



**CORPORACIÓN MEXICANA DE INVESTIGACIÓN EN MATERIALES**

**DIVISIÓN DE ESTUDIOS DE POSGRADO**



***Método para el control de la producción defectuosa dentro del proceso de conformado de metales. Caso ZF Sachs Automotive México.***

**TESIS**

**Que para obtener el Grado Académico de**

**Maestro en Ciencia y Tecnología con Especialidad de Ingeniería Industrial  
y de Manufactura**

**Presenta:  
Jonatan Abraham Limones Lara**

**Saltillo, Coahuila, México, Septiembre del 2012**

*Método para el control de la producción defectuosa dentro del proceso de conformado de metales. Caso ZF Sachs Automotive México.*

*Por*

*Jonatan Abraham Limones Lara*

*Tesis*

*Presentada al Programa Interinstitucional en Ciencia y Tecnología*

*Sede*

**Corporación Mexicana de Investigación en Materiales, S. A. de C. V.**

*Como requisito parcial para obtener el Grado Académico de*

*Maestro en Ciencia y Tecnología*  
***Especialidad en Ingeniería Industrial y de Manufactura***

**Programa Interinstitucional en Ciencia y Tecnología COMIMSA / CONACyT**

*Saltillo, Coahuila. Septiembre del 2012*

**Corporación Mexicana de Investigación en Materiales, S.A. de C.V**

**Gerencia de Desarrollo Humano**

**División de Estudios de Posgrado**

Los abajo firmantes, miembros del Comité Tutorial recomendamos que la Tesis **“Método para el control de la producción defectuosa dentro del proceso de conformado de metales. Caso ZF Sachs Automotive México”**, realizada por el alumno **Jonatan Abraham Limones Lara**, matrícula **0906IM3057** sea aceptada para su defensa como Maestro en Ciencia y Tecnología con Especialidad en **Ingeniería Industrial y de Manufactura**.

El Comité Tutorial

---

Tutor Académico  
Director de Tesis  
**Dr. José Antonio Vázquez López**

---

Tutor en Planta  
**Ing. Erick Gerardo Lozano**

---

Asesor  
**Dr. Ismael López Juárez**

---

Vo. Bo  
**Dr. Felipe Arturo Reyes Valdés**  
Coordinación General de Estudios de Posgrado  
COMIMSA

**Corporación Mexicana de Investigación en Materiales, S.A. de C.V**

**Gerencia de Desarrollo Humano**

**División de Estudios de Posgrado**

Los abajo firmantes, miembros del Jurado del Examen de Grado del alumno, **Jonatan Abraham Limones Lara** una vez leída y revisada la tesis titulada **“Método para el control de la producción defectuosa dentro del proceso de conformado de metales. Caso ZF Sachs Automotive México”**, aceptamos que la referida tesis revisada y corregida sea presentada por el alumno para aspirar al grado de Maestro en Ciencia y Tecnología con Especialidad en Ingeniería Industrial y de Manufactura durante el Examen de Grado correspondiente.

Y para que así conste firmamos la presente a los 27 días del mes de septiembre del año 2012.

Dra. Martha Patricia Guerrero  
Mata

**Presidente**

Dr. Juan Pablo Nieto González

**Secretario**

Dr. Elias Gabriel Carrum Siller

**Vocal**

Dr. Mario César Osorio  
Abraham

**Vocal**

Dr. José Antonio Vázquez  
López

**Vocal**

## DEDICATORIA

A quien merece todo de mí.

“En este mundo todo tiene su hora;  
hay un momento para todo cuanto ocurre”.

## AGRADECIMIENTOS

El agradecimiento más especial a quien dio su vida por mí. Has llenado de fuerza mi espíritu y me has rodeado de personas sorprendentes. Gracias Dios, eres lo máximo.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo económico durante 2 largos años y a la Corporación Mexicana de Investigación en Materiales (COMIMSA) por ser el lugar en el cual he alcanzado este sueño, así como por permitirme utilizar los laboratorios para realizar las pruebas mecánicas que exigió el proyecto.

Agradezco a mi comité tutorial: Dr. José Antonio Vázquez López y al Dr. Ismael López, que hicieron todo lo posible para que esta investigación fuera un éxito, gracias por su paciencia y sus conocimientos científicos. A la empresa ZF Sachs y su personal: gracias por ser parte del proyecto, al Ing. Erick Lozano por su buen carácter y al Ing. Luis A. Rodríguez por compartirme su conocimiento del proceso, sus aportaciones fueron de gran valor.

Gracias a mi familia por su incansable apoyo, ustedes forman mi ejemplo a seguir. Abraham Limones: creatividad sin límites, Raquel Lara: consejos que no pasan de moda y Liliana: siempre en movimiento. Gracias. Agradezco a Mar porque es ella quien me motivó a realizar estos estudios de posgrado y me acompañó en la gran aventura: sin ti no tendría un mundo feliz.

Finalmente agradezco a mis compañeros, ustedes saben lo que vivimos juntos (exámenes, comidas, deporte y hasta debates) y al Dr. Marco Aurelio por ser parte de nosotros.

## RESUMEN

La investigación tiene como objetivo primordial el desarrollo de un método, que permita analizar el estado actual de los procesos de conformado de metales por troquelado y estampado, y que sea guía para enfrentar el problema de la variación en dicho proceso. Los productos que se elaboran a través de esta forma tienen una alta demanda en la industria automotriz, de tal forma que si el proceso se ve afectado por una causa especial de variación entonces, dichos productos se verán afectados convirtiéndose en componentes desechados que a su vez representan una gran cantidad de desperdicio en materia prima, energía y tiempo.

Ante dicha situación, el método planteado dentro de la investigación involucró diferentes herramientas para el análisis numérico y gráfico de los datos, así como las bases estadísticas y de probabilidad del modelo de regresión lineal. Además el método conduce a la selección acertada de un plan de muestreo en el proceso y emplea paquetes computacionales avanzados como los son R Cran y Minitab. La utilización de estos elementos provee un procedimiento sencillo y ordenado para el control de la producción defectuosa en procesos de estampado y troquelado.

En los capítulos finales de esta tesis se muestra que para el caso de estudio la aplicación del método propuesto permitió alcanzar los objetivos planteados en la investigación, así como evaluar el comportamiento de las variables que aportan una mayor variación al producto.



## CONTENIDO

LISTA DE TABLAS .....	ix
LISTA DE FIGURAS .....	x
CAPÍTULO 1 INTRODUCCIÓN .....	12
CAPÍTULO 2 PLANTEAMIENTO DEL PROBLEMA.....	15
2.1 DESCRIPCIÓN DEL PROBLEMA.....	15
2.2 PREGUNTAS DE INVESTIGACIÓN .....	20
2.3 HIPÓTESIS DE INVESTIGACIÓN .....	20
2.3.1 <i>Hipótesis General</i> .....	21
2.3.2 <i>Hipótesis Específicas</i> .....	21
2.4 OBJETIVOS.....	21
2.4.1 <i>Objetivo General</i> .....	21
2.4.2 <i>Objetivos Específicos</i> .....	22
2.5 JUSTIFICACIÓN.....	22
2.6 DELIMITACIONES.....	22
2.6.1 <i>Limitaciones en el caso de estudio</i> .....	23
CAPÍTULO 3 MARCO DE REFERENCIA .....	24
3.1 MARCO TEÓRICO. ....	24
3.1.1 <i>La calidad en el entorno empresarial.</i> .....	24
3.1.2 <i>La calidad y la variación.</i> .....	25
3.1.3 <i>Revisión del Estado del Arte.</i> .....	26
3.2 MARCO CONCEPTUAL. ....	30
3.2.1 <i>Variabilidad dentro de un proceso.</i> .....	30
3.2.2 <i>Análisis numérico de los datos.</i> .....	31
3.2.3 <i>Análisis gráfico de los datos.</i> .....	36
3.2.4 <i>Modelos y Técnicas estadísticas para el análisis de datos.</i> .....	40
3.2.5 <i>Métodos para el control de calidad y muestreo para aceptación.</i> .	46
CAPÍTULO 4 MATERIALES Y MÉTODOS.....	52
4.1 MÉTODO.....	52

4.1.1 Descripción del método.....	52
4.2 DESARROLLO DEL MÉTODO.....	59
4.2.1 Etapa 2: Diagnóstico del entorno y pruebas iniciales.....	59
4.2.2 Etapa 3: Selección de variables involucradas.....	67
4.2.3 Etapa 4: Recolección y estructuración de la base de datos.....	72
4.2.4 Etapa 5: Verificación de la información obtenida / Reestructurar..	76
CAPÍTULO 5 EXPERIMENTACIÓN Y RESULTADOS .....	83
5.1 ETAPA 6: ANÁLISIS DE LA BASE DE DATOS.....	83
5.1.1 Etapa 6.1: Análisis de la variable de salida .....	84
5.1.2 Etapa 6.2: Análisis de correlación.....	86
5.2 ETAPA 7: APLICACIÓN DEL MODELO DE REGRESIÓN LINEAL.....	88
5.3 ETAPA 8: PLAN DE MUESTREO DE ACEPTACIÓN PARA $x_4$ .....	99
CAPÍTULO 6 CONCLUSIONES Y RECOMENDACIONES.....	102
REFERENCIAS .....	106
ANEXO 1 .....	109
APÉNDICE.....	116

## LISTA DE TABLAS

Tabla 2.1 Clasificación de defectos. ....	18
Tabla 2.2 Esquema general de la empresa para la solución de la problemática. ....	20
Tabla 3.1 Construcción de una tabla del análisis de varianza. ....	35
Tabla 3.2 Observaciones $X \sim N(0,1)$ .....	37
Tabla 4.1 Valores de los parámetros para la distribución binomial. ....	59
Tabla 4.2 Datos para las pruebas iniciales .....	62
Tabla 4.3 Pruebas de normalidad para $y$ en cada grupo de $x$ .....	63
Tabla 4.4 Datos para el análisis de regresión .....	64
Tabla 4.5 Análisis de Varianza para la regresión lineal .....	65
Tabla 4.6 Intervalos de predicción para $y$ .....	65
Tabla 4.7 Rango de funcionamiento de los parámetros de la maquinaria. ....	68
Tabla 4.8 Dimensiones y propiedades mecánicas de la materia prima. ....	70
Tabla 4.9 Propiedades químicas de la materia prima. ....	71
Tabla 4.10 Variables seleccionadas para la investigación. ....	72
Tabla 4.11 Valores de las características que reporta el proveedor. ....	73
Tabla 4.12 Mediciones tomadas referentes a la apertura del troquel. ....	74
Tabla 4.13 Mediciones registradas de la longitud del álabe de cada colada. ...	75
Tabla 4.14 Estructura general de la base de datos.....	76
Tabla 4.15 Resultados y comparación del análisis químico.....	77
Tabla 4.16 Resultados del ensayo de dureza .....	78
Tabla 4.17 Resultados del ensayo de tensión. ....	81
Tabla 5.1 Estructura general de la base de datos / Pruebas mecánicas. ....	83
Tabla 5.2 Análisis de varianza de un solo factor.....	85
Tabla 5.3 Coeficiente de correlación $r$ entre las variables. ....	87
Tabla 5.4 ANOVA para la regresión entre $x_4'$ y $y'$ .....	94
Tabla 6.1 Probabilidades calculadas para $\alpha$ y $\beta$ , izquierda y derecha respectivamente.....	103

## LISTA DE FIGURAS

Figura 2.1 Áreas de complejidad dentro de una operación de conformado. ....	16
Figura 2.2 Álabe 712000226.....	16
Figura 2.3 Diagrama de Pareto para los defectos presentados en el álabe. ....	19
Figura 2.4 Defecto de golpe en los álabes.....	19
Figura 3.1 Histograma para las observaciones de la Tabla 3.2. ....	37
Figura 3.2 Diagrama de caja para las observaciones de la Tabla 3.2 .....	38
Figura 3.3 Diagrama de dispersión .....	39
Figura 3.4 Recta de regresión o recta de mínimos cuadrados. ....	42
Figura 4.1 Esquema general del método. ....	58
Figura 4.2 Desfase en las distribuciones de probabilidad del proceso .....	60
Figura 4.3 Diagrama de dispersión para $x$ y $y$ . ....	63
Figura 4.4 Recta de regresión.....	64
Figura 4.5 Gráfica integral de regresión.....	66
Figura 4.6 Longitud del álabe.....	68
Figura 4.7 Materia prima montada en el alimentador de la prensa. ....	69
Figura 4.8 Croquis de las mediciones de dureza. ....	78
Figura 4.9 Metalografías de los aceros .....	79
Figura 4.10 Comparación del esfuerzo de fluencia: proveedor vs análisis. ....	81
Figura 4.11 Comparación del esfuerzo máximo: proveedor vs análisis .....	82
Figura 5.1 Diagrama de caja para los tres grupos de datos. ....	85
Figura 5.2 Histogramas para la variable $y$ de cada colada. ....	86
Figura 5.3 Recta de regresión entre $y$ y $x^4$ .....	88
Figura 5.4 Distribución no normal de los residuales .....	89
Figura 5.5 Residuales no independientes.....	90
Figura 5.6 Observaciones para la elongación.....	91
Figura 5.7 Observaciones de la longitud.....	91
Figura 5.8 Diagrama de dispersión para las nuevas observaciones.....	92
Figura 5.9 Recta de regresión para $x^4$ y $y'$ . ....	93

Figura 5.10 Recta regresión sin observaciones inusuales.....	94
Figura 5.11 Verificación del supuesto de normalidad. ....	96
Figura 5.12 Verificación del supuesto de independencia.....	97
Figura 5.13 Verificación de independencia con otra variable.....	97
Figura 5.14 Intervalos de confianza para la regresión lineal.....	98
Figura 5.15 Curva de operación para el plan de muestreo. ....	101
Figura 6.1 Diagrama de modelos alternativos. ....	105

## CAPÍTULO 1 INTRODUCCIÓN

Actualmente, debido a la complejidad que los procesos de estampado y troquelado enmarcan en referencia a la recuperación elástica, la predicción de dicha recuperación es un problema clave para los ingenieros que participan en la manufactura de componentes automotrices por medio de estos procesos.

Existen casos donde la recuperación elástica es pequeña, pero significativa. Se dice *pequeña* ya que no es visible a simple vista, y se sabe que es significativa, porque afecta la calidad del producto final y repercute en la interacción que este presenta con otros elementos al cual debe unirse. En estas circunstancias se requiere de instrumentos de medición con gran precisión, de datos que representen lo que ocurre en el proceso, pero principalmente de un análisis que permita: revelar *si existe* un cambio ó una tendencia al cambio en el producto y revelar *qué* lo ha ocasionado.

Es entonces que la investigación presente se enfoca en el estudio de los efectos que ocasionan diversas variables de la materia prima en dichos procesos sobre la calidad y recuperación elástica del producto por medio de herramientas estadísticas.

En el Capítulo 2 se da a conocer el panorama general de la problemática que se presenta en la industria de partes automotrices, manufacturadas a través de procesos de troquelado y estampado. Se explica el problema específico del caso de estudio y en base a esto se plantearon las hipótesis que se consideraron a lo largo de la investigación.

Además se establecen los objetivos a los cuales es factible llegar. Al final de este capítulo se delimitó la temática de investigación y se exponen las limitantes que se presentaron en el desarrollo del caso de estudio en particular.

Dentro del Capítulo 3 se realiza una revisión de la literatura enfocada a las problemáticas de variación que surgen en los procesos a los cuales se encuentra dirigida esta tesis. Esta primer parte se nombra como Marco Teórico. En la segunda parte del capítulo se estudia la información correspondiente a la variabilidad de un proceso, análisis numérico y gráfico de los datos, modelos y técnicas estadísticas para análisis de los mismos y planes de muestreo para aceptación, lo que corresponde al Marco Conceptual. Ambas partes conforman el Marco de Referencia de la tesis.

En el Capítulo 4 se incluye una descripción de las 9 etapas que corresponden al método propuesto, además se presenta un esquema general del mismo con el objetivo de brindar una visión completa del orden de cada una de las etapas. En este mismo capítulo se desarrollan las etapas 2 al 5 por ser información concerniente a los materiales y métodos de la investigación. Finalmente se obtiene una base de datos que sirve de fundamento para el siguiente capítulo.

Para el Capítulo 5 se presenta el análisis de la base de datos, el cual incluye en primera instancia, el análisis de las mediciones correspondientes a la variable de salida, así como un análisis de correlación con el objetivo de identificar aquella variable de entrada que posee mayor fuerza lineal con la variable de salida. De esta forma, se establece un Modelo de Regresión Lineal. Al final de este capítulo se desarrolla un Plan de Muestreo de Aceptación para una variable.

Con respecto al Capítulo 6, se presentan las conclusiones y recomendaciones relacionadas al método propuesto, así como aquellos aspectos relevantes del caso de estudio que surgieron durante la investigación. Finalmente, se proyecta el trabajo futuro que puede ser realizado en otras investigaciones de interés.

Al final del documento se incluyen los anexos con relación a los códigos de programación del paquete R cran realizados durante la investigación, así como diferentes documentos relacionados al Capítulo 4, etapa 5.1. También se incluye un apéndice con información adicional de la sección 3.2.5.



## CAPÍTULO 2 PLANTEAMIENTO DEL PROBLEMA

### 2.1 Descripción del problema

Dentro de los procesos de conformado de metales se llevan a cabo diferentes operaciones sobre la materia prima, en el proceso de interés para esta investigación se consideran las operaciones de estampado y troquelado.

En ambas operaciones existen variables aleatorias que son normalmente difíciles de predecir y analizar, individualmente así como las relaciones entre ellas. De acuerdo con Altan et al. (1986) esta situación se ha identificado desde mediados de los años 80, y según García-Romeu (2005) permanece hasta la actualidad. Al desconocer tales relaciones se llega al punto de carecer de un sistema que permita describirlas, lo que generalmente hace que se presenten áreas de complejidad al momento de examinar cada variable y sus efectos en el proceso.

La Figura 2.1 muestra tres áreas en las cuales se presenta complejidad, estas son: análisis, caracterización y predicción, y que además están entrelazadas entre sí, marcando una fuerte dependencia de las acciones que se ejecutan en cada una de ellas.

Cuando estas áreas no se estudian adecuadamente se desatan diversas situaciones problemáticas en los procesos dentro de las empresas, como se identificó en la empresa ZF Sachs Automotive México, en su proceso de fabricación del álabe 712000226, dicho producto se observa en la Figura 2.2.

El álabe es utilizado en el ensamble de la Turbina X15 la cual forma parte del Convertidor de Torque automotriz fabricado en la planta antes mencionada, dicho convertidor de torque es utilizado en automóviles, autobuses u otros vehículos de trabajo pesado para transmitir de forma eficiente la potencia que genera el motor a la caja de transmisión.

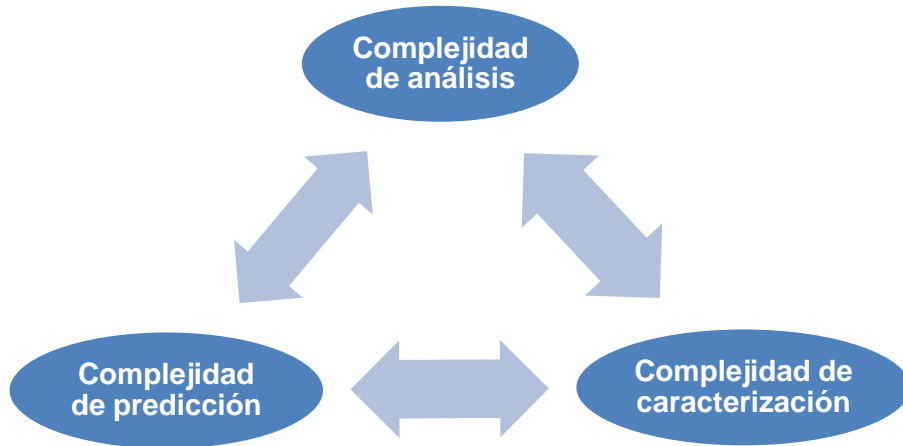


Figura 2.1 Áreas de complejidad dentro de una operación de conformado.



Figura 2.2 Álabe 712000226

En tal proceso se determinó que se trataba de un problema de reducción de la variación en la longitud del álabe y control de unidades defectuosas. Como consecuencia de dicha variación se observó un alto costo por desperdicios a finales del año 2009 ascendiendo a \$170, 648 pesos moneda nacional. Esto representa el 0.22% de las ventas obtenidas en el año para el producto en estudio.

De acuerdo con Montgomery (2007) la complejidad en dicho proceso implica:

- Analizar las variables más importantes que se relacionan con el producto.
- Caracterizar dichas variables como un sistema para describir los efectos que se presentan en el producto.
- Predecir el comportamiento que se tendrá en el producto según el estado de las variables de entrada en el sistema.

Estas tres acciones se encuentran directamente encaminadas a la solución de la problemática con un fundamento científico y estadístico.

En el proceso de manufactura del álabe se identifican 4 elementos principales, los cuales son: geometría, tolerancias, tasas de producción y los factores humanos. El porcentaje de desperdicios se encuentra relacionado principalmente con la geometría y las tasas de producción, ya que el álabe posee una curvatura compleja la cual se monitorea con poca frecuencia en comparación con la cantidad de piezas que se producen, sin embargo una característica que resume en buena forma su geometría y además es utilizada para monitorear la variación en este, es su longitud, característica que se explica en el capítulo 4.

En cuanto a la tasa de producción el proceso funciona a grandes velocidades de ciclo, 180 piezas son producidas en un minuto, en tal situación cuando se presenta una causa especial de variación en el proceso, esta será detectada cuando ya se han producido una gran cantidad de piezas lo que genera un aumento en la fracción de productos defectuosos.

En la Tabla 2.1 se muestran los defectos que se presentan en el álabe así como también se menciona la familia de defectos que comúnmente ocurren dentro de los procesos de conformado de metales por estampado y troquelado.

Tabla 2.1 Clasificación de defectos.

Producto	Defecto
Producto de estudio: <u>álabe de turbina</u>	Golpes
	Piezas de ajuste
	Marcas de Rebabas (M.R.)
	Pruebas de ingeniería (P.I.)
Otros productos	Rayas (arrastre de material)
	Alturas fuera de especificación
	Rebabas
	Run out fuera de especificación

Se identificó a través del diagrama de Pareto mostrado en la Figura 2.3, el defecto principal que se presenta en el álabe, siendo este el defecto de “golpes” y representa el 48.49 % de los defectos con lo cual se observa que la mayor cantidad de productos defectuosos es ocasionada por una cuarta parte de los defectos confirmándose el principio de Pareto: pocos vitales, muchos triviales.

Para ilustrar el principal defecto al cual está enfocada esta investigación, se muestra en la Figura 2.4 la falta de paralelismo entre los álabes al momento de agruparlos uno encima del otro ya que sufrieron una deformación producida por un impacto durante el ensamble, este impacto es ocasionado por la variación que presenta el álabe en su geometría.

Las piezas que presentan dicho defecto no tienen oportunidad de ser re-trabajadas, así que no existe recuperación en la cantidad de producto dañado.

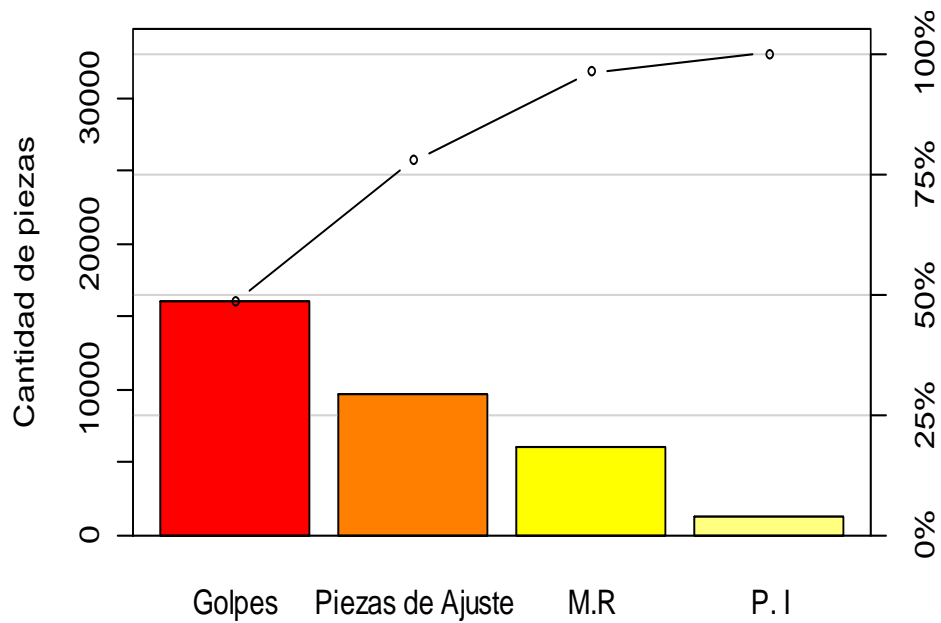


Figura 2.3 Diagrama de Pareto para los defectos presentados en el álabe<sup>1</sup>.



