

This file has been cleaned of potential threats.

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

2cc77693e67454f25e8b169bbcb266d8e4aa77d739144aa3a3d6f1f52c1e2518

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



## **UNIVERSIDAD ESTATAL DE MILAGRO**

### **VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO**

**PROYECTO DE INVESTIGACIÓN PREVIO A LA OBTENCIÓN DEL  
TÍTULO DE:**

**MAGÍSTER EN MATEMÁTICA, MENCIÓN EN MODELACIÓN MATEMÁTICA.**

**TÍTULO DEL PROYECTO:**

**PROPUESTA DE UN MODELO MATEMÁTICO APLICADO AL PRONÓSTICO DE  
PRODUCCIÓN UTILIZANDO REGRESIÓN MULTIVARIADO APLICADO A UNA  
FÁBRICA DE GALLETA WAFER 40\*204 g.**

**TUTOR**

**MSC. CHELE DELGADO SANTIAGO JOSÉ.**

**AUTOR**

**ING. JUAN DAVID SOLÍS RIVERA.**

**MILAGRO, septiembre del 2022**

**ECUADOR**

## ACEPTACIÓN DEL TUTOR

En calidad de Tutor de Proyecto de Investigación, nombrado por el Comité Académico del Programa de Maestría en Matemática, mención en Modelación Matemática.

## 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 WAFER 40\*204 g**, elaborado por el **ING. JUAN DAVID SOLIS RIVERA**, el mismo que reúne las condiciones y requisitos previos para ser defendido ante el tribunal examinador, para optar por el título de **MAGÍSTER EN MATEMÁTICAS MENCIÓN MODELACIÓN MATEMÁTICA**.

Milagro, 29 de noviembre del 2021.



Firmado electrónicamente por:  
**SANTIAGO JOSE  
CHELE DELGADO**

---

**Msc. CHELE DELGADO SANTIAGO JOSÉ.**  
**0922712260**

## **DECLARACIÓN DE AUTORÍA DE LA INVESTIGACIÓN**

El / la autor/a de esta investigación declara ante el Comité Académico del Programa de Maestría en Matemática mención modelación 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, a los 28 días del mes de septiembre del 2022



Firmado electrónicamente por:  
**JUAN DAVID  
SOLIS**

---

**Ing. JUAN DAVID SOLÍS RIVERA  
C.I. 0918769993**

## CERTIFICACIÓN DE LA DEFENSA EI TRIBUNAL CALIFICADOR

Previo a la obtención del título de Máster en Ciencias Matemáticas otorga al presente proyecto de investigación las siguientes calificaciones:

TRABAJO DE TITULACION	57.33
DEFENSA ORAL	38.67
<b>PROMEDIO</b>	<b>96.00</b>
<b>EQUIVALENTE</b>	<b>Excelente</b>



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

---

Phd. VALENZUELA COBOS JUAN DIEGO  
**PRESIDENTE/A DEL TRIBUNAL**



firmado electrónicamente por:  
**EDWIN  
EVARISTO**

---

Phd. LEON PLUAS EDWIN EVARISTO  
**VOCAL**



firmado electrónicamente por:  
**LUIS HENRY  
TORRES  
ORDONEZ**

---

Mee. TORRES ORDOÑEZ LUIS HENRY  
**SECRETARIO**

## DEDICATORIA

A mis hijos, David y Matías, fuente de inspiración.

Ing. Juan David Solís Rivera

## **AGRADECIMIENTO**

A mis padres Ana y Juan, por su apoyo incondicional.

Ing. Juan David Solís Rivera

## CESIÓN DE DERECHOS DE AUTOR

Señor Doctor

**Fabricio Guevara Viejó.**

**Rector de la Universidad Estatal de Milagro Presente.**

**Ciudad. -**

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 WAFER 40\*204 g”** y que corresponde al Instituto de Posgrado y Educación Continua

Milagro, a los 28 días del mes de septiembre del 2022



Firmado electrónicamente por:  
**JUAN DAVID  
SOLIS**

---

**Ing. JUAN DAVID SOLÍS RIVERA  
C.I. 0918769993**

## ÍNDICE GENERAL

<b>INTRODUCCIÓN</b> .....	1
<b>CAPÍTULO I</b> .....	2
1.1 Planteamiento Del Problema.....	2
1.2 Delimitación del problema.....	4
1.3 Formulación del problema .....	4
1.4 Sistematización del problema .....	4
1.5 Objetivos .....	5
1.5.1 Objetivo General .....	5
1.5.2 Objetivos Específicos .....	5
1.6 Justificación e importancia .....	5
1.6.1 Justificación Teórica .....	6
1.6.2 Justificación metodológica .....	6
1.6.3 Justificación practica .....	7
1.7 Hipótesis.....	8
1.7.1 Hipótesis General .....	8
1.7.2 Hipótesis Particular .....	8
1.7.3 Conceptualización y operacionalización de variables .....	8
<b>CAPÍTULO II</b> .....	9
2.1 Marco Teórico .....	9
2.1.1 Modelo matemático.....	9
2.1.2 Pronóstico de producción.....	11
2.2 Marco Conceptual.....	21
2.4 Marco Referencial .....	23
<b>CAPÍTULO III</b> .....	24
<b>MARCO METODOLÓGICO</b> .....	24
3.1 Tipos de Investigación .....	24
3.2 Diseño de la Investigación .....	26
3.3 Población y muestra.....	26
3.4 Variables de investigación .....	26
3.5 Método, Técnica y Uso de Software de Tratamiento y Análisis de Datos .....	28
<b>CAPITULO IV</b> .....	29

<b>PROPUESTA .....</b>	<b>29</b>
<b>4.1 Formulación del modelo de regresión lineal múltiple .....</b>	<b>29</b>
<b>4.2 Diagnósis y validación del modelo de regresión lineal múltiple .....</b>	<b>29</b>
<b>4.3.- Supuesto de Normalidad.....</b>	<b>30</b>
<b>4.4.- Supuesto de No Autocorrelación o de Independencia de las observaciones .....</b>	<b>32</b>
<b>4.5.- Supuesto de Homocedasticidad.....</b>	<b>33</b>
<b>4.6.- Supuesto de linealidad.....</b>	<b>34</b>
<b>4.7.- Supuesto de Multicolinealidad.....</b>	<b>36</b>
<b>4.8 El Modelo Matemático de Regresión Multivariada .....</b>	<b>37</b>
<b>4.9 Comparación de los datos reales con los datos pronosticados .....</b>	<b>40</b>
<b>CONCLUSIONES .....</b>	<b>42</b>
<b>RECOMENDACIONES.....</b>	<b>43</b>
<b>BIBLIOGRAFÍA.....</b>	<b>44</b>

## ÍNDICE DE TABLAS

Tabla 1 .....	9
Tabla 2 .....	27
Tabla 3 .....	27
Tabla 4 .....	28
Tabla 5 .....	28
Tabla 6 .....	30
Tabla 7 .....	33
Tabla 8 .....	35
Tabla 9 .....	37
Tabla 10 .....	38
Tabla 11 .....	39
Tabla 12 .....	40
Tabla 13 .....	40
Tabla 14 .....	41

## ÍNDICE DE GRÁFICOS

Figura 1 .....	31
Figura 2 .....	32
Figura 3 .....	34
Figura 4 .....	35
Figura 5 .....	41

## ÍNDICE DE ANEXOS.

<b>ANEXO A. Base de datos del modelo .....</b>	<b>48</b>
<b>ANEXO B. Reunión de tutorías .....</b>	<b>49</b>

## RESUMEN

El trabajo de investigación tiene como objetivo el desarrollo de un modelo matemático que utiliza regresión múltiple para predecir la producción de galleta Wafer en presentación de 40\*204 g en la microempresa; para ello, se evaluó las características subjetivas de las principales variables que intervienen en el proceso productivo, teniendo en cuenta la información, se distinguen estas variables de otras. El método utilizado es el cuantitativo y las técnicas utilizadas para recopilar la información en este caso primaria las cuales son importantes. A continuación, para construir un modelo matemático se usará la regresión múltiple en el software estadístico SPSS, y luego usando criterios de suposición de regresión múltiple para confirmar y validar el modelo, los cuales son: supuesto de normalidad de residuos, supuesto de correlación de variables, supuesto de multicolinealidad, supuesto de independencia de errores, supuesto de homocedasticidad. De acuerdo a los resultados y análisis de este trabajo, se puede concluir que el modelo matemático propuesto es aceptable porque cumple con los supuestos de regresión y el valor de R-cuadrado es 99,6% que es un valor aceptable.

### **Palabras Claves:**

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

## ABSTRAC

The research work has as objective the development of a mathematical model that uses multiple regression to predict the production of Wafer cookie in presentation of 40\*204 g in the microenterprise; for this, the subjective characteristics of the main variables involved in the productive process were evaluated, taking into account the information, these variables are distinguished from others. The method used is quantitative and the techniques used to collect the information in this primary case which are important. Next, to build a mathematical model will be used multiple regression in SPSS statistical software, and then using multiple regression assumption criteria to confirm and validate the model, which are: assumption of normality of residuals, assumption of correlation of variables, assumption of multicollinearity, assumption of independence of errors, assumption of homoscedasticity. According to the results and analysis of this work, it can be concluded that the proposed mathematical model is acceptable because it complies with the regression assumptions and the R-squared value is 99.6% which is an acceptable value.

### **Keywords:**

Mathematical model, Forecasting, Multivariate regression, Process, Productivity.

## INTRODUCCIÓN

La correcta toma de decisiones en las microempresas es fundamental para alcanzar sus resultados esperados en el corto, mediano y largo plazo. En los actuales tiempos de crisis es cuando más importancia toma el pronóstico de producción dentro de las microempresas, la incorrecta forma de pronosticar dentro de la microempresa ha causado que estas incurran en pérdidas económicas, debido a producto que no se ha vendido o por no producir lo suficiente para vender.

En consecuencia, resulta necesario poder desarrollar un modelo matemático lo más cercano posible a la realidad de producción de cada pequeña empresa, el cual debe ser capaz de incorporar la mayor cantidad de variables que de una u otra forma afecta la producción, estas variables independientes son necesarias para el desarrollo del modelo matemático y representan el proceso de producción de la mejor manera posible.

El propósito de este estudio es poder desarrollar un modelo matemático mediante regresión multivariada, de forma tal que permita comprender cómo las variables independientes influyen y ayudan a describir el proceso de producción de la microempresa, poniendo de manifiesto la relación de causa y efecto entre las variables independientes y dependientes.

Capítulo I. En el desarrollar este tema, se define la delimitación formulación, sistematización del problema y se organiza detalladamente el planteamiento del problema; objetivos generales y específicos; Así como la justificación teórica, metodológica y práctica.

