del movimiento de rebaño

Se realiza un análisis sobre la construcción y aspectos esenciales del *sistema de restricciones*, teniendo en cuenta los criterios de González & Brezo (1987), así como de la *función objetivo*, exhibidos en el planteamiento general del modelo económico-matemático.

Restricción de tipo I, para el movimiento de las categorías por edad y sexo del rebaño, se representa la correspondencia existente entre los distintos elementos relevantes del movimiento de rebaño para cada una de las categorías por edad y sexo de los animales. Puede expresarse, en forma general, mediante la relación de los elementos siguientes, donde al determinar la cantidad de animales para las categorías seleccionadas del rebaño al final del año, X_{i+3n} , es necesario sustraerle a la existencia del rebaño al inicio del

año, a_i ; los animales destinados a la venta o al sacrificio, $\frac{a_i X_i}{100}$; los animales que desaparezcan por las diferentes causas de pérdida, como mortalidad, $\frac{a_i m_i}{100}$; y los animales transferidos a la categoría inmediata superior, X_{i+2n} ; además, se adicionan los animales que lleguen procedentes de la categoría inmediata inferior, X_{i+n} ; y los animales comprados a terceros, X_{i+4n} .

La relación de los elementos anteriores se puede expresar mediante la ecuación siguiente:

$$a_i \cdot \frac{a_i X_i}{100} - \frac{a_i m_i}{100} - X_{i+2n} + X_{i+n} + X_{i+4n} = X_{i+3n}; \quad i = 1,2,3,4,6,7,8,9,10,11$$

Considerar como término independiente las magnitudes conocidas y constantes por categoría de edad y sexo i e incorporando las variables de desviación d_{li}^+ y d_{li}^- , se obtiene la restricción de meta (restricción de tipo uno):

$$\frac{a_i X_i}{100} - X_{i+n} - X_{i+4n} + X_{i+2n} + X_{i+3n} + d_{li}^- - d_{li}^+ = a_i \cdot \frac{a_i m_i}{100}; \quad i = 1,2,3,4,6,7,8,9,10,11$$

La restricción de tipo I, para el movimiento de las categorías por edad y sexo del rebaño, se debe cumplir para las categorías de edad y sexo: vacas, novillas, añojas, terneras hasta un año; toros, bueyes, sementales, toretes, añojos y teneros hasta un año; no se debe utilizar en terneras y terneros crías.

Restricción de tipo II, para el movimiento de las categorías terneras crías y terneros crías del rebaño.

Todas las terneras crías y terneros crías, excepto las crías que se destinan para el sacrificio y ventas, y las que se pierden por alguna causa, deberán ser transferidas durante el año a la categoría inmediata superior, es decir, trasladadas a las terneras y terneros hasta un año de edad.

En el movimiento de rebaño, las categorías terneras y terneros crías pueden expresarse, en forma general, mediante la relación de los elementos siguientes, donde para determinar el número de cabezas de las categorías terneras y terneros crías, (C_i); se adicionan al número de cabezas del rebaño destinados a la venta o al sacrificio, ($\frac{c_i X_i}{100}$); el número de cabezas que desaparezcan

por las diferentes causas de pérdida, como mortalidad, ($\frac{c_i m_i}{100}$); y el traslado a la categoría inmediata superior, (X_{i+2n}).

$$c_i = \frac{c_i X_i}{100} + \frac{c_i m_i}{100} + X_{i+2n}; i = 5,12$$

Al considerar como término independiente las magnitudes conocidas y constantes, por categoría de edad y sexo i , se obtiene la restricción de tipo dos:

$$\frac{c_i X_i}{100} + X_{i+2n} = c_i - \frac{c_i m_i}{100}; i = 5,12$$

La restricción de tipo II para el movimiento de las categorías terneras crías y terneros crías del rebaño, como su nombre lo refiere, se debe cumplir para las categorías de edad y sexo terneras y terneros crías.

Restricción de tipo III, para garantizar la transferencia de animales hacia las categorías inmediatas superiores.

Como resultado del movimiento de rebaño, se debe asegurar la transferencia o traslado de categorías inferiores hacia categorías inmediatas superiores de todo el rebaño, lo que se puede expresar en forma general, mediante la relación de los elementos siguientes, donde para determinar las cabezas de ganado correspondientes a la categoría de edad y sexo i , transferidas a categorías inmediatas superiores, X_{i+2n} ; es necesario sustraerle a la existencia del rebaño al inicio del año, a_i ; los animales destinados a la venta o al sacrificio, $\frac{a_i X_i}{100}$; y los animales que desaparezcan por las diferentes causas de pérdida, como mortalidad, $\frac{a_i m_i}{100}$.

La restricción de tipo III se considera para las categorías de edad y sexo: novillas, añojas, terneras hasta un año, añojos, toretes y teneros hasta un año.

La relación de los elementos anteriores se puede expresar mediante la ecuación siguiente:

$$X_{i+2n} = a_i - \frac{a_i X_i}{100} - \frac{a_i m_i}{100}; i = 2,3,4,9,10,11$$

Considerar como término independiente las magnitudes conocidas y constantes por categoría de edad y sexo i e incorporando las variables de

desviación d_{2i}^+ y d_{2i}^- , se obtiene la restricción de meta (restricción de tipo tres).

$$\frac{a_i X_i}{100} + X_{i+2n} + d_{2i}^- - d_{2i}^+ = a_i - \frac{a_i m_i}{100}; \quad i = 2, 3, 4, 9, 10, 11$$

La restricción es necesaria para asegurar la transferencia de animales hacia las categorías inmediatas superiores, no incurriendo en redundancias; las restricciones de este tipo acotan el conjunto convexo de las soluciones posibles y su eliminación variaría el conjunto de la solución satisfactoria.

Restricción de tipo IV, para garantizar un determinado volumen o cantidad de animales al finalizar el período de planificación anual. En algunas de las categorías, se puede utilizar la siguiente restricción:

$$X_{i+3n} \geq g_i; \quad i = 1, 2, \dots, n$$

$$X_{i+3n} \leq g'_i; \quad i = 1, 2, \dots, n$$

Aplicación de la restricción a las categorías de vacas y toros.

Aplicación a la categoría vaca:

Con vista a garantizar el cumplimiento del plan de producción de leche previsto para el año, es necesario calcular el rebaño de vacas para lograr dicho propósito, es decir, la cantidad mínima de vacas necesarias para asegurar la producción de leche. Se plantea la expresión matemática que permite calcular el plan de producción de leche para el año, P , mediante la media entre el número de cabezas del rebaño de vaca al inicio del año a_i , y el número de cabezas del rebaño de vaca al finalizar el año X_{i+3n} multiplicado por la producción de litros promedio de leche por vaca previstos, R , resultando:

$$P = \left(\frac{a_i + X_{i+3n}}{2} \right) \cdot R$$

El número de cabezas del rebaño de vaca al finalizar el año X_{i+3n} , debe asegurar el número de cabezas mínima de vacas necesarias, V , para garantizar la producción de leche planificada de las vacas.

Entonces:

$$P = \left(\frac{a_1 + V}{2} \right) \cdot R \text{ despejando a } V, \text{ tenemos que: } V = \frac{2P}{R} - a_1$$

Por tanto, para garantizar el rebaño final de vacas con el que se podrá cumplir el plan de producción de leche, se expresa por la restricción de tipo cuatro:

$$X_{i+3n} \geq V; i = 1$$

Una aplicación a la categoría de toro:

Si se planifica mantener la misma cantidad de toros de la existencia inicial, entonces la restricción quedaría expresada de la siguiente forma:

$$X_{i+3n} = a_i; i = 6$$

Restricción de tipo V, para garantizar el rebaño planificado al final del período de un año.

Para garantizar la obtención, al final del período de un año, del volumen total del rebaño planificado para el período, se considera la relación matemática donde la suma de las cabezas de ganado por categoría de edad y sexo i , al final del año, X_{i+3n} , debe ser mayor o igual que el número de cabezas del rebaño planificadas total al final del año, b , exceptuando las terneras y terneros crías.

En dependencia de las capacidades productivas de la empresa objeto de estudio, se planifica un porcentaje de crecimiento (o decrecimiento) general del rebaño, con respecto al rebaño total al inicio del año. El crecimiento (o decrecimiento) general del rebaño se puede estimar aplicando el análisis estadístico que se exhibe en el epígrafe 2.5.

El término independiente de la restricción que garantiza el rebaño planificado al final del año se puede plantear aplicando la *estimación por intervalo*. Para proporcionar mayor flexibilidad al modelo de programación de meta, se pueden utilizar los límites de la estimación por intervalos de la siguiente forma:

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} \geq a + L_{\text{inferior}}, y$$

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} \leq a + L_{\text{superior}}$$

Se considera como término independiente de la restricción V, el número de cabezas planificadas total del rebaño al final del año b , calculándose como $b = a + L_{\text{superior}}$. Además, a la restricción se le incorpora la variable de desviación d_3^- , obteniéndose la restricción de meta (restricción de tipo cinco):

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} + d_3^- = b$$

donde: $b = a + L_{\text{superior}}$

Restricción de tipo VI para garantizar la correspondencia entre la transferencia de animales de una categoría a otra.

En el movimiento de rebaño es trascendental la relación que debe de existir entre las cantidades de animales transferidos a categorías inmediatas superiores y las cantidades de animales recibidas de categorías inferiores. La relación existente entre estos elementos del movimiento de rebaño se muestra a continuación:

$X_{13} = X_{26}$ Garantiza la transferencia de novillas para vacas.

$X_{14} = X_{27}$ Garantiza la transferencia de añojas para novillas.

$X_{15} = X_{28}$ Garantiza la transferencia de terneras hasta un año para añojas.

$X_{16} = X_{29}$ Garantiza la transferencia de terneras crías para terneras hasta un año.

$X_{21} = X_{34}$ Garantiza la transferencia de añojos para toretes.

$X_{22} = X_{35}$ Garantiza la transferencia de terneros hasta un año para añojos.

$X_{23} = X_{36}$ Garantiza la transferencia de terneros crías para terneros hasta un año.

$X_{18} + X_{19} + X_{20} = X_{33}$ Garantiza la transferencia de toretes para toros, bueyes y sementales.

De manera general, la relación que representa la correspondencia entre la transferencia de animales de una categoría a otra, se muestra mediante la restricción siguiente:

$$X_{i+n} = X_{(i+2n)+1}; i = 1,2,3,4,9,10,11$$

$\sum_{i=6,7,8} X_{i+n} - X_{33} = 0$; para la transferencia de toretes a las categorías superiores.

Considerando las variables en el miembro de la izquierda, se obtiene la restricción de tipo seis:

$$X_{i+n} - X_{(i+2n)+1} = 0; i = 1,2,3,4,9,10,11$$

$\sum_{i=6,7,8} X_{i+n} - X_{33} = 0$; para la transferencia de toretes a las categorías superiores.

Restricción de tipo VII, para regular la extracción de animales para carne o ventas, dentro de cada una de las categorías.

Este tipo de restricción se utiliza si es necesario regular dentro de determinado rango o intervalo la extracción de ganado que se destine para carne o ventas, por cada una de las diferentes categorías.

La restricción se plantea estableciendo una relación entre las cabezas del rebaño destinadas a la venta o al sacrificio por categoría de edad y sexo i , X_i , con respecto a porcentajes mínimos y máximos del ganado extraído para carne o venta establecidos por categoría de edad y sexo i , h_i y h'_i , respectivamente. Generándose las siguientes restricciones:

$$X_i \geq h_i; i = 1,2,\dots,n$$

$$X_i \leq h'_i ; i = 1, 2, \dots, n$$

Este tipo de restricciones garantiza en el modelo la salida de cabezas de ganado por cada categoría y el volumen de producción de carne.

Restricción de tipo VIII, para garantizar la correlación estructural necesaria entre distintas categorías.

Este tipo de restricción expresa las correlaciones estructurales entre las diferentes categorías de edad y sexo del rebaño, lo que es necesario para asegurar una determinada estructura del rebaño al final del año, que posibilite el adecuado proceso de reproducción para los próximos períodos. Este tipo de restricción permite vincular los volúmenes finales de cada categoría del rebaño entre sí.

La restricción se representa estableciendo una relación entre las cabezas de ganado por categoría de edad y sexo, al final del año, X_{i+3n} , con respecto al producto de t' (valor máximo de proporcionalidad de la correlación estructural entre la categoría de edad y sexo superior $(i+3n)$ con respecto a la categoría de edad y sexo inmediata inferior $[(i+3n)-1]$ del rebaño, al final del año), por $X_{(i+3n)-1}$. Se genera la siguiente restricción:

$$X_{i+3n} \geq t' X_{(i+3n)-1} ; i = 1, 2, \dots, n$$

Considerando las variables de decisión en el miembro de la izquierda e incorporando las variables de desviación d_{4i}^+ y d_{4i}^- , se obtiene la restricción de meta (restricción de tipo ocho):

$$X_{i+3n} - t' X_{(i+3n)-1} + d_{4i}^- - d_{4i}^+ = 0 ; i = 1, 2, \dots, n$$

A continuación, algunas situaciones en que se puede utilizar este tipo de restricción:

- *Correlación entre vacas y novillas*: las novillas en el rebaño deben tener un volumen que satisfaga las necesidades anuales para el reemplazo de las vacas y garantice el crecimiento del rebaño de las vacas.
- *Correlación entre novillas, añojas y terneras hasta un año*: para una buena selección de las novillas se debe tener suficiente cantidad de añojas y terneras hasta un año.

- *Correlación entre toros, añojos y terneros hasta un año*: los añojos en el rebaño deben tener un volumen que satisfaga las necesidades anuales para el reemplazo de los toros y garantice el crecimiento del rebaño de los toros. Para asegurar los añojos, es necesario contar con terneros de hasta un año que los sustituyan.

Restricción de tipo IX, para garantizar determinada cantidad de animales comprados a terceros el período de planificación anual, se pueden utilizar las restricciones siguientes:

$$X_{i+4n} \geq d_i; \quad i = 1, 2, \dots, n$$

$$X_{i+4n} \leq d'_i; \quad i = 1, 2, \dots, n$$

Restricción de tipo X, para garantizar el cumplimiento de la producción mercantil de carne en pie del rebaño, a la suma de los términos que representan la producción mercantil de carne en pie de todas las categorías de edad y sexo i , se le incorporan las variables de desviación d_s^+ y d_s^- e igualándolo a la producción mercantil de carne en pie del rebaño prevista, se

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} \frac{a_i X_i}{100} p_i + \sum_{i=5,12} \frac{c_i X_i}{100} p_i + d_s^- - d_s^+ = B$$

Función Objetivo.

Se construye considerando las desviaciones definidas y el orden de prioridad establecido con los decisores, tales como: *primero*, garantizar producción mercantil de carne en pie del rebaño (P_1); *segundo*, garantizar el plan de cabezas del rebaño por categoría de edad y sexo al final del año (P_2); *tercero*, garantizar la adecuada correlación estructural entre distintas categorías (P_3); *cuarto*, garantizar el adecuado movimiento de las categorías por edad y sexo del rebaño (P_4), y *quinto*, garantizar la transferencia de animales hacia las categorías inmediatas superiores (P_5).

$$\text{Mínimo } Z = P_1(d_s^- + d_s^+) + P_2 d_3^- + P_3(d_{4i}^- + d_{4i}^+) + P_4(d_{1i}^- + d_{1i}^+) + P_5(d_{2i}^- + d_{2i}^+)$$

2.4. Determinación y validación matemática de parámetros del modelo

Después de haber detallado el planteamiento matemático del modelo económico-matemático de meta con prioridad y previo a la conformación del modelo numérico, se muestra:

- la estimativa y la inferencia de la función de regresión, proceso en que se determina y valida la cantidad de cabezas del rebaño planificadas total al final del año (b), con el propósito de robustecer la solución del modelo, y
- el análisis estadístico, para comprobar la correspondencia entre la estructura satisfactoria obtenida y la planificada del rebaño para el año.

2.4.1. Estimativa e inferencia de la función de regresión

Para realizar la estimativa y la inferencia de la ecuación de regresión, se tienen en cuenta los siguientes aspectos, que se encuentran desarrollados teóricamente en este epígrafe:

- I- Estimativa de la función de regresión. Estimadores mínimos cuadráticos.
- II- Validación de la función de regresión estimada. Prueba F de Fisher y prueba t de Student.
- III- Coeficiente de determinación, coeficiente de correlación y prueba de hipótesis en el análisis de correlación.
- IV- Pruebas de los supuestos de la función de regresión estimada.
- V- Análisis del intervalo de confianza para β_j .
- VI- Validación de la función de regresión estimada a través del intervalo de confianza para β_2 .

En el estudio teórico y práctico de la estimativa, inferencia y validación de la función de regresión, se toma como fundamento los procedimientos metodológicos abordados en los temas de econometría relacionados con la regresión lineal simple y con los supuestos básicos, desarrollados por Gujarati (2003); Wooldridge (2008); Gujarati & Porter (2010); y Espallarga & Solís (2012).

I- Estimativa de la función de regresión. Estimadores mínimos cuadráticos

El propósito de obtener los estimadores mínimos cuadráticos consiste en determinar la cantidad de cabezas del rebaño planificadas total al final del

año (b). Se procede como sigue:

- *Primeramente*, se estima el ritmo de crecimiento o decrecimiento de la masa ganadera ($\hat{\beta}_2$) estimador mínimo cuadrático o coeficiente de regresión de la función de regresión lineal estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$.
- *Segundo*, a partir de $\hat{\beta}_2$ se determina cantidad de cabezas del rebaño planificadas total al final del año (b), como $b = a + L_{\text{superior}}$.

Primero: determinación de los estimadores mínimos cuadráticos de la función de regresión.

A partir de una muestra simple aleatoria de la serie histórica del comportamiento anual de la cantidad total de animales del rebaño, es posible ajustar una función continua representada por la función de regresión lineal estimada, $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$; el signo de $\hat{\beta}_2$ depende del comportamiento de la serie de los datos ajustados, donde, \hat{Y}_t es la cantidad de cabezas del rebaño total al final del año, estimada por la función de regresión lineal en el período t (en cabezas), representa la variable dependiente o regresada; X_t es la variable independiente o explicativa tiempo (en año); $\hat{\beta}_1$ y $\hat{\beta}_2$ son los estimadores mínimos cuadráticos, siendo, $\hat{\beta}_1$ el intercepto con el eje Y o valor que toma Y como promedio, cuando $X = 0$, y $\hat{\beta}_2$ es la pendiente de la recta.

Los estimadores mínimos cuadráticos de la función de regresión, $\hat{\beta}_1$ y $\hat{\beta}_2$, se determinan por las expresiones matemática obtenidas a partir del sistema de ecuaciones normales, como se expone a continuación.

Aplicando el Método de los Mínimos Cuadrados Ordinarios para realizar el ajuste al modelo $E_{(Y_t/X_t)} = Y_t = \beta_1 + \beta_2 X_t + U_t$, conduce a un sistema de ecuaciones normales de dos ecuaciones con dos incógnitas, como:

$$\sum_{i=1}^n Y_i - n\beta_1 - \beta_2 \sum_{i=1}^n X_i = 0$$

$$\sum_{i=1}^n X_i Y_i - \beta_1 \sum_{i=1}^n X_i - \beta_2 \sum_{i=1}^n X_i^2 = 0$$

El sistema de ecuaciones normales tendrá tantas ecuaciones e incógnitas como parámetros tenga el modelo.

Resolviendo el sistema de ecuaciones normales, se obtienen los estimadores mínimos cuadrados de β_1 y β_2 , representados por $\hat{\beta}_1$ y $\hat{\beta}_2$; las expresiones resultantes son las siguientes:

- Para el estimador mínimo cuadrático $\hat{\beta}_1$.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n Y_i}{n} - \hat{\beta}_2 \frac{\sum_{i=1}^n X_i}{n}$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

- Para el estimador mínimo cuadrático $\hat{\beta}_2$.

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2}$$

Realizando algunas transformaciones, $\hat{\beta}_2$ se puede expresar por las siguientes expresiones:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_i Y_i - \frac{\sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n}}{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}} \quad \text{o} \quad \hat{\beta}_2 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Se puede comprobar que los puntos críticos de la función $\hat{\beta}_1$ y $\hat{\beta}_2$, son estimadores que cumplen con la condición de mínimo, constituyendo estimadores *mínimos cuadrados*. Por tanto, la función $\hat{E}_{(Y_t|X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$ también cumplirá con la condición de mínimo, llamándose *función de regresión mínimo cuadrática*. Dicha función proporciona una estimación del valor medio esperado de Y , para cada valor fijo de X .

Los estimadores mínimos cuadráticos de la función de regresión $\hat{E}_{(X/Y)} = \hat{Y} = \hat{\beta}_1 \pm \hat{\beta}_2 X$ se pueden determinar por la vía matricial, empleando el mismo procedimiento desarrollado en el epígrafe 1.6.1, que es válido para cualquier modelo que presente linealidad con respecto a los parámetros. En el epígrafe 2.5 se presenta una situación práctica para la determinación de los estimadores mínimos cuadráticos de la función de regresión $\hat{E}_{(X/Y)} = \hat{Y} = \hat{\beta}_1 \pm \hat{\beta}_2 X$, por la vía matricial.

