

**CORPORACIÓN MEXICANA DE INVESTIGACION EN MATERIALES
(COMIMSA)**

DIVISIÓN DE ESTUDIOS DE POSGRADO



***MODELADO DEL PROCESO DE SOLDADURA
POR FRICCIÓN AGITACIÓN CONSIDERANDO
MÚLTIPLES DEFECTOS CATEGÓRICOS***

TESIS

PRESENTA

Edel Kerim Hernández Flores

**MAESTRÍA EN CIENCIA Y TECNOLOGÍA EN
INGENIERÍA INDUSTRIAL Y DE MANUFACTURA**

Saltillo, Coahuila. Diciembre del 2016

**MODELADO DEL PROCESO DE SOLDADURA
POR FRICCIÓN AGITACIÓN CONSIDERANDO
MÚLTIPLES DEFECTOS CATEGÓRICOS**

Por

Edel Kerim Hernández Flores

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

*Maestría en Ciencia y Tecnología
en Ingeniería Industrial y de Manufactura*

Programa Interinstitucional en Ciencia y Tecnología COMIMSA / CONACyT

Saltillo, Coahuila diciembre de 2016

Dedicado

EN ESPECIAL A MI ABUELO QEPD:

Abraham Flores Olvera

A MI ABUELA:

Antonia Rentería de Flores

A MIS TIOS:

Luis Daniel Pérez Molina

Francisca Concepción Flores Rentería

A DOS PERSONAS MUY ESPECIALES

Paola Fernández Contreras

José Ángel Ramírez Fernández

A MIS AMIGOS

Agradecimientos

A mis abuelos **Abraham Flores Olvera** y **Antonia Rentería de Reyes**, por darme la vida, por su apoyo incondicional y por brindarme los ánimos suficientes para continuar adelante, pero especialmente a mi Abuelo por la sabiduría, amor, alegría y enseñanza que me brindó y que al fallecer me quedé sin el principal pilar de mi vida, que a pesar de cometer muchos errores y perder mis metas y cordura, le dedico con lágrimas y sufrimiento este trabajo por todo su amor.

A mis tíos **Luis Daniel Pérez Molina** y **Francisca Concepción Flores Rentería** por el apoyo, soporte y ánimos otorgados en cada etapa de éstos dos años en el programa de Maestría.

A **Paola Fernández Contreras** y **José Ángel Ramírez Fernández** por procurar mis avances, mi estado de ánimo, por sacar adelante mi mente y por el impulso otorgado durante este trayecto.

A **mis amigos** por apoyarme en todo tipo de situación presentada en ésta etapa y anteriores para salir adelante.

A mi **tutor, Dr. David Salvador González González** por compartir sus conocimientos, tiempo, espacio y paciencia, por los regaños e impulsos que me brindada durante todo el trayecto en el que se desvió mi mente.

A mis **asesores, Dr. Rolando Javier Praga Alejo** y **Dra. Argelia Fabiola Miranda Pérez** por las palabras, regaños, conocimientos y espacio otorgado en cada una de las etapas del programa de Maestría.

A los académicos que de forma directa o indirecta contribuyeron a mi formación académica, **Dr. Pedro Pérez, Dr. Arturo Reyes.**

Al **personal administrativo de posgrado** de COMIMSA por su apoyo en cada una de las actividades realizadas en mi formación académica.

Agradezco a **El Consejo Nacional en Ciencia y Tecnología (CONACYT)** y a la **Corporación Mexicana de Investigación en Materiales** por todos los apoyos

recibidos durante mi formación en el programa de maestría ya que sin su ayuda mi progreso y desarrollo académico, actividades de movilidad e investigación y producción científica no hubieran sido posibles.

Índice

Introducción.....	1
Planteamiento del problema.....	3
2.1 Descripción del problema.....	3
2.2 Preguntas de investigación.....	4
2.3 Hipótesis.....	5
2.4 Objetivo general.....	5
2.5 Objetivos específicos.....	5
2.6 Justificación.....	6
2.7 Alcance y delimitaciones.....	7
Estado del arte.....	8
Marco teórico.....	12
4.1 Proceso de soldadura por fricción – agitación.....	12
4.1.1 Inspección y control de calidad en FSW.....	14
4.1.2 Pruebas no destructivas.....	14
4.1.3 Inspección visual.....	15
4.2 Modelos lineales generalizados.....	17
4.2.1 Escala de medición y notación.....	17
4.2.2 Distribucion chi – cuadrada y algunos resultados.....	19
4.2.3 Distribuciones de la familia exponencial.....	21
4.2.4 Propiedades de las distribuciones.....	23
4.2.5 Función Log Verosimilitud: Valor esperado y varianza de sus derivadas.....	23
4.2.6 Función score o estadística score.....	24
4.3 Modelo lineal generalizado.....	24
4.3.1 Estimación de los parámetros por máxima verosimilitud.....	26
4.4 Inferencia.....	27
4.4.1 Estadístico score para caso binomial.....	28
4.4.2 Aproximaciones por series de Taylor.....	29
4.4.3 Log likelihood ratio.....	29
4.4.4 Devianza para un modelo binomial.....	30
4.5 Pruebas de hipótesis.....	31

4.5.1 Prueba individual para los coeficientes del modelo.....	31
4.5.2 Devianza.....	31
4.6 Construcción del modelo lineal generalizado.....	32
4.6.1 Variables Binarias.....	32
4.6.2 Modelo logístico.....	34
4.6.3 Regresión logística multinomial.....	37
4.6.5 Regresión logística nominal.....	39
4.7 Optimización multiobjetivo.....	40
4.7.1 Algoritmo Genético de clasificación No – Dominado (NSGA II).....	42
Aplicación.....	50
6.1 Identificar variables de proceso y defectos categóricos.....	50
6.2 Experimentación, observación y medición de defectos.....	51
6.3 Ajustar modelo: regresión multinomial.....	59
6.4 Comparar valor AIC y seleccionar modelos.....	60
6.5 Inferencia.....	62
Conclusiones.....	65

Capítulo 1

Introducción

En la industria aeronáutica existen materiales con altos estándares de calidad y buenas propiedades mecánicas, un ejemplo de ello son los compuestos de matriz metálica (por sus siglas en inglés MMC). En este sentido, es de interés adecuar el uso de éstos materiales en la industria automotriz, sin embargo se tiene una limitante, los procesos de soldadura convencionales utilizados en la rama automotriz provocan que las propiedades mecánicas de los MMC se pierdan debido al aporte de calor que generan. De tal manera, es de interés establecer un proceso adecuado dadas las características de dichos materiales; uno de éstos procesos es el de soldadura en estado sólido, en especial la soldadura por fricción agitación (FSW por sus siglas en inglés, friction stir welding), el cual reúne las características indicadas.

El proceso de FSW utiliza una herramienta cilíndrica giratoria con un pin en la punta, al comenzar el proceso la herramienta baja hasta llegar a la junta de los materiales a unir aplicando presión hasta cierta profundidad, una vez que penetra, la herramienta avanza a través de la junta de los materiales generando

una deformación plástica por el aumento de calor provocado por el giro de la herramienta, esto unirá los materiales.

Considerando lo anterior, es necesario estudiar el comportamiento de este proceso así como las características de las uniones para determinar si son capaces de cumplir con los requerimientos de calidad establecidos en la industria automotriz, es decir, que estén libres de defectos. Los defectos encontrados comúnmente en las uniones soldadas mediante FSW son: hueco, hoyo y exceso de rebaba.

Estas anomalías dependen de la configuración en los niveles de las variables de proceso involucradas, por ello es necesario explicar el comportamiento de los defectos a través de las variables de entrada involucrada mediante un modelo.

Con el fin de representar el comportamiento de los defectos, es común utilizar modelos de regresión. Una de las técnicas comúnmente utilizadas en el análisis y estudio de datos, es el análisis de regresión. Éstos modelos expresan la relación entre la variable de salida (respuesta) con las variables de entrada (explicativas). Considerando la naturaleza de los defectos, se propone utilizar modelos lineales generalizados, un modelo de regresión multinomial que es adecuado para representar múltiples respuestas categóricas.

Una vez establecido el modelo, se pretende encontrar y determinar la región de parámetros en donde se minimice la probabilidad de encontrar esos defectos, pero como son múltiples respuestas categóricas, es necesario utilizar métodos de optimización multiobjetivo que tienen la tarea de encontrar más de una solución o un conjunto de soluciones óptimas con respecto a un objetivo, para este caso el objetivo es encontrar la región de parámetros donde la probabilidad de que se presenten los defectos simultáneamente, sea mínima.

Capítulo 2

Planteamiento del problema

2.1 Descripción del problema

Los procesos de soldadura convencionales o de fusión no son adecuados para unir los materiales MMC debido a que el calor producido modifica sus características y por lo tanto afectan sus propiedades estructurales. Es por eso que se propone el uso de la soldadura por fricción agitación (FSW) que es un proceso de soldadura en estado sólido compatible con las características de dichos materiales.

FSW proporciona ventajas significativas sobre los métodos convencionales:

- Presenta menos defectos.
- Alta retención de propiedades mecánicas y menor distorsión,
- Suelda aleaciones que son difíciles o imposibles de unir utilizando técnicas convencionales (como soldadura por fusión).

La presencia de defectos en soldadura, provocan fallas e inestabilidad, limitan las propiedades mecánicas y reducen la vida útil en las uniones soldadas. En FSW comúnmente se presentan los siguientes: exceso de rebaba, hoyo y hueco (AWS D17.3 - 2010):

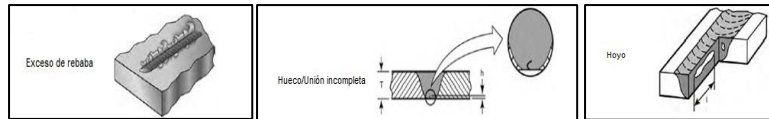


Figura 2.1 Defectos comunes en FSW – Exceso de Rebaba / Hueco / Hoyo de Gusano, respectivamente

Por ello, es de interés comprender su comportamiento y modelarlo a través de las variables de entrada involucradas en el proceso.

Para modelar este tipo de comportamiento, es necesario recurrir a modelos de regresión más generales por el tipo de respuesta o respuestas que suelen presentarse en procesos con más de una característica de calidad a cumplir (una o más variables discretas con solo dos posibles respuestas), los modelos tradicionales son incapaces de representar adecuadamente el proceso con este tipo de respuestas, es por eso que se opta por utilizar otra herramienta estadística acorde con la naturaleza de los defectos.

Para más de una respuesta del tipo binomial, la técnica de regresión logística multinomial resulta apropiada en la comprensión y modelación del comportamiento del proceso, esto con el fin de explicar las características de calidad a través de las variables de entrada y así obtener los niveles de las variables explicativas donde se encuentre la mayor probabilidad de ausencia de los defectos en las uniones soldadas del proceso de soldadura por fricción – agitación.

2.2 Preguntas de investigación

- ¿Es posible establecer un modelo capaz de indicar con certeza estadística las variables significativas de proceso considerando múltiples respuestas categóricas?

- ¿Puede modelarse la relación de las variables de proceso con las respuestas mediante una distribución multinomial dado que se consideran múltiples respuestas categóricas?
- ¿Qué supuestos se deben considerar en el modelo multinomial para modelar el proceso de FSW?
- ¿Es posible encontrar la región de parámetros en donde las uniones soldadas por FSW estén libres de defectos comunes?
- ¿Existe algún método de optimización adecuado dadas las respuestas categóricas múltiples?

2.3 Hipótesis

Es posible obtener uniones de aluminio MMC en el proceso de FSW, libres de hoyos, huecos y sin exceso de rebaba mediante el uso de un modelo multinomial generalizado en conjunto con un método de optimización multi objetivo.

2.4 Objetivo general

Encontrar los niveles de los parámetros del proceso FSW donde las uniones de aluminio de MMC estén libres de defectos, lo anterior mediante un modelo de regresión multinomial y un método de optimización multiobjetivo adecuado dadas las características del modelo.

2.5 Objetivos específicos

- Definir las variables involucradas en FSW.
- Identificar los defectos en FSW y cómo identificarlos.
- Determinar y estudiar las pruebas para medir los defectos.

- Experimentar con el proceso de FSW considerando los múltiples defectos categóricos.
- Estudiar la regresión logística, regresión multinomial y su inferencia.
- Estudiar los métodos de estimación de parámetros para la regresión multinomial.
- Establecer la región óptima de niveles donde las uniones soldadas estarán libres de defectos.

2.6 Justificación

Actualmente, la industria automotriz busca principalmente el uso de mejores materiales debido a los requerimientos exigidos por el cliente con el fin de mejorar los estándares de calidad de sus productos. Esto se logra mediante la búsqueda de materiales más resistentes y de menor espesor a los usados tradicionalmente.

Es de interés la implementación de materiales utilizados en la aeronáutica dentro de la industria automotriz debido a que sus estándares de calidad son más elevados, al igual que sus características de calidad.

Uno de estos materiales es el compuesto de matriz metálica (MMC), los cuales son una clase de materiales con excelentes propiedades y altos estándares de calidad que los hace muy atractivos para una amplia variedad de aplicaciones. Adicionalmente, se han identificado propiedades mecánicas interesantes como la resistencia al desgaste, la resistencia mecánica, estabilidad química, estabilidad dimensional en temperaturas elevadas, la capacidad de soportar cargas pesadas y la eliminación de peso excedente en productos.

Para lograr que las características de los MMC estén presentes en las uniones soldadas, es necesario que la soldadura se encuentre libre de defectos. Por esto, es determinante modelar el comportamiento de los defectos a través de las variables de entrada y obtener los niveles que maximicen la probabilidad de ausencia de dichos defectos. La presencia de defectos en las uniones afectan significativamente en la resistencia de la unión soldada, en la presión

interna (ya que genera grietas, fisuras), reduce significativamente su vida útil, disminución de propiedades mecánicas

Un ejemplo claro es en la estructura de un producto automotriz ya que éste debe soportar diferentes cargas, lo cual se traduce en fatiga. Las superficies de dicha estructura deben estar libres de aquellas discontinuidades que afectan su resistencia. Si el metal que rodea la unión soldada no es lo suficientemente dúctil, la presión interna genera fisuras en la soldadura, por consecuencia fallas y rupturas tempranas en la estructura, debido a los defectos que se puedan presentar. En este sentido, es posible denotar la importancia de la predicción del éxito en las uniones soldadas y evitar la presencia de defectos comunes.

2.7 Alcance y delimitaciones

El proyecto está dividido en:

- 1- La parte de Soldadura por Fricción Agitación
- 2- Modelado del Proceso

La presente investigación se enfoca únicamente en la recepción de datos de proceso así como en la modelación, proporcionar los niveles de los parámetros para asegurar una probabilidad alta de éxito en las uniones soldadas. Además de obtener la significancia de los parámetros de modelo y la correcta elección del modelo que representa adecuadamente el comportamiento del proceso. Adicionalmente, el modelo está enfocado específicamente para el aluminio A436 y MMC, no se considera un análisis de costos y el modelo solo funciona para múltiples respuestas categóricas.

Capítulo 3

Estado del arte

Meker y Hahn (1977), mencionan cómo estimar los parámetros u obtenerlos a partir de experimentos previos. Otro camino es a partir de la experiencia del investigador. A estas aproximaciones se les denomina estimaciones iniciales o aproximaciones iniciales sin importar cómo se hayan obtenido.

Hintze (2007), demuestra que el test de razón de verosimilitud es una prueba adecuada de elección en los modelos de regresión logística. En la razón de verosimilitud, cuando el modelo completo es el modelo saturado, likelihood ratio es análoga a la suma residual de cuadrados en regresión múltiple. De hecho, cuando la desviación se calcula en múltiples regresiones, es igual a la suma del cuadrado del residuo.

González et al, (2010), comprueban que la estadística de Wald disminuye a cero cuando la distancia entre la estimación del parámetro y el valor verdadero crece, por lo que el poder de predicción disminuye.

Reza (2014) emplea la regresión logística y propone un modelo para la clasificación de los tipos de alteración (argílicos y propilítica). El modelo final clasifica con éxito las muestras en la formación de datos (90,50% de una clasificación correcta), además incluye una constante y 9 variables explicativas. El estadístico Wald sugirió que las variables seleccionadas son significativas y se probó la significancia del modelo a través de la prueba de chi-cuadrado y pruebas de Hosmer y Lemeshow (todos en el 5% de nivel de significancia).

Urbano et al., (2014) menciona que mediante superficie de respuesta se pueden proporcionar los niveles de la velocidad de rotación y avance que producen una mejor resistencia a la tensión en las uniones soldadas a través de FSW.

Pando (2004) explica la estimación de los parámetros mediante el método de máxima verosimilitud y se establecen los test estadísticos adecuados del modelo así como el efecto de cada regresor para regresión logística.

Dawes y Thomas (1996) mencionan que en la industria automotriz como en la aeroespacial, el proceso de soldadura por fricción agitación tiene muchas ventajas, entre ésta que se conservan las propiedades mecánicas (tensión y fatiga) del material soldado. Además, los autores comprueban que trae mejoras en los procesos, falta de insumos, problemas de salud y medio ambiente, y ventajas operativas.

Karthikeyan et al., (2010) muestra que en FSW, los diferentes parámetros tales como la velocidad de rotación, velocidad de soldadura, la fuerza axial, geometría de la herramienta, material de la herramienta, el ángulo de título de la herramienta, la fuerza de fijación y la geometría tienen un efecto importante en la MMC, pero los cuatro primeros parámetros juegan un papel significativo en la soldadura.

En la norma AWS D17.3 – 2010 se menciona que las variables más importantes en la soldadura por fricción – agitación son: velocidad de rotación,

velocidad de soldadura y presión aplicada sobre la herramienta. Éstas aportan características de manera significativa a la soldadura.

P. Podržaj, et al, (2014) mencionan los defectos que se presentan en la soldadura por fricción – agitación (FSW): falta de penetración, falta de fusión, huecos, ranuras superficiales, rebaba, excoiación, porosidad, hoyos de gusano.

Loes et al., (2015) presentan un modelo multinomial para representar la importancia relativa de cada predictor y calcular la probabilidad para cada categoría de enfermedad en un paciente.

García et al, (2008) proponen un modelo usando regresión multinomial para reconocer los factores que influyen en la productividad en los sectores industriales. Mediante el modelo de Regresión Multinomial, se determinó que la probabilidad de pertenecer al sector químico aumenta a medida que aumenta el porcentaje de materia prima importada, y disminuye la de pertenecer a los sectores metalmeccánico y plástico.

