

This file has been cleaned of potential threats.

If you confirm that the file is coming from a trusted source, you can send the following SHA-256 hash value to your admin for the original file.

dd0749ee4d95cc21b73c40da9f3c71e9702a7262702a142bd3e7562804da34a7

To view the reconstructed contents, please SCROLL DOWN to next page.



UNIVERSIDAD ESTATAL DE MILAGRO

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO

PROYECTO DE INVESTIGACIÓN PREVIO A LA OBTENCIÓN DEL

TÍTULO DE:

MAGÍSTER EN MATEMÁTICAS MENCIÓN MODELACIÓN MATEMÁTICA

**CASO: PROPUESTA DE UN MODELO MATEMÁTICO APLICADO AL
PRONÓSTICO DE PRODUCCIÓN UTILIZANDO REGRESIÓN
MULTIVARIADA APLICADA A UNA FÁBRICA DE BARRAS DE
CHOCOLATE DE 20 g.**

DIRECTOR TFM

Ing. Edwin Evaristo León Plúas. PhD.

AUTOR

Ing. Kleber Andrés Mora Guevara

MILAGRO, junio del 2022

ECUADOR

ACEPTACIÓN DEL TUTOR

Por la presente hago constar que he analizado el proyecto de grado presentado por **el Sr. Kleber Andrés Mora Guevara**, para optar al título de **Magister en Matemáticas Mención Modelación Matemática** y que acepto tutorizar al estudiante, durante la etapa del desarrollo del trabajo hasta su presentación, evaluación y sustentación.

Milagro, a los 24 días del mes de junio del 2022



Firmado electrónicamente por:

**EDWIN
EVARISTO**

Ing. Edwin Evaristo León Plúas

Ph.D.

C.I. 0918303165

DECLARACIÓN DE AUTORÍA DE LA INVESTIGACIÓN

El autor de esta investigación declara ante el Comité Académico del Programa de **Maestría en Matemáticas Mención Modelación Matemática** de la Universidad Estatal de Milagro, que el trabajo presentado es de mi propia autoría, no contiene material escrito por otra persona, salvo el que está referenciado debidamente en el texto; parte del presente documento o en su totalidad no ha sido aceptado para el otorgamiento de cualquier otro Título o Grado de una institución nacional o extranjera.

Milagro, a los 24 días del mes de junio del 2022



Firmado electrónicamente por:
**KLEBER ANDRES
MORA GUEVARA**

Ing. Kleber Andrés Mora Guevara

C.I. 0926301326

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO
CERTIFICACIÓN DE LA DEFENSA EI TRIBUNAL CALIFICADOR

El TRIBUNAL CALIFICADOR previo a la obtención del título de **MAGÍSTER EN MATEMÁTICA MENCIÓN MODELACIÓN MATEMÁTICA**, otorga al presente proyecto de investigación en las siguientes calificaciones:

TRABAJO DE TITULACION	58.67
DEFENSA ORAL	38.67
PROMEDIO	97.33
EQUIVALENTE	Excelente



Firmado electrónicamente por:
**JUAN DIEGO
VALENZUELA
COBOS**

Phd. VALENZUELA COBOS JUAN DIEGO
PRESIDENTE/A DEL TRIBUNAL



Firmado electrónicamente por:
**EDWIN
EVARISTO**

Phd. LEON PLUAS EDWIN EVARISTO
DIRECTOR/A DE TFM



Firmado electrónicamente por:
**JAIR MANUEL
VISTIN**

Mgs VISTIN VISTIN JAIR MANUEL
SECRETARIO/A DEL TRIBUNAL

DEDICATORIA

A Dios, que cada enseñanza de su divina misericordia y cada circunstancia que ha permitido atravesar en mi vida, me han traído a este momento.

A Elba Cecilia, mi madre, que pese a las adversidades nos crió, a mis hermanos y a mí, con ética y valores, y que después de su muerte, mis acciones siguen siendo pensando en lo que la hará sentir orgullosa. A mi padre, que nos enseñó con el ejemplo, y que, con el pasar de los años, sigue velando por mí.

Finalmente, a mi familia por estar siempre presente y alentando a una constante superación.

Ing. Kleber Andrés Mora Guevara

AGRADECIMIENTO

A mi madre, quien siempre fue mi todo, y que dejó los cimientos para poder continuar creciendo en mi vida personal y profesional.

Agradezco de manera muy especial a mi tutor el PhD. Edwin León, su guía y el tiempo que me ha dedicado; gracias a estos aspectos se convirtió en un aporte invaluable a la finalización de mi proyecto previa a la obtención del título.

Adicionalmente agradezco a mis amigos y personas conocidas, que de una u otra manera colaboraron para materializar esta idea, contribuyendo así en mi formación profesional.

Ing. Kleber Andrés Mora Guevara

CESIÓN DE DERECHOS DE AUTOR

Señor Doctor

Fabricio Guevara Viejó

Rector de la Universidad Estatal de Milagro

Presente.

Mediante el presente documento, libre y voluntariamente procedo a hacer entrega de la Cesión de Derecho del Autor del Trabajo realizado como requisito previo para la obtención de mi Título de Cuarto Nivel, cuyo tema fue **“PROPUESTA DE UN MODELO MATEMÁTICO APLICADO AL PRONÓSTICO DE PRODUCCIÓN UTILIZANDO REGRESIÓN MULTIVARIADA APLICADA A UNA FÁBRICA DE BARRAS DE CHOCOLATE DE 20 g”** y que corresponde al Vicerrectorado de Investigación y Posgrado.

Milagro, a los 24 días del mes de junio del 2022



Firmado electrónicamente por:
**KLEBER ANDRES
MORA GUEVARA**

Ing. Kleber Andrés Mora Guevara

C.I. 0926301326

ÍNDICE GENERAL

INTRODUCCIÓN	1
CAPÍTULO I	3
1.1 Planteamiento Del Problema	3
1.2 Delimitación del problema	4
1.3 Formulación del problema	5
1.4 Sistematización del problema	5
1.5 Objetivos	5
1.5.1 Objetivo General	5
1.5.2 Objetivos Específicos	6
1.6 Justificación e importancia	6
1.6.1 Justificación Teórica	6
1.6.2 Justificación metodológica	7
1.6.3 Justificación practica	7
1.7 Hipótesis	8
1.7.1 Hipótesis General	8
1.7.2 Hipótesis Particular	8
1.7.3 Conceptualización y operacionalización de variables	8
CAPÍTULO II	10
2.1 Marco Teórico	10
2.1.1 Modelo matemático	10
2.1.2 Pronóstico de producción	11
2.1.3 Regresión Multivariada	13
2.2 Marco Conceptual	19
2.3 Marco Referencial	21
CAPÍTULO III	23

MARCO METODOLÓGICO	23
3.1 Tipo de Investigación.....	23
3.2 Diseño De La Investigación.....	24
3.3 Población y Muestra.....	24
3.4 Variables de investigación	25
3.5 Método, Técnica y Uso de Software de Tratamiento y Análisis de Datos	26
CAPITULO IV.....	28
PROPUESTA.....	28
4.1 Formulación del modelo de regresión lineal múltiple.....	28
4.2 Diagnósis y validación del modelo de regresión lineal múltiple.....	28
4.3 Supuesto de Normalidad de los Residuos.....	29
4.4 Supuesto de Independencia de las observaciones.....	32
4.5 Supuesto de Homocedasticidad	33
4.6.- Supuesto de linealidad	34
4.7 Supuesto de multicolinealidad.....	37
4.8 El Modelo Matemático de Regresión Multivariada	38
4.9 Comparación de los datos reales con los datos pronosticados	40
CONCLUSIONES	43
RECOMENDACIONES	44
BIBLIOGRAFÍA	45

ÍNDICE DE TABLAS

Tabla 1	8
Tabla 2	25
Tabla 3	25
Tabla 4	26
Tabla 5	26
Tabla 6	29
Tabla 7	33
Tabla 8	34
Tabla 9	37
Tabla 10	38
Tabla 11	39
Tabla 12	40
Tabla 13	40
Tabla 14	41

ÍNDICE DE GRÁFICOS

Figura 1	29
Figura 2	30
Figura 3	31
Figura 4	33
Figura 5	35
Figura 6	41

RESUMEN

El estudio actual presenta el desarrollo de un modelo matemático, utilizando una regresión multivariada, aplicado a una microempresa dedicada a la fabricación de barras de chocolate 20 g, lo que permite la predicción del pronóstico de producción en esta; para hacer esto se caracterizó las principales variables relacionadas o que se considera que tienen incidencia con el proceso de producción, teniendo en cuenta la información proporcionada por la compañía. El tipo de método utilizado es cuantitativo y técnico, implementado para la recopilación de información principal. Después de eso, la construcción de modelos matemáticos se realizó mediante regresión multivariable, en el programa Statistical Package for the Social Sciences (SPSS), luego se procedió a realizar la validación del modelo empleando el criterio de los supuestos de regresión multivariada los cuales son: supuesto de normalidad, supuesto de correlación de variables, supuesto de multicolinealidad, supuesto de independencia de errores, supuesto de homocedasticidad, supuesto de normalidad de residuos. Como resultado y análisis de este trabajo, podemos concluir que el modelo matemático propuesto es aceptable porque es adecuado para las hipótesis sobre regresión, el valor de la R es admisible.

Palabras Claves:

Modelo matemático, Pronósticos, Regresión multivariada, Proceso, Productividad.

ABSTRAC

The current study presents the development of a mathematical model using a multivariate regression, applied to a micro-company dedicated to the manufacture of 20 g chocolate bars, which allows the prediction of the production forecast in this one; to do this, the main variables related or considered to have incidence with the production process were characterized taking into account the information provided by the company. The type of method used is quantitative and technical implemented for the collection of main information. After that, the construction of mathematical models was performed by multivariate regression in the Statistical Package for the Social Sciences (SPSS) program, then proceeded to perform the validation of the model using the criteria of multivariate regression assumptions which are: assumption of normality, assumption of correlation of variables, assumption of multicollinearity, assumption of independence of errors, assumption of homoscedasticity, assumption of normality of residues. As a result and analysis of this work, we can conclude that the proposed mathematical model is acceptable because it is adequate for the regression hypothesis, the R value is admissible.

Keywords:

Mathematical model, Forecasting, Multivariate regression, Process, Productivity

INTRODUCCIÓN

Las microempresas, durante muchos años, han usado algunos métodos poco efectivos para predecir cuanto es lo que van a producir al final de un período de tiempo determinado; estos métodos no han sido del todo satisfactorios o han sido menos precisos, causando a las microempresas que incurran en pérdidas económicas.

Por lo que es necesario lograr un modelo matemático que esté más relacionado con la realidad de cada microempresa, este modelo debe poder tener en cuenta la mayor cantidad de variables independientes que aporten al modelo matemático y que representen lo mejor posible el proceso de producción.

El objetivo de estudio actual es proponer un modelo matemático utilizando una regresión multivariada que permite comprender cómo afectan las variables independientes o proporcionan una descripción de un proceso de producción de la microempresa, comprendiendo la relación causa - efecto entre las variables independientes y la variable subordinada.

Capítulo 1. En el proceso de desarrollo se aclara el método, definición, formulación y sistematización del problema, objetivos generales y específicos, así como la justificación de la teoría, método y práctica.