Capítulo II. Constituido de cuatro partes, el marco teórico, el marco conceptual y el marco de referencia; Esta sección refuerza la investigación de modelos matemáticos, regresión multivariante y pronósticos de producción que se estudian en libros, artículos, tesis, etc.

Capítulo III. Aquí se exponen los tipos y proyectos de investigación utilizados para desarrollar y analizar este trabajo, y se presentan las herramientas para recolectar y analizar muestras, poblaciones y datos.

Capítulo IV. En este capítulo se presenta una propuesta de modelo matemático basado en algunos supuestos de regresión multivariante tales como: relación lineal, independencia de las observaciones, varianzas, distribución normal de partes, residuos y otras hipótesis; Además, muestra la calidad del modelo matemático al comparar el valor de la producción real con el valor obtenido con la producción real esperada que se obtiene a través del modelo matemático desarrollado.

## **CAPÍTULO I**

### **1.1 Planteamiento Del Problema**

En las operaciones de una empresa existen variables aleatorias que resultan normalmente difíciles de predecir y analizar, individualmente, así como las relaciones entre ellas (Berkowitz, 1998). Al no poder conocerse las relaciones entre estas variables se llega a la situación de carecer de un modelo que permita describir la relación existente, lo que produce que se presente dentro de la empresa situaciones de complejidad al momento de examinar cada variable y sus efectos en el proceso de producción (Rangel Montes de Oca, GarcíaPereira, & Hernández Gómez, 2009).

En la actualidad las empresas presentan un déficit en la producción de productos frente a las demandas del mercado, esto resulta por los constantes cambios en la preferencia de los consumidores y la alta competencia o sobre ofertas que existen en el mercado de productos iguales o similares (Méndez Giraldü & Lopez Santana, 2014).

En un estudio de Análisis Econométrico de la Producción en la Industria de Alimentos y Bebidas del Ecuador Durante el Periodo 2007 –2017 (Cacho & Espinoza-Layana, 2019), se presenta un modelo el cual presenta como variable

dependiente a la producción, y como variables independientes: capital, número de empleados, inventario inicial, materia prima, consumo de combustible, gastos de servicios básicos. (López Rodríguez & Zapata Zuluaga, 2018), (Montes Albarracin, 2016).

El desarrollo de modelos matemáticos en la actualidad es una de las herramientas más utilizadas en el estudio de problemas; sus objetivos primordiales son describir, explicar y predecir las variaciones en los procesos productivos (Montesinos-López & Hernández-Suárez, 2007); es por ello que podemos decir que modelar permite analizar el comportamiento de numerosos fenómenos en forma aproximada (ARAVENA, CAAMANO, & GIMENEZ, 2008). Sin embargo, su aplicación se ve limitada con frecuencia por la falta de conocimientos e información acerca de los principios básicos del modelamiento matemático (Rodríguez Velázquez & Steegmann Pascual, Modelos Matemáticos, 2013).

El principal problema que enfrenta una empresa al pronosticar la demanda, es la baja eficacia que puede tener el pronóstico. 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 (Zafra Mejía & Gutiérrez Gil, 2015).

Por lo tanto, al momento de pronosticar se debe de seguir siete pasos: determinar el uso del pronóstico, seleccionar los aspectos que se deben pronosticar, determinar el horizonte del pronóstico, seleccionar los modelos de pronóstico, reunir los datos necesarios para elaborar el pronóstico, obtener el pronóstico, validar e implantar los resultados (Render & Heizer, 2007).

Todo esto debe de motivar a la empresa a poner en práctica el estudio de un modelo matemático que permita garantizar un mejor uso o aprovechamiento de las materias primas e insumos de la empresa para la elaboración de su producto, con propósito tender a la reducción al mínimo de los desperdicios. (Villada, Muñoz, & García, Aplicación de las Redes Neuronales al Pronóstico de Precios en el Mercado de Valores, 2012), (Villada, Muñoz, & García-Quintero, Redes Neuronales Artificiales aplicadas a la Predicción del Precio del Oro, 2016).

Por ello nace la pregunta, ¿se puede desarrollar un modelo matemático que permita pronosticar la producción de una fábrica de galleta Wafer de presentación 40\*204 g utilizando regresión multivariada?

## **1.2 Delimitación del problema**

### **Espacio**

El presente trabajo de investigación se realiza en Ecuador en la Región Costa, Provincia del Guayas, Ciudad de Guayaquil, en una pequeña empresa.

### **Tiempo**

La información obtenida está dentro del rango previsto de validez de una investigación científica, de 10 años.

### **Universo**

El estudio se basará en la propuesta de un Modelo Matemático para la producción, para lo cual se usarán los datos históricos de producción de los últimos 2 años, 2020 – 2021.

## **1.3 Formulación del problema**

¿De qué manera el uso de la regresión multivariada ayudará en el desarrollo de un modelo matemático que permita pronosticar de manera ágil la producción en una fábrica de galleta Wafer de presentación 40\*204 g?

## **1.4 Sistematización del problema**

¿Mediante el análisis de los componentes de un modelo matemático se podrá identificar los factores que inciden en la producción de galleta Wafer de presentación 40\*204 g?

¿Será posible realizar un análisis de la variabilidad de cada uno de los factores que inciden en la producción de galleta Wafer de presentación 40\*204 g?

¿Será posible construir un modelo matemático apropiado que permita pronosticar de manera ágil la producción aplicándolo a una fábrica de galleta Wafer de presentación 40\*204 g utilizando Regresión Multivariada?

## **1.5 Objetivos**

### **1.5.1 Objetivo General**

Proponer un modelo matemático apropiado que permita poder pronosticar de manera ágil la producción en una fábrica de galleta Wafer de presentación 40\*204 g utilizando Regresión Multivariada

### **1.5.2 Objetivos Específicos**

- Identificar los factores que inciden en la producción de galleta Wafer de presentación 40\*204 g.
- Realizar un análisis de la variabilidad de cada uno de los factores que inciden en la producción de galleta Wafer de presentación 40\*204 g.
- Construir un modelo matemático apropiado que permita pronosticar de manera ágil la producción aplicándolo a una fábrica de galleta Wafer de presentación 40\*204 g utilizando Regresión Multivariada.

## **1.6 Justificación e importancia**

Estudios realizados determinan que para poder ajustar mejor los modelos de regresión multivariados aplicados al pronóstico de producción debe existir una correcta identificación y relación entre las variables independientes y las variables dependientes (FAO, 2021)

De estas comparaciones se obtiene el error del pronóstico de producción que resulta de la diferencia entre lo real y lo dado por el modelo matemático, estos errores son inevitables y siempre estarán presentes dado que el modelo es una representación de la problemática; lo importante es que este error sea siempre

lo más pequeño posible, dado que esto es una medida de que tan bien aproxima el modelo a la realidad (Villada, Muñoz, & García, Aplicación de las Redes Neuronales al Pronóstico de Precios en el Mercado de Valores, 2012).

Consecuentemente, esta investigación tiene como finalidad aportar con un modelo matemático que ayude a la empresa en el proceso de pronosticar de forma más precisa, reduciendo el error. (Nations, 2020)

La presente investigación presenta una estrategia que permitirá la construcción de un modelo aplicado a la realidad de la microempresa, sin volverse una camisa de fuerza, sino que sea flexible y permita su actualización según los cambios que se producen en la demanda del mercado.

#### **1.6.1 Justificación Teórica**

El presente trabajo de investigación es un análisis de la incidencia de los factores que están ligados al pronóstico de producción en una microempresa, que permita a través de un modelo de regresión multivariado estudiar la relación entre la variable dependiente y las variables independientes o explicativas, logrando así determinar los componentes más influyentes en la variación del proceso de producción (Pineda Jaimes, Bosque Sendra, Gómez Delgado, & Plata Rocha, 2009).

Por lo tanto, resulta necesario poder entender cómo se elabora y desarrollo un modelo matemático aplicado a la producción, y con el análisis de la big data de la empresa resulta importante entender el uso de la Regresión Multivariada

#### **1.6.2 Justificación metodológica**

El presente trabajo de investigación se basa en el análisis metódico de los distintos factores que pueden afectar la producción en una pequeña empresa, lo que ocasiona que esta no cumpla con la producción planificada.

Algunos de los factores o variables que pueden influir de manera directa o indirecta en el pronóstico de producción pueden ser: horas reales trabajadas (Hr), horas de paras no programadas (Hr), horas de paras programadas (Hr), horas hombres trabajadas (HH), retraining (kg).

En el presente estudio se planifica el poder identificar cuáles son las variables que en mayor o menor grado aportan en la elaboración del modelo matemático aplicado al pronóstico de producción de la pequeña empresa; con este objetivo se realizará un análisis de la información cuantitativa suministrada por la pequeña empresa.

Al utilizar la regresión multivariada será importante lograr determinar el grado de significancia de cada una de las variables independientes que participan en el modelo, y el valor de R cuadrado dado que este valor explica que tan bien el modelo de ajusta a los valores suministrado para su entrenamiento.

### **1.6.3 Justificación practica**

En este trabajo de investigación se justifica porque con el modelo matemático que se obtenga, permitirá a la empresa realizar un mejor pronóstico de producción, dado que se identifica de manera precisa cuales son las variables que inciden en el proceso de producción, lo cual conlleva a un ahorro económico para la pequeña empresa.

Dado el proceso de elaboración del modelo matemático, este podrá ser actualizado por el departamento de producción en base a las producciones reales que la pequeña empresa mes a mes realice.

Se espera que este modelo matemático sea mucho más apegado a la realidad de la producción en la pequeña empresa, y podría ser escalable a cualquier otro producto que la planta elabore; además que podrá ser utilizado por cualquier tipo de pequeña empresa.

## **1.7 Hipótesis**

### **1.7.1 Hipótesis General**

Se diseña un modelo matemático utilizando Regresión Multivariada, se logrará pronosticar de manera eficiente la producción de galletas Wafer de presentación 40\*204 g.

### **1.7.2 Hipótesis Particular**

Es posible poder identificar los factores que inciden en la producción de galleta Wafer de presentación 40\*204 g.

Es posible poder discriminar que factores influyen en la variabilidad de un modelo de pronósticos de producción de galleta Wafer de presentación 40\*204 g.

Es posible construir un modelo matemático apropiado que permita pronosticar de manera ágil la producción aplicándolo a una fábrica de galleta Wafer de presentación 40\*204 g utilizado Regresión Multivariada

### **1.7.3 Conceptualización y operacionalización de variables**

Variable independiente: Capacidad de Producción Programada CPP (Kg.), Horas Laborares Planificadas HLP (hr), Paradas No Programadas PNP.