Pando y San Martín (2004) presentan la regresión logística multinomial como extensión multivariante de la regresión logística binaria, ampliamente utilizada en el ámbito forestal. A partir de la formulación matemática del modelo estadístico se explica la estimación de los parámetros mediante el método de máxima verosimilitud y se establecen los test estadísticos adecuados para la significancia global del modelo y de cada regresor.

Jiang et al., (2016) utilizan NSGA II para buscar múltiples soluciones óptimas con frente de Pareto para el proceso de soldadura por láser. Se analizan los efectos principales y de los parámetros del proceso en el perfil de cordón de soldadura. Los resultados de los experimentos de verificación indican que los parámetros de proceso óptimos son eficaces y fiables para la producción.

Moradpour et al., (2015) proponen una aproximación basada en lógica difusa y el algoritmo genético de clasificación no dominada II (NSGA-II) además de un nuevo enfoque para predicción de los parámetros del proceso de soldadura. Los parámetros del proceso se optimizan para alcanzar los valores deseados de

convexidad y de penetración simultáneamente utilizando el algoritmo NSGA-II. Como resultado se obtiene un conjunto de vectores óptimos.

Tutum et al., (2010) investigan los parámetros del proceso y geometrías de herramienta óptima en Friction Stir Welding (FSW) para minimizar la diferencia de temperatura entre el borde delantero de la sonda de la herramienta y el material de la pieza de trabajo en la parte delantera del hombro de la herramienta. Al mismo tiempo se busca maximizar la velocidad de soldadura, lo que entra en conflicto con el primer objetivo. Utilizaron un método de optimización multi-objetivo como el NSGA-II para encontrar las soluciones óptimas, seguido de una búsqueda local basada en el gradiente SQP para mejorar la convergencia del frente de Pareto.

Capítulo 4

Marco teórico

4.1 Proceso de soldadura por fricción – agitación

Es un proceso de unión de dos piezas en estado sólido. Cuenta con una herramienta cilíndrica con un pin en su extremo, el cual se introduce en la junta entre las dos piezas que se van a soldar. La herramienta adquiere velocidad y aumenta la temperatura en la zona donde penetra debido a la fricción, dicha herramienta pasa a lo largo de la junta desplazando el material gracias a la rotación.

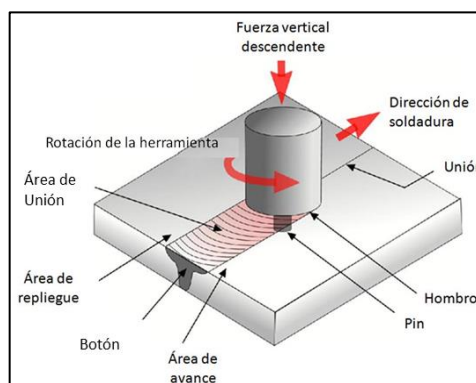


Figura 4.1: Soldadura por Fricción Agitación

Lohwasser y Chen (2010) mencionan las principales características de las mejoras que atrae este tipo de soldaduras en comparación otros procesos:

Mejora la soldabilidad. Algunas aleaciones de aluminio presentan dificultades al soldar por medio de fusión. Para estos materiales, el proceso de FSW es el más adecuado ya que elimina las limitantes que la soldadura por fusión provoca.

Mejora de la fatiga, corrosión y corrosión bajo tensión. Las uniones generadas por FSW tienen resistencia a la fatiga (mejor que los métodos de fusión). Y debido a la baja temperatura con la que se sueldan las uniones, mejoran la corrosión y la corrosión bajo tensión en las aleaciones de aluminio.

Mejora la apariencia estética. En la unión de materiales, la superficie es extremadamente lisa y plana.

La eliminación de metal de relleno. No necesita de un material extra para la unión, lo cual ayuda en la reducción de costos.

Mejora la resistencia estática y ductilidad. En aleaciones con un espesor de hasta 75 mm ha dado lugar a una mejora significativa en la fuerza y ductilidad en comparación con soldadura por arco.

Es un proceso mecanizado. Puede llegar a tener un costo mayor que soldadura por fusión, sin embargo este tipo de trabajo mecanizado ayuda a una operación con variabilidad controlable de fuerza para materiales finos y muy gruesos, generando una calidad de unión muy buena con los materiales que por fusión es imposible.

Tiene requisitos especiales. Requiere una pieza rígida de trabajo y en una sola posición al soldar para lograr una soldadura lisa.

Tiene limitaciones de diseño. Se utiliza para producir soldadura a tope, de esquina y de vuelta.

Su justificación económica es el tiempo de procesamiento, no necesita de licencias o permisos para su uso, capital invertido, robustez (tasa baja de defectos).

4.1.1 Inspección y control de calidad en FSW

Los principales motivos para elegir o inclinarse por el uso de FSW sobre los procesos de soldadura convencional, son: calidad de la soldadura, eficiencia del proceso, ahorro en costos y la eliminación de defectos en las uniones.

Para los requerimientos de la calidad de soldadura en FSW, se tienen 3 fases:

1. Evaluación de la Viabilidad del uso de FSW
2. Desarrollo y Calificación del Proceso: Se desarrolla la herramienta y el equipo, se ajusta el programa y se preparan las piezas con respecto a la herramienta. Se verifica que cada parte del equipo cumpla las características necesarias para este proceso.
3. Implementación para la Producción: Una vez probada cada parte del equipo de éste proceso, pasa a su aplicación industrial y se observa el resultado del proceso.

Se pretende controlar las variables de entrada para el proceso de FSW, las cuales son: presión, velocidad de rotación, avance de soldadura, ángulo de herramienta, entre otras... Con el objetivo de comprender el comportamiento de alguna característica de calidad o salida elegida.

En ésta parte es importante realizar un buen diseño experimental (DOE) ya que el resultado de éste proporcionará la confianza de que la soldadura resultante será significativa en la obtención de los requisitos de calidad de soldadura utilizando las variables significativas y los parámetros adecuados (Lohwasser y Chen, 2010).

4.1.2 Pruebas no destructivas

No solamente es importante realizar el DOE, también es necesario realizar ensayos no destructivos en la parte soldada debido a que pueden existir defectos

que a simple vista no se pueden detectar como: falta de penetración, huecos, falta de unión, defectos en la raíz de soldadura (Lohwasser y Chen, 2010)

4.1.3 Inspección visual

La más simple y primera inspección realizada a la soldadura. Algunos defectos como exceso de flash y falta de penetración se pueden ver fácilmente. Gracias a éste tipo de inspección es posible observar, sin necesidad de análisis más detallados, que los parámetros de la soldadura necesitan mejorarse y corregirse para obtener mejor calidad al soldar (AWS D17.3, 2010).

Existen otros tipos de inspecciones de calidad con un equipo específico y con diferentes funciones para analizar defectos diversos:

- **Radiografía.** Se utiliza en el examen de piezas fundidas y soldadas, sobre todo cuando hay la necesidad de asegurar que no existen defectos internos, como falta de fusión, falta de penetración, corrosión, huevos, grietas, inclusiones, etc.
- **Líquidos penetrantes.** Se utiliza para localizar defectos de superficie en los materiales no porosos (metales, plásticos o cerámicos). Sirve para detectar defectos de fundición y forja, grietas y fugas en las uniones y grietas de fatiga. El revelador utilizado en este análisis ayuda a hacer visible el defecto al inspector ya que se realiza bajo luz ultravioleta.
- **Ultrasonido.** Sirve para detectar defectos, mediciones dimensionales y caracterización de materiales, entre otros. Mediante la emisión y recepción de impulsos o señales, es posible detectar los defectos.

Para este caso de estudio, se utilizará la primera de las inspecciones en las uniones soldadas con el motivo de detectar, con ayuda de los expertos en este tipo de inspección, que las piezas unidas mediante FSW muestren ausencia o presencia de defectos, debido también a que es una de las inspecciones que

rápidamente nos proporciona una decisión en cuanto a si los niveles de los parámetros están siendo correctos o están muy alejados de la calidad deseada (Lohwasser y Chen, 2010).

En la aeronáutica, son utilizados materiales con estándares de calidad elevados y cumplen con todas y cada una de las características de calidad requeridas para la manufactura de sus productos. Además de dichas exigencias, satisfacen la necesidad de los clientes que requieren de productos terminados con una alta calidad de fabricación y de características. De aquí la importancia de fabricar un producto sin defectos.

En los últimos 20 años, se han estudiado un grupo de materiales con características peculiares, éstos materiales son los llamados compuestos de matriz metálica. Dichas características son las propiedades que tienen ya que son difíciles de conseguir con un material monolítico, por tal motivo se desarrollaron este tipo de materiales compuestos principalmente para componentes aeroespaciales y de motores de automoción. Están constituidos por una matriz y un refuerzo, y sus propiedades van a depender del refuerzo seleccionado así como de la matriz (Lohwasser y Chen, 2010).

Estos materiales cuentan con una variedad de características destacables, las cuales son:

- Soporte de altas tensiones ejercidas en el compuesto.
- Aumento de las características mecánicas (como dureza y resistencia al desgaste).
- En altas temperaturas, mitigar los fallos de estas características.
- Detener la propagación de grietas.
- Más ligero.
- El proceso de soldadura con el que se unen este tipo de materiales (FSW) es amable con el ambiente.

Considerando lo anterior, es de interés predecir el comportamiento del proceso de soldadura por fricción – agitación para llevar a cabo la unión exitosa de piezas soldadas y así obtener la mayor probabilidad de éxito en este proceso. Así mismo, producir la mayor cantidad de piezas sin defectos y a la par, piezas con este tipo de características distintivas de los MMC.

4.2 Modelos lineales generalizados

En este apartado se presenta la familia exponencial con distribuciones tales como Normal, Poisson, Binomial, en términos de funciones lineales de tal manera que se crea un “**enlace**” mediante una relación algebraica.

El Modelo Lineal Generalizado nace cuando las variables de “**Y**” y “**X**” no están relacionadas de una manera directa y utilizando la familia exponencial se crea una función de “enlace”, la cual permite métodos como mínimos cuadrados y máxima verosimilitud para estimar sus parámetros. En este caso las ecuaciones no tienen solución explícita, sino una solución implícita lo que hace que se necesite un método numérico. Existen algunos métodos que permiten resolver esta situación entre los cuales se encuentran el método de Newton – Raphson, entre otros.

4.2.1 Escala de medición y notación

Para el tipo de variables con las que se trabajará, existen diferentes tipos de medición las cuales son las siguientes (McCullagh y Nelder; 1989):

- Nominal: Respuestas Binarias
 - o Existen solo 2 categorías
 - o Para más de 2 categorías se llama “multinomiales”
- Ordinal: Clasificaciones
 - o Donde existe un orden natural

- Ejemplo: Joven, edad media, viejo; Grupos ≤ 70 , $71 - 90$, $91 \geq 130$.
- Continua: Incluye escalas de intervalos y proporciones de intervalo.

Los elementos de un modelo lineal generalizado son:

- Una variable explicatoria “cualitativa” es llamada “factor” y sus categorías “niveles” (discretas).
- Una variable explicatoria “cuantitativa” es llamada covariable.
- La variable de respuesta “ y ” es aleatoria, las observaciones en forma matricial se pueden representar de la forma:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

- La ecuación que enlaza “ y ” con “ x_1, x_2, \dots, x_n ” tiene la forma:

$$g[E(y)] = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n \quad (1)$$

que en notación matricial es de la forma:

$$g[E(y)] = X\beta$$

Donde:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} : \text{Vector de respuestas.}$$

$X\beta$: Componente lineal.

$$g[E(y)] = \begin{bmatrix} g[E(y_1)] \\ g[E(y_2)] \\ \vdots \\ g[E(y_n)] \end{bmatrix}$$

Vector de funciones en términos de $E(y)$ con función “ g ”

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} : \text{Vector de parámetros donde } \hat{\beta} \text{ es el estimador del parámetro } \beta.$$

X : Matriz de elementos constantes de las variables explicativas.

Algunas características de estos modelos:

- Existen “ p ” parámetros en el modelo y “ N ” observaciones.
- “ y ” es un vector de $N \times 1$, β es de “ $p \times 1$ ” y “ x ” es de “ $n \times p$ ”
- La transpuesta de una matriz se denota por “ T ”, por ejemplo:

$$y = [y_1, y_2, \dots, y_n]^T$$

- La función de densidad, es referida como función de probabilidad y se denota $f(x; \theta)$, donde θ representa el vector de parámetros.
- En ocasiones se utiliza “ \cdot ” para denotar la suma:

$$\bar{y} = \frac{1}{N} \sum y_i = \frac{1}{N} y$$

- El valor esperado y la varianza para la variable aleatoria “ y ” se denotan como:

$$E(Y), Var(Y)$$

- Si W representa una combinación lineal de las y_i 's.

$$W = a_1 y_1 + a_2 y_2 + \dots + a_n y_n \quad (2)$$

4.2.2. Distribucion chi – cuadrada y algunos resultados

La distribución chi – cuadrada “central” se define de la siguiente manera:

$$x^2 = \sum_{i=1}^n Z_i^2 \sim \chi^2_{(n)} \quad ; \quad Z \sim N(0, 1) \quad (3)$$

En notación matricial:

$$z = [z_1, \dots, z_n]^T \quad (4)$$

Entonces:

$$z^T z = \sum_{i=1}^n z_i^2$$

$$x^2 = z^T z \sim \chi^2_{(n)} \quad (5)$$

Nota: La forma cuadrática también se distribuye como $\chi^2_{(n)}$. El valor esperado se define:

$$E[x^2] = n \quad (6)$$

Si $y_i \sim N(\mu_i, \sigma^2)$, entonces "Z" puede escribirse como:

$$x^2 = \sum_{i=1}^n z_i^2 = \sum_{i=1}^n \left(\frac{y_i - \mu_i}{\sigma_i} \right)^2 \sim \chi^2_{(n)} \quad (7)$$

Sea z_1, \dots, z_n variables aleatorias independientes con distribución $N(0, 1)$ y sea $y_i = z_i + \mu_i$ con al menos una $\mu_i \neq 0$:

$$\sum_{i=1}^n y_i^2 = \sum_{i=1}^n (z_i - \mu_i)^2 = \sum_{i=1}^n z_i^2 + 2 \sum_{i=1}^n z_i \mu_i + \sum_{i=1}^n \mu_i^2 \quad (8)$$

La distribución de $\sum_{i=1}^n y_i^2$ tiene media mayor a $(n + \lambda)$ y varianza mayor a $(2n + 4\lambda)$ en comparación con una $\chi^2_{(n)}$, donde:

$$\lambda = \sum \mu_i^2 \quad (9)$$

Se distribuye $z_i \sim N(0, 1)$

$$= \sum_{i=1}^n z_i^2 + 2 \sum_{i=1}^n z_i \mu_i + \sum_{i=1}^n \mu_i^2 \quad (10)$$

$$= \lambda + N \left(2 \sum_{i=1}^n \mu_i, 4 \sum_{i=1}^n \mu_i^2 \right) + \chi^2_{(n)} \quad (11)$$

$$E[\cdot] = E[\lambda] + E[\chi^2_{(n)}] + E[2] E \left[\sum_{i=1}^n \mu_i \right] E[z_i] (2) \left(\sum_{i=1}^n \mu_i \right) (0) \quad (12)$$

Donde:

$$E\left[\chi^2_{(n)}\right] = n$$

$$E[z_i] = 0$$

La varianza sería:

$$Var(\cdot) = Var(\lambda) + Var\left[N\left(2\sum_{i=1}^n \mu_i, 4\sum_{i=1}^n \mu_i^2\right)\right] + Var\left(\chi^2_{(n)}\right) \quad (13)$$

$$= 4\sum_{i=1}^n \mu_i^2 + 2n = 4\lambda + 2n \quad (14)$$

$$\sum y_i^2 \sim \chi^2(n, \lambda) \quad (15)$$

Es posible obtener la propiedad reproductiva de chi – cuadrada (McCullagh y Nelder, 1989) cuando x_1^2, \dots, x_m^2 son “m” variables aleatorias independientes con $\chi^2_i \sim \chi^2_{(n_i, \lambda_i)}$, se tienen $\sum n_i$ grados de libertad en la suma con una distribución chi – cuadrada con $\sum \lambda_i$ parámetros de no centralidad (**véase demostración 1, apéndice A**):

$$\therefore \sum x_i^2 \sim \chi^2_{(\sum n_i, \sum \lambda_i)} \quad (16)$$

4.2.3 Distribuciones de la familia exponencial

Si Y la consideramos como una variable aleatoria donde su distribución depende de un solo parámetro θ , pertenece a la familia exponencial si es posible escribirla de la siguiente forma (McCullagh y Nelder, 1989):

$$F(y; \theta) = S(y) + (\theta)e^{a(y)b(\theta)}$$

Son funciones conocidas las literales a, b, s, t. Es posible reescribir la función de la siguiente manera (Tattar, 2016):

$$F(y; \theta) = \exp < a(y)b(\theta) + c(\theta) + d(y) > \quad (17)$$

Los elementos de la forma de la familia exponencial son:

- Si $a(y) = y$, la distribución se llama “estándar canónica”
- $b(\theta)$ es llamado “parámetro natural” de la distribución; que nos proporciona en sí el “enlace” que se utilizará más adelante; especifica los parámetros necesarios para dicha distribución.
- Si existen más parámetros de la distribución se tratan como constantes.
- Las distribuciones pertenecientes a esta familia son: Normal, Binomial, Poisson, Exponencial, entre otras.

Por interés, se mostrarán resultados únicamente para la distribución binomial debido a que las respuestas a estudiar son solamente de 2 posibilidades (como se vio en el planteamiento del problema).