Figura 2.4 Defecto de golpe en los álabes.

Se identificó que esta problemática no se analiza científicamente, sino que se aborda por medio de acciones que se ejecutan de forma empírica, a prueba y error. En la Tabla 2.2 se menciona la acción ejecutada por la empresa para la solución de la problemática y la desventaja que se tiene en dicha acción.

<sup>1</sup> El diagrama fue generado con los registros que se tomaron de la empresa en un lapso de 3 meses.

Tabla 2.2 Esquema general de la empresa para la solución de la problemática.

Departamento	Defecto	Acción	Desventaja
Producción e Ingeniería de procesos	Golpes	Ajustes constantes en máquina de ensamble	Se presenta sobreajuste en los parámetros de proceso.

## 2.2 Preguntas de Investigación

Al inicio de la investigación se plasmaron diversas situaciones a través de las siguientes preguntas con el objeto de identificar el panorama general, así como el punto meta de la investigación.

- ¿Cuál es la principal causa de variación en el producto considerado dentro del problema presentado anteriormente?
- ¿La distribución binomial y el modelo de Regresión Lineal son adecuados para realizar el análisis y caracterización del problema?
- ¿Cuáles son las variables que presentan un efecto mayor en la variabilidad del producto?
- ¿Cuál plan de muestreo de aceptación por variables es más conveniente proponer para el caso de estudio?

## 2.3 Hipótesis de Investigación

En la investigación se plantearon las siguientes hipótesis:

### **2.3.1 Hipótesis General**

Por medio del método propuesto es posible aplicar herramientas estadísticas dentro de un proceso de conformado de metales, con el fin de seleccionar la variable de entrada más significativa para controlar la variación en el producto final.

### **2.3.2 Hipótesis Específicas**

$H_0$ : Con la selección de una variable de entrada, relacionada a la variable de salida en el producto, se tiene un tamaño de muestra inadecuado para su inspección y un criterio de aceptación que discrimina en forma ineficiente las partes que no cumplen con las especificaciones del cliente.

$H_a$ : Con la selección de una variable de entrada, relacionada a la variable de salida en el producto, se tiene un tamaño de muestra adecuado para su inspección y un criterio de aceptación que discrimina en forma eficiente las partes que no cumplen con las especificaciones del cliente.

## **2.4 Objetivos**

Los objetivos planteados a lo largo de la investigación fueron los siguientes:

### **2.4.1 Objetivo General**

Desarrollar un método apropiado para verificar la calidad de la materia prima dentro de los procesos de conformado de metales, que permita tener un control de la producción defectuosa.

### **2.4.2 Objetivos Específicos**

- Explicar el comportamiento de la producción defectuosa a través de la distribución binomial.
- Identificar la variable de entrada con mayor relación sobre la variable de salida.
- Proponer un plan de muestreo de aceptación para la variable significativa en el proceso.

### **2.5 Justificación**

La realización del proyecto es justificable por las siguientes razones:

- Se identifica el comportamiento de la producción defectuosa, explicado por medio de la distribución binomial.
- Se explica de manera científica y a través de modelos estadísticos la variación en el producto manufacturado.
- Se propone un método para el control de la materia prima causante de la producción defectuosa.

### **2.6 Delimitaciones**

El proyecto se encuentra enfocado a desarrollar un método de control de la producción defectuosa, utilizando la distribución binomial, un modelo de regresión lineal y un plan de muestreo de aceptación para variables, específicamente para la media de los datos.



### **2.6.1 Limitaciones en el caso de estudio.**

Durante el desarrollo de la investigación se presentaron limitaciones en diferentes áreas, las cuales se describen a continuación:

En relación al tiempo, ya que la relación con la empresa y el acuerdo de la problemática a tratar se establecieron en un tiempo posterior al marcado por el programa de investigación, esto afectó al desarrollo de actividades posteriores. En relación a la población de los datos, debido a los tiempos establecidos de producción, no fue posible la obtención de una base de datos más amplia.

## **CAPÍTULO 3 MARCO DE REFERENCIA**

### **3.1 Marco Teórico.**

#### **3.1.1 La calidad en el entorno empresarial.**

Las empresas se encuentran sumergidas en un constante cambio, consecuencia de las necesidades y deseos del mercado, para satisfacer dichas necesidades y lograr así el crecimiento que buscan requieren que sus procesos adquieran la capacidad de responder a estas situaciones cambiantes. Los procesos cuentan con recursos humanos, recursos económicos y materiales y además el efecto del tiempo, estos recursos deben ser dirigidos en forma adecuada y designados a alcanzar las metas que forman la prioridad de la empresa: crecimiento y competitividad.

La administración eficiente de estos recursos de producción forma la base para la evolución de la economía y la ciencia, logrando conjuntar ambos conceptos en: investigación, tecnología e innovación, esto forma la clave del éxito empresarial. La calidad es la manera posible en que se puede medir dicho éxito, medir si se han desarrollado los servicios y productos al grado que estos puedan satisfacer a los clientes y aun más; generar en ellos una fascinación por las cosas que se tiene para ellos.

La calidad no es un concepto nuevo, pero si es una idea que ha evolucionado a lo largo de la historia llegando hasta este tiempo con más fuerza y mayor impacto, por tal razón el propósito de esta investigación gira alrededor de la calidad y la administración de la variación en los productos

Manufacturados. A continuación se explica en breves palabras el surgimiento y el desarrollo de la calidad hasta llegar al enfoque de administrar la variación y el uso de herramientas estadísticas para este fin.

### **3.1.2 La calidad y la variación.**

Resulta complicado encerrar el significado de calidad bajo un solo concepto, ya que la calidad es una idea con múltiples definiciones, de las cuales se han marcado dentro de la ingeniería industrial tres principalmente.

Como base se puede partir de la primera definición donde el enfoque es que la calidad significa: adecuado para utilizarse. Dentro de esta definición se han marcado dos aspectos importantes: calidad de diseño y calidad de conformidad.

La calidad de diseño se observa cuando todos los productos y servicios son manufacturados en diferentes niveles de calidad; ya que los productos pueden ser fabricados con un mismo objetivo pero existen diferentes modelos, que presentan un tipo de material diferente, un rendimiento diferente, entre otras cosas. Esta situación es el resultado de un diseño pensado intencionalmente.

La calidad de conformidad se ve como un cumplimiento del producto hacia las especificaciones que se han marcado en el diseño y se encuentra influenciada por un gran número de factores como el proceso de fabricación seleccionado, la capacitación al equipo de trabajo, las herramientas utilizadas para el control del proceso, inspecciones realizadas, y la motivación de la mano de obra para desempeñar las actividades logrando calidad.

Como un segundo concepto utilizado dentro del enfoque estadístico, Montgomery (2009) afirma que: la calidad es inversamente proporcional a la variabilidad. Esto implica que si la variabilidad en la característica importante de un producto decrece, entonces la calidad del producto aumenta.

El tercer concepto que se debe considerar según Montgomery (2009) es que: la mejora de la calidad se define como la reducción de la variabilidad en procesos, con el objetivo de garantizar que las características de calidad de un producto se encuentren en el nivel nominal requerido.

Estos conceptos se encuentran involucrados a lo largo de la investigación, ya que al controlar la variación se logra obtener el valor deseado en la característica de calidad y se genera directamente una disminución en la proporción de productos defectuosos.

### **3.1.3 Revisión del Estado del Arte.**

Montgomery (2007) menciona que un proceso de manufactura se debe caracterizar, es decir, determinar cuáles son las variables de entrada que influyen en la variación del producto. Además es importante identificar los ajustes que deben hacerse en dichas variables para reducir tal variación. Para esto se pueden utilizar diferentes herramientas estadísticas como diseño de experimentos, regresión lineal y análisis de varianza, entre otros; las cuales permiten estimar la magnitud de la variación que proporciona cada variable de entrada, es decir, cuánto cambia la variable de respuesta medida en el producto, cuando se modifica cada variable de entrada. Así mismo proporcionar información acerca de las variables que deberían controlarse con mayor atención durante el proceso de fabricación a fin de evitar los niveles elevados de defectos y el desempeño errático del proceso. Como consecuencia

menciona que después de caracterizar el proceso es posible aplicar planes de muestreo y cartas de control a una o más de las variables involucradas.

De Vin (2000) menciona que anteriormente, se creía que los procesos de prensado carecían de importancia comercial y que estos presentaban pocas complicaciones, sin embargo, Bosga y Segura (2007) muestran que actualmente dentro de la industria automotriz el área que precisa más capital es el taller de prensado. Mencionan que la carrocería de un automóvil está compuesta por varias centenas de piezas metálicas donde la mayoría se fabrican mediante procesos de prensado, quedando claro que las empresas con tecnología de punta trabajan con dichos procesos.

Limones et al. (2011) afirma que el conformado de metales por estampado regularmente es muy redituable en cuestiones de energía, ya que un producto final puede ser manufacturado con un solo ciclo de la prensa y un solo golpe. En otras situaciones el producto se fabrica con una sucesión de golpes, siendo el caso cuando se utilizan troqueles progresivos para fabricar una pieza, esto representa un aprovechamiento de la carga utilizada y el tiempo empleado, ya que se ejecutan diferentes pasos simultáneamente. La contraparte se genera cuando el proceso funciona a grandes velocidades de ciclo, ya que si se presenta una causa especial de variación en el proceso, esta será detectada cuando ya se han producido una gran cantidad de piezas, lo que genera un aumento en la fracción de productos defectuosos.

Dentro de la colaboración de Bosga y Segura (2007) se afirma algo similar: si se aumenta la velocidad de funcionamiento de una prensa se pone en gran riesgo la calidad del producto final, quedando visible que los procesos de prensado constituyen un área con potencial de mejora.

De Vin (2000) explica que para modelar un proceso u operación de estampado se requiere conocer adecuadamente dicha operación e identificar el mayor número de variables involucradas.

Garcia- Romeu (2005), menciona que los elementos que presentan una mayor influencia en la falta de precisión de la geometría final del producto son: la variación en las propiedades mecánicas de la lámina, las desviaciones en el espesor de la misma y el herramental conformado por matriz y punzón principalmente. Su estudio se basa en un proceso de doblado al aire, con lo cual afirma que los procesos de doblado son críticos debido a la variación y/o desviaciones que se presentan en el producto y que se constituyen como las principales causas de rechazo o desperdicio.

Yuqi et al. (2002) analizaron un proceso de doblado de lámina en forma de U, dicho análisis se enfoca en la recuperación elástica que se presenta durante el formado, el análisis involucra las siguientes características: condiciones de lubricación, fuerza de soporte del herramental, dureza del material y anisotropía de la lámina, así como el efecto que cada una tiene en el producto. Mencionan que debido a la complejidad que los procesos de estampado envuelven en cuanto a la deformación plástica-elástica, la predicción precisa de dicha deformación es un problema clave para los ingenieros que participan en la manufactura de componentes automotrices por medio de estos procesos. Concluyen que las características del material y los parámetros técnicos afectan estrictamente la calidad de formado, la precisión de los productos y afectan a los resultados de la recuperación elástica.

Senjuntichai y Pitthayakornpisuth (2010) utilizan el modelo general lineal para describir el comportamiento y efecto de los factores independientes como función de una variable de respuesta. Sugieren condiciones del proceso que permitan reducir el porcentaje de productos defectuosos.

Jin y Guo (2003) mencionan que la variación total de un proceso de fabricación por lotes, se descompone en la variación entre lotes y en la variación dentro de los lotes. La separación y estimación de la variación de los componentes es una forma de mejorar el proceso. Consideran como elementos de variación: las condiciones de la herramienta, posición de la herramienta y composición del material.

Montgomery (2009) comenta que el muestreo de aceptación es muy utilizado en las organizaciones con poca experiencia en el aseguramiento de la calidad; sin embargo, en cuanto la organización madura en este aspecto, presentan menos dependencia del muestreo de aceptación y más hacia el control estadístico del proceso (CEP).

Existen empresas que no cuentan con un método adecuado para el aseguramiento de la calidad de la materia prima, ya que confían en la información del proveedor o en ciertos casos solo inspeccionan una característica sin un estudio previo, que les haya mostrado el impacto que dicha característica tienen en el producto a manufacturar. Otros inconvenientes se presentan cuando la inspección de la materia prima se realiza con instrumentos de medición que no cuentan con una buena calibración.

Una situación similar ocurre en el caso de estudio, donde la empresa aplica CEP en diferentes partes del proceso, pero no presenta madurez en la inspección de la materia prima, tal como se explica en el Capítulo 5 del documento.

## 3.2 Marco Conceptual.

### 3.2.1 Variabilidad dentro de un proceso.

De acuerdo con la Real Academia Española (2003) la variación, se puede definir en su concepto más sencillo como un cambio de forma o estado. Dentro de un proceso industrial ésta se debe a dos tipos de causas nombradas como: causas comunes y causas especiales de variación, donde las primeras ocurren siempre dentro del proceso y las segundas poseen un carácter de ocurrencia esporádico en el mismo, lo que genera desviación. Montgomery (2009) explica que las causas comunes de variación en un proceso ocurren como resultado del diseño original del mismo y además por las condiciones de funcionamiento que éste posee.

Todos los procesos de manufactura se encuentran sujetos a variabilidad, incluyendo los de estampado y troquelado, la cual se presenta en las características del mismo y afecta en el producto final. En la mayoría de los casos esta variabilidad repercute degradando la calidad del producto. Es posible definir a la variabilidad como las transformaciones que ocurren en las cualidades de una cosa, a menudo se identifican estas transformaciones en el momento en que se reflejan cambios en los valores numéricos asignados para cada cualidad.

De aquí surge la definición para *variable* explicada por Mendenhall et al. (2008) la cual se expresa como sigue:

Variable: es una característica que se modifica con el tiempo para diferentes objetos que se han considerado analizar.



Si se tiene en cuenta dicha definición es posible analizar el producto sobre el cual se desea controlar la variación, Montgomery (2009) menciona que para realizar dicho análisis se debe considerar que un proceso de producción cuenta con un conjunto de elementos de entrada que son transformados en una salida, es decir, en un producto terminado que posee varias características de calidad. Cada uno de estos elementos de entrada poseen diferentes variables que se pueden representar por  $x_1, x_2, \dots, x_n$ , de la misma forma la característica de calidad se puede representar como un variable de salida denotada por  $y$ .

### 3.2.2 Análisis numérico de los datos.

Es posible seleccionar una muestra aleatoria de una población para describir o predecir el comportamiento de la población a través de medidas numéricas que se calculan de la muestra, a las cuales se les llama estadísticos.

Una de las medidas más importantes en un conjunto de datos es: *la media* de la muestra, la cual ayuda a identificar la tendencia que poseen los datos de agruparse hacia el centro de la muestra. Dicha medición puede verse en (3.1). Es sencillo identificar que la media muestral  $\bar{x}$  es la media aritmética de las  $n$  observaciones.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (3.1)$$

A través de la información que se obtiene de la muestra es posible estimar el valor de un parámetro poblacional, esto se hace por medio de un estimador que puede utilizarse de forma puntual o por intervalo. La estimación puntual calcula un solo número para estimar un parámetro poblacional, por otra parte la estimación de un intervalo se realiza calculando dos números que

forman un intervalo en el cual se espera que un parámetro poblacional quede dentro.

Si se tiene que  $n$ , el tamaño de la muestra, es grande, entonces  $\bar{x}$  es el mejor estimador puntual para la media poblacional  $\mu$ , con un error estándar SE mostrado en (3.2) donde  $s$  es la desviación estándar muestral.

$$SE = \frac{s}{\sqrt{n}} \quad (3.2)$$

Es posible construir un *intervalo para la media poblacional*  $\mu$ , por medio del estimador  $\bar{x}$  y de su error estándar, teniendo un coeficiente de confianza  $(1 - \alpha) = .95$  de la forma en que se muestra en (3.3)

$$\begin{aligned} \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \\ \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \\ \bar{x} \pm 1.96 \frac{s}{\sqrt{n}} \end{aligned} \quad (3.3)$$

donde  $z$  es una medida de posición relativa, la cual mide la distancia entre una observación y la media, medida en unidades de desviación estándar de la forma:  $z = (x - \bar{x})/s$ . En (3.3) el valor de  $z$  para el coeficiente de confianza dado anteriormente es 1.96. Además  $n$  es el tamaño de la muestra,  $\alpha$  es el nivel de significancia de 0.05 y  $\sigma$  es la desviación estándar de la población muestreada. En el caso en donde no se conoce el valor de  $\sigma$  se puede utilizar  $s$ , es decir la desviación estándar muestral.

Si se desea cambiar el coeficiente de confianza, la expresión (3.3) puede dejarse como se observa en (3.4), la cual indica que la probabilidad de que  $\mu$  se encuentre en dicho intervalo es igual al coeficiente de confianza deseado.

$$P\left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha \quad (3.4)$$

Es posible medir la variabilidad que se presenta en una muestra de datos de tamaño  $n$  a través de la *varianza muestral*, como se muestra en la ecuación (3.5), donde se identifica que esta medida de la variabilidad de los datos es la suma de la desviación cuadrada de cada observación respecto a la media, dividido todo por la cantidad  $n - 1$ .

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (3.5)$$

Si en la muestra no se presenta variación resultaría que cada  $x_i$  es igual a la  $\bar{x}$  y el resultado de esto es que la varianza muestral  $s^2 = 0$ . De la misma forma mientras más grande sea la varianza muestral, los datos de la muestra presentan una mayor variabilidad.

Es muy común utilizar la raíz cuadrada de  $s^2$  para describir la variabilidad en los datos, a esta medida se le conoce como *desviación estándar muestral*  $s$  la cual se visualiza en (3.6). La  $s$  como medida de variabilidad se encuentra expresada en las unidades originales de las mediciones, lo que permite observar la variabilidad en los datos de forma más sencilla, a diferencia de  $s^2$  que se expresa en el cuadrado de las unidades de los datos. Es importante tener en mente que la desviación estándar muestral no representa la magnitud de los datos, solo refleja la dispersión que estos tiene respecto al promedio.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (3.6)$$

A través de un *Análisis de Varianza* (ANOVA, del inglés, Analysis of Variance) es posible separar la variación total de las mediciones de respuesta en diferentes porciones de variación asignadas a distintos factores de interés.

Es importante identificar que se puede utilizar el ANOVA para 3 diseños experimentales distintos: el primero es un diseño completamente aleatorizado (clasificación de un solo factor), el segundo es un diseño aleatorizado por bloques y el tercero es el experimento factorial  $a \times b$ . Dentro de este documento solo se aborda la explicación y el uso del primer caso, por ser el de mayor utilidad para la investigación.

*ANOVA de un solo factor*: un factor es una variable independiente dentro del sistema a analizar. Al efectuar un análisis adecuado se logra identificar el efecto que produce un factor sobre la variable de respuesta.

Para realizar un análisis de varianza se debe considerar que las observaciones tomadas tienen una distribución normal, con una varianza común  $\sigma^2$ . Además, dependiendo del diseño experimental que se utilice para la obtención de datos, se deben cuidar los supuestos que cada uno englobe.

Si se comparan  $k$  medias poblacionales,  $\mu_1, \mu_2, \dots, \mu_k$ , a través de muestras aleatorias independientes de tamaño  $n_1, n_2, \dots, n_k$ , se considera primero la variación total por medio de la Suma Total de Cuadrados (STC) como se muestra en (3.7)

$$STC = \sum (x_{ij} - \bar{x})^2 = \sum x_{ij}^2 - \frac{(\sum x_{ij})^2}{n} \quad (3.7)$$

donde  $x_{ij}$  es la  $j$ -ésima medida de la  $i$ -ésima muestra y  $n = n_1 + n_2 + \dots + n_k$ .

Además se tienen que:  $STC = SCTr + SCE$  donde  $SCTr$  es la suma de cuadrados de los tratamientos que se utiliza para medir la variación entre las  $k$  medias muestrales y se obtienen según (3.8)

$$SCTr = \sum n_i(\bar{x}_i - \bar{x})^2 = \sum \frac{T_i^2}{n_i} - \frac{(\sum x_{ij})^2}{n} \quad (3.8)$$

donde  $T_i$  es la suma total de las observaciones para el tratamiento  $i$ , y  $n_i$  es el número de las observaciones en la muestra  $i$ .

La suma de cuadrados de los errores ( $SCE$ ) es utilizada para medir la variación agrupada dentro de las  $k$  muestras por medio de (3.9)

$$SCE = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2 \quad (3.9)$$

Cada fuente de variación se divide entre los grados de libertad apropiados para obtener los cuadrados medios ( $CM$ ), cantidades que son utilizadas para calcular el estadístico de prueba  $F$ , que permitirá rechazar o aceptar la  $H_0$ . En la Tabla 3.1 se muestra la construcción de la tabla de análisis de varianza para un diseño completamente aleatorizado.

Tabla 3.1 Construcción de una tabla del análisis de varianza.

Fuente	Grados de libertad (gl)	Suma de cuadrados (SC)	Cuadrados medios (CM)	F	Valor p
Tratamientos	$k - 1$	$SCTr$	$CMT_r = STC / (k - 1)$	$CMT_r / CME$	*
Error	$n - k$	$SCE$	$CME = SCE / (n - k)$		
Total	$n - 1$	$STC$			

\*NOTA: valor  $p$  = área de extremo "más allá" del valor observado de la prueba estadística  $z$ .

Para el análisis de varianza mostrado en la tabla anterior se prueban las siguientes hipótesis:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

$H_a$ : uno o mas pares de medias poblacionales son diferentes

El estadístico de prueba F se basa en  $gl_1 = (k - 1)$  y  $gl_2 = (n - k)$  grados de libertad. Se rechaza  $H_0$  si  $F > F_\alpha$ , donde  $F_\alpha$  queda ubicada en el extremo superior de la distribución F, de igual forma se rechaza  $H_0$  si el valor  $p \leq \alpha$ , y los resultados son significativos estadísticamente en  $\alpha$ , el nivel de significancia considerado.

### 3.2.3 Análisis gráfico de los datos.

Existen una gran cantidad de técnicas gráficas utilizadas para describir y representar el comportamiento de una muestra de datos, entre las cuales resaltan el histograma, el diagrama de caja y el diagrama de dispersión.

El *histograma* mostrado en la Figura 3.1 resulta de graficar las 50 observaciones mostradas en la Tabla 3.2 provenientes de una distribución normal estándar con  $\mu = 0$  y  $\sigma = 1$ .

En el histograma se representa la escala de mediciones para los datos sobre el eje horizontal, lo que se conoce como “clases” y sobre el eje vertical se representa la frecuencia de los mismos. Un histograma permite ver la forma de la distribución que siguen los datos y la ubicación respecto a la media de los mismos, además muestra cierta información acerca de la variabilidad inherente de los mismos.

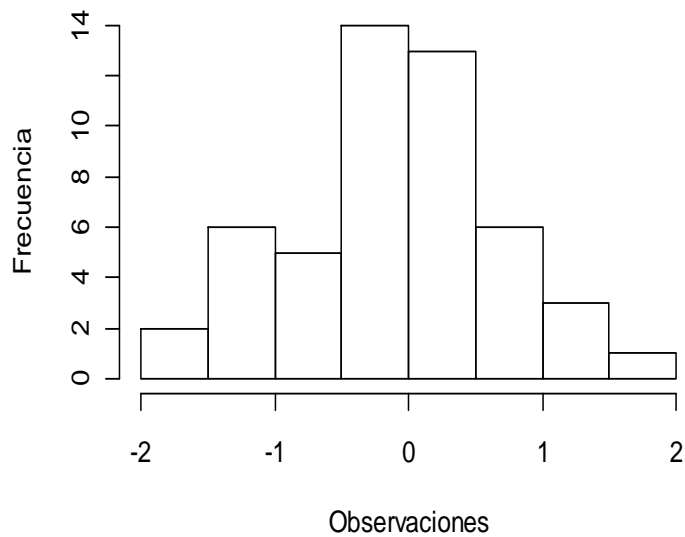


Figura 3.1 Histograma para las observaciones de la Tabla 3.2.

