

CIATEQ, A. C. Centro de Tecnología Avanzada  
Dirección de Posgrado



*Puntos de control para procesos de calidad*

TESIS QUE PRESENTA

**Ing. Adair Rogelio Netzahualcóyotl Correa**  
**Asesor: Dr. Jan Mayén Chaires**

Para obtener el grado de

**Maestría en**  
**Manufactura Avanzada**

San Luis Potosí, San Luis Potosí  
octubre, 2020

## CARTA DE LIBERACIÓN DEL ASESOR



GOBIERNO DE  
MÉXICO



CONACYT  
Consejo Nacional de Ciencia y Tecnología



San Luis Potosí, S.L.P. 26 de abril del 2021.

Mtro. Geovany González Carlos  
Coordinador Académico de Posgrado  
CIATEQ, A.C.

Los abajo firmantes, miembros del Comité Tutorial del **Ing. Adair Rogelio Netzahualcóyotl Correa**, una vez revisado su Proyecto Terminal de tesis, titulado **"Puntos de Control para Procesos de Calidad"**, **autorizo** que el citado trabajo sea presentado por el alumno para su revisión, con el fin de alcanzar el grado de **Maestría en Manufactura Avanzada**.

Sin otro particular por el momento, agradezco la atención prestada.

Firma

---

Dr. Jan Meyer Chaires  
Asesor Académico



## CARTA DE LIBERACIÓN DEL REVISOR



GOBIERNO DE  
MÉXICO



CONACYT  
Consejo Nacional de Ciencia y Tecnología



Villahermosa, Tabasco, a 29 de septiembre de 2021

Dra. María Guadalupe Navarro Rojero  
Directora de Posgrado  
CIATEQ, A.C.

Por medio de la presente me dirijo a usted en calidad de Revisor del proyecto terminal del (la) alumno (a) **Adair Rogelio Netzahualcoyotl Correa**, cuyo título es:

### "Puntos de Control para Procesos de Calidad"

Después de haberlo leído, corregido e intercambiado información con el (la) alumno(a), y realizado los cambios que le fueron sugeridos, puede ser autorizada su impresión, a fin de que se inicien los trámites correspondientes para su defensa.

Sin otro particular por el momento, y en espera de que mis sugerencias sean tomadas en cuenta en beneficio del estudiante y la Institución, agradezco la atención prestada.

Atentamente

Mii. José Carlos Hernández González

F31b Revisión: 01-Mar-2021

Calle 23 de agosto No. 213, Col. Jesús García, CP. 86040, Villahermosa, Tabasco, México.  
+52 (993) 316 83 63 [www.ciateq.mx](http://www.ciateq.mx)



## DEDICATORIA

Quiero dedicar esta tesis a toda mi familia, mi esposa Lucia Baltazar quien me apoyo en todo momento para realizar y completar el programa, además de los tiempos adicionales para revisar y redactar esta investigación. A mi hijo Set Netzahualcoyotl por alentarme y motivarme para seguir creciendo e impulsar mis metas para sobre pasarlas y no rendirme hasta conseguir, esperando sirva también de ejemplo para que no se rinda ante sus objetivos y sueños que desee alcanzar. A mis papas Rogelio Netzahualcoyotl y María del Carmen Correa quienes me brindaron una formación académica y me alinearon por la vida siguiendo siempre las reglas y mejorando en cada paso para contribuir como mejor ser humano para este mundo. A mi hermano Omar Netzahualcoyotl por apoyarme en las etapas importantes de mi vida sabiendo siempre escuchar y acompañar en los momentos más indicados. A mi asesor Dr. Jan Mayen por no perder la esperanza en mi trabajo y desarrollo de la investigación siendo un soporte esencial en la conclusión y formación durante todo el proceso de la Maestría, enfocando siempre su experiencia hacia mejorar cada paso de la investigación realizada en esta tesis.

## AGRADECIMIENTOS

Quiero agradecer a mi asesor por apoyarme y dedicar de su tiempo a las revisiones de esta tesis, así como el involucramiento durante este programa para cumplir con los requerimientos y entregar buenos resultados tanto para la institución como para la empresa.

Agradecer de igual manera a mis líderes en la empresa donde laboro Ethelvina Baños y Maira de Lourdes Vázquez por apoyarme e impulsar el desarrollo y avance de mis entregables para que esta maestría se mantuviera dentro de mis planes profesionales y personales, invitándome incluso a seguir mejorando en cada paso que diese y entregar lo mejor de mi persona.

Agradezco también a Mta. Mary Carmen por ser un apoyo durante el proceso académico, y que siempre estuvo al pendiente de cualquier avance, dando incluso referencias y opciones de todos los entregables para que estos no limitaran mis tiempos de entrega. Así mismo, por tener la confianza y mantener la exigencia con mi persona impulsando siempre a dar lo mejor de mí de manera académica y personal.

\* \* \*

## RESUMEN

Actualmente en el área de calidad recibo no existen procesos o puntos de control que apoyen la toma de decisiones, puesto que los análisis surgen de manera reactiva, sin brindar información sobre las inspecciones de calidad.

Se busca determinar un modelo empírico, con el uso de regresión logística, que apoye la toma de decisiones en procesos de aceptación y rechazo en áreas de calidad recibo en planta, obteniendo como resultado una probabilidad de rechazo en siguientes inspecciones, aportando información en las acciones siguientes para algún material que se vea con riesgo, o caso contrario, reforzar la información necesaria para dejar de inspeccionar algún material. Esta probabilidad resultante es la que da origen a un nuevo indicador, "Estimated Reject Projection", o por sus siglas ERP, que se propone a implementar para calificar y categorizar los riesgos de rechazos que se visualizan para cada uno de los materiales en esta lista de intercepción.

Con las estimaciones que otorga la regresión logística, se puede obtener puntos de control, que apoyan la misma toma de decisiones, soportan al modelo y resuelven dudas sobre algún material y su significancia para el modelo.

Se busca implementar este indicador nuevo ERP para reducir y optimizar el uso del capital humano para las inspecciones, reducir los costos para seguir inspeccionando materiales cuando estos no reflejan ningún riesgo para las líneas de producción, o bien, si estos materiales requieren una mayor intervención por parte de planta o proveedor para resolver hallazgos de calidad.

El correcto seguimiento de estos puntos de control, así como la inferencia objetiva de sus resultados, dan como resultado la reducción de los materiales en la lista de intercepción hasta en un 50%, haciendo del indicador estimado como sus elementos de cálculo, herramientas útiles para soportar decisiones en siguientes pasos, considerando un riesgo futuro.

**Palabras clave:** Ciencias físico matemáticas, Estadística, Teoría y procesos de decisión, Regresión logística.

## ABSTRACT

Actually, in the quality material arrivals department doesn't exist processes or checkpoints that supports decision making, because the analysis done emerge in a reactive way, considering that a well implemented indicator can give information for taking decisions on the quality inspections.

For the investigation in this thesis, it is wanted to determinate an empiric model, with the use of logistic regression, that supports de decision making in acceptance or rejection processes in the quality material arrivals department, obtaining as a result probability for detecting rejects in the next inspections, giving information in the next actions for any material in risk, otherwise, enhanced information needed to stop inspections of any material. This resultant probability is used for a new indicator "Estimated Reject Projection", ERP for the initials, that is proposed to be implemented to qualify and categorized the rejection risks visualized for each of the material in the interception list.

With the estimations gotten from the logistic regression, it can be obtained checkpoints, that besides supporting the same decision making, supports the empiric model, resolving the doubts from any material and its significance for the previous estimated model.

It is wanted to implement this new indicator ERP to reduce and optimize the use of human capital for the inspections, reduce the costs for continuing the inspections on materials that don't show any risk for the production lines, or else, if these materials are required for a bigger intervention for plant members or the supplier in order to solve the quality issues.

The correct tracing of these checkpoints, and the objective inference of the results, conclude in a reduction for the materials in the interception list up to 50%, making the estimated indicator and the calculated elements, a series of useful supporting tools for the decision making in the next steps, considering a future risk.

**Keywords:** Mathematical physical sciences, Statistics, Theory and decision processes, Logistic regression.

## ÍNDICE DE CONTENIDO

RESUMEN .....	v
ABSTRACT .....	vi
ÍNDICE DE CONTENIDO .....	vii
ÍNDICE DE FIGURAS.....	ix
ÍNDICE DE TABLAS .....	x
1. INTRODUCCIÓN .....	1
1.1. ANTECEDENTES .....	3
1.2. DEFINICIÓN DEL PROBLEMA .....	3
1.3. APORTE A PROCESOS DE ACEPTACIÓN Y RECHAZO .....	4
1.4. JUSTIFICACIÓN .....	5
1.5. OBJETIVOS.....	9
1.5.1. Objetivo general.....	9
1.5.2. Objetivos específicos .....	9
1.6. HIPÓTESIS.....	10
2. MARCO TEÓRICO .....	11
2.1. VARIABLES DE ESTUDIO EN PROCESOS DE ACEPTACIÓN Y RECHAZO .....	11
2.1.1. Variables cualitativas .....	12
2.2. REGRESIÓN LOGÍSTICA .....	13
2.2.1. Variables de uso en regresión logística .....	14
2.2.2. Usos de regresión logística.....	15
2.3. USO DE INDICADORES PARA PROCESOS DE ACEPTACIÓN Y RECHAZO .....	22
2.3.1. ¿Qué es un indicador? .....	22
2.3.2. Resultado de indicadores.....	25
2.4. INDICADOR NUEVO .....	26
2.4.1. Necesidad de uso .....	26
2.4.2. Bases para ser calculado .....	27
2.4.3. Información requerida .....	27
2.5. RECOMENDACIONES DE USO DE TESIS .....	28
2.5.1. Alcance y áreas de oportunidad para uso de nuevo indicador .....	28
2.5.2. Áreas de prueba .....	29
2.6. PREGUNTAS PREVIAS AL DESARROLLO DE LA INVESTIGACIÓN.....	30



3. IMPLEMENTACIÓN DE NUEVO INDICADOR.....	32
4. DESARROLLO DEL MODELO .....	34
5. ANÁLISIS DE REGRESIÓN LOGÍSTICA EN NCSS®.....	37
RESULTADOS .....	40
6.1. RESULTADOS OBTENIDOS DE DATOS MENSUALES.....	41
6.1.1. Run Summary .....	41
6.1.2. Y Variable .....	41
6.1.3. Frequency Variable.....	42
6.1.4. Unique Rows (Y and X's) .....	42
6.1.5. Y Variable Summary .....	42
6.1.6. Coefficient Significance Tests.....	43
6.1.7. Coefficient Confidence Intervals.....	44
6.1.8. Odds Ratio .....	46
6.1.9. Estimated Logistic Regression Model(s) in Reading Form.....	47
6.1.10. Row Classification Probabilities.....	49
6.2. MODELO DE REGRESIÓN OBTENIDO .....	52
CONCLUSIONES .....	54
ANÁLISIS DE DATOS MENSUALES .....	58
PREGUNTAS POSTERIORES AL DESARROLLO DE LA INVESTIGACIÓN .....	60
APORTACIÓN DE LA TESIS.....	63
APORTACIÓN SOCIAL DE LA TESIS.....	64
REFERENCIAS.....	65
ANEXO .....	67

## ÍNDICE DE FIGURAS

Figura 1. Ejemplo de gráfico de regresión lineal .....	20
Figura 2. Ejemplo de gráfica de regresión logística .....	21
Figura 3. Carga de datos al software NCSS® .....	36
Figura 4. Seleccionar análisis de regresión logística en software NCSS® .....	37
Figura 5. Selección de variables para análisis de regresión logística en software NCSS®.....	38
Figura 6. Interfaz de resultados software NCSS® .....	40
Figura 7. Sección "Run Summary" dentro de los resultados de software NCSS®. ....	41
Figura 8. Sección "Y Variable Summary" dentro de los resultados del software NCSS®.....	43
Figura 9. Sección "Coefficient Significance Tests" dentro de los resultados del software NCSS® .....	44
Figura 10. Sección "Coefficient Confidence Intervals" dentro de los resultados de software NCSS® .....	45
Figura 11. Sección "Odds Ratio" dentro de los resultados de software NCSS® .....	46
Figura 12. Sección "Estimated Logistic Regression Model(s) In Reading Form" dentro de los resultados de del software NCSS®.....	47
Figura 13. Sección "Row Classification Probabilities" dentro de los resultados del software NCSS®.....	49
Figura 14. Meses de observación utilizando regresión logística nuevo indicador ERP. ....	59

## ÍNDICE DE TABLAS

Tabla 1. Ejemplo de variables binarias Dummy .....	15
Tabla 2. Ejemplo de variables cualitativas y sus respuestas .....	15
Tabla 3. Ejemplo para capturar datos binarios para la investigación .....	27
Tabla 4. Recopilación de datos binarios para la investigación Junio 2018 .....	34
Tabla 5. Acomodo de información para el software NCSS®.....	35
Tabla 6. Resultados del nuevo indicador ERP .....	52

## 1. INTRODUCCIÓN

Esta investigación se realiza con el fin de dar un mejor entendimiento a la estadística y como esta puede aportar a los procesos de calidad de aceptación y rechazo.

En el proyecto se estará investigando las herramientas que se pueden mejor acoplar a procesos de calidad de aceptación y rechazo para cualquiera de los segmentos en los que se encuentre la empresa, ya que se busca la mejor herramienta que pueda aportar de manera general a los demás procesos de aceptación y rechazo que tienen en la industria, inclusive, aportar a otros procesos tales como de producción, mantenimiento, ingeniería industrial, materiales, etc.

Gran parte de esta investigación se centra en una herramienta estadística que mejor puede darnos un modelo de proyección o de pronóstico para procesos de aceptación y rechazo, que es la regresión logística, debido a su proyección de resultados y el significado que pueden tener los mismos.

La regresión logística tiene como propósito brindar probabilidades de éxito y aceptación dadas, unas variables y datos que se irán desarrollando conforme a la investigación y este proyecto. Es decir, que se tendrá como parte de objetivo, determinar las mejores variables que se pueden usar para este tipo de investigación, que no sea complicada de recabar como información y que tenga el impacto necesario en la formulación del modelo para que nos entregue los resultados requeridos, o bien, la mejor visualización de estos.

De esta manera, se podría determinar que un modelo de regresión logística simplemente se acompaña de un proceso de aceptación y rechazo, sin embargo, se busca en base al modelo obtenido, realizar un nuevo indicador de calidad, que apoye al proceso de manera que se pueda medir y además proporcione información a la dirección y gerencia, el estatus que se tiene con dicho proceso.

Una vez que se implemente el indicador se busca observar la diferencia que se tiene con respecto a otros indicadores, ya que, este modelo difiere de los cálculos existentes para los demás indicadores. En el desarrollo de la investigación se busca

explicar de manera clara y concisa cual es la diferencia objetiva entre el nuevo indicador y los indicadores anteriormente utilizados.

Hoy en día, las industrias manejan personalmente sus indicadores, así como los cálculos de estos, sin embargo, no se sabe que tanto se puede apoyar la estadística para predecir o bien evaluar si estos indicadores nos están dando toda la información necesaria. Las conclusiones buscan apoyar en la investigación, para los indicadores, su confiabilidad de cálculo, así como el motivo y razón de ser calculado, para entender de mejor manera como se puede aplicar en la industria o bien como puede ser mejorado.

Como nota importante es necesario resaltar que el título de la investigación "Puntos de control para procesos de calidad" hace una referencia a las herramientas utilizadas durante el procedimiento y cálculo del modelo de regresión logística estimado como betas estimadas, error estándar, odds ratio e intervalos de confianza que, a su vez, complementan cada una de las decisiones tomadas. La probabilidad resultante y el modelo de regresión logística de manera propia definen en su gran mayoría el nuevo indicador, pero se hace el énfasis en observar los demás cálculos y complementos para facilitar la toma de decisión del siguiente paso.

Indirectamente para esta investigación, las variables que se toman para cada uno de los materiales observados se consideran para ser su criterio propio de aceptación y rechazo. Tomando de ejemplos variables de materiales con medidas fuera de especificación que comprenden diámetros más grandes o cortos, se consideran indirectamente como variables de diámetro más grande como rechazo, con el fin de dejar lo más sencillo posible la definición de cada categoría, y que también, apoya a considerar el 95% de confianza que se utiliza en la investigación. Las variables que se visualizan de manera más complicada son las que no tienen como tal una magnitud para medirse, tal el caso como de materiales que presentan golpes en su estructura, en caso de tenerlos, se considera material rechazado. Igual a como se ve anteriormente, esta variable de existencia se considera indirectamente para poder contemplarse entre los dos resultados posibles, aceptado y rechazado, y de esta manera poder considerar el 95% de

confianza que es necesario contemplarse para intervalos de confianza. Aun con todo el intento de juntar todas las variables, puede haber variables que no se consideran en el modelo, o bien, que no están dentro del control para esta investigación. Como variable de cantidad no está contemplada para esta investigación, puesto que la proyección se hace hacia el evento, no hacia la cantidad de material que se va a tener. Adicional, hacer con la cantidad de material rechazado implicaría un desarrollo más complejo que requiere mayor esfuerzo al realizarse individualmente más análisis para cada número de parte ya que las cantidades inspeccionadas no son iguales para cada número de parte dentro de la lista de intercepción. Se recuerda que el modelo se estima de manera general para la lista de intercepción.