En la distribución binomial, como miembro de la familia exponencial, consideramos la variable aleatoria Bernoulli. Utilizando su función de probabilidad (**véase Apéndice A, demostración 2**):

$$f(y; \pi) = \exp < y \ln\left(\frac{\pi}{1-\pi}\right) + n \ln(1-\pi) + \ln\binom{n}{p} > \quad (18)$$

Entonces, utilizando la forma de la familia exponencial:

$$a(y) = y : \textit{Forma Canónica}$$

$$b(\theta) = \ln\left(\frac{\pi}{1-\pi}\right) : \textit{Parámetro Natural}$$

$$c(\theta) = n \ln(1-\pi)$$

$$d(y) = \ln\binom{n}{p}$$

4.2.4 Propiedades de las distribuciones

Se pretende explicar y obtener algunas de las propiedades de esta familia de distribuciones a partir de su forma general, para obtener el valor esperado de la forma canónica $a(y)$ así como su varianza, $E[a(y)]$ y $Var[a(y)]$ de la función de densidad (McCullagh y Nelder, 1989), el valor esperado para la familia exponencial es:

$$E[a(y)] = -\frac{c'(\theta)}{b'(\theta)} \quad (19)$$

Para definir la varianza de la familia exponencial (Shumskii, 2005), utilizamos la segunda derivada obtenida de lo anterior (**Véase apéndice A, demostración 3**), llegamos a:

$$VAR[a(y)] = \frac{b''(\theta) (c'(\theta)) - b'(\theta)c''(\theta)}{[b'(\theta)]^3} \quad (20)$$

Así es como se obtiene el valor esperado y la varianza para las distribuciones pertenecientes a la familia exponencial. Utilizando la forma general del valor esperado para la familia exponencial, en el caso de la binomial, el valor esperado es:

$$E[a(y)] = -\frac{c'(\pi)}{b'(\pi)} \quad (21)$$

$$E[a(y)] = n\pi \quad (22)$$

A partir de la forma general de la varianza, obtenemos para la distribución binomial:

$$Var[a(y)] = n\pi(1 - \pi) \quad (23)$$

4.2.5 Función Log Verosimilitud: Valor esperado y varianza de sus derivadas.

Partiendo de la forma de la familia de la función exponencial, obtenemos la función de log – verosimilitud para cada una de las distribuciones:

$$\ln f(y; \theta) = a(y)b(\theta) + c(\theta) + d(y) \quad (24)$$

Es de importancia obtener la derivada de la función log – verosimilitud para determinar el valor esperado y la varianza de ésta; entonces la derivada de $\ln(\theta; y)$ con respecto a θ y para todas las distribuciones de la familia es:

$$U(y; \theta) = a(y)b'(\theta) + c'(\theta) \quad (25)$$

4.2.6 Función score o estadística score

Una vez obtenida la función de log – verosimilitud, podemos determinar el valor esperado de la función score así como su varianza (McCullagh y Nelder; 1989):

La función Score tiene la siguiente forma:

$$U = a(y)b'(\theta) + c'(\theta) \quad (26)$$

El valor esperado para la función score es:

$$E[U] = 0 \quad (27)$$

Varianza de la función score es:

$$Var[U] = E[U^2] \quad (28)$$

(Ver demostraciones 4 y 5, respectivamente; apéndice A).

4.3 Modelo lineal generalizado

Se define en términos de variables aleatorias independientes y_1, y_2, \dots, y_n , cada una con distribución de la familia exponencial.

Propiedades:

- La distribución de y_i tiene la forma canónica y depende de un solo parámetro θ_i :

$$f(y; \theta) = \exp \langle y_i b_i(\theta_i) + c_i(\theta_i) + d_i(y_i) \rangle \quad (29)$$

- Si las respuestas tienen la misma distribución, la densidad conjunta será:

$$f(y_1, \dots, y_n; \theta_1, \dots, \theta_n) = \prod_{i=1}^n \exp \langle y_i b_i(\theta_i) + c_i(\theta_i) + d_i(y_i) \rangle \quad (30)$$

$$\exp \langle \sum_{i=1}^n y_i b_i(\theta_i) + \sum_{i=1}^n c_i(\theta_i) + \sum_{i=1}^n d_i(y_i) \rangle \quad (31)$$

Usualmente es de interés un conjunto pequeño de parámetros, $p < n$. Suponiendo que $E[y_i] = \mu_i$, donde μ_i es una función de θ_i . Para un modelo lineal generalizado existe una transformación de μ_i tal que:

$$g(\mu_i) = X^T \beta \quad (32)$$

Donde:

- 1) G: Es una función monótona diferenciable llamada "función de enlace".
- 2) El vector x_i son variables explicatorias de dimensiones $p \times 1$ (matriz de diseño):

$$x_i = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{ip} \end{bmatrix}; \quad x_i^T = [x_{i1} \quad \dots \quad x_{ip}]$$

- 3) β es el vector de parámetros de dimensión $p \times 1$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

Funciones de enlace para algunas distribuciones de la familia exponencial (McCullagh y Nelder; 1989):

$$\text{Normal} : \eta = \mu$$

$$\text{Binomial} : \eta = \log\left(\frac{\pi}{1 - \pi}\right)$$

$$\text{Gamma} : \eta = \mu^{-1}$$

4.3.1 Estimación de los parámetros por máxima verosimilitud

Es de interés estimar los parámetros β , los cuales se relacionan con las y_{is} a través de $E[y_i] = \mu_i$ y $g(\mu_i) = X^T \beta$. La función de **log – verosimilitud** es:

$$l_i = y_i b_i(\theta_i) + c_i(\theta_i) + d_i(y_i)$$

Además, al conocer tanto el valor esperado como la varianza de la familia de distribución exponencial, puede asumirse lo siguiente:

$$g(\mu_i) = x_i^T \beta ; \beta = \eta_i$$

La función de **log – verosimilitud** para todas las y_{is} es:

$$l = \sum_{i=1}^n l_i = \sum_{i=1}^n y_i b_i(\theta_i) + \sum_{i=1}^n c_i(\theta_i) + \sum_{i=1}^n d_i(y_i)$$

Obtenemos la primera derivada **“Score”**:

$$U_j = \sum_{i=1}^n \left[\frac{[y_i - \mu_i]}{\text{Var}(y_i)} x_{ij} \left(\frac{d\mu_i}{d\eta_i} \right) \right] \quad (33)$$

El valor esperado y la varianza de U_i es:

$$E[U(\theta)] = 0 \quad (34)$$

$$\text{Var}[U_j(\theta)] = E[U(\theta)U'(\theta)] = I(\theta) \quad (35)$$

Bajo condiciones de regularidad, la matriz de información observada es:

$$I(\theta) = -E \left[\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right] \quad (36)$$

Por lo tanto, para calcular la matriz de segundas derivadas quedaría de la siguiente forma (**demostración 6 en apéndice A**):

$$J_{jk} = \sum_{i=1}^n \left[\frac{x_{ij} x_{ik}}{\text{Var}(y_i)^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2 \right] \quad (37)$$

Utilizando el método de “estimación de Scoring” para esta matriz y obtener los parámetros, partimos de lo siguiente:

$$\theta^m = \theta^{m-1} + \frac{U^{m-1}}{J^{m-1}} \quad (38)$$

y llegamos a la ecuación:

$$x^T W x b^m = x^T W Z \quad (39)$$

Es la ecuación de los “mínimos cuadrados ponderados iterativos”. Despejando b^m de la ecuación (32b) para obtener el valor de los parámetros **(véase la demostración 7, apéndice A; para observar el desarrollo)** queda:

$$b^m = (x^T W x)^{-1} x^T W Z \quad (40)$$

4.4 Inferencia

En este apartado, partiremos de algunos supuestos para realizar a) pruebas de hipótesis y b) Intervalos de confianza. Bajo condiciones de regularidad, si “S” es un estadístico de interés, entonces:

$$\frac{SE[S]}{\sqrt{Var(S)}} \sim N(0,1) \quad (41)$$

Elevando al cuadrado:

$$\frac{(SE[S])^2}{Var(S)} \sim \chi^2_{(1)} \quad (42)$$

De forma matricial:

$$(S - E[S])^T V^{-1} (S - E[S]) \sim \chi^2_{(p)} \quad (43)$$

La distribución del estadístico score es:

$$U_j = \sum_{i=1}^n \left[\left(\frac{d\mu_i}{d\eta_i} \right) \frac{[y_i - \mu_i]}{Var(y_i)} x_{ij} \right] \quad (44)$$

$$E[U_j] = 0$$

La matriz de varianzas y covarianzas del estadístico score es:

$$J_{jk} = E[U_i U_k] \Rightarrow \text{Matriz de Información } J$$

Hay que notar que si existe un solo parámetro, se distribuye como:

$$\frac{U}{\sqrt{J}} \sim N(0,1) \text{ Distribución Asintótica}$$

Elevando al cuadrado será:

$$\frac{U^2}{J} \sim \chi^2_{(p)}$$

Para más de un parámetro:

$$U \sim MVN(0, J)$$

$$U^T J^{-1} U \sim \chi^2_{(p)}$$

4.4.1. Estadístico score para caso binomial

Para obtener el estadístico Score de la distribución binomial, partimos de:

$$Y \sim \text{Bin}(n, \pi)$$

Éste es el estadístico score para el caso binomial:

$$= \frac{y - n\pi}{\pi(1 - \pi)} \Rightarrow U \tag{45}$$

Valor esperado para el caso binomial:

$$E[U] = 0$$

Varianza para el caso binomial:

$$\text{Var}[U] = \frac{y - n\pi}{\sqrt{n\pi(1 - \pi)}} \sim N(0,1) \tag{46}$$

Elevando al cuadrado, quedaría de la siguiente forma. La expresión sigue una distribución chi – cuadrada (**véase demostración 9, apéndice A**):

$$\frac{(y - n\pi)^2}{n\pi(1 - \pi)} \sim \chi^2_{(1)}$$

4.4.2 Aproximaciones por series de Taylor

Mediante aproximaciones por series de Taylor, se puede determinar el estadístico de Wald, que se utiliza como variante de la prueba ANOVA para la significancia de los parámetros. Primeramente, hay que obtener la distribución asintótica de otros estadísticos. La aproximación $f(x)$ con una variable alrededor de t :

$$f(x) = f(t) + (x - t) \left[\frac{df}{dx} \right]_{x=t} + \frac{1}{2} (x - t)^2 \left[\frac{d^2f}{dx^2} \right]_{x=t} + \dots$$

Para el caso de la función de Log verosimilitud de un parámetro único:

$$l_{\beta} = l_{(b)} + (\beta - b)U(b) + \frac{1}{2} (\beta - b)^2 U'(b)$$

Llegamos entonces a la siguiente ecuación:

$$(b - \beta)^T J(b - \beta) \sim \chi^2_{(p)} \quad (47)$$

Lo anterior es el llamado “estadístico de Wald”, se utiliza como variante de la prueba ANOVA en la regresión tradicional para prueba de hipótesis de los parámetros, pero en este caso para modelos lineales generalizados (**véase demostración 10, apéndice A**). En el caso de un parámetro:

$$b \sim N(b, J^{-1})$$

$$\frac{b}{\sqrt{J^{-1}}} \sim N(0,1)$$

4.4.3 Log likelihood ratio

Se denomina “Estadístico de Razón de Verosimilitud” y lo componen:

m : Máximo número de parámetros

β_{\max} : Parámetros del modelo saturado.

b_{\max} : Estimador MLE de β_{\max}

$L(b_{\max}; y)$: Función de verosimilitud saturado evaluada en MLE β_{\max} .

$L(b; y)$: Para el modelo de interés, representa el valor máximo de dicha función.

$$\lambda = \frac{L(b_{\max}; y) \text{ Saturado}}{L(b; y) \text{ Interés}}$$

$$\log(\lambda) = \log(b_{\max}; y) - l(b; y) \quad (48)$$

El estadístico comúnmente utilizado para determinar qué tan bueno es el modelo, es llamado *Deviance* (D). Para valores grandes de éste estadístico, el modelo es pobre (**véase demostración 11, apéndice A**).

$$D = 2[l(b_{\max}; y) - l(b; y)] \quad (49)$$

Por lo tanto, la distribución del estadístico D (véase demostración 10, apéndice A) es:

$$D \sim \chi^2_{(p)}$$

En el caso de la Binomial y Poisson, “ D ” puede utilizarse como bondad de ajuste.

4.4.4 Devianza para un modelo binomial

Una vez determinada la distribución del estadístico de devianza en general, es necesario obtener la ecuación para determinar la devianza para el caso binomial:

$$D = 2[l(b_{\max}; y) - l(b; y)] \quad (50)$$

Partiendo de que tiene una distribución binomial $y \sim Bin(n_i, \pi_i)$ puede representarse de la siguiente forma:

$$D = \sum \left[(n_i - y_i) \log \left(\frac{n_i - y_i}{n_i - \hat{y}_i} \right) + y_i \log \left(\frac{y_i}{\hat{y}_i} \right) \right] \quad (51)$$

La ecuación (51) representa la función de la devianza para el modelo (véase demostración 12, apéndice A).

4.5 Pruebas de hipótesis

Utilizando el estadístico de Wald, es posible probar la hipótesis para el conjunto de parámetros β de la siguiente manera:

$$(\hat{\beta} - \beta)^T J(\hat{\beta} - \beta) \sim \chi_p^2 \quad : \quad \text{Prueba Global}$$

El estadístico score es:

$$U^T J^{-1} U \sim \chi_p^2 \quad : \quad \text{Distribución asintótica de una forma cuadrática}$$

4.5.1 Prueba individual para los coeficientes del modelo

$$\frac{(b - \beta_0)^2}{\text{Var}(b)} \sim \chi_1^2$$

Aplicando raíz:

$$\frac{(b - \beta_0)}{\sqrt{\text{Var}(b)}} \sim N(0,1)$$

4.5.2 Devianza

Si comparamos la bondad de ajuste de dos modelos: H_0 vs H_1 utilizamos la diferencia de los estadísticos de Devianza, y se representa de la siguiente forma:

$$\Delta D = D_0 - D_1$$

$$\Delta D = 2[l(b; y) - l(b_0; y)] \quad (52)$$

Si el modelo ajusta bien:

$$D_0 \sim \chi_{(N-q)}^2 \quad ; \quad D_1 \sim \chi_{(N-p)}^2 \quad \Rightarrow \quad \Delta D \sim \chi_{(p-q)}^2$$

Donde:

H_0 : μ_0 describe bien los datos (D_0){*Interés*}

H_1 : μ_1 describe bien los datos (D_1){*Saturado*}; todos los parámetros.

*** Si ΔD se encuentra en la región crítica, rechazar H_0 .

q : Cantidad de parámetros de interés.

p : Parámetros

En cuanto a la bondad de ajuste del modelo, se utiliza la devianza para indicar qué tan bueno es el modelo planteando la siguiente hipótesis nula:

H_0 : Se ajusta adecuadamente el modelo a los datos .

H_a : No se ajusta adecuadamente el modelo a los datos.

Utilizando el nivel de significación de alpha, la regla de decisión es la siguiente: Rechazar H_0 si la Devianza es significativa (esto es, está asociada a un valor $p < 0.05$) (Agresti, 2006; Agresti 2013).

4.6 Construcción del modelo lineal generalizado

4.6.1 Variables Binarias

Este tipo de variables se encuentran dentro de una escala 0 y 1, solo 2 posibles respuestas.

$$Z \begin{cases} 1 : \text{Si la respuesta es un éxito} \\ 0 : \text{Si la respuesta es un fracaso} \end{cases}$$

Además, se basan en un Experimento de Bernoulli, que es de la siguiente manera:

$$Z \sim Ber(\pi)$$

$$f(Z) = \pi^Z(1 - \pi)^{1-z} \quad (53)$$

Si usamos para “n” extracciones:

$$y = \sum_{i=1}^N Z_i$$

Colocamos las covariables en el parámetro π del modelo binomial. Lo siguiente es, pensar gráficamente en la función logística de crecimiento:

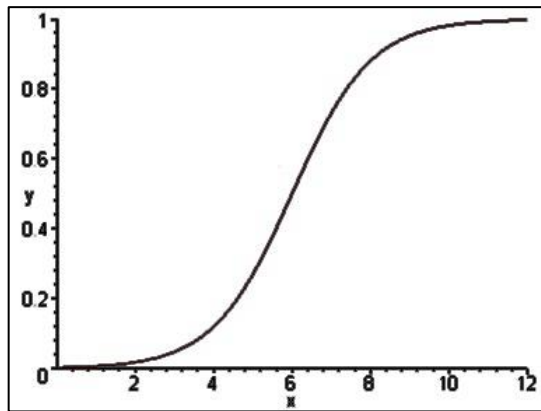


Figura 4.3 Gráfica S - Shape Función Logística

$$\Pr[Y = y] = \binom{n}{y} \pi^y (1 - \pi)^{n-y}$$

Es la probabilidad de que “**Y**” tome un valor “**y**” dado “**n**” cosas tomadas de “**y**” en “**y**”. Recordando capítulos anteriores sobre la varianza, valor esperado y el estimador de log verosimilitud, el objetivo en este apartado es describir la proporción de éxitos $p_i = \frac{y_i}{n_i}$ de cada subgrupo en términos de variables explicativas.

$$E[p_i] = E\left[\frac{y_i}{n_i}\right] = \frac{1}{n_i} E[y_i] = \frac{n_i \pi_i}{n_i} = \pi_i \quad (54)$$

¿Por qué (i)? Porque se hacen experimentos en cada subgrupo. Entonces se modelan las probabilidades π como:

$$g(\pi) = X\beta$$

$\pi \in [0,1]$, este intervalo es continuo porque está en término de probabilidad. Por lo tanto, se utiliza una función de distribución acumulada de probabilidad.

$$\pi = \int_{-\infty}^t f(s)ds = F(t) = pr[T \leq t]$$

$$f(s) \geq 0$$

$$\int_{-\infty}^t f(s)ds = 1$$

$f(s)$: *Función de Tolerancia*

Los modelos lineales generalizados tienen diferente función donde se modelan las covariables. Las opciones que tenemos para modelar el tipo de respuesta binomial son las siguientes:

- Función Logit: Modelo de Crecimiento Logístico.
- Función Probit: Modelo de Distribución Normal.
- Función Log Log: Distribución del Valor Extremo (Logaritmo de Weibull).

Restricción: No se encuentra en el eje de los reales sino en el eje de Probabilidad.