El estimador mínimo cuadrático o coeficiente de regresión $\hat{\beta}_2$ es una constante que expresa la tasa media de variación de \hat{Y}_t al producirse un incremento unitario en X_t :

- Si $\hat{\beta}_2 > 0$, la función es creciente, y \hat{Y}_t aumenta en $\hat{\beta}_2$ unidades, por cada unidad que aumente X_t .
- Si $\hat{\beta}_2 < 0$, la función es decreciente, y \hat{Y}_t disminuye en $|\hat{\beta}_2|$ unidades, por cada unidad que aumente X_t .

La función de regresión $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$, para la situación objeto de estudio, está definida solo para los números reales positivos, o sea, el **Dom f** = $\{x \in \mathfrak{R} : x > 0\}$ y la función toma solamente valores positivos, es decir, la **Im f** = $\{y \in \mathfrak{R} : y > 0\}$.

Segundo: a partir de $\hat{\beta}_2$ se determina la cantidad de cabezas del rebaño planificadas total al final del año (b).

La cantidad (número) de cabezas del rebaño planificadas total al final del año (b) constituye el término independiente de la restricción de tipo V del modelo

de Programación de Meta Lineal,
$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} + d_3^- = b,$$

restricción encargada de garantizar el rebaño planificado al final del período de un año, siendo $b = a + L_{\text{superior}}$, donde: L_{superior} es el límite superior del intervalo de confianza para $\hat{\beta}_2$, y a es la cantidad (número) total de

cabezas del rebaño a inicio del año.

Para la estimación y validación del *modelo de regresión lineal simple* y *por tanto del ritmo de crecimiento o decrecimiento del rebaño* ($\hat{\beta}_2$), se utilizan las bondades que proporciona la estimativa y la inferencia del análisis de regresión. Para determinar el modelo matemático de mejor ajuste, se tienen en cuenta: la prueba de hipótesis de F de Fisher, t de Student; así como las diferentes pruebas de los supuestos básicos de autocorrelación, homocedasticidad y normalidad. La estimación y la inferencia del modelo se realiza mediante cálculos manuales y mediante la aplicación del paquete de programa Eviews versión 3.1, y del paquete de programa estadístico SPSS versión 23.0, que posibilitan obtener los resultados estadísticos de los modelos matemáticos a verificar.

II- Validación de la función de regresión estimada. Prueba F de Fisher y prueba t de Student

Verificando la validez de la función de regresión estimada, mediante la prueba F de Fisher

Planteada la relación funcional $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$ donde se supone la posible relación funcional lineal entre la variable explicativa y la variable respuesta, dicho supuesto debe ser verificado. La prueba de hipótesis para comprobar este tipo de supuesto se obtiene mediante el análisis de varianza en la regresión, que permite plantear una prueba de hipótesis para determinar la validez del ajuste realizado (Espallarga & Solís, 2012).

Se conoce que el análisis de varianza, desde el punto de vista de la regresión, se basa en la descomposición de la suma de cuadrados total; asociado a la suma de cuadrados se encuentran sus grados de libertad, donde para la relación funcional considerada presenta las siguientes características:

- SC_T tiene $n-1$ grados de libertad.
- SC_R tiene $n-p$ grados de libertad, para el caso de una relación funcional de dos variables, existen $n-2$ grados de libertad.
- SC_E tiene $p-1$ grados de libertad, para el caso de una relación funcional de dos variables, existe 1 grado de libertad.

donde:

n es el número de observaciones.

p es el número de parámetros que se estiman.

Como se ha expresado, para verificar la validez de la función de regresión estimada, se hace a partir del análisis de varianza en la regresión, cuyo resultado para una función de regresión de dos variables se muestra en la siguiente tabla de Análisis de Varianza (ANOVA).

Tabla 2.2. Análisis de Varianza (ANOVA) para una función de regresión de dos variables.

Fuente de Variación	Suma de Cuadrado SC	Grados de libertad $g/$	Suma de Cuadrados Medios SCM
Debido a la regresión (SC_E)	$SC_E = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$ $SC_E = \hat{\beta}_2 \left(\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y} \right)$	$\frac{(p-1)}{1}$	$SCM \text{ de } SC_E = \frac{SC_E}{1} = S_E^2$
Debido a los residuos (SC_R)	$SC_R = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ $SC_R = SC_T - SC_E$	$\frac{(n-p)}{n-2}$	$SCM \text{ de } SC_R = \frac{SC_R}{n-2} = S_R^2$
Total SC_T	$SC_T = \sum_{i=1}^n (Y_i - \bar{Y})^2$ $SC_T = \sum_{i=1}^n Y_i^2 - n\bar{Y}^2$	$(n-1)$	$//////$

Fuente: elaborada por los autores.

donde:

SC_E es la suma de cuadrados explicada por la regresión.

SC_R es la suma de cuadrados debido a los residuos (al error).

$p-1$ son los grados de libertad del numerador.

$n-p$ son los grados de libertad del denominador.

S_E^2 es la varianza explicada por la regresión.

S_R^2 es la varianza debido a los residuos.

Partiendo del supuesto de varianza constante en las Y_t , se tiene que la varianza debido a los residuos (S_R^2) es un estimador insesgado, y la varianza explicada por la regresión (S_E^2) solo será un estimador insesgado si $\beta_2 = 0$.

Cuando $\beta_2 = 0$, las variables X y Y son independientes, ya que si $Y_t = \beta_1 + \beta_2 X_t$, y $\beta_2 = 0$, entonces la variable X no explica el comportamiento de la variable dependiente Y .

Para determinar la precisión de la función de regresión lineal estimada, se emplea la prueba de hipótesis F de Fisher, que se apoya en la descomposición de la variación total. Comprobando si $\beta_2 = 0$, se puede validar el modelo estimado, planteando las siguientes hipótesis para verificar el supuesto de relación lineal entre las variables, siendo la formulación de la prueba de hipótesis $H_0: \beta_2 = 0$ versus $H_1: \beta_2 \neq 0$; donde: β_2 representa el valor paramétrico de la ecuación de regresión estimada.

La hipótesis nula, $H_0: \beta_2 = 0$, es la hipótesis de las diferencias nulas, la que representa la igualdad, siendo la que se somete a prueba, por tanto, es la que se rechaza o no se rechaza. La hipótesis alternativa $H_1: \beta_2 \neq 0$, representa la hipótesis de la investigación, lo que se va a contrastar, es la hipótesis que deberá ser aceptada si la hipótesis nula es rechazada.

A partir de los resultados del análisis de varianza de la regresión, se determina el estadístico de prueba o estadígrafo F .

El valor de F calculado, estimado, obtenido u observado, es

$$F = \frac{S_E^2}{S_R^2} \quad \text{o} \quad F = \frac{SC_E}{p-1} \bigg/ \frac{SC_R}{n-p} = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R}$$

La región crítica para el estadígrafo F , dado el nivel de significación α , se encuentra dada por $W = \{F \in \mathfrak{R} : F > F_{1-\alpha}(p-1, n-p)\}$.

El valor crítico, tabulado o teórico del estadístico F está definido por $F_{1-\alpha}(p-1, n-p)$.

Si el valor obtenido o calculado del estadístico de prueba $F = S_E^2/S_R^2$

es mayor que el valor crítico o teórico $F_{1-\alpha}(p-1, n-p)$ se rechaza la hipótesis nula $\beta_2 = 0$ al nivel de significación establecido, por tanto, no existen razones para rechazar el supuesto de linealidad entre la variable dependiente y la explicativa, y se acepta la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$ como buena predictora de la función de regresión poblacional $E_{(Y_t/X_t)} = Y_t = \beta_1 \pm \beta_2 X_t$, y que el ajuste hecho es bueno. Entonces, la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$ se puede utilizar para predecir el valor de la variable dependiente Y , a partir de valores conocidos de la variable explicativa X .

Alternativamente, si la Prob (F -statistic) o valor p es menor que la probabilidad de rechazar la hipótesis nula siendo cierta, o probabilidad de cometer el error de tipo I (nivel de significación α), se rechaza igualmente la hipótesis nula.

Resumiendo la prueba de hipótesis.

Formulación de la hipótesis nula y alternativa:

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

Región crítica para el estadígrafo t , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} : F > F_{1-\alpha}(p-1, n-p) \right\}.$$

$$\text{donde: } F = \frac{S_E^2}{S_R^2} \text{ o } F = \frac{\frac{SC_E}{p-1}}{\frac{SC_R}{n-p}} = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R}.$$

Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha}(p-1, n-p)$.
- No rechazar H_0 , si $F \leq F_{1-\alpha}(p-1, n-p)$.

Alternativamente:

- Rechazar H_0 , si la Prob (F -statistic) o valor $p < \alpha$.
- No rechazar H_0 , si la Prob (F -statistic) o valor $p > \alpha$.

La decisión se puede tomar indistintamente con el estadístico de prueba o estadígrafo F obtenido del análisis de varianza, o con la Prob (F -statistic) o valor p .

Verificando la validez de la función de regresión estimada mediante la prueba t de Student

Para verificar la confiabilidad de la función de regresión ajustada y la validez del ajuste realizado, también se emplea la prueba de hipótesis t de Student sobre el coeficiente de regresión β_2 mediante una prueba bilateral o de dos colas.

Formulación de la hipótesis nula y alternativa:

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

Región crítica para el estadístico de prueba o estadígrafo t , dado el nivel de significación α , es

$$W = \left\{ t \in \mathcal{R} : t < -t_{1-\frac{\alpha}{2}}(n-p) \quad \text{o} \quad t > t_{1-\frac{\alpha}{2}}(n-p) \right\}.$$

El estadístico de prueba se calcula por $t = t_{\hat{\beta}_2} = \frac{\hat{\beta}_2}{S_{\hat{\beta}_2}}$, siendo

$$S_{\hat{\beta}_2} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n (X_i - \bar{X})^2}} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}}$$

Regla de decisión:

- Rechazar H_0 , si $t < -t_{1-\frac{\alpha}{2}}(n-p)$ o $t > t_{1-\frac{\alpha}{2}}(n-p)$.

- No rechazar H_0 , si $t \geq -t_{1-\frac{\alpha}{2}}(n-p)$ o $t \leq t_{1-\frac{\alpha}{2}}(n-p)$,

o sea, $-t_{1-\frac{\alpha}{2}}(n-p) \leq t \leq t_{1-\frac{\alpha}{2}}(n-p)$.

Alternativamente:

- Rechazar H_0 , si la Prob (t -statistic) o valor $p < \alpha$.

- No rechazar H_0 , si la Prob (t -statistic) o valor $p > \alpha$.

donde:

S_R^2 es la varianza debido a los residuos.

$S_{\hat{\beta}_2}$ es el error estándar asociado al estimador mínimo cuadrático $\hat{\beta}_2$.

p es el número de parámetros de la función ($p = 2$).

Si el valor obtenido o calculado del estadístico de prueba $t_{\beta_2} = \hat{\beta}_2 / S_{\beta_2}$ es menor que valor t crítico o teórico $-t_{1-\frac{\alpha}{2}}(n-2)$ o mayor que el valor crítico o teórico $t_{1-\frac{\alpha}{2}}(n-2)$, se rechaza la hipótesis nula $\beta_2 = 0$ al nivel de significación establecido. Por tanto, no existen razones para rechazar el supuesto de linealidad entre la variable dependiente y la explicativa, y se acepta la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$ como buena predictora de la función de regresión poblacional $E_{(Y_t/X_t)} = Y_t = \beta_1 \pm \beta_2 X_t$, y que el ajuste hecho es bueno. Entonces, la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$ se puede utilizar para predecir el valor de la variable dependiente Y , a partir de valores conocidos de la variable explicativa X .

Para tomar la decisión, al igual que en la prueba de F , también se puede utilizar alternativamente el valor p asociado al t -statistic de $\hat{\beta}_2$.

Las pruebas de t y F proporcionan dos formas alternas, pero complementarias, de probar la hipótesis nula $\beta_2 = 0$ en los modelos con una variable explicativa.

No obstante, en los modelos de regresión múltiple, la prueba F tiene varias aplicaciones que hacen que sea un método útil y poderoso para demostrar hipótesis estadísticas (Gujarati, 2003).

III- Coeficiente de determinación, coeficiente de correlación y prueba de hipótesis en el análisis de correlación

Coeficiente de determinación

Se conoce del epígrafe 1.6.1, que a partir de la descomposición de la suma de cuadrados, $SC_T = SC_E + SC_R$, y dividiendo cada término entre la SC_T se obtiene $1 = \frac{SC_E}{SC_T} + \frac{SC_R}{SC_T}$, donde el coeficiente de determinación viene dado por:

$$R^2 = \frac{SC_E}{SC_T} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SC_R}{SC_T}$$

El coeficiente de determinación también se puede expresar en función de las fórmulas abreviadas de las sumas de cuadrados,

$$SC_E = \hat{\beta}_2 \left(\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y} \right) \text{ y } SC_T = \sum_{i=1}^n Y_i^2 - n\bar{Y}^2, \text{ obteniéndose}$$

$$R^2 = \frac{\hat{\beta}_2 \left(\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y} \right)}{\sum_{i=1}^n Y_i^2 - n\bar{Y}^2}$$

Coeficiente de correlación lineal

El coeficiente de correlación es una cantidad estrechamente relacionada con el coeficiente de determinación, pero conceptualmente diferente, ya que es una medida del grado de asociación entre la variable dependiente y explicativa (Gujarati, 2003).

El coeficiente de correlación lineal se calcula, a partir de su definición, como:

$$r_{xy} = \frac{\sum_{i=1}^n X_i \cdot Y_i - \frac{\left(\sum_{i=1}^n X_i\right)\left(\sum_{i=1}^n Y_i\right)}{n}}{\sqrt{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}} - \sqrt{\sum_{i=1}^n Y_i^2 - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n}}$$

El coeficiente de correlación lineal se calcula a partir del coeficiente de determinación R^2 , como: $r_{xy} = \pm\sqrt{R^2}$

A continuación, algunas propiedades del coeficiente de correlación, teniendo en cuenta lo expuesto por Gujarati (2003); y Gujarati & Porter (2010):

- Puede ser positivo o negativo, en dependencia del signo del término en el numerador de la expresión de su definición.
- Se encuentra entre los límites de -1 y +1; es decir, $-1 \leq r \leq +1$.
- Es simétrico por naturaleza, es decir, el coeficiente de correlación entre X y Y , r_{xy} ; es lo mismo que el coeficiente de correlación entre Y y X , r_{yx} .
- Si X y Y son estadísticamente independientes, el coeficiente de correlación entre ellos es cero. Si $r_{xy} = 0$, esto no significa que las dos variables sean independientes.
- Es una medida de asociación lineal o de independencia lineal solamente; su uso en la descripción de relaciones no lineales no tiene significado.

Prueba de hipótesis en el análisis de correlación

El planteamiento de la prueba de hipótesis paramétrica se realiza con respecto al coeficiente de regresión poblacional $\rho_{(x,y)}$, siendo la hipótesis nula, $H_0: \rho_{(x,y)} = 0$, y la hipótesis alternativa a contrastar, $H_1: \rho_{(x,y)} \neq 0$.

La hipótesis nula indica que no existe correlación lineal entre las variables poblacionales, pues el coeficiente de correlación poblacional es igual a cero, y la hipótesis alternativa expresa que existe relación lineal entre las variables poblacionales.

Con la hipótesis, se pretende comprobar, de la población de referencia,

el grado de asociación y relación lineal entre la cantidad de cabezas del rebaño total al final del año, variable dependiente o regresada Y , y el tiempo, variable explicativa X , y para esto, se verifica si el coeficiente de correlación poblacional ($\rho_{(x,y)}$) es igual o distinto de cero, a partir de la información correspondiente al comportamiento tendencial de la cantidad de cabezas, en al menos quince años. La prueba que se formula es una prueba bilateral o de dos colas, es decir, se está en presencia de una alternativa bilateral.

La prueba de hipótesis se realiza mediante la prueba de significación, utilizando el estadístico de prueba o estadígrafo t de Student, con $n-p=n-2$

grados de libertad. El valor del estadígrafo t se calcula en función del

coeficiente de correlación lineal r : $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ (Espallarga & Solís, 2012).

La región crítica o de rechazo para el estadígrafo t , dado el nivel de significación α , se encuentra dada por:

$$W = \left\{ t \in \mathcal{R} : t < -t_{1-\frac{\alpha}{2}}(n-p) \quad \text{o} \quad t > t_{1-\frac{\alpha}{2}}(n-p) \right\}.$$

El valor crítico, tabulado o teórico del estadístico t está definido por el percentil

$t_{tab} = t_{1-\frac{\alpha}{2}}(n-p)$ de orden $1 - \frac{\alpha}{2}$ de la distribución t de Student, con $n-p=n-2$ grados de libertad.

donde:

r es el coeficiente de correlación lineal muestral.

n es el número de observaciones.

$p = 2$ es el número de parámetros de la función.

Regla de decisión:

- Si, $t < -t_{1-\frac{\alpha}{2}}(n-p)$ o $t > t_{1-\frac{\alpha}{2}}(n-p)$: Se rechaza la hipótesis nula, es decir, existe correlación significativa entre las variables dependiente y explicativa, para el nivel de significación α establecido. El valor de t

calculado se sitúa en la región crítica.

- Si, $t \geq -t_{1-\frac{\alpha}{2}}(n-p)$ o $t \leq t_{1-\frac{\alpha}{2}}(n-p)$, o sea,

$-t_{1-\frac{\alpha}{2}}(n-p) \leq t \leq t_{1-\frac{\alpha}{2}}(n-p)$: No se rechaza la hipótesis nula. Por tanto; no existe correlación significativa entre las variables dependiente y explicativa para el nivel de significación α establecido.

Como resultado de aplicar la prueba de hipótesis, se comprobó que se rechaza la hipótesis nula, por tanto, existe una relación significativa entre las variables dependiente o regresada Y , y la variable explicativa X , para un nivel de significación de 0,05.

Resumiendo la prueba de hipótesis.

Formulación de la hipótesis nula y alternativa:

$$H_0: \rho_{(x,y)} = 0$$

$$H_1: \rho_{(x,y)} \neq 0$$

Región crítica para el estadístico de prueba o estadígrafo t , dado el nivel de significación α , es

$$W = \left\{ t \in \mathfrak{R} : t < -t_{1-\frac{\alpha}{2}}(n-2) \quad \text{o} \quad t > t_{1-\frac{\alpha}{2}}(n-2) \right\}.$$

$$\text{donde: } t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Regla de decisión:

- Rechazar H_0 , si $t < -t_{1-\frac{\alpha}{2}}(n-2)$ o $t > t_{1-\frac{\alpha}{2}}(n-2)$.

- No rechazar H_0 , si $t \geq -t_{1-\frac{\alpha}{2}}(n-2)$ o $t \leq t_{1-\frac{\alpha}{2}}(n-2)$,

$$\text{o sea, } -t_{1-\frac{\alpha}{2}}(n-2) \leq t \leq t_{1-\frac{\alpha}{2}}(n-2) .$$

Determinación del coeficiente de correlación lineal muestral

Para calcular a $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ en la prueba de hipótesis del análisis de correlación, primeramente se determina el coeficiente de correlación lineal muestral r , a partir de su definición o del coeficiente de determinación.

IV- Pruebas de los supuestos de la función de regresión estimada

Para verificar la validez de las pruebas estadísticas F de Fisher y t de Student, así como la eficiencia de la estimación obtenida a partir de la función de regresión estimada, se procede a verificar los supuestos básicos del modelo, aplicando procedimientos semejantes a los desarrollados en el epígrafe 1.6.2 del capítulo 1, mediante la comprobación de las siguientes pruebas:

- Para detectar la presencia de normalidad, se emplean las pruebas Jarque-Bera y Shapiro-Wilks.
- Para detectar la presencia de no autocorrelación, se aplican las pruebas de Durbin-Watson y Breusch-Godfrey.
- Para detectar la presencia de homocedasticidad, se utilizan las pruebas de Goldfeld-Quandt y la general de heterocedasticidad de White.

De ser necesario, se realizan las correcciones correspondientes con el objetivo de lograr el cumplimiento de la prueba analizada, y de esta forma, seleccionar la mejor función de regresión que sirva de base para el pronóstico.

Si como resultado de aplicar al parámetro *cantidad de cabezas del rebaño total al final del año*, las pruebas: Jarque-Bera, Shapiro-Wilks, Durbin-Watson, Breusch-Godfrey, Goldfeld-Quandt y White, se comprueba que existe normalidad en la serie de valores de la variable dependiente o regresada y en los residuos, que no existe autocorrelación, y que existe homocedasticidad en las varianzas del error, para el nivel significación seleccionado, se concluye

que la función de regresión estimada representada por la función lineal $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 \pm \hat{\beta}_2 X_t$ constituye el modelo matemático adecuado y permitirá realizar las proyecciones correctas y eficientes de la cantidad de cabezas del rebaño.

Para hallar el comportamiento del ritmo de crecimiento o decrecimiento anual del rebaño, se toman como referencia los últimos quince años, con el propósito de poder verificar las pruebas de hipótesis de referencia.

V- Análisis del intervalo de confianza para β_j

Para proporcionarle mayor flexibilidad al modelo de programación de meta lineal con prioridades, se realizan inferencias a partir de la estimación del intervalo de confianza al coeficiente de regresión, que expresa el ritmo de crecimiento o decrecimiento del rebaño al final del año, $\hat{\beta}_2$.

El intervalo de confianza es uno de los procedimientos generales de la inferencia estadística que puede aplicarse a los problemas de análisis de regresión, para realizar estimaciones y para el contraste de hipótesis.