### **1.1. ANTECEDENTES**

Para que un material deba ser inspeccionado por calidad recibo, debe de pasar por una revisión previa del Ingeniero de calidad recibo SQA (Supplier Quality Assurance), ya que el indica si el material en cuestión se encuentra dentro de especificaciones o no para su ensamble en las líneas de producción. No hay procesos o puntos de control que apoye actualmente a esta decisión, debido a que se hacen análisis de manera reactiva, siendo que un indicador bien implementado podría brindarle información previa y de apoyo para su decisión en las inspecciones de calidad.

### **1.2. DEFINICIÓN DEL PROBLEMA**

Actualmente en el área de calidad recibo, se realizan inspecciones de materiales que provienen directamente de proveedor, los cuales se califican como aceptado o rechazado según los estándares establecidos en planta que definen si estos materiales tienen algún riesgo de impactar o generar algún problema de calidad en las líneas de producción. Como ejemplos de lo anterior, se tiene el caso del pistón de 4ta velocidad para el subensamble "Rotating Clutch", del cual se

detectan casos de un exceso de material vulcanizado generando un problema en la salida de aire al momento de realizar pruebas finales, por lo que se define la presencia de este exceso de material como rechazo en el proceso de inspección. Otro ejemplo se concentra en una válvula del subensamble "Valve Body", la cual se detectan casos, inclusive desde arribos de proveedor, conteniendo un golpe en la zona lateral que tiene como riesgo atascar la válvula en el cuerpo del subensamble por dicho daño, ya que es un material que se encuentra en fricción; un daño visible en el proceso de inspección como este se define como rechazo de este material.

El problema surge al momento de administrar estos números de parte para ser inspeccionados en lo que se denomina la lista de intercepción para los materiales de calidad recibo, ya que, no se tiene una manera de calificarlos o alguna información de soporte que pueda apoyar la decisión de retirar un material de esta lista de intercepción para dejar de inspeccionarlo, o en caso contrario, respaldar la decisión de seguir inspeccionando el material tomando en cuenta su riesgo de encontrar rechazos que afecten a las líneas de ensamble.

### **1.3. APOORTE A PROCESOS DE ACEPTACIÓN Y RECHAZO**

Con este nuevo indicador se dará como aporte importante la probabilidad de que un material sea rechazado, mediante el proceso de aceptación y rechazo definido. Aporta con información de un futuro que bien puede esperarse prevenir cuando este muy crítico, o pueda continuarse observando o inspeccionando de manera correcta debido a su proyección a determinado tiempo adelante.

Otro aporte importante serían los factores de riesgo (Odds), que dan como resultado primario, el riesgo de rechazo de un material o se ensamble un producto de manera que genere un rechazo.

Adicional a los Odds y su información brindada, se puede calcular el Odds Ratio, que ahora en vez de dar una probabilidad resultante para un proceso de aceptación y rechazo, estaría dando lo contrario, estaría diciendo un estimado de cuantos materiales inspeccionados se rechazarían siguiendo este mismo proceso.

Este coeficiente es muy arriesgado para su implementación ya que, hay variables que no se pueden controlar que impactaran significativamente el modelo y el indicador a su vez, por lo que un estimado de "cantidades" no sería apropiado proyectar por su variabilidad tan grande en estos tipos de procesos, como por ejemplo el inspector, el entendimiento del proceso, criterio de aceptación y rechazo, etc.

La interpretación de los coeficientes también aporta información importante para el indicador, dando información de impacto a este mismo. Por ejemplo, teniendo un coeficiente del modelo beta positivo, se obtiene un aumento a rechazar un material o un ensamble de producto contra la variable que este estimando ese coeficiente, y por su lado contrario uno negativo, se tiene como impacto disminuir la probabilidad de rechazo de un material o producto. Para ambos casos, y siempre para este tipo de indicador, sería tomando en cuenta las variables del proceso de aceptación y rechazo.

#### **1.4. JUSTIFICACIÓN**

Actualmente en GM Transmisiones se tienen indicadores de calidad que reflejan las necesidades del cliente, es decir, los indicadores están regulando las principales necesidades que tiene el cliente para decir que la producción, los procesos, los sistemas e incluso el personal involucrado en la producción, está realizando un producto de calidad.

El indicador que actualmente es utilizado es llamado FTQ (First Time Quality) y FTA (First Time Assembly), sin embargo, estos indicadores no se han utilizado actualmente para procesos de aceptación y rechazo, como lo es en los procesos de inspección de calidad por parte de proveedores, o mejor conocido como calidad recibo SAC (Supplier Action Center).

En los procesos de calidad recibo no se cuenta con un indicador que demuestre la cantidad de rechazos o aceptación que ha tenido sus materiales en inspección, ya que actualmente solo se reporta la cantidad de rechazos, los cuales se contienen en cuarentena, y los de aceptación, los cuales se entregan



directamente a almacén para su entrega a las líneas de producción. De esta manera no se tiene un control óptimo en aceptación y rechazo de un proveedor y sus materiales, el cual puede dar información acerca de sus procesos y apoyar en calificar sus materiales que se reciben si así se requiere.

Es importante recordar que esto indirectamente afecta a las líneas de producción, ya que estas esperan un material, que, en caso de ser inspeccionado, tiene un tiempo de surtido mayor a las líneas, o bien, tienen el riesgo de no estar dentro de especificación. Por consiguiente, un indicador para estos procesos de aceptación y rechazo, pueden aportar gran visión, proyección e información en general a diferentes áreas, así como mejorar sus tomas de decisiones y tiempo de respuesta.

Asegurando que se toma la herramienta correcta y que se tiene diferencia con el indicador actual, se pueden originar preguntas como; ¿Puede ser utilizado para procesos de aceptación y rechazo, específicamente en el departamento de calidad recibo? ¿Qué diferencia tiene el indicador que se pretende estimar en esta investigación? ¿Qué tan eficaz sería al utilizarse en procesos de aceptación y rechazo, por la información que proporciona este indicador? Contestando estas preguntas se tiene que mantener el énfasis en la diferencia entre este indicador y el que se pretende estimar, ya que la diferencia más significativa es con el tiempo medido para cada indicador.

Mediante el uso adecuado de la regresión logística y su probabilidad resultante, podemos hacer el cálculo del nuevo indicador enfocado a que este describa la probabilidad de que se detecte un rechazo en siguientes inspecciones. También, con los elementos dentro del desarrollo del modelo, como los coeficientes Betas, se puede tener una complementación de ideas y decisiones, que, a diferencia de los indicadores actuales, pueden ayudar a prever una acción en lugar de ejecutarse de manera reactiva cuando el problema apenas es detectado. Esto busca impactar a los indicadores actuales de planta como el llamado "FTQ" (First Time Quality), para que se tenga una proyección de probabilidades de rechazo y además se cuente con un seguimiento de calidad a los proveedores y en la planta de proceso. El modelo empírico por sí mismo aporta la información necesaria como su estructura de betas estimadas y significancia de cada una de estas en el modelo

representándose en cada uno de los materiales, para así, visualizar una siguiente acción tomando en cuenta el riesgo que se tiene en la lista de intercepción para las inspecciones de cada material.

De esta manera el indicador que se pretende estimar e implementar es el "Estimated Reject Projection" o por sus siglas en inglés ERP, el cual para calcularse contiene la probabilidad resultante del modelo de regresión logística y que otorga como resultado, una probabilidad de encontrar un rechazo en siguientes inspecciones. Para calcular este indicador mediante el modelo de regresión logística, se valida la simplicidad de recolectar datos y de capturarlos, ya que, el área de calidad recibo es un área que labora las 24 horas en tres equipos de trabajo diferentes, haciendo muy simple el recolectar únicamente respuestas binarias, siendo 1 si se encontró algún rechazo del número de parte inspeccionado en ese día, y 0 si no se encontró ningún tipo de rechazo.

El uso de regresión logística para calcular este indicador se ve en la necesidad de calcular una probabilidad de un suceso siguiente, que en este caso se enfoca en la probabilidad de que un material sea rechazado en la siguiente inspección. Existen diferencias con respecto a otras herramientas de predicciones y DSS (Decision Support System), por ejemplo, podría utilizarse la herramienta estadística simulación de Montecarlo, y aunque esta herramienta si nos apoya a realizar predicciones y proyecciones muy exactas de valores resulta muy complejo adaptarla a esta investigación, ya que se manejan variables discretas en sentido binario, por la manera más sencilla de recopilar datos, y la herramienta de Simulación de Montecarlo se enfoca en variables continuas, por lo que al usarla en este enfoque de observaciones binarias no se obtendrán resultados efectivos y con un mayor tiempo de inversión para dominar la herramienta y entender sus inferencias (Ramírez, 2011).

Para tener un mayor énfasis y apreciar mejor la diferencia, se tiene como ejemplo el trabajo de investigación de Jaccard James y Brinberg Miriam, en la que se puede apreciar como la herramienta de regresión logística se utiliza como un complemento de la misma simulación de Montecarlo, ya que con el termino de "Extent data Simulation" se explica como al momento de realizar un análisis de

categorización y probabilidades de ciertos grupos para presentar algún estudio psicológico esperado, la herramienta de Montecarlo no contempla una serie eventos o categorías haciéndolas parecer incluso perfectas para que sucedan, además de que el realizar el modelo para iniciar la simulación resulta muy complicado transformarlo para tener una orientación hacia las variables cualitativas. De manera objetiva se puede definir y comparar con los autores, asentando al igual que ellos, que un modelo de regresión logística es más sencillo incluso de realizar inferencias para sus resultados y que puede ser aprendido fácilmente para ejecutarse por personas con poca experiencia. Igual se hablan de datos binarios ya que al momento de hacer los procedimientos, resulta muy sencillo para capturar y recopilar los datos necesarios (Jaccard & Brinberg, 2021).

Otro ejemplo con lógica difusa es que esta es utilizada en mayor efectividad para grupos y definir sus comportamientos, dificultando la obtención de una probabilidad a menos que se esté especificando en cada una de las funciones, invirtiendo mayor tiempo y complejidad, por lo que es probable que no entregaría una probabilidad resultante para calcular el nuevo indicador. Cabe mencionar que las herramientas antes mencionadas, no presentan ningún elemento en el desarrollo del resultado que pueda apoyar o reforzar la decisión para el resultado obtenido, por otro lado, la regresión logística nos permite trabajar también con coeficientes betas, odds, intervalos de confianza, error estándar que pueden darnos una mayor información del resultado obtenido en la estimación, que, de hecho, se buscan tener como puntos de control para procesos de calidad en varios niveles (Cuevas & Morales-luna, 2002).

Haciendo una referencia a lo anterior, se tiene una investigación realizada por Zuo Weijie, Yuan Haiwen, Shang Yuwei, Liu Yingyi, Chen, Tao; en donde explica que el uso de la regresión logística como una manera más fácil de hacer un análisis para determinar si un transformador se puede categorizar en un nivel de satisfacción como muy bueno, bueno, malo o muy malo, decidiendo de manera binaria como "Sano" e "Insano", en las conclusiones, inclusive con un indicador desarrollado en la investigación. Esto debido a que, al utilizar un modelo de regresión logística, se puede explicar con las variables medidas, inclusive proyectar el resultado que va

a tener y con que probabilidad se podrá calificar el aceite del transformador. En seguida de las conclusiones, se hace la comparativa con el método de utilizar lógica difusa, aunque tienen resultados muy similares, agrega una mayor especificación de estas variables mediante la regresión logística. De esta manera se puede definir junto con los autores, que la lógica difusa es muy difícil de adaptar para cualquier contexto, incluso para los expertos tiene una mayor complejidad para el procedimiento así como el entender su análisis, y al hacer la comparación con la regresión logística, no se requiere una gran cantidad de conocimiento para ejecutar los análisis e interpretar sus resultados, al igual que los procedimientos, que pueden ser ejecutados y recopilados de manera muy sencilla ya que también se trata con datos binarios y da proyección a un indicador (Zuo, Yuan, Shang, Liu, & Chen, 2016).

Para tener una mayor información y justificaciones de uso para la regresión logística por sobre otras herramientas de simulación y DSS para trabajos relacionados, observar ANEXO A.

## **1.5. OBJETIVOS**

### **1.5.1. Objetivo general**

Obtener un modelo empírico por número de parte, mediante el uso de la regresión logística capaz de proyectar la probabilidad de aceptación o rechazo a futuro para procesos de calidad proveedores General Motors "Global Propulsion Systems" (Planta Transmisiones) de aceptación y rechazo.

### **1.5.2. Objetivos específicos**

- Identificar las variables significativas a estudiar en las áreas de calidad recibo por el proceso de aceptación y rechazo que se usa actualmente para identificar las áreas de oportunidad en el estudio y proceso.

- Determinar indicadores que permitan hacer una explicación a los comportamientos de inspecciones dentro de los procesos de calidad recibo en base a los requerimientos de la investigación y así describir su comportamiento.
- Estimar un modelo de regresión logística en procesos de calidad recibo con los datos históricos obtenidos anteriormente de aceptación y rechazo para proyectar una probabilidad futura de aceptación y rechazo, considerando las variables actuales.
- Diseñar un nuevo indicador en procesos de calidad recibo en base al modelo de regresión logística estimado para predecir futura tendencia y probabilidades de aceptación.
- Aplicar modelo de regresión logística diseñado para procesos de calidad recibo en base al requerimiento de investigación para brindar la mayor cantidad de información y justificación al proceso actual de aceptación y rechazo.
- Evaluar el efecto del nuevo indicador en procesos de calidad recibo con base a modelos de estandarización establecidos para ver su aporte en futuros procesos de aceptación y rechazo.

## **1.6. HIPÓTESIS**

Para continuar la investigación para este modelo se tiene que plantear una hipótesis para poder declarar si la investigación nos llevara al resultado esperado de rechazos y su probabilidad de que suceda.

Por lo tanto, se plantea la siguiente hipótesis:

“El indicador obtenido mediante del modelo de regresión logística, puede calcular la probabilidad futura de rechazo considerando el proceso de aceptación y rechazo establecido para concluir que este indicador con el usado actualmente, se diferencian entre sí y aporta más información, en conjunto con sus elementos calculados, para las tomas de decisiones, esperando reducir la cantidad de materiales a inspeccionar en un 50%”.

Con la hipótesis anterior mencionada el objetivo a alcanzar es la reducción de la cantidad de materiales que se encuentran en la lista de intercepción para cada uno de los meses y poder reducir las cargas de inspecciones, optimizar los capitales humanos y movimientos de materiales involucrados.

## **2. MARCO TEÓRICO**

Para desarrollar el marco teórico, se definen los temas en base a las preguntas efectuadas de investigación para los objetivos específicos, es decir, con el marco teórico se busca contestar todas estas preguntas de investigación que surgieron anteriormente.

### **2.1. VARIABLES DE ESTUDIO EN PROCESOS DE ACEPTACIÓN Y RECHAZO**

Existen dos tipos de variables estudios y cálculos estadísticos, las variables cuantitativas y las variables cualitativas.

Las variables cuantitativas son las que se obtienen a través de mediciones, cálculos matemáticos, etc. Y que dan como resultado un número, sea este entero o continuo (Hernández, 2011).

Sin embargo, para procesos de aceptación y rechazo no se tiene como resultado un número, es decir, al momento de inspeccionar un material se pueden obtener resultados que se dirigen a una categoría, aceptado o rechazado, éxito o fracaso, si o no, por lo que los modelos usuales de estadística para variables cuantitativas no pueden apoyar a esta investigación (Larra, 2007).

Para esta investigación, que se centra en procesos de aceptación y rechazo, se hace enfoque en las variables cualitativas, ya que como se dijo anteriormente, se puede obtener un resultado de categoría para cada inspección que se esté haciendo en un proceso de aceptación y rechazo.

### **2.1.1. Variables cualitativas**

Las variables de respuestas se determinan como cualitativas cuando solo pueden clasificarse dentro categorías no numéricas. Ejemplos de estas variables son el color de ojos de unas personas en una muestra seleccionada si estas personas están matriculadas como estudiantes de una Universidad. En este caso sólo se puede representar con gráficas. Su objetivo es dar una percepción visual sencilla de la muestra estudiada. Dentro de la gran variedad de representaciones gráficas existen los diagramas de barras y diagramas de sectores (Hernández, 2011).

A las variables cualitativas se le pueden denominar como binarias, es decir, solo puede tomar dos valores. Ejemplo de esto es si una vaca está preñada o no, si una planta contiene parásitos o no, entre muchos otros ejemplos (Hernández, 2011).

Las variables cualitativas pueden ser independientes o dependientes. Puede llegar a ser cualitativas para ambas variables, o solamente en una, ejemplo, si se aplica una cantidad de antibiótico conocida a un animal y al cabo de un tiempo se revisa si mejora de salud o no, notando que la variable independiente se considera cuantitativa mientras que la variable dependiente se considera cualitativa. Las pruebas más utilizadas para variables cualitativas son la binomial, multinomial y regresión logística (Hernández, 2011).