Variable dependiente: pronóstico de producción

**Tabla 1**

***Operacionalización de variable***

<b>Variable</b>	<b>Definición conceptual</b>	<b>Indicadores</b>	<b>Técnica</b>
Capacidad de Producción Programada CPP(Kg.), Horas Laborares Planificadas HLP (hr), Paradas No Programadas PNP.	Relación entre varias variables independientes (predictoras o explicativas) y otra variable dependiente (criterio, explicada, respuesta)	Valor de los coeficientes de la regresión multivariada.	Prueba de ajuste
Pronóstico de Producción	El pronóstico es la cantidad de kg de galletas Wafer de presentación 40*204 g que debe de producir la empresa	Cantidad (el número total de los datos) Valor máximo (el mayor número de los datos). Valor mínimo (el menor número de los datos)	Registro de datos históricos

Elaborado por: El autor.

## **CAPÍTULO II**

### **2.1 Marco Teórico**

#### **2.1.1 Modelo matemático.**

La elaboración de modelos matemáticos es uno de los instrumentos usados para el estudio y análisis de inconvenientes en diversas áreas del entendimiento; el

propósito de los modelos matemáticos son poder explicar, describir y predecir fenómenos y procesos en las diferentes zonas del saber. (Montesinos-López & Hernández-Suárez, 2007)

Múltiples situaciones en todo el mundo tienen la posibilidad de exponer inconvenientes que necesitan resoluciones y elecciones. Estos inconvenientes poseen un aspecto matemático subjetivamente fácil, involucrando un modelo matemático elemental, otros inconvenientes enfocados en una cierta área del saber, requieren un estudio más preciso de las variables con las que se relacionan. (Salett Biembegut & Hein, 1999)

Un modelo matemático es un modelo simplificado de un escenario real, por medio de la utilización de funcionalidades que describen su comportamiento, o de ecuaciones que representan sus colaboraciones. Para lograr estudiarlo el modelo matemático comienza con la identificación de los puntos primordiales o determinantes del sistema y los caracteriza por medio de las expresiones matemáticas. (Bocco, 2010)

Un grupo de símbolos e interrelaciones matemáticas que traducen, de alguna forma, un fenómeno en cuestión o problema de situación real, lo llamamos modelo matemático. En la ciencia, la idea de modelo es primordial para la obra y expresión del entendimiento. (Salett Biembegut & Hein, 1999)

En esencia, la matemática posibilita la preparación de modelos que posibilitan una mejor comprensión, simulación y previsión del fenómeno estudiado. Un modelo podría ser formulado en términos cotidianos, como, por ejemplo: expresiones numéricas o fórmulas, diagramas, gráficos o representaciones geométricas, ecuaciones algebraicas, tablas, programas computacionales, etc. Por otro lado, una vez que se ofrece un modelo, este procede de aproximaciones llevadas a cabo para lograr comprender mejor un fenómeno, no obstante, no continuamente tales aproximaciones permanecen conforme con la realidad. (Salett Biembegut & Hein, 1999)

En el análisis de un sistema, un modelo matemático empieza con la identificación de los puntos primordiales o determinantes del sistema y los caracteriza por medio de las expresiones matemáticas. La iniciativa en la obra es descubrir un equilibrio entre la simplicidad y una reproducción del comportamiento que posibilite entender, examinar y presagiar, al modificar el costo de la o las cambiantes que lo describen, la contestación del sistema en su grupo. (Bocco, 2010)

El triunfo o fracaso de los modelos matemáticos es un reflejo de la exactitud con que dicho modelo matemático representa al objeto inicial y no de la precisión con que las matemáticas analizan el modelo (Rodríguez Velázquez & Steegmann Pascual, Modelos Matemáticos, 2012).

Las diferentes versiones tecnológicas de la información, y la dinámica cambiante de los negocios obliga, que los modelos de inventarios deban ser todos de revisión continua, lo cual dio como sitio que el cálculo de periodos óptimos de revisión se vuelva innecesario. Los tiempos de abastecimiento tienden por un lado a reducirse debido a las tecnologías de transporte, en lo que, sin embargo, se dificultan gracias a la globalización y los tratados de libre comercio. (Arango Marin, Giraldo Garcia, & Castrillón Gómez, 2013)

### **2.1.2 Pronóstico de producción.**

Los pronósticos se ajustan a un componente fundamental en los procesos de planificación de las empresas, en particular una vez que estas se combaten a ámbitos dinámicos que pide de sus directivos capacidad para visualizar el futuro e incorporarlo a sus proyectos y planes. (Rubio Guerrero, 2017)

El pronóstico de comercialización actualmente se convirtió en una fuente primordial de información para poder predecir la demanda de la manera más apegada a la realidad del mercado. La aleatoriedad propia de la mayor parte de los mercados puede reproducirse por modelos probabilísticos que poseen superiores modalidades de utilización informática de antaño. La integralidad de las resoluciones que tengan en cuenta el problema a partir de la previsión de la

demanda hasta el cálculo de las porciones a solicitar es un imperativo para diseñar modelos confiables y eficientes (Sánchez López, Barreras Serrano, Pérez Linares, Figueroa Saavedra, & Olivas Valdez, 2013)

Contar con un pronóstico es imprescindible para establecer cuánta capacidad o abasto va a ser primario para consumir con la demanda. Los pronósticos son primordiales en la idealización de los negocios, las necesidades de materias primas, las ventas de productos, el capital humano y las necesidades de capacidad que tienen la posibilidad de dañar el futuro de la organización. (Rubio Guerrero, 2017)

Los pronósticos se usan en las empresas para dictaminar si hay suficiente demanda para ingresar al mercado, decidir la capacidad esencial para diseñar instalaciones, identificar fluctuaciones de mediano y corto plazo en las necesidades con el objeto de conservar la programación de producción, programación de mano de obra, planeación de materiales y otras necesidades. (Rubio Guerrero, 2017)

Los procedimientos que se usan para adivinar la demanda, están sujetas a una secuencia de componentes como, por ejemplo: la viable realidad de patrones de tendencia y estacionalidad, sin embargo, más que nada la conducta de la demanda y las razones de ésta. En los procedimientos que se usan para adivinar se poseen las series de tiempo, las regresiones lineal y múltiple, y los procedimientos cualitativos. (Saucedo Castillo, Pérez, Herrera Isla, & Fernández Pérez, 2010)

Las series de tiempo y los procedimientos de regresión son técnicas estadísticas o procedimientos cuantitativos que usan los datos de la demanda histórica para adivinar la demanda futura. Los procedimientos cualitativos usan los juicios de profesionales para la preparación de pronósticos. Hay diversos componentes a considerar al instante de escoger un modelo de pronóstico apropiado a utilizar

en cada situación, se debería de tener en cuenta las diferencias existentes en los comportamientos de la demanda y de la dispensación en todos los puntos de vista de repartición (Saucedo Castillo, Pérez, Herrera Isla, & Fernández Pérez, 2010)

Los componentes que están afectando la demanda cambian una y otra vez; por consiguiente, mientras pasa la época, la exactitud del pronóstico puede reducir. Por esta razón, los pronósticos de demanda tienen que actualizarse periódicamente para conservar su validez. Luego de cada lapso de comercialización, se tienen que comprobar los pronósticos. La obra de un sistema de pronóstico necesita por lo menos ciertos datos iniciales, una técnica de pronóstico y los procedimientos de revisión y seguimiento necesarios para actualizar el modelo. Se han desarrollado numerosas técnicas de pronóstico que tienen la posibilidad de clasificarse primordialmente en 2 categorías: procedimientos cuantitativos y procedimientos cualitativos. (Méndez Giraldu & Lopez Santana, 2014)

Las técnicas cualitativas son de relevancia una vez que no se dispone de datos, en especial en la situación de lanzamientos de productos. Las técnicas cuantitativas se aplican una vez que se dispone de datos históricos y una vez que se necesitan proyecciones futuras a corto o medio plazo. Los procedimientos de pronóstico tienen la posibilidad de ser causales, debido a que integran componentes que tienen la posibilidad de influir la proporción de pronóstico, como la temperatura, la humedad, la estación, etc. y cuyas necesidades tienen la posibilidad de exponer tendencias, estacionalidad y comportamiento repetitivo o aleatorio. (Arroyo García, Sánchez Partida, Juárez García, & Martínez Flores, 2017)

### **2.1.3 Regresión Multivariada**

La investigación de regresión lineal múltiple es una técnica estadística que se usa para aprender la interacción entre cambiantes en varias situaciones y

presagiar diferentes fenómenos. La capacidad de regresión más fácil es la funcionalidad lineal, en la que cada variable participa de manera aditiva y sin cambios para todo el fenómeno observado. (Vilà Baños, Torrado Fonseca, & Reguant Alvarez, 2019)

Con frecuencia se dice que el valor de una variable puede retroceder en relación con el valor de la otra si existe una línea de regresión que coincide exactamente con el valor observado. La regresión se utiliza para determinar posibles relaciones causales o, si existe alguna duda sobre ellas, para predecir una variable relacionada con otra. Si dos variables tienen una relación definida, entonces el valor de una determina exactamente la otra. (Dagnino S., 2014)

En la regresión lineal múltiple se usa por lo general más de una variable explicativa o libre; esto muestra la virtud de usar más información en la elaboración del modelo y, consecuentemente, hacer estimaciones más exactas. La regresión lineal múltiple consiste en ajustar modelos lineales o linealizables entre una variable dependiente y bastante más de unas variables independientes. En esta clase de modelos es fundamental examinar la heterocedasticidad, la multicolinealidad y la explicación. (Montero, 2016)

En un modelo de regresión lineal múltiple, la variable dependiente es una función lineal de  $k$  elementos correspondientes a las variables independientes y una perturbación aleatoria o error. El modelo también incluye un término independiente. Si expresamos por  $y$  a la variable dependiente o variable de respuesta, por  $x_1, x_2, x_3, x_4, \dots, x_k$  a las variables independientes o explicativas; por  $\varepsilon$  al error o perturbación aleatoria, los modelos de regresión multivariada o regresión lineal múltiple viene dado por la siguiente expresión (Montero, 2016):

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

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

$\beta_0$  = intersección en  $Y$

$\beta_1$  = pendiente de  $Y$  con la variable  $X_1$  manteniendo constante las variables  $X_2, X_3, \dots, X_k$

$\beta_2$  = pendiente de  $Y$  con la variable  $X_2$  manteniendo constante las variables  $X_1, X_3, \dots, X_k$

$\beta_3$  = pendiente de  $Y$  con la variable  $X_3$  manteniendo constante las variables  $X_1, X_2, X_4, \dots, X_k$

.