4.6.2 Modelo logístico

Es un modelo que presenta una gráfica de crecimiento Sinoidal, al cual se les llama modelo “en forma S” (S – Shape), su razón de crecimiento de la curva es sinoidal y se modela con el producto del tamaño actual del crecimiento y el crecimiento remanente.

$$df_{\alpha}g(f)[h(\alpha) - h(f)]$$

Donde

- “g” y “h” son funciones crecientes

- $g(0) = h(0) = 0$
- α : *Tamaño Final*
- γ : *Punto de Inflexión*

Una forma simple de la función de crecimiento es:

$$g(f) = h(f) = f$$

$$\frac{df}{dx} = \frac{k}{\alpha} f(\alpha - f)$$

Ecuación Diferencial Ordinaria

$$\Rightarrow \frac{df}{dx} = af(\alpha - f)$$

Donde:

- $k > 0$
- $0 < f < \alpha$
- α : *Tamaño final (Máximo)*

La solución a esta ecuación (**demostración 13, apéndice A**), es la función de Probabilidad de Éxito, escrita de una mejor manera como:

$$\frac{1}{1 + e^{-x\beta}} = \pi \quad (55)$$

Si tomamos en cuenta la posibilidad de fracaso, podemos obtener la función decreciente de Probabilidad de Fracaso:

$$1 - \pi = \frac{e^{-x\beta}}{1 + e^{-x\beta}} \quad (56)$$

Entonces para la función de probabilidad, (**véase demostración 14, apéndice A**):

$$\pi \int_{-\infty}^x f(s) ds = \frac{e^{x\beta}}{1 + e^{x\beta}}$$

El resultado es:

$$b(\pi) = X\beta \quad (57)$$

Entonces, la función de enlace para el caso Binomial es:

$$\log\left[\frac{\pi}{1-\pi}\right] = X\beta \quad (58)$$

Obtenida la función de enlace en esta distribución, es necesario obtener los parámetros β del modelo de la siguiente manera (recordando la función log verosimilitud y la función enlace de la ecuación (58)):

$$l = \sum_{i=1}^n \left\{ y_i X\beta - n_i - \log(1 + e^{X\beta}) + \log\binom{n_i}{y_i} \right\} \quad (59)$$

Ahora, utilizaremos el Método de Scoring para la obtención de los parámetros, partiendo de lo siguiente:

$$X^T W X b^m = X^T W Z \quad (60)$$

El vector de ponderaciones \mathbf{W} es:

$$W = \frac{1}{\text{var}(y_i)} \left(\frac{d\mu_i}{d\eta_i} \right)^2$$

$$w_{ii} = \frac{1}{\text{var}(y_i)} [g'(\pi)]^2$$

$$\Rightarrow \sqrt{w_{ii}} = \frac{1}{\text{std}(y)} g'(\pi) = \sqrt{\text{Var}(y)}$$

La parte \mathbf{Z} de la ecuación es:

$$Z = X\beta + (y - \mu_i) \left(\frac{d\eta_i}{d\mu_i} \right)$$

Con $\eta_i = \text{Función de Enlace}$

Entonces:

$$Z = X\beta + (y - \mu_i) g'(\pi)$$

$$= \ln \left[\frac{\pi}{1 - \pi} \right] + (y - \pi) \left(\frac{1}{\pi(1 - \pi)} \right)$$

Despejando para "b":

$$b = (X^T W X)^{-1} X^T W Z$$

Hay que notar que el valor inicial para π es:

$$\pi_0 = \frac{N_y + 0.5}{N + 1}$$

Prueba de Wald individual:

$$\frac{\hat{\beta}}{\sqrt{\text{Var}(\hat{\beta})}} = N(0,1) (\hat{\beta} - \beta)^T J(\hat{\beta} - \beta) \sim \chi_p^2$$

4.6.3 Regresión logística multinomial

La regresión logística multinomial puede utilizarse de forma nominal u ordinal, cuando una de las variables de respuesta es categórica.

La elección de qué método utilizar en una situación particular depende de si una variable es claramente una "respuesta" o varias variables tienen el mismo estado. Además, la elección puede depender de cómo los resultados deben ser presentados e interpretados.

Las estimaciones de las probabilidades de rendimiento en la regresión nominal y ordinal de regresión logística son relativamente fáciles de interpretar si no hay interacciones (o sólo las interacciones bastante simples). Los modelos log-lineales son buenos para probar hipótesis acerca de las interacciones complejas, pero las estimaciones de los parámetros se interpretan con menos facilidad (Dobson, 2002).

4.6.4 Distribución Multinomial

Considere una variable aleatoria Y con J categorías. Denotamos $\pi_1, \pi_2, \dots, \pi_J$ como las respectivas probabilidades, con $\pi_1 + \pi_2 + \dots + \pi_J = 1$.

Si hay “ n ” observaciones independientes de Y , donde y_1 resulta la categoría 1, y_2 la categoría 2, y así sucesivamente, entonces hacemos (Dobson, 2002):

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_J \end{bmatrix}, \text{ con } \sum_{j=1}^J y_j = n \quad (61)$$

La distribución multinomial es:

$$\frac{n!}{y_1! y_2! \dots y_J!} \pi_1^{y_1} \pi_2^{y_2} \dots \pi_J^{y_J} \quad (62)$$

Si $J = 2$, entonces $\pi_2 = 1 - \pi_1$, $y_2 = n - y_1$ ecuación (62) sería la distribución binomial. En general, no satisface los requerimientos para ser miembro de la familia exponencial, aunque siguiendo una relación con la distribución Poisson, satisface el hecho de que puede ser modelada a través de los modelos lineales generalizados (Dobson, 2002). Denotamos como Y_1, \dots, Y_J como variables aleatorias independientes con distribuciones Poisson. Entonces su distribución de probabilidad conjunta es:

$$f(y) = \prod_{j=1}^J \frac{\lambda_j^{y_j} e^{-\lambda_j}}{y_j!} \quad (63)$$

Si denotamos $n = Y_1 + Y_2 + \dots + Y_J$, **entonces** “ n ” es una variable aleatoria con distribución Poisson $(\lambda_1 + \lambda_2 + \dots + \lambda_J)$. **Entonces**, la distribución de “ y ” condicional en “ n ” es:

$$f(y|n) = \left[\prod_{j=1}^J \frac{\lambda_j^{y_j} e^{-\lambda_j}}{y_j!} \right] / \frac{(\lambda_1 + \lambda_2 + \dots + \lambda_J)^n e^{-(\lambda_1 + \lambda_2 + \dots + \lambda_J)}}{n!} \quad (64)$$

Que puede simplificarse como:

$$f(y|n) = \left(\frac{\lambda_1}{\sum \lambda_k} \right)^{y_1} \dots \left(\frac{\lambda_J}{\sum \lambda_k} \right)^{y_J} \frac{n!}{y_1! \dots y_J!} \quad (65)$$

Si $\pi_j = \lambda_j / (\sum_{k=1}^K \lambda_k)$, para $j = 1, \dots, J$, entonces la ecuación (64) es la misma que la ecuación (69) y $\sum_{j=1}^J \pi_j = 1$, según sea necesario. De tal manera que, la distribución multinomial puede ser considerada como la distribución conjunta de variables aleatorias Poisson. Este resultado provee la justificación del uso de modelos lineales generalizados para su ajuste (Dobson, 2002).

4.6.5 Regresión logística nominal

Estos modelos son utilizados cuando las respuestas o sus categorías no llevan un orden. Cualquiera de las categorías en las respuestas puede ser considerada como categoría de referencia. Por lo general, se supone que la primera categoría es la de referencia (Dobson, 2002). Entonces, los logits para las otras categorías son definidas por:

$$\text{logit}(\pi_j) = \log\left(\frac{\pi_j}{\pi_1}\right) = x_j^T \beta_j \quad \text{para } j = 2, \dots, J \quad (66)$$

Las $(J - 1)$ ecuaciones de logits son utilizadas simultáneamente para estimar los parámetros β . Una vez obtenidos los parámetros estimados b_j , los predictores lineales pueden ser calculados de la ecuación anterior.

$$\hat{\pi}_j = \hat{\pi}_1 \exp(x_j^T \beta_j) \quad \text{para } j = 2, \dots, J \quad (67)$$

Pero si tomamos en cuenta que la suma de probabilidades estimadas $\hat{\pi}_j$ son iguales a 1, entonces:

$$\hat{\pi}_j = \frac{\exp(x_j^T b_j)}{1 + \sum_{j=2}^J \exp(x_j^T b_j)} \quad \text{para } j = 2, \dots, J \quad (68)$$

Los valores ajustados o frecuencias esperadas, para cada patrón de covariable pueden ser calculadas multiplicando las probabilidades estimadas de $\hat{\pi}_j$ por el total de frecuencia de cada patrón de covariable. El resumen estadístico para la bondad de ajuste es análogo para la regresión logística binomial (Dobson, 2002).

4.6.6 La prueba de independencia basada en el estadístico Chi – Cuadrada

Si “n” elementos han sido observados en una población, específicamente dos de sus características (sea X o Y), y se toma una muestra aleatoria simple de éstas poblaciones, entonces se desea contestar si las características de X son independientes de las de Y. Se dividirán los valores de **X** en conjuntos disjuntos A_1, A_2, \dots, A_k mientras que para **Y** será descompuesto en “r” conjuntos disjuntos B_1, B_2, \dots, B_r . Cuando los elementos sean clasificados, aparecerá un cierto número de ellos n_{ij} , en cada una de las $k \times r$ clases así constituidas, y dará lugar a una tabla de contingencia como se muestra en la tabla 4.1:

Tabla 4.1 Clasificación de elementos (Tabla de contingencia)

	A_1	A_2	A_k	Total
B_1	n_{11}	n_{12}	n_{1k}	n_1
B_2	n_{21}	n_{22}	n_{2k}	n_2
B_r	n_{r1}	n_{r2}	n_{rk}	n_r
Total	$n_{.1}$	$n_{.2}$	$n_{.k}$	n

El test de independencia chi – cuadrada utiliza el estadístico de contraste:

$$\chi^{2*} = \sum_{i=1}^r \sum_{j=1}^k \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \quad (69)$$

Los grados de libertad son $(k - 1)(r - 1)$, en donde está igualado e_{ij} con $n_i \cdot n_j/n$

4.7 Optimización multiobjetivo.

Si en un problema de optimización se requiere obtener la mejor solución para cumplir más de un objetivo, a esto se le denomina optimización multi objetivo. Entonces, la atención debe recaer en todos los objetivos y no solamente

en cumplir uno, es por eso que una solución para un objetivo puede que no sea para los demás.

Éste tipo de problemáticas están presentes en muchas áreas y encontrar la solución ha sido un reto constante para los investigadores. La motivación para utilizar algoritmos y solucionar éstos problemas radica en que al estar basados en una población para la búsqueda de soluciones, éstos nos generan un conjunto de soluciones de buena calidad y no solamente una solución (Peñuela, 2007).

Los algoritmos de optimización multiobjetivos requieren métodos matemáticos de optimización sobre una población de soluciones, los algoritmos genéticos son una excelente propuesta para solucionar éstas anomalías, dadas sus características de diversidad y confiabilidad.

El modelo de un problema de optimización multi objetivo es:

$$\min/\max: f_1(x), m = 1, 2, \dots, M$$

s. a.

$$g_1(x) \geq 0, \quad j = 1, 2, \dots, J$$

$$h_k(x) \geq 0, \quad k = 1, 2, \dots, K$$

$$x'_i \leq x \leq x''_i, \quad i = 1, 2, \dots, n$$

Para dar solución a lo anterior, primero se definen las soluciones de calidad y cuáles no se necesitan. Aquí entra el término de dominancia. Es posible decir que $x^{(1)}$ domina a otra solución $x^{(2)}$ si se cumple lo siguiente:

- a) $x^{(1)}$ no es peor que $x^{(2)}$ en todos los objetivos.
- b) $x^{(1)}$ es estrictamente mejor que $x^{(2)}$ en por lo menos un objetivo.

Si no se cumple alguna de las condiciones, la solución $x^{(1)}$ no domina la solución $x^{(2)}$. El significado de dominancia puede extenderse para encontrar un conjunto de soluciones no dominadas dentro de una población. Es posible utilizar el siguiente procedimiento para encontrar el conjunto de soluciones:

- 1) Se hace $j = 1$.
- 2) Para todos los $j \neq i$, comparar las soluciones $x^{(i)}$ y $x^{(j)}$ para determinar dominancia, se usan las condiciones mencionadas.
- 3) Si para alguna j , $x^{(i)}$ es dominado por $x^{(j)}$, marcar $x^{(i)}$ como dominado. Incrementar "i" en uno e ir al paso 2.
- 4) Si todas las soluciones (cuando se llega a $i = j$) en el conjunto son consideradas, ir al paso 5; de lo contrario incrementar "i" en uno e ir al paso 2.
- 5) Todas las soluciones que no son marcadas como dominadas son soluciones no-dominadas.

Si al comparar dos soluciones no se cumple ninguna condición, no es posible concluir acerca de la dominancia de una con respecto a otra. Cuando esto sucede, se dice que las soluciones son no-dominadas. Si se tiene un conjunto finito de soluciones y se realiza una comparación de todos los pares posibles, al final se va a tener un conjunto de soluciones que son no dominadas entre sí y dicho conjunto tiene la propiedad de dominar al resto de soluciones que no pertenezcan a él. Si se tiene un conjunto de soluciones P , el conjunto no dominado de soluciones P' lo forman aquellas que no son dominadas por ningún miembro del conjunto P , este conjunto recibe el nombre de frente de Pareto (Flórez, 2008).

4.7.1 Algoritmo Genético de clasificación No – Dominado (NSGA II)

Propuesto por Deb y sus estudiantes en el año 2000. Se crea una población Q_1 de tamaño N usando una población de "padres" P_1 de tamaño N . Se combinan las poblaciones y mediante un ordenamiento no dominado, se clasifican en diferentes frentes de Pareto.

Una vez hecha la combinación, se genera una nueva población a partir de las configuraciones de los frentes no dominados. Se construye con el mejor frente

no dominado (F_1), sigue con las soluciones del segundo frente (F_2), tercer (F_3) y así sucesivamente. Como la población R_t es de $2N$, y solamente existen N configuraciones que conforma la población descendiente, no todas se podrán acomodar en la nueva población y desaparecen los que no entran en el acomodo.

Considerando el último frente, algunas soluciones sobresalen de las restantes para acomodar a la población de forma descendiente. Entonces, es útil utilizar algún método o estrategia para seleccionar las configuraciones situadas en un área poco poblada (alejada de otras soluciones) para llenar las posiciones restantes de la población descendiente en vez de escogerlas aleatoriamente (Flórez, 2008).

La idea principal es promover las configuraciones que aseguren una variedad dentro del mismo frente de Pareto. Si la población en su totalidad converge al frente de Pareto óptimo, el algoritmo está asegurando que las soluciones estén distanciadas una de otra.

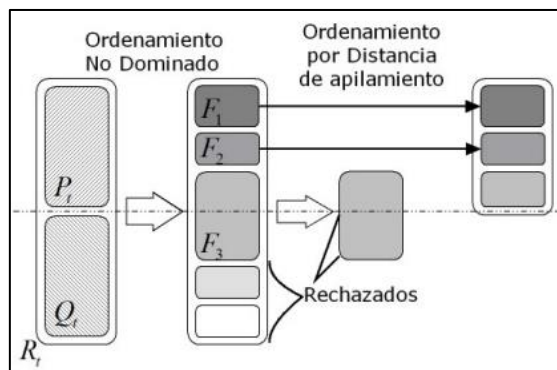


Figura 4.4 Diagrama de promoción de individuos del NSGA II

Los pasos a seguir en el algoritmo NSGA-II son los siguientes (Flórez, 2008):

- 1- Se combinan las poblaciones tanto de padres como descendientes y se crea $R_t = P_t \cup Q_t$. Se realiza el ordenamiento no dominado a R_t y se identifican los frentes $F_t, i = 1, 2, \dots, etc.$
- 2- Se hace $P_{t+1} = \emptyset$, e $i = 1$. Si se cumple que $|P_{t+1}| + |F_i| < N$ entonces se hace $|P_{t+1}| = |P_{t+1}| \cup |F_i|$ e $i = i + 1$.

- 3- El siguiente paso es hacer el ordenamiento por apilamiento ($F'_i < C$ descrito posteriormente) y se van a incluir en P_i las soluciones $N - |P_{i+1}|$ más esparcidas usando los valores de distancia de apilamiento asociadas al frente F_1 .
- 4- Se crea la población descendiente Q_{i+1} a partir de los P_{i+1} mediante el procedimiento de elección por torneo para apilamiento, cruzamiento y mutación.

4.7.1.1 Selección por torneo para apilamiento

Se utiliza un operador ($F'_i < C$) para comparar dos soluciones y elegir un ganador. Se considerará que una solución tiene (r_i) rango de dominancia asociado y una (d_i) distancia de apilamiento que es la medida del espacio de búsqueda en los alrededores de “i” sin ocupar por otra solución.

Si están situadas en el mismo frente de Pareto, gana la que tiene una mejor o distancia más alta de apilamiento. La distancia de apilamiento se determina haciendo uso de la siguiente ecuación en el algoritmo:

$$d_{i,j}^m = d_{i,j}^m + \frac{f_m^{d_{i,j+1}^m} - f_m^{d_{i,j-1}^m}}{f_m^{max} - f_m^{min}} \quad (70)$$

El valor máximo y mínimo son representados por f_m^{max} , f_m^{min} y la función objetivo es “m”, las soluciones vecinas son $f_m^{d_{i,j+1}^m}$, $f_m^{d_{i,j-1}^m}$ para cada una de las funciones objetivo m. Se consideran todas las posibles funciones objetivo en las distancias y asignan un valor infinito a las soluciones extremas del frente de Pareto. La distancia resultante es la suma de las distancias en cada una de las direcciones de las funciones objetivo del problema. Así es como el algoritmo utiliza operadores básicos y promueve las configuraciones que ocupen los mejores frentes de Pareto.

4.7.1.2 Selección

Mediante la lógica del NSGA II, se realizan dos torneos, es decir, se selecciona la configuración mejor ubicada en el mejor frente de Pareto y si las configuraciones involucradas están en el mismo frente se utiliza la mejor distancia de apilamiento para seleccionar la mejor (Flórez, 2008).

4.7.1.3 Cruzamiento y mutación

En esta parte, se utilizan los dos “padres” obtenidos del operador de selección. Para escoger un descendiente, se realiza de la siguiente forma:

- Se introducen los dos hijos de manera temporal en la población,
- La población queda con “N+2” individuos
- Se realiza el ordenamiento no dominado para clasificar las configuraciones de acuerdo al frente al que pertenezcan
- Se asignan las respectivas distancias de apilamiento.

