



UNIVERSIDAD ESTATAL DE MILAGRO

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO

PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO:

DE MAGISTER EN MATEMÁTICAS

TÍTULO DEL PROYECTO

**PROPUESTA DE UN MODELO MATEMÁTICO APLICADO AL
PRONÓSTICO DE PRODUCCIÓN UTILIZADO REGRESIÓN
MULTIVARIADO APLICADO A UNA FÁBRICA DE GALLETA**

COCO CLÁSICO 100 * 25 g.

TUTOR

MSC. KEYLA GUERRERO RUIZ

AUTOR

ECON. HUGO NELSON DONOSO REINOSO

MILAGRO, JUNIO DE 2022

ACEPTACIÓN DEL TUTOR

En calidad de Tutora de Proyecto de Investigación, nombrado por el Comité Académico del Programa de Maestría en Matemática de la Universidad Estatal de Milagro.

CERTIFICO

Que he analizado el Proyecto de Investigación con el tema “PROPUESTA DE UN **MODELO MATEMÁTICO APLICADO AL PRONÓSTICO DE PRODUCCIÓN UTILIZADO REGRESIÓN MULTIVARIADO APLICADO A UNA FÁBRICA DE GALLETA COCO CLÁSICO 100 * 25 g**”, elaborado por el **ECON. HUGO NELSON**

DONOSO REINOSO, el mismo que reúne las condiciones y requisitos previos para ser defendido ante el tribunal examinador, para optar por el título de **MAGÍSTER EN MATEMÁTICA MENCIÓN MODELACIÓN MATEMÁTICA**.

Milagro, 7 de junio de 2022



Firmado electrónicamente por:

**KEYLA STEFANIA
GUERRERO RUIZ**

MSc. Keyla Stefanía Guerrero Ruiz

C.I. 0929607711

DECLARACIÓN DE AUTORÍA DE LA INVESTIGACIÓN

El / la autor/a de esta investigación declara ante el Comité Académico del Programa de Maestría en Matemática de la Universidad Estatal de Milagro, que el trabajo presentado 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 de una institución nacional o extranjera

Milagro, 24 de junio de 2022



Firmado electrónicamente por:

**HUGO NELSON
DONOSO REINOSO**

Econ. Hugo Nelson Donoso Reinoso

C.I. 0919995720

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO
APROBACIÓN DEL TRIBUNAL

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	57.00
DEFENSA ORAL	39.67
PROMEDIO	96.67
EQUIVALENTE	Excelente



Firmado electrónicamente por:
**ARISTIDES
REYES**

Mgtr. REYES BACARDI ARISTIDES
PRESIDENTE/A DEL TRIBUNAL



Firmado electrónicamente por:
**KEYLA STEFANIA
GUERRERO RUIZ**

Megcp GUERRERO RUIZ KEYLA STEFANIA
DIRECTOR/A DE TFM



Firmado electrónicamente por:
EDWIN EVARISTO

Phd. LEON PLUAS EDWIN EVARISTO
SECRETARIO/A DEL TRIBUNAL

DEDICATORIA

A mis padres quienes me han apoyado siempre y en todo instante, siendo un baluarte para culminar el proceso de la maestría; a mi esposa e hijos que son el soporte y el motor para continuar adelante en mi preparación profesional.

Econ. Hugo Nelson Donoso Reinoso

AGRADECIMIENTO

A la Universidad Estatal de Milagro y al cuerpo docente de la Maestría, ya que han sido los facilitadores de nuevos conocimientos muy enriquecedores y que nos proyectan hacia nuevos retos profesionales. También agradezco a mi tutora del proyecto por su guía en el proceso de elaboración del documento.

Econ. Hugo Nelson Donoso Reinoso

CESIÓN DE DERECHOS DE AUTOR

Señor Doctor

Fabrizio 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 UTILIZADO REGRESIÓN MULTIVARIADO APLICADO A UNA FÁBRICA DE GALLETA COCO CLÁSICO 100 * 25 g”** y que corresponde a la Dirección de Investigación y Posgrado.

Milagro, 24 de junio de 2022



Firmado electrónicamente por:
**HUGO NELSON
DONOSO REINOSO**

Econ. Hugo Nelson Donoso Reinoso

C.I. 0919995720

ÍNDICE GENERAL

PORTADA	i
ACEPTACIÓN DEL TUTOR.....	ii
DECLARACIÓN DE AUTORÍA DE LA INVESTIGACIÓN.....	iii
PRESIDENTE/A DEL TRIBUNAL	iv
DEDICATORIA	v
AGRADECIMIENTO	vi
CESIÓN DE DERECHOS DE AUTOR.....	vii
Glosario de Términos.....	xii
RESUMEN.....	xiii
ABSTRACT.....	xiv
INTRODUCCIÓN	1
CAPÍTULO I.....	2
1.1 Planteamiento Del Problema.....	2
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	5
1.6 Justificación e importancia	6
1.6.1 Justificación Teórica.....	6
1.6.2 Justificación metodológica.....	7
1.6.3 Justificación práctica	8
1.7 Hipótesis	8
1.7.1 Hipótesis General.....	8
1.7.2 Hipótesis Particular	8
1.7.3 Conceptualización y operacionalización de variables	9
2.1 Marco Teórico	10
2.1.1 Estado del arte	10
2.2 Fundamentos Teóricos.....	12
2.2.1 Modelo matemático.....	12
2.2.2 Pronóstico de producción	15

2.2.3 Regresión Multivariada.....	19
CAPÍTULO III	27
MARCO METODOLÓGICO	27
3.1 Tipo de Investigación	27
3.2 Diseño de la Investigación	28
3.3 Población y muestra.....	28
3.4 Variables de investigación.....	29
3.5 Método, Técnica y Uso de Software de Tratamiento y Análisis de Datos ..	31
CAPITULO IV.....	32
RESULTADO Y DISCUSIÓN	32
4.1 Formulación del modelo de regresión lineal múltiple	32
4.2 Diagnósis y validación del modelo de regresión lineal múltiple.....	32
4.3 Supuesto de Normalidad de los Residuos.....	33
4.4 Supuesto de Independencia de las observaciones.....	37
4.5 Supuesto de Homocedasticidad	37
4.6.- Supuesto de linealidad.....	39
4.7 Supuesto de no colinealidad	41
4.8 El Modelo Matemático de Regresión Multivariada	42
4.9 Comparación de los datos reales con los datos pronosticados	45
CONCLUSIONES	48
RECOMENDACIONES	49
BIBLIOGRAFÍA.....	50
ANEXOS	56

ÍNDICE DE TABLAS

Tabla 1.....	9
Tabla 2.....	29
Tabla 3.....	29
Tabla 4.....	30
Tabla 5.....	30
Tabla 6.....	34
Tabla 7.....	37
Tabla 8.....	39
Tabla 9.....	42
Tabla 10.....	43
Tabla 11.....	44
Tabla 12.....	45
Tabla 13.....	46
Tabla 14.....	46
Tabla 15.....	56

ÍNDICE DE FIGURAS

Figura 1.....	34
Figura 2.....	35
Figura 3.....	35
Figura 4.....	38
Figura 5.....	40
Figura 6.....	47

Glosario de Términos

Los siguientes conceptos son tomados del libro “Funciones elementales para construir modelos matemáticos” (Bocco, 2010):

- **Modelo matemático:** representa la realidad de un caso trasladado al mundo abstracto de la simbología matemática por medio de funciones o ecuaciones que ejemplifican la dinámica del fenómeno estudiado.
- **Pronóstico:** Expresa la tendencia o información futura de un suceso estudiado a través de la utilización de herramientas estadísticas:
- **Regresión lineal:** Comprende a la técnica que utiliza variables continuas o lineales teniendo en cuenta uno o varios predictores.
- **Regresión lineal múltiple:** Es un modelo que facilita la explicación de variables dependientes frente a variables explicativas; por ello en la misma participan varias variables interrelacionándose.
- **Variable endógena:** El valor está determinado por las relaciones que se establecen dentro del modelo en el que está contenido.
- **Variable exógena:** El valor está determinado por factores externos al modelo en el que está contenido.

RESUMEN

La presente investigación parte de la necesidad que tienen las empresas sobre la predicción de datos vinculados con su accionar económico, necesitando para ello la utilización de modelos matemáticos que permitan de forma eficiente realizar proyecciones, las cuales son necesarias en la toma de decisiones, teniendo en cuenta un margen de error sea muy leve, dando mayor confiabilidad a los resultados obtenidos. El objetivo de la investigación es plantear un modelo matemático que facilite la predicción en la producción de galleta coco presentación clásico 100 * 25 g con la utilización de la regresión multivariada; para ello se analiza cada uno de los factores que inciden en la producción. El tipo de investigación es exploratoria, documental, de campo, aplicada y cuantitativa; con una población y muestra del 100% de la producción, lo que a su vez podrá permitir el estudio de la variable dependiente “cantidad de producción real” y de las variables independientes “Cantidad de producción programada”, “horas laborales planificadas” y “paradas no programadas”; verificando la calidad de ajuste mediante coeficientes de determinación y análisis de la varianza. El modelo matemático es sometido a varias pruebas para la verificación de los supuestos de linealidad, normalidad, homocedasticidad, independencia de los errores y colinealidad; para luego obtener estadísticas de muestras emparejadas que permitieron cotejar el margen de similitud y/o mínimo error entre la producción real y la producción generada por el modelo de regresión multivariado, validándose de forma eficaz la predicción del modelo ya que tiene una relación significativa del 92,4%.

Palabras Claves: Regresión lineal múltiple, Predicción, Modelo matemático Variables

ABSTRACT

This research is based on the need that companies have for the prediction of data related to their economic actions, requiring the use of mathematical models that allow them to efficiently make projections, which are necessary in decision making, taking into account a very slight margin of error, giving greater reliability to the results obtained. The objective of the research is to propose a mathematical model that facilitates the prediction in the production of coconut cookie classic presentation 100 * 25 g with the use of multivariate regression; for this purpose, each of the factors that affect production is analyzed. The type of research is exploratory, documentary, field, applied and quantitative; with a population and sample of 100% of the production, which in turn will allow the study of the dependent variable "quantity of real production" and the independent variables "Quantity of programmed production", "planned working hours" and "unscheduled stops"; verifying the quality of adjustment through determination coefficients and variance analysis. The mathematical model is subjected to several tests for the verification of the assumptions of linearity, normality, homoscedasticity, independence of errors and collinearity; to then obtain paired sample statistics that allowed to compare the margin of similarity and/or minimum error between the actual production and the production generated by the multivariate regression model, effectively validating the prediction of the model since it has a significant relationship of 92.4%.

Keywords: Multiple linear regression Prediction Mathematical
model Variables

INTRODUCCIÓN

Los problemas relacionados con el pronóstico de eventos o sucesos tienen una relevancia en muchos aspectos de la vida cotidiana. En efecto, la utilización de herramientas de modelación matemática en la esfera productiva es un aspecto que va teniendo mayor relevancia frente a un mundo con mercados globales, regionales y locales muy competitivos. De esta forma se buscó articular un modelo matemático que facilite la predicción eficientemente, ágil y cuyos resultados tengan una utilidad pragmática para la empresa y/o empresas en general que se dediquen a esta actividad productiva.

Adicionalmente la investigación ha servido para utilizar herramientas matemáticas aplicadas en un modelo de regresión lineal múltiple estableciendo las variables explicativas y variable explicada con su respectiva medición del nivel de significancia, por lo que fue necesario el uso del valor R cuadrado, obteniéndose el ajuste del modelo con la data registrada.

Los resultados del modelo han servido para pronosticar la producción, analizando la incidencia de cada una de las variables que lo componen, por lo que es útil para la toma de decisiones empresariales y a su vez siendo un medio para mejorar los niveles de producción y productividad al optimizar o combinar adecuadamente las variables que componen la regresión. Por ejemplo, en el modelo matemático la variable “paradas no programadas” se puede modificar para variar el rendimiento del modelo.

De esta forma, el problema de la investigación se basó sobre cómo la regresión multivariada contribuye a un pronóstico adecuado en la producción de galleta de coco presentación clásico 100*25g, analizando para ello las relaciones entre variables. El estudio se realizó en la ciudad de Guayaquil en una fábrica dedicada a la fabricación de galletas, por lo que se han empleado una data histórica de producción de los últimos dos años. Por lo que las preguntas de investigación partieron de criterio de factibilidad del modelo, y la capacidad de predicción del mismo.

CAPÍTULO I

1.1 Planteamiento Del Problema

Valdéz Díaz de Villegas y Sánchez Soto (2012) indican que “en el mundo económico y empresarial se busca la mejora constante de la productividad, eficiencia y superiores niveles de optimización a través de la aplicación de algoritmos y herramientas matemáticas” (p. 47). Por ello se requiere una adecuada selección de las variables y una precisa elección de los parámetros a medir, en función del tipo de empresa. Render y Heizer (2007) “se tendría en consideración los siguientes aspectos: la proyección de la demanda de los bienes y servicios, administración operativa de inventarios, eficiencia en la asignación de recursos dentro del proceso productivo, operatividad adecuada y óptima tanto en la distribución, proyección de producción y eficiencia del uso de la materia prima y demás materiales” (p. 87).