En el contexto de estimar el parámetro poblacional, el intervalo de confianza representado por $\hat{\beta}_j \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_j}$, y para la situación objeto de estudio $\hat{\beta}_2 \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_2}$, es un rango de valores, calculado a partir de los datos de una muestra, en el cual podría encontrarse el verdadero valor del parámetro, para el nivel o coeficiente de confianza considerado (Gujarati & Porter, 2010).

El nivel de confianza es la probabilidad de que un intervalo calculado a partir de una muestra aleatoria contenga al verdadero valor del parámetro.

El término independiente de la restricción b, número de cabezas del rebaño planificadas total al final del año, es el valor resultante de la suma algebraica de la existencia real inicial del rebaño, a ; y el límite superior del intervalo de confianza para $\hat{\beta}_2$, como se exhibe en la restricción V del análisis de la construcción del modelo económico-matemático en el epígrafe 2.3, donde se

obtiene la restricción
$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} \leq a + L_{\text{superior}}.$$

Se considera como término independiente de la restricción, el número de cabezas del rebaño planificadas total al final del año b , que se calcula como $b = a + L_{\text{superior}}$. Además, a la restricción se le incorpora la variable de desviación d_3^- , obteniéndose la siguiente restricción de meta:

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} + d_3^- = b$$

donde: $b = a + L_{\text{superior}}$

VI- Validación de la función de regresión estimada a través del intervalo de confianza para β_2

La verificación de la validez de la función de regresión estimada es posible mediante el desarrollo de pruebas de hipótesis, a partir del intervalo de confianza para β_2 , planteándose como hipótesis $H_0: \beta_2 = 0$, versus, $H_1: \beta_2 \neq 0$.

Se desarrolla una prueba de hipótesis bilateral o de dos colas, como se exhibe a continuación.

Formulación de la hipótesis nula y alternativa:

La hipótesis de la diferencia nula que se desea probar y someter a prueba es $H_0: \beta_2 = 0$ versus la hipótesis alternativa a contrastar $H_1: \beta_2 \neq 0$.

Región crítica q de rechazo para el estimador mínimo cuadrático o coeficiente de regresión β_2 , dado el nivel de significación α , es

$$W = \left\{ \hat{\beta}_2 \in \mathfrak{R} : \hat{\beta}_2 < - t_{1-\frac{\alpha}{2}} (n-p) \hat{\sigma}_{\hat{\beta}_2} \quad \text{o} \quad \hat{\beta}_2 > t_{1-\frac{\alpha}{2}} (n-p) \hat{\sigma}_{\hat{\beta}_2} \right\}.$$

El estimador mínimo cuadrático o coeficiente de regresión se determina por

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2},$$

o se pueden emplear las expresiones obtenidas anteriormente.

Regla de decisión:

- Rechazar H_0 , si $\hat{\beta}_2 < -t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$ o $\hat{\beta}_2 > t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$.
 - No rechazar H_0 , si $\hat{\beta}_2 \geq -t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$ o $\hat{\beta}_2 \leq t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$
- o sea, $-t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2} \leq \hat{\beta}_2 \leq t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$.

El error estándar asociado a $\hat{\beta}_2$ ($\hat{\sigma}_{\hat{\beta}_2}$) se calcula como:

$$\hat{\sigma}_{\hat{\beta}_2} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}, \text{ siendo } S_R^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-p}$$

- Si $-t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2} \leq \hat{\beta}_2 \leq t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$, no se rechaza la

hipótesis nula $\beta_2 = 0$, afirmándose que lo investigado no es estadísticamente significativo, es decir, β_2 no es significativamente diferente de cero.

- Si $\hat{\beta}_2 < -t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$ o $\hat{\beta}_2 > t_{1-\frac{\alpha}{2}}(n-p)\hat{\sigma}_{\hat{\beta}_2}$, se rechaza la

hipótesis nula $\beta_2 = 0$, afirmándose que lo investigado es estadísticamente significativo, es decir, β_2 es significativamente diferente de cero.

Otra forma de someter a prueba la hipótesis nula $\beta_2 = 0$, es a partir de la construcción del intervalo de confianza, donde se plantea como regla de decisión. "Constrúyase un intervalo de confianza para β_2 al $100(1-\alpha)\%$. Si β_2 bajo H_0 se encuentra dentro de este intervalo de confianza, no se rechace H_0 , pero si está por fuera del intervalo, rechace H_0 ". (Gujarati, 2003, p. 122)

Los resultados derivados con la prueba de hipótesis mediante la aplicación del análisis de varianza para la regresión y con la prueba t de Student, son equivalentes al resultado de la prueba de hipótesis mediante el intervalo de confianza para β_2 .

2.4.2. Análisis estadístico para comprobar la correspondencia entre la estructura satisfactoria y la real alcanzada

La comparación entre la estructura del rebaño derivada de la solución satisfactoria del modelo matemático de meta con prioridades y la estructura del rebaño real alcanzada por la empresa para el año, es un aspecto importante en el análisis de la solución satisfactoria obtenida, y que sirve para medir la validez de la proyección realizada, máxime cuando se estudia el comportamiento y evolución de la masa ganadera, que constituye un proceso económico-productivo complejo.

Para comprobar si existe correspondencia entre la estructura satisfactoria, obtenida a partir de la solución del modelo de meta y la estructura real alcanzada, se aplica la prueba estadística de Rango de Wilcoxon, prueba no paramétrica para dos muestras relacionadas.

Para el planteamiento de la prueba, se formula como *hipótesis nula*, hipótesis de igualdad de promedios, que la estructura satisfactoria derivada de la solución del modelo y la estructura real, no difieren. En términos de la prueba de Wilcoxon, la suma de los rangos positivos es igual a la suma de los negativos. La hipótesis alternativa a contrastar expresa que la estructura satisfactoria derivada de la solución del modelo y la estructura real difieren, es decir, la suma de los rangos positivos no equivale a la suma de los negativos.

Como regla de decisión, si el nivel crítico bilateral del estadístico de prueba Wilcoxon, es decir, el valor de la probabilidad de la prueba estadística es mayor que la probabilidad de rechazar la hipótesis nula, siendo cierta o nivel de significación fijado, no se rechaza la hipótesis nula. Por tanto, se puede afirmar que la estructura satisfactoria derivada de la solución del modelo y la estructura real, no difieren, es decir, no existen diferencias estadísticamente significativas entre ambas estructuras.

2.5. Análisis estadístico matemático realizado al parámetro cantidad de cabezas del rebaño planificadas total al final del año (b)

Para realizar la estimativa y la inferencia de la ecuación de regresión, se tienen en cuenta los siguientes aspectos, que se encuentran desarrollados de forma práctica en este epígrafe.

- I. Estimativa de la función de regresión. Estimadores mínimos cuadráticos. Determinación de $\hat{\beta}_1$ y $\hat{\beta}_2$ por la vía matricial.
- II. Validación de la función de regresión estimada. Prueba F de Fisher y Prueba t de Student.
- III. Prueba de hipótesis en el análisis de correlación. Prueba de hipótesis de relación aplicando el SPSS.
- IV. Pruebas de los supuestos de la función de regresión estimada:
 - IV.1- Verificar el supuesto de normalidad mediante la Prueba de Jarque-Bera y de Shapiro-Wilk.
 - IV.2- Verificar el supuesto de no autocorrelación. Prueba de Durbin-Watson.
 - IV.3- Verificar el supuesto de homocedasticidad mediante la Prueba de Goldfeld-Quandt y la Prueba de White
- V. Determinación de los intervalos de confianza para β_1 y β_2 .
- VI. Verificar la validez del función de regresión a través del intervalo de confianza de β_2 .

I- Estimativa de la función de regresión. Estimadores mínimos cuadráticos. Determinación de $\hat{\beta}_1$ y $\hat{\beta}_2$ por la vía matricial

Para mostrar la estimación del parámetro relevante cantidad (número) de cabezas del rebaño planificadas total al final del año, representado por b , se tiene en cuenta la existencia final anual real del rebaño durante 15 años, donde el ganado se ha desarrollado en condiciones similares de manejo,

explotación y alimentación. Los resultados de la muestra seleccionada se presentan en la siguiente tabla.

Tabla 2.3. Existencia real del rebaño al final del año de los últimos quince años.

Rebaño existente al final de año (cabezas)	Tiempo (años)	Rebaño existente al final de año (cabezas)	Tiempo (años)	Rebaño existente al final de año (cabezas)	Tiempo (años)
Y_t	X	Y_t	X	Y_t	X
14309	2001	13595	2006	10814	2011
15449	2002	14419	2007	9748	2012
14631	2003	12668	2008	9234	2013
14522	2004	10837	2009	9887	2014
14757	2005	11741	2010	9698	2015

Fuente: elaborada por los autores.

Se propone investigar si la variable dependiente, cantidad (número) de cabezas del rebaño al final del año, se encuentra relacionada con la variable independiente, tiempo en años, donde se considera una muestra de 15 años.

La información de la muestra expuesta se representa gráficamente por la nube de puntos que se exhibe en el siguiente diagrama de dispersión, formado por 15 pares ordenados (x_i, y_i) , donde $i = 1, 2, \dots, 15$.

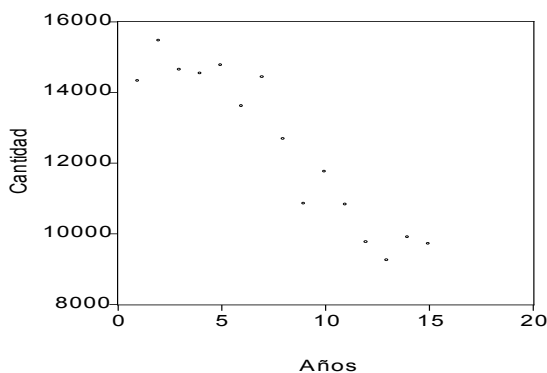


Figura 2.1. Diagrama de dispersión representado por los puntos (x_i, y_i) (EViews).

El diagrama de dispersión da una idea de la forma gráfica de la función, que se obtiene representando en un sistema de coordenadas todos los puntos (x_i, y_i) .

Teniendo en cuenta el comportamiento de la representación de los pares ordenados, se puede ajustar una función racional entera o polinómica de grado uno, o sea, una función lineal, que se expresa en la forma $f(x) = -m x + n$, donde m y n son constantes. Como $m < 0$, la función es decreciente, por cada unidad que aumente x , y disminuye $|m|$ unidades.

Suponiendo que el comportamiento de la cantidad (número) de cabezas del rebaño al final del año (Y) se encuentra en relación lineal con el tiempo (X) en años, entonces, con el propósito de hacer estimaciones, se determina la función $f(x) = -m x + n$ que mejor expresa esta relación.

Conocida la relación funcional entre la cantidad (número) de cabezas del rebaño al final del período (Y) y el tiempo en año (X), se investiga qué modelo matemático, ecuación matemática o función matemática expresa adecuadamente dicha relación. El modelo a estimar será la función de regresión representada por $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 - \hat{\beta}_2 X_t$.

Determinación de los estimadores mínimos cuadráticos de la función de regresión $\hat{\beta}_1$ y $\hat{\beta}_2$ por la expresión matemática resultante del sistema de ecuaciones normales

A continuación, se presenta la información necesaria para calcular los estimadores mínimos cuadráticos.

Tabla 2.4. Información necesaria para calcular $\hat{\beta}_1$ y $\hat{\beta}_2$.

X_i	Y_i	$X_i Y_i$	X_i^2
1	14309	14309	1
2	15449	30898	4
3	14631	43893	9
4	14522	58088	16
5	14757	73785	25
6	13595	81570	36
7	14419	100933	49
8	12668	101344	64
9	10837	97533	81
10	11741	117410	100
11	10814	118954	121
12	9748	116976	144
13	9234	120042	169
14	9887	138418	196
15	9698	145470	225
120	186309	1359623	1240

Fuente: elaborada por los autores.

Cálculo de $\hat{\beta}_2$:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2}$$

$$\hat{\beta}_2 = \frac{1359623 - (15)(8)(12420,6)}{1240 - (15)(64)} = \frac{-130849}{280} \quad \hat{\beta}_2 = -467,3179$$

Cálculo de las medias aritméticas \bar{X} y \bar{Y} :

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad \bar{X} = \frac{120}{15} = 8$$

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} \quad \bar{Y} = \frac{186307}{15} = 12420,6$$

Cálculo de $\hat{\beta}_1$:

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

$$\hat{\beta}_1 = 12420,6 - (-467,3179)8$$

$$\hat{\beta}_1 = 16159,14$$

Por tanto, la función de regresión mínimo cuadrática o ecuación de predicción estimada es $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$.

Determinación de los estimadores mínimos cuadrados de la función de regresión $\hat{\beta}_1$ y $\hat{\beta}_2$ por la vía matricial

Dado el modelo de regresión lineal $Y = \beta_1 + \beta_2 X + U_i$, se tendrá para cada par ordenado (x_i, y_i) , donde $i = 1, 2, 3, \dots, 15$, un conjunto de ecuaciones simultáneas, representado por el siguiente sistema de 15 ecuaciones:

$$\begin{array}{rclcl} Y_1 & = & \beta_1 & - & \beta_2 X_1 & + & U_1 \\ Y_2 & = & \beta_1 & - & \beta_2 X_2 & + & U_2 \\ Y_3 & = & \beta_1 & - & \beta_2 X_3 & + & U_3 \\ Y_4 & = & \beta_1 & - & \beta_2 X_4 & + & U_4 \\ Y_5 & = & \beta_1 & - & \beta_2 X_5 & + & U_5 \\ Y_6 & = & \beta_1 & - & \beta_2 X_6 & + & U_6 \\ Y_7 & = & \beta_1 & - & \beta_2 X_7 & + & U_7 \\ Y_8 & = & \beta_1 & - & \beta_2 X_8 & + & U_8 \\ Y_9 & = & \beta_1 & - & \beta_2 X_9 & + & U_9 \\ Y_{10} & = & \beta_1 & - & \beta_2 X_{10} & + & U_{10} \\ Y_{11} & = & \beta_1 & - & \beta_2 X_{11} & + & U_{11} \\ Y_{12} & = & \beta_1 & - & \beta_2 X_{12} & + & U_{12} \\ Y_{13} & = & \beta_1 & - & \beta_2 X_{13} & + & U_{13} \\ Y_{14} & = & \beta_1 & - & \beta_2 X_{14} & + & U_{14} \\ Y_{15} & = & \beta_1 & - & \beta_2 X_{15} & + & U_{15} \end{array}$$

Representación en notación matricial detallada del sistema de ecuaciones:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \\ Y_8 \\ Y_9 \\ Y_{10} \\ Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{15} \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ 1 & X_4 \\ 1 & X_5 \\ 1 & X_6 \\ 1 & X_7 \\ 1 & X_8 \\ 1 & X_9 \\ 1 & X_{10} \\ 1 & X_{11} \\ 1 & X_{12} \\ 1 & X_{13} \\ 1 & X_{14} \\ 1 & X_{15} \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \\ U_7 \\ U_8 \\ U_9 \\ U_{10} \\ U_{11} \\ U_{12} \\ U_{13} \\ U_{14} \\ U_{15} \end{bmatrix}$$

Representación en notación matricial condensada del sistema de ecuaciones, que representa el modelo de regresión lineal $Y = \beta_1 + \beta_2 X + U_i$, mediante: $Y = X \cdot \beta + U$, donde: Y , X , β y U son matrices.

A partir de la aplicación de $\beta = (X^T X)^{-1} (X^T Y)$ o $\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (X^T X)^{-1} (X^T Y)$ se obtienen los estimadores mínimos cuadráticos $\hat{\beta}_1$ y $\hat{\beta}_2$. Para su determinación, se utiliza el asistente

matemático Derive versión 6.0.

Representación de la matriz columna Y de orden o tamaño de la matriz 15×1 , correspondiente a la variable dependiente cantidad (número) de cabezas del rebaño al final del año, y de la matriz rectangular X de orden o tamaño de la matriz 15×2 , correspondiente a la variable explicativa tiempo.

$$Y = \begin{bmatrix} 14309 \\ 15449 \\ 14631 \\ 14522 \\ 14757 \\ 13595 \\ 14419 \\ 12668 \\ 10837 \\ 11741 \\ 10814 \\ 9748 \\ 9234 \\ 9887 \\ 9698 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \\ 1 & 9 \\ 1 & 10 \\ 1 & 11 \\ 1 & 12 \\ 1 & 13 \\ 1 & 14 \\ 1 & 15 \end{bmatrix}$$

Matriz transpuesta X^T .

La matriz transpuesta de la matriz X se obtiene a partir de X , intercambiando las filas por las columnas.

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \end{bmatrix}$$

Multiplicación de la matriz transpuesta X^T por la matriz X .

Como X^T y X son dos matrices de órdenes 15×2 y 2×15 , respectivamente, el producto de las matrices X^T y X es la matriz $X^T X$ de orden dos o de segundo orden, cuyos elementos para la situación objeto de estudio son números reales positivos, resultantes de la suma de los productos de los elementos de la fila i de X^T por los elementos correspondientes de la columna j de X , resultando:

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ X_1 & X_2 & X_3 & \dots & X_{15} \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ \dots & \dots \\ 1 & X_{15} \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \\ 1 & 9 \\ 1 & 10 \\ 1 & 11 \\ 1 & 12 \\ 1 & 13 \\ 1 & 14 \\ 1 & 15 \end{bmatrix} :$$

$$X^T X = \begin{bmatrix} 15 & 120 \\ 120 & 1240 \end{bmatrix}$$

La matriz cuadrada $(X^T X)$ de orden dos es simétrica, porque se cumple que la matriz $(X^T X)$ es igual a la transpuesta de la matriz $(X^T X)$, es decir, $(X^T X) = (X^T X)^T$.

El determinante de la matriz cuadrada $(X^T X)$ de orden dos o de segundo orden es el número real que se obtiene de la diferencia del producto de los elementos de la diagonal principal menos el producto de los elementos de la diagonal secundaria, o sea,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Resultando el determinante de la matriz $(X^T X)$:

$$\det (X^T X) = \left| X^T X \right| = \begin{vmatrix} 15 & 120 \\ 120 & 1240 \end{vmatrix} = 15 \cdot 1240 - 120 \cdot 120 = 18600 - 14400 = 4200$$

Como el determinante de la matriz cuadrada $(X^T X)$ igual a 4200 es diferente de cero, la matriz cuadrada es no singular o regular, entonces se puede calcular la inversa por el método de la matriz adjunta.

Cálculo de la matriz inversa $(X^T X)^{-1}$ mediante el método de la matriz adjunta.

$$(X^T X)^{-1} = \frac{1}{\det (X^T X)} (X^T X)^+ = \frac{1}{\left| X^T X \right|} (X^T X)^+$$

La matriz adjunta de la matriz cuadrada $(X^T X)$ es:

$$(X^T X)^+ = \begin{bmatrix} 1240 & -120 \\ -120 & 15 \end{bmatrix}$$

Resultando como matriz inversa:

$$(X^T X)^{-1} = \frac{1}{4200} \begin{bmatrix} 1240 & -120 \\ -120 & 15 \end{bmatrix} = \begin{bmatrix} \frac{31}{105} & -\frac{1}{35} \\ -\frac{1}{31} & \frac{1}{280} \end{bmatrix}$$

Multiplicando de la matriz transpuesta de X (X^T) por la matriz columna Y , se obtiene la matriz $(X^T Y)$.

Como X^T y Y son dos matrices de órdenes 2×15 y 15×1 , respectivamente, el producto de las matrices X^T y Y es la matriz $(X^T Y)$ de orden 2×1 , resultando:

$$X^T Y = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ X_1 & X_2 & X_3 & \dots & X_5 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_5 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{2i} Y_i \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \end{bmatrix} \begin{bmatrix} 14309 \\ 15449 \\ 14631 \\ 14522 \\ 14757 \\ 13595 \\ 14419 \\ 12668 \\ 10837 \\ 11741 \\ 10814 \\ 9748 \\ 9234 \\ 9887 \\ 9698 \end{bmatrix} = \begin{bmatrix} 186309 \\ 1359623 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 186309 \\ 1359623 \end{bmatrix}$$

Determinando a $\hat{\beta}_1$ y $\hat{\beta}_2$ como resultado de la multiplicación de las matrices $(X^T X)^{-1}$ y $(X^T Y)$.

$$(X^T X)^{-1} (X^T Y) = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \frac{31}{105} & -\frac{1}{35} \\ -\frac{1}{31} & \frac{1}{280} \end{bmatrix} \begin{bmatrix} 186309 \\ 1359623 \end{bmatrix} = \begin{bmatrix} \frac{113114}{7} \\ -\frac{130849}{280} \end{bmatrix}$$

$$(X^T X)^{-1} (X^T Y) = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \frac{113114}{7} \\ -\frac{130849}{280} \end{bmatrix} = \begin{bmatrix} 16159,1428 \\ -467,3178 \end{bmatrix}$$

Valores de los estimadores mínimos cuadrados obtenidos por la vía matricial.

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 16159,1428 \\ -467,3178 \end{bmatrix}$$

Por tanto, la función de regresión mínimo cuadrática o ecuación de predicción estimada es $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$.

Representación matricial de la ecuación de predicción estimada \hat{Y} .

$$\hat{Y} = [1 \quad X_t] \begin{bmatrix} 16159,1428 \\ -467,3178 \end{bmatrix}$$

Las matrices $(X^T X)$ y $(X^T Y)$ también se pueden obtener calculando sus elementos como se exhibe a continuación.

$$(X^T X) = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \quad (X^T Y) = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

Siendo $(X^T X)$ una matriz cuadrada de segundo orden y simétrica.

