

**DESARROLLO DE UN ENTORNO PROGRAMADO PARA EL ANÁLISIS
DE CONFIABILIDAD CON BASE EN DISTRIBUCIÓN WEIBULL
BIPARAMÉTRICA**

ANDRES FELIPE CASTAÑO REYES

LUIS MIGUEL MONROY OSORIO

UNIVERSIDAD TECNOLÓGICA DE PEREIRA

FACULTAD DE INGENIERÍAS

PROGRAMA DE INGENIERÍA ELÉCTRICA

PEREIRA, 2011

**DESARROLLO DE UN ENTORNO PROGRAMADO PARA EL ANÁLISIS
DE CONFIABILIDAD CON BASE EN DISTRIBUCIÓN WEIBULL
BIPARAMÉTRICA**

**ANDRES FELIPE CASTAÑO REYES
LUIS MIGUEL MONROY OSORIO**

**Proyecto de grado para optar por el título de
Ingeniero Electricista**

Director

M.Sc MAURICIO HOLGUÍN LONDOÑO

UNIVERSIDAD TECNOLÓGICA DE PEREIRA

FACULTAD DE INGENIERÍAS

PROGRAMA DE INGENIERÍA ELÉCTRICA

PEREIRA, 2011

AGRADECIMIENTOS

A nuestra familia por su apoyo constante, su preocupación, entrega y dedicación.

A nuestro director M.Sc Mauricio Holguín Londoño, por sus enseñanzas, consejos y participación en la ejecución de este proyecto.

A todos quienes participaron de este proceso.

TABLA DE CONTENIDO

| | | |
|-------|--|----|
| 1. | INTRODUCCIÓN | 1 |
| 1.1 | DESCRIPCIÓN DEL PROBLEMA | 2 |
| 1.2 | JUSTIFICACIÓN | 4 |
| 1.3 | OBJETIVOS | 6 |
| 1.3.1 | OBJETIVO GENERAL..... | 6 |
| 1.3.2 | OBJETIVOS ESPECÍFICOS | 6 |
| 2. | FORMULACIÓN MATEMÁTICA | 7 |
| 2.1 | VARIABLES ALEATORIAS CONTINUAS Y DISTRIBUCIONES DE PROBABILIDAD | 7 |
| 2.2 | DISTRIBUCIÓN DE PROBABILIDAD Ó LEYES DE FALLA | 9 |
| 2.3 | ESTIMADORES DE PARÁMETROS PARA ALGUNAS DISTRIBUCIONES | 14 |
| 2.4 | PRUEBAS DE BONDAD DE AJUSTE..... | 15 |
| 2.5 | NIVELES DE CONFIANZA..... | 22 |
| 3. | METODOLOGÍA PROPUESTA | 24 |
| 3.1 | RECOLECCIÓN DE LA INFORMACIÓN | 24 |
| 3.2 | TRATAMIENTO DE LOS DATOS | 24 |
| 3.3 | PROCEDIMIENTO DEL MODELO..... | 27 |
| 3.4 | INTERFAZ DE USUARIO..... | 33 |
| 4. | RESULTADOS Y ANÁLISIS | 39 |
| 4.1 | RESULTADOS INICIALES..... | 39 |
| 4.2 | RESULTADOS FINALES..... | 47 |
| 5. | CONCLUSIONES | 56 |
| 6. | BIBLIOGRAFÍA | 58 |
| | ANEXOS..... | 61 |

ÍNDICE DE ILUSTRACIONES

| | |
|--|----|
| Ilustración 1. Función densidad de probabilidad | 7 |
| Ilustración 2. Función de distribución acumulativa | 8 |
| Ilustración 3. Función densidad de probabilidad de la distribución exponencial para diferentes valores de λ | 9 |
| Ilustración 4. Función densidad de probabilidad para la distribución Rayleigh para diferentes valores de α | 11 |
| Ilustración 5. Función densidad de probabilidad para la distribución Weibull para diferentes valores de α y β | 13 |
| Ilustración 6. Interfaz de Usuario..... | 34 |
| Ilustración 7. Mensaje aproximación distribución Exponencial | 35 |
| Ilustración 8. Mensaje aproximación distribución Rayleigh..... | 35 |
| Ilustración 9. Mensaje de advertencia para tiempos negativos | 36 |
| Ilustración 10. Mensaje de advertencia para ingreso de confiabilidad negativa..... | 37 |
| Ilustración 11. Mensaje de advertencia cuando se ingresa una Confiabilidad mayor que 1. | 37 |
| Ilustración 12. Presentación de los resultados mediante la interfaz gráfica..... | 38 |

ÍNDICE DE TABLAS

| | |
|---|----|
| Tabla 1. Estimadores de máxima verosimilitud | 14 |
| Tabla 2. Números aleatorios oscilando entre 0 y 1 | 17 |
| Tabla 3. Datos seleccionados para aplicar Kolmogorov-Smirnov | 17 |
| Tabla 4. Datos organizados de menor a mayor para prueba KS | 18 |
| Tabla 5. Resultados finales para diferentes tipos de distribuciones Weibull | 49 |
| Tabla 6. Errores entre el valor real y el calculado de los parámetros α y β | 53 |
| Tabla 7. Resultados para distribuciones que no pueden ser modeladas por la distribución Weibull..... | 55 |

ANEXOS

| | |
|---|----|
| Anexo 1. Distribución del estadístico de Kolmogorov-Smirnov | 61 |
| Anexo 2. Tabla de distribución Chi-Cuadrado | 62 |
| Anexo 3. Código fuente entorno programado | 70 |

1. INTRODUCCIÓN

La ejecución en las labores de mantenimiento en una empresa, son a menudo permeadas por la visión empírica que posee el estamento encargado de las labores de mantenimiento, dando paso a prácticas deficientes que conllevan a la disminución del ciclo útil de las máquinas.

Se hace necesario considerar las características principales y la forma en la cual los equipos están desarrollando sus funciones, debido a que el conocimiento o desconocimiento de estos elementos puede llevar a la correcta y clara anticipación ante posibles fallas, que en forma gradual fomenten un desgaste y un daño irreparable en las máquinas.

Lo cual hace necesario implementar planes de conservación para los datos que contienen la información de fallas en los equipos, específicamente los tiempos en los cuales se presentaron y así a partir de estos registros entrar a analizar de manera precisa la confiabilidad que poseen los equipos, especialmente el tiempo en el cual entraría a fallar, sin la obligación de orientarse respecto a la visión experimental del organismo encargado de las funciones de mantenimiento en la empresa.

A partir de bases de datos generadas aleatoriamente, así como registros históricos en los cuales se consideran los tiempos de falla de ciertos equipos, se desarrolla e implementa un entorno programado para el análisis de confiabilidad con base en la distribución Weibull biparamétrica, otorgando de este modo suficientes herramientas que permitan considerar planes de mantenimiento predictivo al grupo encargado de esta labor.

1.1 DESCRIPCIÓN DEL PROBLEMA

El mantenimiento tiene como objetivo principal el aprovechamiento adecuado de equipos, sistemas concretos y su restauración. El crecimiento de las industrias, llevó consigo un avance significativo en la visión y concepto que se tenía del mantenimiento, siendo éste de vital importancia para todas las líneas de producción, generando gran cantidad de beneficios tales como costos y mejoras en producción. Por esto, se hizo necesario facilitar el proceso de análisis de sistemas o equipos, haciendo a un lado las antiguas técnicas de mantenimiento tales como la experiencia y la noción de estado de los mismos. Debido a lo anterior, surgió la necesidad de establecer diseños y modelos para la confiabilidad de un sistema [1]. Grosso modo, las técnicas de análisis y pertinentes soluciones en el área de mantenimiento y confiabilidad, se han ejecutado principalmente en contextos vigilados, donde prima el uso de variables controladas así como el desarrollo de prácticas individuales y específicas [2].

Al aumentar las funciones en el área de mantenimiento y sus necesidades, se están conformando nuevos departamentos técnicos especializados que en parte contrarrestan la funcionalidad de los antiguos departamentos de mantenimiento; aumentando los gastos por contratación de nuevo personal y/o la implementación de herramientas sistematizadas a costos elevados. Inconvenientes así afectan a las empresas de procesos continuos con grandes instalaciones automatizadas para las que el alargamiento de la vida útil de los equipos es un imperativo de primer orden. Por ejemplo, en el caso de algunas centrales eléctricas se viene discutiendo si este departamento de nueva creación puede ejercer sus funciones de mantenimiento predictivo o de auditoría técnica externa e independiente, si tiene competencias para paralizar o ralentizar el ritmo de producción de megavatios [3].

Los análisis ejecutados de confiabilidad y representación de los datos de falla, son ajustados netamente a representaciones fijas, sin la posibilidad de ser comparados con nuevos casos posibles o diferentes representaciones ya que han sido desarrollados en contextos uniformes y poco dinámicos. Existen bases de datos para situaciones controladas tales como las realizadas en “Sistema Integrado de Confiabilidad de modelado de herramientas de deformación en frío de plástico usados en la industria de coches” [4]. y “Goodness-of-Fit Tests for the Power-Law Process”[5] donde se realiza el análisis de confiabilidad pero con situaciones previas definidas tales como la temperatura.

El comportamiento probable de una población, equipo o sistema está determinado por la función densidad de probabilidad las cuales requieren de algunos parámetros para ser especificadas en su totalidad. Lograr determinar la función densidad de probabilidad (FDP) de algún equipo partiendo de datos históricos tales como fallas, garantiza una aproximación y/o modelo de dicho comportamiento, obteniendo así la probabilidad o confiabilidad a un determinado tiempo o periodo, además de su vida útil y adecuado mantenimiento; sin embargo, encontrar un modelo únicamente a partir de sus tiempos de falla no es tarea fácil debido a la escases de datos importantes como lo son los parámetros (En algunos casos estos son parámetros de forma y parámetros de escala). Por esto se hace primordial acudir a métodos que realicen una estimación de dichos parámetros y ajusten dicho modelo.

Hoy por hoy, conocer la forma de una distribución de los fallos es importante debido a la cantidad de información que la misma proporciona. Es claro que, a partir de una distribución que se asemeje a la real se puede obtener un acercamiento más preciso al comportamiento real de un equipo determinado. Por lo general, independientemente del comportamiento de los datos, la distribución se aproxima a una distribución exponencial, cometiendo así errores a la hora de calcular probabilidades futuras y más aun, determinar el mantenimiento a realizar.

1.2 JUSTIFICACIÓN

Los modelos y diseños en entornos programados que nos permiten hallar la confiabilidad de equipos y sistemas, se convierten en valiosas herramientas de fácil acceso y operación, obteniendo en los casos en estudio la respuesta deseada, sin intervenir en complejos montajes que lleven a disfuncionalidades y restricciones de tiempo.

Es deseable convertir las anomalías a las que se somete un equipo, ya sean fallas en su construcción, manipulación, mantenimientos previos, entre otros; en claras particularidades para estudios de confiabilidad, dejando a un lado las pruebas enmarcadas por medios controlados y así establecer modelos en tiempo real que poseen las características de lo que realmente sucede, no partiendo de hipótesis tal vez valederas, sino obteniendo datos fiables y certeros.

Los costos de software en confiabilidad, en ocasiones se convierten en un sueño inalcanzable para las pequeñas y medianas empresas, sin importar sus servicios y productos a entregar, lo que acrecienta de manera significativa los dineros entregados a reparaciones inoficiosas y por consiguiente ausentes de total viabilidad. Convirtiéndose estos en gastos, más no en inversiones que permitieran el crecimiento de las compañías. El costo en herramientas programadas y llevadas a aplicaciones industriales mediante plataformas de uso general, disminuyen los sobrecostos presentados y eliminan tediosas y extensas capacitaciones en el manejo de estas.

De esta manera se facilita el análisis en cuanto a fallos por medio de un entorno programado que determina la vida útil y la probabilidad de fallos a determinado tiempo.

Dicho entorno, es una ayuda en diferentes campos ya que facilita el proceso de análisis de diversidad industrias interesadas en efectuar un mantenimiento adecuado a sus equipos, suprimiendo factores externos que influyen dicho proceso, eliminando a su paso variables como la experiencia o el juicio de quien ejecuta el mantenimiento y concentrándose netamente en sus datos históricos, sus fallas, su operatividad y su confiabilidad. Del mismo modo, puede utilizarse con fines académicos y explicativos, mostrando así el comportamiento de los equipos e ilustrando las diferentes leyes de falla, así como el significado de las mismas y algunas aplicaciones en la vida

cotidiana. A su vez, la plataforma programada, evalúa los datos de falla y mediante pruebas de bondad de ajuste se determina, si estos siguen distribuciones tales como la Exponencial, Rayleigh ó Weibull. El entorno se encarga automáticamente de ajustar los datos a la distribución que mejor los describa.

1.3 OBJETIVOS

1.3.1 OBJETIVO GENERAL

Desarrollo de un entorno programado para el análisis y estudio de confiabilidad en sistemas donde sea aplicable la distribución biparamétrica Weibull.

1.3.2 OBJETIVOS ESPECÍFICOS

- ✓ Determinar los parámetros que representan la función densidad de probabilidad para una distribución Weibull biparamétrica, tomando como base los datos tabulados de fallos.
- ✓ Determinar mediante Test de Bondad de Ajuste, intervalos de confianza, la correcta descripción de los datos experimentales por la función densidad de probabilidad obtenida.
- ✓ Permitir determinar la correcta aproximación de la distribución obtenida a otras, tales como la distribución exponencial y la Rayleigh.
- ✓ Con base en la distribución obtenida, permitir el análisis y estudio de datos de confiabilidad.
- ✓ Desarrollar un entorno programado a modo de interfaz hombre/máquina que permita la adecuada interacción entre usuario y el sistema diseñado.

2. FORMULACIÓN MATEMÁTICA

2.1 VARIABLES ALEATORIAS CONTINUAS Y DISTRIBUCIONES DE PROBABILIDAD

2.1.1 Variable aleatoria

Una variable aleatoria es una variable cuyos valores se obtienen de experimentos o eventos donde la medición o el valor de dicha variable se encuentran en un conjunto de posibles valores ya que abarca un intervalo de números haciendo que, la cantidad de valores que pueda tomar la variable sea incontablemente infinito.

Una variable aleatoria x es una función real definida en el espacio muestral asociado a un experimento aleatorio, Ω

$$x : \Omega \rightarrow R$$

2.1.2 Función densidad de probabilidad

La función densidad de probabilidad es una función que describe la probabilidad relativa de una variable aleatoria que se produzca en un momento dado, es decir, caracteriza el comportamiento probable de una población. La probabilidad de una variable aleatoria a caer dentro de una región particular está dada por la integral de la densidad de esta variable en dicha región.

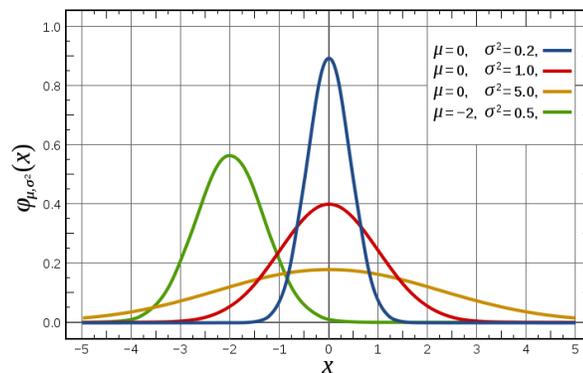


Ilustración 1. Función densidad de probabilidad

Para una variable aleatoria continua x , una función densidad de probabilidad es una función de tal manera que:

1. $f(x) \geq 0$

2. $\int_{-\infty}^{\infty} f(x)dx = 1$

3. $P(a \leq X \leq b) = \int_a^b f(x)dx$

2.1.3 Función de distribución acumulativa

Corresponde a la probabilidad de que un valor de una variable aleatoria X con una distribución de probabilidad dada se encuentra en un valor menos o igual a x . Representa el acumulo o área hasta el momento de la distribución de probabilidad. Se define de la siguiente manera:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u)du$$

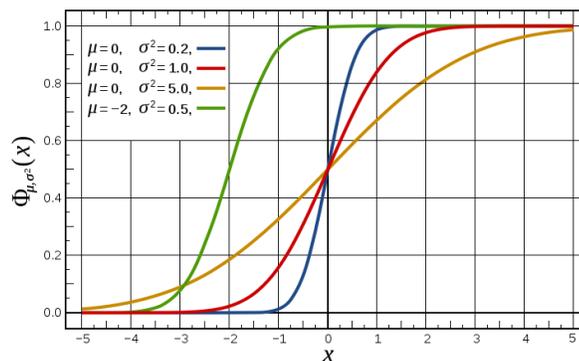


Ilustración 2. Función de distribución acumulativa

2.2 DISTRIBUCIÓN DE PROBABILIDAD Ó LEYES DE FALLA

2.2.1 Distribución Exponencial

Es una distribución de probabilidad continua. Recibe su nombre de la función exponencial que aparece en la función densidad. Tiene como parámetro λ , el cual es un parámetro de escala (también se conoce como tasa de fallas).

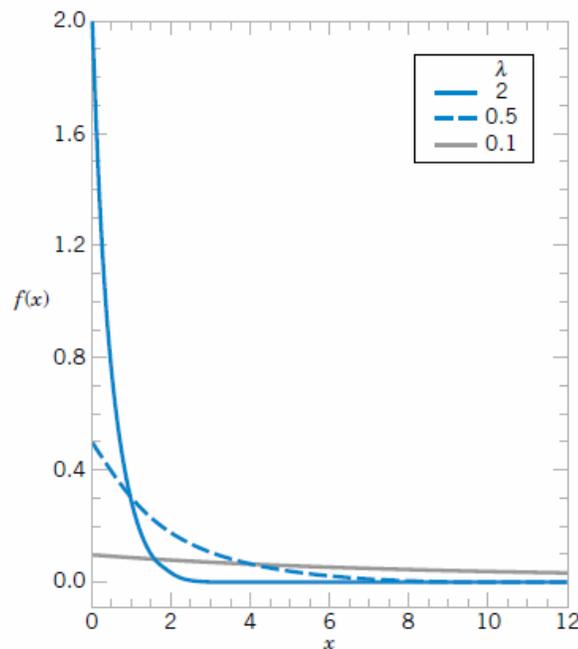


Ilustración 3. Función densidad de probabilidad de la distribución exponencial para diferentes valores de λ

La función densidad de probabilidad para la distribución exponencial está dada por:

$$f(t) = \lambda e^{-\lambda t} \text{ para } 0 \leq t \leq \infty$$

Donde:

λ : Número de fallas por unidad de tiempo (Parámetro de escala)

t : Variable aleatoria que mide el tiempo de fallas.