#### **2.1.1.1. Variables binarias**

Los estudios cualitativos suelen aparecer a menudo bajo información de tipo binaria, por ejemplo, un individuo si es de sexo masculino o femenino, un individuo tuvo participación o no dentro de un programa de formación a profesionales, si en la empresa observada se ofrece o no un plan de pensiones para trabajadores, etc. (Larra, 2007).

En estas variables cualitativas su representación como variables binarias toman los valores cero y uno, denominándose también como variables ficticias o “Dummy” como lo indica su término en inglés.

Para simplificar el estudio en esta investigación, así como los cálculos del modelo, se usarán las variables binarias para recaudar la información y observar el modelo

con los criterios de aceptación y rechazo. Esto mediante la aplicación de variables Dummy 1 y 0 para darles los valores de aceptación y rechazo a cada pieza que se inspecciona (Alberto & Bosque, 2018).

### **2.1.1.2. Variables categóricas**

Las variables categóricas se definen como una evaluación de rasgos que no se pueden representar en magnitudes, ejemplo de esto son los colores de caballos, formas de hoja de ciertas plantas, razas de cabras, etc. A estas variables les puede ser asignado un valor numérico, pero no por ello representan una magnitud, por ejemplo, lugar de procedencia de un grupo de estudiantes: 1) Aguascalientes 2) Campeche 3) Chihuahua, etc. (Pérez Hoyos, 1996).

Las variables también pueden ser determinadas como nominales, es decir, que el número asignado es arbitrario, como el medio de transporte: 1) Bicicleta, 2) Motocicleta 3) Camión 4) Automóvil, etc.

También pueden ser determinadas como ordinales, que, aunque si manejan una secuencia, no representan magnitudes, ejemplo, Niveles de satisfacción 1) Muy satisfecho 2) Satisfecho 3) Poco satisfecho (Descriptiva & Variable, 2018).

## **2.2. REGRESIÓN LOGÍSTICA**

La regresión logística es parte de métodos estadísticos con esta denominación y es la variante que corresponde a valorar cuanta contribución se tiene de diferentes factores al ocurrir un evento determinado. La regresión logística es adecuada cuando una variable Y de respuesta es politómica, es decir, admite varias categorías de respuesta, como en un evento y que la condición de este empeora, mantiene sus condiciones o mejoran. Sin embargo, es útil cuando en particular se tienen solo dos posibles respuestas, binarias, que es el caso más comúnmente aplicado. La regresión logística es una de las técnicas estadístico-inferenciales que más se utiliza en las producciones científicas contemporáneas (NCSS & LLC, 2018).

La regresión logística puede contestar preguntas esenciales como: (McGee, 2013).



- ¿Se puede realizar una predicción anticipadamente si un cliente solicitando un préstamo a un banco será cliente moroso?
- ¿Se puede predecir que una empresa va a la bancarrota?
- ¿Se puede predecir que un paciente corra riesgo de un infarto según sus condiciones de estudio iniciales?

Para identificar el mejor modelo de regresión logística, se realiza una comparación de modelos haciendo uso del cociente de verosimilitud, indicando a partir de los datos de la muestra cuanto más probable es el modelo seleccionado frente a los demás. La diferencia obtenida de los cocientes de verosimilitud entre dos modelos tiene una distribución de Chi-cuadrada con los grados de libertad correspondientes a la diferencia en el número de variables entre ambos modelos. Si partiendo de este coeficiente no puede demostrarse que un modelo tiene un mejor resultado que otro, se considera como el más adecuado al más sencillo (McGee, 2013).

### **2.2.1. Variables de uso en regresión logística**

Debido a que la metodología para la estimación del modelo de regresión logística basa en utilizar variables cuantitativas, de la misma manera que en cualquier otro método de regresión, es necesario considerar incorrecto que en él se intervengan variables cualitativas, sean éstas nominales u ordinales. Aun así, asignar un número a cada categoría no resuelve el problema. La solución es crear la cantidad de variables artificiales como número de respuestas. Estas nuevas variables, creadas artificialmente, pueden conocerse también como variables internas, indicadoras, o variables de diseño. Para efectos de esta investigación se conocerán a estas como variables "Dummy" (Zekic-Susac, Sarlija, Has, & Bilandzic, 2016).

Si una variable toma datos, por ejemplo, de un estudio de uso de tabaco con respuestas como Nunca fumó, Ya no fuma, Fuma 20 o más cigarrillos diarios, hay 4 respuestas posibles por lo que se construyen  $(4-1=3)$  variables "Dummy" (valores 0, 1), con diferentes posibilidades de codificación, las cuales conducen a interpretar de diferentes maneras, por ejemplo, mostrado en la siguiente imagen Tabla 1:

Tabla 1. Ejemplo de variables binarias Dummy

	I1	I2	I3
Nunca fumó	0	0	0
Ex- fumador	1	0	0
< de 20 cigarrillos diarios	0	1	0
≥ 20 cigarrillos diarios	0	0	1

(Hernández, 2011)

Para esta codificación, el coeficiente de la ecuación de regresión para cada una de las variables “Dummy”, transformándose con la función exponencial, en el ejemplo se ve como cuantifica el cambio del riesgo respecto a no haber fumado nunca (Hernández, 2011).

Una posibilidad diferente es aquella con una variable cualitativa de tres respuestas, y ejemplo de esto se puede observar en la siguiente imagen Tabla 2:

Tabla 2. Ejemplo de variables cualitativas y sus respuestas

	I1	I2
Respuesta 1	0	0
Respuesta 2	1	0
Respuesta 3	1	1

(Hernández, 2011)

### 2.2.2. Usos de regresión logística

El modelo de regresión logística ha sido utilizado por muchos años para modelos cualitativos tratando de explicar cómo se relacionaban las variables categóricas entre sí, sin embargo, en la aplicación para diversos estudios se ha podido apreciar el poder y la aplicación de estos modelos.

Muchas de las técnicas para aplicar el método e interpretación de los resultados y sus coeficientes pueden ser solamente encontrados en literatura estadística, lo que está fuera de la comprensión de muchos usuarios potenciales, que, a su vez, fuera de muchas áreas de oportunidad para su uso (Manotas & Ibarra, 2014).

Tiene la ventaja de no requerir supuestos como el de normalidad, por lo que la estimación de los datos, así como la distribución de estos, no influye en el cálculo de este modelo. Al ser multivariable no hace referencia a la interacción entre estas y no afecta al modelo; y cumpliendo con homocedasticidad, la igualdad de varianzas no afecta al residuo o perturbaciones (Goodman, 2017).

Todos estos supuestos son difíciles de verificar además de requerir un tiempo elaborado para calcularlos. El modelo de regresión logística es más potente que el análisis discriminante cuando estos supuestos no se cumplen.

Otra ventaja radica en su similitud con la regresión múltiple: permite el uso de variables independientes continuas y categóricas, el cual es muy útil para realizar diagnósticos. Tiene una amplia aplicación en estudios con observaciones, de encuesta y con experimentos, así también como en estudios epidemiológicos como se han usado anteriormente dando una explicación de manera estadística como influyen las variables y su probabilidad de sucesos (Peng, Lee, & Ingersoll, 2002).

Numerosas investigaciones muestran las ventajas de utilizar el análisis de regresión logística cuando los datos dentro de las muestras no son uniformes, mixtos y cuando no se cuenta con muestras grandes para la investigación (Zekic-Susac, 2016).

Para este caso de la investigación, se busca encontrar o estimar una probabilidad resultante de los procesos de calidad recibo para poder tener información de apoyo en la toma de decisiones para un futuro arribo por parte de proveedor. También se busca con la estimación de sus coeficientes interpretar que tanto influye cada variable en el proceso de aceptación y rechazo, así como sus residuos, los cuales indicarían cual sería la perturbación mínima del modelo aun no encontrando ningún rechazo dentro de las inspecciones.

### 2.2.2.1. Modelo de regresión logística

Sea una variable  $Y$  dependiente binaria, cuyos dos valores posibles etiquetados son 0 y 1, y sean  $X_1, \dots, X_K$  un conjunto de variables independientes observadas con una finalidad de dar una explicación y/o hacer una predicción del valor de  $Y$  (Shalizi, 2013).

El objetivo es determinar la probabilidad de que se suceda un evento seleccionado, para este caso la probabilidad de detectar un rechazo. Por lo tanto, se determina con la siguiente notación:

$$P[Y = 1 | X_1, \dots, X_K] \quad (1)$$

por lo tanto

$$P[Y = 0 | X_1, \dots, X_K] = 1 - P[Y = 1 | X_1, \dots, X_K] \quad (2)$$

Se construye el modelo con una forma:

$$P[Y = 1 | X_1, \dots, X_K] = p(X_1, \dots, X_K; \beta) \quad (3)$$

Donde

$$p(X_1, \dots, x_K; \beta) \in \quad (4)$$

es una función que recibe nombre de función de probabilidad y que su valor depende de un vector de parámetros

$$\beta = (\beta_0, \beta_1, \dots, \beta_K)' \quad (5)$$

Con el fin de hacer una estimación de  $\beta$  y hacer un análisis del comportamiento del modelo considerado, observamos una muestra aleatoria simple de tamaño  $n$  definida por

$$\{(x'_i, y_i): i = 1, \dots, n\} \quad (6)$$

Donde

$$x_i = (1, x_{i1}, \dots, x_{iK})' \quad (7)$$

es el valor de las variables independientes

$y_i = \{0,1\}$  es el valor Y observado en el i-ésimo elemento de la muestra determinada (Shalizi, 2013).

$$Y|(X_1, \dots, X_K) \sim \text{Bernoulli}(1, p(Y = 1|X_1, \dots, X_K; \beta)) \quad (8)$$

Con el hecho de que la variable respuesta puede tomar sólo dos resultados (aceptación y rechazo), cuando el número de éxitos en n repeticiones tiene una distribución binomial  $B(n, p)$ , la función de verosimilitud es:

$$L(\beta|(x'_1, y_1), \dots, (x'_n, y_n)) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \quad (9)$$

Donde

$$p_i = p_i(x'_i; \beta) = p(x_{i1}, \dots, x_{iK}; \beta): i = 1, \dots, n \quad (10)$$

(NCSS & LLC, 2018).

Sea

$$p(X_1, \dots, X_k; \beta) = G(\beta_1 X_1, \dots, \beta_k X_k) \quad (11)$$

donde  $G(x) = \frac{e^x}{1+e^x}$  es aquella función de densidades acumuladas (función logística) (Shalizi, 2013).

El modelo normalmente se conoce como:

$$\ln\left(\frac{p(x_1, \dots, x_K; \beta)}{1-p(x_1, \dots, x_K; \beta)}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_K x_K \quad (12)$$

Conocido como modelo "logit". Cuando la variable cualitativa toma el valor 1, la expresión se define como:

$$\frac{P(y=1|x_1, \dots, x_K)}{P(y=0|x_1, \dots, x_K)} = \frac{p(x_1, \dots, x_K; \beta)}{1-p(x_1, \dots, x_K; \beta)} \quad (13)$$

(Statistical & Ncss, 2021).

Se le llama como "Odds", a este que es un factor de riesgo. Ejemplo de esto es, en el mundo de la medicina, la variable Y indicaría normalmente la presencia de una enfermedad objeto de estudio y en caso de su ausencia tomaría el valor de 0.

### 2.2.2.2. Resultados del modelo de regresión logística

Coefficientes Betas

$$(\beta_0, \beta_1, \dots, \beta_K)$$

El coeficiente representa un incremento o un decremento en el modelo, es decir, que afecta al aumentar o disminuir la probabilidad resultante, el cual puede favorecer dependiendo del contexto de la investigación (Park, 2013).

Así mismo, estos coeficientes afectan a otros resultados como Odds, y Odds-Ratio, los cuales por su significado en la investigación puede decirse que aumenta o disminuye la probabilidad de que suceda un evento en particular estudiado (Park, 2013).

Probabilidad resultante

$$P[Y = 1|x_1, \dots, x_K] = \frac{1}{1+e^{(-\beta_0-\beta_1x_1,\dots,-\beta_Kx_K)}} \quad (14)$$

Esta es la probabilidad de que se rechace el objeto de estudio, dados los coeficientes estimados Betas (McGee, 2013).

Odds

$$Odds = \frac{P[Y=1|x_1,\dots,x_K]}{1 - P[Y=1|x_1,\dots,x_K]} = \frac{p[Y=1|x_1,\dots,x_K;\beta]}{1-p[Y=1|x_1,\dots,x_K;\beta]} = e^{\beta_0+\beta_1x_1+\dots+\beta_Kx_K} \quad (15)$$

Este cociente indica la presencia de un rechazo en el proceso de aceptación y rechazo objeto de estudio (McGee, 2013).

Odds Ratio

$$OddsRatio = \frac{P}{Q} = \frac{Probabilidad\ de\ Exito}{1-Probabilidad\ de\ Exito} \quad (16)$$

Este cociente representa la cantidad de aceptación que se tendrá por cada rechazo que se encuentre (Goodman, 2017).

### **2.2.2.3. Probabilidad en modelos de regresión logística**

La probabilidad resultante que se obtiene en un modelo de regresión afecta directamente la ocurrencia de un evento dada una serie de variables puestas.

Para esta investigación y como parte del objeto de estudio de interés, la probabilidad de rechazo en este caso se vería reflejado en la probabilidad de que un material sea rechazado dentro de las especificaciones del modelo.

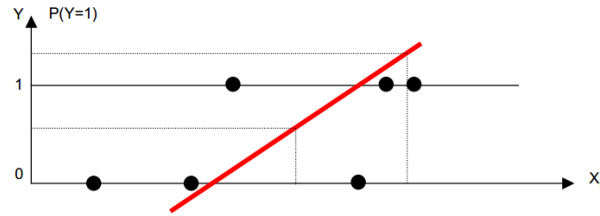
A su vez la probabilidad tendrá como oportunidad de indicarnos cuanto porcentaje de rechazo se puede visualizar para un futuro, que puede aportar información importante para toma de decisiones en inspecciones de calidad recibo, apoyando incluso al recurso que se usa para estos casos.

### **2.2.2.4. Diferencias con modelos de regresión lineales**

Las diferencias entre los modelos lineales y los modelos radican en su proyección o extrapolación de probabilidad, ya que en una tendencia ya antes vista o bien en variables cuyo valor es cuantitativo, los modelos de regresión lineales son lo que mejor se ajustan a ese tipo de variables. Esto debido a que la capacidad del modelo lineal revela que se tiene una influencia significativa, así como el modelo puede decir un valor en particular para un futuro cualquiera con solo conocer algunos datos previos.

Sin embargo, al usarse para los casos de variables cualitativas, crea una inconsistencia significativa, dando como resultado una proyección errónea de cómo se vería afectado el modelo con respecto al tiempo (Alejandro, Pérez, Kizys, Manzanedo, & Hoyo, 2016).

Figura 1. Ejemplo de gráfico de regresión lineal



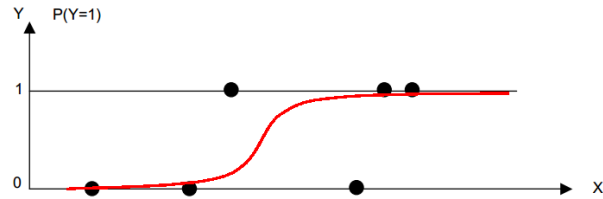
En la Figura 1 se muestra un intento de ajuste de modelo lineal para el caso de variables cualitativas, pero como se puede observar en la imagen, este no adopta ni se ajusta correctamente a los valores que pueden tomar las variables.

Dicho esto, se puede caer en un error donde la probabilidad resultante nos dé una proyección inadecuada, y es por eso, que se usa un ajuste de modelo no lineal para este caso de la investigación en particular, lo cual apoya a que se proyecte dicha probabilidad en el transcurso del tiempo (Alejandro, Pérez, Kizys, Manzanedo, & Hoyo, 2016).

Como diferencia clara, no solo se estimará un modelo no lineal para esta investigación, sino que también se proyecta en una serie de "curvas" que se ajustan mejor a una proyección de variables cualitativas, y específicamente a las binarias. Este modelo proyectado queda muy bien definido con la investigación realizada, porque, así como el proceso binario puede tomar esos valores y arrojar una probabilidad resultante, un proceso de aceptación y rechazo de igual manera puede decir cuando un material esta rechazado y con qué probabilidad será rechazado la siguiente vez, tomando en cuenta las mismas variables tomadas con anterioridad. Se explica de la misma manera con una gráfica para un modelo no lineal con cambios de estados de 0 a 1 en la Figura 2 de tal manera que se pueda explicar visualmente la diferencia entre ambos modelos de regresión (Alejandro, Pérez, Kizys, Manzanedo, & Hoyo, 2016).

Figura 2. Ejemplo de gráfica de regresión logística





## 2.3. USO DE INDICADORES PARA PROCESOS DE ACEPTACIÓN Y RECHAZO

### 2.3.1. ¿Qué es un indicador?

Los indicadores de calidad en determinados procesos son medidas estadísticas basadas en cifras o tasas para dar un criterio, y así, juzgar y evaluar el desempeño de la organización, un sistema o un proceso determinado.