Para Tello Cabello (2014) “el fortalecimiento de la economía y del mercado de trabajo, las pequeñas y medianas empresas (PYMES) son consideradas como elementos fundamentales en la generación de ingresos” (p. 12). Siendo un elemento central en su competitividad en el mercado es la capacidad de predecir la producción en función de la demanda (López Rodríguez & Zapata Zuluaga, 2018). Las empresas tienen algunos errores cuando tratan de predecir la producción, Méndez Giraldu y Lopez Santana (2014) consideran que “aplicar unilateralmente una metodología de predecir la demanda, no relacionar eficientemente el consumo histórico con un estudio científico de mercado, no tener un cálculo adecuado en la estimación del error, no administrar la métrica de error, considerar una base de datos no significativa para la investigación y no tener en cuenta que la demanda puede presentar diversos comportamientos como es el nivel de elasticidad” (p. 91).

Se puede considerar muchas definiciones sobre las ciencias matemáticas y estadística, FAO y CEPAL (2020) lo plantea como “el pronóstico se considera todo una ciencia que se aplica a procesos de simulación de eventos y hechos futuros” (p. 61). La construcción de los modelos implican ciertas condiciones para su funcionalidad, así el Banco Central del Ecuador (2017) sostiene que “para lo

cual la recolección de datos y series históricas son la base para crear los algoritmos y modelos matemáticos que faciliten resultados cuantitativos esperados” (p. 107).

Instituciones como la Superintendencia de Compañías, Valores y Seguros (2018) considera “en todo método de pronóstico se debe tener en cuenta una amplitud en cuanto a los tiempos, lo cual está en función del tipo de necesidades del entorno empresarial, existiendo predicciones de carácter corto, mediano y largo plazo que a su vez se emparejan en visiones a nivel estratégico y táctico” (p. 11).

Damián Llatas y Sandoval Santamaría (2018) afirman “la economía, la sociedad y los problemas considerados importantes pueden resolverse a través de la aplicación modelos matemáticos. Modelar es el proceso de edificar una formulación matemática que resuma e interprete eficientemente la realidad o el contexto dado en un momento concreto” (p. 12). Crear un modelo matemático es una tarea compleja, que requiere un nivel de abstracción simbólica de la realidad estudiada, pero a su vez se debe de considerar en el análisis y estructura una perspectiva multilateral, por lo que es común que diversos profesionales con su experticia aporten en su creación. Ramos et al (2018) “un modelo es una herramienta que facilita la toma de decisiones en una serie de aspectos de la actividad económica, ya sea a nivel logístico, productivo, mercadeo, entre otros aspectos” (p. 21).

Sánchez Sánchez (2018) puntualiza que “el principal problema que enfrenta una empresa al pronosticar la demanda es la baja eficacia que se puede tener de la proyección de datos” (p. 6). Un modelo puede tener problemas de precisión, es una relación inversa ya que mientras mayor sea el error menos eficacia mostrará y viceversa. Para Zafra Mejía y Gutiérrez Gil (2015) “la eficacia se la puede medir de acuerdo al error que tiene el pronóstico, es decir, que tan cerca está el pronóstico del nivel de demanda real” (p. 43).

Cabrera González y De León Arias (2015) indica que “resulta importante para las empresas nacionales poder predecir el comportamiento del mercado de forma ágil, oportuna y con el menor error posible; por lo tanto, al pronosticar se

debe de tener en cuenta algunos factores que de una u otra manera pueden influir en el proceso de producción como es: el histórico de la producción, el tiempo de paradas programadas de las máquinas, el tiempo de paradas no programadas de las máquinas, las horas hombres reales que se trabaja, etc.” (p. 47).

En tal sentido Araya Pizarro y Rojas Escobar (2020) afirma que “se tiene como una alternativa es el uso de los modelos multivariados que pueden ser utilizados en el pronóstico de producción, su ventaja radica en poder encontrar relaciones entre las distintas 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” (p. 12).

Frente a lo cual se formula la pregunta, ¿se puede desarrollar un modelo matemático que favorezca el pronóstico la producción de una fábrica de galletas de coco presentación clásico 100 * 25 g utilizado regresión multivariada?

1.2 Delimitación del problema

Espacio

La investigación se realizó en la república del Ecuador en la Región Litoral, correspondiente a la Provincia del Guayas, Ciudad de Guayaquil, en una pequeña empresa.

Tiempo

Por las características de la información, la vigencia de la presente investigación está comprendida dentro del intervalo de 10 años.

Universo

La investigación se fundamenta en el desarrollo de un Modelo Matemático para la producción, para lo cual se usarán los datos históricos de producción de los últimos 2 años.

1.3 Formulación del problema

¿Cómo la utilización de la regresión multivariada contribuirá en la implementación de un modelo matemático que permita pronosticar eficientemente y de forma ágil la producción en una fábrica de galleta de coco presentación clásico 100 * 25 g?

1.4 Sistematización del problema

¿A través del estudio de los elementos de un modelo matemático se logrará determinar los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g?

¿Será factible aplicar un examen de la variabilidad correspondiente a cada uno de los factores que inciden en la producción de galleta de coco presentación clásico 100 * 25 g?

¿Será posible desarrollar un modelo matemático adecuado que facilite pronosticar de forma rápida la producción aplicándolo a una fábrica de galleta de coco presentación clásico 100 * 25 g utilizando regresión multivariada?

1.5 Objetivos

1.5.1 Objetivo General

Plantear un modelo matemático ajustado que facilite el pronóstico de manera ágil en la producción, en una fábrica de galleta de coco presentación clásico 100 * 25 g aplicando regresión multivariada.

1.5.2 Objetivos Específicos

- Determinar los factores que inciden en la producción de galleta de coco presentación clásico 100 * 25 g.
- Analizar la variabilidad de cada uno de los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g.

- Construir un modelo matemático ajustado que facilite predecir de forma eficiente la producción aplicándolo a una fábrica de galleta de coco presentación clásico 100 * 25 g utilizado regresión multivariada.

1.6 Justificación e importancia

La FAO (2021) explica que “las empresas continuamente se enfrentan a distintos problemas relacionados con la precisión y exactitud de los modelos matemáticos que facilitan a observación y comportamiento de los distintos factores que intervienen en el proceso de producción, así como en la exactitud de la información que es dada por los modelos matemáticos empleados” (p. 63). Las Nations (2020) saben “lo crucial que es para las empresas e instituciones con poder de decisión de política económica es saber medir el efecto de sus decisiones, por lo que poder simular a través de variables escogidas de forma metódica, hacen que los modelos matemático tengan una importancia cada vez más actual” (p. 25).

El error dado por los modelos empleados se valora en función de la comparación y cotejo del ámbito real y lo proporcionado por el modelo matemático, el error es una estimación que siempre se debe de considerar y que estadísticamente está justificado debido a la perspectiva compleja de la propia realidad, presentándose dificultades al momento de generalizar y abstraer matemáticamente la situación estudiada; Garcés y Jaimes Barragán (2017) “consideran plausible que el error estimado sea lo menor posible, incrementando por contrapartida el nivel de confianza y de acierto del modelo. El cálculo y análisis del error es parte consustancial del modelo” (p. 32).

Por lo que, el presente trabajo investigativo busca aportar con un modelo matemático que permita a la empresa poder pronosticar de forma más precisa, reduciendo el error, y con ellos que la empresa mejore su posición en el mercado al reducir sus pérdidas.

1.6.1 Justificación Teórica

El Banco Central del Ecuador (2017) considera “los constantes cambios en la preferencia de los consumidores, y los problemas actuales que hacen que

las empresas tengan una gran necesidad de reducción en todos sus costos; es muy importante precisar y comprender cómo se puede elaborar por medio del lenguaje matemático un modelo eficiente destinado al pronóstico de la producción” (p. 82). Hay que tener en cuenta que los distintos factores que afectan a una fábrica son importantes comprenderlos para el uso de la regresión multivariada, para ello es necesario asimilar los fundamentos teóricos matemáticos que convaliden dicho modelo.

Aspectos como la competitividad empresarial puede también medirse por medio de regresión lineal múltiple. Para ello Vicuña Pino, Cortez Chichande, & Basurto Tovar (2019) “utilizaron el método Best Subset Selection (BSS) tiene como objetivo la formación y selección de indicadores microeconómicos relacionados con la competitividad para la conformación de un modelo eficiente en forma de conjuntos, utilizando como indicadores: índices financieros, mercadotecnia, productividad, recursos humanos y tecnología” (pág. 96).

1.6.2 Justificación metodológica

Desde la perspectiva metodológica, la investigación se enfoca en el análisis de los factores (variables) del modelo que afectan el pronóstico de producción como son: “horas reales trabajadas (Hr)”, “horas de paras no programadas (Hr)”, “horas de paras programadas (Hr)”, “horas hombres trabajadas (HH)”, “retrabajo (kg)”. Por lo que se plantea para cada variable considerar el nivel de aporte al modelo, estudiar su incidencia, garantizando una adecuada predicción.

La utilización de un modelo de regresión multivariada implica la medición del nivel de significancia de las variables independientes del modelo, considerando para dicho propósito el uso del valor R cuadrado, ya que mide el ajuste a través de los datos ingresados.

Para Ortega & Pinto (2021) “la utilización de datos para la construcción de un modelo lineal múltiple para predecir el bienestar en los niños en cuanto al uso de pantallas inteligentes, aplica el método de los mínimos cuadrados para lograr la recta de regresión lineal que minimiza los residuos, donde la variable dependiente de estudio es la salud agregada de un individuo y a su vez se valora

las capacidades de homocedasticidad, multicolinealidad y normalidad de los datos. El análisis de varianza ha sido fundamental para la validación y verificación de la eficiencia del modelo en cuanto a los datos seleccionados” (p. 123).

1.6.3 Justificación práctica

La investigación tiene utilidad para la construcción de un modelo matemático predictivo, favoreciendo la aplicación práctica del análisis del ajuste del modelo como también de la propia verificación de los supuestos que son necesarios validarlos. La bondad de ajuste del modelo está valorado con una R cuadrado ajustado de 0,920, por lo que se procedió a verificar los supuestos del modelo. Por su parte el valor estadístico de prueba F considera una puntuación de 218,862 por lo que se justifica la cada variable independiente por su dinámica explicativa. Por otra parte las diferencias emparejadas entre la cantidad de producción real con la cantidad de producción de la regresión multivariada es mínima siendo apenas del -0.185, por lo que se considera válido la aplicación del modelo predictivo.

El procesamiento práctico de datos ha sido útil ya que el modelo refleje la situación de producción de la empresa, pudiendo ser replicable a otros productos que la compañía oferte, teniendo.

1.7 Hipótesis

1.7.1 Hipótesis General

El diseño de un modelo matemático a través regresión multivariada, permitirá pronosticar de forma óptima la producción de galletas de coco presentación clásico 100 * 25 g.

1.7.2 Hipótesis Particular

Es factible reconocer los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g.

Es factible diferenciar a los factores que inciden en la variabilidad de un modelo de proyección de producción de galleta de coco presentación clásico 100 * 25 g.

Es factible desarrollar un modelo matemático adecuado que facilite el pronóstico de forma rápida de la producción aplicándolo a una fábrica de galleta de coco presentación clásico 100 * 25 g a través de regresión multivariada.

1.7.3 Conceptualización y operacionalización de variables

Variable independiente: regresión multivariada

Variable dependiente: pronóstico de producción

Tabla 1

Operacionalización de variable

Variable	Definición conceptual	Indicadores	Técnica
Regresión Multivariada	Es la relación que hay entre variables, donde una variable explicada o dependiente es influida por variables de carácter independientes; donde matemáticamente es una formulación de relación lineal.	Valor de los coeficientes de la regresión multivariada.	Prueba de ajuste
Pronóstico de Producción	El pronóstico es la cantidad de kg de galletas de coco presentación clásico 100 * 25 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

Nota: La tabla muestra la operacionalización de la regresión multivariada y del pronóstico de producción. Elaborado por el autor.

CAPÍTULO II

2.1 Marco Teórico

2.1.1 Estado del arte

Para los antecedentes del presente proyecto de investigación, se recopiló y revisó trabajos de tesis desarrolladas por diferentes autores los cuales contienen información importante además comparten similitud con el tema estudiado.

Meza y Chafloque (2013) en su trabajo de investigación “Pronóstico para la producción de caña y azúcar terminada con un modelo de regresión lineal múltiple de dos etapas orienta la elaboración de un modelo para la predicción para el pronóstico de la producción de caña y azúcar determinada, cuya información fue tomada de diversos cultivos pertenecientes a la empresa Agroindustrial Casa Grande S.A. El análisis estadístico realizado en la investigación fue el análisis regresión lineal múltiple multivariante con 3 variables dependientes donde se concluyó que las variables asignadas en el proyecto inciden en los niveles de sacarosa, reductores y toneladas de caña por hectárea las cuales fueron la edad, nitrógeno y la estación por temporada” (p. 81).