Por otro lado, la función distribución acumulada se obtiene de la integral de la función densidad de probabilidad evaluada en un intervalo $(0, t)$ y se define como:

$$F(t) = 1 - e^{-\lambda t}$$

Donde:

λ : Número de fallas por unidad de tiempo (Parámetro de escala)

t : Variable aleatoria que mide el tiempo de fallas.

Algunas de las características de la distribución exponencial de un parámetro son:

1. La distribución no tiene parámetro de forma pues tiene una única forma, la exponencial; por lo tanto el único parámetro es la tasa de fallas.
2. A medida que λ disminuye en valor, la distribución se extiende hacia el lado derecho y por el contrario, a medida que λ aumenta en valor, la distribución se acerca al origen.
3. La distribución comienza en $t = 0$, donde $f(t = 0) = \lambda$; a partir de allí, decrece exponencialmente y monótonamente a medida que t se incrementa. Además es convexa.
4. El parámetro de escala $1/\lambda = m = \sigma$ (siendo σ la desviación estándar). Entonces, la confiabilidad para un tiempo de duración $t = m$ es siempre igual a 0,3679 o lo que es lo mismo a un 36,8%. Esto es así pues $R(t) = \exp^{-\lambda t} = \exp^{-\lambda \frac{1}{\lambda}} = \exp^{-1} = 0,368$. Este hecho implica que la confiabilidad es relativamente baja pues sólo el 36,8% de, por ejemplo, componentes en estudio, sobrevivirán.
5. Cuando t tiende a infinito, la función distribución de probabilidades tiende a cero, en consecuencia también tiende a cero la función confiabilidad $R(t)$ [9].

2.2.2 Ley Rayleigh de Falla

En teoría de probabilidad y estadística, la distribución de Rayleigh, es una distribución de probabilidad continua. Una distribución de Rayleigh se observa a menudo cuando la magnitud global de un vector se relaciona con sus componentes de dirección

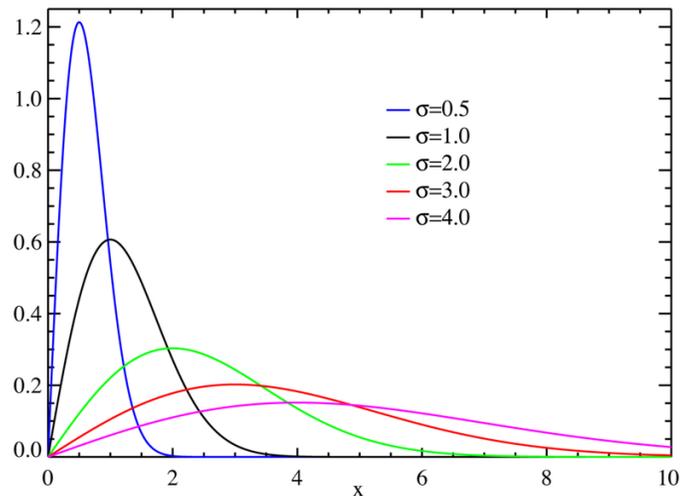


Ilustración 4. Función densidad de probabilidad para la distribución Rayleigh para diferentes valores de α

La función densidad de probabilidad para la distribución Rayleigh está definida por la siguiente expresión:

$$f(t) = \left(\frac{2}{\alpha^2}\right) t e^{-\left(\frac{t}{\alpha}\right)^2}$$

Donde:

α : Parámetro de escala

t: Variable aleatoria que mide el tiempo de fallas.

Por otro lado, la función distribución acumulada se obtiene de la integral de la función densidad de probabilidad evaluada en un intervalo $(0, t)$ y se define como:

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^2}$$

2.2.3 Ley Weibull de Falla

Es una distribución de probabilidad continua aplicable al estudio de la confiabilidad en problemas relativos a la fatiga y vida de componentes y materiales. Los parámetros en la distribución dan una gran idea de la flexibilidad para modelar sistemas, en los cuales el número de fallas se incrementa o decrece con el tiempo. Esta distribución es usada con gran eficacia en los modelos de fallas. Cuenta con dos parámetros, β llamado el parámetro de forma y α corresponde al parámetro de escala.

Su expresión es de la forma:

$$f(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^\beta}$$

Donde:

α : Parámetro de escala

β : Parámetro de forma

t: Variable aleatoria que mide el tiempo de fallas.

Su función de distribución acumulativa es:

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^\beta}$$

Se puede notar, que dependiendo de los parámetros, la función cambia su forma como se puede observar en la figura mostrada a continuación:

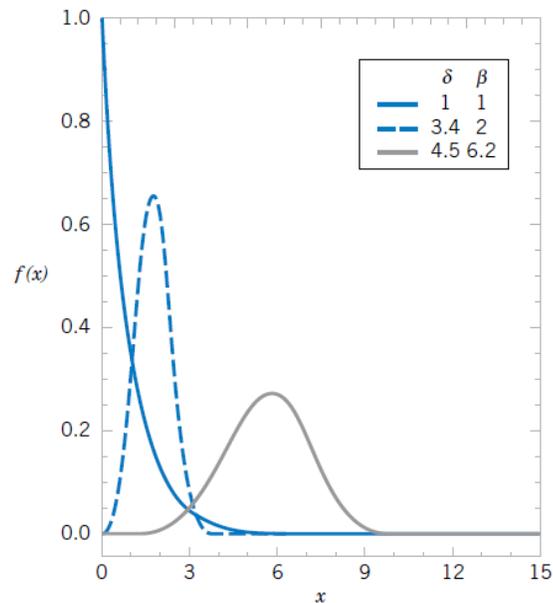


Ilustración 5. Función densidad de probabilidad para la distribución Weibull para diferentes valores de α y β

Para $0 < \beta < 1$:

- A medida que t tiende a cero, la función de densidad de probabilidad tiende a infinito.
- Cuando t tiende a infinito, la función de densidad de probabilidad tiende a cero.
- $f(t)$ decrece monótonamente y es convexa a medida que t aumenta.

Para $\beta = 1$:

- Se puede ver que la distribución exponencial es un caso particular de la distribución de Weibull, por lo tanto la propiedad mencionada en la ley de fallas exponencial de “falta de memoria” es equivalente a la hipótesis de tasa constante.

Para $\beta > 1$:

- $f(t) = 0$ Cuando $t = 0$.
- Para $\beta < 2,6$ la función de densidad de probabilidad de Weibull es asimétrica y posee una cola hacia la derecha.
- Para $2,6 < \beta < 3,7$ la cola desaparece y la forma de la distribución se asemeja a una función de densidad de probabilidad normal.

- Para $\beta > 3,7$, $f(t)$ se vuelve nuevamente asimétrica y aparece una cola en el lado izquierdo.

2.3 ESTIMADORES DE PARÁMETROS PARA ALGUNAS DISTRIBUCIONES

2.3.1 Máxima verosimilitud

El método de máxima verosimilitud suele generar estimadores insesgados de la mínima varianza, siendo este un procedimiento frecuente para ajustar un modelo y encontrar sus parámetros característicos.

Cuando queremos estimar un parámetro de una población de la cual se conoce la familia a la cual pertenece, es decir, su forma, por ejemplo se sabe que dicha población se puede modelar con una distribución Normal, ó Poisson, o Binomial, etc, pero no se conocen sus parámetros, podemos tomar una muestra aleatoria para con base en ella, construir una estadística que nos permita estimar dicho parámetro [12].

| Distribución | Parámetros | Estimadores |
|--------------|---------------------|---|
| Exponencial | β μ | $\hat{\beta} = \bar{x} = \sum \frac{x_i}{n}$ $\hat{\mu} = \sum \frac{x_i}{n}$ |
| Rayleigh | α | $\hat{\alpha} \approx \sqrt{\frac{1}{2N} \sum_{i=0}^N x_i^2}$ |
| Weibull | α β | $\frac{\sum_{i=1}^n X_i^{\hat{\alpha}} \ln X_i}{\sum_{i=1}^n X_i^{\hat{\alpha}}} - \frac{1}{\hat{\alpha}} = \frac{\sum_{i=1}^n \ln X_i}{n}$ $\hat{\beta} = \left(\frac{\sum_{i=1}^n X_i^{\hat{\alpha}}}{n} \right)^{1/\hat{\alpha}}$ |

Tabla 1. Estimadores de máxima verosimilitud

Encontrar las estimaciones para la distribución Weibull, consiste en tomar las derivadas parciales de la función de probabilidad con respecto a los parámetros, organizar las ecuaciones que resultan igual a cero y resolver simultáneamente para determinar los valores de los parámetros estimados [32].

En el artículo enunciado se encuentra la demostración robusta de cómo se llega a los parámetros de α y β . Cabe resaltar que la manera como se hallan los parámetros en el entorno programado es mediante el Toolbox de matlab y su comando *parmhat = wblfit(data)*.

2.4 PRUEBAS DE BONDAD DE AJUSTE

Las pruebas de bondad de ajuste tienen por objetivo determinar si los datos se ajustan a una determinada distribución. Es necesario resaltar, que las pruebas están en capacidad de rechazar con certeza absoluta la hipótesis nula, pero si dicha hipótesis se acepta quiere decir que la distribución escogida es posiblemente la que describe el comportamiento de los datos.

2.4.1 Test de Kolmogorov-Smirnov

Es una prueba no paramétrica que se utiliza para determinar la bondad de ajuste. Estos test se hacen necesarios cuando la distribución no puede ser definida concretamente pues son los datos observados los que la construyen.

La prueba de Kolmogorov-Smirnov se basa en la diferencia entre la Función de distribución acumulativa y la Función de distribución empírica [13]. La discrepancia obtenida entre la función de distribución teórica de la empírica da como resultado la *distancia de Kolmogorov-Smirnov* y con esta se puede determinar si la hipótesis se acepta o se rechaza.

Claramente, la prueba de Kolmogorov-Smirnov se plantea de la siguiente manera

1. Se plantea la hipótesis nula:

$$\mathcal{H}_0 = \text{Los datos tiene por función de distribución a } F_0$$

Donde F_0 es la función de distribución de una ley continua dada

2. Se define la función de distribución empírica

$$\hat{F}(x) = \begin{cases} 0 & \text{para } x < X_{(1)} \\ \frac{i}{n} & \text{para } X_i \leq x < X_{(i+1)} \\ 1 & \text{para } x \geq X_{(n)} \end{cases}$$

3. Se plantea la distancia de Kolmogorov-Smirnov, realizando restas por encima y por debajo de la función entre la función de distribución acumulada y la función de distribución empírica

$$D_{KS}(F_0, \hat{F}) = \max_{i=1, \dots, n} \left\{ \left| F_0(X_{(i)}) - \frac{i}{n} \right|, \left| F_0(X_{(i)}) \right| - \frac{i-1}{n} \right\}$$

Donde :

$F_0(x)$: Función de distribución acumulada

4. Finalmente, una vez obtenida la distancia de Kolmogorov-Smirnov se verifica que dicha distancia sea menor o igual a la Distribución del estadístico de Kolmogorov-Smirnov

$$D_{KS} \leq D_n$$

Si y solo si esta condición se cumple, se puede decir que se cumple la hipótesis nula \mathcal{H}_0

La tabla de distribución del estadístico de Kolmogorov-Smirnov ó los límites de aceptación para el test KS se encuentra en el Anexo 1.

Ejemplo

De una tabla de números aleatorios se eligen los siguientes 50 (divididos entre 100 para que su valor oscile entre 0 y 1)

| | | | | |
|------|------|------|------|------|
| 0.1 | 0.97 | 0.7 | 0.13 | 0.34 |
| 0.37 | 0.04 | 0.64 | 0.74 | 0.24 |
| 0.08 | 0.68 | 0.19 | 0.09 | 0.23 |
| 0.99 | 0.02 | 0.09 | 0.7 | 0.38 |
| 0.12 | 0.99 | 0.8 | 0.36 | 0.64 |
| 0.66 | 0.74 | 0.34 | 0.94 | 0.36 |
| 0.31 | 0.1 | 0.45 | 0.82 | 0.35 |
| 0.85 | 0.77 | 0.02 | 0.65 | 0.68 |
| 0.63 | 0.32 | 0.05 | 0.74 | 0.9 |
| 0.73 | 0.42 | 0.03 | 0.64 | 0.35 |

Tabla 2. Números aleatorios oscilando entre 0 y 1

❖ Paso 1

Se desea probar la hipótesis H_0 : Proviene de una distribución uniforme en $[0,1]$, a un nivel de significancia del 90%

❖ Paso 2

Se selecciona una muestra de tamaño n de números pseudoaleatorios

| | | | | |
|------|------|------|------|------|
| 0.1 | 0.97 | 0.7 | 0.13 | 0.34 |
| 0.37 | 0.04 | 0.64 | 0.74 | 0.24 |
| 0.08 | 0.68 | 0.19 | 0.09 | 0.23 |
| 0.99 | 0.02 | 0.09 | 0.7 | 0.38 |
| 0.12 | 0.99 | 0.8 | 0.36 | 0.64 |
| 0.66 | 0.74 | 0.34 | 0.94 | 0.36 |
| 0.31 | 0.1 | 0.45 | 0.82 | 0.35 |
| 0.85 | 0.77 | 0.02 | 0.65 | 0.68 |
| 0.63 | 0.32 | 0.05 | 0.74 | 0.9 |
| 0.73 | 0.42 | 0.03 | 0.64 | 0.35 |

Tabla 3. Datos seleccionados para aplicar Kolmogorov-Smirnov

❖ Paso 3

3.1 Se organiza la tabla anterior para que se cumpla la condición $X_i \leq X_{i+1}$ para toda posición i .

| | | | | |
|------|------|------|------|------|
| 0.02 | 0.12 | 0.35 | 0.64 | 0.74 |
| 0.02 | 0.13 | 0.36 | 0.65 | 0.77 |
| 0.03 | 0.19 | 0.36 | 0.66 | 0.8 |
| 0.04 | 0.23 | 0.37 | 0.68 | 0.82 |
| 0.05 | 0.24 | 0.38 | 0.68 | 0.85 |
| 0.08 | 0.31 | 0.42 | 0.7 | 0.9 |
| 0.09 | 0.32 | 0.45 | 0.7 | 0.94 |
| 0.09 | 0.34 | 0.63 | 0.73 | 0.97 |
| 0.1 | 0.34 | 0.64 | 0.74 | 0.99 |
| 0.1 | 0.35 | 0.64 | 0.74 | 0.99 |

Tabla 4. Datos organizados de menor a mayor para prueba KS

3.2 Si $i = 1$, entonces $F_n(X_1) = 1/50 = 0.02$. Se construye $F_n(X_i)$ para toda posición i , siendo $n = 50$.

| | | | | |
|--------------------|--------------------|--------------------|--------------------|--------------------|
| $F_n(0.00) = 0.00$ | $F_n(0.12) = 0.22$ | $F_n(0.35) = 0.42$ | $F_n(0.65) = 0.64$ | $F_n(0.82) = 0.88$ |
| $F_n(0.02) = 0.04$ | $F_n(0.13) = 0.24$ | $F_n(0.36) = 0.46$ | $F_n(0.66) = 0.66$ | $F_n(0.85) = 0.90$ |
| $F_n(0.03) = 0.06$ | $F_n(0.19) = 0.26$ | $F_n(0.37) = 0.48$ | $F_n(0.68) = 0.70$ | $F_n(0.90) = 0.92$ |
| $F_n(0.04) = 0.08$ | $F_n(0.23) = 0.28$ | $F_n(0.38) = 0.50$ | $F_n(0.70) = 0.74$ | $F_n(0.94) = 0.94$ |
| $F_n(0.05) = 0.10$ | $F_n(0.24) = 0.30$ | $F_n(0.42) = 0.52$ | $F_n(0.73) = 0.76$ | $F_n(0.97) = 0.96$ |
| $F_n(0.08) = 0.12$ | $F_n(0.31) = 0.32$ | $F_n(0.45) = 0.54$ | $F_n(0.74) = 0.82$ | $F_n(0.99) = 1.00$ |
| $F_n(0.09) = 0.16$ | $F_n(0.32) = 0.34$ | $F_n(0.63) = 0.56$ | $F_n(0.77) = 0.84$ | |
| $F_n(0.10) = 0.20$ | $F_n(0.34) = 0.38$ | $F_n(0.64) = 0.62$ | $F_n(0.80) = 0.86$ | |

❖ Paso 4

Se evalúa:

$$D = \text{Max} |F_n(X_i) - X_i| = 0.12.$$

Es decir, se busca la mayor de las desviaciones en valor absoluto, para lo cual se deben calcular todas las desviaciones entre los valores de probabilidad acumulada teórica y los valores de probabilidad acumulada empírica.

El valor D máximo ocurre para $F_n(0.38)$, que esta en la posición $n = 25$, es decir con una probabilidad acumulada teórica de $0.02 * 25 = 0.50$.

❖ Paso 5

Para un nivel de significancia del 90% y una muestra de 50 números se tiene de la Tabla Kolmogorov – Smirnov (ver anexo 1), un valor de 0.172

Como $D = 0.12 < 0.172$ **se acepta H_0** , o sea, los números si provienen de una distribución uniforme en el intervalo cerrado $[0,1]$. [8]

2.4.2 Prueba Chi-Cuadrado

Es una prueba de bondad de ajuste también conocida como la prueba χ^2 de Pearson o Chi-Cuadrado. La utilización de estos métodos se hace recomendable cuando no se puede asumir que los datos se ajusten a una distribución conocida. Es considerada como una prueba no paramétrica que mide la discrepancia o distancia entre una distribución observada y otra teórica (bondad de ajuste).