El sistema de gestión de la calidad en una empresa define el punto de partida para determinar y llevar a cabo el significado de cada uno de los indicadores que pueden usarse para medir: (Flores Zambada, 2012).

- Medios de comunicación y su seguimiento
- Aplicación de medidas correctoras y ajustes de procesos
- Rendimiento de un proceso específico
- Calidad empresarial y/o organizacional
- Calidad en los servicios
- Preocupaciones potenciales enfocada en la calidad
- Áreas con un mayor requerimiento de estudio e investigación
- Transformaciones y evolución de un objetivo a lo largo del tiempo, para ejercer un mayor control

Debido a esto, los indicadores a elegir se recomiendan que deban ser: (Flores Zambada, 2012).

Viables: Para permitir obtener datos de un modo factible y conllevando los costos proporcionales y un uso adecuado de recursos.

Válidos: Para medir un aspecto relacionado con la calidad.

Comparables: Para que pueda hacerse un análisis comparativo, por áreas, por segmentos de tiempo o por productos.

Cuantificables: Para que puedan ser medidos.

El análisis de todos estos datos debe proporcionar una información relativa a características y tendencias de procesos y productos específicos, incluyendo oportunidades para accionar de manera preventiva. Por otro lado, debe proporcionar la misma información para dar conformidad con respecto a los requisitos del producto, a la satisfacción del cliente o la gestión de proveedores en la empresa y sus procesos (Flores Zambada, 2012).

#### **2.3.1.1. Indicadores actuales**

Actualmente en la empresa GM GPS, no existen los indicadores necesarios para procesos de aceptación y rechazo para áreas de inspección de calidad recibo.

Sin embargo, si hay un indicador utilizado para medir la calidad total y a la primera de un producto, el FTQ. Este indicador lo que infiere es cuantos productos se realizaron bien a la primera, es decir, cuantos productos en las líneas de ensamble fueron ensamblados correctamente de tal manera que no se marcaron como rechazos ni por el departamento de calidad ensamble, o bien, por los bancos de prueba que se tiene para probar los productos en un ambiente más real de tracción, movimiento, rendimiento, etc.

#### **2.3.1.2. Indicador para proponer**

El indicador que se propone es el ERP, el cual indica a futuro cual es la probabilidad de que un material sea rechazado incluyendo las variables establecidas, así como lo ya visto anteriormente.

Con el nuevo indicador no se estaría viendo hacia atrás, es decir, no se estaría contemplando todos los valores del pasado que ya se han observado, que, aunque si se usan para estimarse, este indicador haría una proyección mediante

regresión logística de cómo sería la tasa de rechazo de la siguiente ocasión que se inspeccione un material si las variables se mantienen de la misma manera.

Con este nuevo indicador, se plantea tener mayor control y seguimiento con los materiales que se inspeccionan ya que puede incluso medir y calificar un material que se está inspeccionando actualmente, inclusive, medir y calificar hasta al mismo proveedor. El principal aporte que busca dar este nuevo indicador es el de brindar información, porque si bien este puede dar un valor futuro de rechazo, puede prevenir y alertar de cierta cantidad de rechazo cuando este material ni siquiera ha tocado físicamente planta. Lo anterior dicho con intervalo de confianza del 95% puesto que se considera que hay variables ajenas al control de la herramienta de regresión logística.

Esto apoyaría también con mucha información al departamento de calidad recibo, puesto que se tendrá una información adicional para tomar decisiones con respecto al proveedor, como penalizaciones, retornos de material, inspecciones de proveedor, etc.

Así mismo, a los niveles de gerencia y alta dirección, se plantea brindar información de cuanto material puede estarse proyectando en aceptación, para que se tomen las decisiones importantes de producción o cambios de ingeniería importantes tanto en los materiales y el producto, así como en las líneas de ensamble.

### **2.3.1.3. Diferencia entre indicadores**

Es importante resaltar la diferencia entre estos indicadores, ya que brinda información parecida para una aceptación o rechazo, sin embargo, la diferencia más importante, como adelanto, es el tiempo en el que se ve reflejado el resultado.

Para el cálculo del FTQ se tiene un porcentaje que se obtiene a través de la operación

$$FTQ = \frac{\text{Total Produccion} - \text{Rechazos}}{\text{Total Produccion}} \quad (17)$$

Esta operación como resultante da un porcentaje en donde se ve la cantidad de productos aceptados a la primera vez, es decir, de todos los productos que se

intentaron ensamblar o bien inspeccionar, solo el porcentaje obtenido es el que está calificado como OK (dentro de especificación) cumpliendo sus requerimientos para aceptación.

Indirectamente este indicador lo que da como información es un histórico de cuanto producto está siendo rechazado, y desde su inspección o producción, dice hacia el pasado como se fue comportando el proceso para poder aceptar el producto atacando por supuesto las variables previamente establecidas, así como las causas comunes y raíz de las variables que más afectan al indicador.

Por otra parte, el indicador que se propone al ser calculado por un modelo de regresión logística da una proyección a futuro y cómo se comportará el proceso de aceptación y rechazo a la siguiente inspección o ensamble, en otras palabras, puede decir, a un futuro proyectado, si mediante el proceso actual manejado y sus variables, un material o producto será rechazado o aceptado con una probabilidad obtenida.

De esta manera se puede concluir que FTQ es un indicador que indica cómo se estuvo comportando el proceso desde un histórico obtenido, el cual da una base para poder tomar decisiones o bien atacar variables y causas comunes para disminuir los rechazos. Así mismo, se puede decir que ERP es un indicador que hará visualizar hacia el futuro, se proyectará un resultado pronosticado para decir que mediante el proceso de aceptación y rechazo establecido, así como sus variables, darán una probabilidad de aceptación calculada, lo cual puede apoyar para toma de decisiones en tiempo y forma, además de encontrar otros usos en el modelo como interpretación de coeficientes, y ver que variables son las que están afectando más al modelo.

## **2.3.2. Resultado de indicadores**

### **2.3.2.1. Interpretación de indicadores actuales**

Como ejemplos claros de resultados al calcular estos indicadores se puede decir que el FTQ resultante puede tener el cálculo en la siguiente ecuación, y que tiene una interpretación de la siguiente manera:

$$FTQ = \frac{100-2}{100} \times 100\% = 98\% \quad (18)$$

“El 98% de los productos elaborados el día de ayer, se ensamblaron y fueron aceptados al primer intento, sin mostrar especificaciones de rechazo”.

### **2.3.2.2. Interpretación de indicador calculado**

Y para el siguiente indicador, ERP, con el mismo resultado visualizado en la ecuación se tendría la siguiente interpretación:

$$ERP = 1 - P[Y = 1] = 98\% \quad (19)$$

“En la siguiente jornada de inspección, los materiales que serán inspeccionados tendrán un 98% de no ser rechazados” o también podemos decir “El siguiente material que se inspeccione tiene una probabilidad de 98% de no tener especificaciones de rechazo”.

## **2.4. INDICADOR NUEVO**

### **2.4.1. Necesidad de uso**

La necesidad de uso radica en no poder usar la información completa del indicador actual FTQ para poder decir cómo será el comportamiento del proceso de aceptación y rechazo establecido, ya que como dijimos anteriormente, este solo puede dar información de un histórico o de un tiempo o ciclo en el pasado, es entonces imposible conocer si el proceso actualmente es capaz para seguir aceptando o rechazando el producto ensamblado o el material inspeccionado.

La necesidad de implementarlo se ve en la necesidad de poder encontrar una predicción y así poder dar conclusiones por adelantado para que un proceso sea viable o riesgoso para seguir su funcionamiento. Esto no nada más con la probabilidad resultante, si no con otros indicadores que se obtienen en el modelo mediante recolección de información variada.

### 2.4.2. Bases para ser calculado

Para que este indicador sea calculado de la mejor manera, se necesita como base, un total de rechazos, o bien, el total de piezas que no se pueden considerar como aceptadas u "OK".

La fórmula que se emplea para calcularlo es (Hernández, 2011):

$$P[Y = 1|x_1, \dots, x_K] = \frac{1}{1+e^{(-\beta_0-\beta_1x_1,\dots-\beta_Kx_K)}} \quad (20)$$

$$p[x_1, \dots, x_K: \beta] = G[\beta_1x_1, \dots, \beta_Kx_K] \therefore G(x) = \frac{e^x}{1+e^x} \quad (21)$$

De esta manera también se necesita un conocimiento previo en los estimadores de regresión como las betas, las cuales influyen directamente en el modelo

Como sugerencia para esta investigación también se tendrá que usar como base un software de operaciones matemáticas, debido a que algunos modelos son tan complicados de estimar como para hacerlos a mano alzada (Hernández, 2011).

### 2.4.3. Información requerida

#### 2.4.3.1. Estimación de variables binarias

Para el seguimiento más certero con la estimación del modelo, así como una manera más sencilla de calcular el nuevo indicador, se usarán variables binarias teniendo el uso siguiente como ejemplo en la Tabla 3 siguiente:

Tabla 3. Ejemplo para capturar datos binarios para la investigación

<b>Inspecciones</b>	<b>Dia 1</b>	<b>Dia 2</b>
<b>Material 1</b>	0	1
<b>Material 2</b>	1	0
.	.	.
.	.	.
.	.	.
<i>Material N</i>	<i>Sum Def Dia 1</i>	<i>Sum Def Dia 2</i>

De esta manera se puede ver fácilmente las sumatorias necesarias y los coeficientes betas necesarios.

Todos los procesos de aceptación y de rechazo se llevarán bajo este mismo esquema, para poder calcularse todos de la manera más homogénea, y así ver la diferencia de su proyección mediante los rechazos encontrados.

#### **2.4.3.2. Recaudación de datos cualitativos**

La información recaudada para esta investigación se busca hacer mediante una captura de información para cada característica observada en el proceso de aceptación y rechazo.

En otras palabras, por cada una de las características que harían a un material o producto aceptable o rechazado, en donde se realizarían las sumas de los rechazos observados con las variables previamente observadas.

El efecto más importante sobre el nuevo indicador se verá reflejado por la mayor cantidad de observaciones que se hagan en un turno de inspección, para próximamente calcular el indicador y decirnos cuanto porcentaje se espera rechazar en la siguiente inspección en base a las variables observadas.

### **2.5. RECOMENDACIONES DE USO DE TESIS**

#### **2.5.1. Alcance y áreas de oportunidad para uso de nuevo indicador**

Tratándose este de un indicador que proyecta información a futuro, muchas áreas o usos pueden surgir para este nuevo indicador, que, aunque para esta investigación se usa para procesos de aceptación y rechazo, puede ser modificado de tal manera de ver afectaciones en otras áreas, así como información importante que apoya a otro tipo de decisiones.

Para el caso particular de procesos de calidad, este puede tener un alcance potencial para abarcar garantías de materiales y productos en base a las variables que más impactaron en el modelo, y así calcular su probabilidad de falla estando en uso ya.

También dado que es una función logarítmica, puede ser modificada de tal manera que pueda brindar información de confiabilidad, en el que se puede tener una medición o proyección de la confiabilidad de máquinas, herramientas, detectores en proceso de errores, entre otras tomando en cuenta las variables que están impactando al proceso, y aunque no se tenga una medición binaria como sucede en esta investigación, puede arrojar una probabilidad resultante de categorías en donde se puedan categorizar y evaluar cada una de las maquinas, herramientas, etc. dado sus usos en el procesos y sus variables significativas.

Aun siendo un estimador muy eficiente de proyección a futuro, no podría recomendarse a algunos productos o materiales, ya que, si las variables medidas tienen una desviación o varianza muy grande, no puede brindar un estimador al modelo apropiado, dando estimaciones muy poco certeras al momento de ver las conclusiones. Esto también influenciado por el número y cantidad de inspecciones que se realicen, porque mientras más observaciones se tengan, se puede tener una mejor proyección del modelo que dará más información hacia adelante.

La generación de estos modelos también puede ayudar a reducir costos, que, aunque no sea vean en las primeras mediciones, podrán definir e incluso ampliar un criterio de aceptación, observando que las variables tal vez no tengan un impacto significativo en el modelo, haciendo entonces un reajuste del criterio de aceptación y bien cancelando esa inspección para ahorrar tiempo de inspección y un elemento en el mismo.

### **2.5.2. Áreas de prueba**

Las áreas de prueba donde se realiza esta investigación tendrán como fase inicial el departamento de calidad recibo "SAC" (Supplier Action Center), área donde se ve un área de oportunidad muy significativa para aplicar todos estos indicadores.

Se toma también esta área inicial debido a que sus inspecciones de recibo actuales serán mejor definidas en variables binarias, es decir, la mayoría de sus procesos actuales de aceptación y rechazo para los materiales que se



inspeccionan tienen en su mayoría dos categorías aceptado/rechazado, facilitando así nuestros primeros cálculos y resultados.

Para el departamento de calidad recibo actualmente no se tiene considerado algún indicador para penalizaciones a proveedor, intercepción de material e incluso retención o disposición de material sin cumplir las especificaciones de calidad. El único seguimiento que se le da es el de porcentaje de rechazo y cuanto rechazo por turno se está teniendo.

Esto apoyará a evaluar también, si para el departamento de calidad recibo, resulta útil esta información ya sea para tomar alguna decisión, o para crear un contexto con el proveedor.

La siguiente fase puede ser considerado para el área de calidad ensamble, sin embargo, para esta área tendríamos que considerar más variables binarias, que incluso pueden llegar a las categóricas. Aun si se es más complejo, se puede realizar modelos para ver como es el comportamiento e impacto de éste en el seguimiento de los procesos de ensamble otorgando retroalimentación directa a los ingenieros de calidad para valorar el efecto que se tiene sobre los procesos en observación.

Las siguientes fases de prueba puede considerarse para áreas de producción, principalmente en las áreas donde se detectan defectos y errores de producción, sin embargo, para estas ya tendría una modificación mayor, por lo que su complejidad aumentaría.

## **2.6. PREGUNTAS PREVIAS AL DESARROLLO DE LA INVESTIGACIÓN**

Con el marco teórico desarrollado previamente, se puede contestar a una serie de preguntas que surgen previo a iniciar con el desarrollo de la investigación, estas preguntas se plantean con el fin de resolver todas las dudas que pudieran surgir y que la investigación se dirige para contestar la mayor de estas. Sin embargo, las siguientes preguntas ya pueden tener una respuesta sin tener el desarrollo completo de la investigación.

¿Cómo se puede aplicar la regresión logística en procesos de aceptación y rechazo?

Se puede utilizar en procesos de inspección de calidad donde se pueden calificar materiales como dentro o fuera de especificación, de igual manera en las áreas donde exista alguna calificación o categorización de condiciones, lo que se puede considerar el alcance esperado para estos resultados.

¿Qué puede aportar la regresión logística a procesos de aceptación y rechazo?

Se tiene una mejor conclusión de las siguientes inspecciones en base a los resultados obtenidos, y lo que aporta el modelo de regresión logística como sus derivados, es información complementaria para tomar mejores decisiones para mantener o sacar materiales de inspección de acuerdo con la lista de intercepción definida previamente.

¿Cómo se puede interpretar los resultados obtenidos?

De manera general, se interpretan como la significancia que se tiene para cada material de la lista de intercepción y que tanto influyen de esta manera al resultado del nuevo indicador ERP para definir qué tanto riesgo se tiene para seguir encontrando rechazos en las inspecciones.

¿Qué indicador se propone para esta investigación?

El indicador que se propone para la investigación como se menciona anteriormente, es el ERP, el cual es la probabilidad que se obtiene a través de las observaciones y que proyectan una estimación para encontrar rechazo en la siguiente inspección a realizar.

### 3. IMPLEMENTACIÓN DE NUEVO INDICADOR

Para poder implementar este indicador en el proceso de aceptación y rechazo que actualmente se usa en GM GPS, y para que se tome en cuenta para otras investigaciones, es necesario conocer las variables del proceso, o en su defecto, definir las de manera binaria como se explica anteriormente.

Una vez que se tiene definido las variables binarias, estas deben de ser consideradas en 2 escenarios posibles, en este caso de la investigación fue de aceptación y rechazo. Sin embargo, se pueden usar otros calificativos binarios como si/no, afecta/no afecta, etc.

Como parte de ejemplificar el tipo de especificaciones de aceptación y rechazo para el área de calidad recibo, se puede tomar como ejemplo la válvula "Boost" del cuerpo de válvulas, la cual se valida si presenta algún golpe o daño, clasificando así el criterio de rechazo si hay presencia de estos últimos mencionados. Hay ciertas materiales con tolerancias que dimensionalmente no se encuentran dentro de especificación, sin embargo, para hacer lo más simple posible la inspección para el personal de calidad recibo, se busca algún proceso para categorizarlo en aceptado o rechazado sin la necesidad de realizar alguna medición, como ejemplo se tiene el caso de un rodamiento que contaba con dimensiones de diámetro interior fuera de especificación, por lo que se consigue una parte del sub ensamble afectado para definir si este entra o no en la cavidad requerida, clasificando así el criterio de rechazo si no entra en dicha cavidad del sub ensamble.