Gómez (2017) en su proyecto de investigación denominado “Factores que determinan la rentabilidad del cultivo de cacao en el distrito de Pólvora determina rentabilidad del cultivo de Cacao en el distrito de Pólvora, la metodología de investigación realizada es científica, fáctica y aplicada. La población de estudio fue conformada por 160 productores cacaoteros y la muestra como la unidad de análisis es de 74 productores. Como técnica empleadas fueron el procesamiento de bibliografía (bibliometría), la aplicación de un cuestionario y la inferencia estadística, en la cual mientras la variable costos, rendimiento, precios explican de forma cuantitativa el comportamiento, siendo la estructura de un modelo económico – matemático cuya técnica es la regresión lineal de tipo múltiple, en el cual la utilidad de la cosecha de cacao en el distrito de Pólvora estriba de las variables costos de producción, rendimiento agrícola y precio de venta al público. Concluyendo así que los principales factores que determinan la rentabilidad del

Cacao en el distrito, son los Costos (C), Rendimiento (R), y los Precios (P)” (p. 94).

Cruz (2017) en su trabajo de investigación “Aplicación de la regresión múltiple para predecir las variables que influyen en el peso del racimo en el cultivo de uva RED GLOBE establece la vinculación entre las variables de estudio para la cosecha del cultivo, considerando el peso por racimo de la uva del tipo Red Globe, por lo que en la producción de 2014 de la empresa agroexportadora de uva CAMPOSOL S.A. se empleó la técnica de análisis de regresión múltiple considerando como variable dependiente el peso del racimo y como variable independiente el número de baya, peso de baya, longitud del racimo y grados Brix. Se utilizó el criterio de Schwarz evaluando dos modelos uno con dos variables independientes y el otro con tres variables independientes concluyendo que el peso del racimo queda explicado en un 91,4% por las variables en número de baya y peso de baya” (p. 79).

López (2013) en su proyecto de investigación “Análisis de regresión para la estimación del secuestro de carbono orgánico en suelos presenta evidencias sistematizadas al poner en práctica un modelo de regresión multivariada y analizando sus supuestos, de esta forma se revisan para el caso planteado dos modelos de características similares, los cuales pueden ser usados de forma práctica para resolver este tipo de problemas: Regresión con Componentes Principales (RCP) y Regresión por Mínimos Cuadrados Parciales” (p. 83).

López Pineda (2016) manifiesta “dos métodos tienen en común, que transforman las variables predictoras en componentes ortogonales, expresando las mismas la respuesta al problema de multicolinealidad y facilitan una disminución de la dimensionalidad del espacio de variables independientes” (pág. 23). López Pineda (2016) analiza el objetivo de comparar diferentes métodos de regresión para decidir cuál será el que responda mejor el porcentaje de COS en consonancia a otros aspectos que inciden en la calidad del suelo en la región de la Caldera de Teziutlán en el estado de Puebla” (pág. 31).

Tulcanaza (2015) en su trabajo “Desarrollo de un modelo de caracterización financiera del sector comercio en Ecuador, periodo 2022-2012

plantea que la actividad económica comercial del Ecuador tiene un comportamiento y a su vez lo mide en las finanzas, pudiendo sintetizarse por medio de un modelo matemático de regresión lineal múltiple. Frente a lo cual se empleó un modelo estadístico fundamentado en los mínimos cuadrados ordinarios. La información utilizada como data está en función del formulario 101 impuesto a la renta de sociedades y establecimientos permanentes del organismo de control tributario de Ecuador” (p. 46).

2.2 Fundamentos Teóricos

2.2.1 Modelo matemático.

Las matemáticas facilitan numerosos instrumentos que pueden ser adaptados a diversas situaciones, transformando el problema a resolver a un lenguaje que lo valora de forma cuantitativa, es decir matemáticamente utilizando representaciones y simbologías que sintetizan de esta forma la respuesta a través de una solución con el modelo a aplicar. Hay diversos modelos a utilizar, ya sea en el campo de las ciencias y de la física, como también en el ámbito económico y de producción. Infante Roblejo et al. (2019) afirma que “la modelación requiere iterar a través de algoritmos, lo cual tiene una utilidad en diversos escenarios, como por ejemplo en la teoría de juegos. Sin embargo todos estos modelos en la actualidad vienen acompañado en alguna medida de procesos de programación matemática, facilitando la automatización, veracidad y rapidez de los resultados” (p. 82).

Por ello la modelización matemática es una herramienta para campos como la administración, la logística, la planeación, predicción, economía y en general toda ciencia que requiera cuantificar de forma científica, ordenada y sistematizada aspectos esenciales del conocimiento. El lenguaje matemático es una abstracción de una problemática dada y como tal, su elaboración facilita la comprensión en mejores niveles de las variables en estudio.

La modelación matemática, según Yeste & Barat “busca construir un modelo que simplifique matemáticamente la realidad, a través de parámetros y de la dinámica propia de la técnica matemática escogida, o a su vez de ecuaciones que de sentido a una relación de la realidad concreta estudiada. El

modelo empieza con el reconocimiento de los aspectos centrales de una realidad específica estudiada, los cataloga en variables, y los relaciona a través de expresiones” (p. 151). Se considera que el objeto de estudio realmente no se expresa matemáticamente, por lo que se lo transforma gracias al lenguaje matemático que utiliza simbolizaciones, por lo que para dicho proceso va logrando estructurar un modelo que explique de forma consistente y cuantitativa la dinámica del objeto que se ha estudiado, denominándose este proceso análisis del modelo. Los resultados que se obtuvieron corresponden a la interpretación del análisis matemático dado, verificando y comprobando la hipótesis planteada.

Návar et al. (2017) denominan como modelo matemático “al conjunto de símbolos y relaciones matemáticas que traducen, de alguna manera, un fenómeno en cuestión o problema de situación real” (p. 93). Resolver un problema real por lo regular es muy complejo. En gran medida ello se explica por la complicación del modelo a ejecutar, teniendo la necesidad de registrar cada una de las variables, las cuales pueden ser fácilmente explícitas para considerarlas dentro de la modelación, como también puede que su inclusión requiera de mayor conjetura al momento de considerarla. Morales y Becerra (2018) consideran que “el arte de la abstracción está vinculado con el proceso matemático y por ende con la construcción de modelos, por lo que se requiere comprensión de las variables u observaciones, parámetros y estimaciones” (p. 38). Requiriendo aplicar modelos para verificar que la formulación matemática escogida sea la correcta para explicar a su vez el fenómeno validado por el modelo construido.

Para afrontar de forma constructiva una problemática desde una perspectiva matemática que coadyuve a la generalización de un modelo matemático, siendo necesario observar adecuadamente el hecho o situación relacionada al caso de estudio, valorar cada una de las variables, establecer un método de pertinencia o de utilidad para la construcción del modelo, desechando a aquellos aspectos no relevantes. González et al. Establecen que “al modelo símbolos que representa las operaciones de manera abstracta tiene que tener un nivel de funcionalidad para que el mismo pueda ser validado” (p. 189). De esta

manera juega un papel muy importante la claridad del método matemático a aplicar dentro del contexto del modelo matemático implicado, es decir la plena comprensión del algoritmo, su validación para el caso estudiado, su correcta relación de las variables. Hay que tener en cuenta que para que un modelo tenga el éxito esperado el mismo tiene que expresar cuantitativamente la realidad estudiada, y adicionalmente acompañado de la prolijidad matemática de estudio del modelo.

La mayoría de los problemas en cualquiera de los campos de investigación, requiere que se utilice en mayor o menor medida la abstracción, cuantificando variables y evaluación de impacto. Para García et al. (2019) valoran “necesario estudiar el tipo de relación que las variables tengan entre sí, determinando quienes están en función de otras, el planteamiento o fórmula matemática a desarrollar, el método más “limpio” o con mejor ajuste para la resolución del problema presentado. Tanto el método matemático como las variables a considerar están también función del campo y disciplina de estudio que se presente” (p. 66). La modelación matemática puede adoptar diversos métodos matemáticos, ya sea programación lineal, ecuaciones diferenciales, regresión lineal entre otros; siendo importante que el método seleccionado sea adaptable o explique de mejor forma el objeto de estudio. Las soluciones numéricas, iteraciones y demás formas de cálculo tienen una orientación dependiendo del objetivo de la investigación y por ende de la intencionalidad del modelo, ya sea para maximizar, minimizar, predecir, entre otros aspectos de medición.

En aquellas naciones que se caracterizan por utilizar herramientas de últimas tecnologías requieren de un nivel de implementación matemática importante. Cárdenas Pérez & Benavides Echeverría (2021) afirman que “las formulaciones matemáticas se unifican con los componentes tecnológicos para mejorar aspectos como el control de producción, densidad del tránsito en una ciudad poblada, medición de los cultivos, estimación de la salud de pacientes, entre otros usos en diversas ramas de la producción y del sector servicios”. (p. 62). Vale recordar que actualmente vivimos en una sociedad caracterizada por el conocimiento, la innovación, el desarrollo tecnológico, digital y virtual; donde

las aplicaciones matemáticas son fundamentales, ya que gestionan los algoritmos necesarios para el funcionamiento del mundo digital, el internet de las cosas y campos nuevos como la minería de datos. Los modelos matemáticos han facilitado procesos de automatización, el desarrollo de la informática y virtualidad, comercio electrónico, creación de plataformas, y demás aspectos vinculados con la actual revolución científico técnica. La toma de decisiones en los diversos ámbitos de la sociedad es eficiente si se utilizan formulaciones matemáticas y ellas se ajustan a resolver problemas que requieren rapidez, precisión y eficiencia.

Gutiérrez y Ferreira (2020) expresa que “la matemática es una herramienta cuantitativa para el desarrollo de modelos que faciliten el análisis de un problema u objeto. De esta forma un modelo está estructurado por variables que se tienen que cuantificar a través de la estilización de procesos y ello se refleja en la propia herramienta matemática, ya sean ecuaciones, regresión, y demás variantes matemáticas, siendo su propósito variado, ya sea para estimar, pronosticar, medir la relación e impacto de variables, entre otros aspectos que la investigación requiera” (p. 50). La relevancia de la ciencia matemática se pone de manifiesto con el desarrollo de profesionales relacionados con la misma y la promoción del estudio de carreras STEAM.

2.2.2 Pronóstico de producción

Rubio Guerrero (2017) “el saber pronosticar un escenario es uno de los aspectos que requiere adecuados modelos matemáticos con variables pertinentes que facilitan posibles escenarios cuantitativos e incluso, necesarios para que las decisiones a implementar sean las más acertadas posibles, siendo fundamental al momento de establecer la dimensión del mercado, la demanda efectiva, los suministros necesarios, el volumen de producción y la proyección tanto de los ingresos, gastos y utilidades (p. 27). En el mundo de los negocios, los inversionistas realizan este tipo de análisis para saber si es lucrativo invertir en algún segmento, producto o servicio.

El administrador o el departamento de planificación de cualquier ente tienen que establecer parámetros y proyecciones realizándolas de forma científica y en función de las condiciones reales. Rubio Guerrero (2017) sostiene que “una planificación se la realiza en cuanto a los implementos, suministros y producción a realizar dentro de una escala de tiempo establecida, por lo que un modelo matemático que favorezca dichos cálculos resulta ser fundamental (p. 29).

Rubio Guerrero (2017) “el pronóstico en el ambiente empresarial es necesario para una programación óptima de la actividad económica, desarrollando una investigación operativa de optimización de procesos para asumir menos costos pero a su vez facilitando una maximización de beneficios, aprovechando al máximo no sólo los recursos materiales y de materia prima si fuera el caso, sino también el propio recurso humano” (p. 31).

Los pronósticos se utilizan en los entes para resolver si hay suficiente cantidad de consumidores para ingresar al mercado ya sea de bienes o servicios, valorar los recursos con que se cuenta para diseñar instalaciones, detectar fluctuaciones de mediano y corto plazo en las necesidades con el fin de mantener la programación de producción, programación de mano de obra, planificación de materiales y otras necesidades (Rubio Guerrero, 2017).

Hernández et al. (2019) puntualiza que “las empresas buscan mejorar los procesos en todas las áreas de su estructura organizacional, utilizando de alguna manera definiciones y métodos vinculados a la administración de procesos y logística. Por lo que, ahora se ubica la relevancia de dominar una forma integrada, con especial énfasis en los enlaces de suministro (p. 57). Al cambiarse a condiciones administrativas y productivas de mejora, se requiere una gestión adecuada de los puntos críticos como la valoración de la cantidad demandada, la estimación de suministros e inventarios necesarios para la producción y/o comercialización, la planificación de los materiales para el área de producción, la planificación de la distribución y la producción desde una perspectiva de las necesidades del mercado.

Los pronósticos se definen como las herramientas de carácter objetivo y científico de predecir sucesos, requiriendo la asunción de información y data previa (histórica) para de esta manera estructurar un modelo matemático que previo al análisis de sucesos e información ocurrida de forma previa se podría obtener información post, por lo que se buscará siempre evitar aspectos subjetivos al momento de formular el desarrollo matemático validando de forma oportuna las variables, las cuales a su vez son las evidencias cuantitativas del desarrollo del objeto estudiado en cuanto a su comportamiento, lo que traducido al ámbito administrativo se reflejará en la optimización de decisiones. Arellano y Peña (2020) consideran que “rara vez un método de estimación matemático es superior de forma absoluta a otro, considerando que funcione para un negocio bajo una amplia gama de condiciones puede resultar un desastre total en otro, o incluso dentro de un conjunto diferente de circunstancias, otra división de la misma empresa” (p. 31).