Tabla 3.2 Observaciones  $X \sim N(0,1)$

0.0848	-0.0995	0.2195	-1.0476	0.9034
0.2265	-0.8183	0.5006	0.7008	0.3968
0.4326	-0.4693	1.6693	0.7578	-0.1451
0.5581	0.5750	-1.0035	-0.7781	-0.1021
0.0598	-1.8717	0.3942	-1.3132	-0.3280
-0.1146	-0.6318	-1.4971	-0.3713	-0.0642
-1.0206	-0.0424	1.2393	-0.0847	-1.8648
-0.2971	1.4421	0.0376	-0.2587	0.2774
0.1682	-0.9209	-1.0396	0.2179	-0.5039
1.4199	-0.0156	0.1537	0.3699	-0.1302

El *diagrama de caja* es un gráfico que muestra de forma simultánea 5 características importantes de un conjunto de datos: muestra el cuartil inferior  $Q_1$ , la mediana  $m$ , el cuartil superior  $Q_3$ , el valor máximo y el valor mínimo del conjunto de datos.

De acuerdo con Morgenthaler (2009) es importante notar que en este diagrama se utiliza la mediana como medida de tendencia central y no la

media, ya que la mediana posee mayor resistencia, cuando dentro del conjunto de datos se encuentran valores muy extremos y/o atípicos, la mediana no se ve afectada. Los cuartiles  $Q_1$  y  $Q_3$ , se encuentran sobre una caja rectangular, como se observa en la Figura 3.2, mientras que los valores máximo y mínimo se grafican fuera de la caja en apariencia como de bigotes. La línea central es el segundo cuartil  $Q_2$  que representa la mediana de los datos,  $Q_2 = m$ .

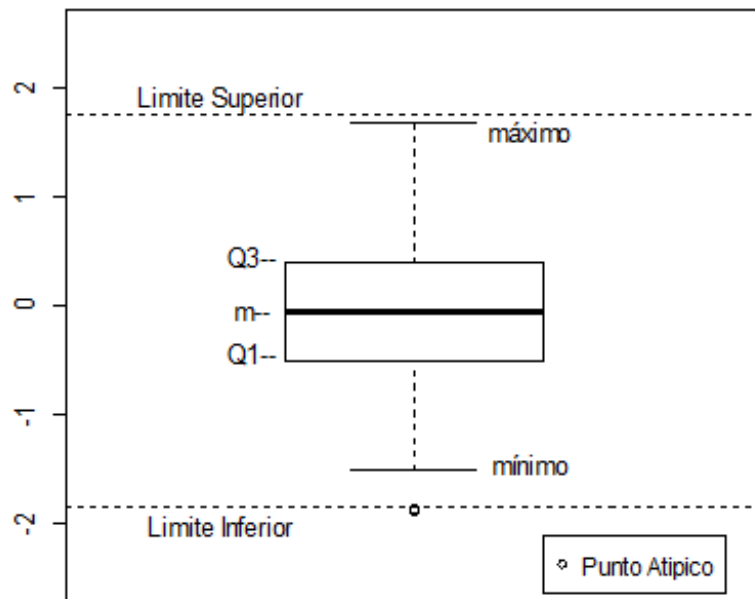


Figura 3.2 Diagrama de caja para las observaciones de la Tabla 3.2

El rango intercuartílico  $Q_3 - Q_1$  es utilizado para separar los valores extremos y generar los límites inferior y superior en el gráfico, actuando como un indicador para la variación de los datos, cualquier medición fuera de estos límites se considera como un valor atípico. El límite inferior y el límite superior están dados por  $Q_1 - 1.5(Q_3 - Q_1)$  y  $Q_3 + 1.5(Q_3 - Q_1)$  respectivamente.

La gráfica de caja es útil para describir la forma de una distribución de datos, lo cual se hace a través de observar la posición de la  $m$  comparada con  $Q_1$ , y  $Q_3$ , ya que si la  $m$  se encuentra ubicada en la mitad de la caja, entonces



la distribución es simétrica. Si la  $m$  se ubica en la parte baja de la caja con referencia del centro, la distribución se encuentra sesgada hacia la derecha y viceversa.

Cuando se tiene un conjunto de datos pertenecientes a dos variables, formando  $n$  pares de observaciones  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , esto se conoce como *datos bivariados*, se puede pensar de forma natural que ambas variables  $x$  y  $y$  tienen una relación positiva o negativa, esto ocurre si los valores grandes de las  $x$ 's se combinan formando pares con valores grandes de las  $y$ 's o si los valores pequeños de las  $x$ 's forman pares con valores pequeños de las  $y$ 's, respectivamente. Dicha relación entre las variables, se puede describir formando un *diagrama de dispersión* y observando aquellos patrones reflejados en dicho diagrama. Cada par de datos se grafican con un punto en el plano, de la forma en que se muestra en la Figura 3.3 Diagrama de dispersión Figura 3.3.

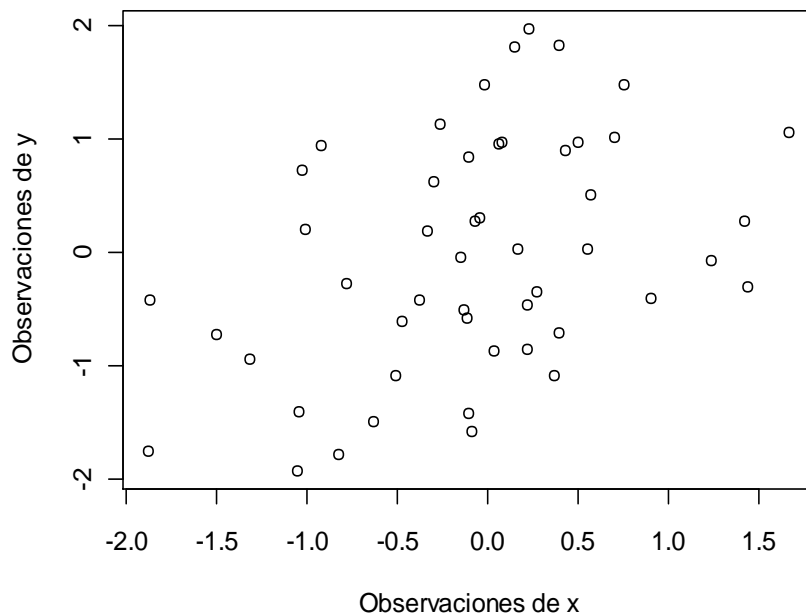


Figura 3.3 Diagrama de dispersión

Uno de los patrones más comunes que se visualizan en los diagramas de dispersión es una razón de incremento constante o decrecimiento constante, al cual se le identifica como un patrón lineal o como una relación lineal.

El coeficiente de correlación muestral  $r$  es una medida utilizada para describir la fuerza de la relación entre las mediciones de  $x$  y  $y$ , dicho coeficiente se define como se muestra en (3.10), donde las cantidades  $s_x$  y  $s_y$  son las desviaciones estándar para las variables  $x$  y  $y$  respectivamente, según (3.6). La parte  $s_x s_y$  se llama covarianza entre las variables y se define según la ecuación (3.11).

$$r = \frac{s_{xy}}{s_x s_y} \quad (3.10)$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1} \quad (3.11)$$

El valor que toma el coeficiente de correlación  $r$  se encuentra siempre entre -1 y +1, cuando se presenta que  $r = 0$  se concluye que no existe una relación lineal entre las dos variables, sin embargo puede existir otra forma en que se relacionen las variables. Mientras que  $r$  se aproxime más a los valores de 1 o -1 la relación lineal entre las dos variables es muy fuerte. A menudo las dos variables se clasifican de la siguiente forma: variable dependiente  $y$  y variable independiente  $x$ .

### 3.2.4 Modelos y Técnicas estadísticas para el análisis de datos.

El *Modelo de Regresión Lineal Simple* es utilizado para modelar una variable de respuesta  $y$  que es una función de una variable independiente. Suponiendo que  $x$  y  $y$  poseen una relación que puede ser descrita por medio de una recta con ecuación (3.12)

$$y = \beta_0 + \beta_1 x \quad (3.12)$$

donde  $\beta_0$  es el valor de  $y$  cuando  $x$  es igual a 0 y  $\beta_1$  es la pendiente de la recta, ambos son parámetros desconocidos. Este modelo es conocido como un modelo determinista porque no incluye ningún error de predicción en la variable  $y$ .

En contraste con el modelo determinista, el modelo probabilístico mostrado en (3.13) proporciona una descripción más precisa de la realidad.

$$y = \beta_0 + \beta_1 x + \epsilon \quad (3.13)$$

donde  $\epsilon$  es una variable aleatoria que posee una distribución de probabilidad normal con media cero.

Para estimar la recta de mejor ajuste mostrada en (3.14) se utilizan los estimadores de mínimos cuadrados  $\hat{\beta}_0$  y  $\hat{\beta}_1$ , los cuales pueden ser calculados por medio de (3.15) y (3.16)

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \epsilon \quad (3.14)$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \quad (3.15)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (3.16)$$

El componente de error aleatorio  $\epsilon$  es utilizado para explicar las desviaciones de los puntos con respecto a la recta. La recta ajustada para los datos de la Tabla 3.2 se muestra en la Figura 3.4, la cual se obtuvo por medio del paquete R `cran` y es la siguiente:  $\hat{y} = 0.5264 + 0.007157 x + \epsilon$ .

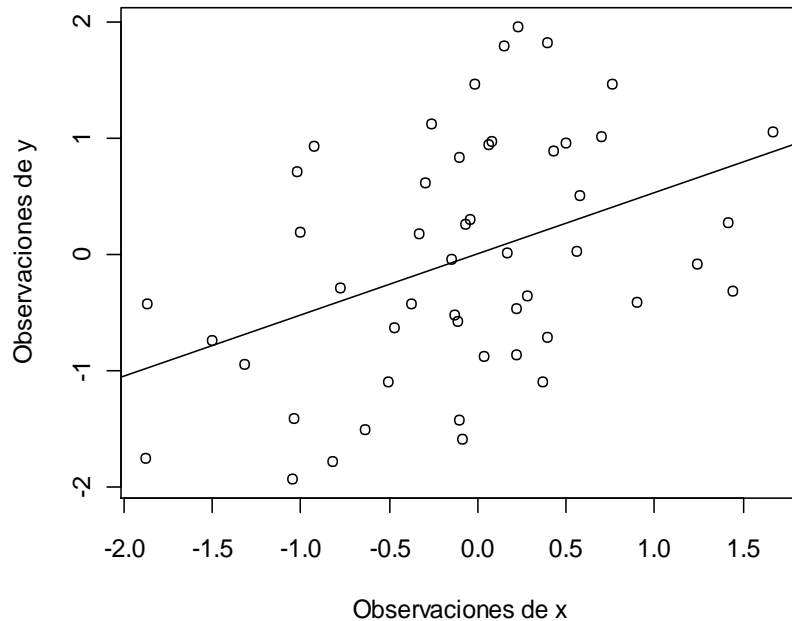


Figura 3.4 Recta de regresión o recta de mínimos cuadrados.

Para determinar que  $x$  es útil para predecir el valor de  $y$  se utiliza un ANOVA para la *regresión lineal*, de forma similar a lo mostrado en la Sección 3.2.2, en donde la respuesta  $y$  se relaciona con la variable independiente  $x$ . De aquí que la variación total en la variable de respuesta  $y$ , está dada por (3.17)

$$STC = s_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n} \quad (3.17)$$

y se divide en dos partes: Suma de Cuadrados para la Regresión ( $SCR$ ) y la Suma de Cuadrados para el Error ( $SCE$ ). La primera mide la cantidad de variación explicada al usar la recta de regresión con una variable independiente  $x$ , y la segunda mide la variación residual en los datos que no es explicada por la variable independiente  $x$ .

Cuando cada una de las fuentes de variación es dividida entre los grados de libertad apropiados, se obtiene una estimación de la variación en el

experimento. Tales estimaciones se conocen como *CM*. Los grados de libertad totales para las  $n$  mediciones son  $(n - 1)$ . Para evaluar la recta de regresión, se requiere estimar un parámetro adicional  $\beta_1$ , por lo que hay un grado de libertad relacionado con *SCR*, dejando  $(n - 2)$  grados de libertad con *SCE*. Las hipótesis relacionadas a este análisis son:  $H_0: \beta_1 = 0$  contra  $H_a: \beta_1 \neq 0$ .

Los resultados de un análisis de regresión lineal son válidos sólo cuando la relación entre  $y$  y  $x$  está dada por (3.14) y los valores de  $\epsilon$  son independientes en el sentido probabilístico, tienen una media de 0 y una varianza común igual a  $\sigma^2$  y poseen una distribución de probabilidad normal, esto es lo que se conoce como supuestos de regresión.

Es posible utilizar la recta de mejor ajuste para estimar el valor promedio de  $y$  cuando  $x = x_0$ , por medio de (3.18)

$$y_{IC95\%, x_0} = \bar{y} + \beta_1(x_0 - \bar{x}) \pm (t_{n-2, \alpha/2})(s) \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{s_x^2(n-1)}} \quad (3.18)$$

Así mismo la recta de mejor ajuste se utiliza para predecir un valor particular de  $y$  cuando  $x = x_0$  como se muestra en (3.19)

$$y_{IP95\%, x_0} = \bar{y} + \beta_1(x_0 - \bar{x}) \pm (t_{n-2, \alpha/2})(s) \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{s_x^2(n-1)}} \quad (3.19)$$

donde en ambas ecuaciones:  $x_0$  representa un valor particular de  $x$ , y  $s$  representa la mejor estimación insesgada de  $\sigma$ , que mide el error residual. La estimación y la prueba se basan en el estadístico mostrado en (3.20)

$$t = \frac{\hat{y} - E(y)}{SE(\hat{y})} \quad (3.20)$$

que tiene una distribución  $t$  con  $(n - 2)$  grados de libertad, y  $SE$  indica el error estándar de  $\hat{y}$ .

### *Distribución binomial.*

En los procesos industriales, existen muchas situaciones que presentan las características de un *ensayo Bernoulli*, es decir presentan solo dos resultados posibles cuando se estudia su comportamiento.

Un único ensayo Bernoulli clasifica a una variable aleatoria  $X$ ; que mide el número de éxitos, en dos posibles categorías: éxito o fracaso, en tal caso la variable aleatoria  $X$  se distribuye como una Bernoulli de parámetro  $p$ ,  $X \sim B_e(p)$ , donde  $p$  es la probabilidad de éxito, tal como se muestra en (3.21)

$$f(X = x) = p^x(1 - p)^{1-x} \quad (3.21)$$

por su parte  $x$  solo puede tomar dos valores posibles que son 0 y 1, La función de distribución de probabilidad para este caso se define como:

$$f(x; p) = \begin{cases} p; & \text{si } x = 1 \\ q; & \text{si } x = 0 \\ 0; & \text{en cualquier otro caso} \end{cases}$$

El desarrollo de *la distribución binomial* parte de la repetición de los ensayos Bernoulli, es decir consta con  $n$  ensayos idénticos con probabilidad de éxito  $p$ , la cual es la misma de un ensayo a otro y la probabilidad de fracaso es  $q = (1 - p)$ . Los ensayos deben ser independientes y se tiene interés en  $x$ , que es el número de éxitos observado. De esta forma la probabilidad de  $k$  éxitos en  $n$  ensayos es determinada por (3.22) para valores de  $k = 0, 1, 2, \dots, n$ .

$$P(x = k) = C_k^n p^k q^{n-k}$$

$$P(x = k) = C_k^n p^k (1 - p)^{n-k} \quad (3.22)$$

donde:

$$C_k^n = \frac{n!}{k!(n-k)!}$$

En este caso la variable aleatoria  $X$ , tiene una distribución de probabilidad con media y varianza indicadas en (3.23) y (3.24), respectivamente.

$$\mu = np \quad (3.23)$$

$$\sigma^2 = \mu(1 - p) \quad (3.24)$$

Una aplicación de la variable aleatoria binomial  $X$ , es cuando se desea conocer cuántos productos defectuosos están presentes en un proceso, en este caso se utiliza una muestra aleatoria de  $n$  productos para lograr estimar la fracción de productos defectuosos  $p$  en la población. Si  $x$  de los productos muestreados resultan defectuosos, entonces se puede utilizar la proporción muestral dada por (3.25), para estimar la proporción poblacional  $p$ .

$$\hat{p} = \frac{x}{n} \quad (3.25)$$

Dado que  $\hat{p}$  es el valor de  $x$ , expresado como una proporción en (3.25), la distribución muestral de  $\hat{p}$  es muy parecida a la distribución de  $x$ , pero posee una escala diferente en el eje horizontal, lo cual genera una media y varianza mostradas en (3.26) y (3.27).

$$\mu = p \quad (3.26)$$

$$\sigma^2 = \frac{p(1-p)}{n} \quad (3.27)$$

### 3.2.5 Métodos para el control de calidad y muestreo para aceptación.

De acuerdo con Canavos (1998), el *muestreo de aceptación* está relacionado con la inspección y la toma de decisiones con respecto a los componentes o materias primas que serán utilizados en un proceso.

Una aplicación típica del muestreo de aceptación es como sigue: una compañía recibe un lote (tarima) de un proveedor. Este producto es a menudo un componente o materia prima usada en el proceso de manufactura de la compañía. Una muestra es tomada del lote y una o varias características de calidad en la muestra son inspeccionadas. En base a la información en la muestra se toma una decisión relacionada a la disposición del lote. Usualmente, esta decisión es aceptar o rechazar el lote.

Algunas veces se nombra esta decisión como: sentencia de lote. Los lotes aceptados son utilizados en el proceso, mientras que los lotes que son rechazados, se regresan al proveedor o pueden ser sometidos a alguna otra disposición del lote. En ocasiones se puede pensar que el muestreo de aceptación es utilizado como una actividad de inspección al recibir elementos del proveedor, sin embargo existen otros métodos de muestreo. Por ejemplo, una empresa puede muestrear e inspeccionar su propio producto en varias etapas de la producción.

De acuerdo con Montgomery y Runger (2007) existen 3 enfoques para declarar la disposición del lote: la primera es aceptar el lote sin inspeccionarlo. La segunda es inspeccionar cada producto del lote, es decir realizar una inspección al 100% y la tercera es utilizar un muestreo por aceptación.



El muestreo de aceptación es útil en muchas situaciones, por ejemplo:

- Cuando el costo de una inspección al 100% es extremadamente alto.
- Cuando se requiere un programa para monitorear el producto continuamente, aunque se sabe que el proceso del proveedor tiene un historial de calidad bueno.
- Cuando la inspección al 100% impacta al tiempo planeado de la producción.

Los planes de muestreo por aceptación se clasifican en varias formas, Ojeda (2005) menciona que la principal clasificación utilizada es por variables y atributos. Los planes de muestreo por variables requieren mediciones en una escala numérica, ya que una medición física es probable que proporcione mucho más información útil con respecto a la calidad de un producto. Por el contrario los planes de muestreo por atributos observan unidades disconformes o defectuosas, ya que muchas de las características de calidad sólo son observables como atributos.

Para la aplicación de un *plan de muestreo por variables*, R Development Core Team (2009) afirma que es necesario que la distribución de la característica de calidad sea conocida. Por cada característica de calidad que se va a inspeccionar, se requiere un plan de muestreo diferente, para que éste sea adecuado a cada una en particular.

El desarrollo de planes efectivos incluye tanto al proveedor y al cliente, en donde el riesgo del proveedor (riesgo  $\alpha$  o cometer el error tipo I)<sup>2</sup> consiste en

---

<sup>2</sup> El nivel de significancia  $\alpha$  es el riesgo que se desea tener al tomar una decisión incorrecta. Para un plan de muestreo:  $\alpha = P(\text{error tipo I}) = P(\text{Rechazar materia prima buena}) = \text{Riesgo de proveedor}$ .

que el cliente rechace una materia prima que es buena, por su parte el riesgo del cliente (riesgo  $\beta$  o cometer el error tipo II) consiste en aceptar una materia prima que es inaceptable.

Dado que  $\alpha$  es el riesgo del proveedor y  $\mu_\alpha$  es la media de alguna característica de calidad de la materia prima, la probabilidad de aceptación es  $1 - \alpha$ . De manera similar,  $\beta$  es el riesgo del cliente y  $\mu_\beta$  es la media de alguna característica de calidad de la materia prima para la cual la probabilidad de aceptación es  $\beta$ . Dados los valores numéricos de  $\alpha$ ,  $\beta$ ,  $\mu_\alpha$  y  $\mu_\beta$ , el plan de muestreo por variables para medias consiste en determinar el tamaño de la muestra  $n$  y un valor de aceptación  $\bar{x}_a$ , de forma que, cuando el valor medido de la media de la muestra  $\bar{X}$  es mayor que  $\bar{x}_a$ , la materia prima será aceptada.

Con el fin de obtener  $\bar{x}_a$  y  $n$ , se considera el riesgo del proveedor de la forma mostrada en (3.28) y (3.29):

$$P(\bar{X} \leq \bar{x}_a) = \alpha \quad (3.28)$$

$$P\left(Z \leq \frac{\bar{x}_a - \mu_\alpha}{\sigma/\sqrt{n}}\right) = \alpha \quad (3.29)$$

donde

$$\frac{\bar{x}_a - \mu_\alpha}{\sigma/\sqrt{n}} = z_\alpha \quad (3.30)$$

así mismo para el riesgo del cliente se tiene (3.31) y (3.32):

$$P(\bar{X} > \bar{x}_a) = \beta \quad (3.31)$$

$$P\left(Z \leq \frac{\bar{x}_a - \mu_\beta}{\sigma/\sqrt{n}}\right) = \beta \quad (3.32)$$

donde

$$\frac{\bar{x}_a - \mu_\beta}{\sigma/\sqrt{n}} = z_{1-\beta} \quad (3.33)$$

Las ecuaciones obtenidas en (3.30) y (3.33) dependen de las incógnitas  $\bar{x}_a$  y  $n$ . Al despejar  $\bar{x}_a$  de las ecuaciones anteriores se tiene que:

$$\bar{x}_a = (\sigma/\sqrt{n}) z_\alpha + \mu_\alpha \quad (3.34)$$

$$\bar{x}_a = (\sigma/\sqrt{n}) z_{1-\beta} + \mu_\beta \quad (3.35)$$

Al igualar (3.34) y (3.35) y resolver para la incógnita  $n$ , de acuerdo con el apéndice al final del documento, se obtiene:

$$n = \left[ \frac{\sigma(z_{1-\beta} - z_\alpha)}{\mu_\alpha - \mu_\beta} \right]^2 \quad (3.36)$$

Cuando se emplea (3.36) para obtener el tamaño de la muestra, el valor de aceptación  $\bar{x}_a$  se obtiene, ya sea de la ecuación (3.34) o de (3.35).

### *Curva Característica de Operación*

Ramos (2010) explica que una gráfica de la probabilidad de aceptación  $P(A)$  contra la media de una característica, es la curva de operación característica (CO). Mediante el empleo de la distribución de probabilidad normal dada por (3.37) y la estandarización de la variable aleatoria normal  $x$  señalada en (3.38), se obtiene la probabilidad de aceptar diferentes valores de la media de la característica analizada.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \quad -\infty < x < \infty \quad (3.37)$$

$$z = \frac{x - \mu}{\sigma} \quad (3.38)$$

Por su parte Kiermeier (2008) señala que una CO es afectada por el tamaño  $n$  de la muestra y por el criterio de aceptación  $\bar{x}_\alpha$ . Conforme  $n$  aumenta, la pendiente de la curva se torna más pronunciada. De esta forma, para tamaños grandes de la muestra, la probabilidad de aceptación aumenta muy lentamente conforme el valor de la media aumenta. Si el valor de  $n$  es fijo, un aumento en el criterio de aceptación  $\bar{x}_\alpha$  tendera a desplazar la curva hacia la derecha. Esto implica que para una media dada en la característica analizada la probabilidad de aceptación es baja conforme  $\bar{x}_\alpha$  aumenta. Además los planes de muestreo con valores grandes de  $\bar{x}_\alpha$  son mejores siempre que el tamaño de la muestra sea grande. El comportamiento de la curva se verifica en el Capitulo 5 utilizando los datos correspondientes al caso de estudio de esta investigación.