Capítulo 2. Consta de tres partes, un marco teórico, un marco conceptual y un marco de referencia; en este apartado, la revisión de modelos matemáticos, regresión multivariable, apoyo a la previsión de producción tal como se estudia en libros, artículos, tesis, etc.; los documentos son la fuente de nuestra investigación.

Capítulo 3. Se expone los tipos y los diseños de investigación que se utilizan en este desarrollo y análisis de este trabajo, se presentan muestras y herramientas de recolección de datos, así como el análisis de los mismo.

Capítulo 4. Este capítulo propone la propuesta de modelos matemáticos, apoyados en ciertas hipótesis sobre regresión multivariada, de la siguiente manera:

- Relación lineal
- Independencia de las observaciones
- Homocedasticidad
- Distribución normal de residuos

Además de demostrar el beneficio de los modelos matemáticos, comparando los cálculos de valor de producción real con los valores obtenidos con el volumen de producción real provisto con el modelo matemático desarrollado.

CAPÍTULO I

1.1 Planteamiento Del Problema

La industria alimentaria desarrolla un papel importante en la economía del mundo, las distintas compañías que se derivan de este sector abarcan un enorme área de actividades que van desde las que están dirigidas al tratamiento, la transformación, la preparación, la conservación y el envasado de productos alimenticios (FAO y CEPAL, 2020).

En el mundo las MYPYMES representan el 90% de las micro y pequeñas empresas que existen, estas emplean el 50% de la mano de obra y contribuyen en el desarrollo del 50% de PIB mundial (Valdés Díaz de Villegas & Sánchez Soto, 2012). El crecimiento de las MYPYMES está sujeto a las fuerzas productivas, predicción de los requerimientos del mercado.

Las MYPYMES en América Latina resultan muy importante por su contribución en la generación de empleo (Tello Cabello, 2014), el principal reto que tienen estas MYPYMES es poder predecir la producción (López Rodríguez & Zapata Zuluaga, 2018). Estas empresas se enfrentan a algunos errores al momento de intentar predecir la producción: si se usa solo un método de pronóstico de la demanda, no se asocia la demanda histórica con la información del mercado, no se manejan errores, se puede elegir unas bases de datos incompletas, no se considera la demanda elástica, se ignora el largo del ciclo de vida del producto, etc (Méndez Giraldu & Lopez Santana, 2014).

El mayor problema que enfrenta una MYPYMES al intentar pronosticar la demanda, es la baja eficacia que obtiene el valor pronosticado (Render & Heizer, 2007). La eficacia se la puede medir debido al cálculo del error que tiene el valor pronosticado, en otras palabras que tan cerca está el valor pronosticado con respecto de la demanda real (Zafra Mejía & Gutiérrez Gil, 2015). Algunas de las variables que se presentan en el área de producción, y que pueden afectar el volumen de producción pueden ser: las demandas del mercado, que número de

trabajadores deben tener y cuantas horas deben trabajar para poder producir lo planificado (Damián Llatas & Sandoval Santamaría, 2018).

Uno de los aportes de los modelos de pronóstico es que el error sea lo menor o lo más mínimo posible (Cabrera González & De León Arias, 2019), para ello se puede trabajar con (Morales Tabares y otros, 2016), error estándar de la estimación (SEE), desviación media absoluta (MAD), error cuadrático medio (RMSE), porcentaje del error medio absoluto (MAPE), error medio absoluto (MAE), Desviación Standard (SD), entre otros. (Sánchez Sánchez, 2018), (Damián Llatas & Sandoval Santamaría, 2018).

Una alternativa es el uso de modelos multivariados para realizar pronósticos, la ventaja de su uso es el poder encontrar relaciones entre las variables de entrada y la variable de salida, esto significa que se trabajará con varias variables de entrada o independiente y una sola variable de salida o dependiente

Por ello nace la pregunta, ¿se puede desarrollar un modelo matemático que permita pronosticar la producción de una fábrica de barra de chocolate de presentación 20 g, utilizando regresión multivariada?

1.2 Delimitación del problema

Espacio

El presente trabajo de investigación se realiza en Ecuador en la Región Costa, Provincia del Guayas, Milagro – Simón Bolívar, en una pequeña empresallamada Luis Medina Idrovo.

Tiempo

La información obtenida está dentro del rango previsto de validez de una investigación científica, de 10 años; más la data histórica se resume a los dos últimos años de funcionamiento.

Universo

El estudio se basará en la propuesta de un Modelo Matemático para la producción, por lo cual se usarán los datos históricos de producción de los últimos 2 años.

1.3 Formulación del problema

¿De qué manera el uso de la regresión multivariada ayudara en el desarrollo de un modelo matemático que permita pronosticar de manera ágil la producción en una fábrica de barras de chocolate de presentación 20 g?

1.4 Sistematización del problema

¿Mediante el análisis de los componentes de un modelo matemático se podrá identificar los factores que afecta la fabricación o elaboración de la barra de chocolate de presentación 20 g?

¿Será posible realizar un análisis de la variabilidad de cada uno de los factores que afecta la fabricación o elaboración de barra de chocolate de presentación 20 g?

¿Será posible construir un modelo matemático apropiado que permitapronosticar de manera ágil la producción aplicándolo a una fábrica de barra de chocolate de presentación 20 g utilizado regresión multivariada?

1.5 Objetivos

1.5.1 Objetivo General

Proponer un modelo matemático apropiado que permita pronosticar de manera ágil la producción en una fábrica de barras de chocolate de presentación 20 g., utilizando regresión multivariada.

1.5.2 Objetivos Específicos

- Identificar los factores que inciden en la producción de barra de chocolate de presentación 20 g.
- Realizar un análisis de la variabilidad de cada uno de los factores que afecta la fabricación o elaboración de barra de chocolate de presentación 20 g.
- Construir un modelo matemático apropiado que permita pronosticar de manera ágil la producción aplicándolo a una fábrica de barra de chocolate de presentación 20 g utilizando regresión multivariada.

1.6 Justificación e importancia

Las MYPYMES se enfrentan continuamente a distintos problemas que tienen relación con la precisión y exactitud de los resultados dados por los modelos matemáticos, que han sido utilizados para describir los distintos factores que intervienen y afectan al proceso de producción (FAO, 2021).

El error presentado por los modelos empleados en los procesos de producción se mide con base en la diferencia entre la producción real y lo dado por el modelo matemático, estos errores son inevitables y siempre estarán presentes, dado que el modelo es una abstracción de la realidad problemática; por lo que error resultante siempre se desea que sea lo más pequeño posible, entonces se convierte en una medida de que tan bien el modelo se aproxima a la realidad (Garcés & Jaimes Barragán, 2015).

Esta investigación tiene como propósito presentar un modelo matemático que permita a la empresa pronosticar la producción de forma más precisa, reduciendo el producto sobrante no vendido, disminuyendo los costos de inventarios, de forma tal que la compañía mejore su posición en el mercado al reducir sus pérdidas.

1.6.1 Justificación Teórica

Dada la variante demanda de los consumidores y a la necesidad de ahorro en costo de producción por parte de las MYPYMES, resulta muy necesario poder entender cómo se elabora y desarrolla un modelo matemático aplicado a la producción, teniendo en cuenta el análisis de los datos de producción, resulta importante comprender el desarrollo y aplicación de la regresión multivariada (Banco Central del Ecuador, 2017).

1.6.2 Justificación metodológica

El presente trabajo de investigación se basa en el análisis metodológico de los distintos factores que afectan el pronóstico de producción como son: horas reales trabajadas (Hr), horas de paras no programadas (Hpn), horas de paras programadas (Hpp), horas hombres trabajadas (HH), retrabajo (kg). Para ellos se plantea poder discriminar o identificar cuáles son las que más o menos aportan en la elaboración del modelo matemáticos

En el uso de la regresión multivariada será importante poder establecer el grado de significancia de cada una de las variables independientes o intervinientes en el modelo, y el valor de R cuadrado, dado que este valor explica que tan bien el modelo se ajusta a los valores suministrado para su entrenamiento.

1.6.3 Justificación práctica

Este trabajo de investigación se justifica porque con el modelo matemático que se obtenga, permitirá a la MYPYMES realizar un mejor pronóstico de producción, dado que se identificará de manera precisa cuales son las variables internas que inciden en la producción, con lo que se tendrá un mejor control.

Con este modelo se espera que sea más cercano a la producción real de la microempresa y pueda extenderse a cualquier otro producto fabricado o elaborado por la fábrica; siempre que se tengan en cuenta las variables que afectan a la producción de nuevos productos.

1.7 Hipótesis

1.7.1 Hipótesis General

Si se diseña un modelo matemático utilizando regresión multivariada, se logrará pronosticar de manera eficiente la producción de barras de chocolate de presentación 20 g.

1.7.2 Hipótesis Particular

Es posible poder identificar los factores que afecta la fabricación o elaboración de barras de chocolate de presentación 20 g.

Es posible poder discriminar que factores influyen en la variabilidad de un modelo de pronósticos de producción de barras de chocolate de presentación 20 g.

Es posible construir un modelo matemático apropiado que permita pronosticar de manera ágil la producción, aplicándolo a una fábrica de barras de chocolate de presentación 20 g, utilizado regresión multivariada.

1.7.3 Conceptualización y operacionalización de variables

Variable independiente: regresión multivariada

Variable dependiente: pronóstico de producción

Tabla 1

Operacionalización de variable

L. de Trabajo	Definición conceptual	Indicadores	Técnica
Regresión Multivariada	Relación entre varias variables independientes (predictoras o explicativas) y otra variable dependiente (criterio, explicada, respuesta)	Valor de los coeficientes de la regresión multivariada.	Prueba de ajuste

Pronóstico de Producción	El pronóstico es la cantidad de kg de barras de chocolate de presentación 20 g., que debe de producir la empresa.	Cantidad (el número total de los datos) Valor máximo (el mayor número de los datos). Valor mínimo (el menor número de los datos).	Registro de datos históricos.
--------------------------	---	---	-------------------------------

Elaborado por: El autor.

CAPÍTULO II

2.1 Marco Teórico

2.1.1 Modelo matemático.

El desarrollo o elaboración de los modelos matemáticos es una de las herramientas utilizadas para analizar y estudiar problemas de diferentes campos del conocimiento. El objetivo de los modelos matemáticos es poder describir, explicar y predecir fenómenos y procesos en diferentes campos del conocimiento. (Montesinos-López & Hernández-Suárez, 2007).

Es importante señalar que un modelo matemático se define por las relaciones que combina entre las variables independientes y la variable dependiente. Estas relaciones son independientes de los datos que se pretendan incluir en el modelo matemático, dado que el modelo matemático se puede utilizar en diferentes circunstancias y en diferentes contextos (Aravena y otros, 2008).

El modelo matemático es una representación, en lenguaje matemático, de un objeto que existe en un universo no-matemático. Gran parte de las aplicaciones de las matemáticas como el cálculo (por ejemplo, optimización) implican el uso de los modelos matemáticos. (Rodríguez Velázquez & Steegmann Pascual, 2013).

En grandes lineamientos, los modelos matemáticos se pueden desarrollar en tres fases:

- Construcción del modelo. Convierte objetos no matemáticos a lenguaje matemático.
- Análisis de modelos. Investiga modelos matemáticos.
- Explicar el análisis matemático. Aplica los resultados de la investigación matemática para el objeto no matemático original.