Actualmente, el pronosticar se considera como un elemento clave y estratégico que ayuda a la gestión de proyectos de eventos futuros en función del comportamiento histórico. A nivel operativo, las estimaciones a buen plazo hacen que la planificación y la producción sean más eficientes, asegurando la disponibilidad de material, optimizando el inventario y aumentando la satisfacción del cliente. Por el contrario, Arroyo García, et al. (2017) reflexiona que “la imprecisión del pronóstico afecta la escasez o excesos de material, la compra innecesaria de suministros, los elevados egresos de inventario debido a la disminuido movimiento de productos, la pérdida de ingresos por escenarios de reducción de ventas, clientes que se han alejado de nuestros servicios, entre otros aspectos que se vinculan a la parte medular de cualquier empresa” (p. 55).

Los componentes que inciden en la demanda se modifican constantemente; tenido presente que la variable unidad de tiempo es necesario considerarla, porque la precisión del pronóstico puede disminuir. Por esta razón, los pronósticos de demanda deben renovarse en sus estimaciones de una forma periódica para seguir teniendo su validez y certeza. Posteriormente a cada ciclo productivo en una empresa, se deben revisar los pronósticos. Vicuña Pino et al. consideran que “la construcción de un sistema de pronóstico requiere al menos,

algunas informaciones estadísticas como punto de partida, estructurar una metodología de pronóstico y los procesos de examen y seguimiento necesarios para actualizar el modelo. En el proceso investigativo y de desarrollo de las ciencias exactas, ya hay formuladas varias técnicas destinadas a la medición y el desarrollo de pronóstico, que pueden clasificarse esencialmente en dos vertientes: métodos de carácter cuantitativos y métodos de carácter cualitativos” (p. 102).

En condiciones de falta de datos cuantitativos, cobra importante relevancia la aplicación de técnicas de investigación cualitativa, que ayudan a obtener apreciaciones desde una perspectiva basada en las características y causas fenomenológicas. Al disponerse de una data o cifras históricas en base a un record de información sistematizado, lo ideal es la utilización de técnicas de investigación cuantitativas para el desarrollo de tendencias, medición de variables y sus interrelaciones. Los métodos de pronóstico pueden ser causales, ya que incorporan factores que pueden afectar la cantidad de pronóstico, como la temperatura, la humedad, la estación, etcétera y cuyas necesidades pueden presentar tendencias, estacionalidad y comportamiento cíclico o aleatorio (Arroyo García et al. p. 50).

La creciente demanda de una amplia gama de bienes y servicios condiciona a los fabricantes a implementar directrices y planes objetivos, pero también que sean adaptable a escenarios y nuevas condiciones, ello en función de obtener la máxima ganancia con la disponibilidad de material, utilizar mano de obra y cumplir con los plazos. La percepción, el discernimiento y la indagación disponible son requisitos fundamentales a la hora de intentar establecer previsiones de ventas para minimizar la variabilidad. Moreno Castro (2019) analiza que al “evaluar la proyección de las ventas es una tarea nada fácil que se simplifica mediante el uso de información histórica del mercado, movimientos de inventario o facturas de ventas como fuente de datos. Las técnicas de pronóstico generalmente aceptadas se clasifican en 5 categorías: discernimiento

de la gerencia, encuestas, análisis multivariado en categorías de tiempo, regresión lineal y pruebas de mercado” (p. 32).

El escogimiento del método matemático implica de varios aspectos como: estimación de costos, la finalidad de la proyección, nivel de confianza, una data realmente objetiva que facilite dichas proyecciones, tiempo, el tipo de oferta, peculiaridades del mercado, disponibilidad de la información necesaria y la experiencia de los responsables de realizar las previsiones. Moreno Castro (2019) sostiene: “las empresas deben combinar diferentes técnicas de pronóstico y planificación de la producción para minimizar los costos y utilizar nuevas tecnologías con menores requisitos de capital, lo que permite a la organización maximizar las ganancias” (p. 43).

2.2.3 Regresión Multivariada

La regresión lineal múltiple es una formulación matemática – estadística que permite el estudio de variables, su relación y más que todo, la predicción del comportamiento medido en términos cuantitativos, logrando inferir futuros sucesos o situaciones por medio del modelo estructurado. Vilà Baños et al. consideran “a la función lineal una de las opciones sencillas al momento de aplicarla, agregando cada variable de acuerdo al sentido de investigación que se esté desarrollando, es decir por la cantidad de aspectos a medir” (p. 7).

Existen muchas técnicas de regresión en función del tipo de variable y el tipo de función que se realiza entre ellas. Montero Granados (2017) expone que “la regresión lineal asume que la relación entre dos variables es lineal o, linealizada al transformar las variables” (p. 11).

La regresión lineal puede ser de tipo "simple" en la que participan dos variables, pero esto generalmente no es suficiente para entender problemáticas más complejas en los que indican más de dos variables, este tipo de regresión es conocido como "múltiple". En el tipo de regresión lineal múltiple, asumimos que a partir de dos variables afectan o se correlacionan con el resultado de la tercera variable. En la regresión lineal múltiple, Montero Granados (2017) indica que “generalmente hay solo una variable endógena y posiblemente varias

variables exógenas. En otras palabras, el fenómeno observable está individualizado. También es posible que existan algunas variables endógenas, pero resolverlas es difícil por lo que este no es un caso general” (p. 10).

Para Vilá Baños et al. (2019) desarrollar un modelo de regresión lineal, se deben cumplir las siguientes condiciones:

- Linealidad: Que todas las variables sean lineales, es decir no haya ninguna que se pueda expresar cuadráticamente, etc.
- Independencia: El error de medición de las variables explicativas es independiente entre sí.
- Homocedasticidad: Error con varianza constante.
- Normalidad: Las variables siguen una distribución normal.
- No colinealidad: Las variables independientes no están correlacionadas entre sí (pág. 5).

Cuando se trabaja en un modelo multivariante de regresión lineal, contamos con que la variable de carácter dependiente responda a una función lineal compuesta por k elementos que responden a variables de carácter independiente y adicionalmente esté enmarcado dentro de un parámetro de error. También el modelo contempla la inclusión de la variable independiente, por lo que se asigna a y como la variable dependiente, y a $x_1, x_2, x_3, x_4, \dots, x_k$ como variables independientes, es decir que inciden y explican el comportamiento de las dependientes; a ε al margen de error. De esta forma el modelo de regresión lineal (o multivariado) se sintetiza en 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 termino 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

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

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 simultaneas b_j de los coeficientes del modelo poblacional β_j utilizando el método de mínimos cuadrados (Correa Morales & Barrera Causil, 2019, pág. 160).

Un aspecto crucial al elaborar un modelo de regresión lineal múltiple, es determinar del conglomerado de variables independientes: $x_1, x_2, x_3, x_4, \dots, x_k$,

cuáles son las pertinentes que inciden en el comportamiento de la variable dependiente Y c.

Supuestos habituales de la regresión multivariada

La estructura del modelo de regresión múltiple, considerando los parámetros y variables, se conforma de esta manera:

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

Adicionalmente en el modelo considera un n conjunto de observaciones. Como premisas que generalmente se consideran al aplicar la regresión lineal múltiple, tenemos:

1. Las x_{ji} son bien números fijo, 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:

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

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 (Quevedo Urias & Pérez Salvador, 2014, pág. 399).

Método de mínimos cuadrados

El método de mínimos cuadrados para la regresión multivariada calcula los coeficientes estimados para reducir la suma de los cuadrados de los residuos. Recuérdese que el residuo es

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

donde y_i es el valor observado de Y e \hat{y}_i es el valor de Y predicho a partir de la regresión (Escalante Cortina, Mayorga Mogollón, & Vergara Schmalbach, 2012, pág. 72).

Ello se expresa de forma minimizada con SCE :

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

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

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

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

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

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

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

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

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

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

Ximénez et al (2013) indica que “los coeficientes son elegidos de manera que la suma de cuadrados entre los valores observados y los pronosticados sea mínima, en otras palabras, que se va a reducir la varianza residual (p. 286)”.

Coefficiente de determinación ajustado.

Ciertos especialistas proponen que, al tratar con modelos de regresión multivariado, se debe usar **r^2 ajustada**, para reflejar el número de variables independientes en el modelo y el tamaño de la muestra. Es muy importante reportar la **r^2 ajustada** al equiparar dos o más modelos de regresión que pronostican la misma variable dependiente, sin embargo, cuenta con diversos números de variables independientes.

r^2 ajustada

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

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

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

Empleamos esta medida para tener en cuenta el hecho de que las variables independientes irrelevantes provocan una pequeña reducción de la suma de los cuadrados de los errores. Por lo tanto, el r_{aj}^2 permite comparar mejor los modelos de regresión multivariado que tienen diferentes números de variables independientes (Dagnino S., 2014).

Coefficiente 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. Usamos r como otra medida de la fuerza de la interacción entre la variable dependiente y las variables independientes.

Por consiguiente, se compara a la correlación entre Y y X en la regresión lineal simple (Liviano Solís & Pujol Jover, 2019).

CAPÍTULO III

MARCO METODOLÓGICO

3.1 Tipo de Investigación

Investigación Exploratoria

El estudio actual se caracterizó por ser exploratorio, ya que es una temática que aún falta de profundizar, por tal motivo por medio de la recolección de datos se podrá develar los patrones que recibirá la explicación con la teoría y el apoyo pragmático, alcanzándose por ende condiciones para la elaboración de un modelo matemático.

Investigación Documental

La investigación también es documental en el sentido de que se desarrolló utilizando diversos materiales bibliográficos tales como: libros, artículos, artículos académicos y otros documentos que fundamentan la presente investigación en diversos aspectos del modelo matemático.

Investigación de Campo

El trabajo de campo es útil para diseñar estudios. Esto se debe a que la investigación se realizó directamente en territorio (pequeñas empresas) para comprender la situación real del entorno del problema en el que el investigador está directamente involucrado en el estudio y estudia cómo se está estudiando cada variable. La información objeto de estudio se hace pública para que pueda ser interpretada, la cual a su vez puede ser consultada con grupos de personas y fuentes dentro de la pequeña empresa.

Investigación Aplicada

La investigación es aplicada porque tiene como objetivo generar conocimiento empleándolo directamente a los problemas del sector manufacturero. Se espera que se desarrollen modelos matemáticos para ayudar

a mejorar la previsión de producción. Se espera poder predecir pronósticos de producción basados en la incidencia de algunas variables independientes.

Investigación Cuantitativa

El estudio es cuantitativo ya que se trabaja y clasifica datos numéricos, manifestados en, registros de producción semanal. Por lo que permite identificar tendencias o comportamientos de producción, hacer proyecciones, verificar correlaciones y obtener soluciones generales de producción. Gracias a esta información se ha permitido lograr procesos de optimización en la producción facilitando mejores niveles de productividad y eficiencia.

3.2 Diseño de la Investigación

Como meta de la presente investigación fue proponer un modelo matemático apropiado que permita predecir rápidamente la producción en planta de galleta de coco presentación clásico 100 * 25 g utilizando la regresión multivariada. El diseño del presente estudio es no empírico, ya que el fenómeno será observado y analizado sin hacer perturbaciones a las variables. La investigación también es de tipo longitudinal, porque analizamos lo que está pasando o la conducta en diferentes momentos del tiempo de las variables que se caracterizan por ser independientes y, que pueden interferir en la recomendación de la modelación.

3.3 Población y muestra

El presente estudio ha considerado todos los datos, correspondiendo el 100 % de la data recopilada de la producción de galleta de coco presentación clásico 100 * 25 g por una pequeña empresa en los últimos dos años. Cabe señalar que los datos recopilados corresponden a la producción semanal, y hubo estados donde este producto no fue elaborado por pequeñas empresas debido a la baja demanda del mercado.

3.4 Variables de investigación

3.5.1 Variable dependiente:

La variable dependiente se considera a la variable “Cantidad de Producción Real de CPR” en kg, las cifras se toman de la documentación proporcionada por la empresa y representa a los kg reales producidos por la micro empresa a partir de una galleta de coco presentación clásico 100 * 25 g.

Tabla 2

Cantidad de Producción Real (kg)

Variable dependiente	Cantidad de Producción Real
Unidad de medida	Kilogramos

Nota: La tabla muestra la cantidad de producción real, considerando que tipo de variable es y la unidad de medida. Elaborado por el autor.

3.5.2 Variable independiente:

Como primera aproximación, analizamos las cuatro posibles variables independientes, y cuando analizamos las variables de exclusión, esperamos ver el grado, peso o contribución de cada variable al modelo matemático. Poca o ninguna contribución al modelo se considera Atención o se excluye.

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

La cantidad de producción planificada es la cantidad de kilogramos de producto de galleta de coco presentación clásico 100 * 25 g que la microempresa desea o planea producir, es decir, el valor teórico, deseado o planificado.

Tabla 3

Cantidad de Producción Programada (kg)

Variable independiente	Cantidad de Producción Programada
Unidad de medida	Kilogramos

Nota: La tabla muestra la cantidad de producción programada, considerando que tipo de variable es y la unidad de medida. Elaborado por el autor.

b) Horas Laborales Planificadas HLP (hr).

Las horas de trabajo programadas corresponden a las horas de trabajo planificadas o teóricas que todos deben trabajar durante la semana para respetar la producción, esto es lo que creemos que podemos hacer sin contar con los riesgos para lograr el objetivo. Así conocen las empresas o las llaman Horas Hombre.