De acuerdo con Canavos (1998) los riesgos del productor y del consumidor pueden representarse por dos puntos  $(\mu_\alpha, 1 - \alpha)$  y  $(\mu_\beta, \beta)$  sobre una curva característica de operación, en este contexto,  $\mu_\alpha$ , recibe el nombre de nivel aceptable de calidad, y  $\mu_\beta$  el de tolerancia de los defectuosos en el lote. Para tales puntos, comúnmente se escoge la probabilidad de aceptación  $P(A) = 1 - \alpha$  cercano al punto 0.95 de la curva, y la probabilidad de aceptación  $P(A) = \beta$  cercano al punto 0.10 sobre la curva. Es decir, el 95% de los lotes que provienen de un proceso con una  $\mu_\alpha$  serán aceptados, mientras que sólo el 10% de los que provienen de un proceso con una  $\mu_\beta$  serán aceptados. Esto representa que ambos puntos dividen el área bajo la curva en tres zonas, es decir los lotes se dividen en tres categorías:

- Lotes buenos, los cuales poseen una  $P(A) \geq 0.95$ ,
- Lotes intermedios, los cuales poseen varias probabilidades de aceptación,  $0.10 < P(A) < 0.95$ ,

- Lotes malos, los cuales poseen una  $P(A) \leq 0.10$ .

La principal ventaja del plan de muestreo por variables es que la misma curva de operación puede ser obtenida con un tamaño de muestra más pequeño, que el que se requiere para un plan de muestreo por atributos. Además cuando se emplean pruebas destructivas para verificar un lote o producto, el muestreo por variables es mucho más útil ya que reduce los costos de inspección. Una segunda gran ventaja es que las mediciones en escala numérica, como es el caso del plan de muestreo por variables, proveen más información del proceso de manufactura o el lote a verificar. Generalmente, una medición numérica de la característica de calidad es más útil que una clasificación simple como la de conforme y no-conforme.

## **CAPÍTULO 4 MATERIALES Y MÉTODOS.**

### **4.1 Método.**

A continuación se explica de forma general cada una de las etapas que involucra el método propuesto por esta investigación, el cual se observa en la Figura 4.1 (pág. 58) en forma esquemática. La etapa 1 del método se desarrolla en el capítulo 2, las etapas 2 al 5 se desarrollan dentro de este capítulo y finalmente las etapas 6 al 9 corresponden a los capítulos 5 y 6.

#### **4.1.1 Descripción del método.**

*ETAPA 1: Definición de una problemática relacionada al sector industrial.*

La problemática se establece en base a las necesidades presentes dentro de la empresa, así como en relación a los intereses particulares de la misma, esto con el fin de dar solución a una situación dentro del sector industrial. En esta etapa se realizan las siguientes actividades:

- Se identifican los tipos de defectos en el producto.
- Se reconoce la parte del proceso con mayor ocurrencia de defectos.
- Se registra la repercusión monetaria que ocasionan los defectos.
- Se identifican las acciones correctivas que los responsables han tomado hasta el momento.

Así mismo es importante analizar si esta problemática se ha presentado

en otros entornos, esto se realiza por medio de una revisión de la literatura. Para el caso de estudio, esta etapa del método se desarrolla en el Capítulo 2 del documento.

*ETAPA 2: Diagnóstico del entorno y pruebas iniciales.* En esta etapa se realiza un análisis previo del entorno, el objetivo es encontrar evidencia que permita expresar al problema en forma científica. Un enfoque dentro de esta etapa es modelar la proporción diaria de productos defectuosos, esto se realiza por medio de la información histórica disponible de dichos datos, la mayoría de las empresas que cuentan con procesos de conformado de metales poseen información de este tipo, lo que permite su análisis. Se utiliza la distribución binomial para alcanzar este primer paso de la etapa, como se muestra en la sección 4.2.1

Como segundo enfoque se busca describir la variación presente en el producto, para esto se toman datos de una variable de entrada y de la variable de salida; Montgomery (2007) menciona que es necesario identificar que la variable de entrada y la variable de salida proporcionen en realidad información útil acerca del proceso en estudio. Enseguida se realizan pruebas estadísticas para determinar el efecto que la variable de entrada produce en la variable de salida, para desarrollar lo anterior, se llevan a cabo una prueba de dispersión de datos bivariados y un análisis de regresión lineal entre las variables.

La información que entregan las pruebas realizadas dentro de esta etapa, establece el punto de partida para la selección de aquellas variables que conformarán la base de datos que será analizada. En la sección 4.2.1 se desarrollan las herramientas utilizadas en el diagnóstico para el caso de estudio.

*ETAPA 3: Selección de variables involucradas.* Se identifican aquellos elementos o partes del proceso que se desea considerar para el análisis, así mismo se prosigue con una descripción de todas las variables relacionadas a dichos elementos. La descripción a realizar debe mostrar de forma ordenada el nombre y la especificación del valor numérico de cada una de ellas. Se debe explicar la relación que tienen con el proceso y finalmente seleccionar, con apoyo del personal de la empresa, aquellas en las cuales se tiene la certeza de que son útiles para el estudio de la problemática.

*ETAPA 4: Recolección y estructuración de la base de datos.* En esta etapa se reúnen datos de cada una de las variables que fueron seleccionadas en la etapa anterior, estos son tomados directamente del proceso. Debe registrarse el nivel de las entradas para cada medición en la variable de salida.

Es importante tener en mente que en los procesos de estampado y troquelado es común que por cada configuración en los valores de las entradas, se obtengan cientos de medidas en la variable de salida, o que incluso existan repeticiones en las mediciones. Según sea factible pueden obtenerse la mayor cantidad de datos provenientes de diferentes configuraciones, pero se debe tener en cuenta que esto es costoso, además de las implicaciones de tiempo que esto conlleva.

Cuando se han tomado las muestras de tamaño  $n$ , se prosigue a reunir todos los datos en un solo archivo de Excel, se ordena cada valor en la variable de entrada con su valor correspondiente en la variable de salida. Una vez que se tenga el archivo en Excel, es exportado como archivo tipo “nombre.csv” al paquete R para su análisis. Existen otros tipos de archivos que son compatibles con diferentes paquetes estadísticos, entre los más utilizados se encuentran los archivos “nombre.txt”, para esta investigación se utilizaron los primeros.



*ETAPA 5: Verificación de la información obtenida.* Dentro de esta etapa ya se dispone de una base de datos, sin embargo se debe considerar verificar la totalidad de la base de datos, ya que es posible que existan errores de medición, calibración inadecuada de la herramienta de medición, fallas en la máquina, información no confiable del proveedor o cualquier otro caso que ponga en duda la información en la base de datos.

No obstante, si se tiene la seguridad de que la información es confiable, entonces se puede emplear la base de datos actual para continuar con la etapa 6 del método.

*Etapa 5.1 Reestructurar.* En caso de que se presente alguna de las situaciones anteriores es necesario reestructurar esa parte de la base de datos, realizar lo necesario para obtener una nueva información que sea coherente y confiable. Una vez concluido esto, se puede proceder al análisis de la base de datos como se indica en la etapa 6.

*ETAPA 6: Análisis de la base de datos.* Al momento de disponer de una base de datos confiable, se nombra a cada una de las variables de entrada como  $x_i$ , donde  $i = 1, 2, \dots, 8$ , y a la variable de salida como  $y$ . Dentro de esta etapa se realiza un análisis tanto numérico como gráfico, con el propósito de que los resultados sean respaldados. Se inicia con el análisis de la variable de salida de forma individual, y un segundo enfoque es el de involucrar las variables de entrada con relación a la variable de salida.

*Etapa 6.1 Análisis de la variable de salida.* Es importante que en este punto del estudio se analice la variable  $y$  por medio de un ANOVA de un solo factor, para identificar el tratamiento que produce mejores mediciones en la variable analizada, es decir, dentro del rango de especificación. Además se

pueden utilizar herramientas visuales como el diagrama de caja e histogramas para confirmar los resultados numéricos del ANOVA.

*Etapa 6.2 Análisis de correlación.* Para esta etapa se calcula el coeficiente  $r$  para cada  $x_i$  con  $y$ , posteriormente se debe identificar el coeficiente con un valor numérico absoluto más alto para determinar cual variable de las  $x_i$  presenta una mayor relación lineal con  $y$ , dicha relación será descrita a través de un modelo de regresión lineal en la siguiente etapa.

Las condiciones de análisis del problema planteado permiten realizar estudios de regresión lineal múltiple, entre otros; sin embargo, para obtener bases de datos de todas las variables involucradas existe la siguiente limitante: in-factibilidad práctica de disponer de los equipos en la empresa para tal fin. No obstante, la metodología de análisis de regresión lineal simple permite estudiar los datos disponibles, dentro de la cual se seleccionó la variable con mejor correlación.

*ETAPA 7: Aplicación del Modelo de Regresión Lineal.* Se realiza el análisis de regresión lineal con el fin de explicar el efecto de la variable de entrada sobre la variable de salida. Posteriormente se calculan los coeficientes  $\beta_0$  y  $\beta_1$  para establecer la ecuación de la recta de regresión y se estudia si el modelo posee un buen ajuste, de lo contrario se debe verificar si existen anomalías, tal como ocurrió en el caso de estudio donde se encontraron puntos atípicos.

Así mismo se verifican los supuestos del modelo y se realiza un ANOVA para la regresión. Finalmente se establecen los intervalos de confianza. Se recomienda utilizar las gráficas correspondientes para mostrar los resultados en forma visual y no solo numéricamente.

*ETAPA 8: Diseño del plan de muestreo de aceptación para una variable.*

En base a la variable seleccionada en la etapa 6, se diseña un plan de muestreo de aceptación para la misma, con el fin de obtener un tamaño de muestra factible a inspeccionar y un criterio de aceptación. Se debe considerar los riesgos tanto de la empresa como del proveedor. Además se requiere realizar una curva de operación característica para ver el desempeño del plan de muestreo propuesto.

*ETAPA 9: Análisis de resultados.* En esta etapa se comparan los resultados obtenidos en los análisis contra los objetivos que se plantearon al inicio del método. Si los resultados son los deseados, se establecen conclusiones y futuras aplicaciones del método para otros productos.

En caso de que los resultados no sean los esperados, se debe revisar cada una de las etapas del método, para verificar que no se hayan cometido errores, si se tiene confianza en que ciertas etapas se ejecutaron correctamente, se pueden revisar solo las etapas significativas, tales como: la Etapa 5 y la Etapa 6.

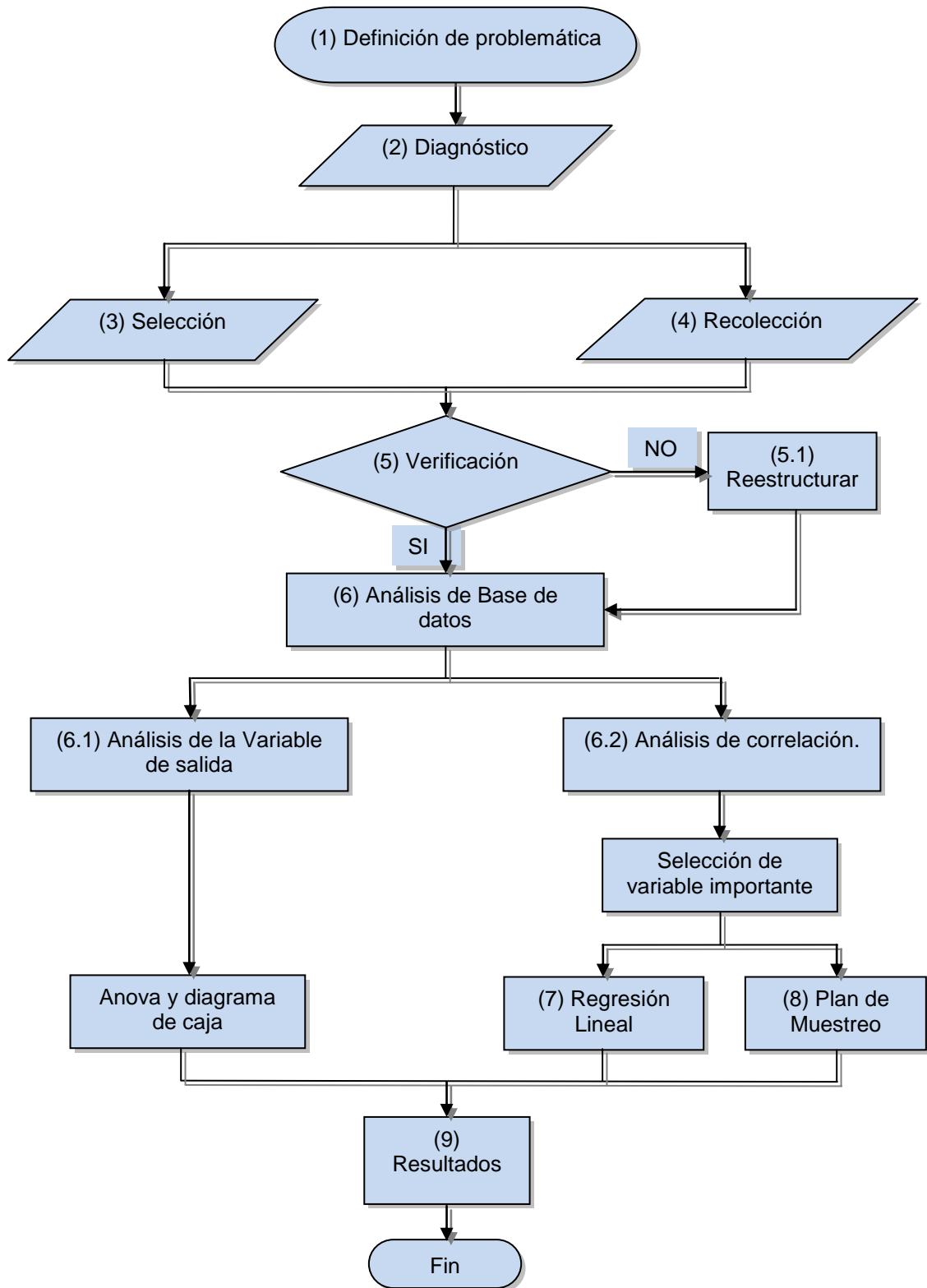


Figura 4.1 Esquema general del método.

## 4.2 Desarrollo del método.

A continuación se describen las etapas relacionadas con los materiales y métodos de la investigación.

### 4.2.1 Etapa 2: Diagnóstico del entorno y pruebas iniciales.

Se realizó un análisis de diagnóstico de la problemática, lo que permitió identificar en mejor forma lo que ocurre con el porcentaje de defectuosos y la variación del producto. Se utilizó la distribución binomial para modelar la fracción de defectuosos  $p$ , que se presentan en la producción de álabes, la cual es ocasionada por la ocurrencia del defecto principal: Golpes. Cada pieza que presenta dicho defecto es considerada defectuosa.

Se utilizaron registros de la producción diaria que comprenden un historial de 5 meses (114 días hábiles) y de acuerdo a los datos se obtuvo la información mostrada en la Tabla 4.1.

Tabla 4.1 Valores de los parámetros para la distribución binomial.

Parámetro	Valor
$K$	40 695
$N$	13 509 000
$p$	0.0030124

Se selecciona una muestra aleatoria de  $n = 118,500$  álabes (producción media por día) de una población binomial con parámetro  $p$ , y sobre ésta se efectúa la revisión para determinar la producción defectuosa, entonces, el número de defectivos sigue una distribución binomial; la probabilidad de encontrar  $k$  productos defectuosos en cualquier día se determina por medio de

(4.1). Las estimaciones puntuales de la media y la varianza de la variable aleatoria binomial,  $p$ , son 356.97 y 355.89, respectivamente.

$$P(x = k) = \binom{118500}{k} 0.0030124^k (1 - 0.0030124)^{118500-k} \quad (4.1)$$

Por su parte,  $\hat{p}$ , como fracción de defectuosos muestral, tiene media y varianza de  $p = 0.0030124$  y  $2.53445 \times 10^{-8}$  respectivamente. Considerando un valor objetivo que desea la empresa para  $p$  de 0.0015, entonces, la media y varianza deseadas son, 0.0015 y  $1.26392 \times 10^{-8}$  respectivamente, bajo una producción diaria de 118500 piezas. Así, la distribución de probabilidad meta para  $k$  tiene media de 177.5 y varianza de 177.23, respectivamente.

Utilizando este enfoque para representar la fracción de productos defectuosos se logró deducir que existe un desfase entre las dos distribuciones de probabilidad consideradas como se muestra en la Figura 4.2, estas distribuciones son: distribución deseada y distribución actual.

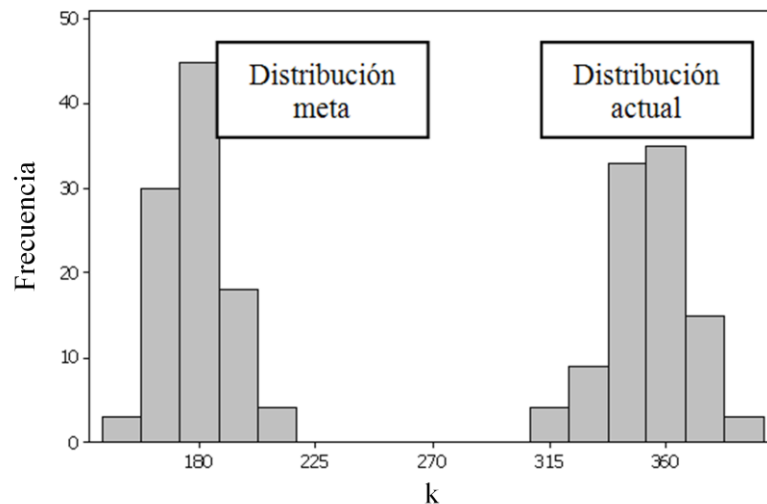


Figura 4.2 Desfase en las distribuciones de probabilidad del proceso

Como segunda prueba de diagnóstico, se involucraron dos variables del proceso dentro del análisis estadístico con las cuales se llevaron a cabo pruebas de dispersión, análisis de regresión e intervalos de confianza. Se presenta una descripción detallada de los elementos y las variables que intervienen en el proceso en la Sección 4.2.2.

Se consideró el *espesor* de la lámina, la cual se presenta en rollos, como variable de entrada en este análisis, ya que es una característica importante de la materia prima, al cual se le asignó la nomenclatura:  $x$ , posee un intervalo de variación de: 1.17 – 1.27 mm.

Se analizaron 236 registros de la variable de salida  $y$  que es la longitud del álabe, la cual tiene un intervalo de variación especificado en: 3.525 – 3.555 pulgadas.

La Tabla 4.2 muestra parte de los registros observados en el proceso. Con dichos datos se tuvo el interés de demostrar la relación que existe entre la variable de salida  $y$  y la variable de entrada  $x$ . Dentro de los registros se identifica el momento en que ocurren cambios de rollo en la materia prima. Dicho cambio de rollo hace referencia a que cuando un rollo se acaba de manufacturar, es necesario detener la producción, tomar un rollo nuevo y montarlo en la prensa para comenzar a producir álabes de nuevo.

Tabla 4.2 Datos para las pruebas iniciales

Medición	Espesor (mm)	Longitud del álabo (pulg.)	Cambio de rollos (*)
1	1.20	3.539	
2	1.20	3.544	
3	1.20	3.543	
4	1.20	3.544	
5	1.20	3.543	
6	1.21	3.540	*
7	1.21	3.538	
⋮	⋮	⋮	⋮
225	1.19	3.536	
230	1.19	3.537	
231	1.19	3.539	*
232	1.19	3.538	
233	1.19	3.540	
234	1.19	3.540	
235	1.19	3.541	
236	1.19	3.540	

Al generar el diagrama de dispersión en la Figura 4.3, se observa que para cada valor puntual de  $x$  existe dispersión importante en los valores de  $y$ , por esta razón se analizó la forma de la distribución de  $y$ , realizando pruebas de normalidad Anderson-Darling.



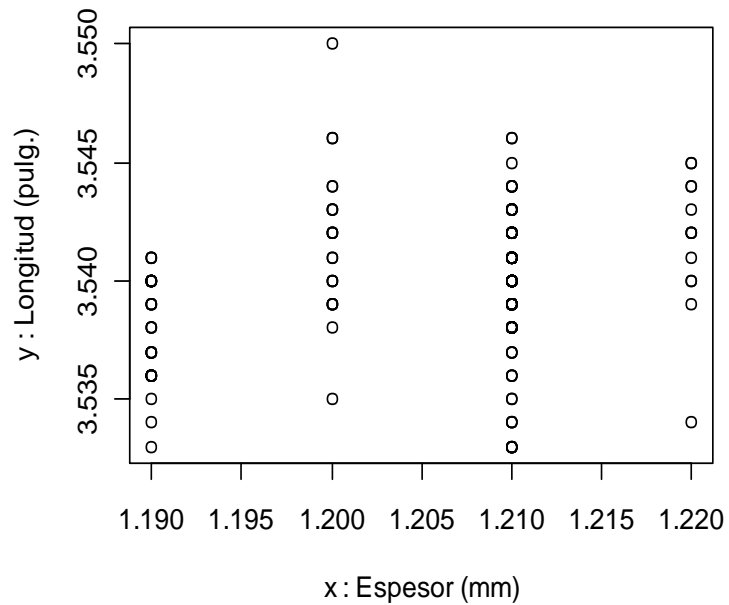


Figura 4.3 Diagrama de dispersión para  $x$  y  $y$ .

El análisis de las pruebas de normalidad para definir la forma de la distribución de  $y$  se muestra en la Tabla 4.3 donde se indica que  $y$  no presenta una distribución normal en algunos de los valores de  $x$ . Se relacionó que esta situación se debe a causas especiales de variación, ya que se puede esperar que un proceso de producción se comporte con un patrón aleatorio y normal si este se encuentra libre de causas especiales de variación que distorsionen su estado.

Tabla 4.3 Pruebas de normalidad para  $y$  en cada grupo de  $x$

Categoría en $x$ : Espesor (mm)	Cantidad de Datos	Distribución de $y$
1.19	36	No normal
1.20	23	Normal
1.21	164	No normal
1.22	13	Normal

El análisis de los 236 registros como se muestra en la Tabla 4.3 genera la siguiente observación importante: existe un factor que afecta a la normalidad

en los datos, se planteó como dicho factor: el cambio de rollos de la materia prima ya que esto ocurre frecuentemente.

Del diagrama de dispersión en la Figura 4.3 se desarrolló el análisis de regresión con los valores medios de  $y$  por cada valor puntual de  $x$ , los cuales se observan en la Tabla 4.4, la desviación estándar para cada grupo se visualiza en la tercer columna de la misma tabla. Al observar la Figura 4.4 se identificó una tendencia positiva en la recta de regresión.