Una vez obtenido el acomodo y determinación de las variables binarias es importante saber que se tiene que recaudar la sumatoria de todos los rechazos durante el tiempo de inspección, cualquier rechazo encontrado puede alterar completamente un modelo, por lo que todo rechazo encontrado debe de ser sumado y cualquier espacio en blanco puede afectar significativamente el modelo.

La sumatoria debe dar como consecuencia una beta por cada una de las variables que estamos definiendo, para así calcular el modelo de regresión

logística. Sin embargo, al ser cálculos muy elaborados si será necesario un software como MiniTab® o en su defecto una Macro en Excel® para realizar estos cálculos, también se pueden proponer otros softwares como R®, MathLab®, entre otros, pero si se requieren hacer cálculos automáticos.

Por último, es saber el significado de los coeficientes obtenidos en el modelo, ya que, al presentar los resultados, se debe de conocer la afectación de cada uno de estos, así como el indicador ERP, que da la probabilidad de rechazo a futuro.

## 4. DESARROLLO DEL MODELO

Para la recopilación de los datos se acomodan los números de parte que se encuentran en lo que se denomina la lista de intercepción, los cuales varían durante el inicio de cada mes (Junio 2018 = 15, Julio 2018 = 21, Agosto 2018 = 28, Septiembre 2018 = 31, Octubre 2018 = 36, Noviembre 2018 = 23, Diciembre 2018 = 24, Enero 2019 = 17, Febrero 2019 = 16, Marzo 2019 = 16, Abril 2019 = 15, Mayo 2019 = 18), se vacían los materiales que se estaban inspeccionando de manera diaria, haciendo así un recopilado total de 12 meses (junio 2018 a mayo del 2019).

Si en el día se encontró algún defecto para ese número de parte, se vacía en el archivo colocando un numero "1", caso contrario se vacía un "0" para decir que ese día no se encontró rechazos para ese material en el lapso del día marcado. De manera visual tenemos como ejemplo la siguiente Tabla 4:

Tabla 4. Recopilación de datos binarios para la investigación Junio 2018

NumerodeParte	01-jun	02-jun	03-jun	04-jun	05-jun	06-jun	07-jun	08-jun	09-jun	10-jun	11-jun	12-jun	13-jun	14-jun	15-jun	16-jun	17-jun	18-jun	19-jun	20-jun	21-jun	22-jun	23-jun	24-jun	25-jun	26-jun
24288511	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24291933	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24291929	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24286406	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24286407	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24288501	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24284996	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24284994	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24282500	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
24290840	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24268389	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24268025	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
11548287	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
24289021	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
24268026	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(Elaboración propia)

Siguiente a vaciar toda la información recolectada por mes, se acomodan los totales de las detecciones de aceptación y rechazo por mes para poder vaciarse y tener una inferencia mejor concluyente en el software NCSS®. Esto se realiza en un concentrado como el mostrado en la Tabla 5:

Tabla 5. Acomodo de información para el software NCSS®

<b>Materiales</b>	<b>Categoria</b>	<b>Totales</b>
<b>24282500</b>	Aceptado	16
<b>24282500</b>	Rechazado	10
<b>24290840</b>	Aceptado	25
<b>24290840</b>	Rechazado	1
<b>24268389</b>	Aceptado	26
<b>24268389</b>	Rechazado	0

(Elaboración propia)

De esta manera se procede a un análisis con el software de regresión logística usando una variable categórica que es la que se estudia, para este caso los materiales inspeccionados; una variable Y que define los hallazgos y datos encontrados en el estudio, siendo en este caso “aceptado” o “rechazado”; y por último tenemos una variable de frecuencia que es la que nos indica la cantidad de hallazgos que se tuvo en el mes para cada uno de los posibles resultados.

El software por default estaría realizando el análisis de regresión logística Categórica, que al momento de su interpretación se ve resumido ante que tan probable es que se realice el cambio de “0” a “1”, para este caso de estudio, que tan probable es que se detecte un material defectuoso en la siguiente inspección.

Para acomodar los datos en el software NCSS® se vacían las columnas antes mencionadas y se nombran de acuerdo con su mes de observación y año, adicional se coloca el tipo de información que contiene para ubicar de manera sencilla las variables al momento de correr el análisis. En la Figura 3 se tiene un ejemplo de la visualización de los datos ya acomodados en el software.

Figura 3. Carga de datos al software NCSS®

The screenshot shows the NCSS 2020 Data software interface. The main window displays a data table with 18 columns and 16 rows. The columns are labeled with month and year abbreviations (e.g., Jun18Mat, Jun18Cat, Jun18Tot) and the rows contain numerical data and status labels (Aceptado/Rechazado). The software interface includes a menu bar (File, Edit, View, Data, Analysis, Graphics, Tools, Window, Help) and a toolbar with various icons for file operations and data manipulation. The status bar at the bottom indicates 'Summary Statistics' and 'Last Row: 76, Col: 36'.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Name	Jun18Mat	Jun18Cat	Jun18Tot	Jul18Mat	Jul18Cat	Jul18Tot	Ago18Mat	Ago18Cat	Ago18Tot	Sep18Mat	Sep18Cat	Sep18Tot	Oct18Mat	Oct18Cat	Oct18Tot	Nov18Mat	Nov18Cat	Nov18Tot
Label																		
Data Type	Text	General	General	Text	General	General	Text	General	General	Text	General	General	Text	General	General	Text	General	General
Format																		
Value Labels																		
Value Order																		
Transformation																		
Note																		
Filter																		
1	11548287	Aceptado	25	24288511	Aceptado	22	24288511	Aceptado	21	24288511	Aceptado	26	24288511	Aceptado	24	24288511	Aceptado	27
2	11548287	Rechazado	1	24288511	Rechazado	4	24288511	Rechazado	6	24288511	Rechazado	2	24288511	Rechazado	7	24288511	Rechazado	3
3	24268025	Aceptado	25	24291933	Aceptado	26	24291933	Aceptado	27	24291933	Aceptado	28	24291929	Aceptado	31	24291929	Aceptado	29
4	24268025	Rechazado	1	24291933	Rechazado	0	24291933	Rechazado	0	24291933	Rechazado	0	24291929	Rechazado	0	24291929	Rechazado	1
5	24268026	Aceptado	26	24291929	Aceptado	25	24291929	Aceptado	24	24291929	Aceptado	24	24242622	Aceptado	31	24242622	Aceptado	30
6	24268026	Rechazado	0	24291929	Rechazado	1	24291929	Rechazado	3	24291929	Rechazado	4	24242622	Rechazado	0	24242622	Rechazado	0
7	24268389	Aceptado	26	24286406	Aceptado	26	24286406	Aceptado	27	24286406	Aceptado	27	24290840	Aceptado	24	24290840	Aceptado	25
8	24268389	Rechazado	0	24286406	Rechazado	0	24286406	Rechazado	0	24286406	Rechazado	1	24290840	Rechazado	7	24290840	Rechazado	5
9	24282500	Aceptado	16	24286407	Aceptado	26	24286407	Aceptado	27	24286407	Aceptado	28	24267848	Aceptado	30	24242620	Aceptado	30
10	24282500	Rechazado	0	24286407	Rechazado	0	24286407	Rechazado	0	24286407	Rechazado	0	24267848	Rechazado	1	24242620	Rechazado	0
11	24284994	Aceptado	26	24288501	Aceptado	26	24288501	Aceptado	26	24288501	Aceptado	28	24242620	Aceptado	31	24242621	Aceptado	30
12	24284994	Rechazado	0	24288501	Rechazado	0	24288501	Rechazado	1	24288501	Rechazado	0	24242620	Rechazado	0	24242621	Rechazado	0
13	24284996	Aceptado	26	24288150	Aceptado	26	24290840	Aceptado	19	24242622	Aceptado	28	24242621	Aceptado	31	24282500	Aceptado	20
14	24284996	Rechazado	0	24288150	Rechazado	0	24290840	Rechazado	8	24242622	Rechazado	0	24242621	Rechazado	0	24282500	Rechazado	10
15	24286406	Aceptado	26	24290840	Aceptado	24	24289021	Aceptado	24	24290840	Aceptado	22	24282500	Aceptado	27	24288243	Aceptado	22
16	24286406	Rechazado	0	24290840	Rechazado	2	24289021	Rechazado	3	24290840	Rechazado	6	24282500	Rechazado	4	24288243	Rechazado	8

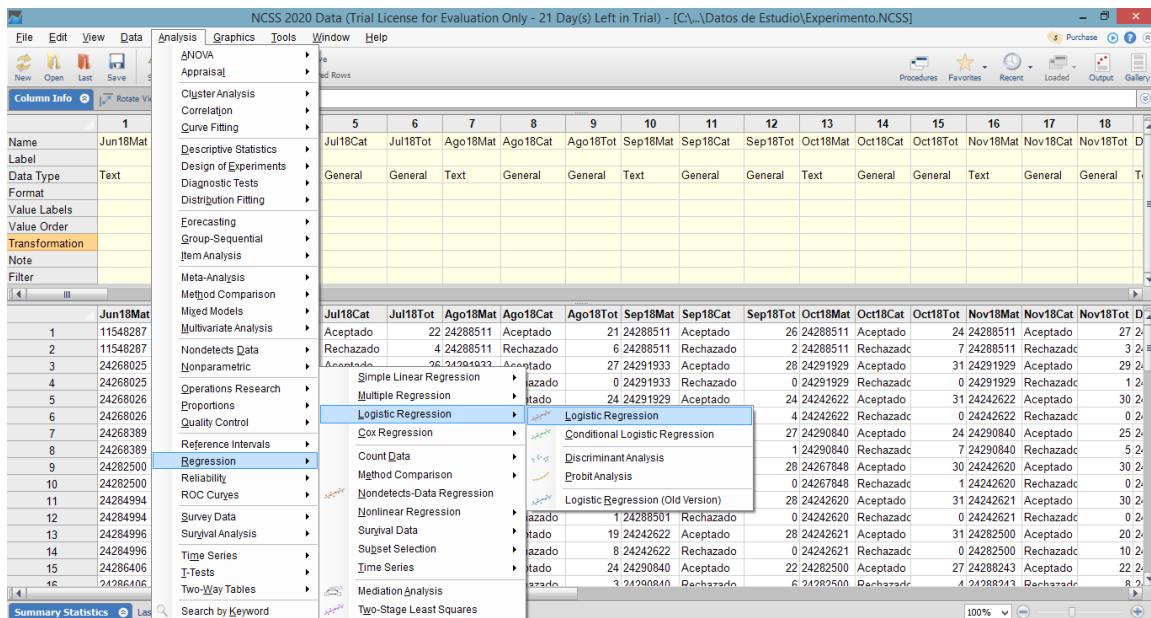
(Elaboración propia)

Para ingresar los datos de manera correcta y no sean confundidas con datos numéricos, se cambian el tipo de dato a texto para los numero de parte, y así se tomen como variables y no como valores.

## 5. ANÁLISIS DE REGRESIÓN LOGÍSTICA EN NCSS®

Para realizar el análisis de regresión logística en el software NCSS® se comienza habilitando la plantilla de regresión logística el cual se debe de llenar correctamente para evitar discrepancias. A este se accede desde el menú de Analysis > Logistic Regression > Logistic Regression, tal como se puede apreciar en la imagen siguiente (Ver Figura 4).

Figura 4. Seleccionar análisis de regresión logística en software NCSS®

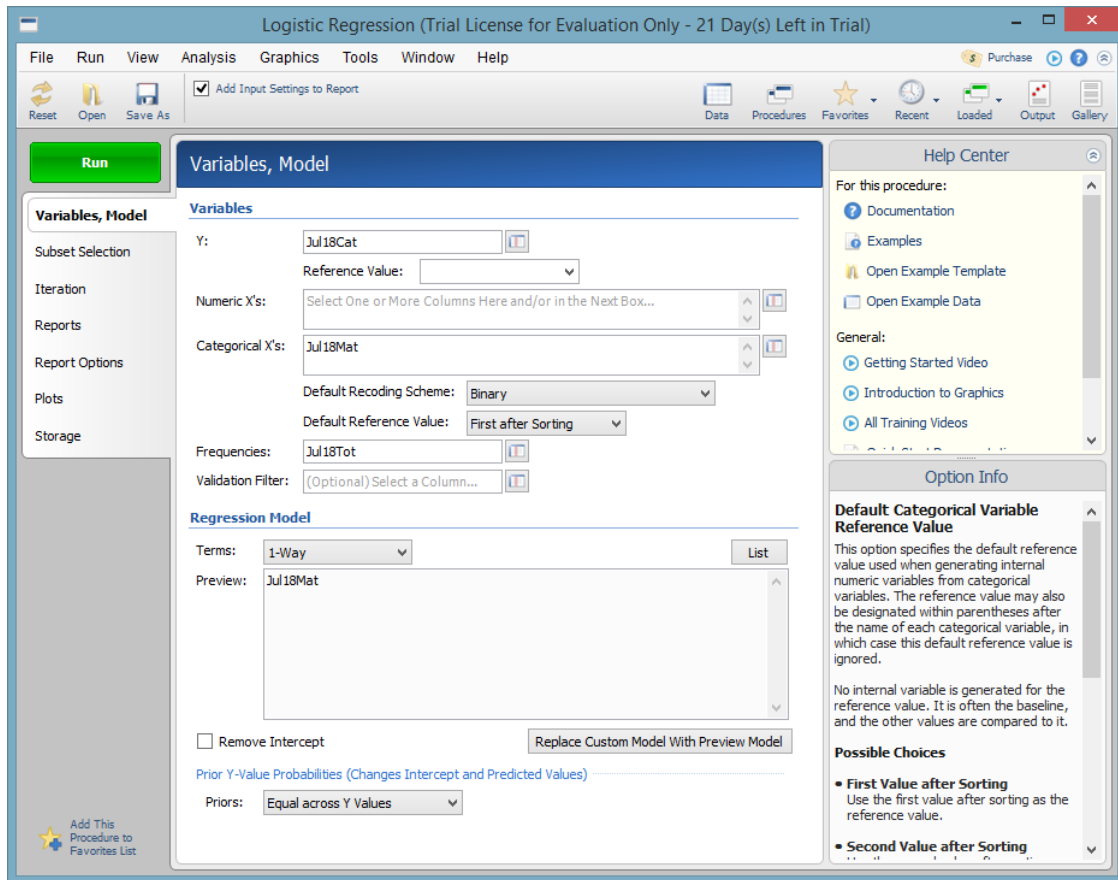


(Elaboración propia)

Cuando se tiene abierto la plantilla para colocar las variables a analizar, hay que realizar el vaciado de las variables en los campos correctos para que se ejecute un análisis de regresión correcto. Tal como se presenta en la imagen siguiente (Ver Figura 5).



Figura 5. Selección de variables para análisis de regresión logística en software NCSS®



(Elaboración propia)

Los datos, como ya se mencionó anteriormente, se acomodan de tal manera que se pudieran identificar de manera sencilla para el campo donde son requeridos.

En el campo de variable Y se colocan aquellas columnas donde se encuentran las categorías de Aceptado y Rechazado, debido a que esta será la variable que defina su estatus como pieza rechazada.

En el campo de Categorical X, se coloca la columna que contiene los materiales o números de parte que se están inspeccionando, ya que estos son las variables que pueden tener las categorías de Aceptado o Rechazado.

Por último, en el campo de frecuencias se coloca la columna que contenga los totales que se tuvieron de los hallazgos detectados en los datos. El software, al

colocar una variable en "Categorical X", asume de manera predeterminada que es un estudio binario, que solo contiene dos casos posibles de estado entre las observaciones. El estudio es de carácter binario al tener solo resultados de "aceptado" y "rechazado" por lo que se dejara ese campo tal y como se desglosa de manera predeterminada en el software NCSS®. Cabe resaltar que para que este tipo de análisis pueda realizarse, es requerida una columna de totales para sus frecuencias, ya que de lo contrario los resultados obtenidos carecerían de sentido para su inferencia.