Tabla 4

Horas Laborales Planificadas (h)

Variable independiente	Horas Laborales Planificadas
Unidad de medida	Hora

Nota: La tabla muestra las Horas Laborales Planificadas, considerando que tipo de variable es y la unidad de medida. Elaborado por el autor.

c) Paradas No Programadas PNP (hr).

El tiempo de inactividad no planificado es el tiempo perdido debido al tiempo de inactividad de la máquina debido a un mal funcionamiento, falla o daño, es decir, la máquina entra en servicio para su reparación y sale de la línea de producción. Este tiempo afecta a la producción, por lo que su valor es el menor posible.

Tabla 5

Paradas No Programadas (h)

Variable independiente	Paradas No Programadas
Unidad de medida	Hora

Nota: La tabla muestra la cantidad de producción real, considerando que tipo de variable es y la unidad de medida. Elaborado por el autor.

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

El procedimiento aplicado para la investigación es cuantitativo, formulado matemáticamente en el método de análisis de regresión multivariante o regresión lineal múltiple, el cual evalúa las condiciones subyacentes del modelo de regresión y realiza un análisis de las variables consideradas excluidas para asegurar que su contribución al modelo sea adecuada, por lo que se decidió si se debe excluir a través de cualquiera de las variables independientes. La calidad de ajuste se evidencia mediante coeficientes de determinación (R^2) y análisis de varianza (anova), encontrando al final del proceso los coeficientes de regresión multivariante y significación para las variables especificadas.

La data se procesa utilizando el software matemático – estadístico IBM SPSS Versión 24 para un adecuado procesamiento y el respectivo examen de la información.

CAPITULO IV

RESULTADO Y DISCUSIÓN

4.1 Formulación del modelo de regresión lineal múltiple

Hay que considerar que en un modelo matemático en base a la regresión multivariada primero se deben establecer cuáles son las variables, ya sea independientes (X) y dependiente (Y), por lo que se podrá estructurar la ecuación respectiva.

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.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

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.

Es sumamente importante realizar de forma sistemática y científica la validación respectiva del modelo, ello se logra gracias a un adecuado proceso de diagnóstico o verificación, evitando sesgos al mismo. De esta forma se valora si las hipótesis indicadas al inicio de la investigación y que han servido para la construcción del modelo matemático, son realmente ciertas, siendo necesario:

- **Linealidad:** Este supuesto expresa que las variables independientes y dependientes tanto para su gráfica como en la correlación de Pearson muestran una correlación positiva entre las variables, de esta forma variables independientes como CPP y HLP tienen valores superiores a 0,88 con respecto a la variable dependiente.
- **Normalidad de los errores:** La modelación matemática aplicada a las variables escogidas demuestra a través del Gráfico P-P Normal de Regresión Residuos Estandarizados los datos están por encima y debajo de la línea, existiendo no normalidad de los residuos.
- **Homocedasticidad:** La métrica de la varianza del error resulta ser constante durante las observaciones. Para ello la gráfica de dispersión se demuestra con puntos aceptablemente distribuidos.
- **Independencia de los errores:** Para ello se aplica la prueba Durbin-Watson logrando tener un valor de 2,266 siendo aproximado a 2, confirmando por ende el supuesto.
- El supuesto de no multicolinealidad se cumple ya que según factor de varianza inflada facilita el diagnóstico de colinealidad respectivo el cual demuestra el supuesto ya que ningún índice es mayor a 30.

4.3 Supuesto de Normalidad de los Residuos

La distribución normal tiene que darse ya sea en variables independientes y variable dependiente. Cumplir este supuesto es indicar que hay normalidad en la información obtenida.

Teniendo una muestra que supera los 50 datos fue necesario aplicar la prueba Kolmogorov – Smirnov (K-S), para que nos indique el nivel de significancia, siendo condición de que sea mayor a 0.05 para indicar que la muestra cumple con ser distribución normal.

La prueba se la realiza a una muestra y cotejó el nivel de distribución acumulada analizada en la variable en la que se pone a prueba el carácter uniforme de los datos.

Tabla 6

Prueba de Kolmogorov-Smirnov para una muestra

		Standardize d Residual
N		58
Parámetros normales ^{a,b}	Media	0,000000
	Desviación estándar	0,97332853
Máximas diferencias extremas	Absoluta	0,129
	Positivo	0,129
	Negativo	-0,106
Estadístico de prueba		0,129
Sig. asintótica (bilateral)		0,018 ^c
a. La distribución de prueba es normal.		
b. Se calcula a partir de datos.		
c. Corrección de significación de Lilliefors.		

Nota: La tabla muestra la aplicación de la prueba Kolmogorov – Smirnov (K-S).

Elaborado por el autor.

Figura 1

Resumen de prueba de hipótesis

	Hipótesis nula	Prueba	Sig.	Decisión
1	La distribución de Standardized Prueba de Residual es normal con la media	Kolmogorov-Smirnov para una muestra	18,000 ¹	Rechazar la hipótesis nula.

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

¹Lilliefors corregida

Nota: La figura resume la prueba Kolmogorov – Smirnov (K-S). Elaborado por: El autor.

Se valora que los resultados expuestos en la tabla 6 tiene un nivel de significancia de 0,018; siendo el valor de p (sig) < 0.05; quedando claro que no cumple con tener normalidad de los datos.

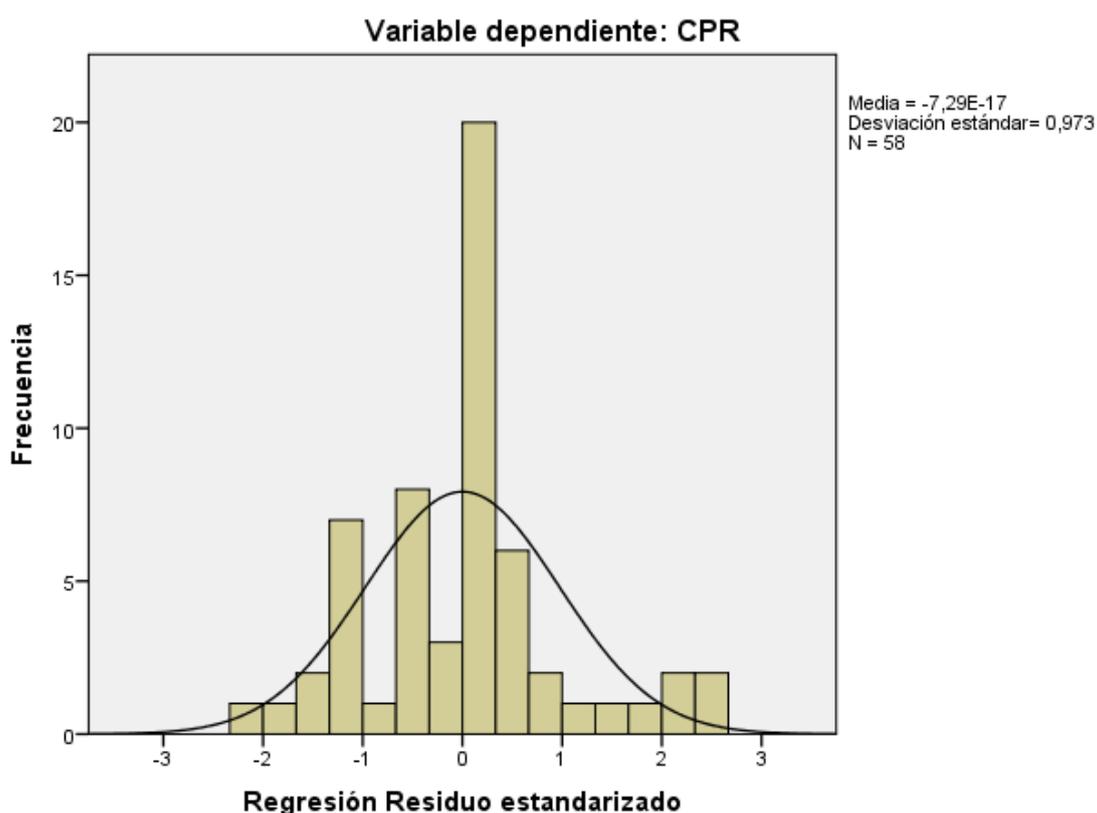
Se aprecia que en la tabla 6 como en la figura 1 los residuos no cumplen la condición de una distribución normal, por lo que se debe rechazar la hipótesis de cumplimiento. A su vez fue necesario analizar con los demás supuestos del modelo los datos y valorarlos dentro de la distribución señalada por la muestra.

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

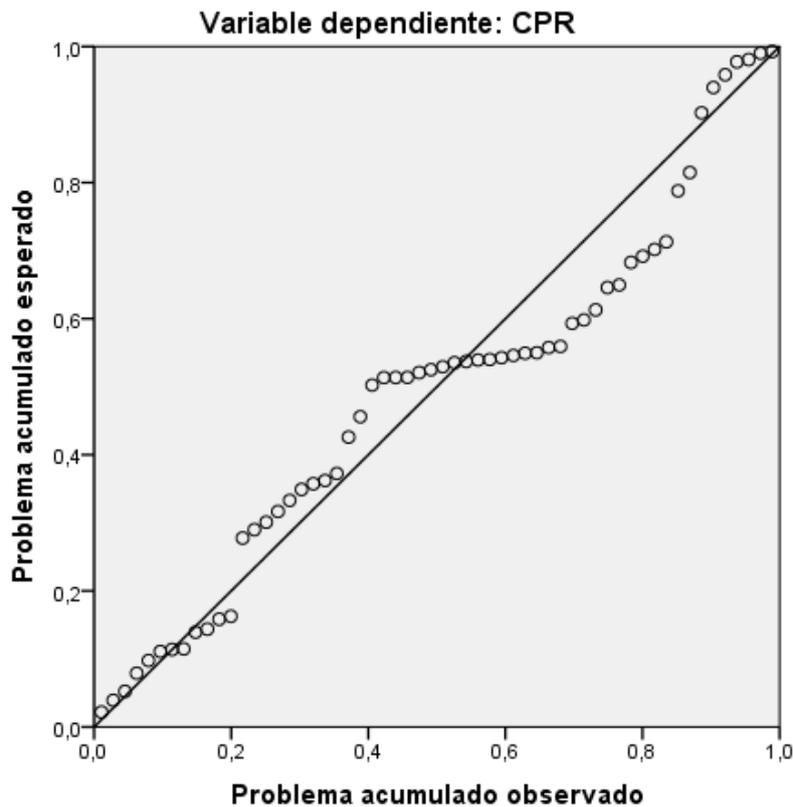


Nota: La figura muestra el Histograma de Regresión de Residuos Estandarizados para el análisis de la distribución. Elaborado por: El autor.

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 derecho donde existe un pequeño grupo de datos.

Figura 3

Gráfico P-P Normal de Regresión Residuos Estandarizados



Nota: La figura muestra Gráfico P-P Normal de Regresión Residuos Estandarizados. Elaborado por: El autor.

Tal como expresa la figura 3, Gráfico P-P Normal de Regresión Residuos Estandarizados, los datos en su mayoría están por encima y por debajo de la línea, siendo pocos los que están algo apegado a esta, con lo que se puede decir que estos no cumplen la condición de ser una distribución normal.

La no normalidad de los residuos se puede deber a la aleatoriedad de las paradas no programadas; por lo que se podría tener como opción no tomar en cuenta estos datos para el modelo.

Con lo expuesto se verifica el objetivo: “Analizar la variabilidad de cada uno de los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g”, ya que la prueba Kolmogorov – Smirnov y el Gráfico P-P Normal de Regresión Residuos Estandarizados la cual observa la variabilidad y/o distribución de cada uno de los factores expresados en las variables que componen el modelo.

4.4 Supuesto de Independencia de las observaciones

Se aplica la prueba de Durbin-Watson para comprobar el supuesto de la independencia frente a los errores. Por lo que se consideró que los resultados de aplicar la prueba de Durbin- Watson estén cifrados entre 2 con movimientos que fluctúen entre + - 1. Se comprende que aquellos valores que están en área de aceptación, es decir entre 1 y 3, son aceptables para confirmar tienen un carácter independiente los residuos.

Tabla 7

Resumen del modelo^b

Mod elo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin- Watson
1	0,961 ^a	0,924	0,920	126,74784	2,266

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

b. Variable dependiente: CPR

Nota: La tabla el resumen del modelo y el Coeficiente de Durbin-Watson.

Elaborado por el autor.

La tabla 7 expresa cómo el coeficiente de Durbin-Watson resulta tener un valor de 2,266; considerándose el mismo un valor aproximado a 2, quedando confirmado el supuesto de independencia de las observaciones.

Dando también por satisfecho el objetivo “Analizar la variabilidad de cada uno de los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g”.