Se tabula la información de las variables Y y X para calcular los valores correspondientes a los elementos de las matrices $(X^T X)$ y $(X^T Y)$.

Tabla 2.5. Información para determinar las matrices $(X^T X)$ y $(X^T Y)$.

Y_i	X_i	X_i^2	$X_i Y_i$
14309	1	1	14309
15449	2	4	30898
14631	3	9	43893
14522	4	16	58088
14757	5	25	73785
13595	6	36	81570
14419	7	49	100933
12668	8	64	101344
10837	9	81	97533
11741	10	100	117410
10814	11	121	118954
9748	12	144	116976
9234	13	169	120042
9887	14	196	138418
9698	15	225	145470
186309 $\sum Y_i$	120 $\sum X_i$	1240 $\sum X_i^2$	1359623 $\sum X_i Y_i$

Fuente: elaborada por los autores.

$$\left(X^T X \right) = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} = \begin{bmatrix} 15 & 120 \\ 120 & 1240 \end{bmatrix}$$

$$\left(X^T Y \right) = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \begin{bmatrix} 186309 \\ 1359623 \end{bmatrix}$$

Los resultados de los estimadores mínimos cuadrados calculados de forma manual también se obtienen mediante la aplicación del paquete el estadístico EViews.

Tabla 2.6. Resultado de los estimadores mínimos cuadrados de la función de regresión lineal (EViews).

Dependent Variable: Y Method: Least Squares Sample: 2001 2015 Included observations: 15					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	
C	$\hat{\beta}_1 = 16159.140$	438.4172	36.85791	0.0000	
X	$\hat{\beta}_2 = -467.3179$	48.21947	-9.691477	0.0000	
R-squared	0.878419	Mean dependent var	12420.60		
Adjusted R-squared	0.869067	S.D. dependent var	2229.854		
S.E. of regression	806.8660	Akaike info criterion	16.34776		
Sum squared resid	8463425.	Schwarz criterion	16.44216		
Log likelihood	-120.6082	F-statistic	93.92473		
Durbin-Watson stat	1.508921	Prob(F-statistic)	0.000000		

Por tanto, el comportamiento del rebaño al final del año se ajusta a la función de regresión mínimo cuadrática o ecuación de predicción estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$, con un coeficiente de determinación asociado a un valor relativamente alto de 0,878.

En la tabla 2.7 y la figura 2.2, se representa la estimativa del ritmo de decrecimiento anual de la existencia final del rebaño, obtenidas a partir de la función de regresión mínimo cuadrática $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$.

Tabla 2.7. Valores estimados del rebaño al final del año (EViews).

Tiempo (años)	Rebaño existente al final de año (cabezas)	Rebaño existente al final de año estimado (cabezas)	Tiempo (años)	Rebaño existente al final de año (cabezas)	Rebaño existente al final de año estimado (cabezas)
X	Y	\hat{Y}	X	Y	\hat{Y}
2001	14309	15691,82	2009	10837	11953,28
2002	15449	15224,50	2010	11741	11485,96
2003	14631	14757,18	2011	10814	11018,64
2004	14522	14289,87	2012	9748	10551,32
2005	14757	13822,55	2013	9234	10084,01
2006	13595	13355,23	2014	9887	9616,62
2007	14419	12887,91	2015	9698	9149,37
2008	12668	12420,60	-	-	-

Fuente: elaborada por los autores.

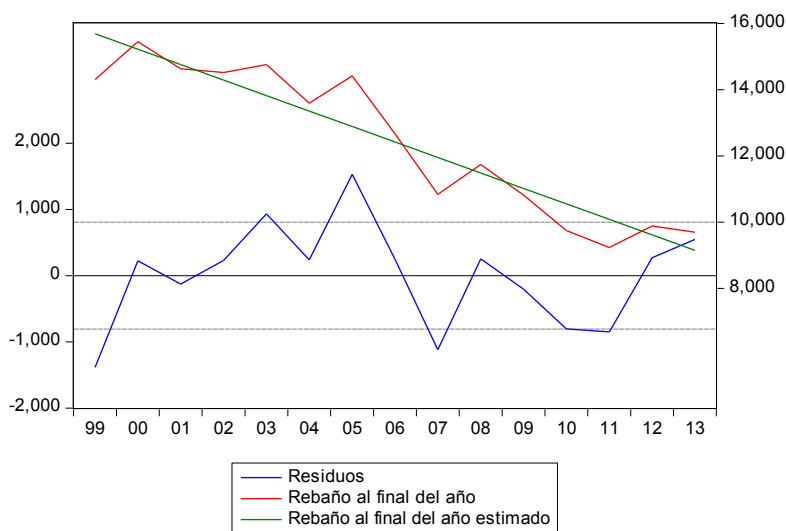


Figura 2.2. Representación de los valores estimados del rebaño al final del año y los residuos (EViews).

II- Validación de la función de regresión estimada. Prueba F de Fisher y prueba t de Student

Prueba F de Fisher para verificar la validez del modelo

La verificación de la confiabilidad de la función de regresión ajustada y la

validez de las proyecciones realizadas se efectúa a partir del análisis de varianza en la regresión, utilizando la prueba de hipótesis de F de Fisher. A continuación, se aplica la prueba de hipótesis.

1- Formulación de las hipótesis nula y alternativa:

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

2- Región crítica para el estadígrafo F , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0 : F > F_{1-\alpha}(p-1, n-p) \right\}$$

$$\text{donde: } F = \frac{S_E^2}{S_R^2} \quad \text{o} \quad F = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R}$$

3- Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha}(p-1, n-p)$.

- No rechazar H_0 , si $F \leq F_{1-\alpha}(p-1, n-p)$.

Alternativamente:

- Rechazar H_0 , si la Prob (F -statistic) o valor $p < \alpha$.

- No rechazar H_0 , si la Prob (F -statistic) o valor $p > \alpha$.

3.1- Cálculo del estadístico de prueba F :

$$F = \frac{S_E^2}{S_R^2} \quad \text{o} \quad F = \frac{SC_E}{p-1} \cdot \frac{n-k}{SC_R}$$

Para determinar al estadístico de prueba o estadígrafo F de Fisher, a partir del análisis de la varianza en la regresión, inicialmente se calcula la: SC_E , SC_R , S_E^2 y S_R^2 .

Tabulando la información correspondiente para la determinación de las sumas de cuadrados.

Tabla 2.8. Cálculo de la información necesaria para determinación de las sumas de cuadrados.

X_i	Y_i	$X_i Y_i$	\hat{Y}_i	$(\hat{Y}_i - \bar{Y})^2$	Y_i^2	$(Y_i - \bar{Y})^2$	$(Y_i - \hat{Y}_i)^2$
1	14309	14309	15691,82	10700880,29	204747481	3566054,56	1912191,150
2	15449	30898	15224,50	7861855,21	238671601	9171206,56	50400,250
3	14631	43893	14757,18	5459606,09	214066161	4885868,16	15921,392
4	14522	58088	14289,87	3494170,33	210888484	4415881,96	53884,336
5	14757	73785	13822,55	1965463,80	217769049	5458764,96	873196,803
6	13595	81570	13355,23	873533,23	184824025	1379215,36	57489,652
7	14419	100933	12887,91	218378,63	207907561	3993602,56	2344236,590
8	12668	101344	12420,60	0	160478224	61206,76	61206,760
9	10837	97533	11953,28	218387,98	117440569	2507788,96	1246081,040
10	11741	117410	11485,96	873551,92	137851081	461856,16	65045,401
11	10814	118954	11018,64	1965491,84	116942596	2581163,56	41877,529
12	9748	116976	10551,32	3494207,71	95023504	7142790,76	645323,022
13	9234	120042	10084,01	5459652,82	85266756	10154419,56	722517,000
14	9887	138418	9616,62	7862303,84	97752769	6419128,96	73105,344
15	9698	145470	9149,37	10700945,71	94051204	7412550,76	300994,877
120	186309	1359623	186308,86	61148429,46	2383681065	69611499,60	8463471,150

Fuente: elaborada por los autores.

Determinación de la suma de cuadrados.

Fórmulas abreviadas:

Fórmulas de la variación:

$S C_E = \hat{\beta}_2 \left(\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y} \right)$ $S C_E = -467,31785 (1359623 - 15 \cdot 8 \cdot 12420,6)$ $S C_E = 61148074,27$	$S C_E = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$ $S C_E \approx 61148429,46$
$S C_T = \sum_{i=1}^n Y_i^2 - n \bar{Y}^2$ $S C_T = 2383681065 - 15 \cdot (12420,6)^2$ $S C_T = 69611499,6$	$S C_T = \sum_{i=1}^n (Y_i - \bar{Y})^2$ $S C_T = 69611499,6$
$S C_R = S C_T - S C_E$ $S C_R = 69611499,6 - 61148074,27$ $S C_R = 8463425,33$	$S C_R = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ $S C_R \approx 8463471,1$

Varianza asociada a la regresión y a los residuos:

$$SCM \text{ de } SC_E = \frac{SC_E}{p-1} = \frac{61148074,289}{2-1} = 61148074,289 = S_E^2$$

$$SCM \text{ de } SC_R = \frac{SC_R}{n-p} = \frac{8463425,311}{15-2} = 651032,716 = S_R^2$$

Cálculo del estadístico de prueba o estadígrafo F :

$$F = \frac{S_E^2}{S_R^2} = \frac{61148074,289}{651032,716} = 93,925$$

$$F = \frac{SC_E}{p-1} \cdot \frac{n-p}{SC_R} = \frac{61148074,289}{2-1} \cdot \frac{15-2}{8463425,311} = 93,925$$

Se compendia la información anterior en tabla de análisis de varianza, para el modelo de regresión de dos variables.

Tabla 2.9. Análisis de Varianza (ANOVA) para el modelo de regresión de dos variables.

Fuente de Variación	Suma de Cuadrado SC	Grados de libertad	Suma de Cuadrados Medios SCM	F
Debido a la regresión (SC_E)	$SC_E = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$ $SC_E = \hat{\beta}_2 \left(\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y} \right)$ 61148074,289	$(p-1)$ 2-1=1	$SCM \text{ de } SC_E = \frac{SC_E}{1} = S_E^2$ 61148074,289	93,925
Debido a los residuos (SC_R)	$SC_R = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ $SC_R = SC_T - SC_E$ 8463425,311	$(n-p)$ 15-2=13	$SCM \text{ de } SC_R = \frac{SC_R}{n-2} = S_R^2$ 651032,716	

Total (SC_T)	$SC_T = \sum_{i=1}^n (Y_i - \bar{Y})^2$	$(n - 1)$		
	$SC_T = \sum_{i=1}^n Y_i^2 - n\bar{Y}^2$	69611499,600	15-1=14	////

Fuente: elaborada por los autores.

Para calcular las sumas de cuadrados: SC_E , SC_T y SC_R , en el modelo de regresión lineal simple, también existe la opción de utilizar la vía matricial utilizando las siguientes expresiones:

$$SC_E = \underline{\beta}^T \underline{X}^T \underline{Y} - n\bar{Y}^2$$

$$SC_R = \underline{Y}^T \underline{Y} - \underline{\beta}^T \underline{X}^T \underline{Y}$$

$$SC_T = \underline{Y}^T \underline{Y} - n\bar{Y}^2$$

El resultado del estadístico de prueba o estadígrafo F calculado de forma manual aplicando el análisis de varianza en la regresión, también se obtiene mediante el paquete estadístico EViews.

Tabla 2.10. Resultado del estadístico de prueba F (EViews).

Dependent Variable: Y				
Method: Least Squares				
Sample: 2001 2015				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16159.14	438.4172	36.85791	0.0000
X	-467.3179	48.21947	-9.691477	0.0000
R-squared	$R^2 = 0.878419$	Mean dependent var		12420.60
Adjusted R-squared	0.869067	S.D. dependent var		2229.854
S.E. of regression	806.8660	Akaike info criterion		16.34776
Sum squared resid	8463425.	Schwarz criterion		16.44216
Log likelihood	-120.6082			
Durbin-Watson stat	1.508921	F-statistic	$F = \frac{S_E^2}{S_R^2} = 93.92473$	
		<i>Prob(F-statistic)</i>		Valor $p = 0.000000$

El valor del estadístico de prueba F se encuentra representado por F - statistic,

siendo el mismo 93,925. El p valor o valor p asociado al estadístico de prueba F se encuentra representado por Prob (F -statistic), siendo el mismo 0,00.

3.2- Determinación del valor crítico o teórico de F :

$$F_{1-\alpha}(p-1, n-p) = F_{1-0,05}(2-1, 15-2) = F_{0,95}(1, 13) = 5,99.$$

Como:

- $F > F_{1-\alpha}(p-1, n-p)$, o sea $93,925 > 5,99$, se rechaza H_0 .
- Prob (F -statistic) o valor $p < \alpha$, o sea $0,00 < 0,05$, se rechaza H_0 .

Basado en los resultados de la muestra y con un nivel de 0,05, a partir del análisis de varianza en la regresión y con la aplicación de la prueba F de Fisher, así como con los resultados del informe del paquete EViews:

- Se puede comprobar que el valor del estadístico de prueba F igual a 93,92 es superior al valor crítico o teórico $F_{0,95}(1;13)$ de 5,99, y se ubica dentro de la región crítica, por lo que se rechaza la hipótesis nula, $\beta_2 = 0$, siendo el valor de F altamente significativo.
- No existen razones para rechazar el supuesto de linealidad entre la variable y la dependiente (Y) explicativa (X), y se acepta la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$ como buena predictora de la función de regresión poblacional $E_{(Y_t/X_t)} = Y_t = \beta_1 \pm \beta_2 X_t$, y que el ajuste hecho es bueno.
- Entonces, la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$ se puede utilizar para predecir la cantidad (número) de cabezas del rebaño al final del período (Y), a partir de valores conocidos del tiempo en año (X).
- *Alternativamente*, se corrobora lo analizado, ya que el valor de la probabilidad de F -statistic o valor p igual a cero (extremadamente pequeño), es menor que el nivel de significación de 0,05. Por tanto, la función de regresión lineal simple $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$ es de buen ajuste

y se considera como buena predictora del valor del ritmo de decrecimiento medio de la masa ganadera, a partir de su comportamiento anual.

Los resultados satisfactorios obtenidos demuestran que la función de regresión lineal simple, $\hat{E}_{(Y/X)} = \hat{Y} = 16159,14 - 467,32X$, representa el modelo adecuado para proyectar la existencia del rebaño al final del año.

Análisis del coeficiente de determinación

En el análisis de regresión, el coeficiente de determinación es una medida con un significado más relevante que el coeficiente de correlación, ya que informa qué proporción de la variación en la variable dependiente es explicada por la variable independiente.

Cálculo del coeficiente de determinación:

$$R^2 = \frac{SC_E}{SC_T} = \frac{\hat{\beta}_2 \left(\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y} \right)}{\sum_{i=1}^n Y_i^2 - n \bar{Y}^2} \quad R^2 = \frac{61148074,2}{69611499,6} \frac{89}{00} = 0,8784$$

A partir del resultado de la relación anterior y del informe del EViews el coeficiente de determinación es de 0,8784, próximo a 1.

El resultado obtenido expresa que:

- La variable predictora, tiempo (en años), explica aproximadamente el 88% de la variación de la existencia final de rebaño durante el período objeto de estudio, es decir que aproximadamente el 88% de la variación de la cantidad de cabezas del rebaño total al final del año es explicado por el tiempo.
- Se puede considerar el tiempo como una variable buena predictora, con el propósito de determinar el ritmo de decrecimiento de la masa ganadera anualmente.
- El porcentaje es considerado de alto poder explicativo. Por lo que la función de regresión estimada de la muestra se ajusta bien a los datos.

Prueba *t* de Student para verificar la validez del modelo

Como complemento a la pruebas de *F* de Fisher, se aplica la prueba *t* de Student para probar la hipótesis nula $\beta_2 = 0$.

1- Formulación de las hipótesis nula y alternativa:

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

2- Región crítica para el estadígrafo t , dado el nivel de significación α , es

$$W = \left\{ t \in \mathfrak{R} : t < -t_{1-\frac{\alpha}{2}}(n-2) \quad \text{o} \quad t > t_{1-\frac{\alpha}{2}}(n-2) \right\}$$

$$t = t_{\hat{\beta}_2} = \frac{\hat{\beta}_2}{S_{\hat{\beta}_2}}, \text{ donde: } S_{\hat{\beta}_2} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n (X_i - \bar{X})^2}} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}}$$

3- Regla de decisión:

- Rechazar H_0 , si $t < -t_{1-\frac{\alpha}{2}}(n-2)$ o $t > t_{1-\frac{\alpha}{2}}(n-2)$.

- No rechazar H_0 , si $t \geq -t_{1-\frac{\alpha}{2}}(n-2)$ o $t \leq t_{1-\frac{\alpha}{2}}(n-2)$,

o sea, $-t_{1-\frac{\alpha}{2}}(n-2) \leq t \leq t_{1-\frac{\alpha}{2}}(n-2)$.

Alternativamente:

- Rechazar H_0 , si la Prob (t -statistic) o valor $p < \alpha$.

- No rechazar H_0 , si la Prob (t -statistic) o valor $p > \alpha$.

3.1- Cálculo del estadístico de prueba t .

$$t = t_{\hat{\beta}_2} = \frac{\hat{\beta}_2}{S_{\hat{\beta}_2}}, \text{ donde}$$

$$S_{\hat{\beta}_2} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n (X_i - \bar{X})^2}} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}}$$

Tabulando la información necesaria para calcular el error estándar $S_{\hat{\beta}_2}$ asociado a $\hat{\beta}_2$.

Tabla 2.11. Cálculo de la información necesaria para determinación el error estándar $S_{\hat{\beta}_2}$.

X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	X_i^2
1	-7	49	1
2	-6	36	4
3	-5	25	9
4	-4	16	16
5	-3	9	25
6	-2	4	36
7	-1	1	49
8	0	0	64
9	1	1	81
10	2	4	100
11	3	9	121
12	4	16	144
13	5	25	169
14	6	36	196
15	7	49	225
120	-	280	1240

Fuente: elaborada por los autores.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{120}{15} = 8$$

$$\sum_{i=1}^n (X_i - \bar{X})^2 = 280 \quad \sum_{i=1}^n X_i^2 = 1240$$

$$S_R^2 = \frac{SC_R}{n-2} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-2} = \frac{8463425,311}{13} = 651032,716, \text{ se obtiene de}$$

la tabla ANOVA.

Cálculo del $S_{\hat{\beta}_2}$:

$$S_{\hat{\beta}_2} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n (X_i - \bar{X})^2}} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}}$$

$$S_{\hat{\beta}_2} = \sqrt{\frac{651032,716}{280}} = \sqrt{\frac{651032,716}{1240 - 15(8)^2}} = 48,2194$$

Cálculo del estadístico de prueba $t = t_{\hat{\beta}_2}$:

$$t = t_{\hat{\beta}_2} = \frac{\hat{\beta}_2}{S_{\hat{\beta}_2}} = \frac{-467,3179}{48,2194} = -9,69$$

Los resultados del $S_{\hat{\beta}_2}$ y t calculados de forma manual también se obtienen aplicando el paquete estadístico EViews.

Tabla 2.12. Resultado del estadígrafo t , error estándar $S_{\hat{\beta}_2}$, y valor p (EViews).

Dependent Variable: Y Method: Least Squares Sample: 2001 2015 Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16159.14	438.4172	36.85791	0.0000
X	-467.3179	$S_{\hat{\beta}_2} = 48.21947$	$t = t_{\hat{\beta}_2} = -9.691477$	0.0000
R-squared	0.878419	Mean dependent var		12420.60
Adjusted R-squared	0.869067	S.D. dependent var		2229.854
S.E. of regression	806.8660	Akaike info criterion		16.34776
Sum squared resid	8463425.	Schwarz criterion		16.44216
Log likelihood	-120.6082	F-statistic		93.92473
Durbin-Watson stat	1.508921	Prob(F-statistic)		0.000000

3.2- Determinación del valor crítico o valor del percentil $t_{1-\frac{\alpha}{2}}(n-2)$ de orden $1-\frac{\alpha}{2}$ de la distribución t de Student con $n-2$ grados de libertad:

$$\text{El valor crítico es } t_{1-\frac{\alpha}{2}}(n-2) = t_{1-\frac{0,05}{2}}(15-2) = t_{0,975}(13) = 2,16$$

Como: $t < -t_{1-\frac{\alpha}{2}}(n-2)$, o sea $-9,69 < -2,16$, se rechaza la hipótesis nula.

A partir del reporte de salida del paquete estadístico EViews, se obtiene que la Prob (t -statistic) o valor p es menor que el nivel de significación, o sea $0,00 < 0,05$, por tanto se rechaza la hipótesis nula.

Los resultados obtenidos con la aplicación de la prueba de t de Student al parámetro β_2 , coinciden con los resultados de la prueba F , por tanto la función de regresión lineal simple $\hat{E}_{(Y_t|X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$ es de buen ajuste y se considera como buena predictora.

III- Prueba de hipótesis en el análisis de correlación. Prueba de hipótesis de relación aplicando el SPSS

Prueba de hipótesis en el análisis de correlación

Para el análisis de la correlación mediante la prueba de hipótesis, se plantea una prueba de hipótesis bilateral o de dos colas.

1- Formulación de la hipótesis nula y alternativa:

$$H_0: \rho_{(x, y)} = 0$$

$$H_1: \rho_{(x, y)} \neq 0$$

Hipótesis nula: no existe relación significativa entre la cantidad (número) de cabezas del rebaño al final del año, variable dependiente (Y), y el tiempo en años, variable explicativa (X).