Tabla 4.4 Datos para el análisis de regresión

Valor puntual para el espesor (mm)	Valor medio de la longitud en cada grupo (pulg.)	Desviación estándar de la longitud.
1.19	3.5381	0.002122
1.20	3.5417	0.003157
1.21	3.5399	0.002363
1.22	3.5416	0.002987

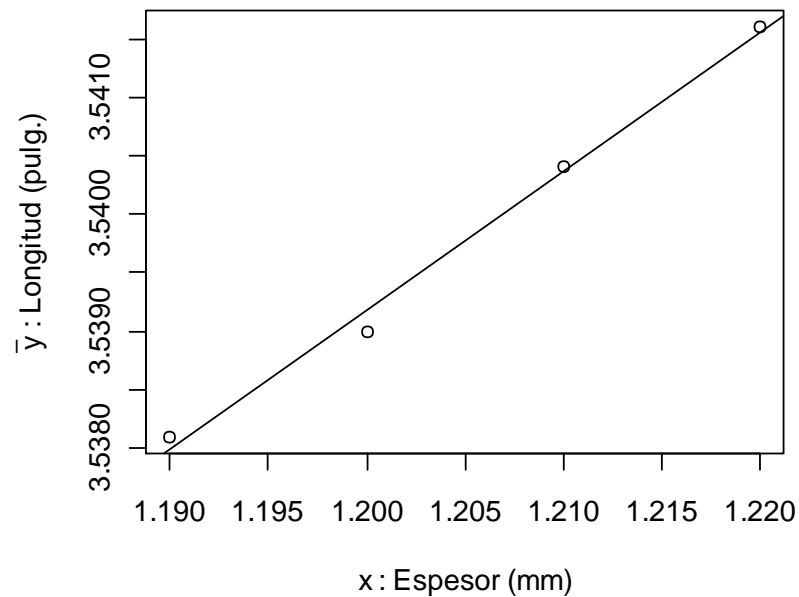


Figura 4.4 Recta de regresión

De esta forma se obtuvo el modelo de regresión mostrado en la ecuación (4.2), donde el coeficiente de determinación  $R^2$  es 0.993. Para probar la significancia de la regresión se realizó el análisis de varianza mostrado en la Tabla 4.5 donde se puede observar el valor del estadístico F y el valor p.

$$\hat{y} = 3.40833 + 0.10917x \quad (4.2)$$

Tabla 4.5 Análisis de Varianza para la regresión lineal

Fuente	gl	SC	CM	F	P
Regresión	1	5.98418E <sup>-6</sup>	5.98418E <sup>-6</sup>	70.19	0.014
Error Residual	2	1.70520E <sup>-7</sup>	8.52600E <sup>-8</sup>		
Total	3	6.15470E <sup>-6</sup>			

NOTA: Se ha realizado el análisis de varianza para la regresión teniendo pocos grados de libertad para el error, sin embargo esto se explica porque la variable  $x$  contiene solo 4 valores puntuales, lo que genera 2 grados de libertad para SCE y 3 para STC.

Una vez mostrada la importancia del análisis de regresión se generaron los intervalos de predicción de los valores medios de  $y$  en cada valor puntual de  $x$ , además se consideró el porcentaje de datos contenidos dentro de cada intervalo, estos datos pueden ser vistos en la Tabla 4.6.

Teniendo el modelo de regresión lineal y los intervalos de predicción al 95% de confianza se obtuvo la predicción media de  $y$ , lo que permitió obtener la gráfica integral de regresión lineal mostrada en la Figura 4.5.

Tabla 4.6 Intervalos de predicción para  $\hat{y}$

x	$\hat{y}$	Intervalo	Porcentaje de puntos dentro de cada intervalo de predicción
1.19	3.53825	(3.53661, 3.53988)	47%
1.20	3.53934	(3.53791, 3.54077)	78%
1.21	3.54043	(3.53900, 3.54186)	67%
1.22	3.54152	(3.53988, 3.54316)	54%

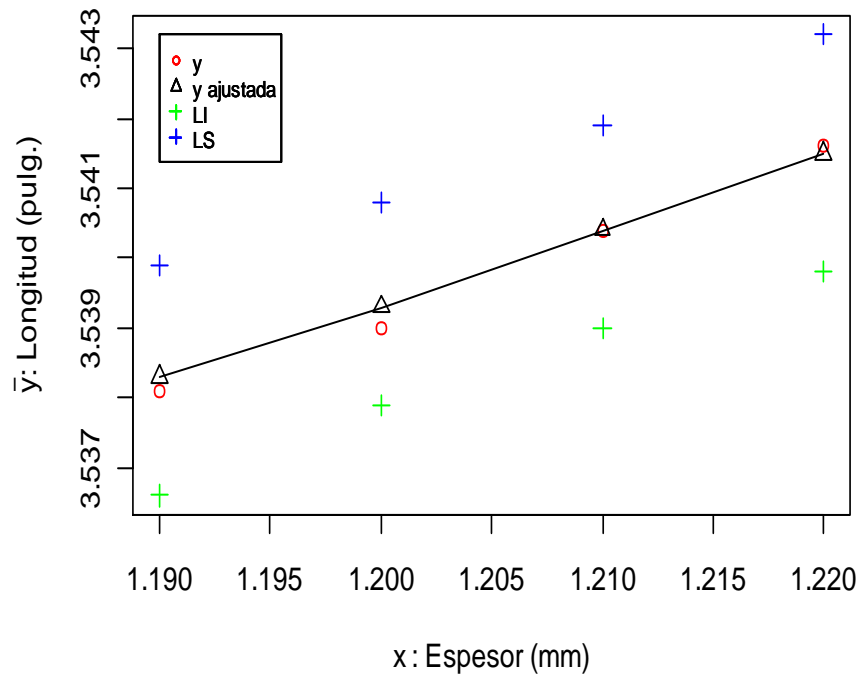


Figura 4.5 Gráfica integral de regresión.

Al realizar este análisis de diagnóstico fue posible definir que la variación en la longitud del álabe está fuertemente relacionada con la variación que se presenta en los lotes de la materia prima, por lo cual estos primeros resultados fueron tomados como fundamento para realizar un análisis que involucró una mayor cantidad de variables de entrada del producto. El desarrollo y los resultados se explican en las etapas siguientes del método.

Utilizar esta información como punto de partida para analizar el caso de estudio correspondiente, permitió identificar que existe una relación fuerte en la variación de la longitud del álabe ocasionada por la variación de la materia prima, esta variación ocasiona un promedio diario de productos defectuosos mayor al deseado.

### 4.2.2 Etapa 3: Selección de variables involucradas.

Dentro de esta investigación se consideró involucrar los siguientes elementos del proceso: Producto manufacturado, maquinaria, materia prima y herramental, debido a que estos son los principales dentro del mismo. Las características de estos elementos y sus especificaciones de trabajo son determinadas por la empresa y se explican a continuación.

En relación al *producto* las características importantes son las siguientes:

- Longitud del álabe, la cual es la medida que presenta el álabe en su abertura final considerada de punta a punta en el mismo.
- Modelo Matemático, que son las mediciones que se realizan en 24 puntos del interior del álabe para verificar su curvatura.
- Ángulos de pestañas: mediciones del ángulo en las 5 pestañas que el álabe posee las cuales permiten el ensamblado del álabe con la carcasa y el anillo.

De las tres características en el producto, se seleccionó para su estudio la primera: *Longitud del álabe*, ya que es la variable de calidad que se monitorea con mayor frecuencia en el producto.

En cuanto al modelo matemático y los ángulos de pestaña su inspección es más costosa para la empresa y su frecuencia de monitoreo es menor a la primera, por tales razones no fue posible su estudio en esta investigación.

En la Figura 4.6 se muestra la medición que se realiza para la variable de Longitud y el valor de la medición, así como su tolerancia la cual le permite variar dentro de un rango de [3.525,3.555] pulgadas.

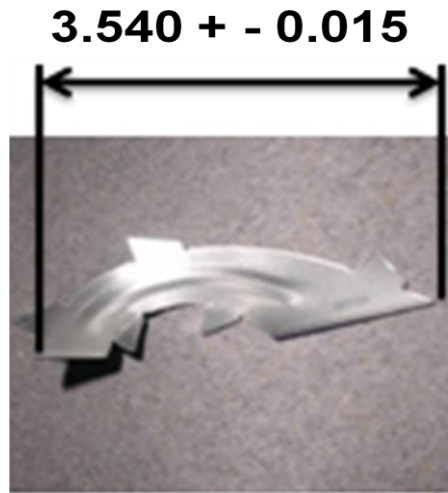


Figura 4.6 Longitud del álabo.

Correspondiente a la *maquinaria*, la cual es una prensa con capacidad de 250 kN, se tienen los parámetros mostrados en la Tabla 4.7, junto con el rango del valor que puede tomar cada parámetro para que el álabo sea fabricado dentro del rango de la variable  $y$ .

Tabla 4.7 Rango de funcionamiento de los parámetros de la maquinaria.

Característica	Rango	Unidad
Altura del RAM:	314 +/- 0.2	mm
Velocidad:	100-180	gpm
Avance de lámina:	41 +/- 0.5	mm
Centrado de rodillos:	122.5 +/- 4	mm
Lubricación de lámina		
Frecuencia	4 a 15	s
Pulsos	Max 1	

*Nota: gpm representa golpes por minuto.*

Se observó en el proceso que dichos parámetros de la maquinaria se mantienen fijos durante el funcionamiento del proceso, esto significa que no

presentan variación, por lo cual se descartó el elemento de maquinaria en la investigación.

La *materia prima* se presenta en rollos, los cuales son montados en el alimentador de la prensa según se muestra en la Figura 4.7. Para dicha materia prima se identificó el tipo de acero que se utiliza en la fabricación del álabe, el cual es un acero SAE 1008, laminado en frío, acabado mate, libre de suciedad, óxido y defectos en la superficie.



Figura 4.7 Materia prima montada en el alimentador de la prensa.

En la Tabla 4.8 se visualizan las características físicas y propiedades mecánicas de la materia prima, en la segunda columna se identifica el rango o valor en el cual se ha determinado se encuentre cada propiedad y en la última columna se señalan las unidades en las que se mide cada una de ellas.

Tabla 4.8 Dimensiones y propiedades mecánicas de la materia prima.

Característica	Valor numérico / Intervalo de variación	Unidades de medición
Ancho	120 +0.1	mm
Espesor	1.17 - 1.27	mm
Esfuerzo de Fluencia	172-317	MPa
	25-46	ksi
Esfuerzo de Tensión	270 mín.	MPa
	39 mín.	ksi
Elongación en 2"	30 mín.	%
Dureza	60-75	HRB

Se observa en la Tabla 4.9 lo relacionado a las propiedades químicas del acero, se utiliza el porcentaje máximo en peso que cada elemento debe contener. La especificación interna marca que la micro-estructura debe ser ferrita equiaxial, con tamaño de grano 7-10 ASTM.

En relación al *herramental* se identificó que se trabaja con un troquel progresivo; en el cual al momento de ser montado en la prensa, se verifican el cierre y la apertura que éste presenta en los postes frontales, izquierdo y derecho, antes de iniciar la producción de álabes. De esta acción no se tiene un valor establecido como parámetro que permita la comparación con lo que registra el operario. Sin embargo es importante identificar y relacionar la apertura del troquel con el aplastamiento que sufre una muestra de alambre de estaño, al ser colocado entre las cubiertas de los postes del *herramental*. Si dicha muestra de estaño no sufre un aplastamiento adecuado indica que el troquel se encuentra abierto, ocasionando que los formadores del mismo presenten un contacto deficiente con la lámina de la materia prima y a su vez el



álabe no sea formado hasta la curvatura deseada. El espesor final de las dos muestras de estaño se mide por medio de vernier y se registra junto con los parámetros de proceso antes mencionados, la medición es en milímetros y es evaluada a criterio del Ingeniero de Procesos de la planta. Esta característica es considerada dentro de la investigación para ver el efecto que presenta la apertura del troquel sobre la variable de salida en el producto

Tabla 4.9 Propiedades químicas de la materia prima.

Elemento químico	% en peso
C	0.1 máx.
S	0.05 máx.
Mn	0.5 máx.
P	0.03 máx.
Si	0.1 máx.
Al	0.02-0.06
Ni	0.007 máx.
V	-
Ti	-
Nb	-
B	0.0005 máx.
Otros elementos	Elemento solo: máx. 0.1 Elementos juntos: máx. 0.3

Desglosar las características de cada elemento en el proceso, permitió seleccionar con las cuales se trabajaría, quedando como se muestra en la Tabla 4.10 a continuación:

Tabla 4.10 Variables seleccionadas para la investigación.

Variables de entrada			Variable de salida
Maquinaria	Materia Prima	Herramental	Producto
-	Esfuerzo de fluencia	Apertura de troquel	Longitud del álabe
-	Esfuerzo de cedencia	-	-
-	Espesor	-	-
-	Elongación	-	-
-	Dureza	-	-
-	Químico: carbón	-	-

El siguiente paso dentro de la metodología está enfocado hacia la estructura de una base de datos.

#### 4.2.3 Etapa 4: Recolección y estructuración de la base de datos.

La recolección de datos para el caso de estudio se llevó a cabo dentro del área de Prensas de la empresa ZF Sachs, en la prensa E-P27.

Se tomaron mediciones relacionadas de la materia prima y la apertura del troquel, que corresponden a las variables de entrada. Como se mencionó en la sección anterior se consideró como variable de salida a la *Longitud del álabe*, de la cual se obtuvieron mediciones adicionales a las registradas en la etapa de diagnóstico.

Se recolectaron 3 muestras de tamaño  $n = 100$  mediciones para la variable de salida. Se utiliza este tamaño de muestra porque una cantidad menor de datos no revela lo que ocurre en el proceso, y además personal de la

empresa considera que una cantidad por encima de 100 implica un mayor costo y tiempo en la recolección de la muestra. Se registraron los niveles en las variables de entrada con las cuales se produjeron las 300 observaciones de la variable de salida. El procedimiento en que se recolectaron fue el siguiente:

En primer lugar se observaron 3 lotes de materia prima con características diferentes, los cuales se identifican por el número de colada que presenta el proveedor: 14559, 26016 y 26015. En la Tabla 4.11 se visualizan los valores que reporta el proveedor para la materia prima en cada lote. Se indican tanto las propiedades químicas como las propiedades mecánicas exceptuando el valor de la dureza.

Tabla 4.11 Valores de las características que reporta el proveedor.

Característica	Colada		
	14559	26016	26015
Esfuerzo de Fluencia (MPa)	239	226	251
Esfuerzo de Tensión (MPa)	327	329	329
Espesor (mm)	1.17	1.17	1.17
Elongación (%)	38	34	36
Propiedades Químicas (% en peso)			
C	0.060	0.050	0.050
S	0.002	0.001	0.002
Mn	0.190	0.160	0.180
P	0.008	0.008	0.006
Si	0.020	0.040	0.020
Al	0.041	0.040	0.035
Ni	0.010	0.010	0.010
B	0.000	0.002	-
Cu	0.020	0.020	0.020
Cr	0.010	0.010	0.010
Ti	0.001	0.001	0.001

En la Tabla 4.12 se presentan los registros correspondientes a la apertura del troquel con la cual se ejecuto la producción de alabes según el lote de materia prima correspondiente.

Tabla 4.12 Mediciones tomadas referentes a la apertura del troquel.

Apertura del troquel en poste:	LOTE		
	14559	26016	26015
izquierdo (mm)	0.48	0.44	0.50
derecho (mm)	0.38	0.40	0.43
Promedio	0.43	0.42	0.465 ≈0.47

Finalmente se tomaron las mediciones en la longitud de 100 álabes producidos con cada uno de los 3 lotes de materia prima. Es posible observar en la Tabla 4.13 una parte de los registros que se tomaron directamente del proceso.

En los procesos de estampado y troquelado se presenta la situación de que por cada configuración en los valores de las entradas, es posible tomar una gran cantidad de medidas para la variable de salida y que incluso algunas de ellas se repiten. Para el caso de estudio los datos que se tomaron se encuentran dentro de este panorama, se realizó de esta forma con el fin de no perder información del comportamiento real del proceso.

Tabla 4.13 Mediciones registradas de la longitud del álabe de cada colada.

Medición (n)	Longitud del álabe (pulg.)		
	14559	26016	26015
1	3.522	3.519	3.535
2	3.530	3.523	3.534
3	3.532	3.521	3.534
4	3.532	3.524	3.535
5	3.528	3.521	3.534
6	3.529	3.521	3.534
7	3.532	3.522	3.527
8	3.535	3.523	3.533
9	3.529	3.519	3.538
10	3.520	3.524	3.541
⋮	⋮	⋮	⋮
91	3.527	3.523	3.526
92	3.529	3.521	3.531
93	3.526	3.522	3.531
94	3.528	3.522	3.530
95	3.529	3.520	3.530
96	3.528	3.521	3.526
97	3.530	3.523	3.531
98	3.527	3.521	3.531
99	3.527	3.523	3.530
100	3.529	3.524	3.530

Se estructuró la base de datos en un solo archivo de Excel como se muestra en la Tabla 4.14, con el fin de utilizarla adecuadamente. Esto permite que la base de datos sea exportada al paquete R Cran para el análisis correspondiente.

Tabla 4.14 Estructura general de la base de datos.

# colada	Variables de entrada						Variable de Salida
	Fluencia	Cedencia	Espesor	Elongación	Carbón	Apertura	Longitud
14559	239	327	1.17	38	0.06	0.43	n= 100
26015	251	329	1.17	36	0.05	0.47	n= 100
26016	226	329	1.17	34	0.05	0.42	n= 100

#### 4.2.4 Etapa 5: Verificación de la información obtenida / Reestructurar.

Para integrar una base de datos con información fiable dentro de la investigación, y debido a que la empresa desea corroborar la información que el proveedor de materia prima reporta, se decide llevar a cabo las siguientes pruebas mecánicas con las cuales es posible es caracterizar la lámina que se utiliza para la fabricación del álabe.

Las pruebas realizadas fueron:

- Análisis químico.
- Dureza.
- Análisis metalográfico.
- Ensayo de tensión.

A continuación se explica en qué consisten dichas pruebas y se detallan los resultados obtenidos.

Todas las pruebas se llevaron a cabo en los laboratorios que pertenecen a la Corporación Mexicana de Investigación en Materiales S. A. de C. V. (COMIMSA) y bajo los procedimientos y normas que están establecidos en

cada una de ellas. Para las pruebas de Dureza y Análisis metalográfico se tuvo la oportunidad de estar presente en la realización de las mismas, lo que amplió el conocimiento y la perspectiva de las pruebas para beneficio de la investigación.

El *análisis químico* se realizó a partir de una muestra metálica de 1 *pulg*<sup>2</sup> provenientes de cada uno de los 3 lotes de materia prima. En la Tabla 4.15 se visualizan los resultados, se muestra el porcentaje en peso del elemento químico carbón (C) en las muestras, así como la especificación y el reporte del proveedor con el fin de comparar los datos.

Únicamente se seleccionó el porcentaje en peso del carbón por ser el más importante, ya que ayuda a verificar el tipo de acero con el cual se manufactura el componente así como también afecta a las propiedades mecánicas de la lámina. En este caso se conoce que es un acero 1008 y su porcentaje en peso promedio de carbón es 0.08, si un acero contiene mayor porcentaje de carbón entonces dicho acero adquiere mayor dureza pero su ductilidad se ve afectada y éste se torna frágil.

Tabla 4.15 Resultados y comparación del análisis químico.

# colada	% C		
	Especificación	Proveedor	Resultados Análisis
14559	0.10 max	0.06	0.10
26016	0.10 max	0.05	0.06
26015	0.10 max	0.05	0.06

Se observó del resultado del análisis: que la colada 14559 tiene mayor porcentaje de carbón y presenta una diferencia significativa contra lo que reporta el proveedor. Para las pruebas siguientes se esperaba que esta colada presentara mayor dureza.

Enseguida se realizó el *ensayo de dureza Rockwell*, para el cual se tomaron 6 mediciones en la secuencia que se muestra en la Figura 4.8, donde 1 es el centro del área en donde se realizaron las mediciones, de 2 a 5 se tomaron las durezas en las esquinas de la muestra y 6 es una medición en el área de la sección transversal.

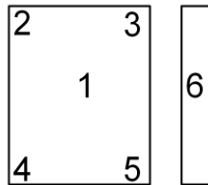


Figura 4.8 Croquis de las mediciones de dureza.

Para determinar la dureza que presenta cada muestra se tomaron los promedios de las 6 mediciones. La especificación marca un rango de dureza entre 60 a 75 en escala Rockwell B (HRB). Los resultados se muestran en la Tabla 4.16, donde se calculó el promedio de las mediciones para considerar este valor en la base de datos.

Tabla 4.16 Resultados del ensayo de dureza

# de medición	Dureza (HRB)		
	14559	26016	26015
1	55	58	57
2	55	59	61
3	53	57	62
4	55	60	57
5	55	60	56
6	56	60	59
Promedio	55	59	59



En las muestras los resultados revelan que la materia prima no cumple con el mínimo especificado en su dureza, además la colada 14559 que se esperaba tuviera mayor dureza por su contenido de carbón relativamente alto, resulta ser en donde se registra menor dureza, por esta razón se planteó el siguiente cuestionamiento: ¿Por qué si la muestra 14559 presenta mayor porcentaje de carbón, tiene menor dureza en comparación con las otras dos muestras? Se logró obtener una respuesta a la pregunta antes planteada analizando la micro-estructura de las muestras a través de un *análisis metalográfico*, de esta forma se identificó que la micro-estructura principal en los aceros es una matriz de ferrita con cementita como se muestra en las metalografías de la Figura 4.9

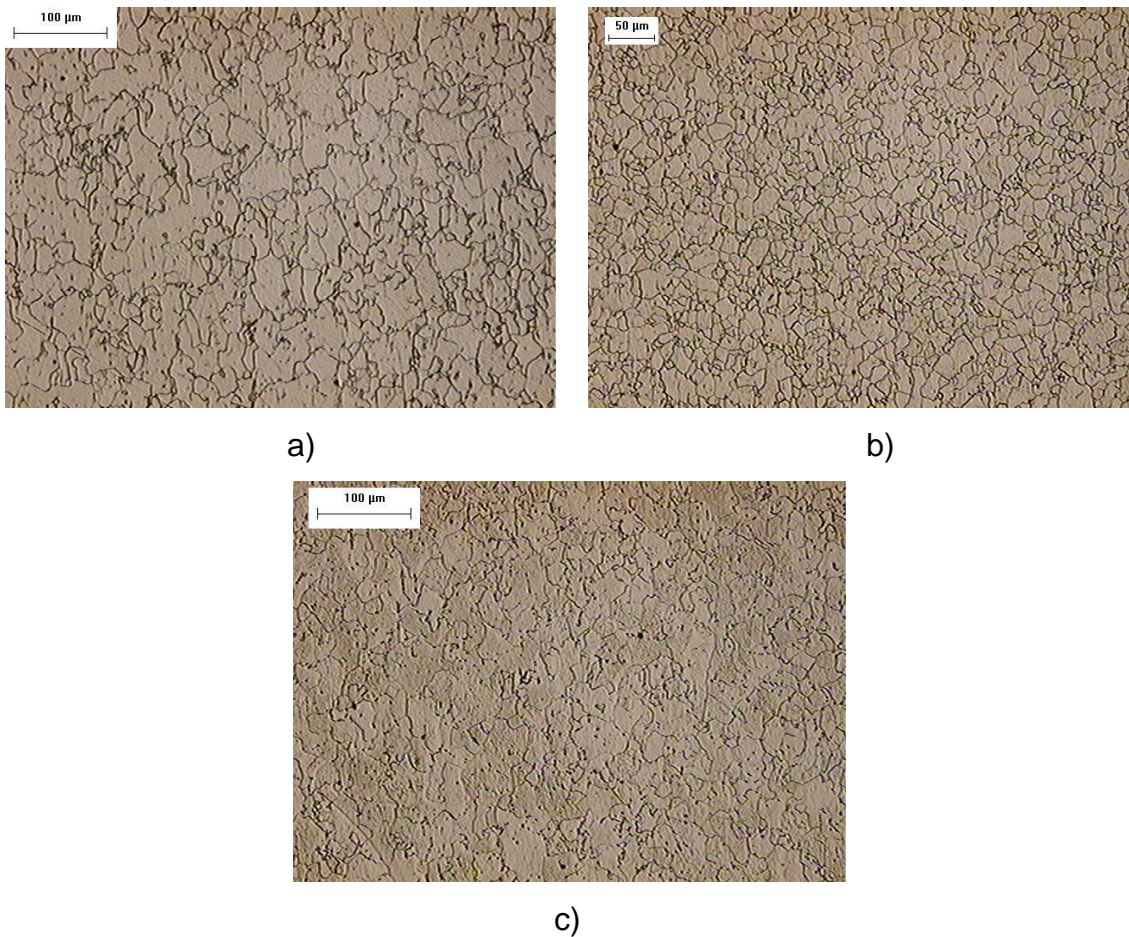


Figura 4.9 Metalografías de los aceros

La Figura 4.9a corresponde a la muestra de la colada 14559 en la cual se determinó que presenta un tamaño de grano 7 ASTM, la Figura 4.9b la cual corresponde a la muestra de la colada 26016 presenta un tamaño de grano 9 ASTM y la Figura 4.9c es la metalografía resultante de la muestra y colada 26015 con tamaño de grano 7.5 ASTM.