4.5 Supuesto de Homocedasticidad

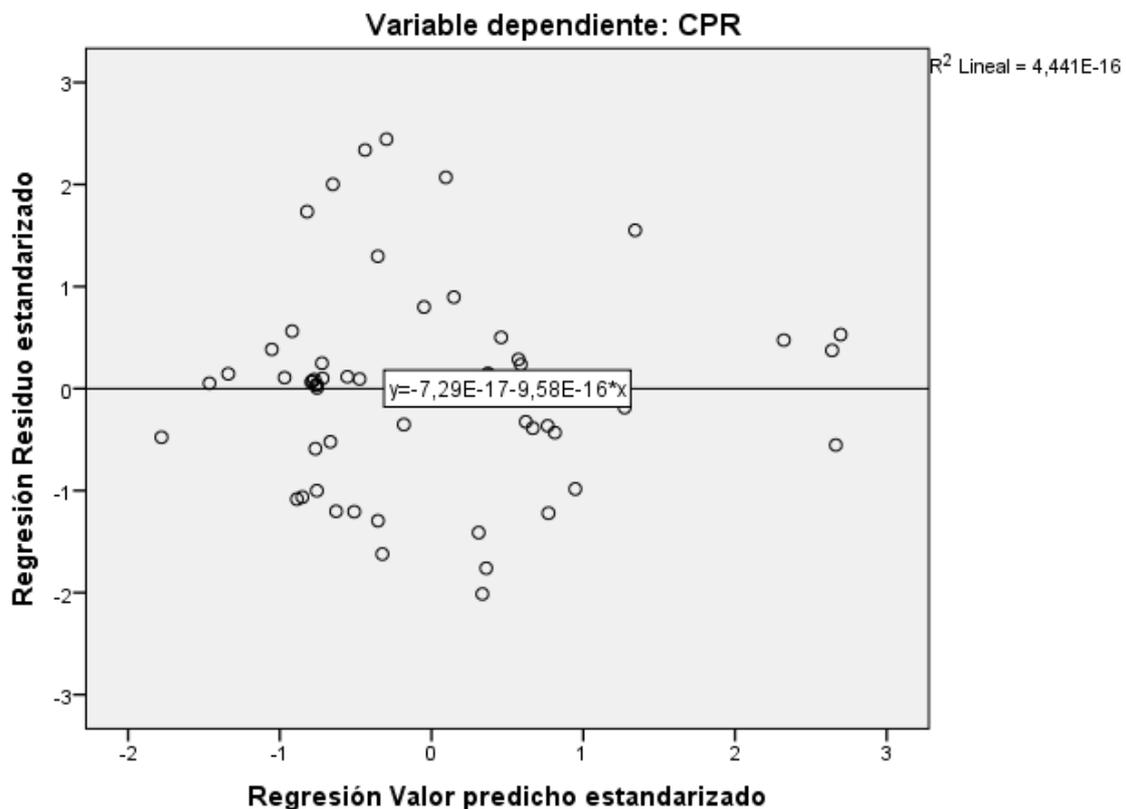
La homocedasticidad valora principalmente la fiabilidad de un modelo matemático en cuanto a su predicción. Para ello se instrumentaliza la observación a lo largo del tiempo de la varianza de los errores, observando si cumple con la condición de ser constante. También si a su vez se considera que

la varianza es pequeña, habrá mejores niveles de fiabilidad al momento de predecir.

A través de la gráfica se evidenció el cumplimiento de la condición de homocedasticidad; por lo que hay una distribución homocedástica si los residuos se distribuyen de tal forma en las puntuaciones indicadas.

Figura 4

Gráfico de dispersión



Nota: La figura muestra Gráfico de Dispersión para el análisis de la distribución de los puntos. Elaborado por: El autor.

El gráfico de la figura 4 sobre el examen de dispersión, constata el hecho de que los puntos estén razonablemente distribuidos, situación que expresa nula relación sistemática entre los residuos tipificados y los valores de pronóstico y cantidad de producción.

Lo anteriormente indicado expresa que se cumple con el supuesto de homocedasticidad ya que la gráfica expresa igual dispersión, por lo que el carácter constante de las variables permiten que el modelo predictor tenga adecuados niveles de fiabilidad.

Por lo que se ha cumplido con el objetivo “Analizar la variabilidad de cada uno de los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g”.

4.6.- Supuesto de linealidad

Este supuesto expresa que debe de haber matemáticamente una vinculación lineal entre las variables, ya sea dependientes e independientes.

Tabla 8

Correlaciones

		CPP	HLP	PNP	CPR
CPP	Correlación de Pearson	1	0,828**	0,698**	0,881**
	Sig. (bilateral)		0,000	0,000	0,000
	N	58	58	58	58
HL P	Correlación de Pearson	0,828**	1	,737**	0,945**
	Sig. (bilateral)	0,000		0,000	0,000
	N	58	58	58	58
PN P	Correlación de Pearson	0,698**	0,737**	1	0,720**
	Sig. (bilateral)	0,000	0,000		0,000
	N	58	58	58	58
CP R	Correlación de Pearson	0,881**	0,945**	0,720**	1
	Sig. (bilateral)	0,000	0,000	0,000	
	N	58	58	58	58

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

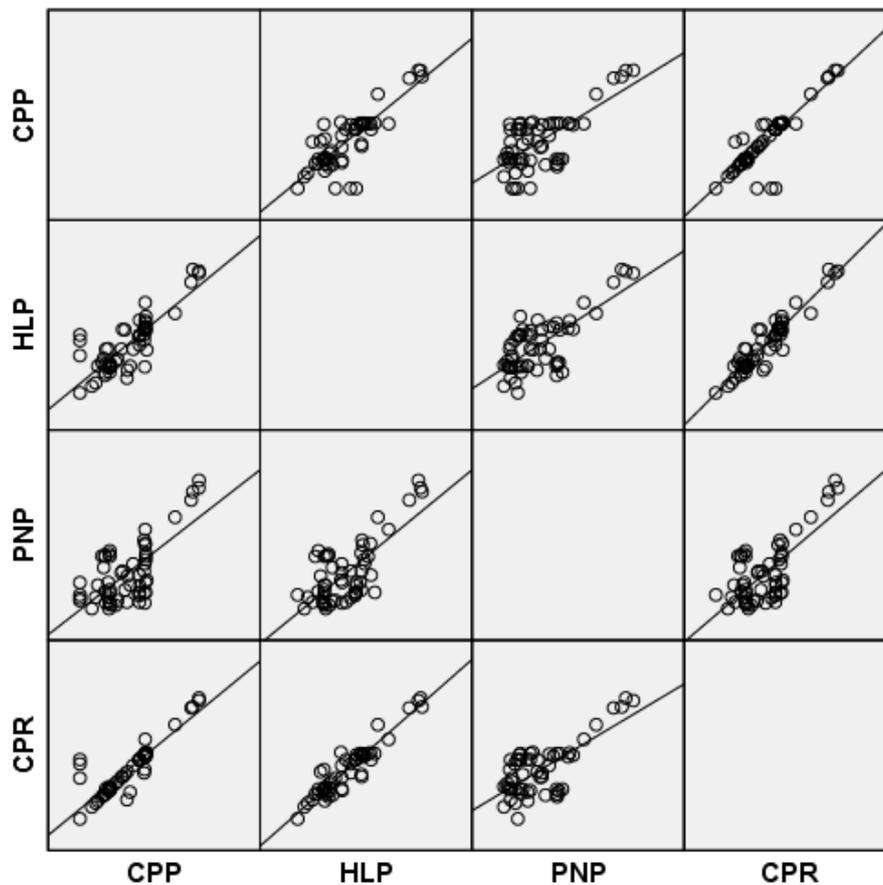
Nota: La tabla indica las correlaciones existentes entre las variables independientes con la variable dependiente. Elaborado por el autor.

Según lo indicado en las correlaciones de la tabla 8 de, entre las variables dependiente e independientes hay una positiva correlación. Se muestran en la matriz todos los valores de correlación de las variables, siendo que CPP tiene

0,881; HLP 0,945 mostrando una alta correlación con la variable dependiente CPR mayor a 0,88; y la variable independiente PNP manifiesta tener una media de 0,720 de correlación con la variable de salida, por lo tanto, se puede afirmar que si existe correlación entre las variables predictoras y la variable explicada.

Con ello se constata el cumplimiento del objetivo: “Identificar los factores que inciden en la producción de galleta de coco presentación clásico 100 * 25 g” ya que cada factor expresado en las variables independientes con la dependiente son adecuadas para el modelamiento de regresión lineal multivariante lo cual quedó patentizado con los resultados de la Correlación de Pearson y en la semejanza de la distribución de variables en la curva normal.

Figura 5
Correlación de variables



Nota: La figura muestra las correlaciones entre las variables independientes y dependiente, siendo base para el análisis de linealidad. Elaborado por el autor.

Como indica la figura 5, se verifica la premisa (supuesto) de linealidad entre la variable explicada CPR y las variables explicativas CPP, HLP, PNP, lo que se puede corroborar en la figura al momento de percatarse sobre la correlación positiva entre las variables ya sea independientes como dependiente.

Se puede afirmar que se cumple de forma irrestricta con el supuesto de linealidad de las variables. Dando también por satisfecho el objetivo: “Analizar la variabilidad de cada uno de los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g”.

4.7 Supuesto de no colinealidad

Para el análisis del supuesto de colinealidad se aplicará el “factor de varianza inflada” (VIF), para saber si hay multicolinealidad, es decir, una vinculación de dependencia lineal elevada entre dos o más variables independientes del modelo de regresión múltiple incumpliendo el enunciado de Gauss-Markov en condiciones de exactitud. Ninguna variable explicativa debe de superar el valor de 10 en la medición, sino simplemente no habría multicolinealidad en las diferentes variables.

Tal cual se observa en la tabla 10, los resultados del modelo indican que no hay valores del VIF que sean mayor a 10.

En esta situación se aplica la matriz de Diagnósticos de Colinealidad, por lo que se debe aplicar el siguiente proceso:

1. Señalar aquellos índices que estén superando la medida de 30
2. Una vez señalados los índices, hay que identificar las variables cuyas varianzas superen el 90%; por lo que se considerará multicolineal si ello sucede a dos o más coeficientes.

Tabla 9**Diagnósticos de colinealidad^a**

Modelo	Dimensión	Autovalor	Índice de condición	Proporciones de varianza			
				(Constante)	CP	HL	PNP
1	1	3,619	1,000	0,01	0,01	0,00	0,01
	2	0,249	3,812	0,43	0,00	0,00	0,35
	3	0,094	6,208	0,44	0,34	0,05	0,59
	4	0,038	9,765	0,12	0,66	0,94	0,05

a. Variable dependiente: CPR

Nota: La tabla muestra información relacionada con la colinealidad en la variable dependiente con respecto de las demás variables. 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.

Según indica la tabla 9, ningún índice de condición es mayor a 30 lo cual evidencia que no se puede descartar el supuesto, por lo tanto, se infiere que No hay multicolinealidad expresada entre dos o más variables, por lo que las variables independientes no están correlacionadas entre sí, ante lo cual se da por satisfecho el supuesto de no colinealidad. De igual forma se ha cumplido con el objetivo: “Analizar la variabilidad de cada uno de los factores que influyen en la producción de galleta de coco presentación clásico 100 * 25 g”.

4.8 El Modelo Matemático de Regresión Multivariada

Se considera como coeficiente de determinación el valor de R cuadrado, el cual está oscilando entre de 0 a 1 es 0 al 100%, lo que a su vez se comprende que solo se puede pronosticar un % de la ecuación lineal.

El valor del R-cuadrado ajustado considera número de variables independientes que se trabajan para proyectar a la variable dependiente.

Modelo. – El modelo fue desarrollado con tres variables independientes (explicativas): la cantidad de producción planificada CPP, las horas laborales planificadas HLP, paradas no programadas PNP, y como variable dependiente

(explicada) la cantidad de producción real CPR, tal cual está expresado en el modelo el R^2 de 0,924, esto indica que con las tres variables independientes se puede explicar un 92,4% de la varianza; y se obtiene un R^2 ajustado del 0,920.

De esta manera se justifica aplicar un Modelo Matemático compuesto por 3 variables, observando que el mismo se adapta adecuadamente al contexto, por lo que las tres variables inciden de buena manera en la proyección de la variable independiente.

Análisis de la Varianza

Tabla 10

ANOVA^a

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	10548047,330	3	3516015,778	218,862	0,000 ^b
	n					
	Residuo	867510,822	54	16065,015		
	Total	11415558,160	57			

a. Variable dependiente: CPR

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

Nota: La tabla registra los datos de la prueba de análisis de la varianza comparando las variables independientes y la dependiente. Elaborado por el autor.

La tabla 10 representa la prueba de análisis de varianza, comparando las varianzas explicadas y no explicadas y a su vez indicando si la varianza explicada por medio de la regresión es significativamente diferente a la varianza no explicada.

La modelación matemática aplicada muestra un valor del estadístico de prueba $F = 218,862$ y tiene un P -valor igual a $0 < 0.05$, de esta manera el modelo matemático aplicado se permite rechazar la hipótesis nula y se infiere que la dependencia lineal es estadísticamente significativa por lo que se deduce que el modelo es correcto dentro del contexto desarrollado.

Tabla 11***Coefficientes^a***

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados			Estadísticas de colinealidad	
	B	Error estándar	Beta	T	Sig.	Tolerancia	VIF
(Constante)	-0,343	37,261		-0,009	0,993		
1 CPP	0,298	0,065	0,317	4,610	0,000	0,298	3,360
HLP	40,521	4,281	0,690	9,465	0,000	0,265	3,775
PNP	-7,024	39,785	-0,010	-0,177	0,861	0,432	2,314

a. Variable dependiente: CPR

Nota: La tabla muestra el cálculo de los coeficientes del modelo matemático.