La fórmula del estadístico es la siguiente:

$$X^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Teniendo en cuenta que $E_i = nP_i$

Donde:

O_i = Frecuencia observada de la muestra

E_i = Frecuencia esperada según la distribución teórica

n = Tamaño de la muestra

Para calcular el valor de Chi-Cuadrado se realiza lo siguiente:

1. Se divide la muestra de datos en intervalos y se obtienen la frecuencias observada para cada intervalo (Una buena aproximación para la cantidad de intervalos necesarios puede ser \sqrt{n} donde n es el tamaño de la muestra).
2. Se calcula la frecuencia esperada para cada intervalo basados en la función distribución de probabilidad o acumulada.
3. Con la fórmula del estadístico se obtiene el valor de Chi-Cuadrado
4. Finalmente, con el valor de χ^2 se verifica que sea menor o igual al valor obtenido de las Tablas de Distribución Chi-Cuadrado (Anexo 2)

Ejemplo

Digamos que 900 estudiantes expresan su voluntad por celebrar el aniversario de la institución organizando uno de dos eventos: un acto solemne en el templo universitario o una actividad deportiva en el estadio de fútbol. Una vez hecha la encuesta se tiene que 495 alumnos prefieren la actividad deportiva y 405 se inclinan por el acto solemne. ¿Existe una diferencia significativa entre los estudiantes en su preferencia por la actividad deportiva?

La prueba estadística para determinar la significatividad de la diferencia en las frecuencias observadas es la prueba llamada Chi Cuadrada. Para el caso que nos ocupa, se supone que si no hay diferencia en la preferencia de los alumnos de una manera perfecta, tendríamos 450 alumnos eligiendo el acto solemne y otros 450 eligiendo las actividades deportivas.

Esa es la frecuencia de respuestas esperadas en el caso de una igualdad absoluta. Pero tenemos frecuencias observadas un poco diferentes en un caso son 495 y en el otro 405, lo que deseamos saber es si esa diferencia observada es significativa.

Lo que se hace al aplicar la fórmula de chi cuadrada es restar al número de frecuencias observadas, el número de frecuencias esperadas; elevar esta diferencia al cuadrado, lo que hace que todos los valores asuman un valor positivo, y luego se divide el cuadrado obtenido entre el las frecuencias esperadas. Esto se hace de manera independiente para cada una de las categorías. Una vez terminado este paso, se suman los resultados obtenidos en cada categoría y ese valor resultante de la suma es el valor Chi cuadrada observado, el cual deberá ser comparado con el valor Chi cuadrada crítico según el nivel alpha de significatividad escogido y los grados de libertad correspondientes.

En el caso de nuestro ejemplo se trata de dos categorías, lo que conduce a un grado de libertad. A continuación el proceso para calcular el valor Chi cuadrada

1. A favor del acto solemne:

Frecuencias observadas = 405

Frecuencias esperadas = 450

$(\text{Frecuencias observadas} - \text{frecuencias esperadas})^2 / \text{frecuencias esperadas}$

$$(405 - 450) / 450 = (-45)^2 / 450 = 2025 / 450 = 4.5$$

2. A favor del acto deportivo:

Frecuencias observadas = 495

Frecuencias esperadas = 450

$(\text{Frecuencias observadas} - \text{frecuencias esperadas})^2 / \text{frecuencias esperadas}$

$$(495 - 450) / 450 = (45)^2 / 450 = 2025 / 450 = 4.5$$

3. Se suman los valores obtenidos en cada grupo para obtener el valor de chi cuadrada.

$$4.5 + 4.5 = 9.00$$

4. Se compara este valor con el valor correspondiente a un grado de libertad en la tabla de Chi cuadrado (ver anexo 2) y se encuentra que el valor crítico de χ^2 para un grado de libertad a un nivel $\alpha = .05$ a dos colas es = 3.8941.

Siendo que el valor Chi cuadrada (χ^2) obtenido es mayor que el valor crítico, se desacredita la hipótesis nula que afirma que no existe diferencia significativa entre las frecuencias observadas y se concluye que la diferencia es significativa. Esto quiere decir que en menos de 5 casos de cada cien, una diferencia como la del valor igual o mayor al observado de Chi cuadrado en este caso ($\chi^2 = 9$), puede ser atribuida a la selección de la muestra (azar). [8]

2.5 NIVELES DE CONFIANZA

El nivel de confianza, es la probabilidad a priori de que el intervalo de confianza a calcular contenga al verdadero valor del parámetro. Se indica por $1-\alpha$ y habitualmente se da en porcentaje ($1-\alpha$). Hablamos de nivel de confianza y no de probabilidad ya que una vez extraída la muestra, el intervalo de confianza contendrá al verdadero valor del parámetro o no, lo que sabemos es que si repitiésemos el proceso con muchas muestras podríamos afirmar que el $(1-\alpha)$ de los intervalos así construidos contendría al verdadero valor del parámetro. Que tanta confianza tenemos que la estimación que hicimos de un intervalo, incluya la mayor parte de la muestra, es decir los casos analizados [14].

El nivel de significancia de un test es un concepto estadístico asociado a la verificación de una hipótesis. Se define como la probabilidad de tomar la decisión de rechazar la hipótesis nula cuando ésta es verdadera. Diferentes valores de α tienen ventajas y desventajas. Valores pequeños de α otorgan mayor confianza en la determinación de dicha significancia, pero hacen correr mayores riesgos de equivocarse al rechazar la hipótesis nula falsa.

❖ Error tipo I.

Consiste en rechazar una hipótesis nula cuando ésta en realidad, es verdadera.

Recomendaciones para disminuir el error de tipo I:

- Disminuir el número de test estadísticos llevados a cabo en el estudio
- Depurar la base de datos para evitar errores de valores extremos que puedan producir hallazgos significativos.
- Utilizar valores de alfa más reducidos (0.01 ó 0.001).
- Reproducir el estudio. Si al reproducir el estudio se obtienen resultados similares, estaremos más seguros de no estar cometiendo el error de tipo I.

❖ **Error tipo II**

Este tipo de error, sucede cuando se afirma una hipótesis que de antemano, es falsa.

Recomendaciones para disminuir el error de tipo II:

- Incrementar el tamaño de la muestra.
- Estimar el poder estadístico del estudio.
- Incrementar el tamaño del efecto a detectar.
- Incrementar el valor de alfa.
- Utilizar test paramétricos (más potentes) en lugar de test no paramétricos.

3. METODOLOGÍA PROPUESTA

3.1 RECOLECCIÓN DE LA INFORMACIÓN

La recolección de información, fue ejecutada en dos partes:

- ✓ Inicialmente se realizó una investigación a fondo sobre confiabilidad, distribuciones de probabilidad y los ajustes necesarios para obtener así un modelo confiable de un equipo basado en datos de fallas registrados.
- ✓ Se generaron bases de datos con números aleatorios obtenidos de “Excel” y una herramienta para el análisis de confiabilidad “EasyFit”; consultas a libros e internet con la finalidad de obtener gran variedad en cuanto a bases de datos y así observar el comportamiento del entorno programado cuando se somete a pruebas con diferentes características.

Los diferentes casos ejecutados fue determinante para la evaluación del entorno programado ya que se puso en evidencia el funcionamiento del mismo.

3.2 TRATAMIENTO DE LOS DATOS

Una vez se obtiene una muestra de tamaño n , la organización de los mismos es fundamental a la hora de realizar los análisis respectivos, por esto, es necesario ordenarlos de forma ascendente como primera medida.

El ingreso de los datos para el respectivo análisis es fundamental para obtener el modelo, sin embargo, son de entera disposición y manipulación por parte del usuario o persona que ejecuta el entorno programado. Básicamente, el usuario determina el tratamiento que desee hacerle a los datos, los cuales tienen diferentes opciones, tales como la organización o el nivel de significancia requerido para el análisis deseado.

Organización de los datos:

- ❖ *Organización de datos S1.* Incluye todos los valores de tiempos tomados para el análisis de confiabilidad. Los n valores tomados para el estudio y aproximación, son los n valores ingresados.
- ❖ *Organización de datos S2.* Una vez sean cargados los datos de tiempos de fallas, realiza un recorrido fila a fila (los cuales deben estar en orden ascendente), verificando si cada dato actual es diferente al anterior. En caso tal que dicho dato sea igual al anterior, lo elimina y continúa con la siguiente fila. Por esto, la cantidad de datos i analizados para el modelo de confiabilidad es menor o igual que el tamaño total de la muestra de fallos n .
- ❖ *Organización de datos S3.* Se realiza una compilación de la base de datos mediante intervalos representativos. El número de intervalos se toma a partir de una convención tomada en estadística donde se recomienda que, para establecer dichos intervalos un buen número de intervalos es \sqrt{n} donde n es el tamaño de la muestra. Cada intervalo tiene una frecuencia relativa y absoluta, las cuales indican el porcentaje de datos incluidos en su intervalo respectivo sobre la muestra total y la cantidad de datos que se encuentran en dicho intervalo respectivamente.

Organización de datos S1
Para todos los valores

| N_o | Valor |
|-------|-------|
| 1 | t_1 |
| 2 | t_2 |
| ... | ... |
| i | t_i |
| ... | ... |
| n | t_n |

$t_1 < t_2 < \dots < t_i < \dots < t_n$
 $N = n =$ Tamaño de la Muestra

Organización de datos S2
Eliminando valores repetidos

| N_o | Valor |
|-------|-------|
| 1 | t_1 |
| 2 | t_2 |
| ... | ... |
| i | t_i |
| ... | ... |
| v | t_v |

$t_1 < t_2 < \dots < t_i < \dots < t_v$
 $v < N$ donde N : Tamaño de la muestra

Organización de datos S3
Agrupando por intervalos

| N_o | Valor |
|-------|------------------|
| 1 | $[d_0, d_1)$ |
| 2 | $[d_1, d_2)$ |
| ... | ... |
| i | $[d_{i-1}, d_i)$ |
| ... | ... |
| m | $[d_{m-1}, d_m)$ |

Se recomienda una cantidad de intervalos de \sqrt{N} donde N es el tamaño de la muestra

Las tablas anteriores muestran las 3 organizaciones de datos diferentes posibles en el entorno programado.

Para la organización de los datos S3 es necesario realizar un tratamiento posterior ya que los valores se encuentran en intervalos y no se tiene un valor estándar

| No. | Intervalo | Valor Convencional | Frecuencia Absoluta n_i | Frecuencia Relativa f_i |
|-----|------------------|---------------------------------|---------------------------|---------------------------|
| 1 | $[d_0, d_1)$ | $t_1 = \frac{d_0 + d_1}{2}$ | n_1 | $f_1 = \frac{n_1}{N}$ |
| ... | ... | ... | ... | ... |
| i | $[d_{i-1}, d_i)$ | $t_i = \frac{d_{i-1} + d_i}{2}$ | n_i | $f_i = \frac{n_i}{N}$ |
| ... | ... | ... | ... | ... |
| m | $[d_{m-1}, d_m)$ | $t_m = \frac{d_{m-1} + d_m}{2}$ | n_m | $f_m = \frac{n_m}{N}$ |
| | | | $\sum_{i=1}^m n_i = N$ | $\sum_{i=1}^m f_i = 1$ |

Según lo anterior, para el tratamiento de los datos S3 es necesario establecer un valor convencional para cada intervalo que haga la representación correspondiente.

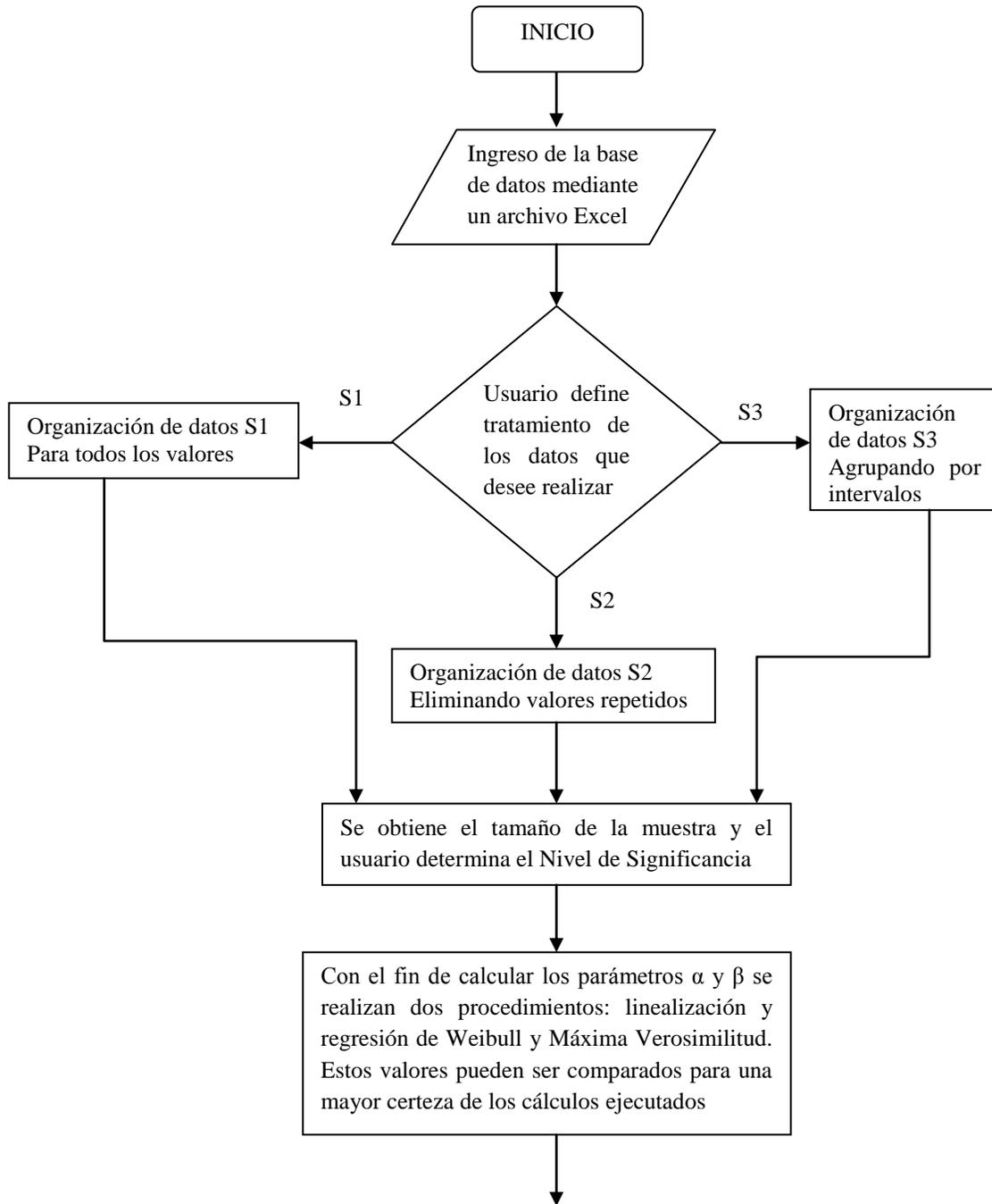
Nivel de significancia:

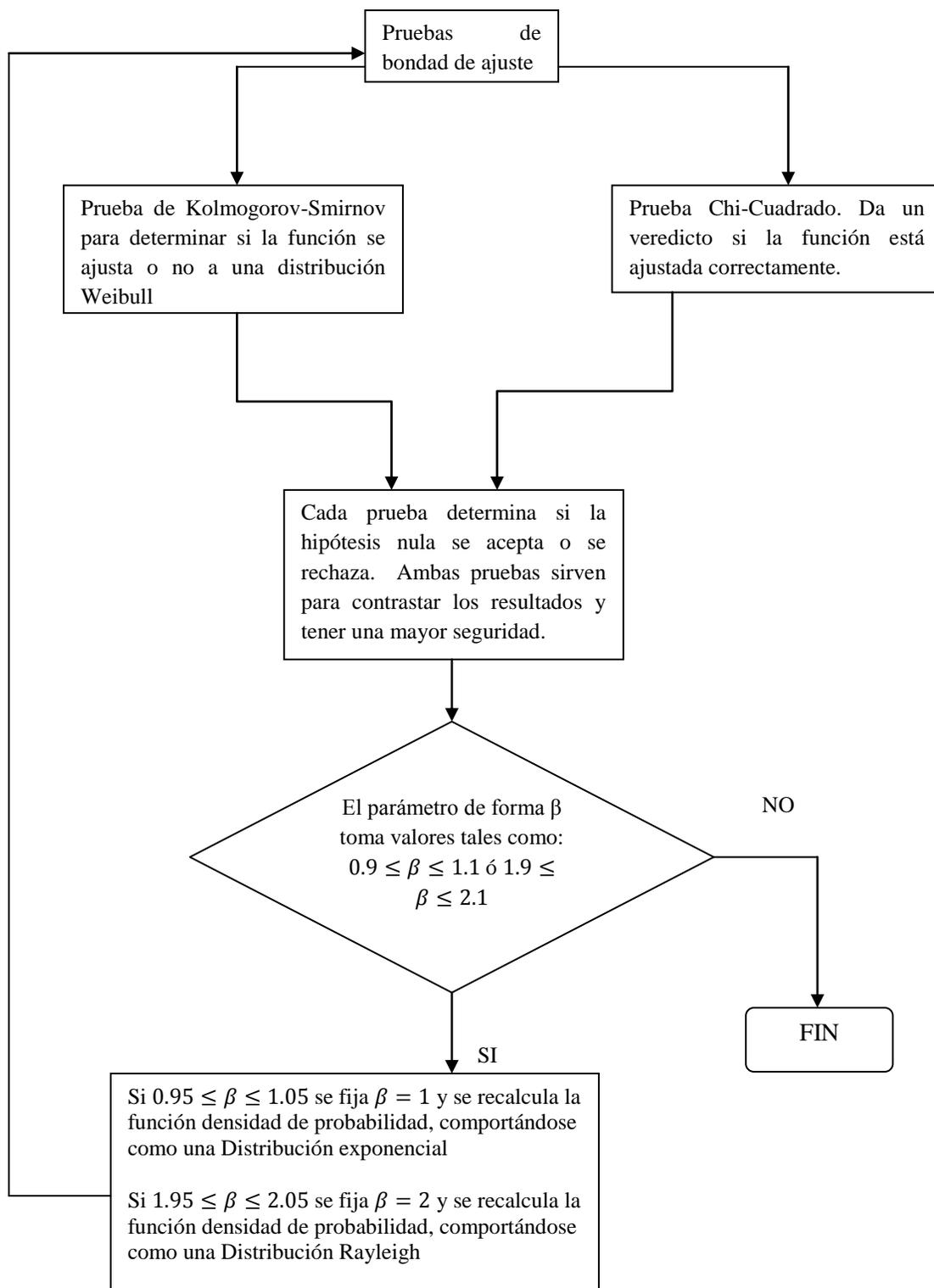
Es la máxima cantidad de error que estamos dispuestos a aceptar para dar como validez la hipótesis nula. Se establece convencionalmente en 5%, sin embargo, el entorno programado cuenta con una serie de posibilidades en cuanto a niveles de significancia los cuales son decisión del usuario para el tratamiento de los datos que desea realizar

- ❖ Nivel de significancia del 1%
- ❖ Nivel de significancia del 5%
- ❖ Nivel de significancia del 10%
- ❖ Nivel de significancia del 20%

El nivel de significancia, es una selección del usuario, la cual se encuentra disponible en la interfaz de usuario, siendo esta un parámetro de entrada.