.

.

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

$\varepsilon_i$  = es el termino de error aleatorio en  $Y$  para la observación en  $i$ , que tiene la media 0 y la varianza  $\sigma^2$

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  $\beta_j$  utilizando el método de mínimos cuadrados

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

## Supuestos habituales de la regresión multivariada

El modelo de regresión poblacional múltiple es:

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

y suponemos que hay  $n$  conjuntos de observables disponibles. Se hacen las siguientes suposiciones comunes para el modelo.

1. Las  $x_{ji}$  son bien números fijo, o bien relaciones de variables aleatorias,  $X_j$ , que son independientes de los términos de error,  $\varepsilon_i$ . 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  $\sigma^2$ . 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  $\varepsilon_i$ , no están correlacionados entre sí, por lo que

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

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

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

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

## **Matriz de Correlación. –**

La matriz de correlación muestra los valores de correlación de Pearson, que miden el grado de relación lineal entre cada par de elementos o variables. Los valores de correlación se pueden ubicar entre -1 y +1. Sin embargo, en la práctica, los elementos por lo general tienen correlaciones positivas. Si los dos elementos tienden a aumentar o disminuir al mismo tiempo, el valor de correlación es positivo.

Se utiliza la matriz de correlación para evaluar la fuerza y dirección de la relación entre dos elementos o variables. Un valor de correlación alto y positivo indica que los elementos miden la misma destreza o característica. Si los elementos no están altamente correlacionados, entonces los elementos pudieran medir diferentes características o no estar claramente definidos.

Las variables con valores de correlación mayores que 0.7 se pueden considerar altamente correlacionadas.

## **Método de mínimos cuadrados**

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

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

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

En términos formales, minimizamos  $SCE$ :

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

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

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

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

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

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

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

La variabilidad del modelo puede dividirse en los componentes

$$STC = SCR + SCE$$

donde estos componentes se definen de la siguiente forma

Suma total de los cuadrados

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

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

Suma de los cuadrados de los errores

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

Suma de los cuadrados de la regresión

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

Esta descomposición puede interpretarse de la forma siguiente:

Variabilidad muestral total = variabilidad explicada + variabilidad no explicada

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

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

y se deduce que

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

Los coeficientes se eligen de modo que se minimice la suma de cuadrados entre los valores observados y predichos, es decir, se minimice la varianza residual.

### **Coefficiente de determinación ajustado.**

Algunos académicos sugieren que, cuando se trata de modelos de regresión multivariable, se debe usar el  **$r^2$  ajustada**, para reflejar el número de variables independientes en el modelo y el tamaño de la muestra. Es muy importante informar la  **$r^2$  ajustada** cuando se comparan dos o más modelos de regresión que predicen la misma variable dependiente, pero con un número diferente de variables independientes.

### **$r^2$ ajustada**

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

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

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

Utilizamos 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.

### **Coefficiente de correlación múltiple.**

El coeficiente de correlación múltiple es la 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 relación entre las variables dependientes e independientes.

Por lo tanto, es comparable a la correlación entre  $Y$  y  $X$  en la regresión lineal simple.

## 2.2 Marco Conceptual

Los siguientes conceptos han sido tomados de los libros: “Funciones elementales para construir modelos matemáticos” (Bocco, 2010), “Probabilidad y estadística para ingeniería y ciencias” **Fuente especificada no válida.**

- **Análisis exploratorio de datos:** El análisis exploratorio de datos (EDA por sus siglas en inglés) implica el uso de gráficos y visualizaciones para explorar y analizar un conjunto de datos.
- **Análisis multivariante:** El análisis multivariante permite predecir para una observación determinada en base a su patrón de covariables, el valor de una variable continua o la probabilidad de ocurrencia de una variable dicotómica.
- **Heterocedasticidad:** Se produce cuando la varianza (del error) no es constante en las observaciones llevadas a cabo.
- **Media aritmética:** La media aritmética es un tipo de media que otorga la misma ponderación a todos los valores.
- **Medidas de dispersión:** Las medidas de dispersión tratan, a través del cálculo de diferentes fórmulas, de arrojar un valor numérico que ofrezca información sobre el grado de variabilidad de una variable.
- **Medidas de tendencia central:** Las medidas de tendencia central son medidas estadísticas que pretenden resumir en un solo valor a un conjunto de valores. Representan un centro en torno al cual se encuentra ubicado el conjunto de los datos
- **Modelo linealizable:** Un modelo de relación entre ambas variables ( $x$ ,  $y$ ) puede ser linealizable si una de las dos variables logra convertirse en relación lineal.

- **Modelo matemático:** Es una simple representación de la realidad, mediante el uso de funciones que describen su comportamiento o ecuaciones que expresan sus relaciones.
- **Multicolinealidad:** Es la relación de carácter dependiente fuertemente lineal entre más de 2 variables.
- **Pronóstico:** El pronóstico es una estimación del comportamiento de variables estadísticas en eventos futuros.
- **Regresión lineal múltiple:** El modelo de regresión lineal múltiple se usa como explicación hacia las variables de respuesta continua relacionadas con varios factores o variables explicativas continuas.
- **Regresión lineal:** Es una técnica de modelado estadístico que se utiliza para describir variables de respuesta continua en función de uno o más predictores.
- **Representaciones Gráficas:** Una representación gráfica o grafico estadístico es un resumen visual de la tabla de frecuencias, que nos permite la representación en forma más resumida y total del experimento o fenómeno en estudio.
- **Variable Continua:** Una variable continua es aquella que puede tomar un número infinito de valores entre dos valores cualesquiera de una característica.
- **Variable Cuantitativa:** Una variable cuantitativa es aquella variable estadística que, a diferencia de la cualitativa, puede expresarse a través de cifras. Por esta razón, puede analizarse con métodos estadísticos.
- **Variable dependiente:** variable objeto del estudio y que sus resultados procuran explicar por medio de las variables denominadas independientes.
- **Variable Discreta:** Una variable es discreta cuando no puede tomar ningún valor entre dos consecutivos.
- **Variable Nominal:** Una variable es nominal cuando sus valores representan categorías que no obedecen a una clasificación intrínseca
- **Variable Ordinal:** Una variable es ordinal cuando sus valores se pueden clasificar en categorías y estas se pueden ordenar en jerarquías con respecto a la característica que se evalúa

- **Varianza:** La varianza es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media.

## 2.4 Marco Referencial.

Los estudios que se mencionan a continuación han servido como referente para este trabajo de investigación, de una forma u otra, abordan el tema de la investigación y hacen interesantes aportaciones en cuanto a la metodología como en el estudio.

Zambrano (2016) en su trabajo “Estimación de la función oferta y demanda para el mercado de la harina de trigo en el cantón Machala provincia de el Oro en el año 2013” establece el comportamiento del mercado de Harina de Trigo de la Ciudad de Machala, estimando para ello la oferta y demanda de dicho producto, utilizando como técnica la regresión lineal múltiple, que es un método estadístico el cual permite conocer las variables que influyen sobre dichas curvas y la sensibilidad de las mismas ante cambios en el precio. (Zambrano Macas, 2016)

Manrique (2007) en su trabajo” Diseño para el sistema estadístico forestal, basado en componentes de software libre” analiza un software estadístico denominado R-Commander, para la validación del paquete estadístico se utilizan datos del actual Sistema Nacional de Información Estadísticas Forestales del país donde se aplicaron diferentes métodos estadísticos como la estadística descriptiva, prueba de hipótesis y análisis de regresión lineal simple y múltiple. (Manrique Leonett, 2007)

Rivera (2011) en su trabajo denominado “Impacto de las variables meteorológicas en la formación de ozono en la cuenca de Santiago” presenta el grado de importancia de las variables meteorológicas respecto al contenido de ozono troposférico en la cuenca de Santiago, considerando el periodo estival de los años 1998 al 2006, en el marco de su formación química. El estudio contempló un análisis del ozono respecto a las variables meteorológicas por medio de tres técnicas estadísticas: método de regresión lineal simple, regresión

lineal múltiple y componentes principales. Además, se realizó un análisis del comportamiento atmosférico de escala sinóptica en días de altas y bajas concentraciones de ozono. (Rivera Garcés, 2011)

López (2013) en su trabajo de investigación “Factores que determinan el consumo nacional aparente de Cártamo en México, 1995-2009” analiza la información relacionada con la situación de la producción de cártamo a nivel internacional, contextualizado en la producción de cártamo en México, aportando gran información en la creación de condiciones para realizar investigación de vanguardia, generando información verdaderamente importante para satisfacer la demanda nacional y las preferencias del consumidor y de la industria, para mejorar la situación agrícola y proporcionando elementos que ayudarán a reformar las políticas agrícolas en el país. El trabajo de investigación determino coeficientes de correlación de variables en relación con el cultivo de cártamo que indiquen en el CNA del cártamo, mediante el modelo de regresión lineal múltiple. (López López, 2013)

Vargas & Rojas (2017) en su trabajo “Análisis de regresión de la producción lechera en la empresa CAMAY, periodo 2000-2016” desarrolla contenidos matemáticos del análisis y métodos numéricos y se enfoca en mostrar su aplicación económica en la toma de decisiones gerenciales en la empresa, mediante explicaciones teóricas. El trabajo se ve reforzado con la implementación computacional del software SPSS 22,0. Asimismo se presentan los contenidos teóricos y prácticos del análisis de regresión multivariada. (Vargas Orbegoso & Rojas Primo, 2017)

## **CAPÍTULO III**

### **MARCO METODOLÓGICO**

#### **3.1 Tipos de Investigación**

### **Investigación Exploratoria. -**

La presente investigación es exploratoria, ya que se acerca a un tema poco estudiado, por lo que mediante la recopilación de datos es posible descubrir el patrón que sigue, y así brindar explicaciones teóricas y prácticas para el desarrollo de modelos matemáticos que lograrán los objetivos de este estudio.

### **Investigación Documental. -**

Este estudio es documental porque ha sido compilado utilizando una variedad de fuentes como libros, artículos, artículos de investigación y otros documentos que respaldan los diversos temas tratados en la investigación.

### **Investigación de Campo. -**

Se utilizará trabajo de campo para desarrollar la investigación, ya que la investigación debe realizarse directamente en el sitio del evento (microempresa) para conocer el estado del entorno problema, el investigador participará directamente en el estudio y estudiará cómo se encuentra cada variable de investigación desarrollado para poder evaluarlo, interpretarlo y a su vez consultar con un grupo de personas y fuentes de la microempresa.