Elaborado por el autor.

Valorando datos de la tabla 11 se infiere que el modelo matemático en base a la regresión multivariada para el modelo de pronóstico es

$$Y = -0,343 + 0,298 X_1 + 40,521 X_2 - 7,024 X_3$$

Donde:

$$Y = \text{CPR}$$

$$X_1 = \text{CPP}$$

$$X_2 = \text{HLP}$$

$$X_3 = \text{PNP}$$

La ecuación nos indica que se ha podido cimentar la elaboración de un algoritmo que facilita la predicción de la cantidad de producción real por medio de tres variables de entrada.

La ecuación de regresión en formas estandarizadas expone el modelo de las variables en la misma dimensión como se indica a continuación:

$$Z = 0,317 Z_1 + 0,690 X_2 - 0,010 X_3$$

De esta forma se logra cumplir el objetivo “Construir un modelo matemático ajustado que facilite predecir de forma eficiente la producción aplicándolo a una fábrica de galleta de coco presentación clásico 100 * 25 g utilizando regresión multivariada”; quedando conformada la ecuación del modelo de predicción con sus coeficientes .

4.9 Comparación de los datos reales con los datos pronosticados

La comparación de datos se la realiza a partir de los datos de producción real CPR con los datos recabados por medio del modelo matemático propuesto CPR_RM.

Tabla 12

Estadísticas de muestras emparejadas

		Media	N	Desviación estándar	Media de error estándar
Par 1	CPR_RM	825,9349	58	430,07681	56,47185
	CPR	826,1207	58	447,51866	58,76208

Nota: La tabla muestra el cálculo de la media tanto en la producción real con el calculado con el modelo. Elaborado por el autor.

Tal como está indicado, la media de los datos logrados con la modelización matemática (modelo) CPR_RM es de 825,9349 y la media de los datos reales observados CPR es de 826,1207, siendo estas medias muy similares o análogas, por lo que, al indicar que existe una divergencia mínima entre ellas. Por lo que se puede concluir que el modelo matemático es muy bueno y acorde a los supuestos de la investigación planteados.

Tabla 13**Correlaciones de muestras emparejadas**

		N	Correlación	Sig.
Par	CPR_RM & CPR	58	0,961	0,000
1				

Nota: La tabla considera el cálculo de la correlación entre las variables.

Elaborado por el autor.

El modelo matemático tiene tanto entre la data procesada como de la información real observada su nivel de correlación es de 0,961, siendo una cifra cercana a 1 y tiene un P_ valor igual a $0 < 0.05$, deduciendo que es afirmativa la correlación entre variables.

Tabla 14**Prueba de muestras emparejadas**

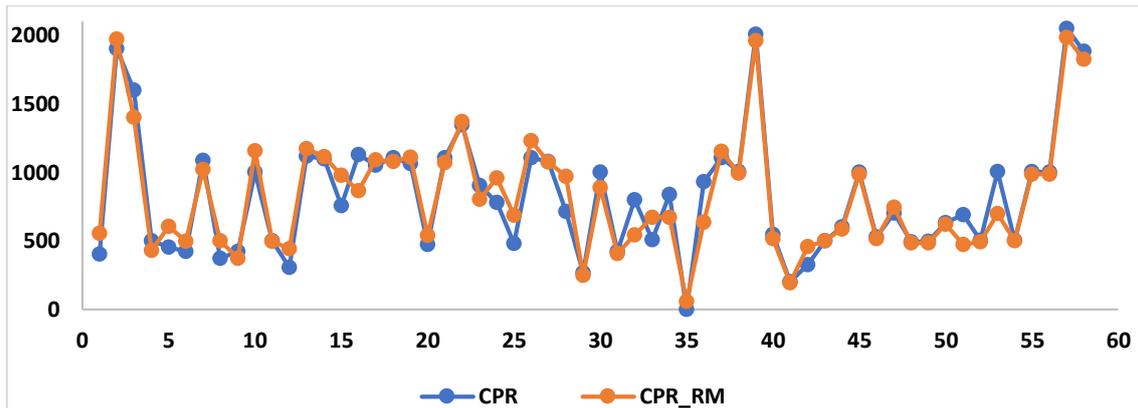
		Diferencias emparejadas							
		Media	Desviación	Media de	95% de intervalo de				
			n	error	confianza de la				
			estándar	estándar	diferencia				
					Inferior	Superior	t	gl	Sig.
					r				(bilateral)
Pa	CPR_R	-	123,36734	16,19893	-32,62356	32,2520	-	57	0,99
r l	M - CPR	0,1857				2	0,011		1
			7						

Nota: La tabla muestra las diferencias emparejadas entre las medias. Elaborado por el autor.

La tabla 14 indica la diferencia emparejada entre las medias, la cual es muy baja, es de -0,18577, quedando como el error que el modelo matemático aplicado realiza en su ejecución.

Figura 6

Comparación entre CPR_RM - CPR



Nota: La figura muestra la comparación entre la Capacidad de Producción Real con el logrado por el modelo matemático sobre la colinealidad en la variable dependiente con respecto de las demás variables. Elaborado por el autor.

La figura 6 aprecia el cotejo visual entre los valores reales obtenidos de la planta de producción es decir Capacidad de Producción Real CPR, con lo valores logados a través del uso del modelo matemático planteado por medio de la regresión multivariado CPR_RM.

Con los indicadores planteados se puede valorar que el modelo expresa un nivel adecuado de aproximación, un adecuado ajuste, infiriendo que el modelo es útil para predecir.

CONCLUSIONES

A lo largo de esta investigación, se presentó de forma breve la importancia de los pronósticos, los modelos matemáticos y se mencionaron algunas áreas de aplicación dentro de la microempresa, y en la propuesta se presenta de manera práctica esta idea en el desarrollo de todos los supuestos y finalmente con la obtención del modelo matemático.

Se han identificado y caracterizado las variables más relevantes relacionadas con el proceso productivo de la microempresa dedicada a la fabricación de galletas de coco clásica en presentación de 100*25 g que son: Cantidad de producción planeada CPP, Horas laborales planificadas HLP, Paradas no programadas PNP, Cantidad de producción real CPR.

La variabilidad de cada uno de los factores que influyen en la producción de la galleta se logra por medio de la construcción del modelo de regresión lineal multivariante considera las variables escogidas y la información cuantitativa se completa con el software estadístico SPSS, para posteriormente validar el modelo de regresión en base a sus supuestos, considerando un nivel de significancia de 0,018 en la prueba Kolmogorov – Smirnov, en la prueba de Durbin – Watson obtuvo 2,266 confirmando la independencia de observaciones, se cumple con la linealidad de las observaciones, el diagnóstico de colinealidad manifiesta no haber problemas de variables independientes correlacionadas entre sí.

La validación del modelo de regresión multivariado para microempresas dedicada a la producción de galletas de coco clásica en presentación de 100*25 g usando las variables independientes seleccionadas: cantidad de producción planeada CPP, horas laborales planificadas HLP, paradas no programadas PNP, permiten predecir la cantidad de producción real CPR y el modelo matemático tiene una relación significativa del 92,4%, por lo que se puede afirmar que el modelo construido permite pronosticar de forma adecuada.

RECOMENDACIONES

En este estudio se recomienda implementar el modelo matemático obtenido por regresión multivariante. El cual actuará como un insumo para la microempresa, permitiéndole predecir de manera óptima la cantidad galletas de coco clásica en presentación de 100*25 g. De forma que los modelos matemáticos sirven como herramienta para mejorar el proceso.

La microempresa debe de contar con una base de datos referente a todas las posibles variables que afecten o formen parte del proceso productivo, porque en base a datos históricos se puede pronosticar su posible futuro, trabajando con modelos matemáticos que permitan lograr el éxito de la microempresa desde una competitividad estratégica.

La microempresa debe comenzar a utilizar el análisis de sus datos históricos del proceso de producción, para mejorar el proceso de toma decisiones, de forma tal que esta sea más productiva, pueda lograr una mayor rentabilidad que sus competidores.

BIBLIOGRAFÍA

- Yeste, P., & Barat, R. (2019). Modelación matemática del tratamiento del efluente de un Biorreactor Anaerobio de Membranas (AnMBR). *Ribagua Revista Iberoamericana del Agua*, 148 - 157. doi:10.1080/23863781.2019.1695551
- ARAVENA, M., CAAMANO, C., & GIMENEZ, J. (2008). Modelos matemáticos a través de proyectos. *Revista latinoamericana de investigación en matemática educativa*, 11(1), 49-92. Obtenido de <http://www.scielo.org.mx/pdf/relime/v11n1/v11n1a3.pdf>
- Araya Pizarro, S., & Rojas Escobar, L. (2020). CONSUMO RESPONSABLE E INTENCIÓN DE COMPRA EN SECTORES POPULARES: UNA APROXIMACIÓN MULTIVARIANTE. *Ciencias Administrativas*(16). doi:<https://doi.org/10.24215/23143738e062>
- Arellano, A., & Peña, D. (2020). Modelos de regresión lineal para predecir el consumo de agua potable. *Revista Digital Novasinería*, 3(1), 27 - 36. doi:<https://doi.org/10.37135/ns.01.05.03>
- Arroyo García, D., Sánchez Partida, D., Juárez García, E., & Martínez Flores, J. (2017). Suavizamiento de la demanda del producto final con impacto en el inventario de materias primas de una empresa embotelladora. *Revista de Ingeniería Industrial*, 1(1), 48-62. Obtenido de https://www.ecorfan.org/republicofperu/research_journals/Revista_de_Ingenieria_Industrial/vol1num1/Revista_de_Ingenier%c3%ada_Industrial_V1_N1_4.pdf
- Banco Central del Ecuador. (2017). *Cuentas Nacionales*. Quito: Banco Central del Ecuador.
- Bocco, M. (2010). *Funciones elementales para construir modelos matemáticos*. Buenos aires: Instituto Nacional de Educación Tecnológica. Obtenido de <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. Obtenido de <https://www.redalyc.org/journal/1250/125065445002/125065445002.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. Obtenido de <http://econoquantum.cucea.udg.mx/index.php/EQ/article/view/7120/6215>
- Cacho, M. J., & Espinoza-Layana, D. (2019). Análisis Econométrico de la Producción en la Industria de Alimentos y Bebidas del Ecuador Durante el Periodo 2007 –2017. *X-Pedientes Económicos*, 3(6), 6-22. Obtenido de

https://ojs.supercias.gob.ec/index.php/X-pedientes_Economicos/article/view/85/27

- Cárdenas Pérez, A., & Benavides Echeverría, I. (2021). EXPLICACIÓN DEL CRECIMIENTO ECONÓMICO EN LA ECONOMÍA POPULAR Y SOLIDARIA MEDIANTE LA APLICACIÓN DEL MODELO ECONOMÉTRICO DE REGRESIÓN LINEAL Y MÚLTIPLE. *revistapublicando.org*, 74 - 84.
- Correa Morales, J., & Barrera Causil, C. (2019). *Introducción a la estadística bayesiana*. Medellín: Instituto Tecnológico Metropolitano.
- Cruz Córdova, S. (2017). *Aplicación de la regresión múltiple para predecir las variables que influyen en el peso del racimo en el cultivo de uva RED GLOBE*. Piura: Universidad Nacional de Piura. Obtenido de <https://repositorio.unp.edu.pe/bitstream/handle/UNP/543/EST-CRU-COR-16.pdf?sequence=1&isAllowed=y>
- Dagnino S., J. (2014). Regresión lineal. *Revista Chilena de anestesia*, 43(2), 143 - 149. doi:<https://doi.org/10.25237/revchilanestv43n02.14>
- 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. Obtenido de <http://repositorio.unprg.edu.pe/handle/UNPRG/2192>
- Escalante Cortina, R., Mayorga Mogollón, W., & Vergara Schmalbach, J. (2012). *Manual de aplicación del modelo de regresión lineal múltiple con correcciones de selecciones, usos de STATA 9.0, STATA 10.0, EVIEWS 5.0, SSPS 11.0*. Madrid: EUMED.
- FAO. (07 de Octubre de 2021). FAO. (Situación Alimentaria Mundial) Recuperado el 08 de Octubre de 2021, de <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. doi:<https://doi.org/10.4060/ca8677es>
- Garcés, D., & Jaimes Barragán, F. (2017). Ronda clínica y epidemiológica. Introducción al análisis multivariable (parte II). *Iatreia*, 28(1), 87-96. Obtenido de <https://www.redalyc.org/articulo.oa?id=180533008009>
- García, S., Arguello, A., & Parra, R. (2019). Factores que influyen en el pH del agua mediante la aplicación de modelos de regresión lineal. *INNOVA Research Journal*, 4(2), 59 - 71. doi: <https://doi.org/10.33890/innova.v4.n2.2019.909>
- Gómez Pintado, L. (2017). *Factores que determinan la rentabilidad del cultivo de cacao en el distrito de Pólvora*. Universidad Nacional Agraria de la Selva. Obtenido de

<https://repositorio.unas.edu.pe/bitstream/handle/UNAS/1014/ECO2016001.pdf?sequence=1&isAllowed=y>