Al determinar el tamaño de grano de las muestras se logró ver que existe una diferencia significativa de la colada 26016 respecto a las otras dos coladas, ya que esta posee un tamaño de grano 9 lo que significa que es más fino, cuando el grano es más pequeño se requiere mayor carga para su deformación plástica. De esta forma se explica también que la colada 14559, la cual tiene un porcentaje mayor de carbón, presenta menor dureza, ya que el grano es grande y requiere menor carga para ser deformado.

Los resultados de la *prueba de tensión* muestran la carga de fluencia y la carga máxima a la cual fueron sometidas las probetas llegando a la ruptura. En la Tabla 4.17 se puede ver esta información la cual es indispensable para determinar el Esfuerzo de Fluencia y el Esfuerzo Máximo que soportan el acero así como el porcentaje de alargamiento de cada probeta.

El Esfuerzo de fluencia es importante porque marca el “punto” al cual el metal pasa de la zona de elasticidad y entra en la zona de deformación plástica. Es importante señalar la fluencia de la colada 26016 porque es el valor mayor de los tres, esto significa que se requiere mayor carga para que el material sea deformado permanentemente.

Tabla 4.17 Resultados del ensayo de tensión.

# colada	L. Inicial (mm)	Área mm <sup>2</sup>	Carga Fluencia Kgf	Carga Máxima Kgf	Esfuerzo Fluencia	Esfuerzo Máximo	Alargamiento %
	L. Final (mm)				MPa	MPa	
					psi	psi	
14559	50.8	15.61	270	528	170	332	44
	73.15				24598	48104	
26016	50.8	15.59	318	529	200	333	40
	71.31				29018	48271	
26015	50.8	15.29	276	493	177	316	47
	74.83				25680	45870	

Se graficaron los resultados de los valores del Esfuerzo de Fluencia, los cuales se observan en la Figura 4.10, donde además se comparó con la información que entrega el proveedor observando mucha variación entre lo que reporta y los resultados del ensayo.

Existe una diferencia significativa en las coladas ya que las tres se encuentran posicionadas cerca del valor mínimo aceptable que marca la especificación.

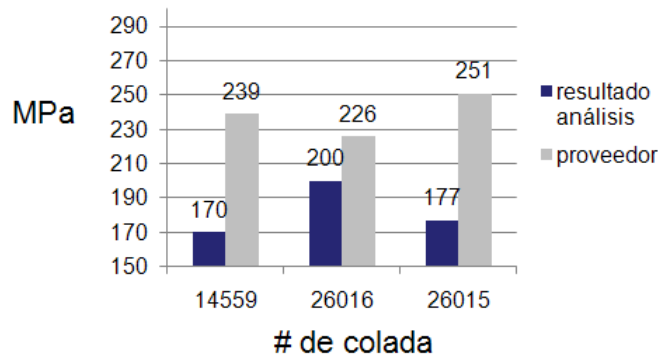


Figura 4.10 Comparación del esfuerzo de fluencia: proveedor vs análisis.

Es posible ver en la Figura 4.11 que existe una diferencia significativa de 13 MPa entre el valor que reporta el proveedor y el resultado del ensayo, para la colada 26015 respecto al Esfuerzo Máximo. Aunque las coladas soportan un Esfuerzo Máximo similar, la transición de elasticidad a plasticidad ocurre a diferentes cargas, ocasionando diferentes valores para el esfuerzo de fluencia.

Después de obtener los resultados de las pruebas mecánicas, se reestructura la base de datos, con información que resulta ser confiable y coherente. De esta manera se puede continuar con la etapa 6 del método, la cual se desarrolla en el Capítulo 5 de la tesis.

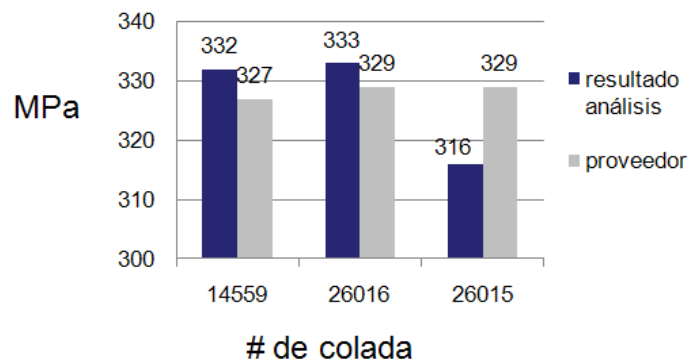


Figura 4.11 Comparación del esfuerzo máximo: proveedor vs análisis

## CAPÍTULO 5 EXPERIMENTACIÓN Y RESULTADOS

### 5.1 Etapa 6: Análisis de la base de datos.

Como se mostró en la Tabla 4.14 la base de datos inicial contiene la información entregada por el proveedor, dado que la empresa requirió corroborar dicha información, la base de datos fue reestructurada como se muestra en Tabla 5.1 donde se incluyen los resultados de las pruebas mecánicas realizadas en el capítulo anterior.

Tabla 5.1 Estructura general de la base de datos / Pruebas mecánicas.

	Variables de entrada								Variable de Salida
colada	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$y$
14559	170	332	1.22	44	0.10	7	55	0.43	n= 100
26015	177	316	1.23	47	0.06	7.5	59	0.47	n= 100
26016	200	333	1.24	40	0.06	9	59	0.42	n= 100

Donde  $x_1$  representa la fluencia,  $x_2$  la cedencia,  $x_3$  el espesor,  $x_4$  la elongación,  $x_5$  el porcentaje de carbón,  $x_6$  el tamaño de grano,  $x_7$  la dureza y  $x_8$  la apertura. La longitud del álabe se denota como  $y$ . Nótese que se incluyen las variables  $x_6$  y  $x_7$ , además las observaciones de la variable  $x_8$  no cambian con respecto a la base de datos inicial.

### 5.1.1 Etapa 6.1: Análisis de la variable de salida

Se utilizó el análisis de varianza para determinar si existe diferencia en las medias poblacionales de la longitud del álabe de cada una de las tres coladas. Se probó la hipótesis nula contra la hipótesis alterna:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

*H<sub>a</sub>: por lo menos una de las medias es distinta de las otras*

Se tiene interés en investigar si el tipo de colada afecta de manera significativa a la longitud del álabe.

A través de un ANOVA de un solo factor es posible dividir la variación total de las medidas de la longitud del álabe, en porciones atribuibles a los 3 tipos de colada que se estudiaron. De esta manera se puede medir el efecto que tiene el tipo de colada sobre la longitud del álabe. Se recolectó una muestra de 100 mediciones de la longitud para cada colada, donde las coladas se identificaron como 14559, 26015 y 26015.

En la Tabla 5.2 se muestran los resultados del ANOVA, donde el estadístico de prueba  $CMT_r/CME$  tiene una distribución F con  $gl_1 = 2$  y  $gl_2 = 297$ . Con un valor  $\alpha = 0.05$ , puede rechazarse  $H_0$  cuando  $F > F_{0.05} = 3.00$ . Ya que el valor observado,  $F = 348.77$ , sobrepasa el valor crítico, es posible ver que existe suficiente evidencia para indicar que por lo menos una de las tres coladas es diferente de las otras, es decir se rechaza  $H_0$ .

Fue posible obtener la misma conclusión utilizando el valor  $p = 0.0$ , y dado que dicho valor es menor que  $\alpha = 0.05$ , los resultados son significativos con un 95% de confianza, lo que significa que por lo menos la media de una de las tres coladas es diferente de las otras.

Tabla 5.2 Análisis de varianza de un solo factor.

Fuente	gl	SC	CM	F	Valor p
Factor (Colada)	2	0.0070510	0.0035255	348.77	0.000
Residuales	297	0.0030022	0.0000101		
Total	299	0.0100532			

Se pueden obtener conclusiones similares utilizando diagramas de cajas, en donde se grafican las observaciones de  $y$  para cada factor conocido como colada. En la Figura 5.1 se muestran las tres gráficas de caja correspondientes a cada colada. También se observan los puntos atípicos de cada colada así como la forma de la distribución de los datos en base a la línea central o mediana.

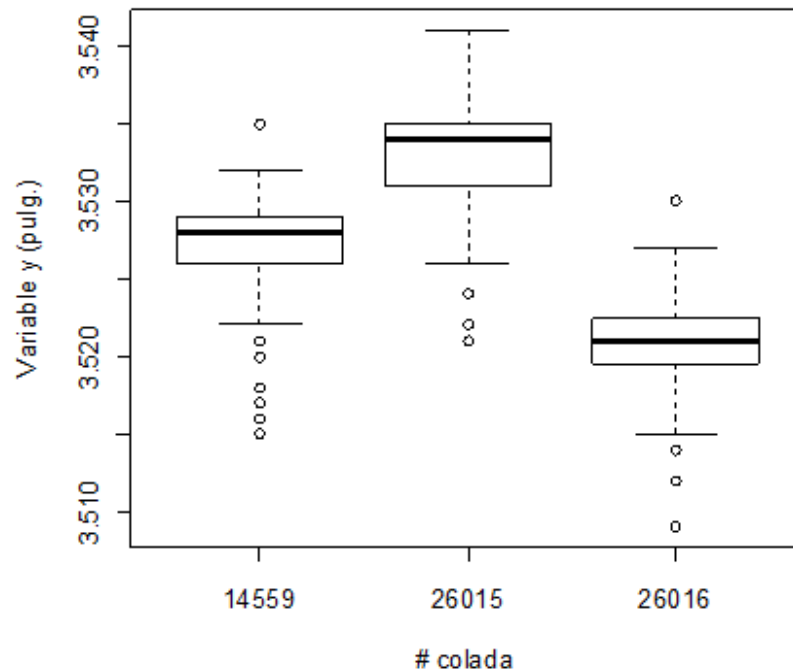


Figura 5.1 Diagrama de caja para los tres grupos de datos.

El diagrama correspondiente a la colada 26016 muestra que la mediana está centrada en la caja, lo que significa que la distribución de los datos es

simétrica. Por el contrario los diagramas para las otras dos coladas muestran que la línea mediana está arriba del centro de la caja, lo que indica que la forma de la distribución se encuentra sesgada hacia la derecha.

En conjunto los diagramas presentan desfase entre sí, coincidiendo con la hipótesis alterna que se probó en el ANOVA, es decir, las medias de las muestras son diferentes.

La distribución de frecuencias para cada grupo de datos se puede observar en la Figura 5.2 donde se muestran los histogramas correspondientes a los tres tipos de colada.

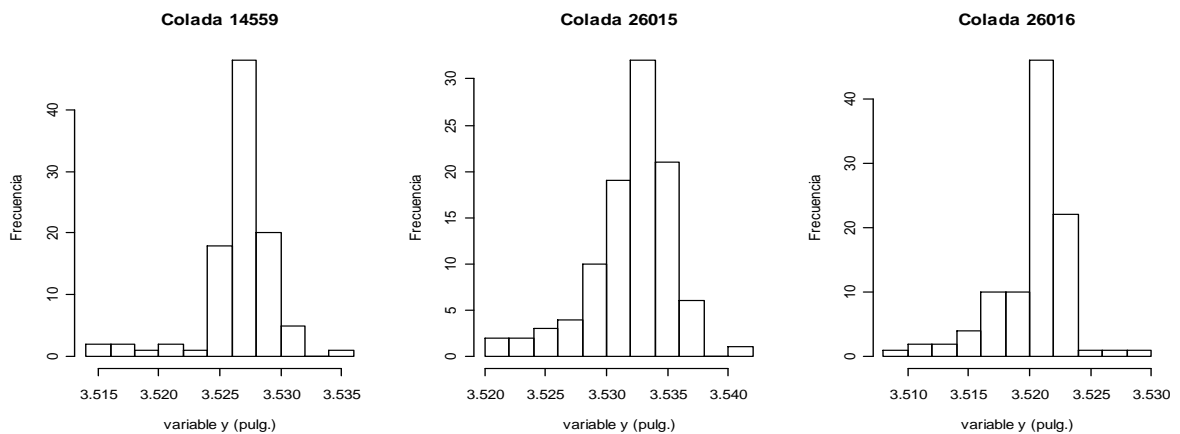


Figura 5.2 Histogramas para la variable  $y$  de cada colada.

### 5.1.2 Etapa 6.2: Análisis de correlación.

Dado que en la Etapa 6.1 se considera solamente analizar la variable de salida  $y$  para esta etapa es necesario considerar también las 8 variables de entrada disponibles en la base de datos.



Se dispone de  $x_i$  variables de entrada, donde  $i = 1, 2, \dots, 8$  y de la variable de salida  $y$ . Se obtienen los coeficientes de correlación de Pearson entre cada  $x_i$  con  $y$ , los cuales se muestran en la Tabla 5.3, tales coeficientes muestran la fuerza de la relación lineal entre las variables respectivamente.

Tabla 5.3 Coeficiente de correlación  $r$  entre las variables.

Variables	Coeficiente de correlación
$x_1 - y$	-0.6416
$x_2 - y$	-0.7260
$x_3 - y$	-0.4548
$x_4 - y$	<b>0.8370</b>
$x_5 - y$	0.0423
$x_6 - y$	-0.6320
$x_7 - y$	-0.0423
$x_8 - y$	0.7764

En la tabla se observa que hay valores cercanos a 1 para algunas combinaciones de variables  $x_i$  y  $y$ , los cuales indican que hay una relación lineal positiva entre ambas. También se puede notar que hay valores cercanos a -1, lo que indica relación lineal negativa entre las  $x_i$  y  $y$  correspondientes. Los valores muy cercanos a cero de las variables  $x_5$  y  $x_7$  indican que no hay una relación lineal entre  $x_5$  y  $y$ , y entre  $x_7$  y  $y$ .

El valor de  $r$  entre  $x_4$  y  $y$  de 0.8370, es bastante cercano a 1 y además es el valor más grande que indica una relación lineal positiva bastante fuerte entre la elongación y la longitud del álabe, por lo tanto la elongación es la variable más importante de la base de datos para la longitud del álabe.

## 5.2 ETAPA 7: Aplicación del Modelo de Regresión Lineal.

Considerando el análisis de correlación en la sección anterior, se realiza un análisis de regresión lineal para obtener un modelo que permita describir a la variable  $y$  utilizando la variable  $x_4$ . En la Figura 5.3 se muestra la recta de regresión para dichas variables, la cual tiene un coeficiente de determinación  $R^2 = 0.7007$  y ecuación (5.1)

$$\hat{y} = 3.4531 + 0.00169x_4 \quad (5.1)$$

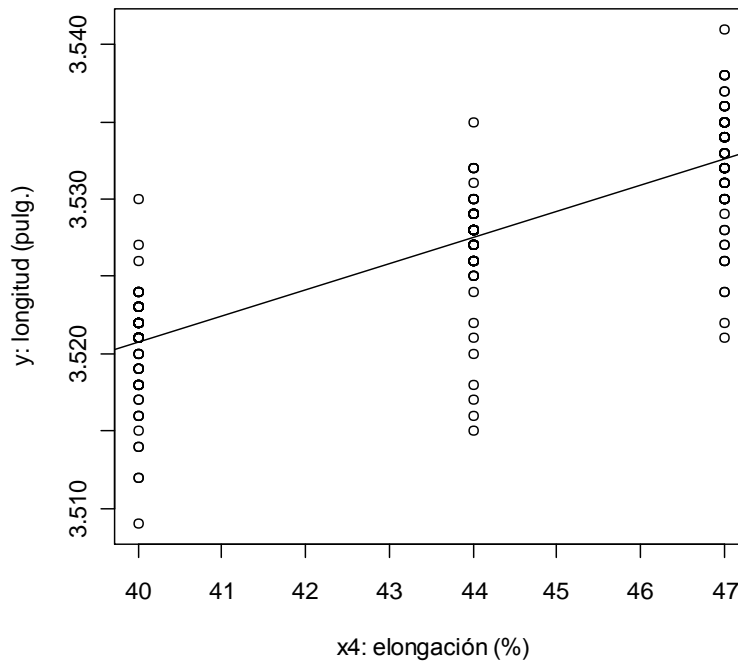


Figura 5.3 Recta de regresión entre  $y$  y  $x_4$

Considerando el ajuste de este modelo se determina que  $x_4$  es útil para predecir el valor de  $y$ . Al verificar el supuesto de normalidad para los residuales del modelo de regresión en (5.1) se prueban las siguientes hipótesis con la prueba Anderson-Darling:

$H_0$ : los residuales tienen una distribución normal

$H_a$ : los residuales tienen otra distribución

Para esta prueba el valor resultante de  $p= 0.000$  es menor que el nivel de significancia  $\alpha = 0.05$ , por lo cual se rechaza  $H_0$  y se acepta  $H_a$ , es decir los residuales no se distribuyen normalmente. Se obtiene la misma conclusión mediante la gráfica de la Figura 5.4.

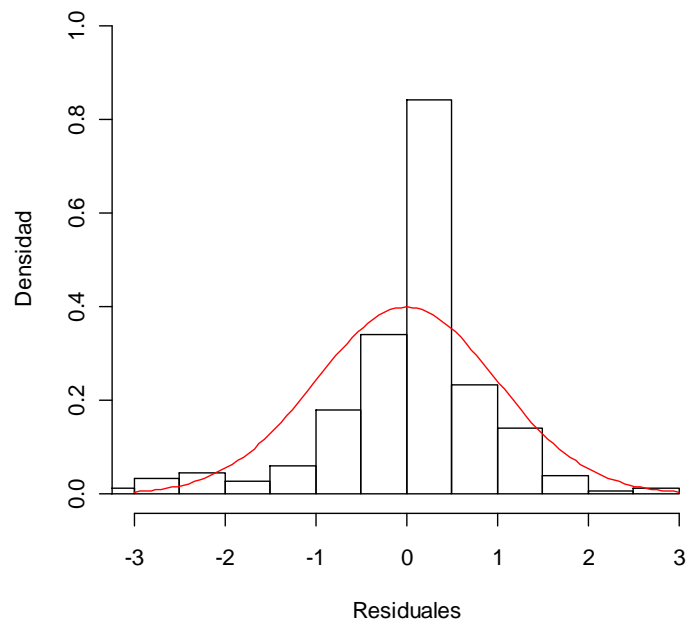


Figura 5.4 Distribución no normal de los residuales

Posteriormente se verifica el supuesto de independencia entre los residuales, por medio de la Figura 5.5 se visualiza que una observación por encima de la media tiende a ser seguida por otra observación encima de la media y viceversa. Esto indica que los residuales no son independientes entre sí, ya que una observación presenta dependencia de las anteriores.

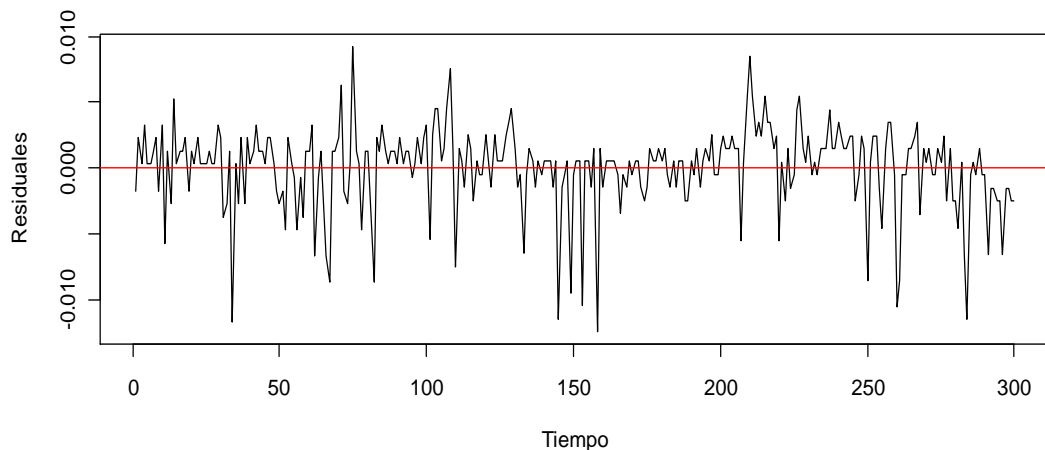


Figura 5.5 Residuales no independientes.

Como se menciona en la Sección 3.2.4, los resultados del análisis de regresión sólo son válidos cuando los datos satisfacen los supuestos de regresión necesarios, que para el caso del modelo (5.1) dichos supuestos no se cumplen. Por lo cual se requiere recolectar nuevas observaciones, sin embargo, la forma en que se obtienen las observaciones de la variable  $x_4$  es por medio de pruebas mecánicas realizadas por un laboratorio especializado, esto representa una opción no factible ya que genera un costo económico alto. Por tal razón se recurre a producir nuevas observaciones mediante el paquete estadístico R cran, de manera que los datos generados poseen la misma media que los datos originales, sólo se ajusta su desviación estándar para que presenten mayor dispersión.

Como se describe en el resto de esta sección, las observaciones generadas permiten obtener un modelo de regresión que cumple satisfactoriamente con los supuestos del análisis de residuales.

En la parte izquierda de la Figura 5.6, se observa la gráfica generada con los datos originales de la variable  $x_4$ , donde se aprecia que no hay dispersión en cada uno de los 3 subgrupos. Mientras que en el lado derecho se observan

los datos generados denotados como  $x'_4$ , los cuales presentan un incremento de 1.95 en la desviación estándar de cada subgrupo.

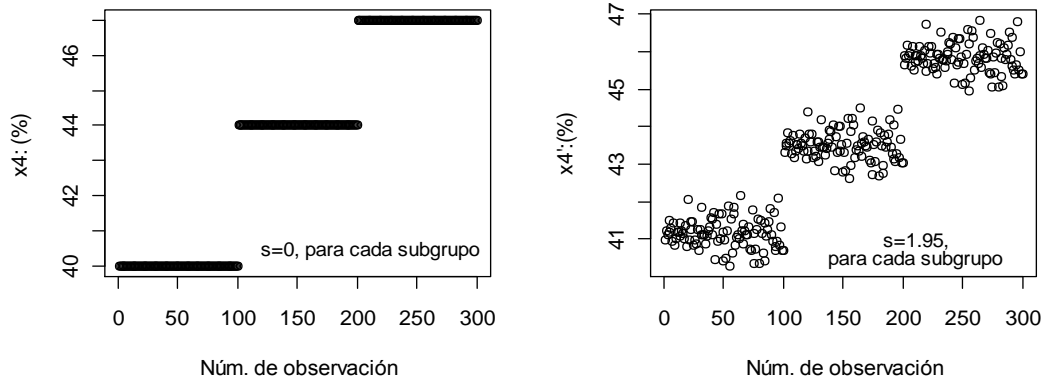


Figura 5.6 Observaciones para la elongación.

En la Figura 5.7 en el lado izquierdo se grafican las observaciones extraídas del proceso de la variable  $y$ , mientras que en la parte derecha se grafican las nuevas observaciones, las cuales se denotan como  $y'$ . Se aclara que los datos para  $y'$ , son generados a partir de una distribución con las mismas propiedades de la distribución original, es decir,  $\bar{x} = 3.526$  y  $s = 0.0057$ .

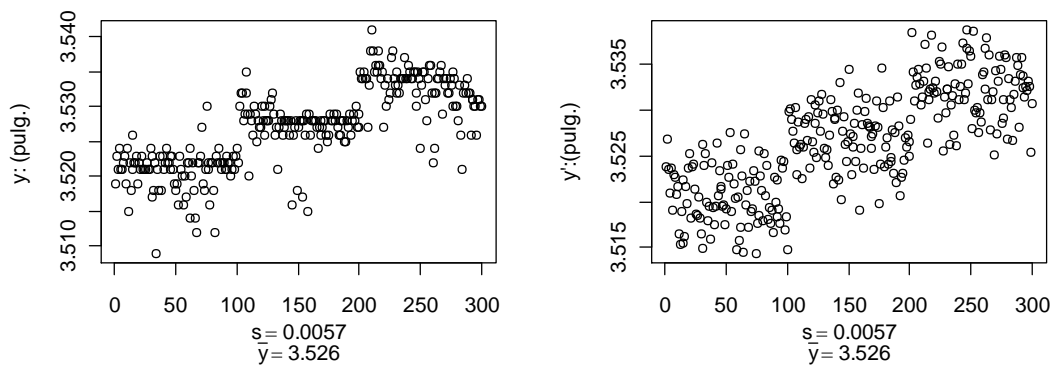


Figura 5.7 Observaciones de la longitud.