A partir de los valores ajustados anteriormente, se elige el mejor hijo, es decir, el que se encuentre en el mejor frente no dominado o posea menor distancia de apilamiento. En el operador de mutación sucede de forma parecida, donde la configuración obtenida después de la mutación es comparada con la original y así elegir correctamente (Flórez, 2008).

Capítulo 5

Metodología

Considerando el problema planteado en el capítulo 2 y a revisión de la literatura referente a los modelos lineales generalizados, el presente capítulo se enfoca en la metodología propuesta para modelar los defectos del proceso de soldadura por fricción agitación a través de sus variables de entrada para obtener los niveles adecuados en donde la probabilidad de ausencia de defectos sea máxima.

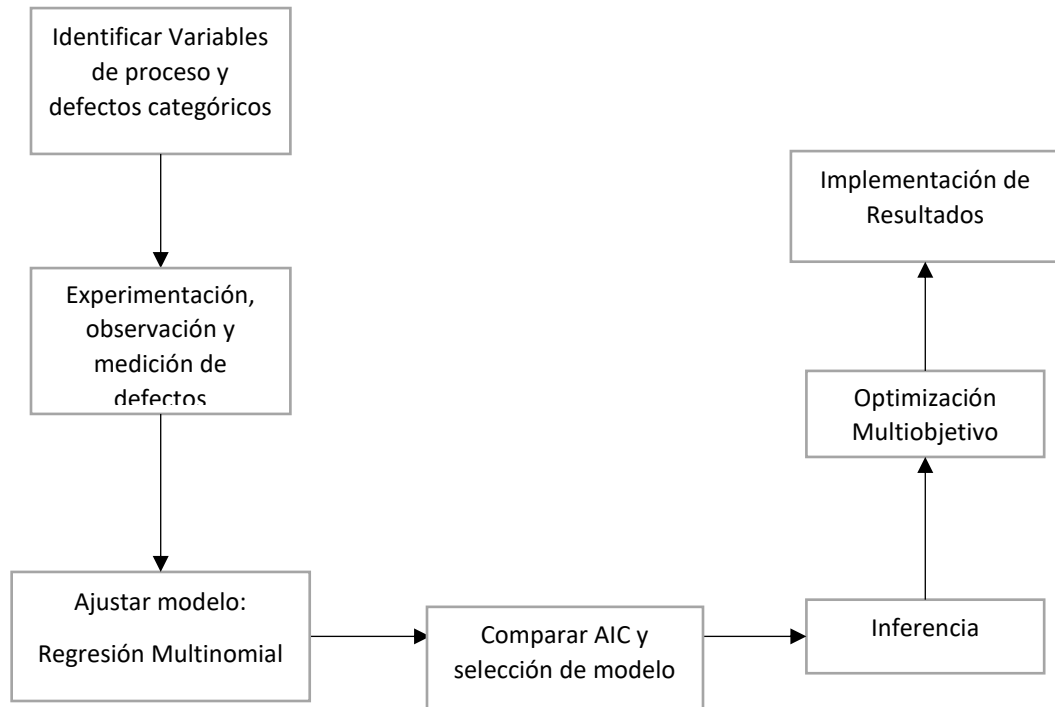


Figura 5.1 Metodología

Pasos generales de la metodología:

1- Identificar variables de proceso y defectos categóricos

Es necesario identificar las variables de entrada del proceso de FSW así como los defectos que se presentan y el procedimiento a seguir para detectarlos, lo cual puede ser realizado a través de la consulta de la literatura.

2- Experimentación, observación y medición de los defectos

Al identificar las variables de entrada y salida de FSW, el siguiente paso es la observación y medición de los defectos, lo cual puede ser realizado, primeramente, a través de la experimentación inicial utilizando la información proporcionada por la literatura tanto para los valores de las variables de entrada

como los métodos y herramientas a utilizar para la medición de los defectos en FSW.

3- Ajustar modelo: regresión multinomial

En este paso, se determinará si existe una relación entre las variables de respuesta (defectos), se utilizará una prueba de independencia, en este caso el test de independencia de Chi – Cuadrado (**ver capítulo 4.6.3**).

Una vez determinada la existencia o no de independencia de los datos, se ajustará la matriz de diseño experimental a través de un modelo lineal generalizado, para el caso de estudio se utilizó la técnica de regresión logística multinomial al tener más de una respuesta a explicar, y cada respuesta del tipo binomial (**ver capítulo 4.6.3**).

4- Comparar valor AIC y seleccionar modelos.

Una vez obtenidos los conjuntos de modelos, se realizará la prueba de AIC para elegir el conjunto de modelos que represente de una mejor manera el comportamiento de los defectos a través del valor de Akaike (el menor valor de AIC es más adecuado).

5- Inferencia

Se determinará la significancia de las variables y el buen ajuste del modelo, se analizarán los resultados obtenidos y posteriormente obtener una posible optimización de parámetros y para continuar con las conclusiones previas.

6- Optimización Multiobjetivo

Se encontrará y determinará la región de niveles de las variables explicativas en los que las uniones estarán libres de defectos (**ver capítulo 4.7**).

7- Implementación de resultados

Obtenidos los parámetros del modelo, el modelo más adecuado y la región de niveles de las variables de los factores de proceso, se espera la validación en campo para concluir satisfactoriamente el caso de estudio y posteriormente determinar las conclusiones finales del proyecto.

Capítulo 6

Aplicación

6.1 Identificar variables de proceso y defectos categóricos

A partir de las corridas experimentales realizadas a las placas de matriz metálica mediante soldadura por fricción agitación, se obtienen las muestras y resultados de dichas corridas para el análisis de los datos. **Cuevas, 2014** sugiere que los niveles de las variables de entrada a utilizar para las placas de compuesto de matriz metálica para proceso de soldadura por fricción agitación son:

- Velocidad de rotación de 1000 rpm – 1600 rpm
- Velocidad de soldadura (Avance) con 40mm/min – 75mm/min

Como se menciona en el capítulo 4.1, la respuesta se obtuvo mediante inspección visual, la más simple y la primera inspección realizada en la soldadura ya que detecta defectos como rebaba, falta de penetración evidente, discontinuidad, huecos. Con ésta inspección es posible observar si los parámetros de entrada necesitan mejorarse o corregirse para obtener una mejor calidad (Lohwasser y Zhan, 2001).

Este tipo de inspección se puede realizar en cualquier etapa de un proceso productivo o durante las operaciones de mantenimiento, dado que muestra las discontinuidades más grandes y señala otras que se pueden detectar mediante otros métodos (Lohwasser y Zhan, 2001).

Los defectos que se consideraron en la soldadura por fricción agitación son: exceso de rebaba, huecos y hoyos de gusano.

Y los factores de proceso considerados fueron: velocidad de rotación, velocidad de soldadura, profundidad de penetración, geometría y ángulo de la herramienta, presión (AWS D17.3 - 2010).

Se consideraron solo dos variables de entrada, velocidad de rotación y velocidad de avance, ya que son las únicas variables de las que se dispone un control para variar sus niveles en el equipo utilizado.

6.2 Experimentación, observación y medición de defectos

Las muestras a unir se cortaron en placas de 80mm x 30mm. El proceso se llevó a cabo en el equipo CNC marca HAAS, modelo VM2, de la Celda de Manufactura Avanzada de COMIMSA. En la figura 6.1 se muestra un ejemplo de las placas a unir así como su medida:

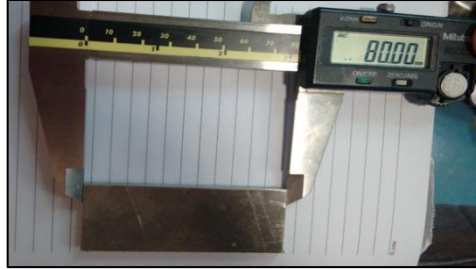


Figura 6.1 Placas de material compuesto A359/SiC/20

La figura 6.2 muestra el equipo del CNC que se utilizó para realizar la experimentación inicial de FSW.



Figura 6.2 CNC de la Celda de Manufactura Avanzada de COMIMSA

Las herramientas utilizadas en el proceso de FSW para unir las placas se muestran en las figuras 6.3, 6.4, 6.5, comenzando con la representación esquemática de la herramienta en la figura 6.3.

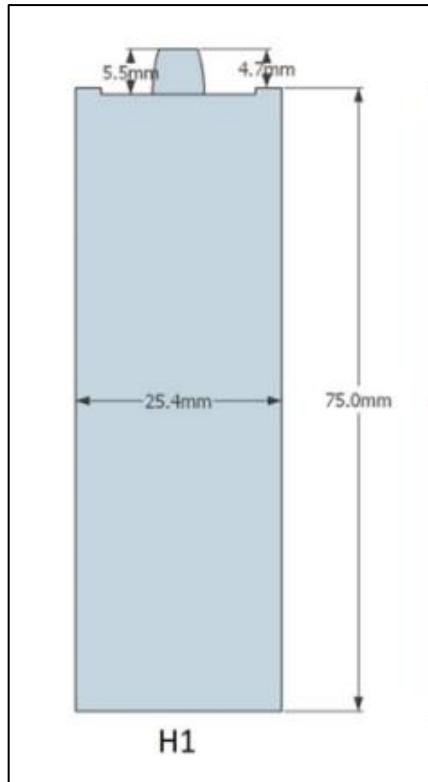


Figura 6.3 Representación esquemática de la herramienta para FSW

En la figura 6.4 se muestra una vista de la punta de la herramienta utilizada en la experimentación, posteriormente su medida en la figura 6.5.

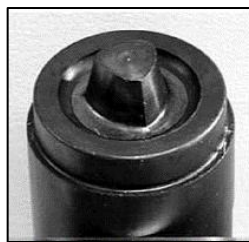


Figura 6.4 Vista de la punta y hombro de la herramienta

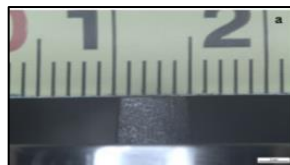


Figura 6.5 Medida de la Herramienta

Para la preparación del proceso, se siguieron los siguientes pasos:

1. Rectificar la probeta (queda en las medidas necesarias para el sistema de sujeción). El rectificado se realizó en un torno.
2. El sistema de sujeción, es un aditamento el cual se encarga de sostener correctamente las probetas y soldarlas mediante FSW (para evitar que se muevan).
3. Las piezas se cortaron en una máquina cortadora marca Struers.
4. La superficie de la máquina se limpió para que estuviera libre de grasa y polvo, únicamente con un trapo del alcohol.

Se realizaron las corridas experimentales según las corridas mostradas en la tabla 6.1.

Tabla 6. 1 Corridas experimentales

Muestra	Vel. avance (mm/m)	Vel. rotación (rpm)
1	75	1000
2	75	1250
3	50	1000
4	50	1400
5	75	1000
6	50	1400
7	40	1600
8	40	1600
9	50	1500

La medición de las respuestas fue a través de la ausencia y/o presencia de defectos en cada unión soldada, apoyado de una persona con experiencia en éste tipo de inspección para determinar cuál de las uniones tiene o no defectos.

La figura 6.6 muestra los defectos más comunes.

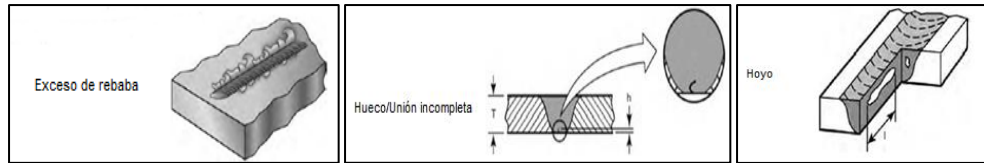


Figura 6.6 Defectos comunes en FSW

Para la medición de los defectos de la figura 6.6, se requiere el apoyo de los expertos en la inspección de uniones en FSW, así se logra determinar cuáles de estos defectos están presentes o ausentes en las muestras.

En la figura 6.7 se muestran las uniones soldadas obtenidas en la experimentación inicial de las cuales se procede a la detección de defectos.

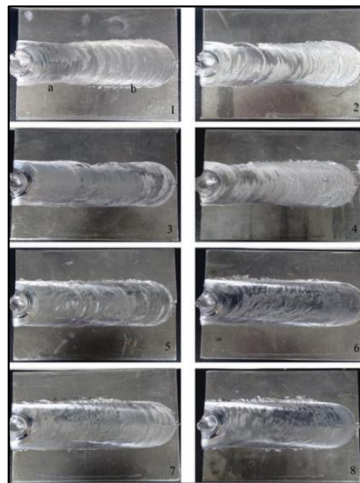


Figura 6.7 Piezas soldadas mediante FSW (Experimentación Inicial)

Es evidente que la característica de calidad requerida para la pieza soldada en este proceso depende de varios criterios como el hecho de que pueden estar ausentes o presentes los defectos comunes.

Dado que en la experimentación inicial se obtiene una respuesta del tipo binomial basada en el éxito o fracaso de varias características de calidad de la pieza, es necesario modelar dichas características de calidad en términos de las variables de entrada. Entonces, lo siguiente es modelar no solo una respuesta,

si no todas las respuestas observadas de manera simultánea (defectos en la pieza). Como se mencionó en un principio, se determina como $y = 1$ la presencia del defecto, $y = 0$ como la ausencia del defecto. Los defectos de cada unión soldada se muestran en las figuras 6.8 a 6.17:

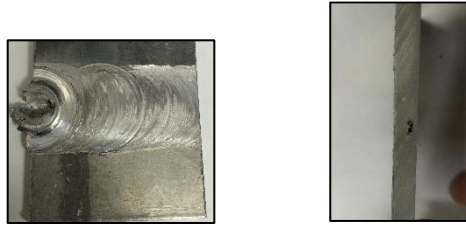


Figura 6.8 Presencia de rebaba y hueco, muestra 1

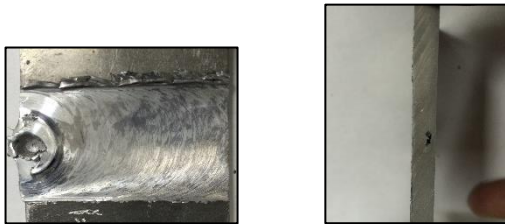


Figura 6.9 Presencia de rebaba y hueco, muestra 2

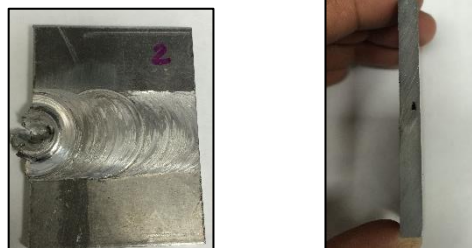


Figura 6.10 Ausencia de rebaba y presencia de hueco, muestra 3



Figura 6.11 Ausencia de rebaba y hueco, muestra 4

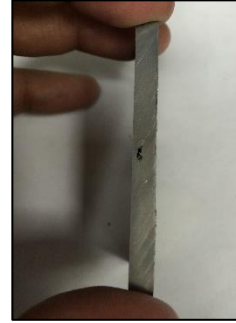


Figura 6.12 Ausencia de rebaba y presencia de hueco, muestra 5



Figura 6.13 Ausencia de rebaba y presencia de hueco, muestra 6

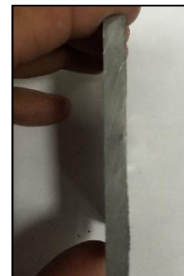


Figura 6.14 Ausencia de rebaba y hueco, muestra 7



Figura 6.15 Presencia de rebaba y ausencia de hueco, muestra 8



Figura 6.16 Ausencia de rebaba y hueco, muestra 9



Figura 6.17 Uniones 1 a 5 y unión 8 presentan hoyo de gusano.

Una vez identificados los defectos para cada corrida, se crea la tabla 6.2 en el que se muestran los defectos que están presentes o ausentes en cada corrida experimental.

Tabla 6.2 Tabla de datos con defectos

Muestra	Velocidad de rotación	Velocidad de avance	Rebaba	Hueco	Hoyo
1	1000	75	1	1	1
2	1000	50	1	1	1
3	1000	75	0	1	1
4	1250	75	0	0	1
5	1400	50	0	1	1
6	1400	50	0	1	0
7	1500	50	0	0	0
8	1600	40	1	0	1
9	1600	40	0	0	0

6.3 Ajustar modelo: regresión multinomial.

Es importante determinar la independencia de las variables de respuesta, ya que en caso de que sean dependientes, el ajuste de los datos tiene que realizarse con otro tipo de herramienta estadística. La prueba seleccionada para la independencia de las respuestas es el test de independencia chi – cuadrado, que se muestra en la tabla 6.3.

Tabla 6.3 Prueba de independencia

	Hueco	Rebaba	Hoyo	Total
1	5	3	6	14
	4.67	4.67	4.67	
	Hueco	Rebaba	Hoyo	Total
2	4	6	3	13
	4.33	4.33	4.33	
	0.026	0.641	0.410	
Total	9	9	9	27
Chi Cuadrado	2.077	GL = 2	Valor P	0.354

Considerando la prueba mostrada en el capítulo 4.6.6 es posible considerar el valor de probabilidad “p” para realizar la prueba de independencia Chi-Cuadrada. Dado que el valor de probabilidad fue $p=0.354$, es posible concluir que las observaciones son independientes. Por lo tanto, es posible utilizar el modelo multinomial ya que se cumple con el supuesto de independencia entre las variables de respuesta.

Como se puede apreciar en el capítulo 4.7.3, a partir de la tabla de datos con la presencia defectos como respuestas posibles, se toma una categoría base para realizar el ajuste de los modelos a través de la estimación de los parámetros de la regresión logística nominal. Los parámetros estimados se muestran en la tabla 6.4.

Tabla 6.4 Parámetros de modelo para cada respuesta

Referencia rebaba				
	B₀	B_{1X1}	B_{2X2}	B_{3X1X2}
Sin hueco	-40.4740249	0.0006113	-0.6290553	0.001020274
Sin rebaba	0.05271604	-0.01511011	-0.4915208	0.000692285
Sin hoyo	19.4573781	-0.03054164	-1.0902162	0.001144479
	Devianza	75.83959	AIC	115.8396
Referencia Hueco				
	B₀	B_{1X1}	B_{2X2}	B_{3X1X2}
Sin hueco	-70.262934	0.03392935	0.18304399	0.000163326
Sin rebaba	-16.072564	0.0086369	0.08584057	3.21672E-06
Sin hoyo	23.668707	-0.02140748	-1.08123267	0.000859653
	Devianza	74.79382	AIC	114.7938
Referencia Hoyo				
	B₀	B_{1X1}	B_{2X2}	B_{3X1X2}
Sin hueco	-22.437753	0.00459383	-0.35128029	0.000464624
Sin rebaba	-3.141181	-0.00030171	-0.08472786	0.000116215
Sin hoyo	10.03143	-0.01342976	-0.74625701	0.000650879
	Devianza	76.73048	AIC	116.7305

Tomando como referencia diferentes categorías, se obtiene un conjunto de modelos diferentes para cada respuesta, para lo cual es necesario realizar la selección del mejor modelo a través del criterio de información de Akaike, el cual menciona que el mejor modelo corresponde al menor valor AIC.