- González, R., Morell, C., & Blanco, A. (2017). Regresión lineal local con reducción de rango para problemas de predicción con salidas compuestas. *Revista cubana de ciencias informáticas*, 184 - 193.
- Gutiérrez, A., & Ferreira, W. (2020). Un modelo de regresión lineal aplicando lógica difusa. *Revista SEXTANTE*, 48 - 54.
- Hernández, J., Espinosa, J., & García, D. (2019). *Sobre el uso adecuado de la regresión lineal: conceptualización básica mediante un ejemplo aplicado a las ciencias de la salud*. Guaranda: Repositorio Universidad Estatal de Bolívar. Obtenido de <https://hdl.handle.net/20.500.12442/5045>
- Infante Roblejo, R., Martínez Serra, J., & Rodríguez Muñoz, R. (2019). *Modelación matemática de procesos económicos : aplicaciones* (Séptima ed.). Madrid, España: Editorial Universo Sur.
- Liviano Solís, D., & Pujol Jover, M. (2019). *Modelos de regresión y análisis multivariante con R commander*. Barcelona, España: Universidad Abierta de Cataluña.
- López Pineda, G. (2016). *Análisis de regresión para la estimación del secuestro de carbono orgánico en suelos*. Puebla: Universidad Autónoma de Puebla. Obtenido de <https://www.fcfm.buap.mx/assets/docs/docencia/tesis/ma/GabrielaLopezPineda.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. Obtenido de <http://repository.ucc.edu.co/handle/ucc/12151>
- 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. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=4778508>
- Meza Rodríguez, A., & Chafloque Chafloque, A. (2013). *Pronóstico para la producción de caña y azúcar terminada con un modelo de regresión lineal múltiple de dos etapas*. Trujillo: Universidad Nacional de Trujillo. Obtenido de <https://dspace.unitru.edu.pe/bitstream/handle/UNITRU/9233/MEZA%20RODRIGUEZ%20Aldo%20Richard.pdf?sequence=1&isAllowed=y>
- Montero Granados, R. (2017). *Modelos de regresión lineal múltiple*. España: Universidad de Granada. Obtenido de https://www.ugr.es/~montero/maticas/regresion_lineal.pdf

- Montesinos-López, O. A., & Hernández-Suárez, C. M. (2007). Modelos matemáticos para enfermedades infecciosas. *Salud Pública de México*, 49(3), 218-226. Obtenido de <http://www.scielo.org.mx/pdf/spm/v49n3/07.pdf>
- Morales Quiroga, M., & Becerra, A. (2018). El efecto de la posición del candidato en la papeleta de votación. El caso de las elecciones locales chilenas de 2008 y 2012. *Colombia Internacional*(96), 29 -55.
- 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 . Obtenido de <https://www.redalyc.org/pdf/3783/378346436007.pdf>
- Moreno Castro, T. F. (2019). *El pronóstico de ventas en los negocios: modelos y aplicaciones*. Santiago de Chile: RIL editores. Obtenido de https://repositorio.uautonoma.cl/bitstream/handle/20.500.12728/3192/Pro_nostico_de_ventas.pdf?sequence=1&isAllowed=yhttp://webcache.googleusercontent.com/search?q=cache:Xw_RbmtUzsQJ:www.scielo.org.co/pdf/seec/v22n52/0120-6346-seec-22-52-161.pdf+&cd=8&hl=es
- Narro Ramírez, A. E. (1996). Aplicación de algunos modelos matemáticos a la toma de decisiones. *Política y Cultura*(6), 183-198. Obtenido de <https://www.redalyc.org/pdf/267/26700614.pdf>
- Nations, O. a. (2020). *OCDE-FAO Perspectivas Agrícolas*. OCDE-FAO. doi:<https://doi.org/10.1787/a0848ac0-es>
- Návar, J., Lizárraga, L., & Jiménez, M. (2017). Modelación de caudales en función de los macroporos del suelo en una microcuenca forestal de Durango, México. *Terra Latinoamericana* , 2(35), 89 - 100.
- Ortega Mohedano, F., & Pinto Hernández, F. (2021). Predicción del bienestar sobre el uso de pantallas inteligentes de los niños. *Revista Científica de Educomunicación*, 119 - 128. doi:<https://doi.org/10.3916/C66-2021-10>
- Quevedo Urias, H., & Pérez Salvador, B. (2014). *Estadísticas para ingenierías y ciencias*. Mexico, México: Grupo Editorial Patria.
- Ramos, A., Sánchez, P., Ferrer , J. M., Barquín, J., & Linares, P. (2018). *Modelos matemáticos de optimización*. Madrid: Universidad Pontificia Comillas. Obtenido de https://d1wqtxts1xzle7.cloudfront.net/34357606/modelado_en_gams-with-cover-page-v2.pdf?Expires=1633687201&Signature=EzhFiG-skxPtDHclAwclQ08Qgji~w3TqbHmhGJKqaRfyW1SAldlZH5DDJXiHRLC8v2Tdb-zncNZYV4WfKu0LVXNlkmLiT6NoocHcf21rjOhupk0blXvhHdDbsWtw~Me7oBaSZ0g4ocy

- Render, B., & Heizer, J. (2007). *Administración de la producción*. México: PEARSON EDUCACIÓN.
- Rodríguez Velázquez, J. A., & Steegmann Pascual, C. (24 de Junio de 2013). *Modelos Matemáticos*. Recuperado el 8 de Octubre de 2021, de https://recursos.salonesvirtuales.com/wp-content/uploads/bloques/2012/08/Modelos_matematicos.pdf
- Rubio Guerrero, G. (2017). Perspectiva multivariante de los pronósticos en las PYMES industriales de Ibagué (Colombia). *Revista Facultad de Ciencias Económicas: Investigación y Reflexión*, 25(2), 25-40. Obtenido de <https://www.redalyc.org/pdf/909/90952679003.pdf>
- Ruiz Pozo, R. G., Soria Torres, L. O., & Vidarte Ureta, A. T. (2020). *Optimización de la producción de petróleo en el Campo Limoncocha mediante el análisis de pruebas de restauración de presión*. Quito: Universidad Central del Ecuador. Obtenido de <http://www.dspace.uce.edu.ec/handle/25000/21549>
- Salett Biembengut, M., & Hein, N. (1999). Modelación matemática: Estrategia para enseñar y aprenderr matemáticas. *Educación Matemática*, 11(1), 119-134. Obtenido de <http://www.revista-educacion-matematica.org.mx/descargas/Vol11/1/10Salett.pdf>
- Sánchez Sánchez, D. A. (2018). *Modelo ARIMA para el pronóstico de la produccion de cacao en el Prú 2012 - 2018*. Facultad de Ciencias Físicas y Matemáticas. Trujillo: Universidad Nacional de Trujillo. doi:<http://190.223.54.254/bitstream/handle/UNITRU/11563/SÁNCHEZ%20SÁNCHEZ%2c%20David%20Alexander.pdf?sequence=1&isAllowed=y>
- Superintendencia de Compañías, Valores y Seguros. (2018). *Panorama de las MIPYMES y Grandes Empresas en el Ecuador 2013-2017*. Quito: Superintendencia de Compañías, Valores y Seguros.
- Tello Cabello, S. Y. (2014). Importancia de la micro, pequeñas y medianas empresas en el desarrollo del país. *LEX*, 12(14), 199-218. doi:<http://dx.doi.org/10.21503/lex.v12i14.623>
- Tulcanaza Prieto, A. (2017). *Desarrollo de un modelo de caracterización financiera del sector comercio en Ecuador, periodo 2022-2012*. Quito: Universidad Andina Simón Bolívar. Obtenido de ervicioenlinea.ug.edu.ec/SIUG/MODULOS/ACADEMICO/ESTUDIANTE/WEB_CONSULTA_HORARIO.aspx?COD_USERNAME=0952811545&ROL_NAME=ACA-ESTUDIANTE&COD_DETALLE_MENU=6
- 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. Obtenido de <https://www.redalyc.org/pdf/2110/211026873005.pdf>

- Vicuña Pino, A., Cortez Chichande, B., & Basurto Tobar, Y. (2019). Determinación de modelos predictivos para los indicadores de competitividad empresarial aplicando regresión lineal. *Revista Ibérica de Sistemas e Tecnologías de Información*, 2(18), 94 - 107.
- Vicuña Pino, A., Cortez Chichande, S., & Basurto Tovar, Y. (Febrero de 2019). Determinación de modelos predictivos para los indicadores de competitividad empresarial aplicando regresión lineal. *Revista Ibérica de Sistemas e Tecnologías de Informação*(18), 94 - 107.
- Vilà Baños, R., Torrado Fonseca, M., & Reguant Alvarez, M. (2019). Análisis de regresión lineal múltiple con SPSS: un ejemplo práctico. *Revista de innovación e investigación en educación*, 12(2), 1-10. doi:<http://doi.org/10.1344/reire2019.12.222704>
- Ximénez, C., & San Martín, R. (2013). *Fundamentos de las técnicas multivariantes*. Madrid: NED - Universidad Nacional de Educación a Distancia.
- 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. doi:<https://doi.org/10.21500/01247492.1365>

ANEXOS

Tabla 15

MODELIZACIÓN

	Ctd. prod. Plan CPP	Horas laborables Planificadas HLP	paradas no programadas estimadas en horas PNP	Ctd. prod. Real CPR
ítem	CPP	HLP	PNP	CPR
1	400	10,917	0,84	403
2	1901,05	35,1	2,384	1902
3	1603,06	23,166	1,866	1600
4	502,01	7,15	1,183	502
5	845,02	8,85	0,483	454
6	422,01	9,35	1,083	422
7	1083,02	17,35	0,317	1087
8	370	9,833	1,076	374
9	422,01	6,3	1,067	422
10	1003,02	21,317	0,684	1003
11	502,01	8,783	1,066	502
12	300	8,85	0,483	307
13	1101,02	21,15	1,33	1121
14	1095,04	19,6	0,916	1100
15	757,03	18,716	0,383	758
16	1125,02	13,217	0,583	1129
17	1100,05	19,117	1,3	1052
18	1110	18,667	1,133	1109
19	1110	19,484	1,001	1064
20	496,98	9,85	1,117	473
21	1108,05	18,417	0,534	1109
22	1099,03	26,1	1,617	1350
23	898,05	13,417	0,917	906
24	1082,05	15,867	0,767	781
25	475,02	13,55	0,534	481
26	1108,98	22,35	0,333	1109
27	1082	18,883	1,4	1082
28	718	18,8	0,767	715
29	264,01	4,267	0,235	268
30	1003,02	14,6	0,133	1003
31	422	7,067	0,35	423
32	800	7,717	0,667	800
33	436,99	13,45	0,116	510
34	1092,02	8,617	0,117	838
35	0,06	1,55	0,283	0
36	0	15,783	0,217	934
37	1108,03	20,55	1,067	1109
38	1003,06	17,25	0,466	1007
39	2007,06	34,1	2,617	2009
40	491,04	9,117	0	547

41	200	3,383	0	203
42	792	5,55	0,133	325
43	502,06	8,75	0,383	502
44	600	10,1	0,067	602
45	1003,02	17,05	0,333	1003
46	528	8,9	0,067	530
47	698	13,45	0,734	703
48	485	8,45	0,234	493
49	494	8,433	0,15	497
50	633,03	10,733	0,166	634
51	0	11,717	0,167	693
52	502	8,517	0,1	505
53	0	17,333	0,533	1008
54	502	8,75	0,317	506
55	1002	17,016	0,266	1005
56	1004	17,034	0,318	1003
57	2006,05	34,681	2,465	2053
58	1874	31,633	2,217	1885

Nota: La tabla muestra los datos trabajados por cada variable. Elaborado por el autor.

Tabla 16

Comparativa entre la cantidad de producción real vs la cantidad de producción del modelo multivariado (predicción)

	CPR	CPR_RM
1	403	555,32
2	1902	1971,71
3	1600	1402,97
4	502	430,67
5	454	606,69
6	422	496,68
7	1087	1023,21
8	374	500,8
9	422	373,2
10	1003	1157,54
11	502	497,66
12	307	444,28
13	1121	1175,44
14	1100	1113,76
15	758	980,95
16	1129	866,38
17	1052	1092,98
18	1109	1078,88
19	1064	1112,92
20	473	539,04
21	1109	1072,38
22	1350	1373,41
23	906	804,51
24	781	959,67
25	481	686,52
26	1109	1233,44
27	1082	1077,42
28	715	970,03
29	268	249,58
30	1003	889,23
31	423	409,32
32	800	546,07
33	510	674,07
34	838	673,43
35	0	60,49
36	934	637,68
37	1109	1155,06
38	1007	994,28
39	2009	1961,15
40	547	515,42
41	203	196,34
42	325	459,63
43	502	501,14
44	602	587,25
45	1003	987,1

46	530	517,17
47	703	747,51
48	493	484,95
49	497	487,53
50	634	622,05
51	693	473,27
52	505	493,67
53	1008	698,26
54	506	501,59
55	1005	985,89
56	1003	986,85
57	2053	1985,45
58	1885	1824,34
max	2053	1985,45

Nota: La tabla muestra la comparación entre la Cantidad de Producción Real y la Cantidad de Producción del Modelo. Elaborado por el autor.