Finalmente, se dispone de nuevos datos para las variables  $x_4$  y  $y$ , los cuales se denotaron como  $x'_4$  y  $y'$ . Es importante aclarar que para  $x'_4$  y  $y'$ , el valor de  $r = 0.8096$  es bastante alto al igual que el mostrado en la Tabla 5.3 para  $x_4$  y  $y$ . En la Figura 5.8 se muestra el diagrama de dispersión correspondiente a  $x'_4$  y  $y'$ , donde se observa una tendencia en línea recta con pendiente positiva.

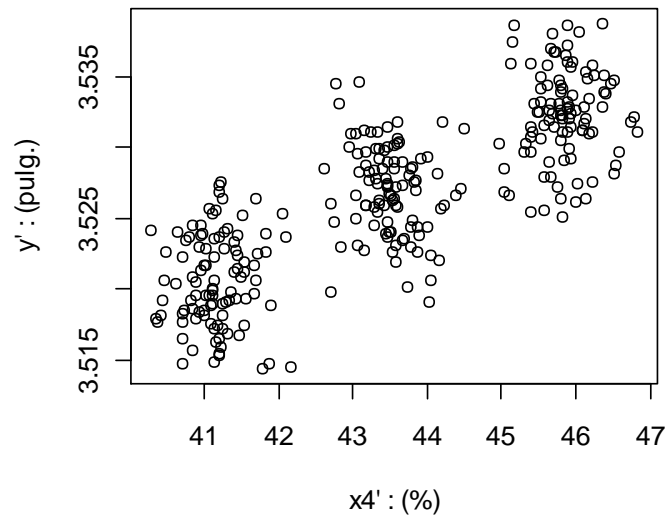


Figura 5.8 Diagrama de dispersión para las nuevas observaciones.

Partiendo del diagrama de dispersión mostrado anteriormente, se realiza un análisis de regresión lineal para obtener un modelo que permita describir a la variable  $y'$  utilizando la variable  $x'_4$ . En la Figura 5.9 se muestra la recta de regresión para dichas variables, la cual tiene un coeficiente de determinación  $R^2 = 0.6556$  y ecuación  $\hat{y}' = 3.45313 + 0.00169x'_4$ .

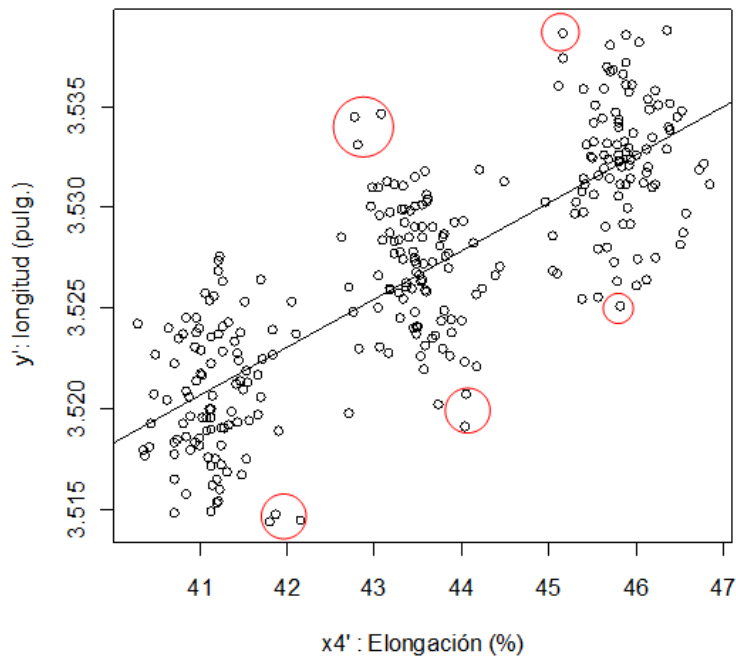


Figura 5.9 Recta de regresión para  $x'_4$  y  $y'$ .

Se detectaron 10 observaciones atípicas, las cuales son señaladas por los círculos en color rojo. Estas observaciones se encuentran alejadas considerablemente de la recta de regresión, esto ocasiona que los residuales correspondientes a estos puntos sean muy elevados y que el ajuste del modelo disminuya. Debido a que la presencia de uno o más puntos atípicos puede introducir serias distorsiones en el ANOVA para la regresión, se retiraron dichas observaciones para realizar nuevamente el análisis de regresión lineal.

En la Figura 5.10 se muestra la gráfica de la recta de regresión para las variables  $x'_4$  y  $y'$  sin observaciones inusuales, su ecuación con la estimación de los coeficientes  $\beta_0$  y  $\beta_1$  se muestra en (5.2) y  $r^2 = 0.7000$ .

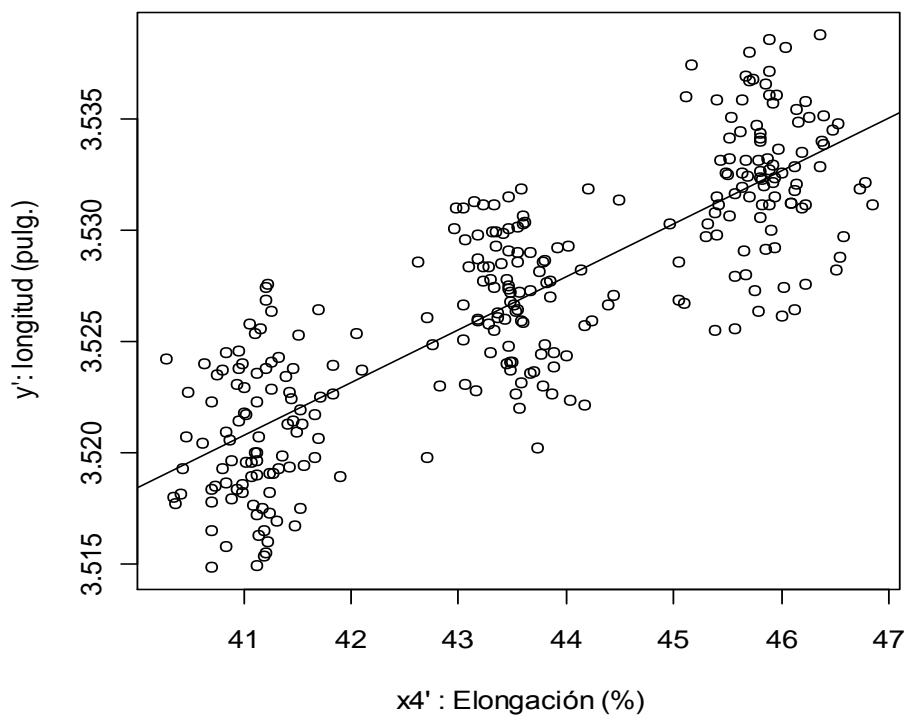


Figura 5.10 Recta regresión sin observaciones inusuales.

$$\hat{y}' = 3.423112 + 0.002381 (x'_4) \quad (5.2)$$

En la Tabla 5.4 se muestra el ANOVA utilizado para aislar la variación debida a la regresión (SCTr) de la variación total en el experimento, además se obtiene la variación residual (SCE) o variación inexplicada. La cantidad STC es igual a 0.0091114.

Tabla 5.4 ANOVA para la regresión entre  $x'_4$  y  $y'$

Fuente	gl	SC	CM	F	Valor p
Regresión	1	0.0063763	0.0063763	671.41	0.000
Error -residual	288	0.0027351	0.0000095		
Total	289	0.0091114			



Para probar la existencia de una relación lineal con un nivel de significancia de 5% entre las variables, se definió la  $H_0: \beta_1 = 0$  contra la  $H_a: \beta_1 \neq 0$ . El valor de tabla del estadístico de prueba F con  $gl_1 = 1$  y  $gl_2 = 288$  es de 3.87, dado que el valor observado  $F = 671.41$  caen en la región de rechazo, se rechaza  $H_0$  y se concluye que hay una relación lineal significativa entre las variables.

Es posible llegar a la misma conclusión observando que el valor p en la Tabla 5.4 del ANOVA es menor que el nivel de significancia  $\alpha = 0.05$ , por lo tanto se apoya la  $H_a$  declarando que  $\beta_1$  es diferente de cero, es decir, existe una pendiente con valor numerico igual a 0.002381 para la recta de regresión.

El coeficiente  $r^2 = SCTr/STC = 0.0063763/0.0091114 = 0.7000$ , lo que significa que el 70% de la variación total es explicada por la regresión lineal de  $y'$  sobre  $x'_4$ , el modelo funciona adecuadamente y existe solidez entre las variables.

### **Verificación de la adecuación del modelo.**

En primer lugar se verificó el supuesto de normalidad, realizando un histograma de los residuales. En la Figura 5.11 se satisface el supuesto de normalidad para los errores, por lo que la gráfica aparece como una muestra de una distribución normal con centro en cero.

Adicionalmente se realizaron pruebas de normalidad numéricas donde se probaron las siguientes hipótesis:

*$H_0$ : los residuales provienen de una distribución normal*

*$H_1$ : los residuales provienen de otra distribución*

De acuerdo con lo prueba Anderson-Darling el valor p es igual a 0.215, y para la prueba Kolmogorov- Smirnov p adquiere un valor igual a 0.3087, es decir, en ambas pruebas dicho valor es mayor que el nivel de significancia  $\alpha = 0.05$  por lo tanto se sabe que la  $H_0$  es aceptada, declarando que los residuales están distribuidos normalmente.

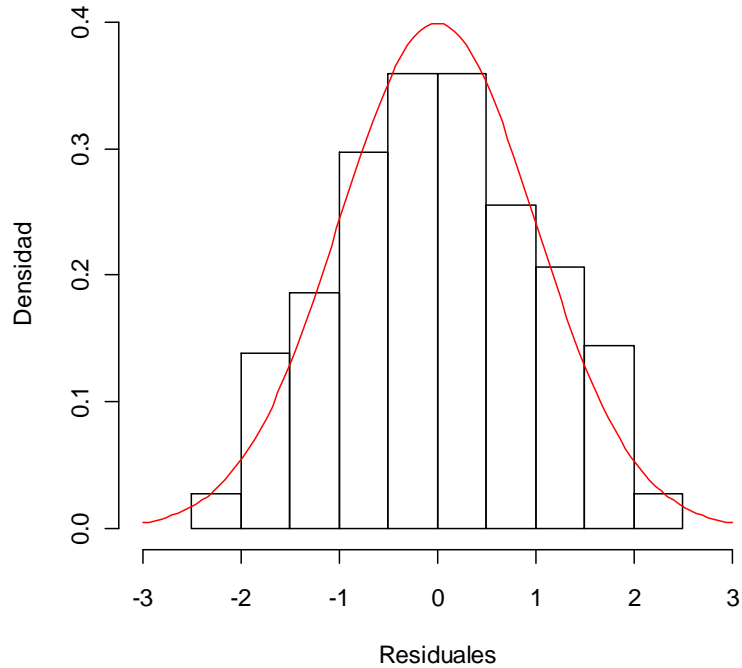


Figura 5.11 Verificación del supuesto de normalidad.

En la Figura 5.12 se presenta una gráfica de los residuales en el orden temporal de la recolección de los datos con el fin de detectar correlaciones entre los residuales. Dado que en la figura no aparecen tendencias o patrones se puede afirmar que el supuesto de independencia entre los residuales se satisface.

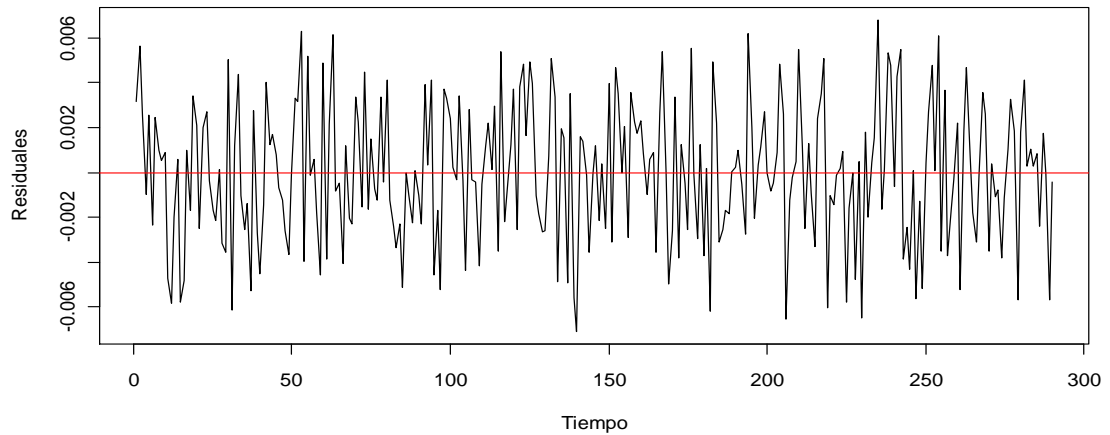


Figura 5.12 Verificación del supuesto de independencia.

En la Figura 5.13 se observa que los residuales no están relacionados con ninguna otra variable, específicamente con los valores ajustados, debido a que no se observa ningún patrón inusual en la gráfica. También se observa que hay homogeneidad de varianzas.

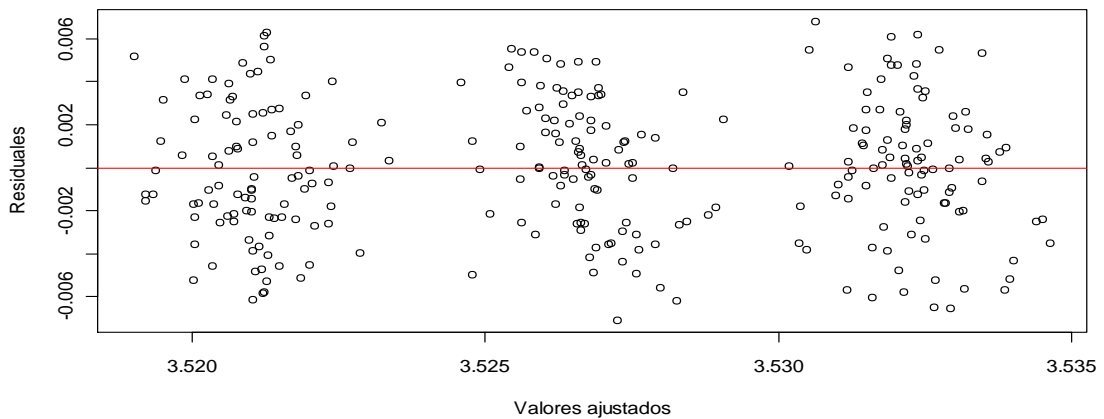


Figura 5.13 Verificación de independencia con otra variable.

En la Figura 5.14, se grafican los intervalos de confianza para la regresión lineal, representados por las líneas punteadas en color azul, los cuales fueron calculados por (3.18). Para los intervalos de predicción, en líneas rojas discontinuas, se utilizó la ecuación (3.19). Ambos intervalos son calculados a un nivel del 95%. Observe que el intervalo de predicción es más

amplio que el intervalo de confianza debido a la variabilidad extra al predecir el valor real de la respuesta  $y'$ . Además el 96.20% de las observaciones se encuentran dentro del intervalo de predicción marcado por las líneas punteadas azules, lo que corresponde a 279 observaciones de 290.

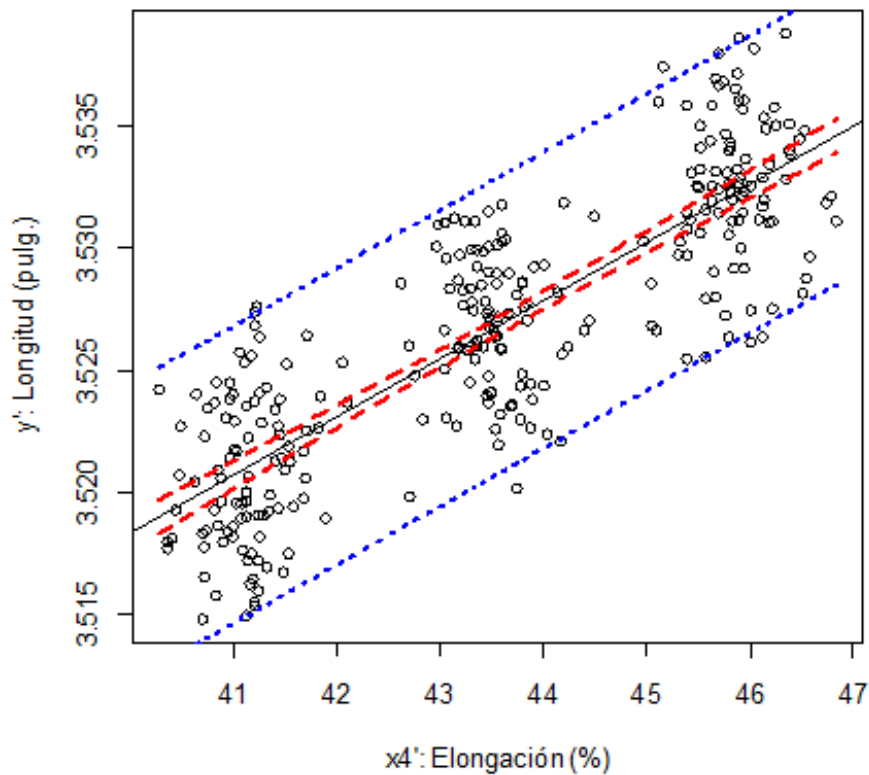


Figura 5.14 Intervalos de confianza para la regresión lineal.

El modelo de regresión lineal obtenido en (5.2) describe a la variable  $y'$  en términos de la variable  $x'_4$  y el coeficiente de correlación entre estas variables es significativamente alto como se mostró con anterioridad. Considerando la relación lineal que tiene el porcentaje de elongación con la característica de calidad  $y$  del álabe se requiere que dicha variable sea inspeccionada con regularidad.

### 5.3 ETAPA 8: Plan de muestreo de aceptación para $x_4$ .

De acuerdo a la sección 5.1.2 se identificó la necesidad de inspeccionar la variable  $x_4$  en la materia prima. Toda inspección puede ser realizada mediante diferentes estrategias. Sin embargo, de acuerdo a la sección 4.2.4, no es factible realizar una inspección del 100% de la variable  $x_4$ , ya que dicha inspección consiste en pruebas mecánicas realizadas por agentes externos como: talleres de maquinado y laboratorios. La intervención constante de tales agentes para la ejecución de dichas pruebas, repercute económicamente y afecta la planeación de la producción. En estas circunstancias, es mejor disponer de una cantidad pequeña de muestras para su inspección, es decir, se requiere de un plan de muestreo para especificar el número de artículos que hay que inspeccionar y el criterio para juzgar si los rollos de materia prima son aceptados o rechazados.

La selección adecuada de la materia prima implica que los rollos que no cumplan con el criterio de aceptación sean sometidos a una disposición de lote alterna, donde tales rollos puedan ser manufacturados según sus características particulares.

Actualmente en el proceso se inspecciona la variable  $x_3$ , pero las observaciones de dicha variable no poseen un coeficiente  $r$  suficientemente alto de acuerdo con la sección 5.1.2, es por eso que una mejor estrategia es que la empresa dirija sus esfuerzos de inspección hacia la variable  $x_4$ .

A continuación se detalla la información requerida para diseñar el plan de muestreo de la variable  $x_4$ .

Los rollos adquiridos con una elongación promedio de 45.8515, se desean aceptar el 95% de las veces. Por otra parte todos aquellos rollos con

una elongación promedio de 41.13592 no son efectivos, y al ser adquiridos se deben rechazar el 90% de la veces. Se estimó la desviación estándar,  $\sigma = 1.973044$  del historial del porcentaje de elongación del proveedor. Se determinó cuántas muestras  $n$  de rollos se deben inspeccionar para el porcentaje de elongación y cuál debe ser su criterio de aceptación,  $\bar{x}_\alpha$ , como se describe a continuación.

Los riesgos del proveedor y del cliente están dados como  $\alpha = 0.05$  para  $\mu_\alpha = 45.8515$  y  $\beta = 0.1$  para  $\mu_\beta = 41.13592$ , respectivamente. Para  $\alpha = 0.05$  y  $1 - \beta = 0.9$ , los valores de  $z$  de la distribución normal estándar son  $z_{0.05} = -1.645$  y  $z_{0.9} = 1.282$ .

Empleando la ecuación (3.31) se obtiene un tamaño de muestra  $n = 1.49985 \approx 2$  y para determinar el criterio de aceptación se utilizan las ecuaciones (3.29) y (3.30), donde  $\bar{x}_\alpha = 43.2013$ .

Lo que representa que el plan de muestreo consiste en inspeccionar para cada colada de materia prima con  $N=15$  rollos, el porcentaje de elongación de 2 rollos y aceptar el rollo siempre que la elongación promedio de estas muestras sea mayor de 43.2013.

En base a este plan de muestreo se construyó la curva de operación característica mostrada en la Figura 5.15, donde se puede observar la probabilidad de aceptación para diferentes valores de la elongación promedio. Donde  $P(1 - \alpha)$  representa el punto de riesgo para el proveedor y  $P(\beta)$  representa el punto de riesgo para el cliente. Además se muestra el criterio de aceptación calculado para el plan de muestreo.

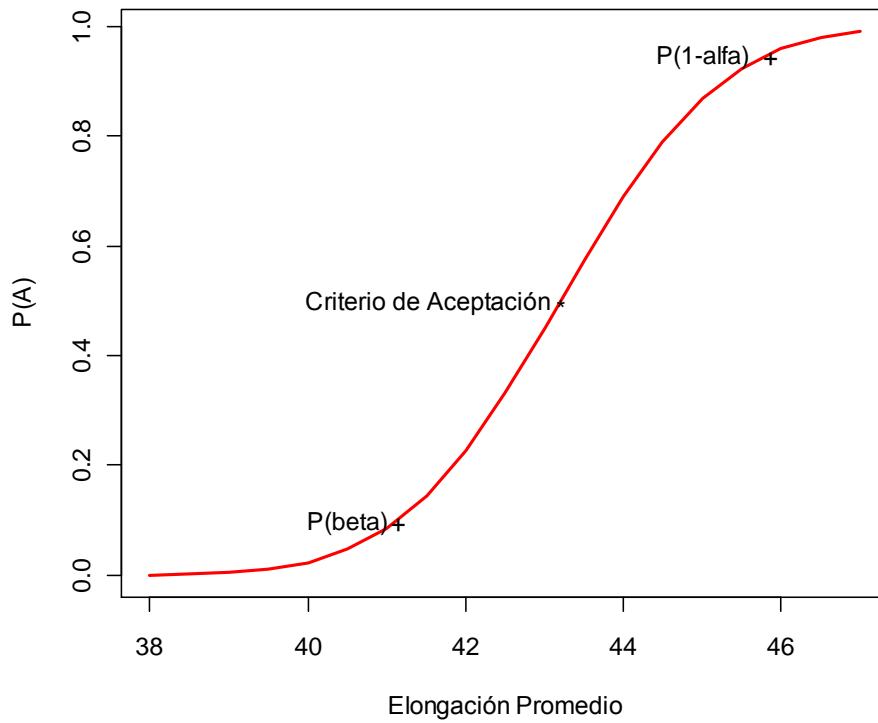


Figura 5.15 Curva de operación para el plan de muestreo.

Tal como se esperaba con el plan de muestreo diseñado, los porcentajes de elongación promedio mayores de  $\bar{x}_a = 43.2013$  tienen una probabilidad de aceptación mayor, mientras que los porcentajes de elongación promedios menores que el criterio de aceptación presentan una probabilidad de aceptación pequeña.

## **CAPÍTULO 6 CONCLUSIONES Y RECOMENDACIONES**

Con la ejecución del diagnóstico del estado del proceso fue posible definir que la variación en la longitud del álabe se encuentra fuertemente relacionada con la variación que poseen los lotes de materia prima. En dicho diagnóstico se detalló el comportamiento del porcentaje de productos defectuosos explicando esta situación mediante el uso de herramientas estadísticas tal como se planteó en los objetivos específicos.