Hipótesis alternativa: existe relación significativa entre la cantidad (número) de cabezas del rebaño al final del año, variable dependiente (Y), y el tiempo en años, variable explicativa (X).

2- Región crítica para el estadígrafo t , dado el nivel de significación α , es

$$W = \left\{ t \in \mathfrak{R} : t < -t_{1-\frac{\alpha}{2}}(n-2) \quad o \quad t > t_{1-\frac{\alpha}{2}}(n-2) \right\}$$

El estadístico de prueba está dado por el valor de t calculado, obtenido u observado $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$.

3- Regla de decisión:

- Rechazar H_0 , si $t < -t_{1-\frac{\alpha}{2}}(n-2)$ o $t > t_{1-\frac{\alpha}{2}}(n-2)$.

- No rechazar H_0 , si: $t \geq -t_{1-\frac{\alpha}{2}}(n-2)$ o $t \leq t_{1-\frac{\alpha}{2}}(n-2)$,

o sea, $-t_{1-\frac{\alpha}{2}}(n-2) \leq t \leq t_{1-\frac{\alpha}{2}}(n-2)$.

3.1- Cálculo del estadístico de prueba t :

Para calcular el valor del estadístico de prueba t , primero hay que determinar el coeficiente de correlación r_{xy} , que se determina a partir de: su definición, el coeficiente de determinación, y la matriz de correlación.

Determinar el coeficiente de correlación, a partir de su definición:

$$r_{xy} = \frac{\sum_{i=1}^n X_i \cdot Y_i - \frac{\left(\sum_{i=1}^n X_i\right)\left(\sum_{i=1}^n Y_i\right)}{n}}{\sqrt{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}} \cdot \sqrt{\sum_{i=1}^n Y_i^2 - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n}}$$

Tabulación de la información necesaria para determinar el coeficiente de correlación.

Tabla 2.13. Cálculo de los elementos de la fórmula del coeficiente de correlación

r_{xy} :

Observaciones	X	Y	$X.Y$	X^2	Y^2
1	1	14309	14309	1	204747481
2	2	15449	30898	4	238671601
3	3	14631	43893	9	214066161
4	4	14522	58088	16	210888484
5	5	14757	73785	25	217769049
6	6	13595	81570	36	184824025
7	7	14419	100933	49	207907561
8	8	12668	101344	64	160478224
9	9	10837	97533	81	117440569
10	10	11741	117410	100	137851081
11	11	10814	118954	121	116942596
12	12	9748	116976	144	95023504
13	13	9234	120042	169	85266756
14	14	9887	138418	196	97752769
15	15	9698	145470	225	94051204
Total	120	186309	1359623	1240	2383681065

Fuente: elaborada por los autores.

Sustituyendo en r_{xy} :

$$r_{xy} = \frac{1359623 - \frac{120 \cdot 186309}{15}}{\sqrt{1240 - \frac{(120)^2}{15}} \sqrt{2383681065 - \frac{(186309)^2}{15}}} = \frac{1359623 - 1490472}{16,73 \cdot 8343,35} = -0,937$$

Determinación el coeficiente de correlación, a partir del coeficiente de determinación R^2 :

$$r_{xy} = \pm \sqrt{R^2}, \text{ toma el signo de } \hat{\beta}_2.$$

Se conoce el coeficiente de determinación $R^2 = 0,8784$.

$$\text{Por tanto: } r_{xy} = \pm \sqrt{R^2} = \pm \sqrt{0,8784} = -0,937$$

Determinación del coeficiente de correlación, a partir de la matriz de correlación:

Los resultados del coeficiente de correlación, calculado de forma manual a partir de su definición y del coeficiente de determinación, también se obtienen mediante una matriz cuadrada simétrica de orden dos, llamada matriz de

correlación. La matriz cuadrada de correlación de orden dos es simétrica, porque se cumple que la matriz de correlación es igual a la transpuesta de la matriz de correlación.

Matriz de correlación obtenida mediante el paquete estadístico EViews.

Correlation Matrix		
	<i>Y</i>	<i>X</i>
<i>Y</i>	1	-0,937
<i>X</i>	-0,937	1

El coeficiente de correlación igual a -0,937 significa que existe una alta correlación negativa entre la variable predictora, tiempo (en años), y la variable predicha cantidad de cabezas del rebaño total al final del año.

Cálculo del estadístico de prueba t en función de r_{xy} :

$$t = \frac{-0,937 \sqrt{15-2}}{\sqrt{1 - (-0,937)^2}} = -9,67$$

El valor obtenido del estadístico de prueba de forma manual coincide con el valor de t - Statistic del EViews, tabla 2.12.

3.2- Determinación del valor crítico o valor del percentil $t_{1-\frac{\alpha}{2}}(n-2)$ de orden $1-\alpha/2$ de la distribución t de Student con $n-2$ grados de libertad:

El valor crítico es $t_{1-\frac{\alpha}{2}}(n-2) = t_{1-\frac{0,05}{2}}(15-2) = t_{0,975}(13) = 2,16$.

Como:

- $t < -t_{1-\frac{\alpha}{2}}(n-2)$, o sea, $-9,67 < -2,16$ se rechaza H_0 .

- Prob (t -statistic) o valor $p <$ nivel de significación, o sea $0,00 < 0,05$, se rechaza H_0 .

Como el valor obtenido del estadístico de prueba t igual a $-9,67$ es menor que el valor crítico de $-2,16$, el valor t calculado se sitúa en la región crítica. Por tanto, se rechaza la hipótesis nula, por lo que existe una correlación significativa entre la variable predicha (Y), cantidad de cabezas del rebaño

total al final del año y la variable predictora (X), tiempo, para un nivel de significación de 0,05, considerándose que existe una relación inversa entre las variables poblacionales.

Alternativamente, a partir del reporte de salida del paquete estadístico EViews se obtiene que la probabilidad de t -statistic o valor p es menor que el nivel de significación. De esta forma, se rechaza la hipótesis nula y se confirma lo expresado sobre la correlación.

En resumen, como el coeficiente de correlación es igual a -0,937, indica que la variable predicha (Y), cantidad de cabezas del rebaño total al final del año y la variable predictora (X), tiempo, tienen un alto grado de asociación, por existir una alta correlación negativa.

Prueba de hipótesis de relación aplicando el SPSS. Prueba del coeficiente de correlación de Pearson

Para determinar la relación entre la variable cantidad (número) de cabezas del rebaño al final del año (Y) y la variable tiempo (X), también se utiliza la prueba de hipótesis paramétrica del coeficiente de correlación de Pearson, mediante la aplicación del paquete de programa estadístico SPSS.

Para aplicar la prueba del coeficiente de correlación de Pearson, se comprueba que se cumplen los supuestos de que variables se encuentran en forma numérica, y que las observaciones o valores numéricos de la variable siguen una distribución normal. El análisis del supuesto de normalidad se exhibe más adelante, mediante la aplicación de las pruebas de Jarque-Berra y Shapiro-Wilk.

Formulación de la hipótesis nula y alternativa.

Hipótesis nula: no existe relación significativa entre la cantidad (número) de cabezas del rebaño al final del año, variable dependiente (Y), y el tiempo en años, variable explicativa (X).

Hipótesis alternativa: existe relación significativa entre la cantidad (número) de cabezas del rebaño al final del año, variable dependiente (Y), y el tiempo en años, variable explicativa (X).

En la tabla 2.14, se aprecia que el coeficiente de correlación de Pearson entre las variables cantidad (número) de cabezas del rebaño al final del año y la variable tiempo es de -0,937, así como el valor p .

Tabla 2.14. Coeficiente de correlación de Pearson entre la cantidad de cabezas del rebaño al final del año y el tiempo en años (SPSS).

Correlaciones			
		Cantidad de cabezas	Tiempo
Cantidad de cabezas	Correlación de Pearson	1	-,937**
	Sig. (bilateral)		,000
	N	15	15
Tiempo	Correlación de Pearson	-,937**	1
	Sig. (bilateral)	,000	
	N	15	15

** . La correlación es significativa en el nivel 0,01 (bilateral).

Como el valor p correspondiente a la prueba del coeficiente de correlación de Pearson igual a 0,00 es menor que el nivel de significación de 0,05, se rechaza la hipótesis nula. Por tanto, existe una relación significativa y una alta correlación negativa entre la variable predicha (Y), cantidad de cabezas del rebaño total al final del año y la variable predictora (X), tiempo (en años). Por tanto, existe una relación inversa entre las variables poblacionales.

El resultado obtenido con la aplicación prueba de hipótesis del coeficiente de correlación de Pearson coincide con el obtenido de la prueba de hipótesis en el análisis de la correlación utilizando al estadígrafo t de Student.

No se aplica la prueba del coeficiente de correlación de Spearman porque las dos variables tienen distribución normal.

IV- Pruebas de los supuestos de la función de regresión estimada

IV.1- Verificar el supuesto de normalidad mediante la Prueba de Jarque-Bera

Prueba de Jarque-Bera para verificar el supuesto de normalidad de la variable dependiente o regresada en el análisis de regresión.

1- Formulación de las hipótesis nula y alternativa:

$$H_0: Y_i \sim N.$$

$$H_1: Y_i \text{ no sigue una distribución normal.}$$

2- Región crítica para el estadístico de prueba JB , dado el nivel de significación α , es

$$W = \left\{ JB \in \mathfrak{R} \text{ con } JB \geq 0 : JB > \chi_{1-\alpha}^2 (2) \right\}$$

$$\text{donde: } JB = n \left[\frac{A^2}{6} + \frac{(K-3)^2}{24} \right] \text{ o } JB = \frac{n-k}{6} \left[A^2 + \frac{1}{4}(K-3)^2 \right]$$

3- Regla de decisión:

- Rechazar H_0 , si $JB > \chi_{1-\alpha}^2 (2)$.

- No rechazar H_0 , si $JB \leq \chi_{1-\alpha}^2 (2)$.

Alternativamente:

- Rechazar H_0 , si el valor de la probabilidad asociada al estadístico de prueba JB o valor $p < \alpha$.

- No rechazar H_0 , si el valor de la probabilidad asociada al estadístico de prueba o valor $p > \alpha$.

3.1- Cálculo del estadístico de prueba JB .

Como el estadístico de prueba JB depende la asimetría y la curtosis, se determinan primeramente la asimetría y la curtosis.

$$\text{Asimetría: } A = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^3 \quad \text{Curtosis: } K = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^4$$

En las expresiones de la asimetría y la curtosis, $\hat{\sigma} = S$ representa el estimador sesgado de la desviación estándar poblacional, y se calcula por:

$$S = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}; \text{ se conoce de la tabla 2.8 la suma del cuadrado de}$$

las desviaciones de las observaciones con respecto a la media

$$\sum_{i=1}^n (y_i - \bar{y})^2. \text{ Entonces, } S = \sqrt{\frac{69611499,6}{15}} = 2154,24387.$$

Tabulando la información correspondiente para la determinación de la asimetría y la curtosis.

Tabla 2.15. Cálculo de $\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^3$ y $\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^4$ para determinar la asimetría y la curtosis.

y_i	$(y_i - \bar{y})$	$\frac{(y_i - \bar{y})}{\hat{\sigma}}$	$\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^3$	$\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^4$
14309	1888,4	0,876595276	0,673592709	0,590468187
15449	3028,4	1,405783274	2,778146317	3,905471625
14631	2210,4	1,026067675	1,080259311	1,108419160
14522	2101,4	0,975469876	0,928200048	0,905431185
14757	2336,4	1,084556875	1,275724791	1,383596093
13595	1174,4	0,545156478	0,162018098	0,088325216
14419	1998,4	0,927657276	0,798293633	0,740542897
12668	247,4	0,114843079	0,001514658	0,000173948
10837	-1583,6	-0,735107117	-0,397239002	0,292013217
11741	-679,6	-0,315470319	-0,031396086	0,009904533
10814	-1606,6	-0,745783717	-0,414799945	0,309351045
9748	-2672,6	-1,240620914	-1,909489589	2,368952720
9234	-3186,6	-1,479219713	-3,236667284	4,787742052
9887	-2533,6	-1,176097115	-1,626782731	1,913254477
9698	-2722,6	-1,263830914	-2,018677410	2,551266917
186309	////	////	-1,937302480	20,954913300

Fuente: elaborada por los autores.

A continuación, se exhibe el comportamiento de la variable dependiente (Y_t) rebaño al final del año durante 15 años.

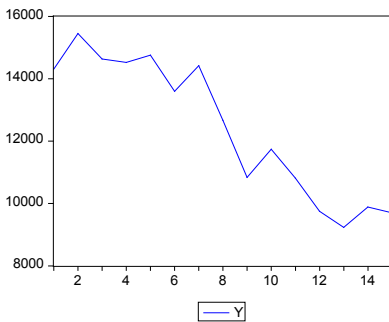


Figura 2.3. Gráfico de línea que representa el comportamiento de la variable dependiente (EViews).

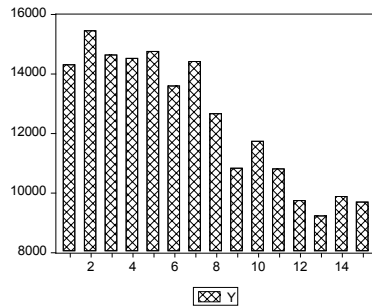


Figura 2.4. Gráfico de barra que representa el comportamiento de la variable dependiente (EViews).

Cálculo de la asimetría y la curtosis:

$$A = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^3 \quad A = \frac{1}{15} (-1,9373025) = -0,129153$$

$$K = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^4 \quad K = \frac{1}{15} (20,954913) = 1,396994$$

Cálculo del estadístico de prueba JB :

$$JB = n \left[\frac{A^2}{6} + \frac{(K - 3)^2}{24} \right]$$

$$JB = 15 \left[\frac{(-0,129153)^2}{6} + \frac{(1,396994 - 3)^2}{24} \right] = 15(0,002780 + 0,107068) = 1,647718$$

Los resultados de los cálculos manuales obtenidos de la asimetría, la curtosis y del estadístico Jarque-Bera se corroboran con los resultados de la prueba Jarque-Bera para la variable regresada, obtenidos a través del programa EViews, como se muestra en la siguiente figura:

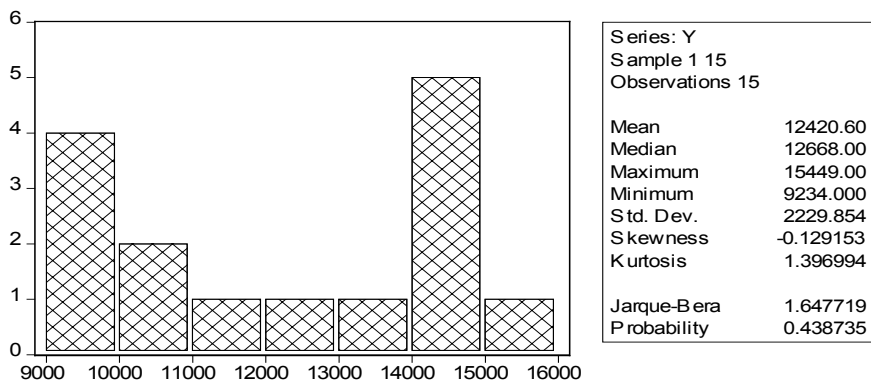


Figura 2.5. Resultado de la prueba Jarque-Bera para la variable regresada. Histograma y estadística de la prueba de normalidad (EViews).

3.2- Determinación del valor crítico o valor del percentil $\chi^2_{1-\alpha} (2)$.

Asintóticamente, el estadístico Jarque-Bera sigue una distribución ji-cuadrado, con dos grados de libertad.

$$\chi_{1-\alpha}^2(2) = \chi_{1-0,05}^2(2) = \chi_{0,95}^2(2) = 5,99$$

Como:

- $JB \leq \chi_{1-\alpha}^2(2)$, o sea, $1,64 < 5,99$, no se rechaza la hipótesis nula.
- El valor de la probabilidad asociada al estadígrafo JB o valor p es mayor que el nivel de significación, o sea, $0,438 > 0,05$, no se rechaza la hipótesis nula.

Toma de decisión:

- Como el estadístico JB igual a 1,64 es menor al valor crítico de $\chi_{1-\alpha}^2(2)$ igual a 5,99, no existen elementos para rechazar la hipótesis nula, lo que implica que la *variable regresada* rebano al final del año sigue una distribución normal para un nivel de significación del 0,05. Alternativamente, se comprueba que la probabilidad asociada al estadístico JB de 0,438 es superior al nivel de significación del 0,05, arribándose a la misma conclusión sobre la existencia de normalidad.
- Como el coeficiente de asimetría igual a -0,129 es menor que cero, la serie de observaciones es asimétrica negativa, presentando la distribución normal deformación a la izquierda.
- La curtosis es igual a 1,396, es menor que tres, por lo que la distribución es menos apuntada que la normal, es más achatada, y se denomina platicúrtica. Existe una baja concentración de los valores en la región central de la distribución.

Prueba de Jarque-Bera para verificar el supuesto de normalidad de los residuos en el análisis de regresión.

1- Formulación de las hipótesis nula y alternativa:

$$H_0: U_i \sim N.$$

$$H_1: U_i \text{ no sigue una distribución normal.}$$

2- Región crítica para el estadístico de prueba JB , dado el nivel de significación α , es

$$W = \left\{ JB \in \mathfrak{R} \text{ con } JB \geq 0 : JB > \chi_{1-\alpha}^2(2) \right\}.$$

$$\text{donde: } JB = n \left[\frac{A^2}{6} + \frac{(K-3)^2}{24} \right] \text{ o } JB = \frac{n-k}{6} \left[A^2 + \frac{1}{4}(K-3)^2 \right]$$

3- Regla de decisión:

- Rechazar H_0 , si $JB > \chi^2_{1-\alpha} (2)$.

- No rechazar H_0 , si $JB \leq \chi^2_{1-\alpha} (2)$.

Alternativamente:

- Rechazar H_0 , si el valor de la probabilidad asociada al estadístico de prueba JB o valor p es menor que el nivel de significación.

- No rechazar H_0 , si el valor de la probabilidad asociada al estadístico de prueba JB o valor p es mayor que el nivel de significación.

3.1 Cálculo del estadístico de prueba JB .

Primero se determina la asimetría y la curtosis, donde: $A = -0,081642$ y $K = 2,621866$.

$$JB = n \left[\frac{A^2}{6} + \frac{(K - 3)^2}{24} \right]$$

$$JB = 15 \left[\frac{(-0,081642)^2}{6} + \frac{(2,621866 - 3)^2}{24} \right]$$

$$JB = 15(0,0011109 + 0,0059577) = 0,106029$$

Los resultados de los cálculos manuales obtenidos del estadístico Jarque-Bera se corroboran con los resultados de la prueba Jarque-Bera para los residuos, obtenidos a través del programa EViews, como se muestra en la siguiente figura:

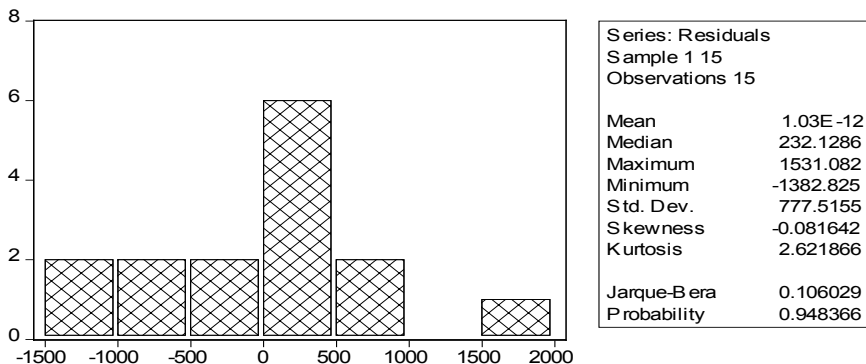


Figura 2.6. Resultado de la prueba Jarque-Bera para los residuos. Histograma y estadística de la prueba de normalidad (EViews).

3.2- Se determina el valor crítico o valor del percentil $\chi^2_{1-\alpha} (2)$:

Asintóticamente el estadístico Jarque-Bera (JB) sigue una distribución ji-cuadrado, con dos grados de libertad.

$$\chi^2_{1-\alpha} (2) = \chi^2_{1-0,05} (2) = \chi^2_{0,95} (2) = 5,99$$

Como:

- $JB \leq \chi^2_{1-\alpha} (2)$, o sea, $0,106 < 5,99$, no se rechaza la hipótesis nula.
- El valor de la probabilidad asociada al estadígrafo JB o valor p es mayor que el nivel de significación, o sea, $0,948 > 0,05$, no se rechaza la hipótesis nula.

Toma de decisión:

- Como el estadístico JB igual a 0,106 es menor al valor crítico de $\chi^2_{1-\alpha} (2)$ igual a 5,99; no existen elementos para rechazar la hipótesis nula, lo que implica que los residuos siguen una distribución normal para un nivel de significación del 0,05. Alternativamente, se comprueba que la probabilidad asociada al estadístico JB de 0,948 es superior al nivel de significación del 0,05, arribándose a la misma conclusión sobre la existencia de normalidad.
- Como el coeficiente de asimetría igual a -0,081 es menor que cero, la serie de observaciones es asimétrica negativa, presentando la distribución normal deformación a la izquierda.
- Al ser la curtosis igual a 2,62 es menor que tres, la distribución es menos apuntada que la normal, es más achatada, y se denomina platicúrtica. Existe una baja concentración de los valores en la región central de la distribución.