El buen desempeño o la inexactitud de los modelos matemáticos está dado o depende de la exactitud con que el modelo matemático se ajuste a los objetos o

datos originales, y no depende de la exactitud con que las matemáticas analizan al modelo en sí (Rodríguez Velázquez & Steegmann Pascual, 2013).

De esta manera, los modelos matemáticos deben ser presentados como un nexo que relaciona la teoría matemática y el mundo cotidiano, por lo que estos se desarrollan en una opción didáctica para su explicación, con una actitud crítica y sistémica del pensamiento (Rodríguez-Gallegos & Bourguet-Díaz, 2015).

Por lo tanto, se puede decir que un modelo es una representación matemática simplificada de una realidad compleja (Plaza Gálvez, 2016). Modelar es la acción de construir un modelo, de encorsetar la realidad (Montesinos-López & Hernández-Suárez, 2007). Para la realización de los modelos matemáticos se debe de trabajar en equipo multidisciplinario que aportan diferentes perspectivas y conocimiento en la representación de la realidad. Por lo tanto, un modelo se convierte en una herramienta de ayuda en la toma de decisiones (Ramos y otros, 2010).

2.1.2 Pronóstico de producción.

La habilitación de variaciones en la tecnología de la información y la dinámica comercial cambiante, dictan que todos los modelos utilizados para el manejo de inventario deben revisarse y actualizarse de forma continua, lo que hace innecesario calcular los intervalos de revisión óptimos. Por un lado, los tiempos en la entrega de suministro tienden a acortarse debido a la cambiante tecnología del transporte; en cambio, este se ve complicado o perjudicado por la globalización y los tratados de libre comercio que cada día se van dando (Arango Marin y otros, 2013).

El pronóstico de venta en la actualidad se ha convertido en una vital fuente de datos para predecir la demanda de productos, de manera lo más apegada a la realidad del mercado. La aleatoriedad propia de la mayoría de los mercados de productos puede desarrollarse por medio de los modelos probabilísticos que tienen mejores predisposiciones de implementación informática. Todas las soluciones que tengan en cuenta el problema, desde la previsión de la demanda

hasta el cálculo de las cantidades de los pedidos, son necesarias para diseñar modelos fiables y eficientes (Sánchez-López y otros, 2013).

Las pequeñas y medianas empresa necesitan saber la cantidad de compra que el mercado requiere de cada producto que vende, para de esta manera tener suficiente cantidad de cada uno de sus productos, lo que le permitirá satisfacer la demanda de sus consumidores, con poco riesgo de obsolescencia o deterioro por inventario y el mayor costo de mantenimiento de producto no vendido. (Medina Varela y otros, 2009).

Las decisiones referentes al tamaño y ubicación de la planta, así como las de elección de los procesos productivos y equipos a utilizar, son el primer paso que debe de dar la pequeña empresa para intentar alcanzar un equilibrio entre demanda y producción en el largo plazo. Si solo se considerada a corto plazo, la variabilidad de la demanda del mercado con respecto al producto resulta mucho más elevada, por lo que es preciso tomar medidas correctivas para resolver este problema. Es por ello que la planificación y la programación de la producción toma relevancia dentro de cada micro y pequeña empresa (Escobar-Gómez y otros, 2010).

El pronóstico de producción es una predicción o supuesto de lo que podría suceder en un futuro, por lo que se convierte en una extensión del pasado, aunque bajo la acción de cualquier fenómeno aleatorio conlleva a inciertos, es por ello que resulta en un proceso en el cual se utilizan métodos cuantitativos y cualitativos para su desarrollo (Alonso Galicia & Villegas Valladares, 2005).

La demanda de producción puede ser pronosticada por distintos métodos, la selección de estos depende de un grupo de factores o variables tales como: el tiempo de los datos, la existencia de patrones o tendencia y estacionalidad del producto, pero la mayor influencia la tiene el comportamiento o tendencia de la demanda del producto y saber o entender las causas que producen dicha demanda (Saucedo Castillo y otros, 2010).

Entre los métodos más comúnmente utilizados para pronosticar se tienen: las series de tiempo, las regresiones lineales simples y las regresiones lineales múltiple, y los métodos cualitativos. Las series de tiempo y los métodos de regresión resultan ser técnicas estadísticas o métodos cuantitativos, en los cuales se debe utilizar un conjunto de datos de la demanda histórica, con la finalidad de poder predecir cuál será la demanda futura. Por otro lado los métodos cualitativos se basan en el uso los juicios de valores de expertos para la elaboración de los pronósticos de producción o de la demanda del mercado (Saucedo Castillo y otros, 2010).

Existen distintos factores a tener en cuenta al momento de elegir un modelo de pronóstico apropiado a aplicar en cada situación, se debe de considerar las diferencias existentes en los comportamientos de la demanda y de la dispensación en cada uno de los puntos de distribución (Perez y otros, 2012).

2.1.3 Regresión Multivariada

En la regresión multivariada se trata de ajustar modelos lineales o linealizables entre una variable dependiente y dos o más variables independientes. En este tipo de modelos matemático resulta vital poder testar la heterocedasticidad, la multicolinealidad y la especificación (Montero Granados, 2016).

En la regresión multivariada se utiliza más de una variable independiente o explicativa; presenta la ventaja de poder incorporar mayor información en la elaboración del modelo matemático y, consiguientemente, realizar estimaciones o predicciones más apegadas a realidad o precisas (Rojo Abuín, 2007).

En la regresión multivariada, la variable dependiente es una función lineal de k elementos proporcionadas por las variables independientes y una perturbación aleatoria o error. El modelo matemático además incorpora un término independiente. Si designamos por y a la variable dependiente o variable de respuesta, por $x_1, x_2, x_3, x_4, \dots, x_k$ a las variables independientes o explicativas; por ε al error o perturbación aleatoria, el modelo de regresión multivariado o regresión lineal múltiple viene dado por la siguiente expresión:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_k x_{ki} + \varepsilon_i$$

Los parámetros $\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k$ son fijos y desconocidos.

β_0 = intersección en Y

β_1 = pendiente de Y con la variable X_1 manteniendo constante las variables X_2, X_3, \dots, X_k

β_2 = pendiente de Y con la variable X_2 manteniendo constante las variables X_1, X_3, \dots, X_k

β_3 = pendiente de Y con la variable X_3 manteniendo constante las variables $X_1, X_2, X_4, \dots, X_k$

.

.

.

β_k = pendiente de Y con la variable X_k manteniendo constante las variables $X_1, X_2, X_3, \dots, X_{k-1}$

ε_i = es el término de error aleatorio en Y para la observación en i , que tiene la media 0 y la varianza σ^2

El modelo muestral estimado es:

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{3i} + \dots + b_k x_{ki} + e_i$$

donde e_i es el residuo o diferencia entre el valor observado de Y y el valor estimado de Y obtenido utilizando los coeficientes estimados b_j , donde $j = 1, \dots, k$. El método de regresión obtiene estimaciones simultáneas b_j de los coeficientes del modelo poblacional β_j utilizando el método de mínimos cuadrados

Un asunto de gran interés será responder a la siguiente pregunta: de un vasto conjunto de variables explicativas: $x_1, x_2, x_3, x_4, \dots, x_k$, cuáles son las que más influyen en la variable dependiente Y .

Supuestos habituales de la regresión multivariada

El modelo de regresión poblacional múltiple es:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_k x_{ki} + \varepsilon_i$$

y suponemos que se dispone de n conjuntos de observaciones. Se postulan los siguientes supuestos habituales para el modelo.

1. Las x_{ji} son bien números fijos, o bien relaciones de variables aleatorias X_j , que son independientes de los términos de error, ε_i . En el segundo caso, la inferencia se realiza condicionada a los valores observados de las x_{ji} .
2. El valor esperado de la variable aleatoria Y es una función lineal de las variables independientes X_j .
3. Los términos de error son variables aleatorias cuya media es 0 y que tiene la misma varianza σ^2 . Este último supuesto se denomina homocedasticidad o varianza uniforme.

$$E[\varepsilon_i] = 0 \quad \text{y} \quad E[\varepsilon_i^2] = \sigma^2 \quad \text{para } (i = 1, 2, 3, \dots, n)$$

4. Los términos de error aleatorio ε_i , no están correlacionados entre sí, por lo que

$$E[\varepsilon_i \varepsilon_j] = 0 \quad \text{para todo } i \neq j$$

5. No es posible hallar un conjunto de números que no sean iguales a cero, $c_0, c_1, c_2, c_3, \dots, c_k$, tal que

$$c_0 + c_1 x_{1i} + c_2 x_{2i} + c_3 x_{3i} + \dots + c_k x_{ki} = 0$$

esta es la propiedad de la ausencia de relación lineal entre las X_j

Método de mínimos cuadrados

El método de mínimos cuadrados para la regresión multivariada calcula los coeficientes estimados para minimizar la suma de los cuadrados de los residuos.

Recuérdese que el residuo es

$$e_i = y_i - \hat{y}$$

donde y_i es el valor observado de Y e \hat{y} es el valor de Y predicho a partir de la regresión.

En términos formales, minimizamos SCE :

$$SCE = \sum_{i=1}^n e_i^2$$

$$SCE = \sum_{i=1}^n (y_i - \hat{y})^2$$

$$SCE = \sum_{i=1}^n (y_i - (b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{3i} + \dots + b_kx_{ki}))^2$$

Descomposición de la suma de los cuadrados y coeficientes de determinación.

Se comienza con el modelo de regresión multivariada ajustado mediante mínimos cuadrados

$$y_i = b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{3i} + \dots + b_kx_{ki} + e_i = \hat{y} + e_i$$

donde las b_j son las estimaciones por mínimos cuadrados de los coeficientes del modelo de poblacional y las e_i son los residuos del modelo de regresión estimado.

La variabilidad del modelo puede dividirse en los componentes,

$$STC = SCR + SCE$$

donde estos componentes se definen de la siguiente forma

Suma total de los cuadrados

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

$$STC = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Suma de los cuadrados de los errores (residuos)

$$SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2$$

Suma de los cuadrados de la regresión

$$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 +$$

Esta descomposición puede interpretarse de la forma siguiente:

Variabilidad muestral total = variabilidad explicada + variabilidad no explicada

El coeficiente de determinación r^2 de la regresión ajustada es la proporción de la variabilidad muestral total explicada en la regresión.

$$r^2 = \frac{SCR}{STC} = 1 - \frac{SCE}{STC}$$

y se deduce que:

$$0 \leq r^2 \leq 1$$

Los coeficientes son elegidos de forma que la suma de cuadrados entre los valores observados y los pronosticados sea mínima, es decir, que se va a minimizar la varianza residual.

Coeficiente de determinación ajustado.

Algunos expertos indican que, al desarrollar los modelos de regresión multivariado, se debe utilizar **r² ajustada**, para indicar el número de variables independientes en el modelo y el tamaño de la muestra. Es vital indicar la **r² ajustada** al confrontar dos o más modelos de regresión que predicen la misma variable dependiente, pero cuenta con distintos números de variables independientes.

r² ajustada

$$r_{aj}^2 = 1 - [(1 - r^2) \frac{n - 1}{n - k - 1}]$$

donde k es el número de variables independientes en la ecuación de regresión

$$r_{aj}^2 = 1 - \frac{\frac{SCE}{(n - k - 1)}}{\frac{STC}{(n - 1)}}$$

Se utiliza esta medida para tener en consideración el hecho de que las variables independientes irrelevantes provocan una pequeña reducción de la suma de los cuadrados de los errores. Es por ello que, el valor de r_{aj}^2 permite comparar de mejor manera los modelos obtenidos de la regresión multivariada, en los cuales se tienen diferentes números de variables independientes.