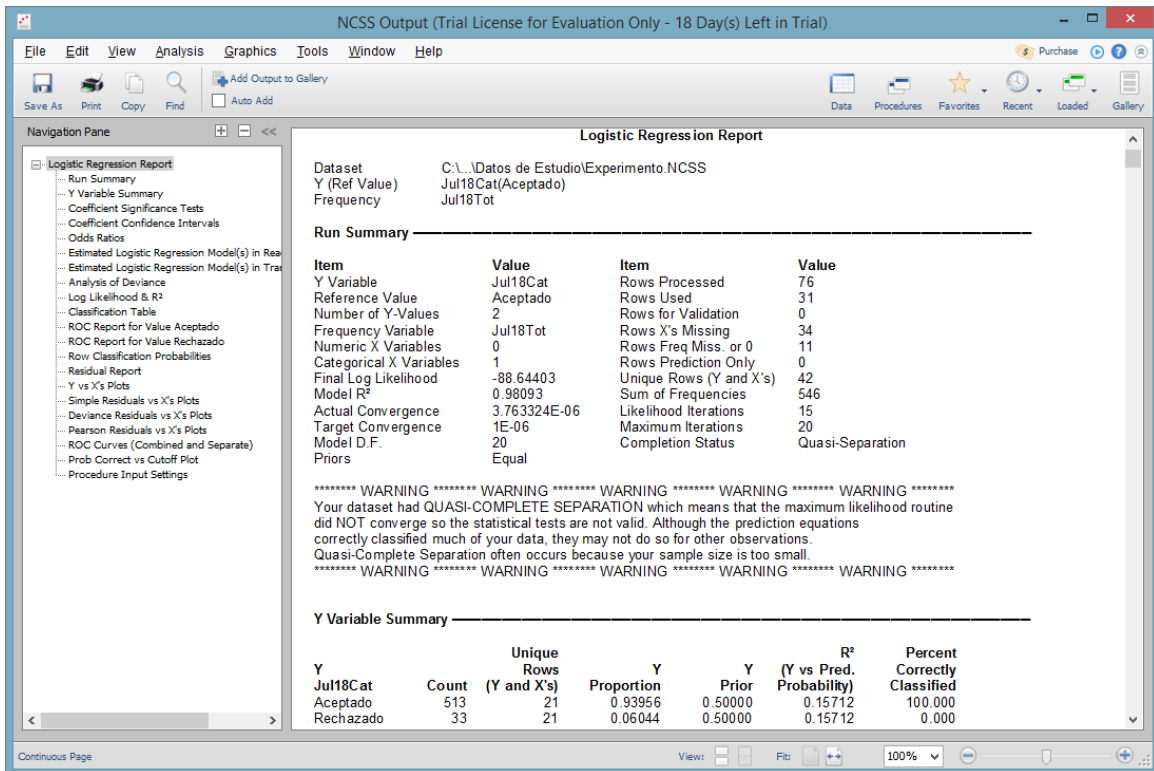
Una vez llenados todos los campos necesarios de la plantilla, se presiona el botón "Run" para que nos entregue los resultados de análisis de regresión logística. Ya que se realiza la división de las columnas de datos por mes, se obtiene un análisis de regresión logística por mes, por lo que se ejecuta todo lo antes mencionado para todos los meses donde se tienen observaciones.

Al tener por mes el análisis se puede ver de manera asertiva en el software, cuáles son los materiales o números de parte que tienen una mayor incidencia en el modelo, o bien, cuáles son los coeficientes para tener una mejor conclusión del análisis. De igual manera se puede definir si el material o número de parte es significativo para seguirse inspeccionado y poder destinar ese recurso para otras actividades del área.

## RESULTADOS

Al momento de que el software NCSS® termina de hacer el análisis de regresión logística, se desglosa un documento con todo el resumen de información que calcula la hacer un análisis categórico y binario. A pesar de contener mucha información y de manera detallada, se hace el enfoque a cierta información arrojada por el software para que se pueda hacer una inferencia con el contexto de la investigación, y simplificando la lectura de datos para concluir en menos pasos o teniendo menos condicionales.

Figura 6. Interfaz de resultados software NCSS®



(Elaboración propia)

Tomando de referencia la imagen anterior Figura 6, se visualiza el reporte arrojado por el software con la información del análisis efectuado. En los siguientes capítulos

se hará referencia a los que se les está considerando con mayor énfasis para interpretar el análisis de regresión logística.

## 6.1. RESULTADOS OBTENIDOS DE DATOS MENSUALES

### 6.1.1. Run Summary

En esta sección del reporte se pueden visualizar el total de datos que se usaron para hacer el análisis y cuáles son las variables que se están mencionando para los siguientes resúmenes. Para referencia visualizar Figura 7.

Figura 7. Sección "Run Summary" dentro de los resultados de software NCSS®.

Logistic Regression Report			
Dataset	C:\_Datos de Estudio\Experimento.NCSS		
Y (Ref Value)	Jul18Cat(Aceptado)		
Frequency	Jul18Tot		
Run Summary			
Item	Value	Item	Value
Y Variable	Jul18Cat	Rows Processed	76
Reference Value	Aceptado	Rows Used	31
Number of Y-Values	2	Rows for Validation	0
Frequency Variable	Jul18Tot	Rows X's Missing	34
Numeric X Variables	0	Rows Freq Miss. or 0	11
Categorical X Variables	1	Rows Prediction Only	0
Final Log Likelihood	-87.94544	Unique Rows (Y and X's)	42
Model R <sup>2</sup>	0.29409	Sum of Frequencies	546
Actual Convergence	3.793218E-06	Likelihood Iterations	15
Target Convergence	1E-06	Maximum Iterations	20
Model D.F.	21	Completion Status	Quasi-Separation
Priors	Equal		
***** WARNING ***** WARNING ***** WARNING ***** WARNING ***** WARNING *****			
Your dataset had QUASI-COMPLETE SEPARATION which means that the maximum likelihood routine did NOT converge so the statistical tests are not valid. Although the prediction equations correctly classified much of your data, they may not do so for other observations.			
Quasi-Complete Separation often occurs because your sample size is too small.			
***** WARNING ***** WARNING ***** WARNING ***** WARNING ***** WARNING *****			

(Elaboración propia)

### 6.1.2. Y Variable

Se debe detectar que se está usando la variable correcta, para este caso de estudio, es la variable que indica los tipos de hallazgos para los materiales inspeccionados (Ej. Jul18Cat).

“Reference Value” es el valor que de manera predeterminada toma como valor “0” para efectuar el análisis. Ya que en este estudio se está tomando el valor “0”

para material "Aceptado" y el valor "1" para material "Rechazado", se debe asegurar que el valor de referencia que ahí aparezca sea "Aceptado".

### **6.1.3. Frequency Variable**

Hay que asegurar que se está tomando la columna con las cantidades totales de todos los hallazgos en los datos, es decir, los totales de días que se encuentran materiales como "Aceptado" y "Rechazado" respectivamente (Ej. Jul18Tot).

### **6.1.4. Unique Rows (Y and X's)**

Aquí se visualiza el total de filas, o para el contexto del estudio, la cantidad de materiales "Aceptados" y "Rechazados". Se debe de asegurar que está tomando todas las filas del mes ya que la cantidad de materiales o números de parte que se inspeccionan en el periodo observado varia con respecto al mes, ya que hay materiales que salen o entran a lista de intercepción cambiando el total de filas a las que se les realizara el análisis de regresión.

A lo largo de las observaciones mensuales se ve un comportamiento poco uniforme en los primeros meses observados, haciendo que la cantidad de materiales en lista de intercepción aumente para los siguientes meses, hasta llegar a un punto donde se estabilizan la cantidad de materiales manteniendo estable los modelos siguientes. Es por lo anterior dicho, que en los primeros meses se encuentra en esta sección de cada mes, más cantidad de filas que se transaccionan haciendo más largo en cada ocasión los reportes de resultados.

### **6.1.5. Y Variable Summary**

En esta sección se encuentra el total de las observaciones que se realizaron para el total de los datos, haciendo así el software la segregación de datos para cada hallazgo que este haya detectado. Es importante resaltar que para efecto de este estudio deben de aparecer dos posibles observaciones, "Aceptado" y "Rechazado"; caso contrario es necesario regresar al software para ver la

distribución de los datos y que estos no contengan errores tanto de escritura como de llenado. Para referencia visualizar Figura 8.

Figura 8. Sección "Y Variable Summary" dentro de los resultados del software NCSS®

Y Variable Summary						
Y	Count	Unique Rows (Y and X's)	Y Proportion	Y Prior	R <sup>2</sup> (Y vs Pred. Probability)	Percent Correctly Classified
Jul18Cat						
Aceptado	513	21	0.93956	0.50000	0.12079	70.370
Rechazado	33	21	0.06044	0.50000	0.12079	90.909
Total	546	42				71.612

(Elaboración propia)

Para esta parte del resumen solo hay que tener correctamente identificado el total Count y Unique Rows. Count nos apoya para ver cuantas veces se "Acepta" o se "Rechaza" un material y Unique Rows para saber que se están leyendo todos los datos de manera correcta para no generar un sesgo a alguna de las categorías establecidas.

En los meses observados para esta investigación, se tiene la misma tendencia que en la sección anterior, donde se va incrementando los totales para cada mes por contener una mayor cantidad de materiales en la lista de intercepción. En los meses se detecta de igual manera que los resultados en los porcentajes de rechazos detectados influyen para tener resultados claros y concisos para los siguientes resultados de las secciones que más adelante se detallan.

### 6.1.6. Coefficient Significance Tests

En esta parte del resumen es donde se encuentran los coeficientes Betas de todos los materiales que se inspeccionan durante el mes. Esta sección se ve con mayor detalle y análisis para cada mes, ya que contiene detalles que pueden influir para el siguiente mes a inspeccionar y los materiales que saldrán o se mantendrán en lista de intercepción. Para referencia ver Figura 9.

Figura 9. Sección "Coefficient Significance Tests" dentro de los resultados del software NCSS®

Coefficient Significance Tests					
Independent Variable X	Regression Coefficient b(i)	Standard Error Sb(i)	Wald Z-Value H0: $\beta=0$	Wald P-Value	Odds Ratio Exp(b(i))
(Jul18Mat=24267696)	-12.45913	0.17957	-69.383	0.00000	0.00000
(Jul18Mat=24267848)	-2.03688	1.62906	-1.250	0.21118	0.13043
(Jul18Mat=24268025)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24268026)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24269647)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24271482)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24273971)	-3.21888	0.98058	-3.283	0.00103	0.04000
(Jul18Mat=24278085)	-2.03688	1.62906	-1.250	0.21118	0.13043
(Jul18Mat=24282500)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24285405)	-3.21888	0.98058	-3.283	0.00103	0.04000
(Jul18Mat=24286406)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24286407)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24288150)	-15.20289	0.00255	-5966.316	0.00000	0.00000
(Jul18Mat=24288243)	-2.48491	1.35873	-1.829	0.06742	0.08333
(Jul18Mat=24288501)	-15.20289	0.00255	-5966.316	0.00000	0.00000

(Elaboración propia)

Adicional en esta sección se pueden empezar a identificar los Odds Ratio calculados para cada uno de los materiales, que de igual manera con los coeficientes Betas, cambian con respecto al mes ya que, al no ser la misma cantidad de hallazgos ni materiales, lleva su detallada inferencia para determinar el estatus de cada material a salir o mantenerse en lista de intercepción.

### 6.1.7. Coefficient Confidence Intervals

De manera muy similar a la sección anterior, aquí se pueden observar los coeficientes Betas que se obtuvieron en el análisis de regresión logística. La diferencia radica en que aquí se muestran los intervalos de confianza para cada uno de los coeficientes obtenidos, así como la desviación estándar calculada para cada uno de ellos. Estos se podrán apreciar de mejor manera con respecto al contexto del mes, puesto que, vistos de manera general, no aporta lo suficiente para tener una conclusión o inferencia de impacto para una siguiente toma de decisión. Para referencia visualizar la imagen anterior (Ver Figura 10).

Figura 10. Sección "Coefficient Confidence Intervals" dentro de los resultados de software NCSS®.

Independent Variable X	Regression Coefficient b(i)	Standard Error Sb(i)	Lower 95% Confidence Limit	Upper 95% Confidence Limit	Odds Ratio Exp(b(i))
(Jul18Mat=24267696)	-12.45913	0.17957	-12.81108	-12.10717	0.00000
(Jul18Mat=24267848)	-2.03688	1.62906	-5.22979	1.15602	0.13043
(Jul18Mat=24268025)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24268026)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24269647)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24271482)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24273971)	-3.21888	0.98058	-5.14078	-1.29697	0.04000
(Jul18Mat=24278085)	-2.03688	1.62906	-5.22979	1.15602	0.13043
(Jul18Mat=24282500)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24285405)	-3.21888	0.98058	-5.14078	-1.29697	0.04000
(Jul18Mat=24286406)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24286407)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24288150)	-15.20289	0.00255	-15.20789	-15.19790	0.00000
(Jul18Mat=24288243)	-2.48491	1.35873	-5.14797	0.17816	0.08333
(Jul18Mat=24288501)	-15.20289	0.00255	-15.20789	-15.19790	0.00000

(Elaboración propia)

Los valores cambian de acuerdo con los hallazgos de rechazos que se tengan durante el mes observado, por lo que en cada mes se tiene un valor diferente aun teniendo una cantidad de hallazgos de rechazo de 0. Los valores que se esperan ver en esta sección son lo más bajo posible, e inclusive, menores a cero, debido a que si se tiene un valor negativo de estos coeficientes se tiene un impacto menor al modelo y puede incluso considerarse para retirarse del modelo, debido a que su significancia no tiene un impacto fuerte dentro del modelo. Por el lado contrario, cuando se tienen valores muy altos y positivos, estos nos da a entender que los materiales con esta condición, de encontrarse un hallazgo, incrementara de manera significativa el modelo incrementando su valor y por ser parte de su cálculo, el indicador, entonces estos materiales se deben considerar para mantenerse en lista de intercepción o bien obtener la mayor cantidad de información con las inspecciones y proveedores de todos sus modos de fallas para validar su riesgo e impacto a las líneas de producción.



### 6.1.8. Odds Ratio

En esta sección se encuentran los Odds Ratios calculados para cada uno de los materiales, tal y como en las secciones anteriores, pero adicionando el intervalo de confianza para cada uno de estos. Al interpretar un Odds Ratio es bastante recomendable hacerlo acompañado de su intervalo de confianza, ya que si se tiene un rango bastante extenso podemos tener altas probabilidades de realizar una decisión con mucho factor de riesgo. Para referencia visualizar Figura 11.

Figura 11. Sección "Odds Ratio" dentro de los resultados de software NCSS®

Independent Variable X	Regression Coefficient b(i)	Odds Ratio Exp(b(i))	Lower 95% Confidence Limit	Upper 95% Confidence Limit
(Jul18Mat=24267696)	-12.45913	0.00000	0.00000	0.00001
(Jul18Mat=24267848)	-2.03688	0.13043	0.00535	3.17727
(Jul18Mat=24268025)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24268026)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24269647)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24271482)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24273971)	-3.21888	0.04000	0.00585	0.27336
(Jul18Mat=24278085)	-2.03688	0.13043	0.00535	3.17727
(Jul18Mat=24282500)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24285405)	-3.21888	0.04000	0.00585	0.27336
(Jul18Mat=24286406)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24286407)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24288150)	-15.20289	0.00000	0.00000	0.00000
(Jul18Mat=24288243)	-2.48491	0.08333	0.00581	1.19502
(Jul18Mat=24288501)	-15.20289	0.00000	0.00000	0.00000

(Elaboración propia)

En el comportamiento de los meses observados, en los casos que se tiene una mayor variabilidad dentro de las observaciones se tienen intervalos más grandes entre sus límites inferiores y superiores, afirmando de igual manera, que una siguiente estimación podrá tener una variación muy grande o pequeña la cual se toma en cuenta para empezar a decidir si el material en cuestión se mantiene o se retira de la lista de intercepción.

El valor que se desea para esta sección es 0 o lo más cercano posible, ya que lo que nos indica, es que se esperarían 0 rechazos para las siguientes inspecciones a realizar, afirmando que el material siendo analizado contiene un riesgo bajo de seguir encontrando rechazos. Por el lado contrario, desde que se tiene un valor de Odds Ratio de 1 ya tenemos un riesgo de seguir encontrando cuando menos un hallazgo de rechazo para las siguientes inspecciones a realizar, y de manera directa, mientras más grande sea este número, mayor es el riesgo de rechazo que tiene el material para todas las inspecciones que se programen.

### 6.1.9. Estimated Logistic Regression Model(s) in Reading Form

En esta importante sección del reporte se indica el modelo que se generó al hacer el análisis de regresión logística con los datos del mes evaluado. Es de percatarse que aquí se incluyen los coeficientes Betas anteriormente calculados, pero adicional, se juntan para hacer una ecuación de estimación para modelos de regresión logística.

Figura 12. Sección "Estimated Logistic Regression Model(s) In Reading Form" dentro de los resultados de del software NCSS®

Estimated Logistic Regression Model(s) in Reading Form

---

Model for Logit(Jul18Cat) = XB when Jul18Cat = Rechazado  
 12.46 \* (Jul18Mat=24267696) - 2.04 \* (Jul18Mat=24267848) - 15.20 \* (Jul18Mat=24268025) - 15.20 \* (Jul18Mat=24268026)  
 - 15.20 \* (Jul18Mat=24269647) - 15.20 \* (Jul18Mat=24271482) - 3.22 \* (Jul18Mat=24273971) - 2.04 \* (Jul18Mat=24278085)  
 - 15.20 \* (Jul18Mat=24282500) - 3.22 \* (Jul18Mat=24285405) - 15.20 \* (Jul18Mat=24286406) - 15.20 \*  
 (Jul18Mat=24286407) - 15.20 \* (Jul18Mat=24288150) - 2.48 \* (Jul18Mat=24288243) - 15.20 \* (Jul18Mat=24288501) - 1.70 \*  
 (Jul18Mat=24288511) - 1.20 \* (Jul18Mat=24289021) - 2.48 \* (Jul18Mat=24290840) - 3.22 \* (Jul18Mat=24291929) - 15.20 \*  
 (Jul18Mat=24291933)

(Elaboración propia)

Para el contexto manejado en este estudio, esta ecuación y modelo es nuestra guía para saber cuáles materiales y números de parte se seguirán inspeccionando para hacer una estimación del siguiente paso, haciendo una suposición de "Rechazo" antes de tener el lote siguiente a inspeccionar. Ver Figura 12 para referencia.