6.4 Comparar valor AIC y seleccionar modelos.

Los modelos correspondientes a la presencia de hueco, exceso de rebaba y presencia de hoyo están dados por las ecuaciones (54), (55) y (56), respectivamente.

$$\pi = 1 - \left(\frac{1}{1 + e^{(-(-70.262934 + 0.03392935x_1 + 0.18304399x_2 + 0.00016333x_1x_2))}} \right) \quad (54)$$

$$\pi = 1 - \left(\frac{1}{1 + e^{(-(-16.072564 + 0.008636899x_1 + 0.08584057x_2 + 0.000003216721x_1x_2))}} \right) \quad (55)$$

$$\pi = 1 - \left(\frac{1}{1 + e^{(-(-23.668707 - 0.021407483x_1 - 1.08123267x_2 + 0.0008596534x_1x_2))}} \right) \quad (56)$$

La tabla 6.5 muestra los niveles a introducir y la probabilidad de que ocurra el defecto con cada modelo.

Tabla 6.5 Probabilidad de ocurrencia (o ausencia) del defecto

x1	x2	Probabilidad de hueco en unión	Probabilidad de rebaba en unión	Probabilidad de hoyo en unión
1000	75	0.999968193	0.680564706	0.999999418
1000	50	0.999999994	0.951791032	0.999851939
1000	75	0.999968193	0.680564706	0.999999418
1250	75	0.233454341	0.187982946	0.973158317
1400	50	0.898145879	0.369050213	0.546778967
1400	50	0.898145879	0.369050213	0.546778967
1500	50	0.115804323	0.195278883	0.12240457
1600	40	0.142011629	0.200052328	0.234110442
1600	40	0.142011629	0.200052328	0.234110442

En la tabla 6.5 se pueden apreciar las regiones en donde la probabilidad de presencia de un defecto es mínima o máxima. La región de interés es en donde combinando un conjunto de parámetros la probabilidad de ocurrencia de un defecto sea mínima, es decir, la región en donde pueda obtenerse una muestra soldada por FSW libre de defectos.

En este sentido, es necesario utilizar un método de optimización del tipo multiobjetivo para cumplir la meta de ausencia de defectos simultáneamente.

6.5 Inferencia

En la tabla 6.6 se muestran los “valores p” para los parámetros en cada modelo, analizados mediante la prueba de Wald con el fin de establecer la significancia individual, así como la devianza para bondad de ajuste del modelo:

Tabla 6.6 Valores de probabilidad P estimados.

	Valor – P (referencia hueco)			
	Bo	B1x1	B2x2	B3x1x2
Sin hueco	0	0	0	0.0003027904
sin rebaba	0	0.0001232779	0	0.9353871197
Sin hoyo	0	0.000001461063	0	0

En el modelo con referencia “hueco” se observa que la interacción entre las dos variables en la respuesta “sin rebaba”, no es significativa.

En cuanto a la bondad de ajuste del modelo, se utiliza la Devianza para indicar qué tan bueno es el modelo como se muestra en el capítulo 4.5.2. Esto, se muestra en la tabla 6.7:

Tabla 6.7 Devianza y bondad de ajuste

	Valor	Grados de Libertad	Chi - Cuadrada	p - Value
Devianza	74.79382	23	3.251905217	0.9999

Considerando los resultados mostrados en la tabla 6.7, es posible concluir que no se rechaza H_0 , es decir, el modelo representa el comportamiento de la presencia de defectos adecuadamente (Agresti, 2006; Agresti, 2013).

6.6 Optimización Multiobjetivo

El procedimiento consiste en determinar los niveles de las variables de entrada en los que, para la ecuación de predicción de probabilidad de cada defecto, la probabilidad de que ocurra alguno de ellos sea cero para cada uno de ellos simultáneamente. Se utiliza el algoritmo NSGA II (Algoritmo genético de clasificación No – Denominada – II) para encontrar la región de parámetros óptimos en donde las piezas estarán libres de defectos. En el algoritmo se determina la función y las restricciones dadas las condiciones del proceso. En la tabla 6.8 se muestran los resultados de la optimización.

Tabla 6.8 Puntos óptimos determinados por medio del NSGA II

Función (Hueco)	Función (Rebaba)	Función (Hoyo)	x_1 : V. Rotación	x_2 : Vel. Avance
0.0000096878	0.0318503870	0.0009435123	1522.1429823633	69.8757156638
0.0000068060	0.0302561126	0.0004952610	1558.9932426261	66.8647977208
0.0000068060	0.0302561062	0.0004952612	1558.9931495848	66.8648091959
0.0000068060	0.0302561029	0.0004952622	1558.9929313914	66.8648316840
0.0000062457	0.0295881403	0.0005251257	1548.5771912186	68.1333895019
0.0000025236	0.0248349654	0.0002588395	1564.7366868917	68.5394207505
0.0000020919	0.0233043352	0.0004117721	1523.4396581909	73.2888113092
0.0000020671	0.0231818073	0.0004366502	1519.7561395300	73.7085175741
0.0000018619	0.0233609770	0.0002141377	1565.3503856985	69.1695132498
0.0000017983	0.0231540691	0.0002184497	1561.6832675483	69.6273041337
0.0000014473	0.0224082960	0.0001507343	1586.0846129823	67.6181817005
0.0000009530	0.0206896205	0.0001056958	1597.3135477354	67.4217625907
0.0000008468	0.0199987668	0.0001212889	1573.4779617425	70.1246918220
0.0000008315	0.0197905828	0.0001391846	1560.9682198799	71.4623726154
0.0000007975	0.0194135613	0.0001727517	1544.1254378449	73.3230106888
0.0000005579	0.0183457223	0.0000973613	1570.3033863789	71.4019701753
0.0000004684	0.0175778559	0.0001027552	1558.2641070789	73.0561316712
0.0000003015	0.0161706433	0.0000685405	1568.4143283524	72.9991597930
0.0000002143	0.0152405645	0.0000443580	1586.4193212752	71.9044880493
0.0000000955	0.0129411161	0.0000262930	1586.9521773460	73.6765180588

Puede observarse en la tabla 6.8, que es posible determinar la región óptima para cada parámetro de proceso. La región de niveles óptimos se muestra en la tabla 6.9:

6.9 Región óptima de parámetros de proceso

Parámetros de proceso	Región de Niveles Óptimos	
x1 : Vel. Rotación	1519.7561	1597.3135
x2 : Vel. Avance	66.8648	73.7085

A través de la herramienta NGSА II, se logró obtener la región óptima de parámetros en donde la probabilidad de encontrar huecos, exceso de rebaba y hoyo en las uniones soldadas mediante FSW sea mínima.

Capítulo 7

Conclusiones

Los modelos de regresión son una herramienta utilizada comúnmente con el objetivo de representar el comportamiento de una característica de calidad de un proceso a través de las variables de entrada. Es posible obtener una diversidad de respuestas las cuales al ser medidas cumplen con ciertas características, pueden ser del tipo continua o discreta, el tipo de modelo de regresión con el que es posible modelar ésta característica depende de su naturaleza. Cuando la variable dependiente es del tipo discreta, no es posible representarla con un modelo de regresión tradicional, para eso resulta útil un modelo lineal generalizado que es adecuado para éste tipo de respuestas.

Los resultados de esta investigación demuestran que la naturaleza de la característica de calidad analizada es del tipo discreta y es por eso que se propuso el uso de un modelo lineal generalizado para determinar el comportamiento que tienen los defectos analizados en la soldadura por fricción agitación. Además, se analizaron los defectos (múltiples defectos) que presentaron en las piezas soldadas y fue necesario el uso de un modelo de regresión multinomial para realizar el ajuste para más de una respuesta del tipo discreta. Entonces, para ajustar un modelo de regresión adecuado tomando en cuenta la naturaleza de los defectos, se propuso utilizar un modelo de regresión generalizado y además, mediante la prueba de Wald para prueba de significancia

individual de parámetros, se logró determinar las variables significativas en el modelo, con lo cual se contestó la primera pregunta de investigación.

No llevar a cabo lo anterior, tendrá efectos negativos en la modelación del comportamiento de los defectos y por consecuencia, las piezas soldadas obtenidas con los parámetros de un modelo inadecuado no serán de una calidad satisfactoria.

Utilizando un modelo de regresión multinomial fue posible modelar la relación de las variables de proceso considerando múltiples respuestas categóricas, con esto se respondió la segunda pregunta de investigación.

Para múltiples respuestas categóricas, fue necesario determinar en primera instancia que éstas son independientes una de otra, para esto se propuso el uso de una prueba de independencia, para este caso se utilizó la prueba de independencia de Chi Cuadrada, con la cual fue posible determinar la independencia de las respuestas (defectos múltiples categóricos), con esto se contestó la tercer pregunta de investigación, una de las pregunta importantes a responder ya que es necesaria la independencia de variables en el modelo de regresión multinomial.

Al observar que era necesario cumplir más de una característica de calidad en las uniones soldadas y encontrar la región de parámetros que lograra cumplir todas las características de calidad simultáneamente, se determinó que era necesaria la aplicación de una herramienta de optimización que lograra este objetivo. A través del uso y aplicación del algoritmo NGSA II, fue posible encontrar la región de los niveles de parámetros de proceso en donde las uniones están libres de defectos, con esto se respondió la cuarta y quinta pregunta de investigación.

Gracias a la literatura (Karthikeyan et al., 2010), (Cuevas, 2014), (Dawes y Thomas, 1996) así como la norma consultada (AWS D17.3, 2010), se obtuvieron tanto los parámetros y los niveles a utilizar para el material MMC en la experimentación, así como los defectos a observar en el proceso de FSW, en

conjunto con las pruebas específicas para la medición de los defectos. Con esto se contestaron los *primeros tres objetivos*.

Una vez detectadas las variables de entrada así como los defectos a considerar durante la experimentación, se procedió a la experimentación inicial, realizando y cumpliendo el *cuarto objetivo*.

Al conocer la forma en que se presentaban los defectos, se determinó su distribución y la clase de modelos que pueden representar su comportamiento. Conociendo la naturaleza de los defectos, se propuso el uso de la regresión multinomial, herramienta adecuada para el ajuste del modelo. Determinado el modelo a utilizar, se estudiaron los métodos y/o herramientas para estimación de los parámetros del modelo. Con esto se iniciaron y cumplieron los *objetivos cinco y seis*.

Para determinar la región de los niveles de los parámetros que cumpla simultáneamente las características de calidad requeridas (uniones soldadas sin defectos), se propuso el uso del algoritmo NSGA II, herramienta utilizada para la optimización multi – objetivo con la cual se logró obtener la región óptima de niveles de parámetros para obtener uniones donde los defectos estén ausentes, con esto se cumplió el *último objetivo*.

Como trabajo futuro, se deja en investigación alguna herramienta estadística que sea capaz de modelar el comportamiento de variables de respuesta dependiente.

Bibliografía

Badri M. et. al., 2015. Predicting unit testing effort levels of classes: An exploratory study based on Multinomial Logistic Regression Modeling. Vol: 62, 2015, pág. 529-538, Universidad de Quebec, Canadá.

Castañeda L. María Eugenia, López R. Víctor, 2010. "VII Coloquio Regional de Estadística", Comparación de Diseños Óptimos para Modelos no Lineales bajo diferentes enfoques. Departamento de Matemáticas, Facultad de Ciencias Exactas y Naturales, Universidad de Antioquia, Medellín, Colombia.

Chernoff H., 1953. "Locally optimal designs for estimating parameters", *Annals Mathematics Statistics*, 24, 586-602.

D.A. Belsley; 1991. *Conditioning Diagnostics: Collinearity and Weak Data in Regression*. John Wiley & Sons.

Dawes and Thomas, 1996 Friction Stir Process Welds Aluminum Alloys, *Welding Journal*, 75: 41-45.

Fernández V. y Fernández San Martín, 2004. *Regresión Logística Multinomial*. Depto. Estadística e Investigación Operativa. E.T.S. de Ingenierías Agrarias. Universidad de Valladolid, España.

Flórez Correa et. al. 2008. Algoritmo multiobjetivo NSGA-II aplicado al problema de la mochila. *Scientia et Technica*, vol. xiv, núm. 39, pp. 206-211

González-González D. S., Cantú-Sifuentes Mario, Praga-Alejo R., López-Cortez Víctor, 2010. A model to reduce the collinearity effects in the logistic regression.

Proceedings of the 16th Annual International Conference on Industrial Engineering Theory, Applications and Practice.

Haines Linda M., Kabera Gaetan, Principal Ndlovu and Timothy E. O'Brie, 2007. D-optimal Designs for Logistic Regression in Two Variables. Pág. 41 – 48

Hintze J. L., 2007. NCSS Users Guide III: Regression and Curve Fitting. NCSS Statistical System. Kaysville, Utah 84037. Copyright ©.

Karthikeyan L. et. al., 2010. On the role of process variables in the friction stir processing of cast aluminium A319 alloy. Mater, 31:761–71.

Kitsos and W.G. Muller, W.G., Heidelberg: Physica Verlag, Heidelberg, D-Optimal Designs for Generalized Linear Models. Statistica Sinica, pág. 87-102.

Lohwasser Daniela, Zhan Chen; 2001. Friction Stir Welding, From Basics to Applications. Woodhead Publishing.

López Víctor Ignacio, 2007. Revista Colombiana de Estadística. Una introducción a los diseños óptimos; Vol. 30, núm. 1; 37-51 0120-1751.

López-Ríos V., 2008. Diseños óptimos para discriminación y estimación en modelos no lineales Tesis Doctoral, Centro de Investigación en Matemáticas, CIMAT, México.

M.A. Moradpour; Multi-objective Optimization of Welding Parameters in Submerged Arc Welding of API X65 Steel Plates; Journal of Iron and Steel Research, International, Volume 22, Issue 9, September 2015, Pages 870-878.

McCullagh, P. and Nelder, J. A., 1989. Generalized Linear Models, 2nd Edition. London: Chapman and Hall.

Meeker y Hahn, 1977. Asymptotically optimum over-stress test to estimate the survival probability at a condition with a low expected failure probability. "Technometrics ", Vol. 19, Pág. 381-404.

Meneses P. y Augusto C., 2007. Optimización Multiobjetivo usando un algoritmo genético y un operador elitista basado en un ordenamiento no-dominado (NSGA-

ii). Scientia et Technia Año XIII, No. 35, Universidad Tecnológica de Pereira, ISSN 0122 – 1701.

Mohammad Hasan Shojaeefard et Al., 2014; Multi objective optimization of friction stir welding parameters using FEM and neural network. International Journal of Precision Engineering and Manufacturing November 2014, Volume 15, Issue 11, pp 2351-2356.

Murata Atsuo, et. al., 2015. Multinomial Logistic Regression Model for Predicting Driver's Drowsiness Using Behavioral Measures. Vol: 3, 2015, pág: 2426-2433. Okayama, Japón.

Norma del Instituto Americano Nacional de Estándares AWS D17.3 / D17.3M:2010; Specification for Friction Stir Welding of Aluminum Alloys for Aerospace Applications.

Ping Jiang et Al., 2016; Optimization of laser welding process parameters of stainless steel 316L using FEM, Kriging and NSGA-II; Advances in Engineering Software Volume 99, September 2016, Pages 147–160.

Reza Mokhtari Ahmad, 2014; Hydrothermal alteration mapping through multivariate logistic regression analysis of lithogeochemical data; Journal of Geochemical Exploration. Vol. 145, Pág. 207- 212.

Rodríguez Carmelo, Ortiz Isabel; 1999. Diseño Óptimo de Experimentos para Modelos de Regresión. Universidad de Almería, Ser vicio de Publicaciones.

Sebastiani Paola y Settini Raffaella, 1992. A Note on D-optimal designs for a logistic regression model. Vol. 59, Pág. 359 – 368.

Shumskii, A. E., 2005. "Parametrization of the Input Output Image of Nonlinear Dynamic Systems". Automation and Remote Control.

Sitter y Torsney, 1995. Optimal Designs for Binary Response Experiments with Two Design Variables. Statistica Sinica, Vol: 5, Pág. 405-419.

Tattar, Prabhanjan Narayanachar, Suresh Ramaiah, and B.G. Manjunath. "Generalized Linear Models", A Course in Statistics With R., 2016.

Téllez Piñerez y Cristian Fernando, 2013. Tesis de Grado, Diseño D-óptimo promediado por una distribución a priori: Metodología para incrementar el número de puntos experimentales.

Tulum Cem Celal et. Al., 2010. Hybrid search for faster production and safer process conditions in Friction Stir Welding. Chapter simulated evolution and learning volume 6457 of the series lecture notes in computer science pp 603-612.

Tutz Gerhard, Pöbnecker Wolfgang y Uhlmann Lorenz, 2015. Variable selection in general multinomial logit models. Vol: 82, pag: 207–222. München, Germany.

Yang Min, Zhang Bin, Huang Shuguang, 2011. Optimal Designs For Generalized Linear Models With Multiple Design Variables. Statistica Sinica, Vol. 21. Pág. 1415 – 1430.