Coeficiente de correlación múltiple.

El coeficiente de correlación múltiple es una correlación entre el valor predicho y el valor observado de la variable dependiente.

$$r = r(\hat{y}, y) = \sqrt{r^2}$$

y es igual a la raíz cuadrada del coeficiente múltiple de determinación. Se debe utilizar r como otra medida de la relación existente entre la variable dependiente y las variables independientes.

Es por ello que resulta similar a la correlación entre Y y X en el modelo de la regresión lineal simple.

2.2 Marco Conceptual

Los siguientes conceptos han sido tomados de los libros: “Funciones elementales para construir modelos matemáticos” (Bocco, 2010), “Probabilidad y estadística para ingeniería y ciencias” (Walpole y otros, 2012)

- **Análisis exploratorio de datos:** El análisis exploratorio de datos (EDA por sus siglas en inglés) implica el uso de gráficos y visualizaciones para explorar y analizar un conjunto de datos.
- **Análisis multivariante:** El análisis multivariante permite predecir para una observación determinada en base a su patrón de covariables, el valor de una variable continua o la probabilidad de ocurrencia de una variable dicotómica.
- **Heterocedasticidad:** se produce cuando la varianza (del error) no es constante en las observaciones llevadas a cabo.
- **Media aritmética:** La media aritmética es un tipo de media que otorga la misma ponderación a todos los valores.
- **Medidas de dispersión:** Las medidas de dispersión tratan, a través del cálculo de diferentes fórmulas, de arrojar un valor numérico que ofrezca información sobre el grado de variabilidad de una variable.

- **Medidas de tendencia central:** Las medidas de tendencia central son medidas estadísticas que pretenden resumir en un solo valor a un conjunto de valores. Representan un centro en torno al cual se encuentra ubicado el conjunto de los datos
- **Modelo linealizable:** un modelo de relación entre ambas variables (x , y) puede ser linealizable si una de las dos variables logra convertirse en relación lineal.
- **Modelo matemático:** es una simple representación de la realidad, mediante el uso de funciones que describen su comportamiento o ecuaciones que expresan sus relaciones.
- **Multicolinealidad:** es la relación de carácter dependiente fuertemente lineal entre más de 2 variables.
- **Pronóstico:** el pronóstico es una estimación del comportamiento de variables estadísticas en eventos futuros.
- **Regresión lineal múltiple:** el modelo de regresión lineal múltiple se usa como explicación hacia las variables de respuesta continua relacionadas con varios factores o variables explicativas continuas.
- **Regresión lineal:** es una técnica de modelado estadístico que se utiliza para describir variables de respuesta continua en función de uno o más predictores.
- **Representaciones Gráficas:** Una representación gráfica o gráfico estadístico es un resumen visual de la tabla de frecuencias, que nos permite la representación en forma más resumida y total del experimento o fenómeno en estudio.
- **Variable Continua:** Una variable continua es aquella que puede tomar un número infinito de valores entre dos valores cualesquiera de una característica.
- **Variable Cuantitativa:** Una variable cuantitativa es aquella variable estadística que, a diferencia de la cualitativa, puede expresarse a través de cifras. Por esta razón, puede analizarse con métodos estadísticos.
- **Variable dependiente:** variable objeto del estudio y que sus resultados procuran explicar por medio de las variables denominadas independientes.

- **Variable Discreta:** Una variable es discreta cuando no puede tomar ningún valor entre dos consecutivos.
- **Variable Nominal:** Una variable es nominal cuando sus valores representan categorías que no obedecen a una clasificación intrínseca
- **Variable Ordinal:** Una variable es ordinal cuando sus valores se pueden clasificar en categorías y estas se pueden ordenar en jerarquías con respecto a la característica que se evalúa
- **Varianza:** La varianza es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media.

2.3 Marco Referencial

Los siguientes trabajos de investigación sirvieron como referencia para el presente trabajo de investigación, dado que de una u otra manera abordan el tema de investigación y presentan interesantes aportes tanto en la metodología como en la investigación.

Alcívar (2000) en su trabajo de investigación “Análisis y valoración del sistema crediticio en el Ecuador, tanto para Bancos como Instituciones financieras”, plantea una interesante comparación entre los modelos de regresión lineal y lo modelo de regresión multivariado, analiza el comportamiento del crédito en el país, en los últimos años de la década del 90, presentando un conjunto de datos extensos e interesantes. Analiza que tan bueno es un modelo en comparación con el otro, utiliza tabla de contingencia, el coeficiente de correlación de Spearman (Alcívar Zavala, 2000)

Limones (2012, en su trabajo “Método para el control de la producción defectuosa dentro del proceso de conformado de metales. Caso ZF Sachs Automotive México”, presenta un método, que permita analizar el estado actual de los procesos de fabricación en la empresa, además sirve para enfrentar el problema de la variabilidad dentro del proceso. El método planteado dentro de su investigación utiliza la probabilidad del modelo de regresión multivariada, así como análisis de datos (Limones Lara, 2012).

Montoya (2017), en su trabajo “Factores socioeconómicos que afectan al consumo doméstico de agua potable usando la regresión múltiple, puno ciudad – 2015”, desarrolla un modelo matemático usando regresión multivariada aplicado al consumo mensual de agua potable en la ciudad de Puno, la población de análisis fueron 19209 viviendas, utilizó muestreo estratificado con afijación óptima. Para su modelo analizó 19 variables independientes, a las cuales se les aplicó un proceso de selección. Posteriormente el modelo matemático fue validado mediante la utilización de la regresión multivariada (MONTTOYA VALER, 2017).

Martínez (2020), en su trabajo de investigación denominado “Modelo de regresión lineal múltiple para el pronóstico de ventas de bolsas ecológicas para la empresa Boleco S.A, en la ciudad de Bogotá D.C.”, plantea una selección de cuáles considera que deben de ser las principales variables intervinientes en el modelo de regresión lineal múltiple, que para este trabajo fueron: Capacidad de la Bolsa, Clientes, Ventas y Costo de ventas; finalmente validó el modelo con los resultados obtenidos (Forero Gómez & Martínez Lozano, 2020).

Rojas (2020) en su trabajo “Principales factores de riesgos psicológicos que influyen en los accidentes de trabajo en la minera aurífera Retamas S.A.”, presenta un modelo matemático utilizando regresión multivariada con variables independientes como: fatiga emocional, la disociación y la ansiedad por el logro personal, para ello analizaron 20 ítems, que miden la identificados como principales riesgos psicológico que se relacionan con los accidentes de trabajo (ROJAS VICTORIO, 2020).

CAPÍTULO III

MARCO METODOLÓGICO

3.1 Tipo de Investigación

Investigación Exploratoria. -

La presente investigación es de tipo exploratoria, debido a que por medio de esta se realiza una aproximación a un tema que no ha sido investigado muy a fondo, por eso a través de la recolección de los datos se podrá detectar patrones a los cuales se dará una explicación con sustento teórico y práctico que permitirá desarrollar una propuesta de un modelo matemático, logrando así el objetivo de esta investigación.

Investigación Documental. -

Esta investigación es documental debido a que para el desarrollo de la misma se emplearon diferentes fuentes documentales como son libros, tesis, artículos científicos entre otros documentos que respaldan los diversos temas que se tratan en la presente investigación.

Investigación de Campo. -

Para el desarrollo de la investigación se empleará la investigación de campo debido a que el estudio debe realizarse directamente en el lugar de los hechos (la microempresa) para comprender la situación real del entorno problemático, donde los investigadores están directamente involucrados en el estudio y examinar cómo se desenvuelve cada variable de estudio para poder interpretarla y a su vez consultar con un grupo de personas y fuentes de la microempresa.

Investigación Aplicada. -

La investigación es aplicada porque se busca la generación de conocimiento con la aplicación directa a un problema del sector productivo. Se espera desarrollar un modelo matemático que ayude a mejorar el pronóstico de la producción. Se espera poder predecir el pronóstico de producción en base a la afectación de varias variables independientes.

Investigación Cuantitativa. -

La investigación es de tipo cuantitativa porque se recolecta y analiza datos numéricos, es decir los datos de producción semanal. Con este método se procederá a identificar una tendencia o comportamiento de la producción, que permita realizar predicciones, comprobar relaciones y obtener resultados generales de la producción. Se espera poder ayudar a mejorar los niveles de producción, con la finalidad de que la microempresa pueda tener una toma de decisiones más precisa.

3.2 Diseño De La Investigación

El objetivo de la presente investigación será proponer un modelo matemático apropiado que permita poder pronosticar de manera ágil la producción en una fábrica de barra de chocolate de presentación 10 g utilizando regresión multivariada

El presente diseño de investigación es No Experimental, dado que se procederá a la observación y análisis del fenómeno sin manipular las variables.

Este trabajo es de tipo longitudinal, dado que estamos analizando lo que pasa o cual es el comportamiento en diferentes puntos de tiempo de las distintas variables independientes que podrían intervenir dentro de la propuesta del modelo matemático.

3.3 Población y Muestra

Para la presente investigación se trabajará con todos los datos, es decir el 100% de los datos recolectados de la producción de la micro empresa de su producto barras de chocolates de 20 g, de los 2 últimos años; los datos recolectados corresponden a la producción semanal, cabe señalar que existieron semanas en las cuales la micro empresa no producía este producto debido a la poca demanda del mercado.

3.4 Variables de investigación

3.4.1 Variable dependiente:

La variable dependiente a considerar es Cantidad de Producción Real CPR (kg), este dato se lo obtuvo de la información suministrada por la empresa, corresponde a los kilogramos reales que la empresa produce de la barra de chocolate de 20 g.

Tabla 2

Variable dependiente

Variable dependiente	Unidad de medida
Cantidad de Producción Real	kilogramos

Elaborado por: El autor.

3.4.2 Variable independiente:

Como primera aproximación se analizarán 3 variables independientes, y con esto, se espera ver el grado, peso o aporte de cada una que estas tienen dentro del modelo matemático. Si existiera una variable que aporte muy poco o no aporte al modelo, deberá no ser tomada en cuenta o eliminada.

- a) Cantidad de Producción Programada CPP (kg).

La Cantidad de Producción Programada, son la cantidad de kg de producto de barra de chocolate de 20 g., que la micro empresa desea o espera producir, ese es el valor teórico, deseado o planificado.

Tabla 3

Cantidad de Producción Programada

Variable independiente	Unidad de medida
Cantidad de Producción Programada.	kilogramos

Elaborado por: El autor.

- b) Horas Laborales Planificadas HLP (hs).

Las Horas Laborales Planificadas corresponden a las horas planificadas o teóricas que las personas tienen que trabajar durante la semana para

poder cumplir con la producción, es lo que se cree que se puede hacer sin tener en cuenta los imprevistos para cumplir con los objetivos. A estas empresas las conoce o las llama las Horas Hombre.