### **Investigación Aplicada. -**

La aplicabilidad de este estudio se basa en que pretende generar conocimiento que sea directamente aplicable a los problemas del sector manufacturero. Se espera que se desarrolle un modelo matemático para ayudar a mejorar la previsión de producción. Es posible hacer un pronóstico de ganancias basado en la influencia de varias variables independientes.

### **Investigación Cuantitativa. -**

La encuesta es cuantitativa porque se recopilan y analizan datos digitales, es decir, datos de producción semanales. Con este método, se identifican tendencias o comportamientos de producción para que se puedan hacer predicciones, se verifiquen las relaciones y se puedan obtener resultados generales de producción. Con lo que se espera en que contribuya a elevar el nivel de producción, permitiendo a los microempresarios tomar mejores decisiones.

Propuesta de un modelo matemático aplicado al pronóstico de producción utilizando regresión multivariado aplicado a una fábrica de galleta wafer 40\*204 g.

### **3.2 Diseño de la Investigación**

El objetivo de este estudio es proponer un modelo matemático apropiado que permita el uso de la regresión múltiple para predecir dinámicamente la producción de una fábrica de galleta Wafer presentación de 40\*204 gramos.

El presente diseño de estudio no es empírico ya que los fenómenos serán observados y analizados sin manipular las variables.

Este trabajo es de largo plazo porque analizamos el comportamiento en diferentes momentos, para diferentes variables independientes sugeridas por un modelo matemático de intervención.

### **3.3 Población y muestra**

Para este estudio utilizaremos todos los datos, es decir, el 100 % de los datos recopilados proviene de los últimos 2 años de una microempresa que produce galleta Wafer presentación de 40\*204 gramos; Los datos recopilados son para la producción referente a 38 semanas, cabe señalar que debido al mercado la demanda es baja y la micro empresa lleva varias semanas sin producir el producto.

### **3.4 Variables de investigación**

#### **3.4.1 Variable dependiente:**

La variable dependiente a tener en cuenta es la Cantidad Producida Real CPR (kg) que se deriva de la información proporcionada por la empresa y corresponde a los kilogramos reales de galleta Wafer presentación de 40\*204 gramos producida por la empresa.

**Tabla 2**

***Variable dependiente***

<b>Variable dependiente</b>	<b>Unidad de medida</b>
Cantidad de Producción Real	Kilogramos

Elaborado por: El autor.

**3.4.2 Variable independiente:**

Como primera aproximación se analizarán 4 posibles variables independientes, y analizando las variables excluidas se espera que el grado, peso o contribución de cada variable al aprendizaje del modelo matemático sea tal que la variable tenga poca o ninguna contribución al modelo, no se consideran ni descartan entradas.

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

La cantidad de producción programada es la cantidad expresada en kilogramos de galleta Wafer presentación de 40\*204 gramos que la microempresa quiere o pretende producir, es decir, teóricamente esperada o planificada.

**Tabla 3**

***Variable independiente***

<b>Variable independiente</b>	<b>Unidad de medida</b>
Cantidad de Producción Programada	kilogramos

Elaborado por: El autor.

**b) Horas Laborales Planificadas HLP (hr).**

Las horas de trabajo planificadas corresponden a las horas planificadas o teóricas que las personas deben trabajar cada semana para igualar la producción y se espera que se puede hacer independientemente de los imprevistos para lograr la meta. Las empresas por lo general lo llaman Horas Hombre.

**Tabla 4**

***Variable independiente***

Variable independiente	Unidad de medida
Horas Laborales Planificadas	Hora

Elaborado por: El autor.

**c) Paradas No Programadas PNP (hr).**

El tiempo de inactividad no planificado es el tiempo perdido por una máquina que no funciona debido a una avería, falla o daño, es decir, la máquina se repara y sale de la línea de producción. Este tiempo afecta a la producción por lo que cuanto menor sea el valor, mejor.

**Tabla 5**

***Variable independiente***

Variable independiente	Unidad de medida
Paradas No Programadas	Hora

Elaborado por: El autor.

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

El método empleado es cuantitativo, la técnica de análisis es regresión múltiple o regresión lineal muchas veces, se evalúan los supuestos básicos del modelo de regresión y se analizan las variables excluidas para ver el grado de cercanía que aportan al modelo, es decir, que tengan alguna independencia, variables a excluir mediante coeficiente de determinación ( $R^2$ ) y análisis de varianza (Anova) para analizar el ajuste, y finalmente obtener coeficientes de regresión múltiple y su significancia para las variables anteriores.

La información será ejecutada por el software estadístico IBM SPSS versión 24 para el procesamiento y análisis de datos.

## CAPITULO IV

### PROPUESTA

#### 4.1 Formulación del modelo de regresión lineal múltiple

Al momento de realizar la formulación del modelo matemático utilizando la regresión multivariada, primero seleccionamos la variable dependiente (Y) y las variables independientes (X), luego, se muestra la selección de las variables y la ecuación.

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

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

Donde:

$Y$  = CPR (Cantidad de producción real)

$X_1$  = CPP (Cantidad de producción programada)

$X_2$  = HLP (Horas laborables planificadas)

$X_3$  = PNP (Paradas no programadas)

#### 4.2 Diagnósis y validación del modelo de regresión lineal múltiple.

Para interpretar correctamente un modelo de regresión debemos acompañar siempre nuestro estudio de la diagnósis y validación del modelo. La diagnósis consiste en analizar si se verifican las hipótesis básicas del modelo:

- Linealidad: Los parámetros y su interpretación carecen de sentido si en realidad los datos no proceden de un modelo lineal, caso en la que además las predicciones pueden ser completamente equivocadas.

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

#### 4.3.- Supuesto de Normalidad

El supuesto de Normalidad tiene como objetivo demostrar que tanto las variables independientes, así como de la variable dependiente, tienen que estar distribuidos normalmente.

##### Normalidad de los datos

Como la muestra es menor de 50 datos entonces debemos usar la prueba Shapiro-Wilk SW, y esta prueba nos indica que la significancia tiene que ser mayor a 0.05 para decir que los datos corresponden a una muestra con distribución normal.

**Tabla 6**

##### *Pruebas de normalidad*

	Kolmogórov-Smirnov			Shapiro-Wilk		
	Estadístico	GI	Sig.	Estadístico	GI	Sig.
CPP	0,151	38	0,029	0,956	38	0,142
HLP	0,147	38	0,037	0,957	38	0,149
PNP	0,122	38	0,166	0,959	38	0,172
CPR	0,152	38	0,027	0,949	38	0,081

a. Corrección de significación de Lilliefors

Elaborado por: El autor.

Podemos observar en la tabla 6, con respecto a los estadísticos Kolmogórov-Smirnov y Shapiro-Wilk, para todas las variables se tiene un valor de p (sig) >

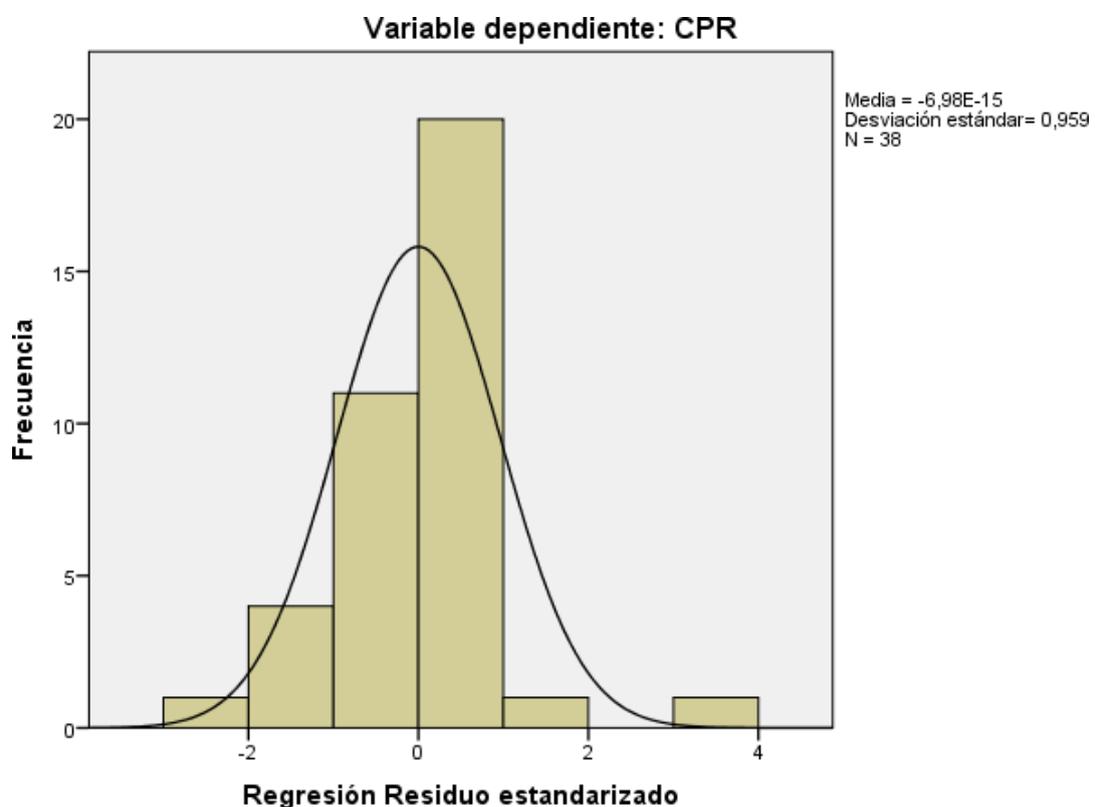
0,05; por lo que queda demostrado que se cumple con el supuesto de normalidad de los datos

### Normalidad de los Residuos

Para demostrar la normalidad de los residuos se usará la gráfica de histograma de Regresión de residuos Estandarizados y la gráfica de Gráfico P-P Normal de Regresión Residuos Estandarizados

**Figura 1**

### *Histograma de regresión de Residuos Estandarizados*

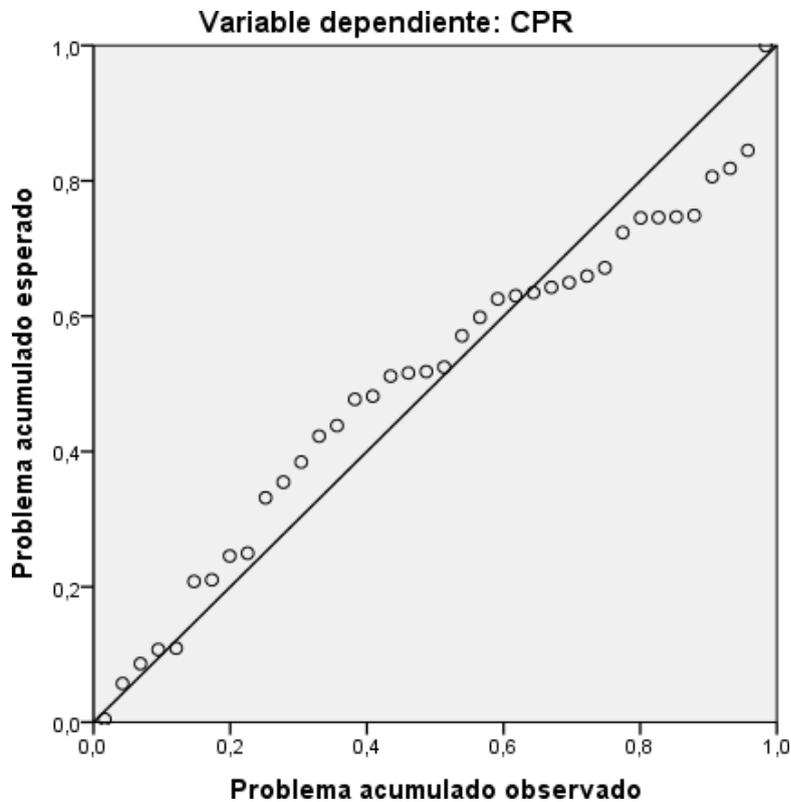


Elaborado por: El autor.