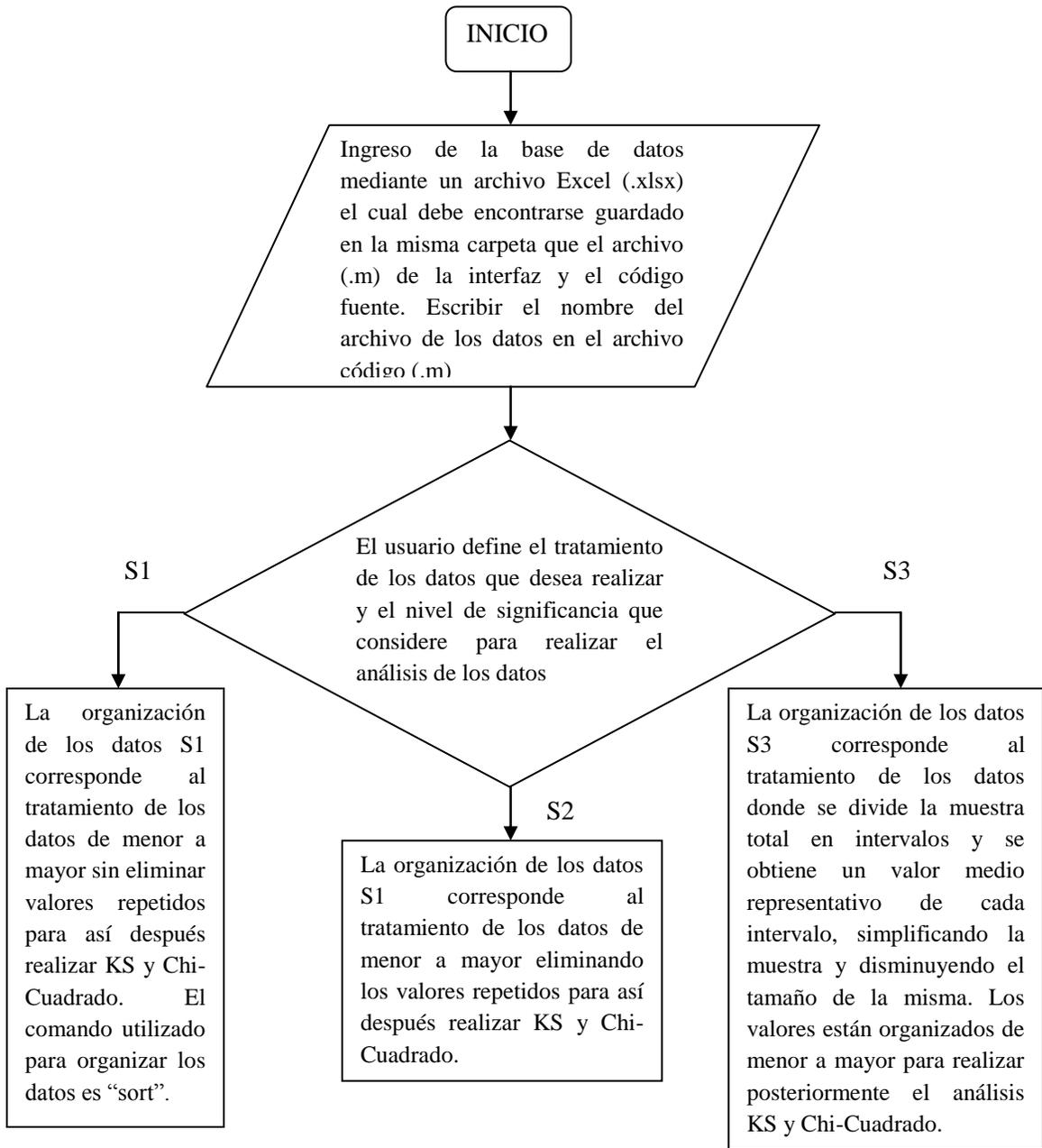
3.3 PROCEDIMIENTO DEL MODELO





Diagramas de flujo discriminados de acuerdo a su función

❖ Tratamiento de los datos:



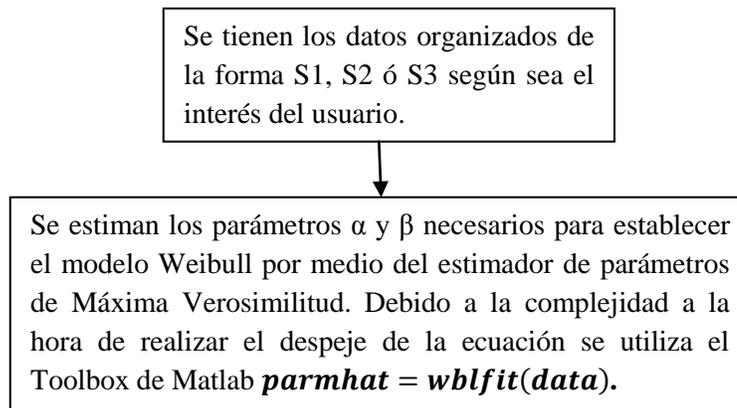
A continuación se presenta el código fuente con el cual se realizan los ajustes necesarios a los datos dependiendo del tratamiento que el usuario seleccionó para el respectivo análisis:

| Código Fuente | Proceso realizado |
|--|--|
| <pre> if selecciondatos == 1, NUMERIC = xlsread('Base1.xlsx'); end; </pre> | <p>En el código, se puede observar “selecciondatos”, que corresponde al tratamiento de los datos seleccionado por el usuario. En este caso, selecciondatos == 1 y corresponde a S1, donde se tienen todos los datos, incluyendo los repetidos. El comando “xlsread” ayuda a cargar la base de datos a analizar la cual debe encontrarse en un archivo de Excel.</p> |
| <pre> if selecciondatos == 2, Numeric1 = xlsread('Base1.xlsx'); Numeric2 = sort(Numeric1); ii = size (Numeric1,1); iii=1; n=1; NUMERIC=zeros(n,2); NUMERIC (n,1) = n; if iii==1, NUMERIC (n,2) = Numeric2 (iii,2); iii=iii+1; n=n+1; NUMERIC (n,1) = n; end; while iii<=ii, if Numeric2 (iii,2) == Numeric2 (iii-1,2); else NUMERIC (n,2) = Numeric2 (iii,2); NUMERIC (n,1) = n; n=n+1; end; iii=iii+1; end end; end; </pre> | <p>En este caso, selecciondatos == 2 y corresponde a S2 donde se eliminan los valores repetidos de la base de datos. Nuevamente el comando “xlsread” carga la base de datos del archivo Excel original de los datos.</p> <p>Después de esto, se toma el tamaño o la cantidad de datos de la base a analizar y además de esto se crea una matriz de ceros donde se organizarán los datos de menor a mayor, verificando siempre el anterior y eliminando el nuevo dato en caso tal de encontrarse repetido.</p> $t_1 < t_2 < \dots < t_i < \dots < t_v$ $v < N$ <p>Donde N: Tamaño de la muestra.</p> <p>Se garantiza que cada valor t de los datos es mayor que su anterior y menor que su predecesor.</p> |

| Código Fuente | Proceso realizado |
|---|--|
| <pre> if selecciondatos == 3, Numeric3 = xlsread('Base1.xlsx'); Numeric4 = sort (Numeric3) iv = size (Numeric4,1); Nintervalo= round(sqrt(iv)); DatoP= Numeric4(1,2) DatoF= Numeric4(iv,2) Tamano= DatoF - DatoP Intervalo = Tamano / Nintervalo FrecAbs = (hist(Numeric4,Nintervalo))'; nn=1; v=1; T = zeros (Nintervalo,1); while v <= Nintervalo, T(v,1)= nn; T(v,2)= (DatoP + (DatoP +Intervalo))/2; DatoP = DatoP + Intervalo; v=v+1; nn= nn+1; end NUMERIC = T end; </pre> | <p>En este caso, selecciondatos == 3 y corresponde a S3 donde se divide la muestra total en una cantidad de intervalos fija igual a \sqrt{n} donde n es la cantidad de datos de la muestra total.</p> <p>Nintervalos: \sqrt{n} (Corresponde al número de intervalos considerados para el tamaño de la base de datos).</p> <p>DatoP: El menor de los datos de la muestra.</p> <p>DatoF: El mayor de los datos de la muestra.</p> <p>Intervalo: Es el tamaño de cada uno de los intervalos, la cual se obtiene dividiendo la longitud desde el primer dato hasta el último, entre el numero de intervalos.</p> <p>T: se establece un valor medio $t_1 = \frac{d_0+d_1}{2}$ del intervalo correspondiente $[d_0, d_1)$.</p> |

Máxima Verosimilitud:

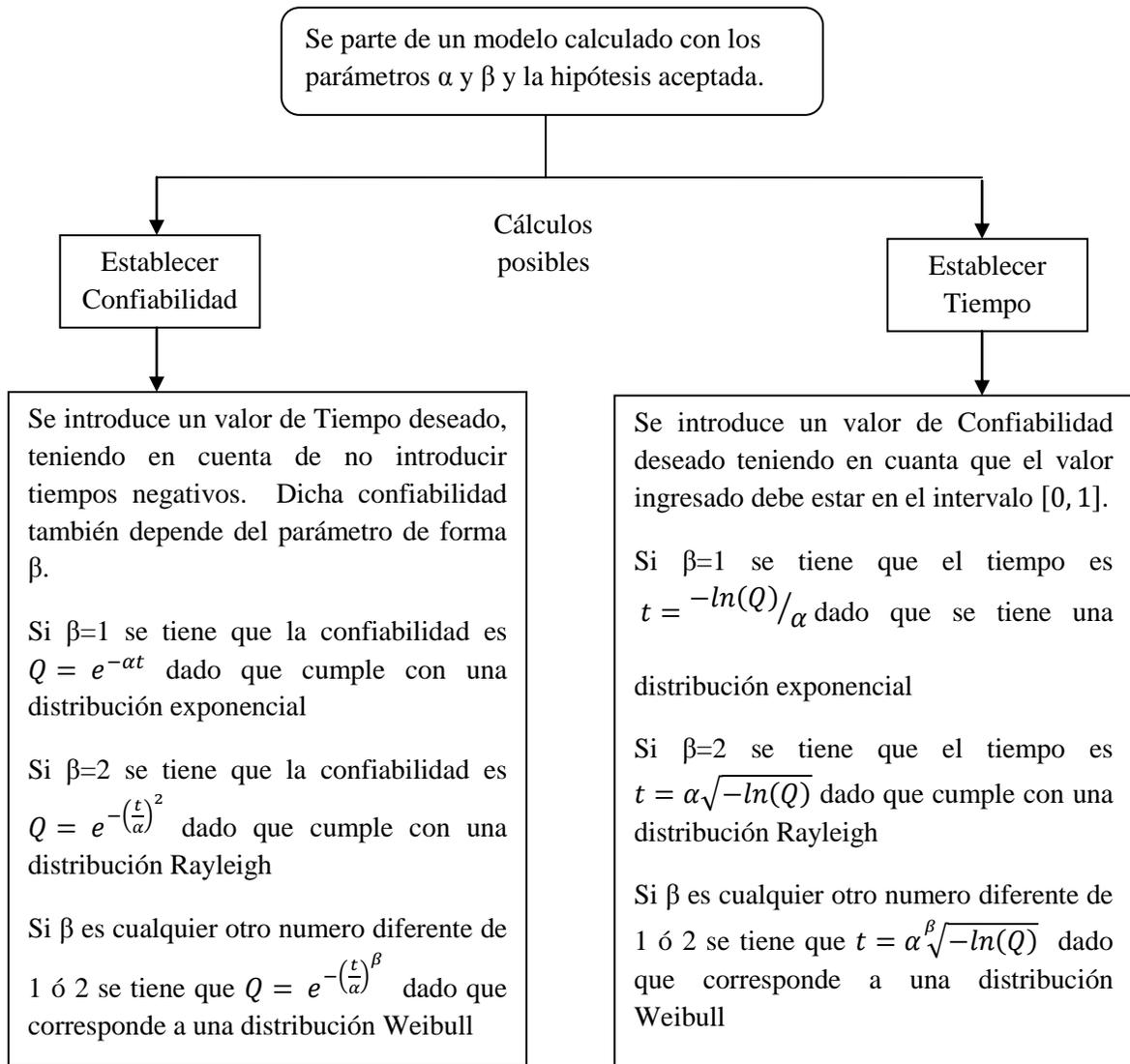
Es un método de estimación de los parámetros de un modelo estadístico. Cuando se aplica un conjunto de datos y teniendo en cuenta un modelo estadístico, la probabilidad máxima de estimación proporciona estimaciones de los parámetros del modelo.



Pruebas de bondad de ajuste:

Con el resultado de ambos parámetros α y β , después de aplicar el Toolbox de Matlab, se realizan las pruebas de bondad de ajuste con el fin de verificar que tan bien ajustado o aproximado se encuentra el modelo calculado de la base de datos a analizar. Los cálculos ejecutados en el programa son iguales a los que se pueden evidenciar en la página 17-22, donde se encuentra un ejemplo respectivo para Kolmogorov-Smirnov como para el test Chi-Cuadrado.

Cálculos posteriores de Confiabilidad y/o Tiempo:



Cabe resaltar, que dado que no se cumplan ninguna de las condiciones necesarias para realizar los cálculos de Confiabilidad y/o Tiempo, se tiene mensajes de advertencia necesarios para evitar errores o equivocaciones por parte del usuario. El manejo de la interfaz se explica en detalle en el subcapítulo siguiente, con cada una de las posibles situaciones a presentarse y la interfaz respectiva.

3.4 INTERFAZ DE USUARIO

La interfaz de usuario, es la comunicación directa entre el usuario y el código fuente del Entorno programado, siendo esta la encargada de mostrar en pantalla todos los resultados obtenidos durante la ejecución y análisis de una base de datos determinada. Allí se presentan los parámetros de la distribución Weibull, si la hipótesis nula se acepta o rechaza, es decir, si el modelo aproximado para una base de datos en cuestión puede o no ser simulado como una Distribución Weibull; y las gráficas de la función densidad de probabilidad y la función de distribución acumulada. Además de esto, emite un mensaje de alerta cada vez que el parámetro de forma β se encuentra en unos rangos permitidos y definidos interiormente en el código fuente para realizar aproximaciones de una distribución Weibull bien sea a una distribución Exponencial o Rayleigh dependiendo de dicho valor numérico.

El usuario, por medio de la interfaz, está en la obligación de fijar algunos datos de entrada por medio de un “Pop-Up Menú” los cuales se realizan directamente en la interfaz. El tratamiento de los datos (S1, S2 ó S3) y el nivel de significancia (0.01, 0.05, 0.1, 0.2) deben ser determinados ya que son fundamentales y decisivos a la hora de realizar un análisis de confiabilidad.

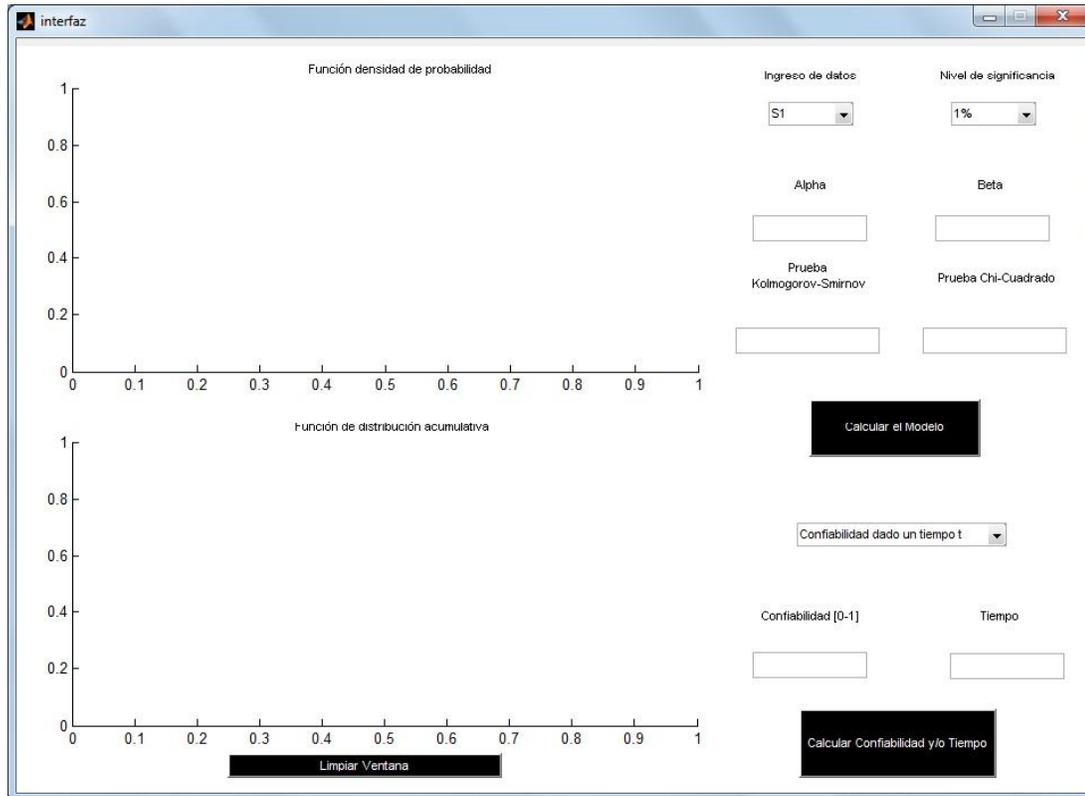


Ilustración 6. Interfaz de Usuario

En la ilustración 6, se observa claramente la presencia de 3 "Pushbutton": Calcular el modelo, Calcular confiabilidad y/o tiempo, Limpiar ventana.