Tabla 4

Horas Laborales Planificadas

Variable independiente	Unidad de medida
Horas Laborales Planificadas	Hora

Elaborado por: El autor.

c) Paradas No Programadas PNP (hs).

Las Paradas No Programadas corresponde al tiempo que se pierde porque la maquina no trabaja debido a una avería, algún desperfecto o daño, es decir la maquina entra en mantenimiento correctivo, y sale de la línea de producción. Este tiempo afecta la producción, por lo que su valor mientras más pequeño sea es mucho mejor.

Tabla 5

Paradas No Programadas

Variable independiente	Unidad de medida
Paradas No Programadas	Hora

Elaborado por: El autor.

3.5 Método, Técnica y Uso de Software de Tratamiento y Análisis de Datos

El método utilizado es el cuantitativo y las técnicas de análisis es la regresión multivariada o regresión lineal múltiple, evaluando los supuestos básicos del modelo de regresión, se realizará un análisis de variables excluidas para ver si el grado de aporte al modelo, es decir si alguna de las variables independientes debe de excluirse; se analiza la bondad de ajuste a través del coeficiente de determinación (R^2) y el análisis de varianza (anova), finalmente se obtendrán los coeficientes de la regresión multivariada y su grado de significancia para las variables mencionadas.

La información será procesada mediante el software estadístico IBM SPSS versión 24, para el tratamiento y análisis de los datos.

CAPITULO IV

PROPUESTA

4.1 Formulación del modelo de regresión lineal múltiple

En el desarrollo de la propuesta del modelo matemático utilizando la regresión multivariada, como primer paso se debe seleccionar o identificar la variable dependiente (Y) y las variables independientes (X), posteriormente, se presenta la selección de las variables y la ecuación.

La variable dependiente o de salida corresponde a la Cantidad de Producción Real (CPR), las variables independientes son: Capacidad de Producción Programada, Horas Laborares Planificadas HLP (hr), Paradas No Programadas PNP.

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3$$

Donde

$Y = \text{CPR}$

$X_1 = \text{CPP}$

$X_2 = \text{HLP}$

$X_3 = \text{PNP}$

4.2 Diagnósis y validación del modelo de regresión lineal múltiple.

Para interpretar bien un modelo de regresión, siempre debemos acompañar nuestra investigación con el diagnóstico y la validación del modelo. Este diagnóstico incluye analizar si se verifican los supuestos subyacentes del modelo.:

- Linealidad: Los parámetros y su interpretación no tienen sentido si los datos no provienen realmente de un modelo lineal, situación en la que las predicciones también pueden ser completamente erróneas.

- Normalidad de los errores: El modelo de regresión lineal asume que la distribución de los errores es Normal.
- Homocedasticidad: La varianza del error es constante.
- Independencia de los errores: Las variables aleatorias que representan los errores son mutuamente independientes.
- Las variables explicativas $X_1; X_2; \dots; X_n$, son linealmente independientes.

4.3 Supuesto de Normalidad de los Residuos

El supuesto de Normalidad tiene como objetivo demostrar que los residuos están distribuidos normalmente.

Como la muestra es mayor de 50 datos entonces debemos usar la prueba de Kolmogorov – Smirnov (K-S), y esta prueba nos indica que la significancia tiene que ser mayor a 0.05 para decir que los residuos corresponden a una muestra con distribución normal.

Tabla 6

Prueba de Kolmogorov-Smirnov para una muestra

		Standardized Residual
N		73
Parámetros normales ^{a,b}	Media	0,0000000
	Desviación estándar	0,97894501
Máximas diferencias extremas	Absoluta	0,234
	Positivo	0,224
	Negativo	-0,234
Estadístico de prueba		0,234
Sig. asintótica (bilateral)		0,000 ^c

a. La distribución de prueba es normal.

b. Se calcula a partir de datos.

c. Corrección de significación de Lilliefors.

Elaborado por: El autor.

Figura 1

Resumen de Prueba de hipótesis

	Hipótesis nula	Prueba	Sig.	Decisión
1	La distribución de Standardized Residual es normal con la media -0,00000 y la desviación estándar 0,979.	Prueba de Kolmogorov-Smirnov para una muestra	,000 ¹	Rechazar la hipótesis nula.

Se muestran significaciones asintóticas. El nivel de significación es de ,05.

¹Lilliefors corregida

Elaborado por: El autor.

Como podemos observar en la tabla 6, el nivel de significancia es de 0,000; por lo que valor de p (sig) < 0,05; así queda demostrado que se cumple con el supuesto de normalidad de los datos.

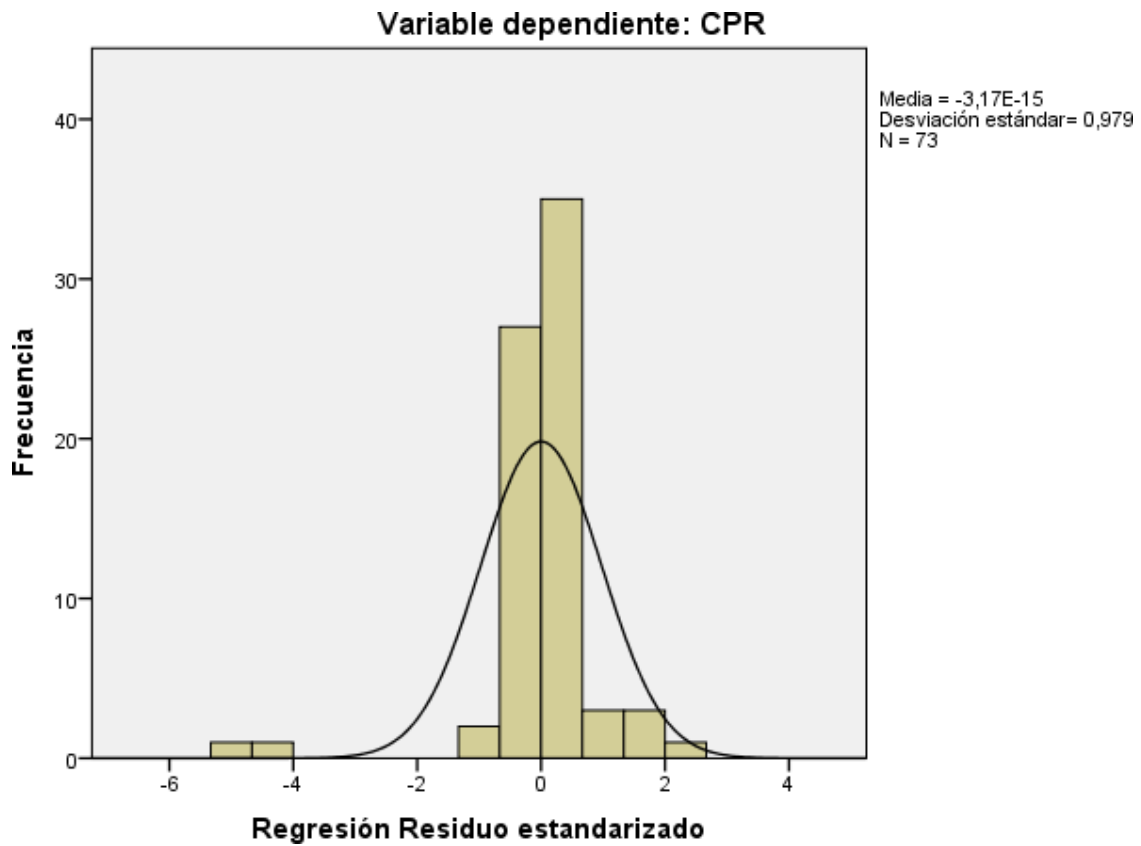
Como se puede observar en la tabla 6 y en la figura 1, se rechaza el supuesto de Hipótesis nula, esto quiere decir que los residuos tienen una distribución normal.

Normalidad de los Residuos

Otra forma de demostrar la normalidad de los residuos es usando la gráfica de histograma de Regresión de Residuos Estandarizados y la gráfica de Gráfico P-P Normal de Regresión Residuos Estandarizados

Figura 2

Histograma de regresión de Residuos Estandarizados

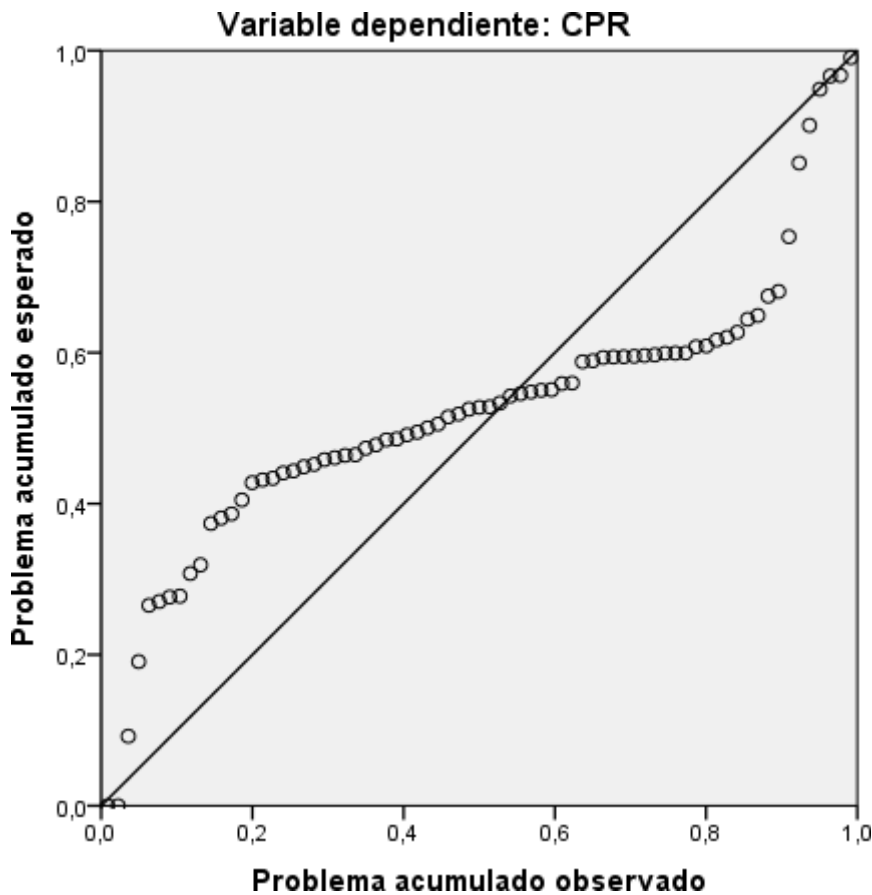


Elaborado por: El autor.

Como se puede observar figura 2, Histograma de regresión de Residuos Estandarizados, estos aparentemente siguen una distribución normal con media cero, pero tienen variabilidad y eso se puede observar en el extremo izquierdo donde existe un pequeño grupo de datos, por lo tanto, existe un peso de datos que alteran la normalidad de los residuos.

Figura 3

Gráfico P-P Normal de Regresión Residuos Estandarizados



Elaborado por: El autor.

Como se puede observar figura 3, Gráfico P-P Normal de Regresión Residuos Estandarizados, los datos se encuentran repartidos tanto por encima como por debajo de la línea, pero están separados de ella, con lo que se puede decir viendo la gráfica que estos no siguen una distribución normal, aunque tienen un comportamiento distribuido.