Verificar el supuesto de la normalidad mediante la Prueba de Shapiro-Wilk

Para comprobar si las observaciones de la variable dependiente (Y) cantidad (número) de cabezas del rebaño al final del año y de la variable independiente (X) tiempo, siguen una distribución normal, existen dos pruebas que también es posible aplicar: la prueba Kolmogorov-Smirnov, para cuando el tamaño de la muestra es mayor a 50, y Shapiro-Wilk, para cuando el tamaño de la muestra es menor a 50. Ambas se encuentran disponibles en el paquete de programa estadístico SPSS.

Considerando la formulación de las hipótesis nula como la cantidad (número) de cabezas del rebaño al final del año sigue una distribución normal, y como hipótesis alternativa a constatar, que la cantidad (número) de cabezas del rebaño al final del año no sigue una distribución normal, es decir difiere de la distribución normal. Como el número de observaciones es menor a 50, se considera el resultado de la prueba de Shapiro-Wilk, expuesto en la tabla 2.16, donde el valor p del estadístico Shapiro-Wilk es 0,051, superior al nivel de significación de 0,05 o margen de error que se corre al momento de rechazar la hipótesis nula siendo cierta, por tanto no se rechaza la hipótesis de normalidad, y se concluye que la variable cantidad de cabeza del rebaño se ajusta a una distribución normal.

De igual manera, considerando las observaciones de la variable el tiempo, se considera el resultado de la prueba de Shapiro-Wilk, expuesto en la tabla 2.16, donde el valor p del estadístico Shapiro-Wilk es 0,75, superior al nivel de significación de 0,05 o margen de error que se corre al momento de rechazar la hipótesis nula siendo cierta. Por tanto, no se rechaza la hipótesis de normalidad, concluyendo que la variable el tiempo no difiere de la distribución normal.

Tabla 2.16. Resultado de la prueba de normalidad Kolmogorov-Smirnov y Shapiro-Wilk (SPSS).

Pruebas de normalidad						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Cantidad de cabezas	,201	15	,103	,882	15	,051
Tiempo	,082	15	,200*	,964	15	,755

*. Esto es un límite inferior de la significación verdadera.
a. Corrección de significación de Lilliefors

IV.2- Verificar el supuesto de no autocorrelación. Prueba de Durbin-Watson

A partir de la cantidad (número) de cabezas del rebaño al final del año, variable dependiente (Y) y del tiempo en año, variable explicativa (X), se comprueba si existe no autocorrelación en los residuos y se sigue el procedimiento expuesto en el epígrafe 1.6.5:

- Estimar el modelo por el Método de los Mínimos Cuadrados Ordinarios, bajo el supuesto de no autocorrelación.

- Determinar el estadístico de prueba d de Durbin-Watson, a partir de los residuos del modelo e_i .
- Contrastar las hipótesis.

La estimación de la función de regresión lineal se realizó de forma manual y mediante el programa estadístico EViews, bajo el supuesto de no autocorrelación, obteniéndose la función de regresión mínimo cuadrática o función de regresión predictora de mejor ajuste

$$\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32 X_t .$$

Determinación del estadístico de prueba d de Durbin-Watson, a partir de los residuos del modelo, mediante su expresión:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Seguidamente, se muestran los residuos obtenidos mediante el programa estadístico EViews.

Tabla 2.17. Error de estimación o residuos y su ploteo (EViews).

Y_i	\hat{Y}_i	$(Y_i - \hat{Y}_i) = e_i$	Residual Plot
Actual	Fitted	Residual	
14309,0	15691,8	-1382,82	* . .
15449,0	15224,5	224,49	. * .
14631,0	14757,2	-126,18	. * .
14522,0	14289,9	232,12	. * .
14757,0	13822,6	934,44	. . *
13595,0	13355,2	239,76	. * .
14419,0	12887,9	1531,08	. . *
12668,0	12420,6	247,40	. * .
10837,0	11953,3	-1116,28	* . .
11741,0	11486,0	255,03	. * .
10814,0	11018,6	-204,64	. * .
9748,0	10551,3	-803,32	* .
9234,0	10084,0	-850,01	* . .
9887,0	9616,7	270,30	. * .
9698,0	9149,4	548,62	. * .

Tabulando la información necesaria para determinar el estadístico de prueba d de Durbin-Watson.

Tabla 2.18. Información necesaria para calcular el estadístico de Durbin-Watson.

Y_i	\hat{Y}_i	$(Y_i - \hat{Y}_i) = e_i$	$e_i - e_{i-1}$	$(e_i - e_{i-1})^2$	e_i^2
14309,0	15691,8	-1382,82	-	-	1912191,152
15449,0	15224,5	224,49	1607,31	2583445,43	50395,760
14631,0	14757,2	-126,18	-350,67	122969,44	15921,392
14522,0	14289,9	232,12	358,30	128378,89	53879,694
14757,0	13822,6	934,44	702,32	493253,38	873178,113
13595,0	13355,2	239,76	-694,68	482580,30	57484,857
14419,0	12887,9	1531,08	1291,32	1667507,34	2344205,966
12668,0	12420,6	247,40	-1283,68	1647834,34	61206,760
10837,0	11953,3	-1116,28	-1363,68	1859623,14	1246081,038
11741,0	11486,0	255,03	1371,31	1880491,11	65040,300
10814,0	11018,6	-204,64	-459,67	211296,51	41877,529
9748,0	10551,3	-803,32	-598,68	358417,74	645323,022
9234,0	10084,0	-850,01	-46,69	2179,96	722517,000
9887,0	9616,7	270,30	1120,31	1255094,49	73062,090
9698,0	9149,4	548,62	278,32	77462,02	300983,904
Total				12770534,10 $\sum_{i=2}^n (e_i - e_{i-1})^2$	8463348,583 $\sum_{i=1}^n e_i^2$

Fuente: elaborada por los autores.

Donde:

$$\sum_{i=2}^n (e_i - e_{i-1})^2 = 12770534,10$$

$$\sum_{i=1}^n e_i^2 = 8463348 = SC_R$$

Entonces:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} = \frac{12770534}{8463348} = 1,508$$

Determinación del estadístico de prueba d de Durbin-Watson, a partir del reporte de salida del programa estadístico EViews.

Tabla 2.19. Estadístico de prueba Durbin-Watson (EViews).

Dependent Variable: Y				
Method: Least Squares				
Sample: 2001 2015				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16159.14	438.4172	36.85791	0.0000
X	-467.3179	48.21947	-9.691477	0.0000
R-squared	0.878419	Mean dependent var		12420.60
Adjusted R-squared	0.869067	S.D. dependent var		2229.854
S.E. of regression	806.8660	Akaike info criterion		16.34776
Sum squared resid	8463425.	Schwarz criterion		16.44216
Log likelihood	-120.6082	F-statistic		93.92473
Durbin-Watson stat	$d = 1.508921$	Prob(F-statistic)		0.000000

De forma manual y mediante el programa estadístico EViews, se obtiene el valor del estadístico de prueba de Durbin-Watson $d = 1,508$.

Planteamiento de la hipótesis nula y alternativa, consideradas:

$$H_0 : \rho_{U_i U_{i-1}} = 0 \Leftrightarrow Cov(U_i U_{i-1}) = 0 \text{ No hay autocorrelación.}$$

$$H_1 : \rho_{U_i U_{i-1}} \neq 0 \Leftrightarrow Cov(U_i U_{i-1}) \neq 0 \text{ Hay autocorrelación.}$$

Determinando los valores tabulados que se encuentran representados por los valores críticos d_u valor máximo y d_L valor mínimo, tabulados en dependencia del número de observaciones $n = 15$ y del número de variables explicatorias $k' = 1$, y en función del nivel de significación, como la prueba de hipótesis es bilateral o de dos colas, entonces para un $\alpha = 0,05$ se tendría que buscar en $\alpha/2 = 0,025$ o 2,5%, obteniéndose como resultados los valores máximo y mínimo, $d_u = 1,23$ y $d_L = 0,95$, respectivamente.

Regla de decisión

Mediante la Prueba de Durbin-Watson, a partir de la siguiente representación, se determinó que el valor del estadístico de prueba d es 1,508, cumple la condición $d_u < d < 4 - d_u$, o sea, $1,23 < 1,508 < 2,77$, situándose en la región o zona de aceptación, por tanto, no se rechaza la hipótesis nula, $H_0 : \rho_{U_i U_{i-1}} = 0 \Leftrightarrow Cov(U_i U_{i-1}) = 0$, aseverándose que no existe autocorrelación serial de primer orden en los residuos.

↓	Rechazar H_0 hay autocorrelación positiva	↓	Zona de indeterminación o de duda	↓	No rechazar H_0	↓	Zona de indeterminación o de duda	↓	Rechazar H_0 hay autocorrelación negativa	↓
0		d_L		d_u	2	$4 - d_u$		$4 - d_L$		4
		0,95		1,23	Valor de $d = 1,508$	2,77		3,05		

IV.3- Verificar el supuesto de homocedasticidad mediante la Prueba de Goldfeld-Quandt y la Prueba de White

IV.3.1- Prueba de Goldfeld-Quandt

1- Formulación de las hipótesis nula y alternativa:

Se utiliza una prueba para determinar si las varianzas de los dos grupos de muestras son iguales, es decir:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

2- Región crítica para el estadístico de prueba $\lambda = F$, dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0: F > F_{1-\alpha} \left[(n-c-2p)/2; (n-c-2p)/2 \right] \right\}$$

$$W = \left\{ \lambda \in \mathfrak{R} \text{ con } \lambda \geq 0: \lambda > F_{1-\alpha} \left[(n-c-2p)/2; (n-c-2p)/2 \right] \right\}$$

$$\text{donde el estadístico de prueba es } \lambda = F = \frac{\frac{SC_{R2}}{(n-c-2p)/2}}{\frac{SC_{R1}}{(n-c-2p)/2}}$$

3- Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha} [(n - c - 2p) / 2 ; (n - c - 2p) / 2]$.

- No rechazar H_0 , si $F \leq F_{1-\alpha} [(n - c - 2p) / 2 ; (n - c - 2p) / 2]$.

3.1- Cálculo del estadístico de prueba $\lambda = F$:

$$\lambda = F = \frac{\frac{SC_{R2}}{(n - c - 2p) / 2}}{\frac{SC_{R1}}{(n - c - 2p) / 2}}$$

Para calcular el estadístico de prueba $\lambda = F$, se determina primeramente la suma de cuadrados de los residuos: SC_{R1} y SC_{R2} .

En la tabla 2.20, se ordena la información de los datos de mayor a menor, tomando como referencia la variable explicativa X , y se dividen las observaciones en dos muestras: una muestra que recoge los menores valores y otra que recoge los mayores valores de la muestra, sin considerar los valores centrales.

Tabla 2.20. Submuestras de los valores mayores y menores prueba de Goldfeld-Quandt.

Submuestra de las primeras observaciones correspondiente a los valores mayores de X .	9698,0	15	Se le determina la suma de cuadrados de los residuos a las primeras observaciones correspondiente a los valores mayores de X . SC_{R2}
	9887,0	14	
	9234,0	13	
	9748,0	12	
	10814,0	11	
	11741,0	10	
Valores centrales C	10837,0	9	
	12668,0	8	
	14419,0	7	
Submuestra de las segundas observaciones correspondiente a los valores menores de X	13595,0	6	Se le determina la suma de cuadrados de los residuos a las segundas observaciones correspondiente a los valores menores de X . SC_{R1}
	14757,0	5	
	14522,0	4	
	14631,0	3	
	15449,0	2	
	14309,0	1	

Fuente: elaborada por los autores.

Estimación de la submuestra de las primeras observaciones, correspondiente a los valores mayores de X y determinar la suma de cuadrados de los residuos SC_{R2} .

Tabla 2.21. Regresión para las primeras observaciones, correspondiente a los valores mayores de X (EViews).

Dependent Variable: Y1				
Method: Least Squares				
Sample(adjusted): 1 6				
Included observations: 6 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	15012.00	1925.571	7.796129	0.0015
X1	-386.0000	152.6277	-2.529029	0.0647
R-squared	0.615236	Mean dependent var		10187.00
Adjusted R-squared	0.519045	S.D. dependent var		920.6624
S.E. of regression	638.4877	Akaike info criterion		16.01728
Sum squared resid	$SC_{R2} = 1630666.$	Schwarz criterion		15.94787
Log likelihood	-46.05185	F-statistic		6.395988
Durbin-Watson stat	1.158910	Prob(F-statistic)		0.064731

Función de regresión estimada $\hat{Y}_1 = 15012,0 - 386,0 X_1$, donde la suma de cuadrados de los residuos SC_{R2} es 1630666.

Estimación de la submuestra de las segundas observaciones, correspondiente a los valores menores de X y determinar la suma de cuadrados de los residuos SC_{R1} .

Tabla 2.22. Regresión para las segundas observaciones, correspondiente a los valores menores de X (EViews).

Dependent Variable: Y2				
Method: Least Squares				
Sample(adjusted): 1 6				
Included observations: 6 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	15119.33	541.8592	27.90270	0.0000
X2	-164.4286	139.1366	-1.181778	0.3028
R-squared	0.258792	Mean dependent var		14543.83
Adjusted R-squared	0.073491	S.D. dependent var		604.6938
S.E. of regression	582.0502	Akaike info criterion		15.83219
Sum squared resid	$SC_{R1} = 1355130.$	Schwarz criterion		15.76278
Log likelihood	-45.49658	F-statistic		1.396599
Durbin-Watson stat	2.425195	Prob(F-statistic)		0.302751

Función de regresión estimada $\hat{Y}_2 = 15119,33 - 164,43X_2$, donde la suma de cuadrados de los residuos SC_{R1} es 1355130.

Calcular el cociente de la suma de cuadrados de los residuos de ambas regresiones.

Para detectar la heterocedasticidad, a partir de la prueba de Goldfeld-Quandt, se determina el valor estimado $\lambda = F$, obtenido como el cociente de la suma de cuadrados de los residuos de las primeras observaciones, correspondiente a los valores mayores de X , con respecto a la suma de cuadrados de los residuos de las segundas observaciones, correspondiente a los valores menores de X , es decir:

$$\lambda = F = \frac{\frac{SC_{R2}}{(n-c-2p)/2}}{\frac{SC_{R1}}{(n-c-2p)/2}}$$

$$\lambda = F = 1630666/1355130 = 1,20$$

3.2- Determinación del valor crítico o teórico de $\lambda = F$

$$F_{1-\alpha} [(n-c-2p)/2 ; (n-c-2p)/2] = F_{0,95} [(15-3-4)/2 , (15-3-4)/2] =$$

$$F_{0,95}(4;4) = 6,39$$

Como: $F \leq F_{1-\alpha} [(n-c-2p)/2 ; (n-c-2p)/2]$, o sea, $1,20 < 6,39$, no se rechaza H_0 .

Como el valor obtenido $\lambda = F = 1630666/1355130 = 1,20$ no excede el valor crítico o teórico de $F_{0,95}(4;4)$ igual a 6,39, por tanto, no se rechaza la hipótesis nula, infiriéndose que existe homocedasticidad en las varianzas del error, pues las varianzas se mantienen constantes para todas las observaciones y no dependen de la observación correspondiente.

IV.3.2- Prueba general de heterocedasticidad de White

Inicialmente a partir de los datos correspondientes de la cantidad de cabezas del rebaño planificadas total al final del año durante quince años, se estima la función de regresión representada por $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_t$, cuyo resultado exhibido anteriormente es el siguiente:

Tabla 2.23. Resultado estadístico del modelo de regresión lineal estimado (EViews).

Dependent Variable: Y				
Method: Least Squares				
Sample: 2001 2015				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16159.14	438.4172	36.85791	0.0000
X	-467.3179	48.21947	-9.691477	0.0000
R-squared	0.878419	Mean dependent var		12420.60
Adjusted R-squared	0.869067	S.D. dependent var		2229.854
S.E. of regression	806.8660	Akaike info criterion		16.34776
Sum squared resid	8463425.	Schwarz criterion		16.44216
Log likelihood	-120.6082	F-statistic		93.92473
Durbin-Watson stat	1.508921	Prob(F-statistic)		0.000000

Como:

- $t < -t_{1-\frac{\alpha}{2}}(n-2)$ y la Prob (t -statistic) o valor p menor que el nivel de significación, se rechaza la hipótesis nula.

- $F > F_{1-\alpha}(p-1, n-p)$ y la Prob (F -statistic) o valor p menor que el nivel de significación, se rechaza la hipótesis nula.

Como resultado de aplicar las pruebas t de Student y F de Fisher, la función

de regresión lineal simple $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$ es de

buen ajuste y significativa para un nivel de significación de 0,05.

A partir de los resultados anteriores, se realiza la prueba general de heterocedasticidad de White, obteniéndose el siguiente resultado:

Tabla 2.24. Prueba de la heterocedasticidad de White (EViews).

White Heteroskedasticity Test:				
F-statistic	$F = 0.225256$	Probability	0.801612	
Obs*R-squared	$nR^2 = 0.542763$	Probability	Valor $p = 0.762326$	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1 15				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	781872.1	702525.9	1.112944	0.2875
X	-19370.01	202048.6	-0.095868	0.9252
X^2	-758.2707	12279.62	-0.061750	0.9518
R-squared	$R^2 = 0.036184$	Mean dependent var	564228.4	
Adjusted R-squared	-0.124452	S.D. dependent var	743779.2	
S.E. of regression	788704.7	Akaike info criterion	30.17103	
Sum squared resid	7.46E+12	Schwarz criterion	30.31264	
Log likelihood	-223.2827	F-statistic	0.225256	
Durbin-Watson stat	2.522230	Prob(F-statistic)	0.801612	

La ecuación de los errores está dada por la regresión auxiliar asociada a la prueba, siendo: $e_i^2 = \alpha_1 + \alpha_2 X_2 + \alpha_3 X_2^2 + v_i$

1- Formulación de las hipótesis nula y alternativa:

H_0 : $\alpha_2 = \alpha_3 = 0$ Existe homocedasticidad (no hay heterocedasticidad).

H_1 : Alguna $\alpha_i \neq 0$ Existe heterocedasticidad.

2- Región crítica para el estadístico de prueba nR^2 , dado el nivel de significación α , es

$$nR^2 \sim \chi^2_{1-\alpha}(k)$$

$$W = \left\{ n \cdot R^2 \in \mathfrak{R} \text{ con } n \cdot R^2 \geq 0 : n \cdot R^2 > \chi^2_{1-\alpha}(k) \right\}$$

k - número de parámetros de los errores al cuadrado menos una constante.

3- Regla de decisión:

- Rechazar H_0 , si $n \cdot R^2 > \chi^2_{1-\alpha}(k)$.

- No rechazar H_0 , si $n \cdot R^2 \leq \chi^2_{1-\alpha}(k)$.

3.1- Cálculo del estadístico de prueba $n.R^2$

$$n.R^2 = 15 \cdot 0,036184 = 0,54276$$

Información del reporte de salida del EViews: Obs*R-squared = 0,542763.

3.2- Determinando el valor crítico o teórico $\chi^2_{1-\alpha}(gl) = \chi^2_{1-\alpha}(k)$

Existen tres parámetros y una constante: $k = p - 1 = 3 - 1 = 2$.

Grados de libertad igual a 2.

$$\text{Entonces: } \chi^2_{1-\alpha}(gl) = \chi^2_{1-\alpha}(k) = \chi^2_{0,95}(2) = 5,99$$

Como:

- $n.R^2 < \chi^2_{(1-\alpha)gl} = \chi^2_{(1-\alpha)k}$, o sea, $0,54276 < 5,99$, no se rechaza H_0 .

- El valor p de $n.R^2 > \alpha$, o sea, $0,762 > 0,05$, no se rechaza H_0 .

Corroborando el resultado anterior, mediante la aplicación de la prueba F de Fisher.

1- Formulación de las hipótesis nula y alternativa:

$H_0: \alpha_2 = \alpha_3 = 0$ Existe homocedasticidad (no hay heterocedasticidad).

H_1 : Alguna $\alpha_i \neq 0$ Existe heterocedasticidad.

2- Región crítica para el estadígrafo F , dado el nivel de significación α , es

$$W = \left\{ F \in \mathfrak{R} \text{ con } F \geq 0: F > F_{1-\alpha}(p-1, n-p) \right\}.$$

3- Regla de decisión:

- Rechazar H_0 , si $F > F_{1-\alpha}(p-1, n-p)$.

- No rechazar H_0 , si $F \leq F_{1-\alpha}(p-1, n-p)$.

3.1- Tomar como valor del estadístico de prueba el que proporciona la prueba de White, $F = 0,225$.

3.2- Determinar el valor crítico o teórico de F :

$$F_{1-\alpha}(p-1, n-p) = F_{1-0,05}(3-1, 15-3) = F_{0,95}(2, 12) = 3,89$$

Como: $F < F_{1-\alpha}(p-1, n-p)$, o sea, $0,225 < 3,89$, no se rechaza la hipótesis nula.

La aplicación de la prueba F de Fisher complementa los resultados obtenidos con el estadístico de prueba nR^2 ; al no rechazar la hipótesis nula, no existe heterocedasticidad, por tanto, se cumple con el supuesto de homocedasticidad.