- ❖ Calcular el modelo (Pushbutton1): Con este pushbutton se logran obtener los resultados del modelo de confiabilidad partiendo del tratamiento de datos y nivel de significancia seleccionados previamente. Básicamente se ejecuta el programa que determina los parámetros de la distribución y los resultados de las pruebas de bondad de ajuste.
- ❖ Calcular Confiabilidad y/o Tiempo: El uso de este pushbutton es únicamente para calcular confiabilidad o tiempo. Es claro que, para obtener un resultado de los datos analizados tanto para tiempo o confiabilidad, es obligatorio ejecutar el cálculo del modelo con anterioridad. Cabe resaltar que ambos cálculos son decisión del usuario y se parte del hecho de tener un modelo previamente establecido.

- ❖ Limpiar Ventana: El uso de este pushbutton es recomendable cuando el usuario desea limpiar los valores en los campos de “Confiability” y “Tiempo” de la interfaz gráfica.

Consideraciones Especiales en la Interfaz Gráfica:

Si el parámetro de forma β se encuentra en un rango de tolerancias tal como $0.9 < \beta < 1.1$ el entorno programado emite un mensaje de advertencia como el observado en la Ilustración 8 con el objetivo de dejar a consideración del usuario si desea realizar una aproximación del modelo basado en la distribución Weibull a una distribución Exponencial. Se ha visto anteriormente que la distribución Exponencial es una simplificación o caso particular de la distribución Weibull cuando el parámetro de forma es $\beta = 1$

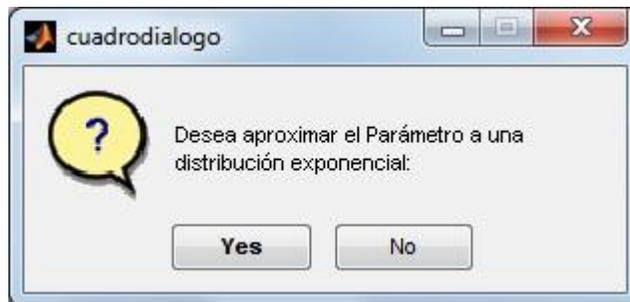


Ilustración 7. Mensaje aproximación distribución Exponencial

Una nueva aproximación se realiza para un rango entre $1.9 < \beta < 2.1$. En este caso, cuando $\beta = 2$, el modelo basado en la distribución Weibull, puede aproximarse a un modelo esta vez basados en la distribución Rayleigh.

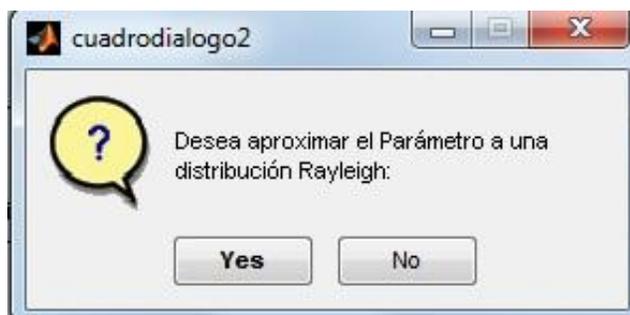


Ilustración 8. Mensaje aproximación distribución Rayleigh

El modelo realiza los cálculos de los parámetros por dos métodos: Máxima verosimilitud y Regresión lineal. Debido a una cierta cantidad de pruebas realizadas, se ha notado que el método de Máxima verosimilitud es más exacto que dicha regresión lineal. Por esto, es mediante los parámetros obtenidos por máxima verosimilitud que se realizan los cálculos y análisis necesarios para obtener el modelo. Los mensajes de aproximación entre distribuciones cuando el parámetro β se encuentra en un rango determinado se realizan basados en el parámetro de forma obtenido por el método de máxima verosimilitud, no por el de regresión lineal.

Calculado el modelo de confiabilidad para una base de datos en cuestión, se tiene la posibilidad de realizar dos cálculos útiles, por medio de una selección la cual se realiza por medio de un Pop-Up Menú, tales como: “Confiabilidad dado un tiempo t” y “Tiempo dada una confiabilidad”.

- ❖ Confiabilidad dado un tiempo t: Con esta herramienta, es posible establecer la confiabilidad del modelo en un tiempo determinado que el usuario ingresa previamente. Es claro que, al no tomarse en cuenta tiempos negativos, el usuario se encuentra inhabilitado para ingresar valores menores que 0. Si esto llegase a ocurrir, el entorno programado arroja inmediatamente un mensaje de advertencia, obligando al usuario cambiar el valor en el campo de “Tiempo”. Para el campo de confiabilidad se puede observar que se tiene un mensaje “Error”.



Ilustración 9. Mensaje de advertencia para tiempos negativos

- ❖ Tiempo dada una confiabilidad: En esta selección, el usuario determina el tiempo (en horas, días, semanas, meses, etc. ó según se encuentre la base de datos) en el cual, el equipo al cual se le está realizando el modelo, llega a una

confiabilidad determinada. Dicha herramienta, es útil para determinar el mantenimiento y el lapso de tiempo donde dicho equipo operará a una mayor confiabilidad. La confiabilidad para este entorno programado, debe estar en el rango $[0, 1]$, por esto se tienen mensajes de advertencia si se llega a ingresar confiabilidades por encima de 1 o negativas.



Ilustración 10. Mensaje de advertencia para ingreso de confiabilidad negativa



Ilustración 11. Mensaje de advertencia cuando se ingresa una Confiabilidad mayor que 1.

Presentación de los Resultados

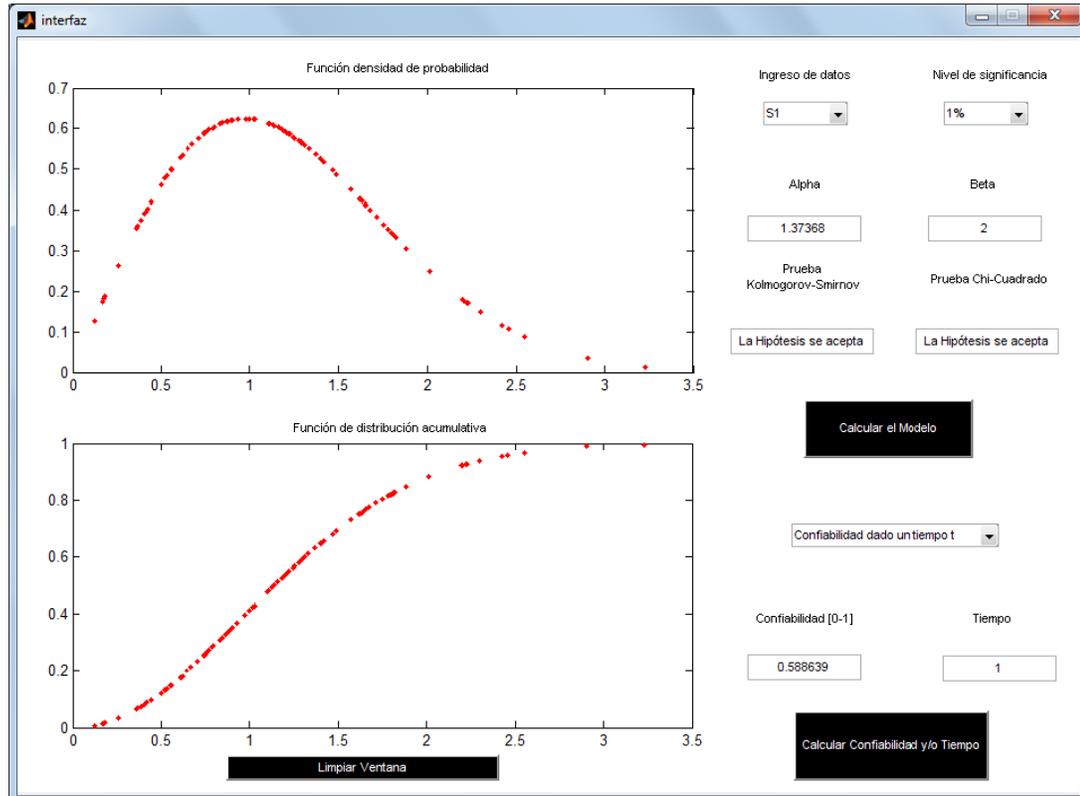


Ilustración 12. Presentación de los resultados mediante la interfaz gráfica

La ilustración 12 evidencia la presentación de los datos. Se puede observar claramente las dos graficas obtenidas, función densidad de probabilidad y función de distribución acumulativa. Como bien se aclaró en los subcapítulos anteriores, el proceso de simulación consta en: cargar datos, calcular el modelo y finalmente calcular confiabilidad y/o tiempo.

4. RESULTADOS Y ANÁLISIS

El entorno programado fue diseñado principalmente para el análisis de confiabilidad, es por esto que fue necesario realizar una cantidad de pruebas aleatorias con el fin de verificar la funcionalidad y consistencia del código de programación. Por tal motivo, se mostrarán una serie de casos simulados con bases de datos generadas aleatoriamente las cuales evidencian el proceso llevado a cabo en la construcción del entorno.

4.1 RESULTADOS INICIALES

En la etapa inicial de simulaciones en masa, se tomaron bases de datos de números generados aleatoriamente, por esto no se tienen referencia de los parámetros de la distribución. El objetivo principal era observar la gráfica obtenida para la función de densidad de probabilidad y la acumulativa debido al cálculo del modelo de los respectivos datos. Por otro lado, se estudió el comportamiento de las pruebas de bondad de ajuste y su respectivo resultado de hipótesis, es decir, verificar si ambas pruebas funcionaban correctamente y arrojaban los mismos resultados.

Los datos se obtienen por medio de “Microsoft Excel”, una aplicación para manejar hojas de cálculo. Se obtuvieron diferentes muestras de datos usando la función de Excel “ALEATORIO.ENTRE”, donde se fija un valor inferior y superior y se devuelve un número aleatorio entre los números que se especifique.

Se resalta que para esta etapa de prueba no se había desarrollado una interfaz gráfica para el usuario, por ende los datos no se presentaban de una manera amigable con el usuario.

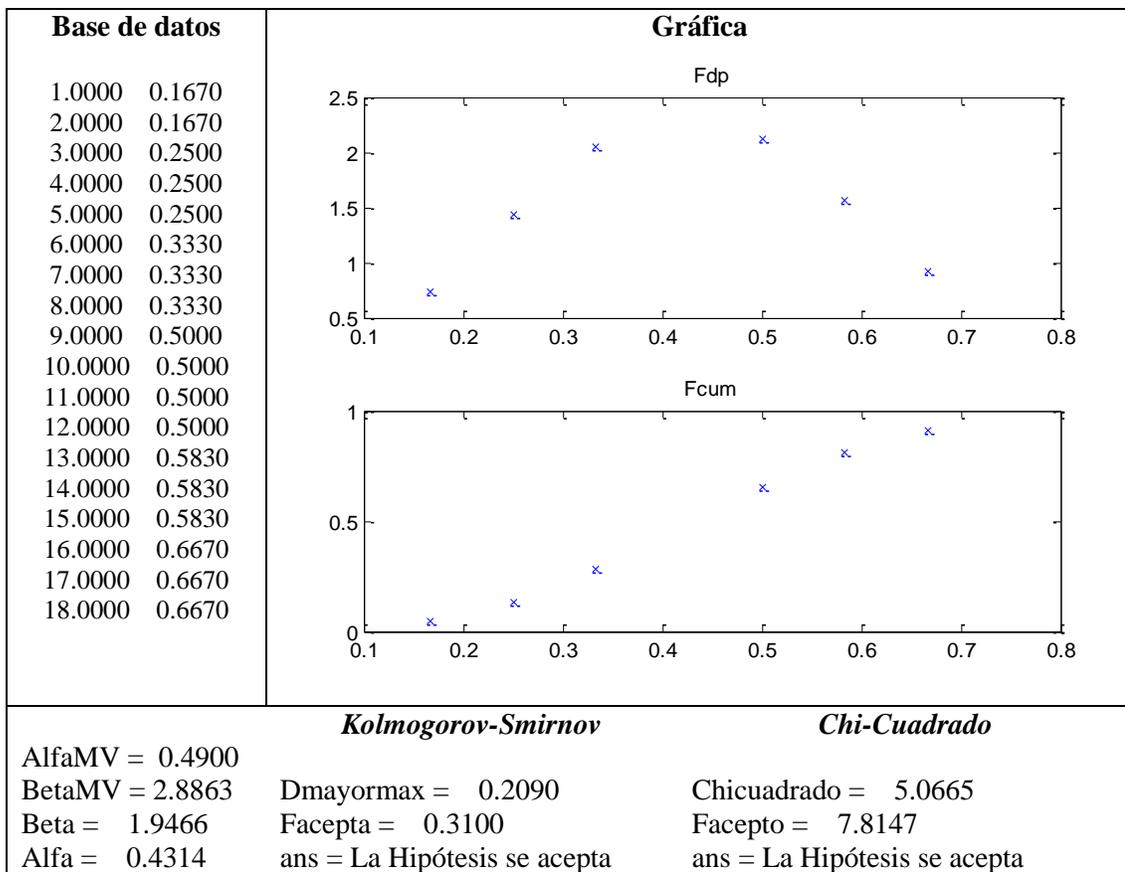
A continuación, se mostrarán 5 casos diferentes de simulación de los 30 realizados, con los respectivos resultados tanto para los parámetros como para sus respectivas pruebas de bondad de ajuste.

Dado el hecho de no poseer un respaldo en cuanto a los parámetros de la distribución con el cual contrastar los valores obtenidos por el estimador de parámetros “Máxima

Verosimilitud”, fue necesario realizar una programación anexa cuyo objetivo fuera arrojar un cálculo diferente de los parámetros α y β , para así comparar ambos resultados y evaluar el funcionamiento del entorno programado.

Los cálculos realizados para contrastar con los obtenidos por Máxima Verosimilitud fueron realizados mediante una Linealización de la función de Weibull.

CASO 1

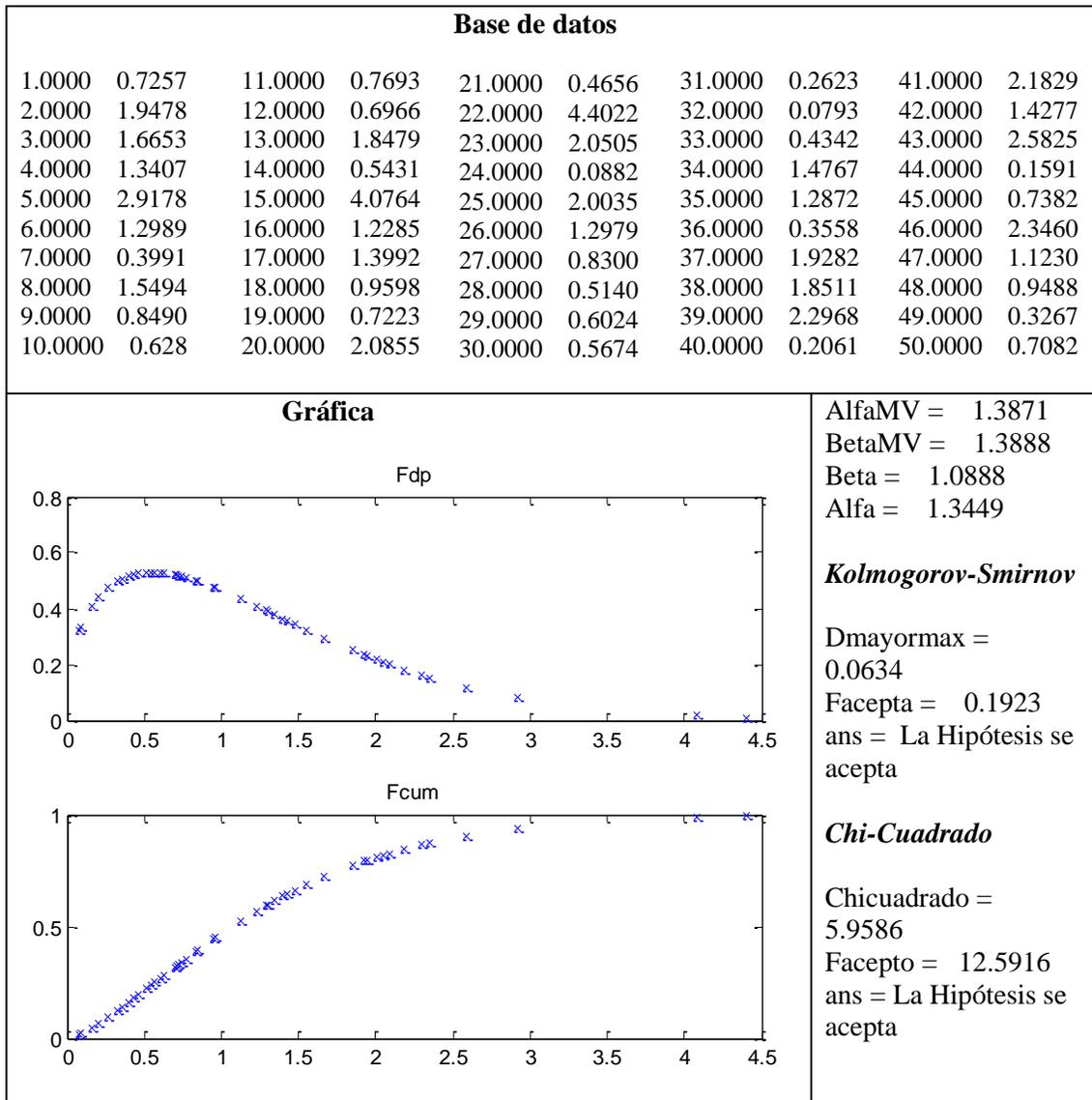


- ✓ El caso 1 se simuló con una muestra de 18 datos.
- ✓ Se encuentran similitudes en el parámetro de escala “ α ” obtenido tanto por el método de Máxima Verosimilitud como por la linealización.
- ✓ El parámetro de forma “ β ” presenta diferencias considerables, sin embargo, se toma como guía el resultado arrojado por el método de Máxima Verosimilitud,

dada la gráfica y la cantidad de datos se concluye que el parámetro no es lo suficientemente exacto.