La no normalidad de los residuos puede ser causa de la aleatoriedad de las paradas no programadas; por lo que se podría tener como opción no tomar en cuenta estos datos para el modelo, pero se debe considerar que las paradas no programadas influyen en la producción de una empresa.

4.4 Supuesto de Independencia de las observaciones

Para responder al supuesto de la independencia de errores se utiliza la Prueba de Durbin-Watson. El criterio para decir que existe independencia de las observaciones, es que el valor de Durbin-Watson debe ser lo más cercana

posible a 2 con una oscilación de + - 1. Por lo tanto, los valores entre 1 y 3 están bien, para aceptar que los residuos son independientes

Tabla 7

Resumen del modelo^b

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin-Watson
1	0,977 ^a	0,954	0,952	96,55973	1,815

a. Predictores: (Constante), PNP, CPP, HLR

b. Variable dependiente: CPR

Elaborado por: El autor.

En la Tabla 7 se puede observar que el coeficiente de Durbin-Watson tiene un valor de 1,815; que es un valor aceptable muy cercano a 2, por lo que se da por satisfecho el supuesto de independencia de las observaciones.

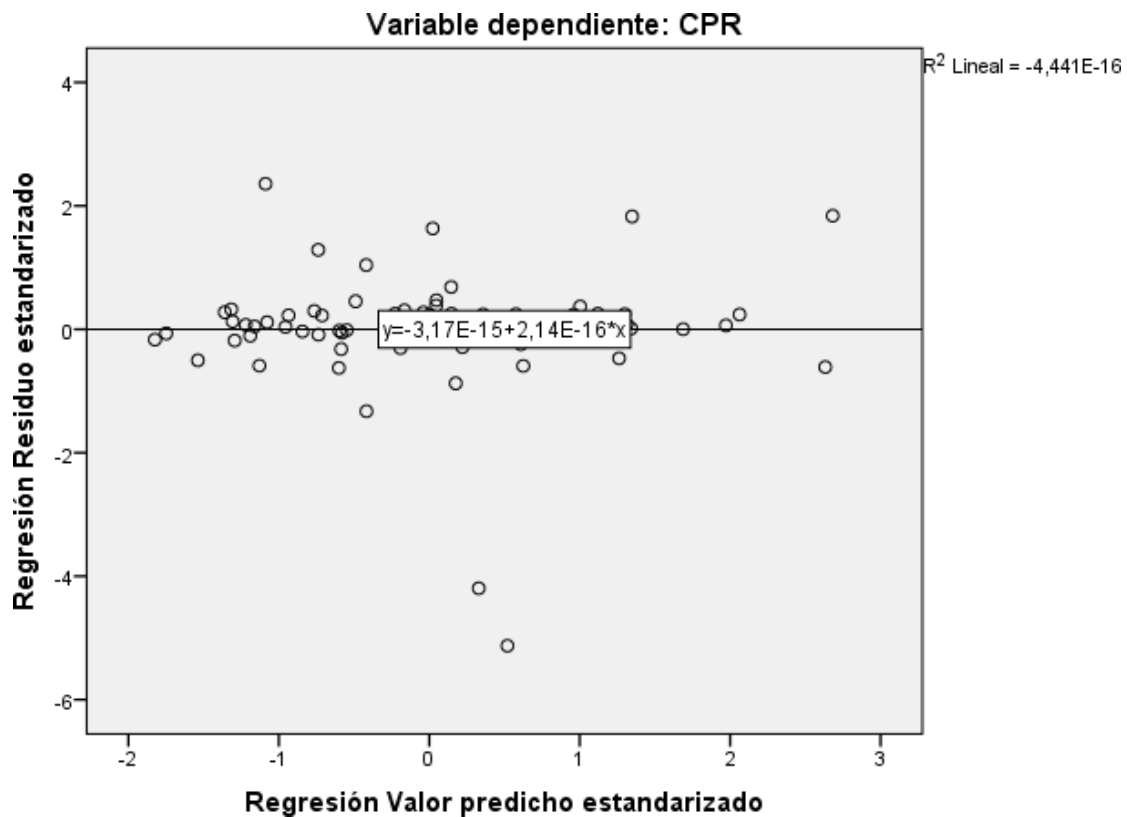
4.5 Supuesto de Homocedasticidad

La homocedasticidad es una característica de un modelo de regresión lineal que implica que la varianza de los errores es constante a lo largo del tiempo. Además, si una varianza, aparte de ser constante es también más pequeña, nos dará como resultado una predicción del modelo más fiable.

El supuesto de homocedasticidad se responderá de forma gráfica; para ello los residuos deben de distribuirse de manera homocedástica a lo largo de las puntuaciones predichas.

Figura 4

Gráfico de dispersión



Elaborado por: El autor.

A partir del examen de gráfico de dispersión de la figura. 4, se comprueba que los puntos están distribuidos razonablemente bien por lo que se puede decir que no hay ninguna relación sistemática entre los residuos tipificados y los valores pronosticados tipificados de la cantidad de producto a producir.

Por tanto, podemos dar por cumplido el supuesto en cuanto a la homocedastidad de los datos

4.6.- Supuesto de linealidad

El supuesto de la linealidad implica que la relación entre la variable dependiente y las independientes debe ser lineal.

Tabla 8

Correlaciones

		CPP	HLR	PNP	CPR
CPP	Correlación de Pearson	1	0,472**	0,163	0,975**
	Sig. (bilateral)		0,000	0,169	0,000
	N	73	73	73	73
HLR	Correlación de Pearson	0,472**	1	0,325**	0,499**
	Sig. (bilateral)	0,000		0,005	0,000
	N	73	73	73	73
PNP	Correlación de Pearson	0,163	0,325**	1	0,191
	Sig. (bilateral)	0,169	0,005		0,105
	N	73	73	73	73
CPR	Correlación de Pearson	0,975**	0,499**	0,191	1
	Sig. (bilateral)	0,000	0,000	0,105	
	N	73	73	73	73

** . La correlación es significativa en el nivel 0,01 (bilateral).

Elaborado por: El autor.

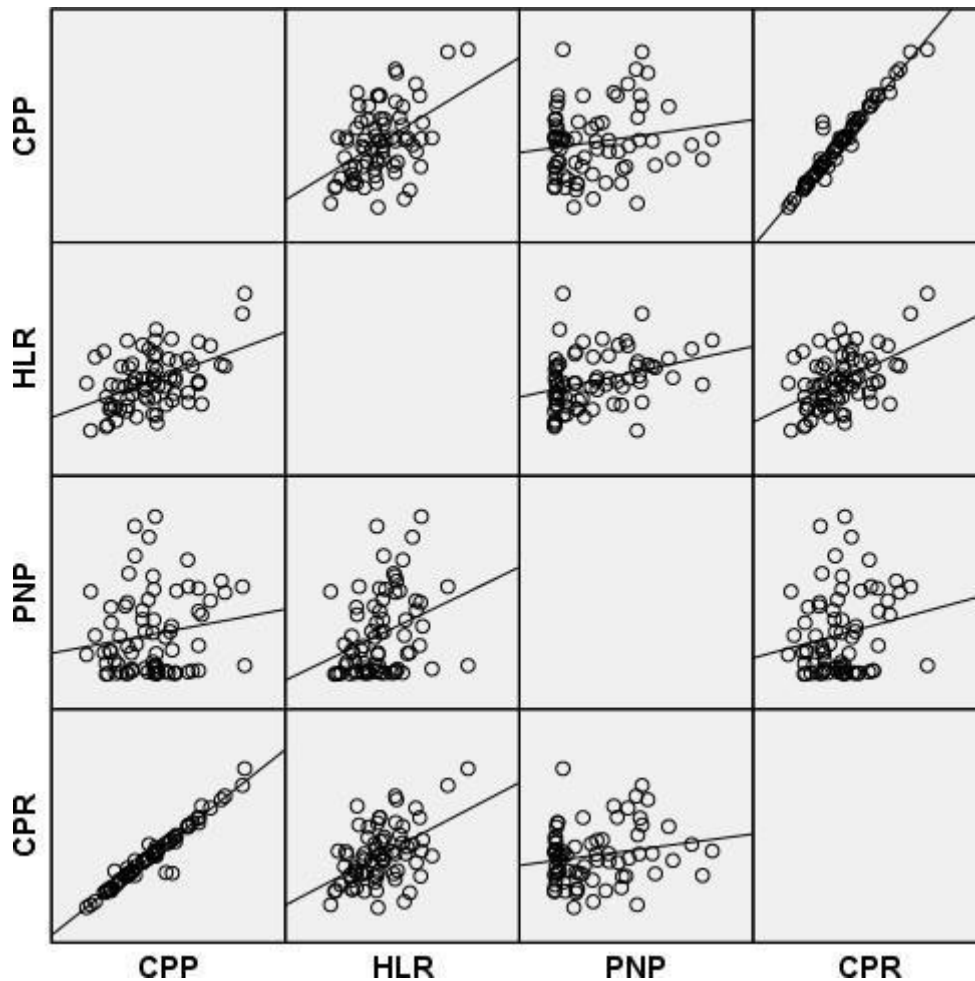
Como se puede observar en la tabla 8 de correlación, existe una correlación positiva entre todas las variables de entrada y la variable de salida.

En la matriz se observa que todas las variables independientes de las variables HLP, CPP, PNP tienen correlación con la variable dependiente CPR.

La variable independiente con mayor relación es la CPP con 0,975, luego la variable HLR con 0,499 y finalmente la variable PNP con 0,191. Por lo tanto, podemos afirmar que si existe correlación entre las variables independientes y la variable dependiente.

Figura 5

Correlación de variables



Elaborado por: El autor.

En la figura 5, se puede observar que se cumple el supuesto de linealidad entre la variable dependiente CPR y las variables independientes CPP, HLP, PNP, además que se puede observar en la figura existe una correlación positiva entre las variables independientes y la variable dependiente

Mediante la matriz de la tabla 8 y la figura 5 de correlación se puede decir que la variable PNP podría ser excluida dado que no aporta mucho al modelo, pero se ha decidido incluirla

Por lo anterior podemos afirmar que se cumple con el supuesto de linealidad de las variables.

4.7 Supuesto de multicolinealidad o supuesto de ausencia de multicolinealidad

Para el diagnóstico de colinealidad se usará el VIF (factor de varianza inflada), este supuesto es para saber si hay o no hay multicolinealidad entre las variables independientes. Ninguna variable independiente puede estar por encima de 10, caso contrario hay multicolinealidad entre las variables.

Para el presente modelo de regresión multivariada se tiene que ningún valor del VIF es mayor a 10.

Para este caso como se va a trabajar con la matriz de Diagnósticos de colinealidad, se debe de seguir los siguientes pasos:

1. Identificar los índices que estén por encima del umbral: 30
2. Para los índices identificados, identificar las variables con proporciones de varianza por encima del 90%: Habrá multicolinealidad si ocurre con dos o más coeficientes.

Tabla 9

Diagnósticos de colinealidad^a

Modelo	Dimensión	Autovalor	Índice de condición	Proporciones de varianza			
				(Constante)	CPP	HLR	PNP
1	1	3,408	1,000	0,01	0,01	0,01	0,03
	2	0,388	2,964	0,03	0,05	0,01	0,90
	3	0,119	5,348	0,47	0,84	0,04	0,02
	4	0,085	6,343	0,50	0,10	0,90	0,03

a. Variable dependiente: CPR

Elaborado por: El autor.