Como se puede observar figura 1, Histograma de regresión de Residuos Estandarizados, estos siguen una distribución normal con media cero, aunque en el extremo derecho existe un pequeño grupo de datos, por lo que se puede decir que se cumple con el supuesto que los residuos siguen una distribución normal.

**Figura 2**

**Gráfico P-P Normal de Regresión Residuos Estandarizados**



Elaborado por: El autor.

Como se puede observar figura 2, Gráfico P-P Normal de Regresión Residuos Estandarizados, los datos en su mayoría están por encima de la línea, y los datos que están por debajo de la línea son muy pocos y están apegados a esta, lo cual se observa también en la figura 1, con lo que se puede decir que estos siguen una distribución normal con media cero, por lo que se cumple con el supuesto de que los residuos siguen una distribución normal.

#### **4.4.- Supuesto de No Autocorrelación o de Independencia de las observaciones**

Para responder al supuesto de la independencia de errores se utiliza la Prueba de Durbin-Watson. El criterio para decir que existe independencia de las

observaciones, es que el valor de Durbin-Watson debe ser lo más cercana posible a 2 con una oscilación de  $\pm 1$ . Por lo tanto, los valores entre 1 y 3 están bien, para aceptar que los residuos son independientes

**Tabla 7**

**Resumen del modelo<sup>b</sup>**

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin-Watson
1	0,998 <sup>a</sup>	0,996	0,996	46,92367	1,507

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

b. Variable dependiente: CPR

Elaborado por: El autor.

**Durbin-Watson**

En la Tabla 7 se puede observar que el coeficiente de Durbin-Watson tiene un valor de 1,507; que es un valor aceptable muy cercano a 2, por lo que se da por satisfecho el supuesto de independencia de las observaciones

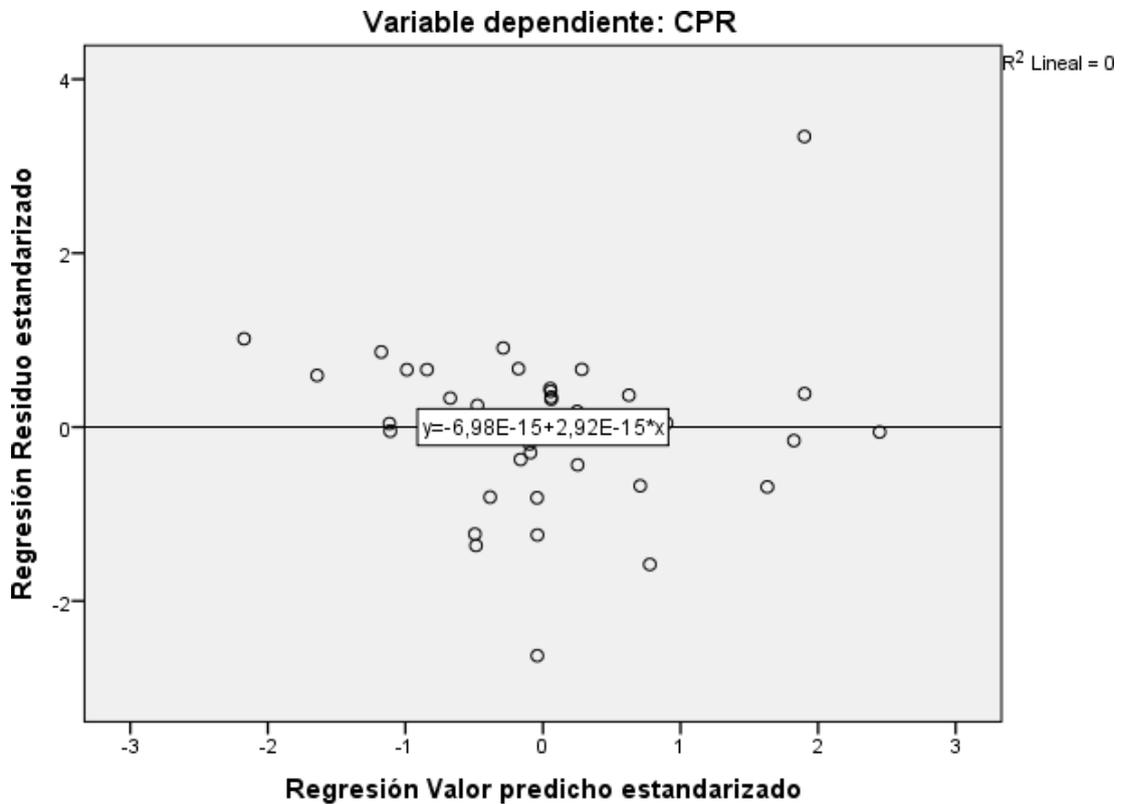
**4.5.- Supuesto de Homocedasticidad**

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

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

**Figura 3**

**Gráfico de dispersión**



Elaborado por: El autor.

A partir del examen de gráfico de dispersión de la fig. 3, se comprueba que los puntos están distribuidos razonablemente bien por lo que se puede decir que no hay ninguna relación sistemática entre los residuos tipificados y los valores pronosticados tipificados de la cantidad de producto a producir.

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

**4.6.- Supuesto de linealidad**

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

**Tabla 8****Correlaciones de variables**

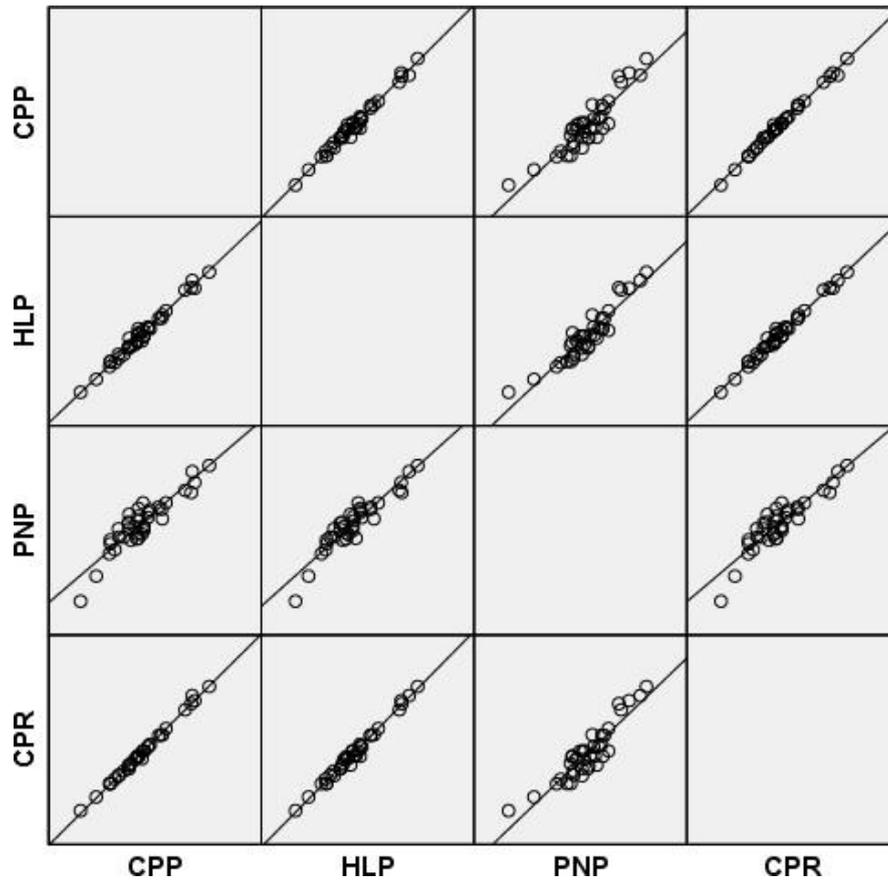
		<b>CPP</b>	<b>HLP</b>	<b>PNP</b>	<b>CPR</b>
CPP	Correlación de Pearson	1	0,991**	0,916**	0,998**
	Sig. (bilateral)		0,000	0,000	0,000
	N	38	38	38	38
HLP	Correlación de Pearson	0,991**	1	0,936**	0,993**
	Sig. (bilateral)	0,000		0,000	0,000
	N	38	38	38	38
PNP	Correlación de Pearson	0,916**	0,936**	1	0,916**
	Sig. (bilateral)	0,000	0,000		0,000
	N	38	38	38	38
CPR	Correlación de Pearson	0,998**	0,993**	0,916**	1
	Sig. (bilateral)	0,000	0,000	0,000	
	N	38	38	38	38

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

Elaborado por: El autor.

Como se puede observar en la tabla de correlación, existe una correlación positiva y fuerte entre las variables independientes y la variable dependiente. En la matriz se observa que todos los valores de correlación de las variables independiente son mayores a 0,91; por lo que se puede decir que se cumple con el supuesto de linealidad

**Figura 4****Correlación de variables**



Elaborado por: El autor.

En la figura 4 se puede observar de manera gráfica que existe una relación lineal entre las variables independientes y la variable dependiente; por lo que se puede decir que se cumple con el supuesto de linealidad

#### 4.7.- Supuesto de Multicolinealidad

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

Como se puede observar en la tabla 9, para el presente modelo de regresión multivariada se tiene que ningún valor del VIF es mayor a 10

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

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

**Tabla 9**

***Diagnósticos de colinealidad<sup>a</sup>***

Modelo	Dimensión	Autovalor	Índice de condición	Proporciones de varianza			
				(Constante)	CPP	HLP	PNP
1	1	3,877	1,000	0,00	0,00	0,00	0,00
	2	0,112	5,877	0,41	0,00	0,00	0,00
	3	0,010	20,121	0,52	0,05	0,01	0,80
	4	0,001	57,779	0,07	0,95	0,99	0,20

a. Variable dependiente: CPR

Elaborado por: El autor.