Este modelo apoya de una manera muy significativa al nuevo indicador propuesto debido a su fuerte impacto de los datos observados, y a su vez, la significación que tienen cada uno de estos en la ecuación final para determinar una probabilidad de rechazo.

El nuevo indicador es calculado en base a un modelo de este tipo, solo que el de esta sección se estimó para generalizar el mes. Por lo que puede ser considerado un indicador mensual de la cantidad estimada de rechazos a tener en la siguiente inspección.

El software NCSS® de manera predeterminada coloca en el modelo los Betas en orden ascendente, por lo que no hay que hacer inferencia por el acomodo que se tienen los coeficientes Betas en el modelo mostrado en el reporte de esta sección.

Durante las observaciones de los meses se tiene un modelo más complicado dependiendo de las cantidades de materiales que se encuentran en lista de intercepción, haciendo el modelo de regresión logística estimado más complicado y largo mientras se tenga una mayor cantidad de materiales y por el lado contrario, cuando sea menor la cantidad de materiales en inspección el modelo permanecerá más simple de visualizar y de analizar para cada uno de los coeficientes mostrados.

Aun siendo el modelo de regresión logística estimado simple o complicado, es importante se tome la inferencia que se puede obtener, porque los coeficientes Betas estimados dan información suficiente para empezar proponer a retirar materiales de la lista de intercepción al no presentar un impacto significativo para el modelo. Haciendo énfasis en lo anterior dicho, es común empezar a proponer los materiales de lista de intercepción a salir que tengan valores menores a cero, puesto que visualmente en el modelo, se pueden interpretar como riesgo menor de desestabilizar el modelo en la siguientes estimación, en otras palabras, si un coeficiente Beta de un material, es menor a cero, aun teniendo hallazgos de rechazos, estaría otorgando un valor de cálculo negativo al modelo, haciendo que su valor no tenga impacto al calcularse al final, recordando que, el modelo arroja

como resultado una probabilidad resultante, la cual no puede ser de valores menores a cero.

### 6.1.10. Row Classification Probabilities

En esta sección es donde encontramos la información más significativa de todo el reporte de resultados. Dentro de esta sección se encuentran las probabilidades de que un material sea "Aceptado" o "Rechazado" para cada una de las filas que analiza el software. Para referencia visualizar Figura 13.

Figura 13. Sección "Row Classification Probabilities" dentro de los resultados del software NCSS®.

Row Classification Probabilities			
Row	Actual Jul18Cat	Estimated Prob. in Aceptado	Estimated Prob. in Rechazado
1*	Aceptado	0.26134	0.73866
2	Rechazado	0.26134	0.73866
3	Aceptado	1.00000	0.00000
4*	Rechazado	1.00000	0.00000
5	Aceptado	0.61659	0.38341
6*	Rechazado	0.61659	0.38341
7	Aceptado	1.00000	0.00000
8*	Rechazado	1.00000	0.00000
9	Aceptado	1.00000	0.00000
10*	Rechazado	1.00000	0.00000
11	Aceptado	1.00000	0.00000
12*	Rechazado	1.00000	0.00000
13	Aceptado	1.00000	0.00000
14*	Rechazado	1.00000	0.00000
15*	Aceptado	0.43564	0.56436
16	Rechazado	0.43564	0.56436
17	Aceptado	1.00000	0.00000
18*	Rechazado	1.00000	0.00000
19	Aceptado	1.00000	0.00000
20*	Rechazado	1.00000	0.00000
21*	Aceptado	0.17657	0.82343
22	Rechazado	0.17657	0.82343
23	Aceptado	1.00000	0.00000
24*	Rechazado	1.00000	0.00000
25	Aceptado	1.00000	0.00000
26*	Rechazado	1.00000	0.00000
27*	Aceptado	0.09332	0.90668
28	Rechazado	0.09332	0.90668
29*	Aceptado	0.33029	0.66971
30	Rechazado	0.33029	0.66971
31	Aceptado	0.61659	0.38341
32*	Rechazado	0.61659	0.38341
33*	Aceptado	0.33029	0.66971
34	Rechazado	0.33029	0.66971
35	Aceptado	1.00000	0.00000
36*	Rechazado	1.00000	0.00000
37	Aceptado	1.00000	0.00000
38*	Rechazado	1.00000	0.00000
39*	Aceptado	0.43564	0.56436
40	Rechazado	0.43564	0.56436
41	Aceptado	0.61659	0.38341
42*	Rechazado	0.61659	0.38341

(Elaboración propia)

Para tener una mejor interpretación de estas probabilidades, debe tomarse en cuenta el acomodo que se hizo de los datos para que se pudiera hacer este tipo de análisis. Recordando así, que se colocaron los materiales para cada Categoría de "Aceptado" y "Rechazado", dando como resultado dos probabilidades resultantes iguales para cada par de filas, por lo que para revisar qué número de parte se está tratando con esas probabilidades, es necesario revisar dentro de los datos vaciados en el software NCSS® la cantidad de filas que hace referencia esta sección. El software por su parte realiza una notoria división de datos alertando que se repiten con un asterisco a lado de la fila donde se detecta un valor duplicado.

Esta probabilidad resultante calculada se estima con un modelo de regresión logística, pero a diferencia del modelo en la sección 6.1.8 ESTIMATED LOGISTIC REGRESSION MODEL(S) IN READING FORM, este modelo se hace para cada uno de los números de parte o materiales que se inspeccionaron. Por esa razón el cálculo se hace por fila, discriminando si el resultado puede hacer referencia a todo un grupo o una categoría de datos.

Esta probabilidad resultante es la que se maneja como el nuevo indicador propuesto para cada uno de los materiales inspeccionados, teniendo como objetivo calificar el material durante el mes que se hizo el análisis. Cabe destacar que en esta parte del análisis se realiza un corte al mes para determinar si el material continua o sale de lista de intercepción para seguir o no con las inspecciones, sin embargo, este nuevo indicador puede calcularse incluso de manera diaria, donde día que vaya pasando se vaya alimentando en la recopilación de datos y sin esperar a que pase más tiempo, tomar la decisión del material conociendo siempre con anterioridad, la estimación que tendrá en el siguiente paso o inspección.

El valor para calificar el material con el nuevo indicador será de 90% como "Sin Riesgo de encontrar Rechazo" (Verde), entre 50% y 89% como "Riesgo Medio de Rechazo" (Amarillo) y menor a 50% como "Riesgo Alto de Rechazo" (Rojo), definiendo así, que todo material con un valor de indicador ERP bajo se siga inspeccionando ya que aun corre un alto riesgo para encontrar un rechazo en la siguiente inspección. Por su lado contrario, si el indicador ERP calculado es alto, el

material se considera con muy poca probabilidad de encontrar un rechazo en la siguiente inspección realizada.

Los indicadores calculados para cada uno de los meses observados en esta investigación se despliegan en una tabla con los resultados obtenidos al término de este mes, esto con el fin de mostrar de manera más gráfica que resultado obtuvo cada material de la lista de intercepción. Al mostrar estos resultados obtenidos a todos los niveles de planta se da por concluido el análisis de regresión logística para el mes y así pasar a las siguientes acciones tanto de proveedor como del mismo departamento de calidad recibo y calidad proveedores. Cabe destacar que en algunos casos los materiales tienen conclusiones que pueden no verse reflejado directamente con el resultado obtenido, debido a que estos materiales tienen una planeación siguiente para las siguientes inspecciones. Ejemplificando algunos casos tenemos materiales que son parte de pruebas piloto que deben inspeccionarse en un periodo de tiempo establecido para asegurar que no se encuentra ningún modo de falla tal que ponga en riesgo el introducirse a las líneas de producción. Otro ejemplo es para los materiales que tienen acciones pendientes con proveedor por tener un modo de falla que no puede ser soportado por inspecciones del departamento de calidad recibo por lo que debe seguirse inspeccionando por residentes del proveedor para encontrar más información de los modos de falla y el riesgo que tendrían a las líneas de producción y producto terminado.

Para mostrar de una manera visual los resultados y los indicadores para cada uno de los números de parte a revisar, se colocan en una tabla como la siguiente (Ver Tabla 6).

Tabla 6. Resultados del nuevo indicador ERP

Numero de Parte	Aceptado	Rechazado	Prob. Reject	ERP (1-P[Y=1])
24288511	22	4	0.73866	0.26134
24291933	26	0	0	1
24291929	25	1	0.38341	0.61659
24286406	26	0	0	1
24286407	26	0	0	1
24288501	26	0	0	1
24288150	26	0	0	1
24290840	24	2	0.56436	0.43564
24268026	26	0	0	1
24268025	26	0	0	1
24289021	20	6	0.82343	0.17657
24271482	26	0	0	1
24282500	26	0	0	1
11548287	16	10	0.90668	0.09332
24267848	23	3	0.66971	0.33029
24285405	25	1	0.38341	0.61659
24278085	23	3	0.66971	0.33029
24267696	26	0	0	1
24269647	26	0	0	1
24288243	24	2	0.56436	0.43564
24273971	25	1	0.38341	0.61659

(Elaboración propia)

## 6.2. MODELO DE REGRESIÓN OBTENIDO

El modelo estimado de regresión logística cambia a lo largo de los meses debido a que no se tienen la misma cantidad de materiales, así como la misma cantidad de hallazgos de rechazado para cada lista de intercepción observada. Sin embargo, en todos los casos se obtiene el modelo con la misma estructura que se puede interpretar en cualquiera de los meses incluso sin saber el origen o cantidad de los hallazgos de rechazado obtenidos.

El modelo de regresión de manera general queda de la siguiente manera:

$$\text{Modelo de Regresion} = \pm \text{Intercepto} \pm \beta_{\text{Material } 1} * \text{Valor Material } 1 \pm \beta_{\text{Material } 2} * \text{Valor Material } 2 \pm \beta_{\text{Material } 3} * \text{Valor Material } 3 \dots \pm \beta_{\text{Material } n} * \text{Valor Material } n \quad (22)$$

Donde n es el número de materiales que se encuentran inspeccionándose en la lista de intercepción para el departamento de calidad recibo.

De una manera general el modelo se compone en primer lugar por el intercepto, que es el número que en primer lugar estará realizando un incremento o

decremento en todo el modelo conforme avanzan las inspecciones. De manera predeterminada el software acomoda los materiales de manera ascendente, por lo que, al momento de calcular el modelo de regresión logística, el primer valor que toma como intercepto es el primer número de parte con el menor número acomodado de manera ascendente. Dicho lo anterior, se pueden obtener conclusiones sobre el intercepto hablando así del material que tiene un menor valor, por lo que si se tiene un intercepto negativo, significa que no tiene un impacto fuerte en el modelo, ya que si el valor que se detecta llegase a ser 1, estaría realizando un decremento en el modelo de tal manera que no estaría dando información a la probabilidad resultante, de manera contraria, al ser positivo se dice que tiene un impacto fuerte ya que si se tiene un hallazgo en el valor del material, estaría incrementando el valor de la probabilidad resultante.

De igual manera para los demás Betas de los demás materiales, se tiene la misma influencia en sus valores negativos y positivos obtenidos, sin embargo, para estos valores de coeficientes Beta también se puede obtener un valor de 0, y esto lo que indica en las conclusiones es que tiene una significancia media para el modelo puesto que sin importar el valor que se obtenga 1 o 0, se seguirá aumentando la misma cantidad de probabilidad resultante en cada observación siguiente que se realice.



## CONCLUSIONES

A lo largo de la investigación en el lapso observado, se encuentra de gran utilidad el detalle de cada uno de los resultados obtenidos, ya que, se complementan entre sí para poder tomar decisiones de una manera más objetiva y con una mayor cantidad de información que respalda las mismas decisiones y las acciones a tomar en el siguiente paso para cualquier de los materiales que entrasen a la lista de intercepción del departamento de calidad recibo.

Se determina ya en esta parte de la investigación, que se cumple el porcentaje de reducción del 50% planteado en la hipótesis, puesto que se tiene con anterioridad al periodo de observación un promedio de 30 materiales en la lista de intercepción siendo inspeccionados durante todo el mes y al final de la implementación se tiene un promedio de 15 materiales. Aunque en el periodo de implementación no se ve una diferencia muy amplia con respecto a la esperada, se logra acelerar la curva de aprendizaje mediante entrenamientos constantes hacia el personal de apoyo, así como el manejo correcto de la información para todos los niveles. Por ejemplo, la correcta definición del indicador obtenido y sus herramientas facilita ver a los ingenieros de calidad recibo y calidad proveedores a tener una percepción de los proveedores que presentan materiales con características fuera de especificación, logrando la respuesta pronta de proveedor para evitar mandar defectos y evitar hacer cualquier tipo de inspección en planta para los materiales detectados. También, facilita la toma de decisiones para materiales de bajo riesgo, logrando únicamente las tareas y actividades de contención, pero que no requieren una inspección prolongada o con tiempo especializado.

Gracias a la correcta implementación de estas herramientas y el indicador, se puede ahora manejar una cantidad menor de personas para inspecciones, así como las horas de trabajo requeridas, adicional, la optimización de recursos humanos para actividades adicionales en el área de calidad recibo como inventarios, cuarentena y disposición de material. Sumado a lo anterior, los tiempos de suministro hacia las líneas de producción indirectamente se reducen en un 50%, por lo que el riesgo de impactar a una línea de ensamble se minimiza ya que su inventario disponible crezca para ser usado en el momento requerido.

La interpretación de datos es fundamental para poder dar una proyección acertada, así como para poder tomar una decisión de manera correcta. Sin embargo, a pesar de no tener una cantidad de experiencia con la estadística y el manejo de información, se encuentra en todos los niveles de la empresa una idea y estimación de los materiales que tuvieron mayor problema dentro de las inspecciones en el mes, por lo que la explicación de estos indicadores ya se podía inferir para proyectar un siguiente paso posible aun sin revisar el detalle de los resultados para ese mes. Esto quiere decir, que la información brindada de estos indicadores y el modelo de regresión logística, están agilizando la toma de decisiones y ayudando a que en todos los niveles se tenga una ligera proyección de que poder seguir haciendo y pidiendo al proveedor para que el indicador llegue a su valor óptimo.

Concluyendo sobre los resultados obtenidos a través de los meses observados de esta investigación, se tienen varios detalles que se pueden validar para mejorar el enfoque en futuros resultados y hacer inferencias más asertivas. Como primer ejemplo tenemos el caso de los Odds Ratio, que para los casos más obvios que no se tuvieron hallazgos de rechazos durante el mes observado, se tiene un valor de 0 o muy cercano a él, puesto que se refiere a que se espera seguir teniendo una cantidad de 0 rechazos para las siguientes observaciones. Caso contrario sucede con los materiales que si tuvieron hallazgos de rechazos en el mes, haciendo que cada vez que se encuentren más rechazos, el Odds Ratio de esos materiales se irá incrementando de manera proporcional a la cantidad de rechazos que se encuentren, de esta manera cuando se tiene Odds Ratio inclusive de 10000+ dentro de los resultados del software, se puede inferir que se tiene una cantidad muy grande de rechazos a detectar en las siguientes inspecciones a realizar.

De una manera muy similar, se tiene un resultado para los coeficientes Betas estimados para los materiales de una manera que se correlacionan con la cantidad de rechazos detectados durante el mes. Cuando el material observado no contiene ningún hallazgo dentro del mes, este tiende a ser menor a 0 alejándose cada vez más si se tienen en su mayoría de las observaciones, materiales con ningún hallazgo de rechazos. En otras palabras, si en el mes se presentan materiales

de la lista de intercepción que no tienen hallazgos de rechazo, estos serán negativos y mientras crezca la cantidad de materiales sin rechazos, el valor de estos será cada vez más negativo con respecto a 0. Por el lado contrario, los materiales que contengan rechazos ya reflejan un valor mayor a 0 dando a entender que al seguir encontrando rechazos para esos materiales, se seguirá modificando el modelo final debido a que tienen un aporte significativo dependiendo del valor que tomase. Dicho de otra manera, para aquellos materiales con rechazos, mientras más rechazos se encuentran dentro de las observaciones del mes, se esperan valores altos mayores a 0 de la misma manera para sus coeficientes Betas, con los cuales se infiere que son los materiales que más están afectando al modelo del mes en investigación, por lo que sacarlos de la lista de intercepción o del modelo en cuestión, podría resultar en una alteración significativa para la siguiente estimación.