Se considera que a partir de un valor de 20 hay cierta multicolinealidad y que ésta es alta a partir de 30.

Como se puede observar en la tabla 9, ningún índice de condición es mayor a 30, por lo tanto, se puede concluir que No existe multicolinealidad dado que esto

con dos o más variables; por lo tanto, se da por satisfecho el supuesto de multicolinealidad.

4.8 El Modelo Matemático de Regresión Multivariada

El valor de R cuadrado es el coeficiente de determinación, este va de 0 a 1 es 0% al 100%, esto quiere decir que con las variables que se tiene solo se puede predecir un % de la ecuación lineal.

El valor del R-cuadrado ajustado tiene en cuenta el número de variables independientes que se utilizan para predecir la variable dependiente.

Modelo. - En este modelo se trabajó con tres variables independientes: la cantidad de producción planificada CPP, las horas laborales planificadas HLP, paradas no programadas PNP, y como variable dependiente la cantidad de producción real CPR, como se observa en este modelo el R^2 de 0,954, esto quiere decir que con esas tres variables independientes ya se puede explicar un 95,4% de la varianza; y se tiene un R^2 ajustado del 0,952

Como se puede observar se sustenta el trabajar con un Modelo Matemático con 3 variables, dado que este presenta un mejor ajuste a la realidad, y las 3 variables seleccionadas aportan positivamente a la predicción de la variable dependiente.

Análisis de la Varianza

Tabla 10

ANOVA^a

Modelo		Suma de cuadrados	GI	Media cuadrática	F	Sig.
1	Regresión	13240605,160	3	4413535,054	473,363	0,000 ^b
	Residuo	643340,949	69	9323,782		
	Total	13883946,110	72			

a. Variable dependiente: CPR

b. Predictores: (Constante), PNP, CPP, HLR

Elaborado por: El autor.

El p – value es menor a 0.05 por lo tanto el modelo es adecuado (se puede nombrar las hipótesis

El modelo matemático utilizado presenta un valor del estadístico de prueba F= 473,363 y tiene un P_ valor igual a $0 < 0.05$, por lo que se rechaza la hipótesis nula y se concluye que la dependencia lineal es estadísticamente significativa por lo que el modelo es adecuado.

Tabla 11
Coeficientes^a

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	T	Sig.	Estadísticas de colinealidad	
		B	Error estándar	Beta			Tolerancia	VIF
1	(Constante)	-27,046	29,549		-0,915	0,363		
	CPP	0,970	0,030	0,952	32,379	0,000	0,777	1,288
	HLR	1,558	1,152	0,041	1,353	0,181	0,714	1,401
	PNP	6,648	8,043	0,023	0,827	0,411	0,894	1,118

a. Variable dependiente: CPR
Elaborado por: El autor.

Teniendo en cuenta los resultados de la tabla 11 se deduce que el modelo matemático en base a la regresión multivariada para el modelo de pronóstico es

$$Y = -27,046 + 0,970 X_1 + 1,558 X_2 + 6,648 X_3$$

Donde:

$Y = \text{CPR}$

$X_1 = \text{CPP}$

$X_2 = \text{HLP}$

$X_3 = \text{PNP}$

Como se muestra en la ecuación, se ha podido construir una ecuación que permita predecir la cantidad de producción real en función de tres variables de entrada.

La ecuación de regresión estandarizadas nos muestra las variables en la misma dimensión de la siguiente manera:

$$Y = 0,952 X_1 + 0,041 X_2 + 0,023 X_3$$

4.9 Comparación de los datos reales con los datos pronosticados

A continuación, se procede a comparar los datos de producción real CPR con los datos obtenidos con el modelo matemático propuesto CPR_RM.

Tabla 12

Estadísticas de muestras emparejadas

		Media	N	Desviación estándar	Media de error estándar
Par 1	CPR_RM	797,9279	73	428,61822	50,16597
	CPR	798,3288	73	439,12707	51,39594

Elaborado por: El autor.

Como se puede observar la media de los datos obtenidos con el modelo matemáticos CPR_RM es de 797,9279, y la media de los datos reales observados CPR es de 798,3288, estas medias son muy semejantes o parecidas, con lo cual al existir una diferencia entre ellas muy pequeñas se puede decir que el modelo matemático es bastante bueno.

Tabla 13

Correlaciones de muestras emparejadas

	N	Correlación	Sig.
Par 1 CPR_RM & CPR	73	0,977	0,000

Elaborado por: El autor.

La correlación entre los datos construidos con el modelo matemático y los datos reales observados es de 0,977 que es un valor muy cercano a 1 y tiene un P_ valor igual a $0 < 0.05$, por lo que podemos decir que si existe correlación entre las variables.

Tabla 14

Prueba de muestras emparejadas

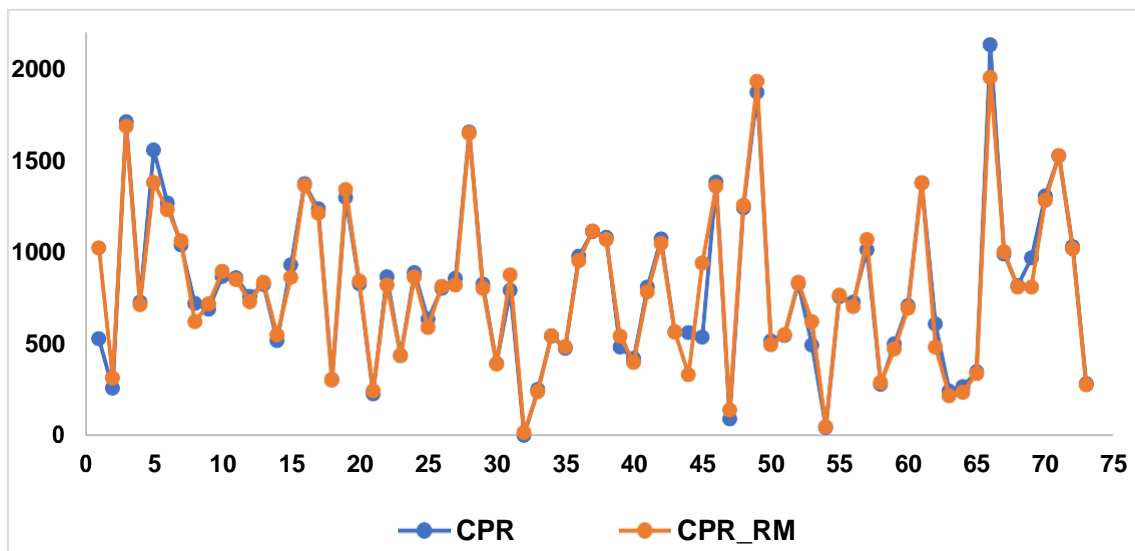
Par 1	Diferencias emparejadas						t	gl	Sig. (bilateral)
	Media	Desviación estándar	Media de error estándar	95% de intervalo de confianza de la diferencia					
				Inferior	Superior				
CPR_RM - CPR	-0,40089	94,52691	11,06354	-22,45565	21,65387	-0,036	72	0,971	

Elaborado por: El autor.

En la tabla 14, se puede observar que la diferencia emparejada entre las medias es muy baja, es de -0,40089, con lo cual se puede decir que es el error que se comete con el modelo matemático propuesto.

Figura 6

Comparación entre CPR_RM - CPR



Elaborado por: El autor.

En la figura 6, se puede observar que se realiza la comparación visual entre los valores reales obtenidos de la planta de producción es decir Capacidad de Producción Real CPR, con lo valores obtenidos mediante el uso del modelo matemático propuesto por medio de la regresión multivariado CPR_RM.

Se puede observar que el modelo presenta una muy buena aproximación, un buen ajuste, por lo que se puede decir que el modelo sirve para predecir.

CONCLUSIONES

Las variables más relevantes relacionadas con el proceso de producción de barras de chocolate de 20 gramos se han identificado y presentadas de la siguiente manera, variables independientes; cantidad de producción planeada CPP, horas laborales planificadas HLP, paradas no programadas PNP, variable dependiente: cantidad de producción real CPR.

La construcción de un modelo de regresión multi-variable se ejecuta tomando en cuenta las variables seleccionadas y luego los datos se ingresan en el programa SPSS.

Al verificar el modelo de regresión múltiple, se concluye que, para la microempresa dentro de su proceso de elaboración de galletas de 20 g, utilizando las variables independientes: cantidad de producción planeada CPP, horas laborales planificadas HLP, paradas no programadas PNP, se puede pronosticar la cantidad de producción real CPR, con el modelo matemático obtenido, dado que este tiene un nivel de significancia del 95,4% que es su R^2

RECOMENDACIONES

Se recomienda utilizar el modelo matemático obtenido por regresión multivariante en este estudio, el cual servirá como herramienta para que la microempresa pueda predecir de manera óptima la cantidad de galletas de barras de chocolate de 20 gramos que la microempresa podrá elaborar según unas condiciones dadas, con la finalidad de que el modelo matemático sirva como una herramienta para mejorar sus procesos.

Se debe tener en cuenta otras posibles variables independientes según lo considere el departamento de producción, para ello la microempresa debe desarrollar una cultura de la toma de datos, de forma tal que puedan servir para un posterior análisis.

Promover el desarrollo de investigaciones donde se promueva el desarrollo de modelos matemáticos mediante el uso de regresión multivariada en las distintas microempresas, dado que esto dará un mayor beneficio al sector al reducir sus diferencias entre lo que el mercado quiere y lo que se produce.

Si se viola el supuesto de normalidad, los resultados bajan en cierto grado su confiabilidad; por lo que se recomienda contar con personal idóneo para el área de productividad, con el fin de obtener muestras representativas de la población.

Cuando una variable independiente presenta una influencia baja en la variable dependiente (p mayor a 0,05); se puede sugerir la transformación logarítmica de los valores o en su defecto aumentar la recolección de datos; aunque en muchos casos esto representa un gasto alto de tiempo y dinero.