Como se puede observar en la tabla 9, existe un valor arriba de 30 que es de la variable PNP, además se observa que existe 2 proporciones de varianza por encima de 90%, por lo que se puede afirmar que existe colinealidad con la variable PNP; esto puede ocurrir debido a que las paradas no programadas, no depende de la producción, si no que se debe a la falta de mantenimiento en los equipos, por lo que resulta casi imposible poder predecir cuándo podrá ocurrir una para.

**4.8 El Modelo Matemático de Regresión Multivariada.**

Como se observa en la tabla 7, el  $R^2$  es de 0,996, lo que significa que el modelo explica un 99,6% de la varianza, que una vez corregido por el efecto de la muestra y de las variables independientes resulta ser el  $R^2$  ajustado es de 99,6%. Por otro lado, el error típico de la estimación (raíz cuadrada de la varianza no explicada) resulta ser de 46,92.

En este modelo se trabajó con tres variables independientes: cantidad de producción planificada CPP, horas laborales planificadas HLP, paradas no programadas PNP

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

### **Análisis de la varianza.**

**Tabla 10**

**ANOVA<sup>a</sup>**

<b>Modelo</b>		<b>Suma de cuadrados</b>	<b>Gl</b>	<b>Media cuadrática</b>	<b>F</b>	<b>Sig.</b>
1	Regresión	20537911,030	3	6845970,343	3109,218	0,000 <sup>b</sup>
	Residuo	74862,234	34	2201,830		
	Total	20612773,260	37			

a. Variable dependiente: CPR

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

Elaborado por: El autor.

La tabla 10 es una prueba de análisis de varianza la cual muestra si la varianza explicada por la regresión es significativamente distinta a la varianza no explicada.

El modelo matemático utilizado presenta un valor del estadístico de prueba  $F=3109,218$  y tiene un P\_ valor igual a  $0 < 0.05$ , por lo que se rechaza la hipótesis nula y se concluye que la dependencia lineal es estadísticamente significativa por lo que el modelo es adecuado, en otras palabras, podemos afirmar que las variables independientes CPP, HLP, PNP, si predicen a la variable dependiente CPR.

**Tabla 11**  
**Coeficientes<sup>a</sup>**

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	T	Sig	Estadísticas de colinealidad	
	B	Error estándar	Beta			Tolerancia	VIF
1 (Constante)	-2,150	29,517		-,073	0,942		
1 CPP	0,758	0,082	0,750	9,269	0,000	0,016	61,306
1 HLP	9,802	3,177	0,283	3,085	0,004	0,013	79,027
1 PNP	-15,520	12,863	-0,036	-1,207	0,236	0,117	8,544

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

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

$$Y = -2,150 + 0,758 X_1 + 9,802 X_2 - 15,520 X_3$$

Donde:

$Y$  = CPR (La cantidad de producción real)

$X_1$  = CPP (La cantidad de producción planificada)

$X_2$  = HLP (Horas laborales planificadas)

$X_3$  = PNP (Paradas no programadas)

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

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

$$Z = 0,750 X_1 + 0,283 X_2 - 0,036 X_3$$

#### 4.9 Comparación de los datos reales con los datos pronosticados.

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

**Tabla 12**

***Estadísticas de muestras emparejadas***

		Media	N	Desviación estándar	Media de error estándar
Par 1	CPR_RM	1631,1351	38	744,90663	120,83981
	CPR	1631,4211	38	746,39264	121,08088

Elaborado por: El autor.

Como se observa en la tabla 12; la media de los datos obtenidos con el modelo matemáticos CPR\_RM es de 1631,135; y la media de los datos reales observados CPR es de 1631,421; estas medias son prácticamente las mismas o con muy poca diferencia, con lo cual al existir una diferencia entre ellas muy pequeñas se puede decir que el modelo matemático es bastante bueno.

**Tabla 13**

***Correlaciones de muestras emparejadas***

		N	Correlación	Sig.
Par 1	CPR_RM & CPR	38	,998	,000

Elaborado por: El autor.

Como se observa en la tabla 13; la correlación entre los datos construidos con el modelo matemático CPR\_RM y los datos reales observados CPR es de 0,998 que es un valor muy cercano a 1 y tiene un P\_ valor igual a  $0 < 0.05$ , por lo que podemos decir que si existe correlación entre las variables.

**Tabla 14**

**Prueba de muestras emparejadas**

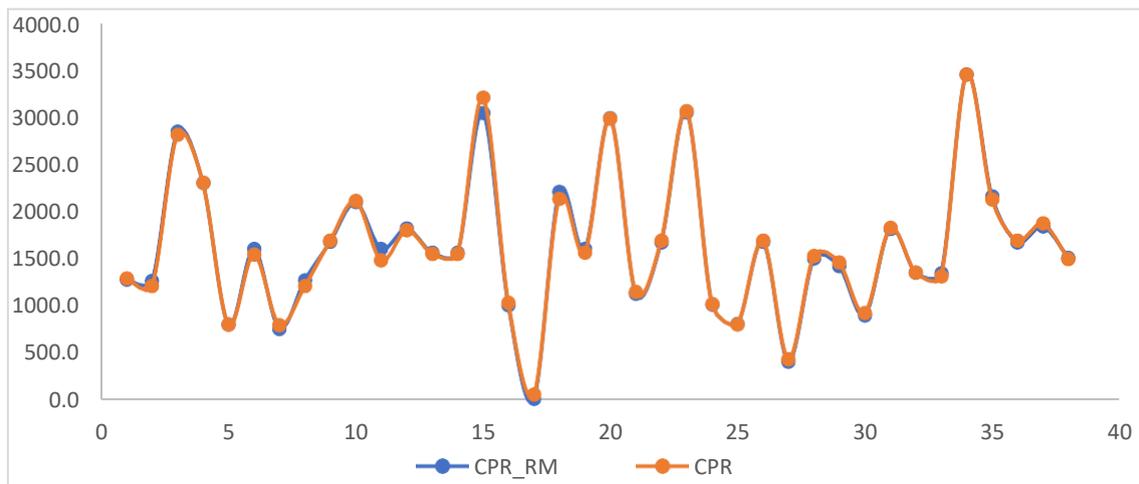
	Media	Desviación estándar	Diferencias emparejadas		T	gl	Sig. (bilateral)	
			Media de error estándar	95% de intervalo de confianza de la diferencia				
			Inferior	Superior				
Par 1 CPR_RM - CPR	-0,28591	44,98133	7,29694	-15,07090	14,49909	-0,039	37	0,969

Elaborado por: El autor.

Como se observa en la tabla 14; la diferencia emparejada entre las medias es muy baja, es de -0,28591; con lo cual se puede decir que es el error que se comete con el modelo matemático propuesto.

**Figura 5**

**Comparación entre CPR\_RM - CPR**



Elaborado por: El autor.

En la figura 5 se puede observar que, se realiza la comparación visual entre los valores reales obtenidos de la planta de producción es decir Capacidad de Producción Real CPR, con lo valores obtenidos mediante el uso del modelo matemático propuesto por medio de la regresión multivariado CPR\_RM.

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

## CONCLUSIONES

Se identificaron y caracterizaron las variables más relevantes para el proceso de producción de galletas Wafer con presentación de 40\*204 g para pequeñas empresas, las cuales son:

- La cantidad de producción planificada CPP.
- Horas laborales planificadas HLP.
- Paradas no programadas PNP.
- La cantidad de producción real CPR.

Se realizó la construcción de un modelo de regresión multivariante tomando en cuenta las variables seleccionadas, luego se ingresaron los datos al programa SPSS, para el análisis de los supuestos de la regresión multivariada.

Al realizar el análisis ANOVA de las variables independientes CPP, HLP, PNP nos da un valor de significancia P\_ valor igual a  $0 < 0.05$ , por lo que se puede concluir que dichas variables si predicen a la variable dependiente CPR.

Al analizar los datos reales de cantidades producidas con los datos reales pronosticados se observa que sus medias tienen muy pocas diferencias con lo que se puede concluir que el modelo planteado es factible.

De la comprobación del modelo de regresión multivariado se concluyó que para esta microempresa que produce galletas Wafer con presentación de 40\*204 g, utilizando las variables independientes seleccionadas; la cantidad de producción planificada de CPP, el tiempo (horas) de trabajo planificado de HLP y las paradas no programadas PNP le permiten predecir la cantidad real de producción CPR, con un modelo matemático que tiene una relación significativa calculada en 99,6%.

Propuesta De Un Modelo Matemático Aplicado Al Pronóstico De Producción Utilizado Regresión Multivariado Aplicado A Una Fábrica De Galleta Wafer 40\*204 g

## RECOMENDACIONES

Se recomienda hacer comparaciones con otras variables que intervienen en el proceso de producción, con el fin de tener un análisis minucioso de toda la cadena productiva de la empresa.

Se recomienda poner en práctica el modelo matemático obtenido por regresión multivariante en este estudio, el cual servirá como insumo para que esta microempresa pueda predecir de manera óptima la cantidad de galletas Wafer con presentación de 40\*204 g, que pueden producir bajo ciertas condiciones, de modo que el modelo matemático sirve como una herramienta para optimizar sus operaciones.

La microempresa debe realizar un registro exhaustivo de la mayor cantidad de datos que tengan relación con su proceso productivo, dado que el uso correcto de esta información le permite tener una ventaja competitiva frente a su competencia, ya que podrá ofrecer su producto acorde a las necesidades y demanda de cada uno de sus clientes.