Apéndice A

Demostraciones

1) Propiedad Reproductiva de la Chi – Cuadrada

$$M_x(t) = \left(\frac{1}{1-pt}\right)^\alpha$$

$$x = \sum y_i \quad y_i \sim \chi^2_n$$

$$M_x(t) = E[e^{tx}] = E[e^{t \sum y_i}]$$

$$= E[e^{ty_1} \cdot e^{ty_2} \dots e^{ty_n}]$$

$$= \prod_{i=1}^n E[e^{ty_i}]$$

$$= \left(\frac{1}{1-pt}\right)^{\alpha_1} \cdot \left(\frac{1}{1-pt}\right)^{\alpha_2} \dots \left(\frac{1}{1-pt}\right)^{\alpha_n}$$

$$= \left(\frac{1}{1-pt}\right)^{\sum \alpha_i}$$

Suponiendo que:

$$\beta = 2 ; \quad \alpha = \frac{n}{2} \quad \Rightarrow \quad \left(\frac{1}{1-pt} \right)^{\frac{1}{2}\sum n_i}$$

$$\therefore \sum x_i^2 \sim \chi^2_{(\sum n_i, \sum \lambda_i)}$$

2) Función de Probabilidad del caso binomial perteneciente a la familia exponencial

Se considera la variable aleatoria Bernoulli, se distribuye de la siguiente forma y de aquí partimos para determinar la función de probabilidad:

Variable aleatoria “y” que se distribuye de forma Bernoulli:

$$y \sim B(n, p)$$

$$f(y; \pi)$$

Aplicamos logaritmo en ambos lados de la ecuación:

$$\ln f(y; \pi) = \ln \binom{n}{p} \pi^y (1 - \pi)^{n-y}$$

Aplicamos propiedades de logaritmos y álgebra:

$$\begin{aligned} &= \ln \binom{n}{p} + \ln \pi^y + \ln(1 - \pi)^{n-y} \\ &= \ln \binom{n}{p} + y \ln \pi + (n - y) \ln(1 - \pi) \\ &= \ln \binom{n}{p} + y \ln \pi + n \ln(1 - \pi) - y \ln(1 - \pi) \\ &= y[\ln \pi - \ln(1 - \pi)] + n \ln(1 - \pi) + \ln \binom{n}{p} \\ &= y \ln \left(\frac{\pi}{1 - \pi} \right) + n \ln(1 - \pi) + \ln \binom{n}{p} \end{aligned}$$

Para terminar, llegamos a su función de probabilidad:

$$f(y; \pi) = \exp \left\langle y \ln \left(\frac{\pi}{1-\pi} \right) + n \ln(1-\pi) + \ln \binom{n}{p} \right\rangle$$

3) Valor esperado para la familia de distribuciones exponencial

Partimos de lo siguiente:

$$\int_y f(y; \theta) dy = 1$$

Se diferencian ambos lados de la ecuación con respecto a θ :

$$\frac{d}{d\theta} \left[\int_y f(y; \theta) dy \right] = \frac{d}{d\theta} (1) = 0$$

Se aplica la segunda derivada y se intercambia el orden de la integración y diferenciación:

$$\frac{d^2}{d\theta^2} \left[\int_y f(y; \theta) dy \right] \Rightarrow \int_y \frac{d^2 f(y; \theta)}{d\theta^2} dy = 0$$

Considerando los resultados para la familia exponencial:

$$f(y; \theta) = \exp \langle a(y)b(\theta) + c(\theta)d(y) \rangle$$

En consecuencia, la derivada de esta función es:

$$\frac{df(y; \theta)}{d\theta} = [a(y)b'(\theta) + c'(\theta)]f(y; \theta)$$

Si tomamos en cuenta que:

$$\int \frac{df(y; \theta)}{d\theta} dy = 0$$

Entonces:

$$\int [a(y)b'(\theta) + c'(\theta)]f(y; \theta) dy = 0$$

Simplificando:

$$\int f(y; \theta) a(y)b'(\theta)dy + c'(\theta) \int f(y; \theta) dy = 0$$

En la que:

$$\int f(y; \theta) a(y)dy = E[a(y)]$$

$$f(y; \theta) = 1$$

Reemplazando:

$$b'(\theta) E[a(y)] + c'(\theta) = 0$$

Por lo tanto, el valor esperado para todas las funciones de la familia exponencial es:

$$E[a(y)] = -\frac{c'(\theta)}{b'(\theta)}$$

4) Varianza para la familia exponencial:

Para definir la varianza, usaremos la segunda derivada que obtuvimos:

$$\begin{aligned} \frac{d^2 f(y; \theta)}{d\theta^2} &= \frac{d}{d\theta} [a(y)b'(\theta) + c'(\theta)]f(y; \theta) \\ &= [a(y)b'(\theta) + c'(\theta)][a(y)b'(\theta) + c'(\theta)]f(y; \theta) + f(y; \theta)[a(y)b''(\theta) + c''(\theta)] \\ &\quad f(y; \theta)[a(y)b'(\theta) + c'(\theta)]^2 + f(y; \theta)[a(y)b''(\theta) + c''(\theta)] \end{aligned}$$

Simplificando

$$f(y; \theta)[a(y)b'(\theta) + c'(\theta)]^2$$

Considerando que:

$$c'(\theta) = -b'(\theta)E[a(y)]$$

$$\langle a(y)b'(\theta) - b'(\theta)E[a(y)] \rangle^2 f(y; \theta)$$

$$(b'(\theta))^2 < a(y) - E[a(y)] >^2 f(y; \theta)$$

Entonces sustituimos las igualdades anteriores:

$$\begin{aligned} \frac{d^2 f(y; \theta)}{d\theta^2} &= f(y; \theta) a(y) b''(\theta) + f(y; \theta) c''(\theta) \\ &+ \dots [b'(\theta)]^2 < a(y) - E[a(y)] >^2 f(y; \theta) \end{aligned}$$

La varianza sería igual a:

$$\begin{aligned} Var &= \int < y - E[y] >^2 f(y) dy \\ \int \frac{d^2 f(y; \theta)}{d\theta^2} dy &= \int f(y; \theta) a(y) b''(\theta) dy + \dots + \int f(y; \theta) c''(\theta) dy \\ &+ \int [b'(\theta)]^2 < a(y) - E[a(y)] >^2 f(y; \theta) dy \end{aligned}$$

Siendo:

$$\begin{aligned} \Rightarrow b''(\theta) \int a(y) f(y; \theta) dy + c''(\theta) \int f(y; \theta) dy \\ + [b'(\theta)]^2 \int < a(y) - E[a(y)] >^2 f(y; \theta) dy \end{aligned}$$

Igualando lo siguiente de la ecuación:

$$\begin{aligned} \int a(y) f(y; \theta) dy &= E[a(y)] \\ \int f(y; \theta) dy &= 1 \\ \int < a(y) - E[a(y)] >^2 f(y; \theta) dy &= VAR[a(y)] \end{aligned}$$

Sustituyendo lo anterior en la ecuación:

$$\Rightarrow b''(\theta) E[a(y)] + c''(\theta) + [b'(\theta)]^2 VAR[a(y)] = 0$$

$$VAR[a(y)] = -\frac{b''(\theta) E[a(y)] + c''(\theta)}{[b'(\theta)]^2}$$

Sustituimos el valor esperado $E[a(y)] = -\frac{c'(\theta)}{b'(\theta)}$ en la ecuación:

$$\begin{aligned} VAR[a(y)] &= \frac{-b''(\theta) \left(-\frac{c'(\theta)}{b'(\theta)}\right) - c''(\theta)}{[b'(\theta)]^2} \\ &= \frac{b''(\theta) (c'(\theta))}{[b'(\theta)]^3} - \frac{c''(\theta)}{[b'(\theta)]^2} \end{aligned}$$

Esta es la varianza para las distribuciones de la familia exponencial:

$$VAR[a(y)] = \frac{b''(\theta) (c'(\theta)) - b'(\theta)c''(\theta)}{[b'(\theta)]^3}$$

5) Valor esperado para la “función score”

La función Score tiene la siguiente forma:

$$U = a(y)b'(\theta) + c'(\theta)$$

Aplicando el valor esperado para la ecuación, es:

$$E[U] = E[a(y)b'(\theta) + c'(\theta)]$$

$$E[U] = b'(\theta)E[a(y)] + E[c'(\theta)]$$

Considerando que:

$$E[a(y)] = -\frac{c'(\theta)}{b'(\theta)}$$

Entonces el valor esperado para $E[U]$ es:

$$E[U] = b'(\theta) \left(-\frac{c'(\theta)}{b'(\theta)}\right) + c'(\theta)$$

$$E[U] = -c'(\theta) + c'(\theta)$$

$$E[U] = 0$$

Varianza de la "función score"

$$Var[U] = Var[a(y)b'(\theta) + c'(\theta)]$$

Si tomamos en cuenta que $Var[y] = \langle y - E[y] \rangle^2$, entonces:

$$Var[U] = \langle [a(y)b'(\theta) + c'(\theta)] - E[a(y)b'(\theta) + c'(\theta)] \rangle^2$$

$$Var[U] = \langle a(y)b'(\theta) + c'(\theta) - c'(\theta) - b'(\theta)E[a(y)] \rangle^2$$

$$= \langle a(y)b'(\theta) - b'(\theta)E[a(y)] \rangle^2$$

$$= \langle b'(\theta)(a(y) - E[a(y)]) \rangle^2$$

Hay que recordar la siguiente igualdad:

$$(a(y) - E[a(y)])^2 = Var[a(y)]$$

$$Var[U] = [b'(\theta)]^2 Var[a(y)]$$

Notar que la $Var[U]$ es llamada la Matriz de Información "J"

$$J = Var[U] = [b'(\theta)]^2 Var[a(y)]$$

Recordar que:

$$Var[U] = \frac{b''(\theta)c'(\theta)}{b'(\theta)} - c''(\theta)$$

Otra propiedad que tiene U es:

$$Var[U] = E[U^2] - E^2[U]$$

Como $E^2[U] = 0$, entonces:

$$Var[U] = E[U^2]$$

6) Estimación de los parámetros por máxima verosimilitud:

La función de log verosimilitud para todas las y_{is} es:

$$l = \sum_{i=1}^n l_i = \sum_{i=1}^n y_i b_i(\theta_i) + \sum_{i=1}^n c_i(\theta_i) + \sum_{i=1}^n d_i(y_i)$$

Si tomamos en cuenta que: $\frac{dl}{d\theta_i}$ es decir, U_j

$$U_j = \sum_{i=1}^n \left[\frac{dl}{d\theta_i} \right]$$

$$\eta_i = X^T \beta$$

$$g(\mu) = \eta_i$$

$$\mu_i = \theta_i$$

Entonces, por la regla de la cadena obtenemos:

$$\sum_{i=1}^n \left[\frac{dl}{d\theta_i} \right] = \sum_{i=1}^n \left[\frac{dl}{d\theta_i} \cdot \frac{d\theta_i}{d\mu_i} \cdot \frac{d\mu_i}{d\beta_i} \right]$$

$$[I] \Rightarrow \frac{dl_i}{d\theta_i} = \frac{d}{d\theta_i} [y_i b_i(\theta_i) + c_i(\theta_i) + d_i(y_i)]$$

Considerando que:

$$\mu_i = -\frac{c'(\theta_i)}{b'(\theta_i)}$$

Entonces:

$$\frac{dl_i}{d\theta_i} = y_i b_i'(\theta_i) - b_i'(\theta_i) \mu_i = b_i'(\theta_i) [y_i - \mu_i]$$

$$[II] \Rightarrow \frac{d\theta_i}{d\mu_i} : \text{Es el inverso de } \frac{d\mu_i}{d\theta_i}$$

$$\frac{d\theta_i}{d\mu_i} = \frac{1}{\frac{d\mu_i}{d\theta_i}}$$

Esto es:

$$\frac{dE[y_i]}{d\theta_i} = \frac{d\mu_i}{d\theta_i}$$

$$E[y] = -c'(\theta_i)[b'(\theta_i)]^{-1}$$

$$\frac{dE[y_i]}{d\theta_i} = -\langle -c'(\theta_i)[b'(\theta_i)]^{-2}b''(\theta_i) + [b'(\theta_i)]^{-1}c''(\theta_i) \rangle$$

$$= \frac{c'(\theta_i)b''(\theta_i)}{[b'(\theta_i)]^2} - \frac{c''(\theta_i)}{[b'(\theta_i)]}$$

$$= \frac{b'(\theta_i)c'(\theta_i)b''(\theta_i) - [b'(\theta_i)]^2c''(\theta_i)}{[b'(\theta_i)]^3}$$

$$= \frac{b'(\theta_i)[c'(\theta_i)b''(\theta_i) - b'(\theta_i)c''(\theta_i)]}{[b'(\theta_i)]^3}$$

$$\Rightarrow \frac{dl_i}{d\theta_i} = b'(\theta_i)Var(y_i)$$

$$\frac{d\theta_i}{d\mu_i} = \frac{1}{b'(\theta_i)Var(y_i)}$$

$$(III) \Rightarrow \frac{d\mu_i}{d\beta_j} = \frac{d\mu_i}{d\eta_i} \cdot \frac{d\eta_i}{d\beta_j} \rightarrow \eta_i = X^T \beta$$

$$\Rightarrow \frac{d\mu_i}{d\eta_i} = (x_{ij})$$

$$\frac{d\eta_i}{d\beta_j} = X_i^T = \langle x_{ij} \rangle \text{ "Elemento Genérico"}$$

Como:

$$U_j = \sum_{i=1}^n \left[\frac{dl_i}{d\theta_i} \cdot \frac{d\theta_i}{d\mu_i} \cdot \frac{d\mu_i}{d\beta_j} \right]$$

$$= \sum_{i=1}^n \left[b'(\theta_i)[y_i - \mu_i] \left(\frac{1}{b'(\theta_i)Var(y_i)} \right) \left(\frac{d\mu_i}{d\eta_i} x_{ij} \right) \right]$$

Esta es la primera derivada **"The Score"**:

$$U_j = \sum_{i=1}^n \left[\frac{[y_i - \mu_i]}{\text{Var}(y_i)} x_{ij} \left(\frac{d\mu_i}{d\eta_i} \right) \right]$$

Entonces el valor esperado y la varianza de U_i es:

$$E[U(\theta)] = 0$$

$$\text{Var}[U_j(\theta)] = E[U(\theta)U'(\theta)] = I(\theta)$$

$$J_{jk} = E[U_i U_k]$$

Bajo condiciones de irregularidad, la matriz de información observada es:

$$I(\theta) = -E \left[\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

Entonces, J es la matriz de segundas derivadas:

$$J_{jk} = E \left\langle \sum_{i=1}^n \left[\frac{[y_i - \mu_i]}{\text{Var}(y_i)} x_{ij} \left(\frac{d\mu_i}{d\eta_i} \right) \right] \sum_{i=1}^n \left[\frac{[y_i - \mu_i]}{\text{Var}(y_i)} x_{ik} \left(\frac{d\mu_i}{d\eta_i} \right) \right] \right\rangle$$

Para $l = i$; con fines de subíndice:

$$= E \left\langle \sum_{i=1}^n \left[\frac{[y_i - \mu_i][y_i - \mu_i]}{\text{Var}(y_i)\text{Var}(y_i)} x_{ij} x_{ik} \left(\frac{d\mu_i}{d\eta_i} \right) \left(\frac{d\mu_i}{d\eta_i} \right) \right] \right\rangle$$

$$= E \left\langle \sum_{i=1}^n \left[\frac{[y_i - \mu_i]^2}{\text{Var}(y_i)^2} x_{ij} x_{ik} \left(\frac{d\mu_i}{d\eta_i} \right)^2 \right] \right\rangle$$

$$= \sum_{i=1}^n \left[\frac{E[y_i - \mu_i]^2}{\text{Var}(y_i)^2} x_{ij} x_{ik} \left(\frac{d\mu_i}{d\eta_i} \right)^2 \right]$$

$$= \sum_{i=1}^n \left[\frac{\text{Var}(y_i)}{\text{Var}(y_i)^2} x_{ij} x_{ik} \left(\frac{d\mu_i}{d\eta_i} \right)^2 \right]$$

Por lo tanto, para calcular la matriz de segundas derivadas quedaría de la siguiente forma:

$$J_{jk} = \sum_{i=1}^n \left[\frac{x_{ij}x_{ik}}{\text{Var}(y_i)^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2 \right]$$

7) Estimación de scoring para parámetros beta

Utilizando el método de Estimación de Scoring para la matriz de segundas derivadas y obtener los parámetros beta, partimos de lo siguiente:

$$\theta^m = \theta^{m-1} + \frac{U^{m-1}}{J^{m-1}}$$

En general:

$$b^m = b^{m-1} + [J^{m-1}]^{-1}U^{m-1}$$

Multiplicando ambos lados de la ecuación por $[J^{m-1}]$:

$$J^{m-1}b^m = J^{m-1}b^{m-1} + J^{m-1}[J^{m-1}]^{-1}U^{m-1}$$

$$\Rightarrow (I) \quad J^{m-1}b^m = J^{m-1}b^{m-1} + U^{m-1}$$

Recordemos que, para la matriz de segundas derivadas J:

$$J_{ik} = \sum_{i=1}^n \left[\frac{x_{ij}x_{ik}}{\text{Var}(y_i)^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2 \right]$$

Los elementos de esta ecuación son:

$$x_{ij} = x^T \quad x_{ik} = x$$

$$W = \frac{1}{\text{Var}(y_i)^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2$$

Por lo tanto, queda de la siguiente forma:

$$J = x^T W x$$

Los términos $J^{m-1}b^{m-1} + U^{m-1}$ pueden ser escritos de la siguiente forma:

$$J^{m-1}b^{m-1} + U^{m-1} = x^T W Z$$

Dado que:

$$x^T : \sum_{i=1}^n x_{ij} + \sum_{i=1}^n x_{ij}$$

$$W_{ij} = \frac{1}{\text{Var}(y_i)^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2$$

$$Z_j = \sum_{i=1}^n x_{ik} b^{m-1} + (y_i - \mu_i) \frac{d\mu_i}{d\eta_i}$$

En la misma expresión (I), el lado derecho se puede reescribir de la siguiente manera:

$$\sum_{i=1}^n \frac{x_{ij}x_{ik}}{\text{Var}(y_i)^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2 b^{m-1} + \frac{y_i - \mu_i}{\text{Var}(y_i)} \left(\frac{d\mu_i}{d\eta_i} \right)$$

Esto es para:

$$X_{n \times p}; b_{p \times 1} \text{ vector de parámetros}$$

Entonces, si lo generalizamos para “p” covariables:

$$\sum_{k=1}^p \left[\sum_{i=1}^n \frac{x_{ij}x_{ik}}{\text{Var}(y_i)^2} \left(\frac{d\mu_i}{d\eta_i} \right)^2 b^{m-1} + \frac{y_i - \mu_i}{\text{Var}(y_i)} \left(\frac{d\mu_i}{d\eta_i} \right) \right]$$