- ✓ Las pruebas de bondad de ajuste aprueban claramente el modelo, permitiendo el tratamiento de los datos partiendo desde la ley de distribución Weibull

CASO 2

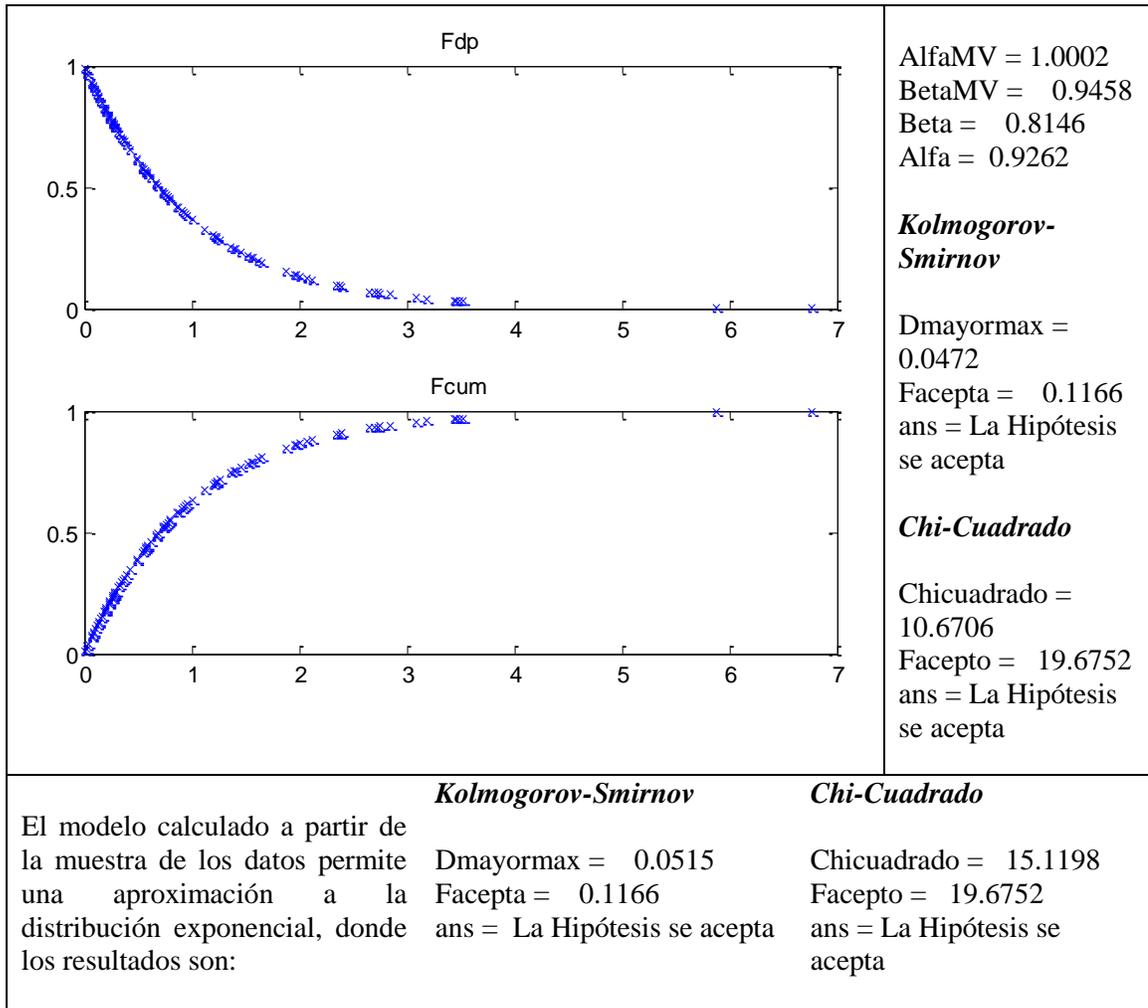


- ✓ El caso 2 se simuló con una muestra de 50 datos.

- ✓ Se encuentran similitudes en el parámetro de escala “ α ” obtenido tanto por el método de Máxima Verosimilitud como por la linealización.
- ✓ El parámetro de forma “ β ” presenta diferencias, sin embargo, se toma como guía el resultado arrojado por el método de Máxima Verosimilitud. Esta vez se puede observar que la cantidad de datos ayudo a aproximarse un poco más que en el caso anterior.
- ✓ Las pruebas de bondad de ajuste aprueban claramente el modelo, permitiendo el tratamiento de los datos partiendo desde la ley de distribución Weibull

CASO 3

| Base de datos | | | | | | | | | |
|---------------|--------|---------|--------|---------|--------|----------|--------|----------|--------|
| 1.0000 | 2.7296 | 28.0000 | 0.1956 | 55.0000 | 2.7042 | 82.0000 | 0.9370 | 109.0000 | 1.2312 |
| 2.0000 | 3.4615 | 29.0000 | 2.0113 | 56.0000 | 0.3201 | 83.0000 | 0.2328 | 110.0000 | 0.3965 |
| 3.0000 | 0.0110 | 30.0000 | 0.5681 | 57.0000 | 0.8004 | 84.0000 | 1.5551 | 111.0000 | 0.3663 |
| 4.0000 | 0.1636 | 31.0000 | 0.2929 | 58.0000 | 0.2750 | 85.0000 | 3.0815 | 112.0000 | 0.1248 |
| 5.0000 | 0.2956 | 32.0000 | 0.6143 | 59.0000 | 0.9335 | 86.0000 | 2.0030 | 113.0000 | 0.0689 |
| 6.0000 | 0.4965 | 33.0000 | 0.2413 | 60.0000 | 0.5774 | 87.0000 | 0.5847 | 114.0000 | 0.0782 |
| 7.0000 | 0.9559 | 34.0000 | 0.3720 | 61.0000 | 0.5783 | 88.0000 | 2.7452 | 115.0000 | 1.8810 |
| 8.0000 | 1.1250 | 35.0000 | 1.3590 | 62.0000 | 0.0434 | 89.0000 | 0.6211 | 116.0000 | 0.7654 |
| 9.0000 | 1.5713 | 36.0000 | 1.9648 | 63.0000 | 0.0745 | 90.0000 | 0.0129 | 117.0000 | 0.6890 |
| 10.0000 | 0.4863 | 37.0000 | 0.0310 | 64.0000 | 1.4579 | 91.0000 | 3.5187 | 118.0000 | 0.6639 |
| 11.0000 | 0.9080 | 38.0000 | 1.3709 | 65.0000 | 1.3916 | 92.0000 | 0.5445 | 119.0000 | 1.4096 |
| 12.0000 | 0.1896 | 39.0000 | 0.6649 | 66.0000 | 1.1165 | 93.0000 | 1.5210 | 120.0000 | 0.0303 |
| 13.0000 | 0.1102 | 40.0000 | 0.7786 | 67.0000 | 0.3917 | 94.0000 | 2.3743 | 121.0000 | 0.1984 |
| 14.0000 | 0.3525 | 41.0000 | 0.0932 | 68.0000 | 0.0075 | 95.0000 | 1.4641 | 122.0000 | 0.5876 |
| 15.0000 | 0.5552 | 42.0000 | 0.2587 | 69.0000 | 1.4660 | 96.0000 | 0.5940 | 123.0000 | 2.8507 |
| 16.0000 | 1.0092 | 43.0000 | 0.8708 | 70.0000 | 0.3967 | 97.0000 | 1.6239 | 124.0000 | 0.2771 |
| 17.0000 | 0.3285 | 44.0000 | 0.0931 | 71.0000 | 0.1316 | 98.0000 | 0.0163 | 125.0000 | 0.7909 |
| 18.0000 | 0.2094 | 45.0000 | 1.2316 | 72.0000 | 2.3439 | 99.0000 | 0.1568 | 126.0000 | 0.6711 |
| 19.0000 | 0.8019 | 46.0000 | 1.2279 | 73.0000 | 0.5876 | 100.0000 | 0.4236 | 127.0000 | 0.2889 |
| 20.0000 | 0.0865 | 47.0000 | 0.2700 | 74.0000 | 3.4847 | 101.0000 | 0.0869 | 128.0000 | 1.5693 |
| 21.0000 | 0.5784 | 48.0000 | 1.2014 | 75.0000 | 3.1789 | 102.0000 | 2.3914 | 129.0000 | 2.0797 |
| 22.0000 | 0.7409 | 49.0000 | 1.9779 | 76.0000 | 0.6747 | 103.0000 | 6.7622 | 130.0000 | 2.6516 |
| 23.0000 | 0.8741 | 50.0000 | 0.0806 | 77.0000 | 0.2620 | 104.0000 | 0.5539 | 131.0000 | 2.1224 |
| 24.0000 | 0.7998 | 51.0000 | 1.6532 | 78.0000 | 0.2829 | 105.0000 | 3.4481 | 132.0000 | 0.2284 |
| 25.0000 | 1.5587 | 52.0000 | 0.1451 | 79.0000 | 1.2705 | 106.0000 | 5.8828 | 133.0000 | 0.3311 |
| 26.0000 | 0.2667 | 53.0000 | 1.9821 | 80.0000 | 0.5471 | 107.0000 | 1.2180 | 134.0000 | 0.7303 |
| 27.0000 | 2.7011 | 54.0000 | 0.0403 | 81.0000 | 0.7364 | 108.0000 | 0.1894 | 135.0000 | 0.1996 |



- ✓ El caso 3 se simuló con una muestra de 135 datos.
- ✓ Se encuentran similitudes en el parámetro de escala “ α ” obtenido tanto por el método de Máxima Verosimilitud como por la linealización.
- ✓ El parámetro de forma “ β ” presenta concordancia alguna por ambos métodos. Es claro que debido al tamaño de la muestra, la linealización arroja un parámetro un poco más cercano al arrojado por máxima verosimilitud el cual presenta gran exactitud.
- ✓ Las pruebas de bondad de ajuste aprueban claramente el modelo, permitiendo el tratamiento de los datos partiendo desde la ley de distribución Weibull
- ✓ Según el resultado de los parámetros, los datos puede ser simulados a partir de la Ley Exponencial. El entorno programado detecta si los datos pueden o no ser

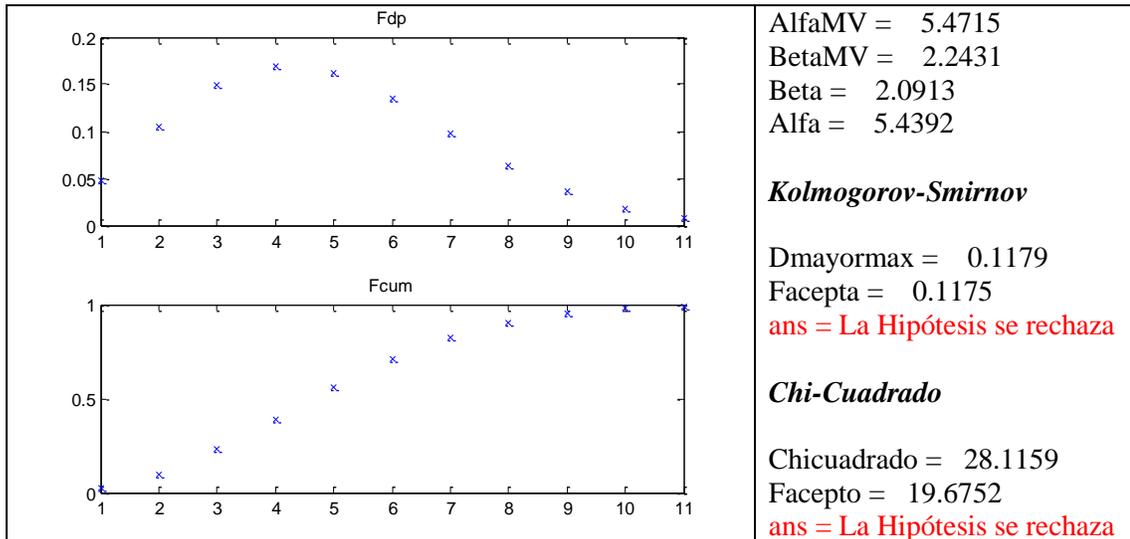
simulados de esta forma pero es el usuario quien decide realizar dicha aproximación.

- ✓ Las pruebas de bondad de ajuste aprueban el modelo calculado por la ley Exponencial.

CASO 4

El caso 4, consta de una distribución diferente a la Weibull, con un tamaño de 134 datos los cuales son simulados bajo el modelo establecido y se puede establecer que los datos no corresponden a la distribución Weibull, dado a la naturaleza de los datos y por otro lado a que ambas pruebas de bondad de ajuste rechazan rotundamente el modelo calculado.

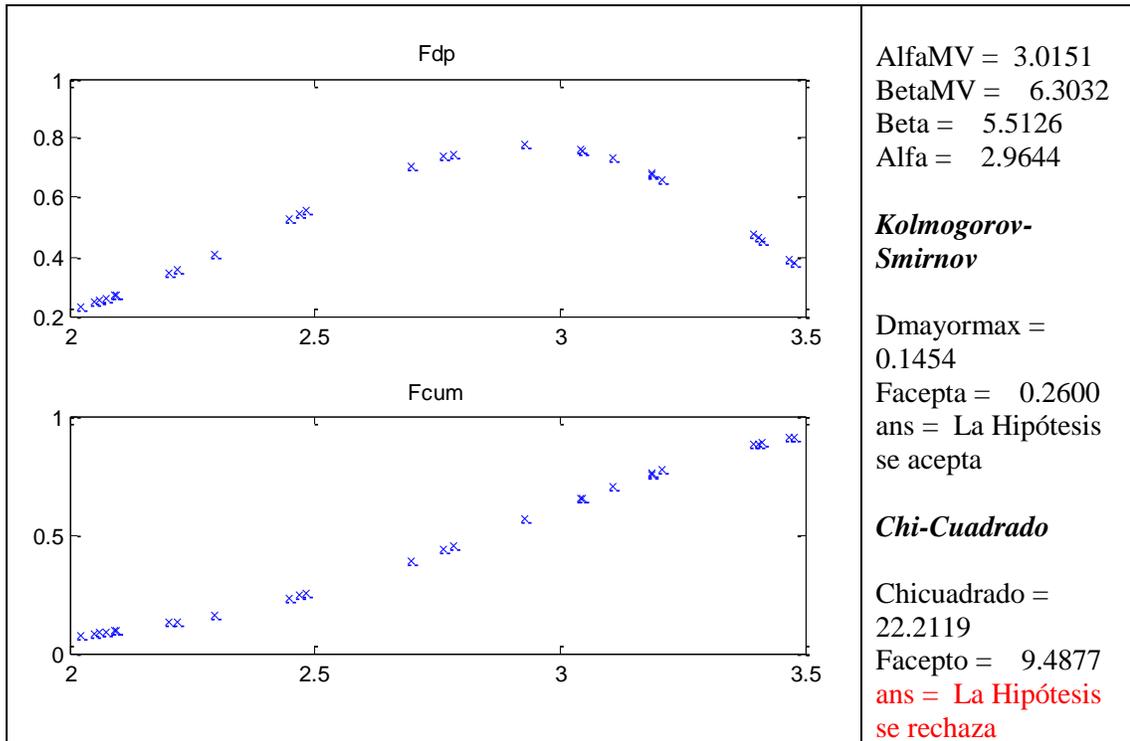
| Base de datos | | | | | | | | | |
|---------------|----|----|---|----|---|-----|----|-----|----|
| 1 | 4 | 28 | 5 | 55 | 5 | 82 | 4 | 109 | 3 |
| 2 | 1 | 29 | 8 | 56 | 6 | 83 | 3 | 110 | 3 |
| 3 | 7 | 30 | 3 | 57 | 1 | 84 | 4 | 111 | 4 |
| 4 | 5 | 31 | 8 | 58 | 5 | 85 | 3 | 112 | 5 |
| 5 | 3 | 32 | 5 | 59 | 7 | 86 | 2 | 113 | 6 |
| 6 | 5 | 33 | 5 | 60 | 1 | 87 | 3 | 114 | 6 |
| 7 | 7 | 34 | 2 | 61 | 7 | 88 | 7 | 115 | 7 |
| 8 | 5 | 35 | 3 | 62 | 6 | 89 | 4 | 116 | 4 |
| 9 | 6 | 36 | 3 | 63 | 5 | 90 | 4 | 117 | 3 |
| 10 | 6 | 37 | 5 | 64 | 5 | 91 | 1 | 118 | 7 |
| 11 | 6 | 38 | 5 | 65 | 4 | 92 | 8 | 119 | 3 |
| 12 | 1 | 39 | 8 | 66 | 2 | 93 | 2 | 120 | 6 |
| 13 | 3 | 40 | 3 | 67 | 2 | 94 | 2 | 121 | 7 |
| 14 | 3 | 41 | 7 | 68 | 2 | 95 | 5 | 122 | 6 |
| 15 | 6 | 42 | 7 | 69 | 3 | 96 | 8 | 123 | 6 |
| 16 | 6 | 43 | 2 | 70 | 1 | 97 | 6 | 124 | 9 |
| 17 | 7 | 44 | 6 | 71 | 7 | 98 | 2 | 125 | 4 |
| 18 | 6 | 45 | 6 | 72 | 4 | 99 | 3 | 126 | 8 |
| 19 | 8 | 46 | 6 | 73 | 5 | 100 | 4 | 127 | 6 |
| 20 | 2 | 47 | 2 | 74 | 5 | 101 | 5 | 128 | 5 |
| 21 | 8 | 48 | 6 | 75 | 6 | 102 | 10 | 129 | 2 |
| 22 | 5 | 49 | 1 | 76 | 3 | 103 | 11 | 130 | 3 |
| 23 | 5 | 50 | 4 | 77 | 7 | 104 | 2 | 131 | 10 |
| 24 | 6 | 51 | 3 | 78 | 3 | 105 | 11 | 132 | 4 |
| 25 | 3 | 52 | 6 | 79 | 1 | 106 | 7 | 133 | 4 |
| 26 | 11 | 53 | 6 | 80 | 6 | 107 | 2 | 134 | 5 |
| 27 | 3 | 54 | 2 | 81 | 8 | 108 | 8 | | |



- ✓ El caso 4 se simuló con una muestra de 134 datos.
- ✓ Se encuentran similitudes tanto el parámetro de escala “ α ” como para el parámetro “ β ” obtenido por los métodos de Máxima Verosimilitud y Linealización de Weibull.
- ✓ El tamaño de la muestra hace más exacto el cálculo de los parámetros por ambos métodos, siendo más cercanos entre sí y respaldando la validez de los valores obtenidos.
- ✓ Las pruebas de bondad de ajuste rechazan el modelo, corroborándose así que dicha muestra no puede ser modelada bajo la Ley de distribución Weibull ya que las pruebas de bondad de ajuste son muy concretas al rechazar una hipótesis.