## BIBLIOGRAFÍA

- Arango Marin, J. A., Giraldo Garcia, J. A., & Castrillón Gómez, O. D. (2013). Gestión de compras e inventarios a partir de pronósticos Holt-Winters y diferenciación de nivel deservicio por clasificación ABC. *Scientia Et Technica*, 18(4), 743-747. Obtenido de <https://www.redalyc.org/pdf/849/84929984023.pdf>
- 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>
- 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](https://www.ecorfan.org/republicofperu/research_journals/Revista_de_Ingenieria_Industrial/vol1num1/Revista_de_Ingenier%c3%ada_Industrial_V1_N1_4.pdf)
- Berkowitz, D. E. (1998). INDUSTRIA ALIMENTARIA. En *ENCICLOPEDIA DE SALUD Y SEGURIDAD EN EL TRABAJO* (págs. 67 - 67.35). Madrid: Chantal Dufresne, BA. Obtenido de <https://www.insst.es/documents/94886/161971/Cap%C3%ADtulo+67.+Industria+alimentaria>
- Bocco, M. (2010). *Funciones elementales para construir modelos matemáticos*. Buenos Aires: Instituto Nacional de Educación Tecnológica. 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 <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](https://ojs.supercias.gob.ec/index.php/X-pedientes_Economicos/article/view/85/27)
- Dagnino S., J. (2014). Regresión lineal. *Revista Chilena de Anestesia*, 43(2), 143-149. Obtenido de <https://revistachilenadeanestesia.cl/PII/revchilanestv43n02.14.pdf>
- 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>
- 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>

- López López, M. (2013). *Factores que determinan el consumo nacional aparente de Cártamo en México, 1995-2009*. Buenavista, Saltillo: Universidad Autónoma Agraria. Obtenido de <http://repositorio.uaaan.mx:8080/xmlui/bitstream/handle/123456789/5314/T19752%20%20LOPEZ%20LOPEZ%2C%20MARICELA%20MARCELINA%20%20TESIS.pdf?sequence=1&isAllowed=y>
- 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>
- Manrique Leonett, A. (2007). *Diseño para el sistema estadístico forestal, basado en componentes de software libre*. Mérida: Universidad de los Andes. Obtenido de [http://bdigital.ula.ve/storage/pdftesis/pregrado/tde\\_arquivos/8/TDE-2007-06-29T04:00:51Z-314/Publico/Ana%20Manrique%20Parte%20I.pdf](http://bdigital.ula.ve/storage/pdftesis/pregrado/tde_arquivos/8/TDE-2007-06-29T04:00:51Z-314/Publico/Ana%20Manrique%20Parte%20I.pdf)
- 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>
- Montero, R. (2016). *Modelos de regresión lineal múltiple*. España: Universidad de Granada. Obtenido de [https://www.ugr.es/~montero/matematicas/regresion\\_lineal.pdf](https://www.ugr.es/~montero/matematicas/regresion_lineal.pdf)
- Montes Albarracin, A. I. (2016). *Uso de redes neuronales artificiales para el pronóstico de demanda de hielo industrial en Hielos Santa Lucía S.R.L.* Tacna: Escuela Profesional de Ingeniería en Informática y Sistemas.
- 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 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>
- Nations, O. a. (2020). *OCDE-FAO Perspectivas Agrícolas*. OCDE-FAO. doi:<https://doi.org/10.1787/a0848ac0-es>
- Pineda Jaimes, N. B., Bosque Sendra, J., Gómez Delgado, M., & Plata Rocha, W. (2009). Análisis de cambio del uso del suelo en el Estado de México mediante sistemas de información geográfica y técnicas de regresión multivariantes. Una aproximación a los procesos de deforestación. *Investigaciones Geográficas (Mx)*(69), 33-52. Obtenido de <https://www.redalyc.org/articulo.oa?id=56912295004>
- Ramos, A., Sánchez, P., Ferrer , J. M., Barquín, J., & Linares, P. (2010). *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~w3TqbHmhGJKqarFyW1SAIdIZH5DDJXiHRLC8v2Tdb-zncNZYV4WfKu0LVXNlkmLiT6NoocHcf21rjOhupk0biXvhHdDbsWtw~Me7oBaSZ0g4ocy](https://d1wqtxts1xzle7.cloudfront.net/34357606/modelado_en_gams-with-cover-page-v2.pdf?Expires=1633687201&Signature=EzhFiG-skxPtDHclAwclQ08Qgji~w3TqbHmhGJKqarFyW1SAIdIZH5DDJXiHRLC8v2Tdb-zncNZYV4WfKu0LVXNlkmLiT6NoocHcf21rjOhupk0biXvhHdDbsWtw~Me7oBaSZ0g4ocy)

- Rangel Montes de Oca, L., GarcíaPereira, A., & Hernández Gómez, A. (2009). Uso de técnicas de análisis multivariable aplicadas en la obtención de modelos de predicción de propiedades relacionadas con los sistemas agrícolas. *Revista Ciencias Técnicas Agropecuarias*, 18(2), 74-77. Obtenido de <https://www.redalyc.org/articulo.oa?id=93215937014>
- Render, B., & Heizer, J. (2007). *Administración de la producción*. México: PEARSON EDUCACIÓN.
- Rivera Garcés, A. (2011). *Impacto de las variables meteorológicas en la formación de ozono en la cuenca de Santiago*. Santiago: Universidad de Chile. Obtenido de [http://www.tesis.uchile.cl/tesis/uchile/2011/cf-rivera\\_ag/pdfAmont/cf-rivera\\_ag.pdf](http://www.tesis.uchile.cl/tesis/uchile/2011/cf-rivera_ag/pdfAmont/cf-rivera_ag.pdf)
- Rodríguez Velázquez, J. A., & Steegmann Pascual, C. (Agosto de 2012). *Modelos Matemáticos*. Obtenido de [https://recursos.salonesvirtuales.com/wp-content/uploads/bloques/2012/08/Modelos\\_matematicos.pdf](https://recursos.salonesvirtuales.com/wp-content/uploads/bloques/2012/08/Modelos_matematicos.pdf)
- 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](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 Biembegut, 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 López, E., Barreras Serrano, A., Pérez Linares, C., Figueroa Saavedra, F., & Olivas Valdez, J. A. (2013). Aplicacion de un modelo arima para pronosticar la producción de leche de bovino en Baja California, México. *Tropical and Subtropical Agroecosystems*, 16(3), 315-324. Obtenido de <https://www.redalyc.org/articulo.oa?id=93929595004>
- 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>
- Saucedo Castillo, O., Pérez, L. V., Herrera Isla, L., & Fernández Pérez, L. (2010). Sistema de pronóstico climático del tizón tardío (phytophthora infestans) en el cultivo de la papa en la provincia de Villa Clara. *REDVET. Revista Electrónica de Veterinaria*, 11(3B), 1-12. Obtenido de <https://www.redalyc.org/articulo.oa?id=63613140038>

- 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>
- 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>
- Vargas Orbegoso, R., & Rojas Primo, V. (2017). *Análisis de regresión de la producción lechera en la empresa CAMAY, periodo 2000-2016*. Huacho: Universidad Nacional José Faustino Sánchez Carrión. Obtenido de <http://repositorio.unjpsc.edu.pe/bitstream/handle/UNJFSC/2753/ROJAS%20PRIMO%20Y%20VARGAS%20ORBEGOSO.pdf?sequence=1&isAllowed=y>
- 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>
- Villada, F., Muñoz, N., & García, E. (2012). Aplicación de las Redes Neuronales al Pronóstico de Precios en el Mercado de Valores. *Información tecnológica*, 23(4), 11 - 20. Obtenido de [https://scielo.conicyt.cl/scielo.php?script=sci\\_arttext&pid=S0718-07642012000400003](https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-07642012000400003)
- Villada, F., Muñoz, N., & García-Quintero, E. (2016). Redes Neuronales Artificiales aplicadas a la Predicción del Precio del Oro. *Información tecnológica*, 27(5).  
doi:<http://dx.doi.org/10.4067/S0718-07642016000500016>
- 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>
- Zambrano Macas, D. (2016). *Estimación de la función oferta y demanda para el mercado de la harina de trigo en el cantón Machala provincia de el Oro en el año 2013*. Machala: Universidad Técnica de Machala. Obtenido de <http://repositorio.utmachala.edu.ec/bitstream/48000/6665/1/TUACE-2016-EC-CD00004.pdf>

## ANEXOS.

### ANEXO A. Base de datos del modelo.

	Ctd. prod. Plan CPP	Horas laborables Planificadas HLP	paradas no programadas estimadas en horas PNP	Ctd. prod. Real CPR
1	1290,03	39,45	5,56	1288
2	1292,03	37,8	5,49	1205
3	2810,1	85,48	7,73	2814
4	2293,08	68,48	6,88	2303
5	800,02	26,87	4,36	802
6	1551,02	53,9	6,5	1543
7	776,02	22,65	3,4	796
8	1292,03	37,8	5,09	1205
9	1677,04	52,38	6,88	1691
10	2099,06	62,23	6,6	2114
11	1647,03	43,63	4,76	1477
12	1809,03	55,22	5,88	1799
13	1551,02	48,75	5,66	1550
14	1551,02	48,75	5,96	1550
15	3000,09	93,33	9,02	3204
16	1034,01	29,75	4,54	1034
17	0	1,75	0,17	60
18	2196,08	65	5,79	2136
19	1550,06	50,55	4,47	1561
20	2971,1	87,48	7,58	2983
21	1163,04	32,67	4,54	1145
22	1697,1	47,82	5,29	1692
23	3068,13	86,95	8,27	3066
24	1018,03	32,82	5,11	1015
25	800,02	26,87	4,1	802
26	1697,1	47,82	4,99	1692
27	420	12,35	1,88	436
28	1518,03	42,68	4,42	1530
29	1453,05	40,65	5,26	1459
30	921	26,02	3,68	925
31	1825,02	54,25	6,23	1825
32	1354,05	39,83	4,33	1349
33	1305,04	46,13	6,1	1307
34	3456,13	100,27	9,43	3452
35	2174,09	62,38	6,44	2126
36	1697,1	47,82	5,19	1692
37	1869,09	53,7	6,41	1873
38	1523,04	44,05	4,76	1493

## ANEXO B. Reunión de tutorías.

En un modelo de regresión lineal múltiple, la variable dependiente es una función lineal de  $k$  elementos correspondientes a las variables independientes y una perturbación aleatoria o error. El modelo también incluye un término independiente. Si expresamos por  $y$  a la variable dependiente o variable de respuesta, por  $x_1, x_2, x_3, x_4, \dots, x_k$  a las variables independientes o explicativas; por  $\varepsilon$  al error o perturbación aleatoria, los modelos de regresión multivariada o regresión lineal múltiple viene dado por la siguiente expresión:

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

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

$\beta_0$  = intersección en  $Y$   
 $\beta_1$  = pendiente de  $Y$  con la variable  $X_1$  manteniendo constante las variables

Zoom meeting interface: Juan Solís, Santiago Chele Delgado, Unemi. Status: 100% Looks good!

Word document: Word - tesis Juan David Solís Rivera 01 - Guardado en OneDrive. Search: Buscar (opción + Q). Footer: Página 27 de 63, 10,876 palabras, español (España - alfabeto tradicional), Marcado del editor: activado. Zoom controls: 100%, Ajustar, Proporcionar comentarios a Microsoft.