El lado izquierdo de (I) $J^{m-1}b^m$, será de la siguiente manera:

$$x^T W x b^m$$

Por lo tanto:

$$x^T W x b^m = x^T W Z$$

Es la ecuación de los “**mínimos cuadrados ponderados iterativos**”.
Despejando b^m de la ecuación para obtener el valor de los parámetros, queda:

$$b^m = (x^T W x)^{-1} x^T W Z$$

8) Estadístico score para caso binomial

Partimos de la distribución Bernoulli para la binomial:

$$Y \sim Bin(n, \pi)$$

$$f(y) = \binom{n}{y} \pi^y (1 - \pi)^{n-y}$$

$$L(\pi|y_i) = \prod_{i=1}^n \binom{n}{y} \pi^y (1 - \pi)^{n-y}$$

Al haber aplicado logaritmo a su función de densidad, podemos llegar a dicho estadístico:

$$= \log \binom{n}{y} + y \log(\pi^y) + (n - y) \log((1 - \pi))$$

$$\frac{dl}{d\pi} = \frac{y}{\pi} - \frac{n}{1 - \pi} + \frac{y}{1 - \pi} = \frac{y}{\pi} - \frac{n - y}{1 - \pi}$$

$$= \frac{y(1 - \pi) - (n - y)\pi}{\pi(1 - \pi)}$$

$$= \frac{y - n\pi}{\pi(1 - \pi)} \Rightarrow U$$

9) Valor esperado y varianza para el caso binomial

Valor esperado:

$$E[U] = 0$$

$$E[U] = \frac{y - n\pi}{\pi(1 - \pi)} = \frac{1}{\pi(1 - \pi)} [y - n\pi]$$

$$= \frac{1}{\pi(1 - \pi)} \langle E[y] - E[n\pi] \rangle$$

$$= \frac{1}{\pi(1 - \pi)} (n\pi - n\pi) = 0$$

Varianza:

$$\begin{aligned}
 \text{Var}(U) &= \text{Var}\left(\frac{y - n\pi}{\pi(1 - \pi)}\right) \\
 &= \frac{1}{(\pi(1 - \pi))^2} < \text{Var}(y) - \text{Var}(n\pi) > \\
 &= \frac{n\pi(1 - \pi)}{(\pi(1 - \pi))^2} = \frac{n}{\pi(1 - \pi)} \Rightarrow J \\
 \frac{u}{\sqrt{J}} &\sim N(0,1) \\
 \frac{\frac{y - n\pi}{\pi(1 - \pi)}}{\sqrt{\frac{n}{\pi(1 - \pi)}}} &= \frac{(y - n\pi)(\pi(1 - \pi))^{1/2}}{(\pi(1 - \pi))n^{1/2}} \\
 &= \frac{y - n\pi}{\sqrt{n}\sqrt{\pi(1 - \pi)}} = \frac{y - n\pi}{\sqrt{n\pi(1 - \pi)}} \sim N(0,1)
 \end{aligned}$$

Elevando al cuadrado sigue una distribución chi – cuadrada y queda así:

$$\frac{(y - n\pi)^2}{n\pi(1 - \pi)} \sim \chi^2_{(1)}$$

10)Aproximaciones por serie de Taylor

Para la inferencia, se requiere un estadístico para determinar la significancia de cada parámetro del modelo, partiendo de la serie de Taylor:

$$f(x) = f(t) + (x - t) \left[\frac{df}{dx} \right]_{x=t} + \frac{1}{2} (x - t)^2 \left[\frac{d^2f}{dx^2} \right]_{x=t} + \dots$$

Para el caso de la función de Log verosimilitud de un parámetro único:

$$l_{\beta} = l_{(b)} + (\beta - b)U(b) + \frac{1}{2}(\beta - b)^2 U'(b)$$

Con:

$$U(b) = \frac{dl}{d\beta} ; \quad U(b) = \frac{d^2l}{d\beta^2} \cong -J$$

Obtenemos lo siguiente al sustituir:

$$l_{\beta} = l_{(b)} + (\beta - b)U(b) - \frac{1}{2}(\beta - b)^2J$$

Sea un vector de parámetros β :

$$l_{\beta} = l_{(b)} + (\beta - b)^T U(b) - \frac{1}{2}(\beta - b)^T J(b)(\beta - b)$$

Si aplicamos para la función Score:

$$U(\beta) = U(b) + (\beta - b)^T U'(b)$$

Con:

$$U'(b) = -J(b)$$

$$U(\beta) = U(b) - (\beta - b)J(b)$$

$$\hat{\beta} = b$$

Sustituyendo para la función Score anterior:

$$U(\beta) = 0 - (\beta - b)J(b)$$

$$U(\beta) = -(\beta - b)J(b)$$

$J(b) = \text{Matriz de Información} : \text{Tiende a ser una constante}$

$$U(\beta) = J(b - \beta)$$

$$(b - \beta) = J^{-1}U$$

$$E[b - \beta] = E[J^{-1}U]$$

$$E[b - \beta] = J^{-1}E[U]$$

Como $E[U] = 0$, entonces:

$$E[b - \beta] = 0$$

La matriz de varianzas y covarianzas para b:

$$J_{jk} = E[U_j U_k] \text{ o } J = E[UU^T]$$

Entonces:

$$UU^T = J(b - \beta)(b - \beta)^T J^T$$

Despejando para $(b - \beta)$:

$$(J^T)^{-1}(UU^T)(J)^{-1} = (b - \beta)(b - \beta)^T$$

Tomando su valor esperado:

$$E[(b - \beta)(b - \beta)^T] = E[(J^T)^{-1}(UU^T)(J)^{-1}]$$

$$= J^{-1}E[UU^T]J^{-1} \quad ; \quad E[UU^T] = J$$

$$J^{-1} = IJ^{-1}$$

$$E[(b - \beta)(b - \beta)^T] = J^{-1}$$

$$(b - \beta)^T J (b - \beta) \sim \chi_{(p)}^2$$

Lo anterior es el llamado “Estadístico de Wald”, que se utiliza como variante de la prueba ANOVA para la significancia de los parámetros.

11) Devianza

Es el estadístico comúnmente utilizado para determinar qué tan bueno es el modelo, para valores grandes de este estadístico, el modelo es pobre.

$$D = 2[l(b_{max}; y) - l(b; y)]$$

Y la distribución de este estadístico sería de la siguiente forma:

$$l(\beta) = l(b) + (\beta - b)^T U(b) - \frac{1}{2}(\beta - b)^T J(b)(\beta - b)$$

$$U(b) = 0$$

$$l(\beta) = l(b) - \frac{1}{2}(\beta - b)^T J(b)(\beta - b)$$

$$l(\beta) - l(b) = -\frac{1}{2}(\beta - b)^T J(b)(\beta - b)$$

$$-2[l(\beta) - l(b)] = (\beta - b)^T J(b)(\beta - b)$$

$$2[l(b) - l(\beta)] = (\beta - b)^T J(b)(\beta - b)$$

Por lo tanto, la distribución del estadístico D:

$$D \sim \chi_{(p)}^2$$

Nota: En el caso de la Binomial y Poisson, "D" puede ser utilizado como bondad de ajuste.

12) Devianza para el caso Binomial

Partiendo de que tiene una distribución:

$$y \sim Bin(n_i, \pi_i)$$

$$l(\beta, y) = \sum_{i=1}^n \left[y_i \log(\pi_i) - y_i \log(1 - \pi_i) + n_i \log(1 - \pi_i) + \log \binom{n_i}{y_i} \right]$$

$$\frac{dl}{d\pi} = \frac{y - n\pi}{\pi(1 - \pi)} = 0$$

$$\Rightarrow y - n\pi = [\pi(1 - \pi)](0)$$

$$y - n\pi = 0$$

$$y = n\pi$$

$$\therefore \hat{\pi} = \frac{y}{n}$$

$$(1 - \pi_i) = \left(1 - \frac{y_i}{n_i}\right) = \frac{n_i - y_i}{n_i}$$

Entonces, sustituyendo $(1 - \pi_i)$ en la ecuación para el máximo valor de $l(\beta, y)$, queda:

$$l(\beta_{max}, y_i) = \sum \left[y_i \log \left(\frac{y_i}{n_i} \right) - y_i \log \left(\frac{n_i - y_i}{n_i} \right) + n_i \log \left(\frac{n_i - y_i}{n_i} \right) + \log \left(\frac{n_i}{y_i} \right) \right]$$

Para una y estimada:

$$\hat{y}_i = \hat{\pi} n_i$$

Entonces:

$$l(b; y) = \sum \left[y_i \log \left(\frac{\hat{y}_i}{n_i} \right) - y_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) + n_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) + \log \left(\frac{n_i}{y_i} \right) \right]$$

Por lo tanto, la Devianza será:

$$D = 2 \left\langle \sum \left[y_i \log \left(\frac{y_i}{n_i} \right) - y_i \log \left(\frac{n_i - y_i}{n_i} \right) + n_i \log \left(\frac{n_i - y_i}{n_i} \right) + \log \left(\frac{n_i}{y_i} \right) \right] \right. \\ \left. - \sum \left[y_i \log \left(\frac{\hat{y}_i}{n_i} \right) - y_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) + n_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) + \log \left(\frac{n_i}{y_i} \right) \right] \right\rangle$$

Uniendo todo en una sola sumatoria:

$$D = 2 \left\langle \sum \left[y_i \log \left(\frac{y_i}{n_i} \right) - y_i \log \left(\frac{\hat{y}_i}{n_i} \right) + y_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) \right. \right. \\ \left. \left. - y_i \log \left(\frac{n_i - y_i}{n_i} \right) + n_i \log \left(\frac{n_i - y_i}{n_i} \right) - n_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) \right] \right\rangle$$

$$D = 2 \left\langle \sum \left[y_i \left(\log \left(\frac{y_i}{n_i} \right) - \log \left(\frac{\hat{y}_i}{n_i} \right) \right) + y_i \left(\log \left(\frac{n_i - \hat{y}_i}{n_i} \right) - \log \left(\frac{n_i - y_i}{n_i} \right) \right) \right. \right. \\ \left. \left. + n_i \left(\log \left(\frac{n_i - y_i}{n_i} \right) - \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) \right) \right] \right\rangle$$

$$D = \left\langle \sum \left[y_i \log \left(\frac{\left(\frac{y_i}{n_i} \right)}{\left(\frac{\hat{y}_i}{n_i} \right)} \right) + (n_i - y_i) \left[\log \left(\frac{n_i - y_i}{n_i} \right) - \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) \right] \right] \right\rangle$$

$$D = \left\langle \sum \left[y_i \log \left(\frac{y_i}{\hat{y}_i} \right) + (n_i - y_i) \left[\log \frac{\left(\frac{n_i - y_i}{n_i} \right)}{\left(\frac{n_i - \hat{y}_i}{n_i} \right)} \right] \right] \right\rangle$$

$$D = \left\langle \sum \left[y_i \log \left(\frac{y_i}{\hat{y}_i} \right) + (n_i - y_i) \log \left(\frac{n_i(n_i - y_i)}{n_i(n_i - \hat{y}_i)} \right) \right] \right\rangle$$

$$D = \left\langle \sum \left[y_i \log \left(\frac{y_i}{\hat{y}_i} \right) + (n_i - y_i) \log \left(\frac{n_i - y_i}{n_i - \hat{y}_i} \right) \right] \right\rangle$$

Esta es la función de Devianza para el modelo Binomial.

13) Solución de la ecuación diferencial ordinaria para el modelo logístico:

Partiendo de la ecuación diferencial ordinaria:

$$\frac{df}{dx} = af(\alpha - f)$$

Utilizando variables separadas:

$$\frac{df}{f(\alpha - f)} = a dx$$

Integramos por ambos lados:

$$\int \frac{df}{f(\alpha - f)} = a \int dx$$

$$(I) \quad (II)$$

Obteniendo la solución para (II): $a \int dx = ax + c$

Solución para (I): Integrar por fracciones parciales:

$$\begin{aligned} \int \frac{df}{f(\alpha - f)} &= \frac{1}{f(\alpha - f)} = \frac{u}{f} + \frac{v}{(\alpha - f)} \\ &= \frac{u(\alpha - f) + vf}{f(\alpha - f)} \end{aligned}$$

$$1 = u(\alpha - f) + vf$$

$$1 = u\alpha - uf + vf$$

$$1 = u\alpha - f(u - v)$$

Como es respecto a “f”:

$$f = 0 ; (u - v) = 0$$

Sustituyendo para “u”:

$$1 = u\alpha = 0$$

$$u = \frac{1}{\alpha}$$

Sustituyendo para “v”:

$$(u - v) = 0$$

$$\frac{1}{\alpha} - v = 0$$

$$v = \frac{1}{\alpha}$$

Sustituyendo para la integral inicial:

$$\int \frac{df}{f(\alpha - f)} = \frac{u}{f} + \frac{v}{(\alpha - f)}$$

$$\int \frac{df}{f(\alpha - f)} = \frac{1}{\alpha} \int \frac{df}{f} + \frac{1}{\alpha} \int \frac{df}{(\alpha - f)}$$

$$= \frac{1}{\alpha} \left[\int \frac{df}{f} + \int \frac{df}{(\alpha - f)} \right]$$

$$= \frac{1}{\alpha} [\ln(f) - \ln(\alpha - f)]$$

$$\int \frac{df}{f(\alpha - f)} = \frac{1}{\alpha} \ln\left(\frac{f}{\alpha - f}\right) + c$$

Encontrando la solución general uniendo (I) y (II):

$$\frac{1}{\alpha} \ln\left(\frac{f}{\alpha - f}\right) = ax + c$$

$$\ln\left(\frac{f}{\alpha - f}\right) = \alpha(ax) + \alpha c$$

$$\frac{f}{\alpha - f} = e^{\alpha(ax) + \alpha c}$$

Despejando para "f":

$$f = (\alpha - f)e^{\alpha(ax) + \alpha c}$$

$$f = \alpha e^{\alpha(ax) + \alpha c} - f e^{\alpha(ax) + \alpha c}$$

$$f + f e^{\alpha(ax) + \alpha c} = \alpha e^{\alpha(ax) + \alpha c}$$

$$f(1 + e^{\alpha(ax) + \alpha c}) = \alpha e^{\alpha(ax) + \alpha c}$$

$$f = \frac{\alpha e^{\alpha(ax) + \alpha c}}{(1 + e^{\alpha(ax) + \alpha c})}$$

$$f = (\alpha e^{\alpha(ax) + \alpha c})(1 + e^{\alpha(ax) + \alpha c})^{-1}$$

$$f = \alpha e^{\alpha(ax) + \alpha c} + \alpha e^{\alpha(ax) + \alpha c} e^{-\alpha(ax) - \alpha c}$$

Notar que:

$$e^{\alpha(ax) + \alpha c} e^{-\alpha(ax) - \alpha c} = 1$$

Entonces:

$$f = \alpha e^{\alpha(ax) + \alpha c} + \alpha$$

$$f = \alpha(e^{\alpha(ax) + \alpha c} + 1)$$

$$f = \frac{\alpha}{1 + e^{-\alpha(ax) - \alpha c}}$$

$$f = \frac{\alpha}{1 + e^{-\alpha(ax+c)}}$$

Para $\alpha = 1$:

$$f = \frac{1}{1 + e^{-(ax+c)}}$$

Como "a" es el parámetro de escala:

$$ax - c = X_d\beta$$

$$ax = X_d\beta + c$$

$$a = \frac{X_d\beta + c}{x}$$

Sustituimos en lo anterior:

$$f = \frac{1}{1 + e^{-\left(\left(\frac{X_d\beta+c}{x}\right)x-c\right)}}$$

$$f = \frac{1}{1 + e^{-(X_d\beta)}}$$

$$f = \frac{1}{1 + e^{-x\beta}} \Rightarrow \pi$$

Es la función de Probabilidad de Éxito, escrita de una mejor manera como:

$$\frac{1}{1 + e^{-x\beta}} = \pi$$

Si tomamos en cuenta la posibilidad de fracaso, podemos obtener la función decreciente de Probabilidad de Fracaso:

$$1 - \pi = \frac{e^{-x\beta}}{1 + e^{-x\beta}}$$

14) Función de probabilidad

Entonces para la función de probabilidad:

$$\pi \int_{-\infty}^x f(s) ds = \frac{e^{x\beta}}{1 + e^{x\beta}}$$

Utilizamos $\frac{e^{-X\beta}}{1+e^{-X\beta}}$ para la obtención de $f(s)$ y se recomienda lo siguiente:

$$\frac{e^{\beta_1+\beta_2x}}{1+e^{\beta_1+\beta_2x}}$$

$$f(s) = \frac{d\pi(x)}{dx} = \frac{d}{dx} \left[\frac{e^{\beta_1+\beta_2x}}{1+e^{\beta_1+\beta_2x}} \right]$$

$$f(s) = \frac{\beta_2 e^{\beta_1+\beta_2x}}{[1+e^{\beta_1+\beta_2x}]^2}$$

Recordar que para la forma de la distribución de la familia exponencial, en el caso de la binomial:

$$a(y) = y \quad (\text{Canónica})$$

$$b(\pi) = \ln\left(\frac{\pi}{1-\pi}\right)$$

$$c(\pi) = n \ln(1-\pi)$$

$$d(y) = \ln\binom{n}{y}$$

Usar el parámetro natural para establecer la función de enlace de esta distribución:

$$b(\pi) = \ln\left(\frac{\pi}{1-\pi}\right)$$

Sustituimos para " π " y " $1-\pi$ " encontrados anteriormente:

$$b(\pi) = \ln\left(\frac{\frac{e^{X\beta}}{1+e^{X\beta}}}{1-\frac{e^{X\beta}}{1+e^{X\beta}}}\right)$$

$$b(\pi) = \ln\left(\frac{\frac{e^{X\beta}}{1+e^{X\beta}}}{\frac{1+e^{X\beta}-e^{X\beta}}{1+e^{X\beta}}}\right)$$

$$b(\pi) = \ln\left(\frac{e^{X\beta}(1 + e^{X\beta})}{(1 + e^{X\beta})(1 + e^{X\beta} - e^{X\beta})}\right)$$

$$b(\pi) = \ln\left(\frac{e^{X\beta}}{(1 + e^{X\beta} - e^{X\beta})}\right)$$

$$b(\pi) = \ln(e^{X\beta})$$

$$b(\pi) = X\beta$$

Entonces, la función de enlace para el caso Binomial es:

$$\log\left[\frac{\pi}{1 - \pi}\right] = X\beta$$