Como resultado de aplicar las pruebas: Durbin-Watson, general de heterocedasticidad de White, y Jarque-Bera, se comprueba que no existe autocorrelación; que existe homocedasticidad en la varianza del error, es decir, las perturbaciones son homocedásticas, y que existe normalidad en la serie de valores de la variable dependiente o regresada y en los residuos. Teniendo en cuenta los resultados de la muestra y con un nivel de significación del 5%, se concluye que la función de regresión estimada $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,14 - 467,32X_t$, constituye el modelo matemático adecuado que permite realizar las proyecciones correctas de la cantidad de cabezas del rebaño.

V- Determinación de los intervalos de confianza para β_1 y β_2 .

Determinada la función de regresión lineal simple, $\hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,142 - 467,317X_t$, y considerada como de buen ajuste y buena predictora, esta se utilizará para proyectar el comportamiento de la masa ganadera al final del período de un año, es decir, permitirá obtener el término independiente de la restricción de tipo cinco, rebaño planificado al final del período.

Con el interés de incorporarle mayor flexibilidad al modelo de programación de meta lineal con prioridad y robustecer sus resultados, se considera la necesidad de realizar inferencias a partir de la estimación del intervalo de confianza al coeficiente de regresión, que expresa el ritmo de decrecimiento del rebaño al final del año, $\hat{\beta}_2$.

V.1- Determinación del intervalo de confianza del 95% para β_2

Partiendo de $\hat{\beta}_j \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_j}$

se tiene que $\hat{\beta}_2 \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_2}$

$$\text{donde: } \hat{\sigma}_{\hat{\beta}_2} = S_{\hat{\beta}_2} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n (X_i - \bar{X})^2}} = \sqrt{\frac{S_R^2}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}}$$

$$\text{siendo } S_R^2 = \frac{SC_R}{n-p} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-p}$$

Los valores correspondientes al error estándar asociado al coeficiente de regresión $\hat{\beta}_2$ ($\hat{\sigma}_{\hat{\beta}_2} = S_{\hat{\beta}_2}$) y a la varianza asociada a los residuos (S_R^2) fueron determinados en el desarrollo de la prueba t de Student y en el análisis de varianza, obteniéndose el siguiente resultado:

$$\hat{\sigma}_{\hat{\beta}_2} = S_{\hat{\beta}_2} = 48,2194$$

$$S_R^2 = 651032,716$$

Determinando el intervalo en cuestión, para un nivel de confianza del 95%:

$$\hat{\beta}_2 \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_2} \quad \hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,142 - 467,317 X_t$$

$$- 467,317 \pm t_{1-0,05/2}(15-2) \cdot 48,2194$$

$$- 467,317 \pm t_{0,975}(13) \cdot 48,2194$$

$$- 467,317 \pm (2,16) \cdot (48,2194)$$

$$- 467,317 \pm 104,154$$

Límite de confianza inferior: $-467,317 - 104,154 = -571,471$

Límite de confianza superior: $-467,317 + 104,154 = -363,163$

Intervalo de confianza para β_2 : $\left[-571,471 ; -363,163 \right]$ o

$$-571,471 \leq \beta_2 \leq -363,163.$$

Dado el coeficiente de confianza de 95% ($1-\alpha=0,95$), intervalos como este $\left[-571,471 ; -363,163 \right]$ contendrán el verdadero valor β_2 .

Para el propósito que se persigue en este epígrafe, solo es interés determinar el intervalo de confianza para β_2 , no obstante se determina también el intervalo de confianza asociado a β_1 por la particularidad que presenta el cálculo de $\hat{\sigma}_{\hat{\beta}_1} = S_{\hat{\beta}_1}$. Además, puede ser de utilidad para el lector de pregrado.

V.2- Determinación del intervalo de confianza del 95% para β_1

Partiendo de $\hat{\beta}_j \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_j}$

se tiene que $\hat{\beta}_1 \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_1}$

$$\text{donde } \hat{\sigma}_{\hat{\beta}_1} = S_{\hat{\beta}_1} = \sqrt{\frac{S_R^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2}}$$

$$\text{siendo } S_R^2 = \frac{SC_R}{n-p} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-p}$$

Cálculo de $\hat{\sigma}_{\hat{\beta}_1} = S_{\hat{\beta}_1}$, a partir de la información de la tabla 2.11 empleada en la prueba t de Student.

$$\hat{\sigma}_{\hat{\beta}_1} = \sqrt{\frac{(651032,717) \cdot (1240)}{15 \cdot 280}} = 438,417$$

$S_R^2 = \frac{8463425,33}{15 - 2} = 651032,717$, resultado obtenido del análisis de varianza.

Determinando el intervalo en cuestión para un nivel de confianza del 95%:

$$\hat{\beta}_1 \pm t_{1-\alpha/2}(n-p) \hat{\sigma}_{\hat{\beta}_1} \quad \hat{E}_{(Y_t/X_t)} = \hat{Y}_t = 16159,142 - 467,317X_t$$

$$16159,142 \pm t_{1-0,05/2}(15-2) \cdot 438,417$$

$$16159,142 \pm t_{0,975}(13) \cdot 438,417$$

$$16159,142 \pm (2,16) (438,417)$$

$$16159,142 \pm 946,981$$

Límite inferior: $16159,142 - 946,981 = 15212,161$

Límite superior: $16159,142 + 946,981 = 17106,123$

Intervalo de confianza para β_1 : $\left[15212,161 ; 17106,123 \right]$ o

$$15212,161 \leq \beta_1 \leq 17106,123$$

Dado el coeficiente de confianza de 95% ($1-\alpha=0,95$), intervalos como este

$\left[15212,161 ; 17106,123 \right]$ contendrán el verdadero valor de β_1 .

Los resultados del intervalo de confianza se obtienen, igualmente, mediante la utilización del programa estadístico EViews, como se muestra a continuación:

Tabla 2.25. Intervalos de confianza para β_j (EViews).

Coefficient Confidence Intervals Sample: 2001 2015			
Variable	Coefficient	95% CI	
		Low	High
C	16159.14	15212.000	17106.286
X	-467.31	-571.490	-363.146

En correspondencia con lo expresado en el análisis de la restricción de tipo cinco en el epígrafe 2.3, esta se acota superiormente, con el valor resultante

de adicionar a la existencia real inicial (a) el límite superior del intervalo de confianza, resultando como el término independiente de la restricción la cantidad de 9335 animales, es decir, 9698 menos 363, $[9698 + (-363) = 9335]$. Generándose a partir de la restricción en forma general

$$\sum_{i=1,2,3,4,6,7,8,9,10,11} X_{i+3n} + d_3^- = b, \text{ donde: } b = a + L_{\text{superior}},$$

la restricción particular

$$X_{37} + X_{38} + X_{39} + X_{40} + X_{42} + X_{43} + X_{44} + X_{45} + X_{46} + X_{47} + d_3^- \leq 9335.$$

VI- Verificar la validez de la función de regresión a través del intervalo de confianza para β_2 .

1- Formulación de la hipótesis nula y alternativa:

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

2- Región crítica o de rechazo para el estadístico de prueba o coeficiente de regresión β_2 , dado el nivel de significación α , es

$$W = \left\{ \hat{\beta}_2 \in \mathfrak{R} : \hat{\beta}_2 < -t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2} \text{ o } \hat{\beta}_2 > t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2} \right\}.$$

3- Regla de decisión:

$$\text{- Rechazar } H_0, \text{ si } \hat{\beta}_2 < -t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2} \text{ o } \hat{\beta}_2 > t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2}.$$

$$\text{- Norechazar } H_0, \text{ si } \hat{\beta}_2 \geq -t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2} \text{ o } \hat{\beta}_2 \leq t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2}$$

$$\text{o sea, } -t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2} \leq \hat{\beta}_2 \leq t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2}.$$

3.1- Cálculo del estadístico de prueba o coeficiente de regresión $\hat{\beta}_2$:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n X_i Y_i - n \overline{XY}}{\sum_{i=1}^n X_i^2 - n \overline{X}^2}$$

$$\hat{\beta}_2 = \frac{1359623 - (15)(8)(12420,6)}{1240 - (15)(64)} = \frac{-130849}{280} = -467,3179$$

3.2- Determinar el valor crítico o teórico $t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2}$:

$$t_{1-0,05/2}(15-2) \cdot 48,2194 = t_{0,975}(13) \cdot 48,2194 = (2,16)(48,2194) = 104,154$$

Como: $\hat{\beta}_2 < -t_{1-\frac{\alpha}{2}}(n-p) \hat{\sigma}_{\hat{\beta}_2}$, o sea, $-467,32 < 104,15$, se rechaza H_0 .

Basado en los resultados de la muestra y con un coeficiente de confianza del 95%, se puede afirmar que β_2 es significativamente diferente de cero.

Asimismo, como $\beta_2 = 0$ se encuentra por fuera del intervalo de confianza $-571,471 \leq \beta_2 \leq -363,163$, por tanto se rechaza la hipótesis nula de que el verdadero valor de β_2 sea cero, con un 95% de confianza. Entonces, como β_2 no puede tomar valor cero, se puede afirmar que β_2 es significativamente diferente de cero.

Con la información expuesta se comprueba que el resultado derivado de la prueba de hipótesis, mediante el intervalo de confianza para β_2 , es coincidente con el resultado de la prueba de hipótesis mediante la aplicación del análisis de varianza para la regresión y con la prueba t de Student.

2.6. Planteamiento del modelo económico-matemático de trabajo

Continuando con la aplicación de la metodología de la modelación económico-matemática que sustenta el desarrollo del contenido del libro, una vez realizado el planteamiento *general* del modelo económico-matemático teórico, procesada la información de partida, y determinados y validados los parámetros del modelo, se plantea el modelo económico-matemático de *trabajo* que posibilita consolidar e integrar el modelo teórico con los resultados numéricos obtenidos de procesar la información de partida y de la determinación de los parámetros.

Seguidamente, se especifican las restricciones y la función objetivo del planteamiento *general* del modelo económico-matemático de meta a la situación práctica objeto de estudio, lo que constituye el *modelo económico-matemático de trabajo*.

Construcción del sistema de restricciones

Restricción de tipo I, para el movimiento de las categorías por edad y sexo del rebaño.

$$\text{Vacas: } \frac{3463 X_1}{100} - X_{13} - X_{49} + X_{26} + X_{37} + d_{11}^- - d_{11}^+ = 3463 - \frac{3463 \cdot 2,6}{100}$$

$$\text{Novillas: } \frac{2280 X_2}{100} - X_{14} - X_{50} + X_{27} + X_{38} + d_{12}^- - d_{12}^+ = 2280 - \frac{2280 \cdot 2,6}{100}$$

$$\text{Añojas: } \frac{557 X_3}{100} - X_{15} - X_{51} + X_{28} + X_{39} + d_{13}^- - d_{13}^+ = 557 - \frac{557 \cdot 9,4}{100}$$

$$\text{Terneras: } \frac{1140 X_4}{100} - X_{16} - X_{52} + X_{29} + X_{40} + d_{14}^- - d_{14}^+ = 1140 - \frac{1140 \cdot 8,8}{100}$$

$$\text{Toros: } \frac{308 X_6}{100} - X_{18} - X_{54} + X_{30} + X_{42} + d_{16}^- - d_{16}^+ = 308 - \frac{308 \cdot 1,5}{100}$$

$$\text{Bueyes: } \frac{195 X_7}{100} - X_{19} - X_{55} + X_{31} + X_{43} + d_{17}^- - d_{17}^+ = 195 - \frac{195 \cdot 2,6}{100}$$

$$\text{Sementales: } \frac{2X_8}{100} - X_{20} - X_{56} + X_{32} + X_{44} + d_{18}^- - d_{18}^+ = 2 - \frac{2 \cdot 2,1}{100}$$

$$\text{Toretas: } \frac{259 X_8}{100} - X_{21} - X_{57} + X_{33} + X_{45} + d_{19}^- - d_{19}^+ = 259 - \frac{259 \cdot 4,6}{100}$$

$$\text{Añojos: } \frac{356 X_{10}}{100} - X_{22} - X_{58} + X_{34} + X_{46} + d_{110}^- - d_{110}^+ = 356 - \frac{356 \cdot 16,4}{100}$$

$$\text{Terneros: } \frac{1138 X_{11}}{100} - X_{23} - X_{59} + X_{35} + X_{47} + d_{111}^- - d_{111}^+ = 1138 - \frac{1138 \cdot 11}{100}$$

Restricción de tipo II, para el movimiento de las categorías terneras crías y terneros crías del rebaño.

$$\text{Terneras crías: } \frac{943 X_5}{100} - X_{29} = 934 - \frac{943 \cdot 11}{100}$$

$$\text{Terneros crías: } \frac{1030 X_{12}}{100} - X_{36} = 1030 - \frac{1030 \cdot 5}{100}$$

Restricción de tipo III, para garantizar la transferencia de animales hacia las categorías inmediata superior.

$$\text{Para las novillas: } \frac{2280 X_2}{100} + X_{26} + d_{22}^- - d_{22}^+ = 2280 - \frac{2280 \cdot 2}{100}$$

$$\text{Para las añojas: } \frac{557 X_3}{100} + X_{27} + d_{23}^- - d_{23}^+ = 557 - \frac{557 \cdot 9,4}{100}$$

$$\text{Para las terneras: } \frac{1140 X_4}{100} + X_{28} + d_{24}^- - d_{24}^+ = 1140 - \frac{1140 \cdot 8,8}{100}$$

$$\text{Para los toretes: } \frac{559 X_9}{100} + X_{33} + d_{29}^- - d_{29}^+ = 559 - \frac{559 \cdot 4,6}{100}$$

$$\text{Para los añojos: } \frac{356 X_{10}}{100} + X_{34} + d_{210}^- - d_{210}^+ = 356 - \frac{356 \cdot 16,4}{100}$$

$$\text{Para los terneros: } \frac{1138 X_{11}}{100} + X_{35} + d_{211}^- - d_{211}^+ = 1138 - \frac{1138 \cdot 11}{100}$$

Restricción de tipo IV, para garantizar un determinado volumen o rango de animales en algunas de las categorías.

Para la categoría vacas, una aplicación importante de la restricción es garantizar el rebaño final de vacas con el que se podrá cumplir el plan de producción de leche, expresado por:

$$X_{i+3n} \geq V; \quad i = 1$$

Cálculo de V :

$$V = \frac{2P}{R} - a_1$$

$$V = \frac{2 \cdot 750000}{247} - 3463 = 2610 \text{ vacas}$$

El valor de V es igual a 2610 y representa el número de cabezas mínimas de vacas necesarias para garantizar la producción de leche planificada de las vacas; por tanto, la restricción para las vacas se plantea como se muestra a

continuación:

$$\text{Para las vacas: } X_{37} \geq 2610, \quad X_{37} \leq 3850$$

Las restricciones para el resto de las categorías se expresan como:

$$\text{Para las novillas: } X_{38} \leq 1800$$

$$\text{Para las añojas: } X_{39} \leq 650$$

$$\text{Para las terneras: } X_{40} \leq 950$$

$$\text{Para los toros: } X_{42} \leq 300$$

$$\text{Para los bueyes: } X_{43} \leq 250$$

$$\text{Para sementales: } X_{44} = 2$$

$$\text{Para los toretes: } X_{45} \leq 250$$

$$\text{Para los añojos: } X_{46} \leq 400$$

$$\text{Para los terneros: } X_{47} \leq 900$$

Restricción de tipo V, para garantizar el rebaño planificado al final del período.

$$X_{37} + X_{38} + X_{39} + X_{40} + X_{42} + X_{43} + X_{44} + X_{45} + X_{46} + X_{47} + d_3 \leq 9335$$

El término independiente 9335, cota superior de la restricción de tipo V, se determina a partir del intervalo de confianza, donde se adiciona a la existencia inicial del rebaño el límite superior del intervalo de confianza, como se analizó en el epígrafe 2.5.

Restricción de tipo VI, para garantizar la correspondencia entre la transferencia de animales de una categoría a otra.

$$\text{Correspondencia entre vacas y novillas: } X_{13} - X_{26} = 0$$

$$\text{Correspondencia entre novillas y añojas: } X_{14} - X_{27} = 0$$

$$\text{Correspondencia entre añojas y terneras: } X_{15} - X_{28} = 0$$

$$\text{Correspondencia entre terneras y terneras crías: } X_{16} - X_{29} = 0$$

$$\text{Correspondencia entre toretes y añojos: } X_{21} - X_{34} = 0$$

Correspondencia entre añojos y terneros: $X_{22} - X_{35} = 0$

Correspondencia entre terneros y terneros crías: $X_{23} - X_{36} = 0$

Correspondencia entre toros, bueyes y sementales con respecto a los toretes: $X_{18} + X_{19} + X_{20} - X_{33} = 0$

Restricción de tipo VII, para regular la extracción de animales para carne o venta dentro de cada una de las categorías.

Para las vacas: $X_1 \geq 8$, $X_1 \leq 12$

Para las novillas: $X_2 \geq 3$, $X_2 \leq 11$

Para las añojas: $X_3 \geq 25$, $X_3 \leq 30$

Para las terneras: $X_4 \geq 10$, $X_4 \leq 28$

Para las terneras, crías: $X_5 = 0$

Para los toros: $X_6 \geq 15$, $X_6 \leq 20$

Para los bueyes: $X_7 \geq 5$, $X_7 \leq 8$

Para sementales: $X_8 = 0$

Para los toretes: $X_9 \geq 45$, $X_9 \leq 85$

Para los añojos: $X_{10} \geq 25$, $X_{10} \leq 35$

Para los terneros: $X_{11} \geq 25$, $X_{11} \leq 35$

Para los terneros, crías: $X_{12} = 0$

Restricción de tipo VIII, para garantizar la correlación estructural necesaria entre distintas categorías.

Correlación entre vacas y novillas: $X_{38} - 0,12 X_{37} + d_{41}^- - d_{41}^+ = 0$

Correlación entre novillas y terneras: $X_{40} - 0,2 X_{38} + d_{42}^- - d_{42}^+ = 0$

Correlación entre toros ceba y toretes: $X_{45} - 0,2 X_{32} + d_{43}^- - d_{43}^+ = 0$

Correlación entre añojos y terneros: $X_{47} - 0,1 X_{46} + d_{44}^- - d_{44}^+ = 0$

Restricción de tipo IX, para garantizar las compras a terceros.

$$\text{Para las vacas: } X_{49} = 0$$

$$\text{Para las novillas: } X_{50} = 0$$

$$\text{Para las añojas: } X_{51} = 0$$

$$\text{Para las terneras: } X_{52} = 0$$

$$\text{Para los toros: } X_{54} = 0$$

$$\text{Para los bueyes: } X_{55} = 0$$

$$\text{Para los sementales: } X_{56} = 0$$

$$\text{Para los toretes: } X_{57} = 0$$

$$\text{Para los añojos: } X_{58} = 0$$

$$\text{Para los terneros: } X_{59} = 0$$

Restricción de tipo X, para garantizar la producción mercantil de carne en pie del rebaño.

$$12466,8X_1 + 5472X_2 + 1225,4X_3 + 1368X_4 + 1293,6X_6 + 780X_7 + 9X_8 + 828,8X_9 + 783,2X_{10} + 1365,6X_{11} + d_5^- - d_5^+ = 530000$$

Restricción de tipo XI, condición de no negatividad de las variables.

$$X_i \geq 0; \quad i = 1, \dots, 12 \quad X_{i+n} \geq 0; \quad i = 13, \dots, 24 \quad X_{i+2n} \geq 0; \quad i = 12, \dots, 36$$

$$X_{i+3n} \geq 0; \quad i = 37, \dots, 48 \quad X_{i+4n} \geq 0; \quad i = 49, \dots, 60 \quad \text{Todo } d_i^- \geq 0; \quad d_i^+ \geq 0$$

Función Objetivo

La función objetivo tiene como propósito la minimización de las desviaciones, dentro de la estructura prioritaria asignadas a estas desviaciones.

$$\begin{aligned} \text{Min } Z = & P_1(d_5^- + d_5^+) + P_2d_3^- + P_3(d_{41}^- + d_{41}^+ + d_{42}^- + d_{42}^+ + d_{43}^- + d_{43}^+) + P_4(d_{11}^- + d_{11}^+ + d_{12}^- + d_{12}^+ \\ & + d_{13}^- + d_{13}^+ + d_{14}^- + d_{14}^+ + d_{16}^- + d_{16}^+ + d_{17}^- + d_{17}^+ + d_{18}^- + d_{18}^+ + d_{19}^- + d_{19}^+ + d_{110}^- + d_{110}^+ + d_{111}^- + d_{111}^+) \\ & + P_5(d_{22}^- + d_{22}^+ + d_{23}^- + d_{23}^+ + d_{24}^- + d_{24}^+ + d_{29}^- + d_{29}^+ + d_{210}^- + d_{210}^+ + d_{211}^- + d_{211}^+) \end{aligned}$$

En el planteamiento matemático del modelo de trabajo, fueron definidas 60 variables de decisión o esenciales de tipo continuas, 53 variables de desviación, y se representaron, en total, 74 restricciones entre los once tipos de restricciones.

Concluida la construcción del modelo económico-matemático de meta lineal con prioridad, antes de llevarlo al proceso de cómputo electrónico, se procedió a efectuar la *prueba del modelo*, que consiste en la revisión y verificación de los datos iniciales que permitieron el planteamiento del modelo. Estas comprobaciones se efectuaron no solo sobre los cálculos numéricos, sino también de las fuentes originales donde fueron tomados. Finalizada la revisión de los datos, se procedió a determinar la solución del modelo mediante el empleo del paquete de programa WinQSB, utilizando el módulo Linear goal programming and Integer linear goal programming, siendo necesario un total de 77 iteraciones.