Al validar la información que el proveedor entrega a la empresa, por medio de pruebas mecánicas, se identificó que existe mucha diferencia entre los reportes del proveedor y los resultados de las pruebas. Dado que dichas pruebas mecánicas fueron realizadas bajo los procedimientos que marca la norma ASTM E8-04 y validadas por especialistas de COMIMSA, se tiene certeza de que dichos resultados son correctos y que los reportes del proveedor no son confiables. De acuerdo con las pruebas mecánicas, la mejor colada de la materia prima es la 26015, porque presenta un mayor porcentaje de elongación y requiere menos carga para ser deformada.

En el Capítulo 5 se concluye que de acuerdo con el ANOVA y el diagrama de caja realizados, la mejor colada es la 26015, porque produce mejores mediciones de la longitud del álabe ( $y$ ), dentro del rango de especificación: [3.525-3.555] pulgadas. Dentro de este capítulo se logró identificar la variable de entrada que posee mayor relación lineal con la variable ( $y$ ), cumpliéndose uno de los objetivos específicos planteados en el Capítulo 2.



Se explicó la relación entre estas variables mediante una ecuación de regresión lineal, presentando ésta un buen ajuste y el cumplimiento de los supuestos de la regresión lineal, con lo cual se afirma que el modelo es útil al explicar la variable  $y$  en términos de  $x_4$ .

De los resultados del plan de muestreo analizado, se asegura que al ejecutar dicho plan, la probabilidad de aceptar materia prima en buenas condiciones,  $P(1 - \alpha)$  se incrementa cuando un valor de elongación promedio sea mayor que la  $\bar{x}_a = 43.2013$  como se observa en la parte izquierda de Tabla 6.1 en la región sombreada de color azul. Por el contrario, la probabilidad de aceptación de materia prima en condiciones inadecuadas,  $P(\beta)$  decrece, cuando la elongación promedio es menor que  $\bar{x}_a$ , como se observa en la parte derecha en la región sombreada.

Tabla 6.1 Probabilidades calculadas para  $\alpha$  y  $\beta$ , izquierda y derecha respectivamente.

Elongación Promedio	$P(\alpha)$	$P(1 - \alpha)$	Elongación Promedio	$P(1 - \beta)$	$P(\beta)$
47	0.00918976	0.99081024	43.2013	0.50000000	0.50000000
46.5	0.02030311	0.97969689	43	0.54971782	0.45028218
46	0.04117817	0.95882183	42.5	0.6683285	0.3316715
45.8515	0.04998495	0.95001505	42	0.77206243	0.22793757
45.5	0.07681577	0.92318423	41.5	0.85451818	0.14548182
45	0.13211115	0.86788885	41.13592	0.9000782	0.0999218
44.5	0.21008882	0.78991118	41	0.91408761	0.08591239
44	0.31003214	0.68996786	40.5	0.95320116	0.04679884
43.5	0.42645554	0.57354446	40	0.97654277	0.02345723
43.2013	0.50000000	0.50000000	39.5	0.98920274	0.01079726

Al disponer de una materia prima que cumple con las especificaciones, se asegura que al manufacturar álabes, la longitud de los mismos quede dentro de las especificaciones de la empresa.

## **Recomendaciones.**

Para la construcción de la base de datos, es importante que la recolección de la información se realice de manera no secuencial, es decir, debe ser aleatoriamente, para evitar dependencia con el tiempo.

Con respecto al plan de muestreo, se debe considerar que las unidades seleccionadas para inspección del lote deben ser tomadas aleatoriamente, y deben ser representativas de todos los artículos del lote. Esto es importante porque de otra forma existiría sesgo en el plan de muestro, dejando oportunidad a que el proveedor haya manipulado la colocación de piezas buenas al inicio del lote, afectando de esta manera los resultados de la inspección.

En recomendación directa a la empresa, los lotes de materia prima que sean rechazados en el plan de muestreo, no necesariamente tienen que devolverse al proveedor. Una estrategia es que dichos lotes sean manufacturados con una configuración de maquinaria diferente.

## **Trabajo Futuro.**

Rencher (2002) menciona que el modelo de regresión múltiple y el de regresión polinomial permiten usar al mismo tiempo otras variables independientes para explicar la variación en  $y$ , lo que sería un recurso útil para extender la investigación.

En la Figura 6.1 se muestran las ecuaciones de los modelos que pueden usarse en un trabajo futuro, para el caso del modelo de regresión polinomial  $y$

no depende de variables independientes diferentes, es decir la variable  $x$  posee diferentes exponentes desde  $x^2, x^3, \dots, x^n$ .

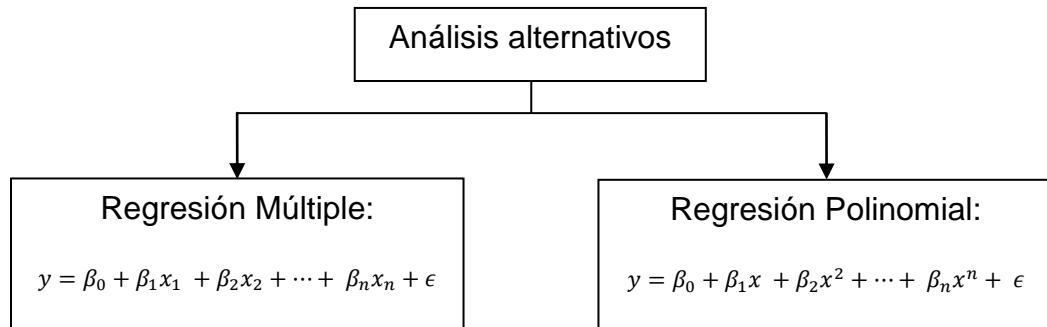


Figura 6.1 Diagrama de modelos alternativos.

Adicionalmente se podría diseñar un plan de muestreo para más de una variable o diseñar planes con una técnica de muestreo diferente, tales como muestreo doble o muestreo múltiple.

## REFERENCIAS

- Automotive Industry Action Group (2005). Statistical Process Control (SPC). 2da ed. Manual de referencia. Daimler Chrysler Corporation, Ford Motor Company y General Motors Corporation.
- Altan, T., Oh, S., y Gegel, H. L. (1986). *Metal Forming Fundamentals and Applications*. 3ra ed. American Society for Metals, E.U.A.
- Bosga, S. y Segura, M. (2007). Pressing challenge. ABB Review, 3, 58-59.
- Canavos, George C. (1998). *Applied Probability and Statistical Methods*. McGraw Hill. U.S.A.
- De Vin, L. J. (2000). Curvature prediction in air bending of metal sheet. Journal of Materials Processing Technology, 100, 1-3, 257-261.
- Devore, J.L. (2000). *Probability and Statistics for engineering and the Sciences*. 5 ed. Duxbury Thomson Learning. U.S.A.
- Di Bucchianico, A. (2008). Lecture Notes 3TU Course Applied Statistics, Technische Universiteit Eindhoven.
- García- Romeu de Luna, M. L. y de Ciurana. Q. (2005). Contribución al estudio del proceso de doblado al aire de chapa. Modelo de predicción del ángulo de recuperación y del radio de doblado final. Tesis Doctoral. Universidad de Girona, España.
- Jin, J. y Guo, H. (2003). Anova Method for variance component decomposition and diagnosis in batch manufacturing processes. The International Journal of Flexible Manufacturing Systems. Holanda, 15, 167-186.
- Kiermeier, A. (2008). Visualizing and assessing Acceptance Sampling Plans: The R Package Acceptance Sampling. Journal of Statistical Software. Australia. Vol. 26, No. 6.
- Limones, J.A., López J., I. y Vázquez, J.A. (2010). Mathematical model for defect variability in the stamping process of a manufactured

subcomponent. 1<sup>st</sup> International Congress on Instrumentation and Applied Sciences, Cancun, Q.R., Mexico.

Limones, J.A., Vázquez, J.A. y López J., I. (2011). Linear regression model to minimize the variation in a stamping product of automotive industry. 16<sup>th</sup> Annual International Conference on Industrial Engineering Theory, Applications & Practice, Stuttgart, Germany.

Mendenhall, W., Beaver, R. J., y Beaver, B. M. (2008). *Introducción a la Probabilidad y Estadística*. 12<sup>a</sup> ed. Cengage Learning, México.

Montgomery, D. C. (2009). *Introduction to Statistical Quality Control*. 6<sup>a</sup> ed. Wiley, U. S. A.

Montgomery, D. C. (2007). *Diseño y análisis de experimentos*. 2da ed. Ed Limusa-Wiley

Montgomery, D. C. y Runger G. C. (2007). *Probabilidad y Estadística aplicadas a la Ingeniería*. 3er ed. Ed. Limusa-Wiley.

Morgenthaler, S. (2009). Exploratory Data Analysis. Wires Computational Statistics. John Wiley and Sons, Inc. Vol. 1, Julio / Agosto 2009, 33-44.

Ojeda, E. (2005). Control de calidad estadístico, parte dos, Muestreo de Aceptación. Universidad de Santiago de Chile.

R Development Core Team (2009). Creation and evaluation of Acceptance Sampling Plans. Manual de referencia. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>

Ramos, R. (2010). El lenguaje de programación R, notas de clase. Centro de Investigación en Matemáticas. Guanajuato, México.

Real Academia Española. (2003). Diccionario de la lengua española. 22a ed. Ed. Espasa Calpe.

Rencher, A. C. (2002). *Methods of Multivariate Analysis*. 2da ed. Ed. Wiley – Interscience. U. S. A.

- Schilling G.E. y Neubauer V. D. (2009). *Acceptance Sampling in Quality Control*. 2da ed. Ed. CRC Press.
- Senjuntichai, A. y Pitthayakornpisuth, A. (2010). Application of General Linear Model to the reduction of defectives in packaging process of soap industry. International MultiConference of Engineers and Computer Scientists. Hong Kong, vol. 3.
- Yuqi, L., Ping, H. y Jincheng, W. (2002). Springback simulation and analysis of strong anisotropic sheet metals in U-channel bending process. Acta Mechanica Sinica. New York, U.S.A. vol. 18, 3, 264-273.
- Tekkaya, A. E. (2000). State of the art of simulation of sheet metal forming. Journal of Materials Processing Technology, 103 (1) 14-22
- Wackerly D. D, Mendenhall W. y Scheaffer L. R. (2008). *Mathematical Statistics with applications*. 7a. ed. Ed. Duxbury.

## ANEXO 1

### Códigos de programación en el paquete R

Figura 2.3 Diagrama de Pareto para los defectos presentados en el álabe.

```
pareto=read.csv("pareto.csv")
def= pareto$cantidad
names(def)<- c("Golpes", "M.R", "P. l", "Piezas de Ajuste")
pareto.chart(def, ylab="Cantidad de piezas", xlab="Defecto", las=0)
```

Figura 3.1 Histograma para las observaciones de la Tabla 3.2.

```
set.seed(200)
a=rnorm(50,0,1)
b=matrix(a,10,5)
hist(b,nclass=10,main="Histograma3.1",
ylab="Observaciones",ylab="Frecuencia")
```

Figura 3.2 Diagrama de caja para las observaciones de la Tabla 3.2

```
a1=boxplot(a, ylim=c(-2.5,2.5))
fivenum= a1$stats
fivenum
q1= fivenum[2,1];q3= fivenum[4,1]
LI=q1 -1.5*(q3-q1); LS=q3 +1.5*(q3-q1); LI;LS
abline(h= LI, lty = 2); abline(h= LS, lty = 2)
text(0.75,-0.44,"Q1--");text(0.75,-0.05,"m--");text(0.75,0.44,"Q3--")
text(1.18,-1.45,"mínimo");text(1.18,1.6,"máximo")
text(0.7,-2,"Limite Inferior"); text(0.7,1.9,"Limite Superior")
legend(1.2,-2.1,"Punto Atipico" , cex=0.9, col= 1, pch=1)
```

Figura 3.3 Diagrama de dispersión

```
set.seed(200)
xi=rnorm(50,0,1)
set.seed(500)
yi=rnorm(50,0,1)
xi,yi
plot(xi,yi,xlab="Observaciones de x", ylab="Observaciones de y")
```

Figura 3.4 Recta de regresión o recta de mínimos cuadrados.

```
lm(yi~xi)
abline(lm(yi~xi))
summary(lm(yi~xi))
```

Figura 4.3 Diagrama de dispersión para  $x$  y  $y$ .

```
datos=read.csv("dispersion.csv")
x=(datos$Espesor)
y=(datos$Longitud)
```

```
plot(x,y,xlab="x : Espesor (mm)", ylab="y : Longitud (pulg.)")
```

Figura 4.4 Recta de regresión

```
mx<- c(1.19,1.20,1.21,1.22)
my<- c(3.5381,3.5390,3.5404,3.5416)
c=c(mx,my)
cm= matrix(c)
cm
plot(mx,my,xlab="x : Espesor (mm)", ylab="y : Longitud (pulg.)")
lm(my~mx)
abline(lm(my~mx))
summary(lm(my~mx))
```

Figura 4.5 Gráfica integral de regresión.

```
inter=read.csv("intervalos.csv")
interx<- as.numeric(inter$x)
nx<-max(inter$x)
xrange <- range(inter$espesor)
yrange <- range(inter$longitud)
plot(xrange, yrange, type="n", xlab="x : Espesor (mm)",ylab="y: Longitud
(pulg.)")
colors <- c(col="red", "black", "green", "blue")
pech <- c(pch= 1,2,3,3)
types <- c(type= "p", "o", "p", "p")
a1<- matrix(c("y","y ajustada", "LI", "LS"))
for (i in 1:nx){
x <- subset(inter, x==i)
lines(x$espesor, x$longitud, type= types[i],pch=pech[i], lwd=1, col=colors[i])
legend(xrange[1], yrange[2],a1 , cex=0.7, col= colors, pch=pech)
}
```

Tabla 5.2 Análisis de varianza de un solo factor.

```
caja=read.table("longitud.txt")
longitud=(caja$V2)
colada=(caja$V1)
anova(lm(longitud~factor(colada)))
```

Figura 5.1 Diagrama de caja para los tres grupos de datos.

```
caja=read.table("longitud.txt")
caja
colada=(caja$V2)
longitud=(caja$V1)
box=boxplot(colada~longitud, xlab="# colada ", ylab="Variable y (pulg.)")
box
```



Tabla 5.3 Coeficiente de correlación  $r$  entre las variables.

```
variables=read.csv("variables.csv")
x1= variables$fluencia
x2= variables$cedencia
x3= variables$espesor
x4= variables$elongacion
x5= variables$carbon
x6= variables$grano
x7= variables$dureza
x8= variables$apertura
y=variables$longitud
cor1=cor(x1,y)
cor2=cor(x2,y)
cor3=cor(x3,y)
cor4=cor(x4,y)
cor5=cor(x5,y)
cor6=cor(x6,y)
cor7=cor(x7,y)
cor8=cor(x8,y)
cor1;cor2;cor3;cor4;cor5;cor6;cor7;cor8
rs=c(cor1,cor2,cor3,cor4,cor5,cor6,cor7,cor8)
max(rs)
rm(list=ls(all=TRUE))
reales= read.csv("300.csv")
reales
xo= reales$elongacion
xo= matrix(xo)
yo= reales$longitud
yo= matrix(yo)
xo;yo
plot(yo)
hist(yo)
mean(yo)
sd(yo)
mean(xo)
sd(xo)
y1= yo[1:100]
y2= yo[101:200]
y3= yo[201:300]
y1;y2;y3
m1= mean(y1)
m2= mean(y2)
m3= mean(y3)
m1;m2;m3
s1= sd(y1)
s2= sd(y2)
```

```

s3= sd(y3)
s1;s2;s3
set.seed(500)
ys1=rnorm(100,m1,s1)
ys2=rnorm(100,m2,s2)
ys3=rnorm(100,m3,s3)
ny<- c(ys1,ys2,ys3)
ny= matrix(ny)
mean(ny)
sd(ny)
plot(ny)
hist(ny)
a= seq(length=300, from= 40, to=47)
a = matrix(a)
m=a[1:100]
m=mean(m)
s= (42.31773-40)/6
set.seed(100)
d1= rnorm(100,m,s)
d1= matrix(d1)
d1
m2=a[101:200]
m2=mean(m2)
s2= (44.65886-42.34114)/6
s2
set.seed(100)
d2= rnorm(100,m2,s2)
d2= matrix(d2)
d2
m3=a[201:300]
m3=mean(m3)
s3= (47-44.68227)/6
s3
set.seed(100)
d3= rnorm(100,m3,s3)
d3= matrix(d3)
d3
nx= matrix(c(d1,d2,d3))
plot(nx)
mean(nx)
sd(nx)

```

Figura 5.6 Observaciones para la elongación.

```

par(mfrow=c(1,2))
plot(xo,xlab= "", ylab="x4: (%) ")
mtext(expression(s==0),side=1,line=2)
mtext(expression(bar(x)==43.66),side=1,line=3)

```

```

plot(nx, xlab="", ylab="x4':(%)")
mtext(expression(s==1.95),side=1,line=2)
mtext(expression(bar(x)==43.66),side=1,line=3)

```

Figura 5.7 Observaciones de la longitud.

```

par(mfrow=c(1,2))
plot(yo,xlab= "", ylab="y: (pulg.) ")
mtext(expression(s==0.0057),side=1,line=2)
mtext(expression(bar(y)==3.526),side=1,line=3)
plot(ny, xlab="", ylab="y':(pulg.)")
mtext(expression(s==0.0057),side=1,line=2)
mtext(expression(bar(y)==3.526),side=1,line=3)

```

Figura 5.8 Diagrama de dispersión

```

par(mfrow=c(1,2))
plot(xo,yo, xlab= " x4 : (%) ", ylab=" y : (pulg.) ")
plot(nx,ny, xlab= " x4' : (%) ", ylab=" y' : (pulg.) ")
ro=cor(xo,yo)
nr=cor(nx,ny)
plot(xo,yo,xlab=" x2: deformation (%) ", ylab=" y: length (mm) ")
z=lm(yo~xo)
abline(lm(yo~xo))
summary(lm(yo~xo))
anova(z)
a=matrix(residuals(z))
acf(residuals(z))
plot(z)
require(nortest)
ad.test(a)
shapiro.test(a)
hist(a)
boxplot(a)
mean(a)
a1=a[1:100]
a2=a[101:200]
a3=a[201:300]
var(a1)
var(a2)
var(a3)
var(a)

```

Figura 5.9 Recta de regresión para  $x'_4$  y  $y'$ .

```

plot(nx,ny,xlab=" x4' : Elongación (%) ", ylab=" y': longitud (pulg.) ")
z=lm(ny~nx)
abline(lm(ny~nx))
summary(lm(ny~nx))
anova(z)

```

```

a=matrix(residuals(z))
acf(residuals(z))
plot(z)
require(nortest)
ad.test(a)
shapiro.test(a)
hist(a)
mean(a)
a1=a[1:100]
a2=a[101:200]
a3=a[201:300]
var(a1)
var(a2)
var(a3)
var(a)
var(nx)

```

Figura 5.10 Recta regresión sin observaciones inusuales.

```

ajustados=read.csv("ajustados10.csv")
nx=ajustados$x4as
ny=ajustados$yas
plot(nx,ny,main="Modelo ajustado",xlab=" x4' : Elongación (%) ", ylab=" y':
longitud (pulg.) ")
z=lm(ny~nx)
abline(lm(ny~nx))
summary(lm(ny~nx))
anova(z)

```

Figura 5.11 Verificación del supuesto de normalidad.

```

a=matrix(residuals(z))
acf(residuals(z), xlab="Retraso",ylab="FAC")
par(mfrow=c(1,2))
plot(z)
require(nortest)
ad.test(a)
lillie.test(a)
qqnorm(a, xlab="Cuantiles teoricos", ylab="Cuantiles muestrales")
astd=(a-mean(a))/sd(a)
hist(astd,freq=F, xlim=c(-3,3),ylim=c(0,0.40),xlab="Residuales",
ylab="Densidad")
curve(dnorm,col=2,add=T)

```

Figura 5.12 Verificación del supuesto de independencia.

```

plot(a,type="l", xlab="Tiempo", ylab="Residuales")
abline(h=0, col=2)

```

Figura 5.13 Verificación de independencia con otra variable.

```
ya=3.423112+ 0.002381*(nx)
plot(ya,a, xlab="Valores ajustados", ylab="Residuales")
abline(h=0, col=2)
```

Figura 5.14 Intervalos de confianza para la regresión lineal.

```
regconf <- function(nx,ny,titulo,titx,tity,qq=0.975){
plot(nx,ny,main=titulo,xlab=titx,ylab=tity)
reg <- glm(ny~nx)
abline(reg)
dif <- ny-reg$fitted.values
Sxy <- sqrt(sum(dif*dif)/(length(dif)-2))
Sx2 <- var(nx)
meanx <- mean(nx)
meany <- mean(ny)
n <- length(nx)
df <- n-2
xs <- sort(nx)
tvalue <- qt(qq,df)
confqqplus <- meany+reg$coefficients[2]*(xs-meanx)+tvalue*Sxy*sqrt(1/n+((xs-
meanx)*(xs-meanx)/((n-1)*Sx2)))
confqqmin <- meany+reg$coefficients[2]*(xs-meanx)-tvalue*Sxy*sqrt(1/n+((xs-
meanx)*(xs-meanx)/((n-1)*Sx2)))
lines(xs,confqqmin,lty="dashed",col="red", lwd= 2)
lines(xs,confqqplus,lty="dashed", col="red", lwd= 2)
confqqplus2 <- meany+reg$coefficients[2]*(xs-
meanx)+tvalue*Sxy*sqrt(1+1/n+((xs-meanx)*(xs-meanx)/((n-1)*Sx2)))
confqqmin2 <- meany+reg$coefficients[2]*(xs-meanx)-
tvalue*Sxy*sqrt(1+1/n+((xs-meanx)*(xs-meanx)/((n-1)*Sx2)))
lines(xs,confqqmin2,lty="dotted", col=" blue",lwd= 2)
lines(xs,confqqplus2,lty="dotted", col=" blue ",lwd= 2)
}
op <- par(mfrow = c(1, 1))
regconf(nx,ny,"title","x4': Elongación (%)","y': Longitud (pulg.)")
```

Figura 5.15 Curva de operación para el plan de muestreo.

```
medias=curva$medias
pacep=curva$paceptacion
yrange=c(38,47,0.5)
plot(medias,pacep,type="l",col=2, xlab="Elongación Promedio",
ylab="P(A)",lwd=2)
text(45.8515,0.95,"+");text(45,0.95,"P(1-alfa)")
text(43.2013,0.5,"*");text(41.5,0.5,"Criterio de Aceptación")
text(41.13592,0.1,"+");text(40.5,0.1,"P(beta)")
```

## APÉNDICE

De las ecuaciones mostradas en la sección 3.2.5

$$\bar{x}_\alpha = (\sigma/\sqrt{n}) z_\alpha + \mu_\alpha \quad (3.29)$$

$$\bar{x}_\beta = (\sigma/\sqrt{n}) z_{1-\beta} + \mu_\beta \quad (3.30)$$

se obtiene  $n$  al igualar y resolver de la siguiente forma:

$$(\sigma/\sqrt{n}) z_\alpha + \mu_\alpha = (\sigma/\sqrt{n}) z_{1-\beta} + \mu_\beta$$

$$\mu_\alpha - \mu_\beta = (\sigma/\sqrt{n}) z_{1-\beta} - (\sigma/\sqrt{n}) z_\alpha$$

$$\mu_\alpha - \mu_\beta = \frac{\sigma}{\sqrt{n}} (z_{1-\beta} - z_\alpha)$$

$$\sqrt{n} = \frac{\sigma(z_{1-\beta} - z_\alpha)}{\mu_\alpha - \mu_\beta}$$

$$n = \left[ \frac{\sigma(z_{1-\beta} - z_\alpha)}{\mu_\alpha - \mu_\beta} \right]^2 \quad (3.31)$$

La cual sirve para conocer el tamaño de la muestra  $n$  con el que se realiza un plan de aceptación para variables.