BIBLIOGRAFÍA

- Alcívar Zavala, J. A. (2000). *Análisis y valoración del Sistema Crediticio en el Ecuador, tanto para Bancos como Instituciones Financieras*. Guayaquil: ESPOL.
<https://www.dspace.espol.edu.ec/bitstream/123456789/4159/1/6687.pdf>
- Alonso Galicia, P., & Villegas Valladares, E. (2005). LA IMPORTANCIA DE LOS CRITERIOS CUALITATIVOS DE LOS PRONÓSTICOS EN LOS AGRONEGOCIOS. *Revista Mexicana de Agronegocios*, 9(16), 464-477.
<https://www.redalyc.org/articulo.oa?id=14101606>
- Arango Marin, J. A., Giraldo Garcia, J. A., & Castrillón Gómez, O. D. (2013). Gestión de compras e inventarios a partir de pronósticos Holt-Winters y diferenciación de nivel de servicio por clasificación ABC. *Scientia Et Technica*, 18(4), 743-747.
<https://www.redalyc.org/pdf/849/84929984023.pdf>
- Aravena, M., Caamaño, C., & Giménez, J. (2008). Modelos matemáticos a través de proyectos. *Revista latinoamericana de investigación en matemática educativa*, 11(1). http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-24362008000100003&lng=es&nrm=iso&tlng=es
- Banco Central del Ecuador. (2017). *Cuentas Nacionales*. Quito: Banco Central del Ecuador.
- Berkowitz, D. E. (1998). INDUSTRIA ALIMENTARIA. En *ENCICLOPEDIA DE SALUD Y SEGURIDAD EN EL TRABAJO* (págs. 67 - 67.35). Madrid: Chantal Dufresne, BA.
<https://www.insst.es/documents/94886/161971/Cap%C3%ADtulo+67.+Industria+alimentaria>
- Bocco, M. (2010). *Funciones elementales para construir modelos matemáticos*. Buenos aires: Instituto Nacional de Educación Tecnológica.
<http://www.bnm.me.gov.ar/giga1/documentos/EL001843.pdf>
- Cabrera González, G., & De León Arias, A. (2019). Modelación markoviana para identificar la dinámica y pronóstico del índice de producción industrial en México de 1980 a 2018. *EconoQuantum*, 16(2), 23 - 41.
<http://econoquantum.cucea.udg.mx/index.php/EQ/article/view/7120/6215>

- Damián Llatas, M. R., & Sandoval Santamaría, N. J. (2018). *Modelo óptimo de Pronóstico del índice mensual de Producción de Electricidad"-Perú en el Periodo 2006 – 2015*. Facultad de Ciencias Físicas y Matemáticas. Lambayeque: Universidad Nacional Pedro Ruiz Gallo.
<http://repositorio.unprg.edu.pe/handle/UNPRG/2192>
- Escobar-Gómez, E. N., Díaz-Núñez, J. J., & Taracena-Sanz, L. F. (2010). Modelo para el ajuste de pronósticos agregados utilizando lógica difusa. *Ingeniería. Investigación y Tecnología*, 11(3), 289-302.
<https://www.redalyc.org/articulo.oa?id=40415987005>
- FAO. (07 de Octubre de 2021). FAO. (Situación Alimentaria Mundial) Retrieved 08 de Octubre de 2021, from <http://www.fao.org/worldfoodsituation/csdb/es/>
- FAO y CEPAL. (2020). *Sistemas alimentarios y COVID-19 en América Latina y el Caribe*. Santiago de Chile: FAO.
<https://doi.org/https://doi.org/10.4060/ca8677es>
- Forero Gómez, G., & Martínez Lozano, J. A. (2020). *Modelo de regresión lineal múltiple para el pronóstico de ventas de bolsas ecológicas para la empresa Boleco S.A., en la ciudad de Bogota D.C*. Bogotá: Universidad Cooperativa de Colombia. <http://hdl.handle.net/20.500.12494/16439>
- Garcés, D., & Jaimes Barragán, F. (2015). Ronda clínica y epidemiológica. Introducción al análisis multivariable (parte II). *Iatreia*, 28(1), 87-96.
<https://www.redalyc.org/articulo.oa?id=180533008009>
- Limones Lara, J. A. (2012). *Método para el control de la producción defectuosa dentro del proceso de conformado de metales. Caso ZF Sachs Automotive México*. Saltillo: CORPORACIÓN MEXICANA DE INVESTIGACIÓN EN MATERIALES.
<https://comimsa.repositorioinstitucional.mx/jspui/bitstream/1022/96/1/Tesis%20Jonathan%20A%20Limones%20L.pdf>
- López Rodríguez, S. M., & Zapata Zuluaga, T. (2018). *Técnicas de pronósticos en revisión sistémica de literatura para empresas de confecciones*. Bogotá: Universidad Cooperativa de Colombia.
<http://repository.ucc.edu.co/handle/ucc/12151>
- Medina Varela, P. D., Restrepo Correa, J. H., & Cruz Trejos, E. A. (2009). PLAN DE PRODUCCION PARA LA COMPAÑÍA DE HELADOS "NATA". *Scientia Et*

- Technica*, 15(43), 311-315.
<https://www.redalyc.org/articulo.oa?id=84917310055>
- Méndez Giraldu, G. A., & Lopez Santana, E. R. (2014). Metodología para el pronóstico de la demanda en ambientes multiproducto y de alta variabilidad. *Tecnura: Tecnología y Cultura Afirmando el Conocimiento*, 18(40), 89-102.
<https://dialnet.unirioja.es/servlet/articulo?codigo=4778508>
- Montero Granados, R. (2016). Modelos de regresión lineal múltiple. *Universidad de Granada. España*.
- Montesinos-López, O. A., & Hernández-Suárez, C. M. (2007). Modelos matemáticos para enfermedades infecciosas. *Salud Pública de México*, 49(3), 218-226.
<http://www.scielo.org.mx/pdf/spm/v49n3/07.pdf>
- MONTOYA VALER, S. K. (2017). *FACTORES SOCIOECONÓMICOS QUE AFECTAN AL CONSUMO DOMÉSTICO DE AGUA POTABLE USANDO LA REGRESIÓN MÚLTIPLE*. Puno: UNIVERSIDAD NACIONAL DEL ALTIPLANO.
https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwio-JDjwKr1AhXtSjABHRhMAOgQFnoECBoQAQ&url=http%3A%2F%2F repositorio.unap.edu.pe%2Fbitstream%2Fhandle%2FUNAP%2F4967%2F Montoya_Valer_Sharom_Kelly.pdf%3Fsequence%3D1&usg=A
- Morales Tabares , Z. E., Cabrera Campos , A., Vázquez Silva , E., & Caballero Mota , Y. (3 de julio de 2016). MPREDSTOCK : Modelo multivariado de predicción del stock de piezas de repuesto para equipos médicos. *Revista Cubana de Ciencias Informáticas*, 10(3), 143-159 .
<https://www.redalyc.org/pdf/3783/378346436007.pdf>
- Nations, O. a. (2020). *OCDE-FAO Perspectivas Agrícolas*. OCDE-FAO.
<https://doi.org/https://doi.org/10.1787/a0848ac0-es>
- Perez, R. A., Mosquera Sanchez, S. A., & Bravo, J. J. (2012). Aplicación de modelos de pronósticos en productos de consumo masivo. *Biotecnología en el Sector Agropecuario y Agroindustrial: BSAA*, 10(2), 117-125.
<https://dialnet.unirioja.es/servlet/articulo?codigo=6117643>
- Plaza Gálvez, L. F. (2016). Modelación matemática en ingeniería. *IE Revista de investigación educativa de la REDIECH*, 7(13), 47-57.
<https://www.redalyc.org/journal/5216/521655237005/>

- Ramos, A., Sánchez, P., Ferrer, J. M., Barquín, J., & Linares, P. (2010). *Modelos matemáticos de optimización*. Madrid: Universidad Pontificia Comillas.
https://d1wqtxts1xzle7.cloudfront.net/34357606/modelado_en_gams-with-cover-page-v2.pdf?Expires=1633687201&Signature=EzhFiG-skxPtDHclAwclQ08Qgji-w3TqbHmhGJKqaRfyW1SAldIZH5DDJXiHRLC8v2Td b-zncNZYV4WfKu0LVXNikmLiT6Noochcf21rjOhupk0bIXvhHdDbsWtw~Me7oBaSZ0g4ocy
- Render, B., & Heizer, J. (2007). *Administración de la producción*. México: PEARSON EDUCACIÓN.
- Rodríguez Velázquez, J. A., & Steegmann Pascual, C. (2013). *Modelo Matemáticos*.
https://recursos.salonesvirtuales.com/wp-content/uploads/bloques/2012/08/Modelos_matematicos.pdf
- Rodríguez-Gallegos, R., & Bourguet-Diaz, R. E. (2015). Identifying Modeling Practices Through Differential Equations and Simulation. *122nd SEE Annual Conference-Exposition*. <https://www.asee.org/public/conferences/56/papers/13153/view>
- ROJAS VICTORIO, C. E. (2020). *Principales factores de riesgos psicológicos que influyen en los accidentes de trabajo en la minera aurífera Retamas S.A.* Cerro de Pasco: UNIVERSIDAD DANIEL ALCIDES CARRION.
https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjrs6jV06r1AhWIRTABHXFECUw4ChAWegQIAhAB&url=http%3A%2F%2Frepositorio.undac.edu.pe%2Fbitstream%2Fundac%2F1876%2F1%2FT026_04067250_M.pdf&usg=AOvVaw0FNN0rOV5PtyuLP1ytZKW-
- Rojo Abuín, J. M. (2007). Regresión lineal múltiple. *Instituto de Economía y Geografía*, 2-33.
https://d1wqtxts1xzle7.cloudfront.net/36787341/REGRESION_LINEAL_MULTIPLE_3-with-cover-page-v2.PDF?Expires=1640663317&Signature=MH9ICvACS9azM8zOaqrxKMjbZKnQ49AngSm4U1Hz2XG-EePgqOo413-c6mu4HBGfaWrBXsHftzJ9KYmY070cTcz6ReXEL2alQyNbk-wa uciFjWwWzIYf7fwji2iiDJo
- Sánchez Sánchez, D. A. (2018). *Modelo ARIMA para el pronóstico de la producción de cacao en el Perú 2012 - 2018*. Facultad de Ciencias Físicas y Matemáticas. Trujillo: Universidad Nacional de Trujillo.

<https://doi.org/http://190.223.54.254/bitstream/handle/UNITRU/11563/SÁNCHEZ%20SÁNCHEZ%2c%20David%20Alexander.pdf?sequence=1&isAllowed=y>

Sánchez-López, E., Barreras-Serrano, A., Pérez-Linares, C., Figueroa-Saavedra, F., & Olivas-Valdez, J. A. (2013). APLICACIÓN DE UN MODELO ARIMA PARA PRONOSTICAR LA PRODUCCIÓN DE LECHE DE BOVINO EN BAJA CALIFORNIA, MÉXICO. *ropical and Subtropical Agroecosystems*, 16(3), 315-324. <https://www.redalyc.org/articulo.oa?id=93929595004>

Saucedo Castillo, O., Pérez, L. V., Herrera Isla, L., & Fernández Pérez, L. (2010). Sistema de pronóstico climático del tizón tardío (*phytophthora infestans*) en el cultivo de la papa en la provincia de Villa Clara. *REDVET. Revista Electrónica de Veterinaria*, 11(38), 1-12. <https://www.redalyc.org/articulo.oa?id=63613140038>

Tello Cabello, S. Y. (2014). Importancia de la micro, pequeñas y medianas empresas en el desarrollo del país. *LEX*, 12(14), 199-218. <https://doi.org/http://dx.doi.org/10.21503/lex.v12i14.623>

Valdés Díaz de Villegas, J. A., & Sánchez Soto, G. A. (2012). LAS MIPYMES EN EL CONTEXTO MUNDIAL: SUS PARTICULARIDADES EN MÉXICO. *Iberóforum*, VII(14), 126-156. <https://www.redalyc.org/pdf/2110/211026873005.pdf>

Zafra Mejía, C. A., & Gutiérrez Gil, V. H. (2015). Análisis de la producción de lixiviado y biogás bajo condiciones de extracción activa. *Ingenium Revista De La Facultad De ingeniería*, 16(31), 9-23. <https://doi.org/https://doi.org/10.21500/01247492.1365>

ANEXOS

Link de acceso a Data Histórica de la Fábrica de barras de chocolate

https://drive.google.com/drive/u/1/folders/1yDsoUhKGitaTYiHqJr_njknuzEjx6UIC