CASO 5

| Base de datos | | | | | |
|---------------|--------|---------|--------|---------|--------|
| 1.0000 | 3.4135 | 11.0000 | 2.0946 | 21.0000 | 3.3942 |
| 2.0000 | 2.7632 | 12.0000 | 3.3946 | 22.0000 | 3.1071 |
| 3.0000 | 2.4834 | 13.0000 | 2.0218 | 23.0000 | 2.0498 |
| 4.0000 | 3.0411 | 14.0000 | 2.7819 | 24.0000 | 3.1876 |
| 5.0000 | 3.4672 | 15.0000 | 2.6970 | 25.0000 | 3.4043 |
| 6.0000 | 2.2189 | 16.0000 | 3.4762 | 26.0000 | 2.2950 |
| 7.0000 | 3.3942 | 17.0000 | 3.4770 | 27.0000 | 3.1859 |
| 8.0000 | 2.0733 | 18.0000 | 3.2079 | 28.0000 | 2.9284 |
| 9.0000 | 3.0467 | 19.0000 | 2.2018 | 29.0000 | 2.0915 |
| 10.0000 | 2.4670 | 20.0000 | 2.0604 | 30.0000 | 2.4461 |



- ✓ El caso 5 se simuló con una muestra de 30 datos.
- ✓ Se encuentran similitudes en el parámetro de escala “ α ” obtenido tanto por el método de Máxima Verosimilitud como por la linealización.
- ✓ El parámetro de forma “ β ” presenta diferencias ambos métodos. Dado el tamaño de la muestra, la linealización arroja un valor no tan exacto, lo cual se puede observar al comparar dicho valor con el obtenido por máxima verosimilitud el cual presenta gran precisión.
- ✓ Las pruebas de bondad de ajuste presentan diferencias ya que la prueba Chi-cuadrado rechaza el modelo. Sin embargo el test de Kolmogorov-Smirnov lo acepta, concluyendo así que se está cometiendo un error de tipo I. La solución a este inconveniente puede ser aumentar el nivel de significancia.

4.2 RESULTADOS FINALES

Para los resultados finales, de antemano se desarrollo la interfaz gráfica, esto con el fin de facilitar la interacción entre el usuario y el entorno programado. Además se introdujo un nuevo cálculo el cual es plenamente decisión del usuario, siendo de gran ayuda ya que se puede establecer la confiabilidad de un equipo o del modelo dado un determinado tiempo, o viceversa, dado un tiempo t la confiabilidad del equipo.

Los datos analizados en la presente sección, son obtenidos mediante el software “EasyFit”, un programa de ajuste de distribuciones. EasyFit permite ajustar automáticamente las distribuciones a los datos de la muestra y seleccionar mejor un modelo en poco tiempo. Está diseñado para hacer el análisis de los datos de la manera más fácil posible. El programa de simulación esta desarrollado por “Mathwave data analysis & simulation” [31].

En esta segunda etapa de simulaciones se ejecutaron 30 casos diferentes. Las bases de datos fueron obtenidas de un programa dedicado a la confiabilidad, es por esto que se tienen diferentes tamaños de muestras y cada uno de los parámetros α y β reales para ser contrastados con los calculados por el método de máxima verosimilitud. En esta ocasión la Linealización de Weibull fue descartada ya que, como se concluyo anteriormente el tamaño de la muestra afecta considerablemente el valor arrojado por este método para el parámetro de forma β .

La tabla 2 mostrada a continuación, es una comparación entre los datos reales y los calculados por el entrono programado. No se anexa la muestra total dado que no es útil para los análisis requeridos.

Una tercera etapa de simulaciones se ejecutó para analizar y revisar el comportamiento del entorno programado cuando la bases de datos no puede ser modelada o simulada bajo una distribución Weibull. En este caso, la hipótesis nula debe ser rechazada por los métodos de pruebas de bondad de ajuste tomados para el análisis de los datos. En la tabla 4 se pueden ver los resultados de dichos análisis.

| BASE DE DATOS | | N (Tamaño muestra) | ORGANIZACIÓN DE LOS DATOS | | | PARÁMETROS | | | | | | Nivel de Significancia [%] | PRUEBAS DE BONDAD DE AJUSTE | | | |
|---------------|-----------|--------------------|---------------------------|----|----|------------|---------|-----------------------|---------|----------------------|---------|----------------------------|-----------------------------|---------|--------------|---------|
| | | | | | | Real | | Calculado por EasyFit | | Máxima Verosimilitud | | | Kolmogorov-Smirnov | | Chi-Cuadrado | |
| | | | S1 | S2 | S3 | α | β | α | β | α | β | | Acepta | Rechaza | Acepta | Rechaza |
| B.1 | Sin aprox | 200 | x | | | 1 | 1 | 1,016 | 1,079 | 1,033 | 1,053 | 0,05 | x | | | x |
| | Con aprox | | | | | | | 1,016 | 1,000 | 1,033 | 1,000 | | x | | x | |
| B.2 | N.A | 155 | x | | | 7 | 1,5 | 7,363 | 1,457 | 7,303 | 1,568 | 0,05 | x | | | x |
| B.3 | N.A | 15 | x | | | 2 | 5,8 | 1,843 | 5,555 | 1,891 | 5,756 | 0,05 | x | | x | |
| B.4 | N.A | 64 | x | | | 3,8 | 2,9 | 3,955 | 3,280 | 4,007 | 3,265 | 0,05 | x | | x | |
| B.5 | N.A | 25 | x | | | 0,5 | 0,5 | 0,296 | 0,592 | 0,372 | 0,541 | 0,05 | x | | x | |
| B.6 | N.A | 46 | x | | | 0,8 | 2 | 0,769 | 2,230 | 0,791 | 2,185 | 0,05 | x | | x | |
| B.7 | N.A | 5 | x | | | 2 | 1,5 | 1,686 | 1,199 | 1,831 | 2,202 | 0,05 | x | | x | |
| B.8 | N.A | 85 | x | | | 0,1 | 4 | 0,106 | 4,592 | 0,107 | 4,387 | 0,05 | x | | x | |
| B.9 | N.A | 74 | x | | | 1,2 | 1,8 | 1,273 | 1,796 | 1,300 | 1,823 | 0,05 | x | | x | |
| B.10 | Sin aprox | 100 | x | | | 1,4 | 2 | 1,364 | 1,866 | 1,374 | 1,966 | 0,05 | x | | x | |
| | Con aprox | | | | | | | 1,364 | 2,000 | 1,374 | 2,000 | | x | | x | |
| B.10 | Sin aprox | 100 | x | | | 1,4 | 2 | 1,364 | 1,866 | 1,374 | 1,966 | 0,01 | x | | x | |
| | Con aprox | | | | | | | 1,364 | 2,000 | 1,374 | 2,000 | | x | | x | |
| B.11 | N.A | 20 | | x | | 3,2 | 2,7 | 2,980 | 2,786 | 3,057 | 3,159 | 0,05 | x | | x | |
| B.12 | Sin aprox | 32 | | x | | 3,2 | 0,96 | 2,383 | 1,046 | 2,618 | 1,016 | 0,05 | x | | x | |
| | Con aprox | | | | | | | 2,383 | 1,000 | 2,618 | 1,000 | | x | | x | |
| B.13 | N.A | 62 | | x | | 6,8 | 2,3 | 6,654 | 2,105 | 6,648 | 2,437 | 0,1 | x | | x | |
| B.14 | N.A | 160 | | x | | 10 | 0,1 | 16,391 | 0,083 | 19,656 | 0,086 | 0,05 | x | | | x |
| B.15 | N.A | 76 | | x | | 0,01 | 1,3 | 0,012 | 1,066 | 0,012 | 1,254 | 0,01 | x | | x | |
| B.16 | N.A | 50 | | x | | 0,5 | 2,4 | 0,499 | 2,105 | 0,504 | 2,301 | 0,05 | x | | x | |

| BASE DE DATOS | | N (Tamaño muestra) | ORGANIZACIÓN DE LOS DATOS | | | PARÁMETROS | | | | | | Nivel de Significancia [%] | PRUEBAS DE BONDAD DE AJUSTE | | | |
|---------------|------------|--------------------|---------------------------|----------|----|------------|------------|-----------------------|--------------|----------------------|--------------|----------------------------|-----------------------------|---------|--------------|----------|
| | | | | | | Real | | Calculado por EasyFit | | Máxima Verosimilitud | | | Kolmogorov-Smirnov | | Chi-Cuadrado | |
| | | | S1 | S2 | S3 | α | β | α | β | α | β | | Acepta | Rechaza | Acepta | Rechaza |
| B.17 | N.A | 270 | | x | | 5 | 5 | 4,980 | 4,369 | 4,952 | 4,972 | 0,05 | x | | x | |
| B.18 | Sin aprox | 57 | | x | | 3,7 | 2 | 3,879 | 1,625 | 3,828 | 1,910 | 0,05 | x | | x | |
| | Con aprox | | | | | | | 3,879 | 2,000 | 3,828 | 2,000 | | x | | x | |
| B.18 | Sin aprox | 57 | | x | | 3,7 | 2 | 3,879 | 1,625 | 3,828 | 1,910 | 0,01 | x | | x | |
| | Con aprox | | | | | | | 3,879 | 2,000 | 3,828 | 2,000 | | x | | x | |
| B.19 | N.A | 45 | | x | | 0,5 | 0,5 | 0,421 | 0,435 | 0,451 | 0,474 | 0,05 | x | | x | |
| B.20 | N.A | 52 | | x | | 2,5 | 6,4 | 2,367 | 5,054 | 2,310 | 5,721 | 0,05 | x | | | x |
| B.21 | N.A | 40 | | | x | 1 | 2 | 1,025 | 1,489 | 1,322 | 7,389 | 0,05 | x | | x | |
| B.22 | N.A | 112 | | | x | 2 | 1 | 1,836 | 0,955 | 1,217 | 4,967 | 0,05 | x | | x | |
| B.23 | N.A | 46 | | | x | 0,5 | 3,8 | 0,373 | 5,889 | 1,025 | 2,734 | 0,05 | x | | x | |
| B.24 | N.A | 90 | | | x | 1,6 | 1,4 | 1,620 | 1,345 | 1,063 | 3,595 | 0,05 | x | | x | |
| B.25 | N.A | 66 | | | x | 3,5 | 2,1 | 3,747 | 2,179 | 5,009 | 7,235 | 0,05 | x | | x | |

Tabla 5. Resultados finales para diferentes tipos de distribuciones Weibull

La fuente de los 25 casos relacionados en la tabla 5 fueron obtenidos mediante el programa de simulación y análisis de confiabilidad EasyFit.

Análisis de resultados finales

Los 25 casos simulados relacionados en la tabla 2 corresponden a bases de datos generadas aleatoriamente. Las 25 bases de datos son distribuciones Weibull y se tiene el valor del parámetro α y β real, por lo tanto los valores calculados por el estimador de parámetros Máxima Verosimilitud pueden ser comparados con los valores reales y así obtener un error relativo y absoluto de los resultados obtenidos. El entorno programado se basa en la distribución Weibull y cada una de los datos de las diferentes bases proviene de una distribución Weibull, por esto se garantiza de antemano que para las 25 bases de datos diferentes las pruebas de bondad de ajuste deben aceptar la hipótesis nula. En caso contrario, se ratifica que se está cometiendo un Error de tipo I donde se rechaza una hipótesis nula verdadera.

Las siguientes bases de datos son mencionadas ya que se presentan como casos especiales de los 25 casos de simulación diferentes:

❖ Base 1

Consta de 200 datos generados aleatoriamente por medio del EasyFit. Se puede observar que la hipótesis nula es aceptada por el test de Kolmogorov-Smirnov, sin embargo la prueba de Chi-Cuadrado la rechaza en su primer análisis cuando el modelo se aproxima a la Ley de distribución Weibull. Debido al valor que toma el parámetro β se puede realizar una aproximación a la Ley exponencial, aceptando el test de Chi-Cuadrado la hipótesis nula. De esta forma se ha eliminado el error tipo I.

❖ Base 2

Tamaño de la muestra de 155 datos. El valor de los parámetros α y β calculados son valores cercanos a los valores reales, pero la hipótesis nula es rechazada por el test Chi-Cuadrado. Esto se convierte en un error de tipo I ya que se está rechazando una hipótesis que de antemano, es verdadera. En este caso se tiene un parámetro de escala alto. Se recomienda basarse en el test de Kolmogorov-Smirnov, siendo este una herramienta más potente.

❖ Base 10

Se tiene una base de datos de 100 valores diferentes. El tratamiento de los datos seleccionado es S1 y se realizaron dos modelos para niveles de significancia

diferentes, 1% y 5%. Debe agregarse que el valor de los parámetros en ambos modelos no cambia, ya que Máxima Verosimilitud es independiente del Nivel de Significancia. La distancia de Kolmogorov-Smirnov y el valor de Pearson o Chi-Cuadrado van a ser comparados con un valor de aceptación diferente ya que este si depende del nivel de significancia. Si con un nivel de significancia del 1% se acepta una hipótesis nula, se garantiza que con los otros niveles (mayores) se acepte también dicha hipótesis.

❖ **Base 11**

En este caso, se tiene un tamaño de muestra igual a 20. Aunque la hipótesis nula se acepta por ambas pruebas de bondad de ajuste, se nota que a medida que la muestra es pequeña (pocos datos), el error o diferencia entre el valor real y el calculado de los parámetros α y β es mayor.

❖ **Base 14**

Base de datos de tamaño 160. El parámetro de escala calculado posee un error del 20% respecto al parámetro de escala real. El valor del parámetro de forma calculado, presenta un error del 3%, por lo que se considera similar al valor real. Pese a ambas circunstancias de la poca exactitud de los cálculos de los parámetros para el modelo de Weibull, la prueba de Kolmogorov-Smirnov acepta el modelo y establece que dichos datos pueden ser tratados como una distribución Weibull bajo los parámetros calculados. El test Chi-Cuadrado rechaza la hipótesis nula, siendo este un error tipo I.

❖ **Base 20**

El tamaño de la muestra de los datos es de 52. En este caso los valores de los parámetros calculados no excede el 15% de error, sin embargo, la prueba Chi-Cuadrado rechaza la hipótesis nula, cometiendo así un error tipo I ya que la prueba de bondad de ajuste de Kolmogorov-Smirnov acepta la hipótesis nula. Se concluye que es error tipo I ya que los datos efectivamente pertenecen a una Ley de distribución Weibull.

❖ **Base 21-25**

Debido al tratamiento de los datos realizado en estas últimas bases de datos, se tiene resultados diferentes para los valores calculados de los parámetros α y β ya que se agrupan los datos de la muestra por intervalos y se trabaja con un único valor representativo, reduciéndose así el tamaño de la muestra. Sin embargo las pruebas de

bondad de ajuste aceptan la hipótesis nula, significando esto que se acepta un modelo basado en la distribución Weibull.