2.7. Análisis económico de la solución del modelo matemático de meta con prioridad

Obtenida la solución del modelo matemático de meta con prioridad, se procede a la presentación de la solución satisfactoria que representa el movimiento del rebaño, organizadas en rebaño a inicio del año, entradas, salidas y rebaño al final de año; por último, se expone de forma resumida el análisis de la solución.

Presentación de la solución satisfactoria obtenida

Para efectuar el análisis económico de la solución satisfactoria obtenida, se interpretan las variables de decisión o esenciales en correspondencia con su definición, desde el punto de vista cualitativo y dimensional en el contexto del problema. Las variables de holgura se interpretan teniendo en cuenta su significado, en conformidad con la restricción donde se encuentren; además, se analizan las variables de desviación teniendo en cuenta su definición, evaluándose el efecto que producen sobre la restricción de meta.

El análisis de la solución obtenida del modelo económico-matemático para la programación del movimiento de rebaño, como técnica cuantitativa de gestión, es un momento importante, pues mediante este estudio se realiza la interpretación económica de la solución satisfactoria obtenida.

Para el análisis de la solución del modelo numérico, se toma como referencia el reporte de salida correspondiente al módulo Linear goal programming and Integer linear goal programming del paquete de programa WinQSB. La solución obtenida del modelo económico-matemático es alterna, su existencia amplía las posibilidades en el análisis de la solución.

Los resultados de la solución satisfactoria que se pretende analizar desde el punto de vista técnico-económico, han sido tabulados de manera resumida

en la tabla 2.26, que constituye una adecuada presentación del balance del movimiento de rebaño, por categorías de edad y sexo, para el período planificado.

Tabla 2.26. Movimiento del rebaño propuesto para el período planificado.

Categorías de edad y sexo	Rebaño al inicio del año		Entradas		
	Animales a_i	Estructura	Nacimientos C_i	Llegadas de las categorías inferiores X_{i+n}	Compras a terceros X_{i+4n}
	(cabezas)	(%)	(cabezas)	$(X_{12} - X_{24})$ (cabezas)	$(X_{48} - X_{60})$ (cabezas)
	2	3	4	5	6
Vacas	3463	35,71	-	1983	0
Novillas	2280	23,51	-	366	0
Añojas	557	5,74	-	721	0
Terneras	1140	11,76	-	896	0
Terneras, crías	0	0,00	943	0	0
Toros Ceba	308	3,18	-	260	0
Bueyes	195	2,01	-	76	0
Sementales	2	0,02	-	0	0
Toretos	259	2,67	-	559	0
Añojos	356	3,67	-	785	0
Terneros	1138	11,73	-	979	0
Terneros, crías	0	0,00	1030	0	0
Total	9698	100	1973	6625	0

(Continuación tabla 2.26)

Categorías de edad y sexo	Salidas					Pérdidas $a_i m_i$ 100 o $c_i m_i$ 100 (cabezas)
	Transferencias a categorías superiores X_{i+2n} $(X_{25} - X_{36})$ (cabezas)	Ventas a otras empresas o para el sacrificio			Total extraído para carne	
		Porcentaje de ganado extraído X_i $(X_1 - X_{12})$ (%)	Animales $\frac{a_i X_i}{100}$ o $\frac{c_i X_i}{100}$ (cabezas)	Peso vivo promedio		
	7	8	9	10	11	

Vacas	0	8	277	360	99734,4	90
Novillas	1983	11	251	240	60192	46
Añojas	366	25	139	220	30635	52
Terneras	721	28	319	120	38304	100
Terneras, crías	896	0	0	0	0	47
Toros Ceba	0	20	62	420	25872	5
Bueyes	0	8	16	400	6240	5
Sementales	0	0	0	450	0	0
Toretas	336	85	220	320	70448	12
Añojos	559	35	125	220	27412	58
Terneros	785	20	228	120	27312	125
Terneros, crías	979	0	0	0	0	51
Total	6625	////	1636	////	386149,4	591

(Continuación tabla 2.26)

Categorías de edad y sexo	Variable de desviación restricción movimiento de rebaño		Término Independiente	Rebaño al final de año	
	d_i^-	d_i^+		Animales	Estructura
				X_{i+3n} ($X_{37} - X_{48}$) (cabezas)	(%)
	13	14	15	16	17
Vacas	1299	0	3373	3850	41,24
Novillas	0	1434	2234	1800	19,28
Añojas	71	0	505	650	6,96
Terneras	0	37	1040	933	9,99
Terneras, crías	0	0	896	0	0,00
Toros Ceba	201	0	303	300	3,21
Bueyes	0	0	190	250	2,67
Sementales	0	0	2	2	0,02
Toretas	0	0	247	250	2,67
Añojos	0	0	298	400	4,28
Terneros	79	0	1013	900	9,64
Terneros, crías	0	0	979	0	0,00
Total	////	////	////	9335	100

Fuente: elaborada por los autores.

Análisis de la solución satisfactoria

El análisis de la solución satisfactoria, es decir, el análisis de la mejor solución obtenida con la aplicación del modelo económico-matemático de meta con prioridad, se realiza teniendo en cuenta las fases que constituyen el balance del movimiento de rebaño.

Análisis de las entradas del balance del movimiento de rebaño

Un momento trascendental en el movimiento de rebaño son los cambios de categorías que se originan en correspondencia con la edad y sexo del rebaño, es decir, las transferencias que se producen de la categoría inferior a la categoría inmediata superior, dentro del mismo sexo. En total, se produce el traslado natural entre las categorías de 6625 animales, teniendo un peso importante las transferencias de las categorías de edades inferiores a las categorías vacas, añojas, terneras, añojos y terneros, representando el 80,96% de las transferencias totales.

Análisis de las salidas del balance del movimiento de rebaño

El porcentaje satisfactorio del rebaño extraído para la venta o el sacrificio, con respecto a la existencia inicial, se encuentra formado por nueve categorías por grupos de edad y sexo, como se muestra a continuación: 8% de las vacas, 11% de las novillas, 25% de las añojas, 28% de las terneras, 20% de los toros, 8% de los bueyes, 85% de los toretes, 35% de los añojos, y 20% de los terneros. En total, se extraen para la venta o el sacrificio 1636 cabezas.

La cantidad satisfactoria de cabezas del rebaño destinados a la venta o sacrificio, desglosado por categoría de edad y sexo, permitirá alcanzar una producción mercantil de carne en pie del rebaño total anual de 386520 kg, superior en 37985 kg, es decir, en un 10,90%, a la producción mercantil planificada de carne en pie del rebaño total.

Los terneros, las terneras hasta un año, las vacas y los añojos son las categorías que mayor incidencia tienen en las pérdidas por muerte, constituyendo el 63,11% del total de las muertes.

Análisis de la estructura satisfactoria del rebaño completo al final de año que se planifica

Para el análisis del balance de movimiento de rebaño desde la perspectiva productiva y económica, es indispensable la evaluación de la estructura más adecuada del rebaño al final de año. Siendo la solución satisfactoria, la que constituye la mejor variante en el contexto del modelo matemático planteado, es decir, es la estructura que mejor representa las existencias finales por cada una de las categorías. La misma queda representada por la cantidad de animales y el porcentaje exhibido en las dos últimas columnas de la tabla

2.26, alcanzando un rebaño total de 9335 cabezas.

Se destaca que las vacas y novillas, como categorías estratégicas para el desarrollo de la masa ganadera, representan el 60,52% del rebaño al final de año, lo que contribuirá a garantizar para el año que se planifica, la reproducción y futuro crecimiento del rebaño, así como lograr el nivel satisfactorio de producción de leche planificada anual de 750000 litros.

Comparación entre la estructura propuesta y la real alcanzada por la empresa

La estructura propuesta corresponde a la solución satisfactoria obtenida con la aplicación del modelo económico-matemático de meta planteado.

Al comparar la estructura satisfactoria propuesta del rebaño para final de año, con respecto a la estructura planificada por la empresa para el mismo período, se plantea un crecimiento del 1,46% de cabezas de la propuesta con respecto a lo planificado, equivalente a 134 animales, como se muestra en la siguiente tabla.

Tabla 2.27. Estructura propuesta, planificada y real del rebaño para final de año.

Categorías de edad y sexo	Estructura propuesta del rebaño para final de año		Estructura planificada del rebaño por la empresa		Diferencia de la propuesta respecto al plan	Estructura real del rebaño alcanzada por la empresa		Diferencia del real respecto a la propuesta
	Animales (cabezas)	Estructura %	Animales (cabezas)	Estructura %		Cabezas	Animales (cabezas)	
Vacas	3850	41,24	3728	40,52	122	3902	41,06	52
Novillas	1800	19,28	1840	20,00	-40	1760	18,52	-40
Añojas	650	6,96	571	6,20	79	701	7,38	51
Terneras	933	9,99	856	9,30	77	968	10,19	35
Toros Ceba	300	3,21	198	2,15	102	337	3,55	37
Bueyes	250	2,67	182	2,00	68	291	3,06	41
Sementales	2	0,02	2	0,02	0	2	0,02	0
Toretas	250	2,67	148	1,61	102	295	3,10	45
Añojos	400	4,28	782	8,50	-382	341	3,59	-59
Terneros	900	9,64	894	9,70	6	906	9,53	6
Total	9335	100,00	9201	100,00	134	9503	100,00	168
Porcentaje de crecimiento de la estructura propuesta respecto a la estructura planificada: 1,46 %						////		

Fuente: elaborada por los autores.

Al confrontar la estructura real alcanzada por la empresa con respecto a la estructura satisfactoria propuesta del rebaño para final de año para el mismo período, existe un crecimiento del 1,79% de cabezas del real con respecto a la estructura propuesta, correspondiente a 168 animales, como se exhibe en la tabla.

Para verificar la validez de la *estructura satisfactoria propuesta del rebaño para final del año*, se realiza la prueba estadística no paramétrica de Rango de Wilcoxon. La prueba permite comprobar si existe correspondencia entre la estructura propuesta del rebaño para final del año, con la estructura real del rebaño alcanzada por la empresa al final del año.

En la tabla 2.28 se muestra el estadístico de prueba Wilcoxon (Z) y su nivel crítico bilateral (Asymp. Sig.).

Tabla 2.28. Prueba para dos muestras relacionadas de Wilcoxon (SPSS).

Estadísticos de prueba ^a	
	Estructura Propuesta – Estructura Real
Z	-1,125 ^b
Sig. asintótica (bilateral)	,260
a. Prueba de rangos con signo de Wilcoxon b. Se basa en rangos negativos.	

Como el valor del nivel crítico obtenido de 0,260 es mayor que el nivel de significación de 0,05, no se rechaza la hipótesis nula de la igualdad de promedios, aceptándose la hipótesis de homogeneidad. Por tanto, se puede afirmar que la estructura satisfactoria propuesta del rebaño para final del año y la estructura real del rebaño alcanzada al final de año no difieren significativamente, no existiendo diferencias estadísticamente significativas entre ambas estructuras.

Referencias bibliográficas

- Anderson, D., Sweeney, D., & Williams, T. (2008). *Estadística para administración y economía*. 10^m ed. México: Cengage Learning Editores.
- Arya, J., Lardner, R., & Ibarra. (2009). *Matemáticas aplicadas a la administración y a la economía*. 4^{ta} ed. México: Pearson Educación.
- Ahumada, O., & Villalobos, J. (2009). Application of planning models in the agri-food supply chain. *European Journal Operation Research*, 196(1), 1-20. Recuperado de <https://asu.pure.elsevier.com/en/publications/application-of-planning-models-in-the-agri-food-supply-chain-a-re>
- Bazaraa, M., Jarvis J. & Sherali, H. (2009). *Linear programming and network Flows*. 4^{ta} ed. Nueva York: Wiley.
- Carrascal, A. U., González, G. Y., & Rodríguez, P. B. (2001). *Análisis econométrico con EViews*. México: Alfaomega Grupo Editor.
- Dave, A. (2015). Goal programming applications in agricultural management. *International Research Journal of Engineering and Technology*, 2(6), 883-887. Recuperado de <https://www.irjet.net/archives/V2/i6/IRJET-V2I6137.pdf>
- Díaz, A. (2012). *Estadística aplicada a la administración y a la economía*. México: McGraw-Hill.
- Eppen, G. D., Gould, F. J., Schmidt, C. P., Moore, J. H. & Watherford, L. R. (2000). *Investigación de operaciones en la ciencia administrativa*. 5^{ta} ed. México: Prentice-Hall Hispanoamericana.
- Espallarga, S., & Solís, M. (2012). *Econometría y series temporales. Aplicaciones*. La Habana: Félix Varela.
- Felipe, P., Ortiz, M., Martínez, E., & Arias, E. (2013). *Investigación de operaciones*. La Habana: Editorial Universitaria Félix Varela.
- García, J., Florez, J., Torralba, A., Borrajo, D., Linares, C., García-Olava, A. & Sáenz, J. (2013). Combining linear programming and automated planning to solve intermodal transportation problems. *European Journal Operation Research*: 227 (1), 216-222. Recuperado de <https://www.infona.pl/resource/bwmeta1.element.elsevier-57ad1641-f0ca-3123-899c-8876827fb52f>
- Gil, P. J. (2017). *Estadística e informática SPSS en la investigación descriptiva e inferencial*. (Versión actualizada SPSS 23). Madrid: UNED.
- González, F. C., & Brezo, B. J. (1987). *Modelos matemáticos de los procesos económicos en la agricultura*. La Habana: Pueblo y Educación.

- Gujarati, D. N. (2003). *Econometría*. 4^{ta} ed. México: McGraw-Hill.
- Gujarati, D., & Porter, D. (2010). *Econometría*. 5^{ta} ed. México: Mc Graw-Hill.
- Hillier, F., & Lieberman, G. (2010). *Introducción a la investigación de operaciones*. 9^{ta} ed. México: Mc Graw-Hill.
- Haeussler, E., Paul, R., & Wood, R. (2014). *Matemática para administración y economía*. 13^{ra} ed. México: Pearson Educación.
- International Business Machines Corp. (2017). Guía del usuario de IBM SPSS. Statistics 23 Core System. Recuperado de <http://www.uca.es/serv/ai/formacion/spss/Inicio.pdf>
- Mezghani, M., Rebai, A., Dammak, A., & Loukil, T. (2009). A goal Programming model for aggregate production planning problem. *International Journal of Operational Research*. 4(1), 23 – 33. Recuperado de <https://www.inderscienceonline.com/doi/abs/10.1504/IJOR.2009.021616>
- Moskowitz, H., & Wright, G. (1992). *Investigación de operaciones*. México: Prentice-Hall Hispanoamericana.
- Mussone, L., & Wolfler, R. (2013). An analytical approach to calculate the capacity of a railway system. *European Journal Operation Research*, 228(1), 11-23. Recuperado de <https://www.semanticscholar.org/paper/An-analytical-approach-to-calculate-the-capacity-of-Mussone-Calvo/6d-5612d4a353e039fa3013c09eb198c7d03292e1>
- Ortuño, M. T., & Vitoriano, B. (2011). A goal programming approach for farm planning with resources dimensionality. *Annals of Operations Research*, 190(1), 181–199. Recuperado de <https://link.springer.com/article/10.1007/s10479-009-0524-5>
- Quesada, V., & Vergara, J. (2006). *Análisis cuantitativo con WinQSB*. Recuperado de <http://www.eumed.net/libros/2006c/216>.
- Ribeiro, R., & Morabito, R. (2017). Optimization approaches for sugarcane harvest front programming and scheduling. *Gestão & Produção*. 24(2), 407-422. Recuperado de http://www.scielo.br/scielo.php?pid=S0104-530X2017000200407&script=sci_arttext&tling=en
- Steadie, S. M., Dellaert, N., Nuijten, W., Van Woensel, T., & Raoufi, R. (2014). Multimodal freight transportation planning. *European Journal Operation Research*, 233(1), 1-15. Recuperado de <https://econpapers.repec.org/RePEc:eee:ejores:v:233:y:2014:i:1:p:1-15>

- Stygar, A., & Makulska, J. (2010). Application of mathematical modeling in beef herd management. *Annals of Animal Science*, 10(4), 333–348. Recuperado de <http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.agro-849e4801-31fe-45d0-8878-d8795496bdae>
- Taha, H. (2012). *Investigación de operaciones*. 9^a ed. México: Pearson Educación.
- Wackerly, D., Mendenhall, W., & Scheaffer, R. (2010). *Estadística matemática con aplicaciones*. 7^{ma} ed. México: Cengage Learning Editores.
- Wooldridge, J. M. (2008). *Introducción a la econometría*. 2^{da} Barcelona: Paraninfo.
- Zgajnar, J & Kavcic, S. (2016). Optimal allocation of production resources under uncertainty: Application of the multicriteria approach. *Agri. Econ-Czech*. 62 (12). 556-565. Recuperado de https://www.agriculturejournals.cz/web/agricecon.htm?type=article&id=238_2015-AGRICECON
- Zill, D. G., & Wright, S. W. (2011). *Cálculo. Trascendentes tempranas*. 4^a ed. México: McGraw-Hill/Interamericana Editores.

Índice

Prólogo	8
Capítulo I. Modelación económico-matemática para la planificación óptima del transporte en procesos agroindustriales	10
1.1. Introducción al capítulo. Aspectos esenciales de la metodología de la modelación económico-matemática.....	10
1.2. Características tecnológicas y económicas del proceso de transportación	15
1.3. Formulación matemática, supuestos y elementos básicos para la construcción de un modelo de programación lineal	18
1.4. Generalidades sobre la modelación y formulación de un problema de transporte	25
1.5. Formulación general del problema y modelación económico-matemática para la optimización del transporte en la agroindustria azucarera	28
1.6. Modelización de la capacidad industrial de los centros de recepción. Determinación y validación	40
1.6.1. Generalidades, estimativa e inferencia de la función de regresión.....	40
1.6.2. Pruebas de los supuestos del modelo de regresión.....	55
1.6.3. Esbozo sobre la aplicación de las derivadas para el análisis de las propiedades de funciones	68
1.6.4. Determinación del aprovechamiento máximo de la capacidad industrial	72
1.6.5. Estudio de un caso real sobre la utilización de la capacidad industrial y validación de resultados	75
1.6.6. Análisis de las propiedades de la función de regresión estimada	114
1.7. Modelización económico-matemática de la función de costo. Aplicaciones	117
1.7.1. Determinación y validación de la función de costo.....	117
1.7.2. Determinación de la validez del costo total de transportación calculado	129

1.7.3. Determinación de la distancia económica de los medios de transportes	139
1.7.4. Utilidad práctica de las funciones matemáticas de los costos en la modelización económico-matemática.....	142
1.8. Presentación del modelo numérico	143
1.9. Generalidades sobre la obtención y análisis de la solución del modelo de programación lineal	144

Capítulo II. Modelación económico-matemática para la planificación de procesos económicos en la ganadería 151

2.1. Introducción al capítulo. Generalidades sobre la modelación matemática y aplicaciones a los procesos económicos en la ganadería	151
2.2. Formulación general del problema y modelación económico-matemática de la programación del movimiento de rebaño	156
2.3. Análisis de la construcción del modelo económico-matemático de meta con prioridad para la programación del movimiento de rebaño	172
2.4. Determinación y validación matemática de parámetros del modelo ...	180
2.4.1. Estimativa e inferencia de la función de regresión	181
2.4.2. Análisis estadístico para comprobar la correspondencia entre la estructura satisfactoria y la real alcanzada	199
2.5 Análisis estadístico matemático realizado al parámetro cantidad de cabezas del rebaño planificadas total al final del año (b).....	200
2.6. Planteamiento del modelo económico-matemático de trabajo	250
2.7. Análisis económico de la solución del modelo matemático de meta con prioridad	256

Referencias bibliográficas	262
----------------------------------	-----

En el presente libro, se articula de manera coherente el tratamiento metodológico dado a los contenidos teóricos y prácticos relacionados con las áreas del conocimiento de Programación Lineal, Pruebas de Hipótesis, Análisis de Regresión y Aplicaciones de las Derivadas, además se logra la aplicación de los métodos y modelos económico-matemáticos a procesos económicos del sector agroindustrial y agropecuario. El desarrollo metodológico de los capítulos se sustenta en la aplicación de la metodología de la modelación económico-matemática. En cada capítulo, se hace referencia a las características productivas, económicas y tecnológicas del objeto de estudio práctico, lo que permite lograr una mejor comprensión del proceso de modelación matemática mediante la aplicación de la programación lineal y de meta lineal con prioridades. Un aspecto relevante del libro, que contribuye a robustecer la solución de los modelos económico-matemáticos anteriores, es la determinación y validación de sus parámetros, a través del empleo de los métodos y técnicas de la Estadística, la Econometría y la Matemática. Estos métodos y técnicas, así como los modelos económico-matemáticos, se abordan teóricamente y se realizan sus aplicaciones prácticas a situaciones económicas reales. Se utilizan los paquetes de programas informáticos WinQSB, EViews y SPSS, para la obtención de la solución de los modelos económico-matemáticos y para la estimación e inferencia de los modelos estadísticos. El libro proporciona un adecuado material de consulta para los estudiantes de pregrado de las Ciencias Económicas y Empresariales, en temas de las asignaturas de Matemática Aplicada, Estadística Inferencial, Investigación Operativa y Econometría.

ISBN: 978-959-257-544-8



FUNDACIÓN
METROPOLITANA
Fomentando la Educación Superior