Los intervalos de confianza tanto para los Odds Ratio como para los coeficientes Betas se ven afectados por la cantidad de variación que tienen sus observaciones. Cuando se tiene una serie de valores estables para ambos casos de "Aceptados" y "Rechazados", sus intervalos no se extienden por mucho entre límite inferior y superior, ayudando así a dar inferencias de valores de estimación siguientes muy confiables (al 98%) de que no afectaran una estimación de valores siguientes. Por lo anterior dicho se encuentra también, que cuando se tiene un hallazgo en una serie que no podemos definir como estable, como de ejemplo, no encontrar nada de rechazos durante los primeros 28 días del mes, y encontrar rechazos en los últimos dos, genera una variación muy grande en sus observaciones, por lo que , sus intervalos de confianza se verán afectados estando muy separados los límites interiores y superiores para sus estimaciones de Odds Ratio y Coeficientes Beta, definiéndolos entonces, como materiales con un valor muy cambiante para las siguientes estimaciones que puede tener el riesgo de influir de manera significativa al modelo en una siguiente inspección.

Como parte final y que se considera más importante dentro del análisis de regresión logística para esta investigación y el nuevo indicador, se tiene la probabilidad resultante. Ayuda a estimar la probabilidad de un material a ser rechazado en la

siguiente inspección. Esta probabilidad es la que se determina como filtro final para tomar la decisión de sacar o mantener un material en lista de intercepción para que siga inspeccionándose.

De una manera directa a los indicadores antes mencionados dentro de los análisis de regresión logística, es común ver que la probabilidad resultante o de rechazo cambie con respecto a los hallazgos encontrados dentro de las observaciones, de manera que esta probabilidad tendrá como resultado más cercano a 100% mientras más hallazgos tenga de rechazos durante el mes, y por el lado contrario se acerca al 0% mientras menos rechazos se encuentren dentro de las observaciones. Es importante destacar que estadísticamente las probabilidades de 0% no es posible, y aunque dentro de los resultados se ven varios valores de 0%, el software NCSS® lo trunca por ser un número muy pequeño a 0% para visualizarlo de la mejor manera en el reporte. Con esta probabilidad de rechazo es con la que se calcula el nuevo indicador ERP, es por ello que el indicador está respaldado por un análisis estadístico de regresión logística.

El indicador ERP estimado al proyectar información a futuro, muchas áreas o usos pueden surgir para este nuevo indicador, y aunque el énfasis de esta investigación se enfoca para procesos de aceptación y rechazo, puede ser modificado de tal manera de ver afectaciones en otras áreas de planta con información importante que apoya a una mayor cantidad de decisiones.

Puede tener una dirección hacia proceso de garantías de materiales y productos calculando su probabilidad de falla estando ya en uso, proyectando incluso, probabilidades siguientes de falla.

Puede proyectar la confiabilidad de máquinas, herramientas, detectores en proceso de errores, tomando en cuenta las variables que están impactando al proceso, y aunque no se tenga una medición binaria como fue el caso de esta investigación, puede estimar una probabilidad para que se puedan categorizar y evaluar cada una de las maquinas, herramientas, etc., dado sus usos en el proceso y sus variables de mayor impacto.

Aun siendo un estimador muy eficiente de proyección hacia adelante, no se recomienda usarse en algunas especificaciones de algunos productos o materiales, sabiendo que las variables medidas tienen una desviación muy amplia, no puede brindar un estimador al modelo apropiado, dando estimaciones muy poco certeras al momento de ver las conclusiones. Si la dificultad de las inspecciones o validaciones dan camino a tener pocas observaciones se puede afectar en el resultado final, mientras más observaciones se tengan, se puede tener una mejor proyección del modelo que dará más información hacia adelante.

Estos modelos también pueden ayudar a reducir o ampliar un criterio de aceptación, observando que las variables tal vez no tengan un impacto importante en el modelo, por lo que se puede hacer un reajuste del criterio y tolerancia de aceptación.

Por último, es importante mencionar que aunque como tal no se desarrolla un sistema DSS en tecnologías digitales o con sistemas computacionales en la investigación, se destaca el apoyo de información estadística recaudada en un software para facilitar y apoyar a la toma de decisiones en la lista de intercepción de planta, ayudando también a que los materiales tengan una base analítica para definir sus riesgos de rechazo en las siguientes inspecciones.

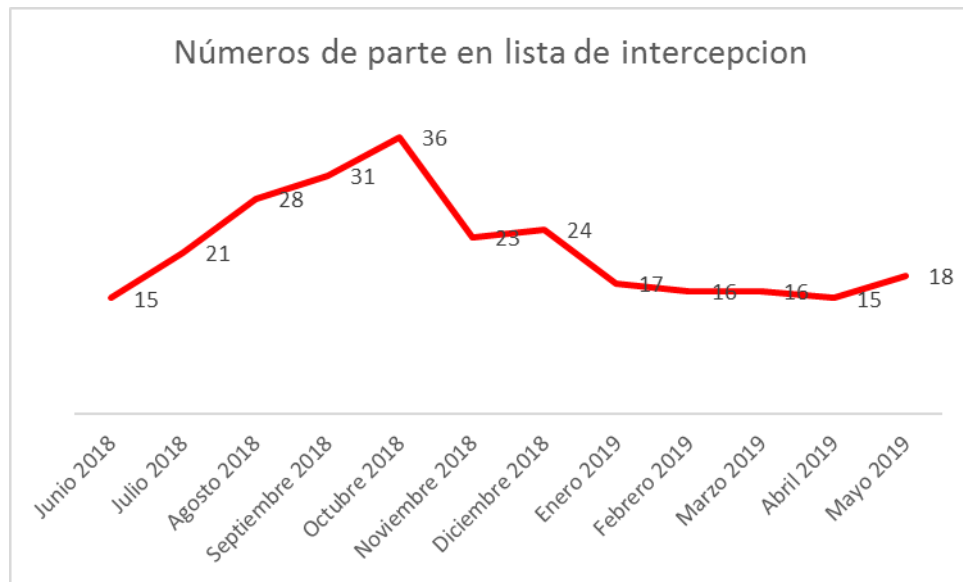
En un siguiente paso para la investigación, se puede considerar manejar ahora un sistema computacional o interfaz de usuario que facilite el desarrollo de las decisiones sin requerir forzosamente de una persona que haga las inferencias arrojadas como resultados, haciéndolo así, un sistema más automatizado y amigable para detectar los materiales y su riesgo de rechazar en inspecciones siguientes.

## **ANÁLISIS DE DATOS MENSUALES**

En el transcurso de los meses observados se ve una notoria diferencia al momento de que se implemente el nuevo indicador y el análisis de regresión logística para apoyar las decisiones sobre los materiales que salen o se mantienen en la lista de intercepción del departamento de calidad recibo. Esto se puede apreciar en un

seguimiento a la cantidad de números de parte en las listas de intercepción para los meses observados, siguiendo de manera directa a la cantidad de materiales que se encuentran en la lista de intercepción, comparando así mes con mes la mejora de las decisiones tomadas para retirar números de parte de los procesos de inspección para el área de calidad recibo. De manera visual ejemplificamos la mejora durante las observaciones e implementación con la siguiente gráfica Figura 14.

Figura 14. Meses de observación utilizando regresión logística nuevo indicador ERP.



(Elaboración propia)

Dentro de las principales mejoras que se ven en los meses de observación, se puede apreciar claramente como la curva de aprendizaje alcanza una madurez en el mes de Diciembre 2018 para estabilizar en los siguientes periodos y tener una menor cantidad de números de parte en lista de intercepción apoyando así a la cantidad de inspecciones y los modos de falla detectados para que no afecten en la liberación de estos materiales y entrega a las líneas de producción, así mismo, evitar saturar las condiciones críticas que puedan causar un rechazo en producción o producto terminado.

Cabe mencionar que un periodo anterior a Junio del 2018 se tiene un promedio de 30 números de parte en la lista de intercepción, alcanzando el pico más alto dentro del año 2018 en Octubre 2018 con 36 números de parte en inspección. Los datos previos al periodo de observación de Junio 2018 no cuentan con una transformación binaria para el estudio, pero se tiene la certeza por la cantidad que se reporta al cierre de mes a planta antes de este periodo para determinar el promedio de 30 números de parte en inspección en promedio por mes, esto se reporta a los niveles gerenciales para poder implementar las herramientas que se describen en esta investigación. Dicho esto, aseguramos que las cantidades de materiales que se inspeccionan en el periodo de observación no corresponden a una tendencia o temporada de demanda a lo largo del año.

A partir de enero 2019 podemos decir que los materiales incluidos en la lista de intercepción tienen ya un respaldo estadístico por el análisis efectuado de regresión logística, afirmando entre todos los niveles de planta que los resultados apoyan a la decisión para cada uno de estos materiales y su permanencia.

### **PREGUNTAS POSTERIORES AL DESARROLLO DE LA INVESTIGACIÓN**

De manera complementaria es importante recapitular algunas preguntas de investigación que surgen durante el desarrollo de la investigación para así conocer las aportaciones que se realizan con la investigación utilizando un modelo de regresión logística y el nuevo indicador ERP estimado.

¿Qué diferencia tiene con respecto a los indicadores usados actualmente?

La diferencia principal se centra en que el nuevo indicador ERP propuesto trabaja con observaciones realizadas para proyectar una estimación a futuro o en un siguiente paso, y por el otro lado, los indicadores actuales utilizan la misma

información de las observaciones para interpretar el histórico que se ha tenido y explicar el pasado para retomarlo en un presente con acciones correctivas.

¿Qué información puede dar como resultado?

Información complementaria para tomar la decisión acerca de los materiales que requieren seguir siendo inspeccionados o detectar aquellos que no generan riesgo de rechazarse y que pueden ahorrar recursos de inspección para ingresarse a las líneas de producción y no generar discrepancias de calidad.

¿Qué mide este indicador?

La probabilidad que tiene un material para rechazarse en la siguiente inspección.

¿Cómo se puede conseguir la información para calcular el nuevo indicador?

De la manera más simple posible, solo se requieren valores binarios, que son valores de 1 y 0, determinando que se tiene un 1 para cuando se encuentra un rechazo y 0 cuando no se encuentra rechazo para la observación que se realiza en la inspección. Por lo tanto, si se encuentra algún rechazo durante el día, se coloca un valor de 1, mientras que si no se encuentra ningún rechazo para algún material durante el día se coloca 0 en la tabla de hallazgos mencionada en secciones anteriores.

¿Qué decisiones se pueden apoyar en los resultados con este nuevo indicador?

Se pueden apoyar decisiones como sacar o mantener materiales en la lista de intercepción, retroalimentar a proveedor de hallazgos y riesgos en modos de falla



detectados, administrar recurso de inspecciones de calidad para actividades de mayor valor, acciones preventivas para minimizar impactos de calidad a líneas de producción, entre otras, en las cuales se centra la decisión para aceptar o no un riesgo de seguir o no inspeccionando materiales.

## **APORTACIÓN DE LA TESIS**

Con el uso y conocimientos adquiridos en esta investigación se encuentra un aporte importante al uso de herramientas estadísticas para la toma de decisiones y los riesgos y beneficios que se tienen por soportar estas decisiones con información e inferencias de herramientas estadísticas como la regresión logística. El entrenamiento y capacitaciones que se llevan a cabo durante esta investigación apoyan a ampliar el conocimiento en todos los niveles de la organización para tomar información necesaria de cada una de las estimaciones, así como el de fundamentar mejor las respuestas para explicar alguna tendencia o comportamiento en el análisis y recopilación de los datos. De igual manera se busca ampliar los criterios en el uso y beneficios que se tiene para la estadística en procesos que no usa variables numéricas y que pueden ser igualmente explicadas con herramientas con las que se utilice variables cualitativas o no numéricas.

## **APORTACIÓN SOCIAL DE LA TESIS**

Mediante el uso correcto y comprensión de las inferencias que se pueden hacer mediante la estadística, y específicamente, la regresión logística, se puede tener un impacto muy significativo en la optimización de variables en las que se mide el estudio, tomando como ejemplo costos, tiempos y materiales. El enfoque resultante de esta investigación aporta de igual manera en el uso correcto del capital humano ya que, además de optimizar la cantidad de miembros en los equipos de trabajo, se hace un significativo apoyo a los conocimientos técnicos y experiencias que se pueden adquirir teniendo en seguimiento una investigación de este tipo. Los entrenamientos y capacitaciones se consideran esenciales para todos los niveles de la empresa, esto por las conclusiones que se tienen de diferentes puntos de vista y que pueden fundamentar aún más la toma de decisión para el paso siguiente.

## REFERENCIAS

- Alberto, J., & Bosque, M. (2018). *Variables Cualitativas 1 Definición*. (4), 3–6. Retrieved from <http://www.uaaan.mx/~jmelbos/cursosma/made1.pdf>
- Alejandro, Á., Pérez, J., Kizys, R., Manzanedo, M., & Hoyo, D. (2016). *T10\_Reg\_Logistica.pdf*. 1–17.
- Cuevas, R. V., & Morales-Luna, G. (2011). *Introducción a la lógica difusa*.
- Descriptiva, I., & Variable, U. N. A. (2018). *Introducción Variables cualitativas Variables cuantitativas*. 1–8.
- Flores Zambada, R. (2012). Factores de la calidad de vida en el trabajo como predictoras de la intención de permanencia. *Acta Universitaria*, 22(2), 24–31. <https://doi.org/10.15174/au.2012.363>
- Goodman, M. S. (2017). Logistic regression. *Biostatistics for Clinical and Public Health Research*, 441–468. <https://doi.org/10.4324/9781315155661-19>
- Hernández, S. de la F. (2011). Regresión logística. *Facultad de Ciencias Económicas y Empresariales UNAM*.
- Jaccard, J., & Brinberg, M. (2021). Monte Carlo simulations using extant data to mimic populations: Applications to the modified linear probability model and logistic regression. *Psychological Methods*, No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/met0000383>
- Larra, P. (2007). *on Logística*. 1–9.
- Manotas, E. N., & Ibarra, K. H. (2014). *Regresión Logística Ordinal Aplicada a la Identificación de Factores de Riesgo para Cáncer de Cuello Uterino Ordinal Logistic Regression Applied to the Identification of Risk Factors for Cervical Cancer*. (17), 87–105.
- McGee, M. (2013). Logistic Regression. *Key Topics in Clinical Research*, 152–154. <https://doi.org/10.3109/9780203450307-30>
- NCSS, & LLC. (2018). *Book Chapter: Conditional Logistic Regression*. 1–26. Retrieved

from [https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Conditional\\_Logistic\\_Regression.pdf](https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Conditional_Logistic_Regression.pdf)

Park, H. A. (2013). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing*, 43(2), 154–164. <https://doi.org/10.4040/jkan.2013.43.2.154>

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *Journal of Educational Research*, 96(1), 3–14. <https://doi.org/10.1080/00220670209598786>

Pérez Hoyos, S. (1996). Introducción a la regresión logística. *Bioestadística*, (7), 1–11.

Ramírez, J. U. (2011). *Introducción a Simulación en la Industria. Definición.*

Shalizi, C. R. (2013). Advanced data analysis from an elementary point of view. *Book Manuscript*, 801. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.371.4613&rep=rep1&type=pdf>

Statistical, N., & Ncss, S. (2021). *Logistic Regression ( Old Version )*. 1–68.

Zekic-Susac, M., Sarlija, N., Has, A., & Bilandzic, A. (2016). Predicting company growth using logistic regression and neural networks. *Croatian Operational Research Review*, 7(2), 229–248. <https://doi.org/10.17535/crorr.2016.0016>

Zuo, W., Yuan, H., Shang, Y., Liu, Y., & Chen, T. (2016). *Calculation of a Health Index of Oil-Paper Transformers Insulation with Binary Logistic Regression*. 2016.

## **ANEXO**

## **ANEXO A**

El uso de la regresión logística se ve intensificado cuando se trata de explicar o hacer una proyección sobre una categoría dentro de un grupo específico para poder tener referencia de cuanta significancia tiene el modelo estimado de regresión logística y algunas de sus herramientas que conlleva.

Como ejemplo inicial se tiene un caso de estudio donde, mediante diferentes categorías y división de características de una muestra, se intenta determinar el factor de riesgo para presenciar cáncer de cuello uterino en una población de mujeres de estudio con el fin de anticipar estadísticamente si pueden padecer esta enfermedad (Manotas & Ibarra, 2014).

Otro ejemplo que se puede observar es una investigación que se realiza en unos transformadores a base de aceite. En este estudio se busca determinar con la regresión logística, si se puede categorizar como "Sano" o "Insano" la condición del aceite mediante los valores estimados y observados de agua en el aceite, acidez del aceite, gases de combustión, voltaje de transferencia en el aceite, factor de disipación y un factor determinado de degradación de polímeros. Sus conclusiones con el uso de la regresión logística se basan con el uso de un indicador HI (Health Index), por lo que califican con este indicador cada transformador que se observa, y con el modelo de regresión logística binaria estimada, proyectar un valor HI con una serie de variables ya obtenidas (Zuo, Yuan, Shang, Liu, & Chen, 2016).

Un último ejemplo que se explora para el uso de la regresión logística se explica en una investigación donde se busca medir y analizar diferentes comportamientos psicológicos e incluso sus probabilidades de suceso extendiendo su análisis con esta herramienta. Aquí se le denomina como "Extant Data Simulation", que tiene como procedimiento, extender el análisis de una herramienta como la simulación de Monte Carlo, con las mismas conclusiones, con un procedimiento más fácil y con mejor comprensión de la estimación (Jaccard & Brinberg, 2021).