Errores de cálculo

La tabla 3 mostrada a continuación, se puede resumir como el error entre el parámetro dado por EasyFit y el parámetro calculado. Se tiene tanto el error absoluto como el error relativo, los cuales se definen de la siguiente manera:

- ❖ Error Absoluto: Es la diferencia entre el valor real y el valor calculado. En este caso es la resta entre el parámetro α de la muestra con el valor del parámetro calculado α por el estimador de parámetros de máxima verosimilitud.
- ❖ Error Relativo: Se define como el cociente entre el error absoluto y el valor real del parámetro. Dicho error se representa en porcentaje, permitiendo así al usuario definir qué tan cerca o tan lejos se encuentra un valor calculado del valor real.

| | Valor EasyFit | | Valor Calculado | | Error absoluto | | Error relativo [%] | |
|------|---------------|---------|-----------------|---------|----------------|---------|--------------------|---------|
| | α | β | α | β | α | β | α | β |
| B.1 | 1,016 | 1,079 | 1,033 | 1,053 | 0,017 | 0,026 | 1,7 | 2,4 |
| B.2 | 7,363 | 1,457 | 7,303 | 1,568 | 0,0605 | 0,1117 | 0,8 | 7,7 |
| B.3 | 1,843 | 5,555 | 1,891 | 5,756 | 0,0488 | 0,2006 | 2,6 | 3,6 |
| B.4 | 3,955 | 3,280 | 4,007 | 3,265 | 0,052 | 0,0151 | 1,3 | 0,5 |
| B.5 | 0,296 | 0,592 | 0,372 | 0,541 | 0,07566 | 0,05111 | 25,5 | 8,6 |
| B.6 | 0,769 | 2,230 | 0,791 | 2,185 | 0,02196 | 0,0443 | 2,9 | 2,0 |
| B.7 | 1,686 | 1,199 | 1,831 | 2,202 | 0,1449 | 1,003 | 8,6 | 83,7 |
| B.8 | 0,106 | 4,592 | 0,107 | 4,387 | 0,00098 | 0,2051 | 0,9 | 4,5 |
| B.9 | 1,273 | 1,796 | 1,300 | 1,823 | 0,0266 | 0,0275 | 2,1 | 1,5 |
| B.10 | 1,364 | 1,866 | 1,374 | 1,966 | 0,0097 | 0,1002 | 0,7 | 5,4 |
| B.11 | 2,980 | 2,786 | 3,057 | 3,159 | 0,0768 | 0,3736 | 2,6 | 13,4 |
| B.12 | 2,383 | 1,046 | 2,618 | 1,016 | 0,2347 | 0,0295 | 9,8 | 2,8 |
| B.13 | 6,654 | 2,105 | 6,648 | 2,437 | 0,0063 | 0,3319 | 0,1 | 15,8 |
| B.14 | 16,391 | 0,083 | 19,656 | 0,086 | 3,265 | 0,00254 | 19,9 | 3,0 |
| B.15 | 0,012 | 1,066 | 0,012 | 1,254 | 0,00017 | 0,1876 | 1,4 | 17,6 |
| B.16 | 0,499 | 2,105 | 0,504 | 2,301 | 0,00515 | 0,1961 | 1,0 | 9,3 |

| | Valor EasyFit | | Valor Calculado | | Error absoluto | | Error relativo [%] | |
|------|---------------|---------|-----------------|---------|----------------|---------|--------------------|---------|
| | α | β | α | β | α | β | α | β |
| B.17 | 4,980 | 4,369 | 4,952 | 4,972 | 0,0282 | 0,6032 | 0,6 | 13,8 |
| B.18 | 3,879 | 1,625 | 3,828 | 1,910 | 0,0512 | 0,2854 | 1,3 | 17,6 |
| B.19 | 0,421 | 0,435 | 0,451 | 0,474 | 0,0303 | 0,03911 | 7,2 | 9,0 |
| B.20 | 2,367 | 5,054 | 2,310 | 5,721 | 0,0566 | 0,6671 | 2,4 | 13,2 |
| B.21 | 1,025 | 1,489 | 1,322 | 7,389 | 0,297 | 5,9 | 29,0 | 396,2 |
| B.22 | 1,836 | 0,955 | 1,217 | 4,967 | 0,6188 | 4,01153 | 33,7 | 419,8 |
| B.23 | 0,373 | 5,889 | 1,025 | 2,734 | 0,65179 | 3,1549 | 174,6 | 53,6 |
| B.24 | 1,620 | 1,345 | 1,063 | 3,595 | 0,5565 | 2,2499 | 34,4 | 167,3 |
| B.25 | 3,747 | 2,179 | 5,009 | 7,235 | 1,2617 | 5,056 | 33,7 | 232,0 |

Tabla 6. Errores entre el valor real y el calculado de los parámetros α y β

Los errores relativos en rojo para las bases de datos comprendidas entre la 1 y la 20 se consideran especiales por presentar altos valores de error en dichos cálculos. Sin embargo, el modelo puede ejecutarse siguiendo la distribución Weibull sin inconveniente alguno. Se nota que para B.7, debido a un tamaño de muestra $N = 5$, se tiene un error muy alto. Se recomienda tener muestras de datos mayores a 20.

Para las bases de datos desde la 21 hasta la 25 se tienen errores relativos y absolutos significativamente grandes, ya que la muestra se reduce de tamaño a \sqrt{n} donde n es el tamaño de la muestra. Esta reducción de los datos altera los resultados finales para los estimadores de parámetros. Las pruebas no rechazan la hipótesis nula, por eso se acepta el modelo entregado por el entorno programado.

Análisis de resultados para datos que no se ajustan a la Distribución Weibull

En la tabla 4 se pueden observar 5 diferentes casos los cuales no pueden ser modelados bajo la distribución Weibull ya que los datos de la muestra son representados por otro tipo de distribución diferente a la Weibull.

Las simulaciones o modelos realizados fueron realizados bajo las siguientes consideraciones iniciales:

- ✓ La organización de los datos o tratamiento de los datos asumido fue S1. Cabe resaltar que no hay discrepancia alguna al escoger el tipo de tratamiento de los datos y se escogió por simple arbitrariedad.
- ✓ El nivel de significancia siempre fue del 1%, siendo este el más estricto, ya que valores pequeños en el nivel de significancia establece una mayor confianza. Es notable que si el nivel de significancia es pequeño y rechaza una hipótesis nula, a niveles más altos con mayor razón se rechaza ya que disminuye la confianza.

Claramente se puede diferenciar en la tabla que solo la base de datos 27 presenta inconvenientes a la hora de realizar las pruebas de bondad de ajuste, ya que el teste de Kolmogorov-Smirnov está aceptando una hipótesis que sin lugar a duda es falsa. Dicho error es conocido como Error tipo II ó falso negativo. La distribución que describe de una manera más exacta los datos analizados es la Wakeby. Sin embargo, para este análisis de la base de datos 27 se tiene una muestra de tamaño 34. Claramente el tamaño de la muestra es crucial tanto a la hora de realizar la estimación de los parámetros como a la hora de ejecutar pruebas de bondad de ajuste.

La base de datos 28 se considera una ampliación de la base de datos 27 ya que los datos pertenecen al mismo tipo de distribución, con los mismos parámetros pero con un tamaño de muestra mayor que el anterior. La base de datos 27 contiene 34 datos mientras que la base de datos 28 tiene 98 datos. A mayor cantidad de datos más precisión se puede lograr en las pruebas de bondad de ajuste.

Téngase en cuenta que una de las recomendaciones para disminuir el error tipo II o falso negativo es incrementar el tamaño de la muestra, lo cual se verifico al plantear el modelo tanto para la base de datos 27 como para la 28.

| BASE DE DATOS | Tipo de Distribución | Tamaño de la muestra [n] | Parámetros Máxima Verosimilitud | | Nivel de Significancia [%] | PRUEBAS DE BONDAD DE AJUSTE | | | |
|---------------|----------------------|--------------------------|---------------------------------|--------------|----------------------------|-----------------------------|---------|---------------------|----------|
| | | | α | β | | <i>Kolmogorov-Smirnov</i> | | <i>Chi-Cuadrado</i> | |
| | | | | | | Acepta | Rechaza | Acepta | Rechaza |
| B.26 | Pareto | 60 | 0,758 | 2,654 | 0,01 | | x | | x |
| B.27 | Wakeby | 34 | 3,904 | 2,549 | 0,01 | X | | | x |
| B.28 | Wakeby | 98 | 4,895 | 1,043 | 0,01 | | x | | x |
| B.29 | Logarithmic | 150 | 1,124 | 4,003 | 0,01 | | x | | x |
| B.30 | Gamma | 180 | 3,424 | 25,033 | 0,01 | | x | | x |

Tabla 7. Resultados para distribuciones que no pueden ser modeladas por la distribución Weibull

5. CONCLUSIONES

El análisis de confiabilidad destinado a equipos es notablemente una ayuda, siendo esta fundamental al establecer un modelo que determine la “operatividad de dicho equipo”. Establecer el modelo descrito por los diferentes datos obtenidos de las fallas de los equipos en cuestión, requiere de ajustes y tratamientos diferentes para así realizar un modelo más exacto que describa de una mejor manera la naturaleza de los datos. Cabe resaltar, que el entorno programado diseñado permite crear diferentes modelos para diferentes distribuciones posibles, ya que la distribución Weibull puede ser tomada como una “generalización”, la cual comprende diferentes casos y alternativas diferentes. Sin embargo, hay ocasiones en las cuales cierta cantidad de datos no pueden ser modelados bajo una distribución Weibull y es por esto que se tiene el respaldo de las pruebas de bondad de ajuste, cuya funcionalidad es determinar si el modelo realizado es aceptado o rechazado.

Las curvas de densidad de probabilidad y la acumulativa describen el comportamiento del equipo a través del tiempo, observándose así el comportamiento confiable de un equipo y su operatividad.

Efectuando un estudio basado en confiabilidad y supeditado a estos los tiempos de falla, es indispensable establecer los tipos de datos que se poseen para el análisis de la distribución que mejor los describe, es por esto necesario realizar un tratamiento a la información recolectada, para brindar con esto una organización que define la correcta escogencia de la distribución con el fin de tener una aproximación al modelo más cercano y fiel.

Se aproximó la distribución obtenida en comienzo para la base 1 del análisis de resultados finales, a otra como la exponencial, donde se puede constatar que el entorno se ejecuta con los criterios previamente establecidos.

El entorno programado fue diseñado principalmente para el análisis de confiabilidad, es por esto que fue necesario realizar una cantidad de pruebas aleatorias con el fin de verificar la funcionalidad y consistencia del código de programación.

El entorno realiza los cálculos de los parámetros por dos métodos: Máxima verosimilitud y Regresión lineal. Debido a una cierta cantidad de pruebas realizadas, se noto que el método de Máxima verosimilitud es más exacto que la regresión lineal. Por esto, es mediante los parámetros obtenidos por máxima verosimilitud que se realizaron los cálculos y análisis necesarios para obtener el modelo.

La base de datos 27 contiene 34 datos mientras que la base de datos 28 tiene 98 datos. A mayor cantidad de datos más precisión se pudo lograr en las pruebas de bondad de ajuste, tanto en la aceptación de la hipótesis nula como en el acercamiento entre los valores que toman los parámetros, reales y calculados.

6. BIBLIOGRAFÍA

- [1]. TIPPACHON W. BOONRUANG R. BOONPUN S. KHAMTANG C. KLAIUANG N. RERKPREEDAPONG D. HOKIERTI J. Failure Analysis of Protective Devices in Power Distribution Systems for Reliability Purpose. En: IEEE. Department of Electrical Engineering, Kasetsart University. Bangkok, 10900 Thailand.
- [2]. HAJEK Jan. DAHLUND Mats. PETTERSSON Lars. y BENNSTAM Gunnar. La calidad del aceite marca la diferencia ¿Ha descubierto ABB la solución para los fallos de transformadores?. Revista ABB, No 3, 2004. ABB Power Technologies AB Power Transformers Ludvika, Suecia. Págs. 61 - 63
- [3]. FERNÁNDEZ STEINKO Armando. Teoría y práctica del mantenimiento en España. Cuadernos de Estudios Empresariales, No 8, 1998. Universidad Complutense de Madrid. ISSN: 1131-6985. Págs., 125 – 147
- [4]. BABAN Calin Florin RADU Ioan Eugen BABAN Marius Integrated System for Reliability- Modeling Of Cold Plastic Deformation Tools Used in Car Industry. 2002. University of Oradea. En: IEEE. Proceedings, annual reliability and maintainability, Symposium. Págs: 212 – 216
- [5]. PARK Won J. KIM Yoon G. Goodness-of-Fit Tests for the Power-Law Process. En: IEEE transactions on reliability, vol. 41, no. 1, 1992 march.
- [6]. TASAMA Julián y GÓMEZ Santiago. Análisis de confiabilidad aplicado a una conformadora de rollos empleando la técnica de modos de fallo. Scientia et Technica Año XV, No 41, Mayo de 2009. Universidad Tecnológica de Pereira. ISSN 0122-1701
- [7]. MONTGOMERY Douglas C. y RUNGER George C. Applied Statistics and Probability for Engineers . Third Edition. John Wiley & Sons, Inc. ISBN 0-471-20454-4

- [8]. LOPEZ RENDON Juan Fernando, SOTO MEJIA Jose Adalberto, Laboratorios de Simulación Discreta. Postergraph S.A. 2010, ISBN: 978-958-44-6516-0 v. 0 pág. 167
- [9]. EPSTEIN Benjamin y WEISSMAN Ishay. Mathematical Models for Systems Reliability. CRC Press, 2008. ISBN-13: 978-1-4200-8082-7
- [10]. GRINSTEAD Charles M. y SNELL J. Laurie. Introduction to Probability . Second edition. American Mathematical Society.
- [11]. YANG Guangbin. Life Cycle Reliability Engineering. JOHN WILEY & SONS, INC. 2007
- [12]. BIOESTADÍSTICA. Bioestadística: Métodos y Aplicaciones U.D. Facultad de Medicina. Edición 3. Universidad de Málaga. ISBN: 847496-653-1
- [13]. DICKINSON GIBBONS Jean y CHAKRABORTI Subhabrata. Nonparametric Statistical Inference. Fourth Edition. The University of Alabama. Marcel Dekker, INC. ISBN: 0-8247-4052-1
- [14]. LEVIN Richard I. y RUBIN David S. Estadística para administración y economía. Séptima edición. Pearson Prentice Hall. 2004. ISBN: 970-26-0497-4
- [15]. WOODRUFF W. Brian, MOORE H. Albert, DUNNE J. Edward, CORTES Ramon. A Modified Kolmogorov – Smirnov Test for Weibull Distributions with Unknown Location and Scale Parameters. En: IEEE transactions on reliability, vol.r-32, no.2, june 1983
- [16]. MALLOR F, SANTOS J “fiabilidad de sistemas”, [en línea] Disponible: <http://www.unavarra.es/estadistica/LADE/M.O.C./Tema4.PDF>
- [17]. NACHLAS J.A. “Fiabilidad”. Isdefe. 1995, Primera edición
- [18]. IEEE “Power System Reliability Evaluation”, tutorial course 82 EHO 195-8-PWR, IEEE, 1982
- [19]. ZAPATA C. Julio “Confiabilidad de Sistemas Eléctricos”, Pereira 2005 Universidad Tecnológica de Pereira. Escuela de Tecnología Eléctrica
- [20]. R. Billinton, S. Kumar, “A Reliability Test System for Educational Purposes – Basic Results”, IEEE Transactions on Power Systems, Vol. 5 No. 1, February 1990.

- [21]. MEYER PAUL L., "Probabilidad y Aplicaciones Estadísticas
- [22]. TORRES M. Alvaro, "Probabilidad, Variables Aleatorias, Confiabilidad y Procesos Estocásticos en Ingeniería Eléctrica", Universidad de Los Andes, Volumen I, 1993
- [23]. SALVADOR F, M y GARGALLO, P. (2003): "Análisis Exploratorio de Datos", [en línea]
Disponible: <http://www.5campus.com/leccion/aed>
- [24]. PEYTON Z P, "Probability, Random Variables, and Random Principles". Mcgraw-Hill 1987 Second Edition
- [25]. GROEN, F.J. y DROGUETT, E.L. Prior specification for multi-failure mode weibull reliability models. En: IEEE Reliability and Maintainability Symposium, 2006. RAMS '06. Annual, Págs, 201 - 207, enero 2006
- [26]. NAKAGAWA, Toshio; YASUI, Kazumi. Approximate Calculation of Inspection Policy with Weibull Failure Times. En: IEEE transactions on reliability, vol.r-28, agosto 2009
- [27]. ROSS, R. Comparing linear regression and maximum likelihood methods to estimate Weibull distributions on limited data sets: systematic and random errors. En: IEEE electrical insulation and die, vol.1, agosto 2002
- [28]. FABIANI, D. SIMONI, L. Discussion on application of the Weibull distribution to electrical breakdown of insulating materials. En: IEEE Dielectrics and Electrical Insulation, vol.12, febrero 2005
- [29]. HAIBO Li, ZHENGPING Zhang, YANPING Hu, DEQIANG Zheng. Maximum likelihood estimation of weibull distribution based on random censored data and its application. En: IEEE Reliability, Págs. 302 - 304, septiembre 2009
- [30]. SOO-JONG Lee, KYEONG-HO Lee. A weighted distribution of residual failure data using Weibull model. En: IEEE information networking, Págs. 265 - 270, agosto 2002
- [31]. <http://www.mathwave.com/es/home.html>
- [32]. [http://www.fepslutc.org/FEPSEUN/en/departments/statistics/courses/Dr.%20Hanan%20Aly/Lec_8\(MLE\).pdf](http://www.fepslutc.org/FEPSEUN/en/departments/statistics/courses/Dr.%20Hanan%20Aly/Lec_8(MLE).pdf)

ANEXOS

Anexo 1.

| <i>n</i> | $\alpha = 0,2$ | $\alpha = 0,1$ | $\alpha = 0,05$ | $\alpha = 0,01$ |
|----------|-----------------|-----------------|-----------------|-----------------|
| 1 | 0,90 | 0,95 | 0,98 | 0,99 |
| 2 | 0,68 | 0,78 | 0,84 | 0,93 |
| 3 | 0,56 | 0,64 | 0,71 | 0,83 |
| 4 | 0,49 | 0,56 | 0,62 | 0,73 |
| 5 | 0,45 | 0,51 | 0,56 | 0,67 |
| 6 | 0,41 | 0,47 | 0,52 | 0,62 |
| 7 | 0,38 | 0,44 | 0,49 | 0,58 |
| 8 | 0,36 | 0,41 | 0,46 | 0,54 |
| 9 | 0,34 | 0,39 | 0,43 | 0,51 |
| 10 | 0,32 | 0,37 | 0,41 | 0,49 |
| 11 | 0,31 | 0,35 | 0,39 | 0,47 |
| 12 | 0,30 | 0,34 | 0,38 | 0,45 |
| 13 | 0,28 | 0,32 | 0,36 | 0,43 |
| 14 | 0,27 | 0,31 | 0,35 | 0,42 |
| 15 | 0,27 | 0,30 | 0,34 | 0,40 |
| 16 | 0,26 | 0,30 | 0,33 | 0,39 |
| 17 | 0,25 | 0,29 | 0,32 | 0,38 |
| 18 | 0,24 | 0,28 | 0,31 | 0,37 |
| 19 | 0,24 | 0,27 | 0,30 | 0,36 |
| 20 | 0,23 | 0,26 | 0,29 | 0,35 |
| 25 | 0,21 | 0,24 | 0,26 | 0,32 |
| 30 | 0,19 | 0,22 | 0,24 | 0,29 |
| 35 | 0,18 | 0,21 | 0,23 | 0,27 |
| 40 | 0,17 | 0,19 | 0,21 | 0,25 |
| 45 | 0,16 | 0,18 | 0,20 | 0,24 |
| > 45 | $1.07/\sqrt{n}$ | $1.22/\sqrt{n}$ | $1.36/\sqrt{n}$ | $1.63/\sqrt{n}$ |

Anexo 1. Distribución del estadístico de Kolmogorov-Smirnov

Anexo 2.

| <i>n</i> | $\alpha = 0,2$ | $\alpha = 0,1$ | $\alpha = 0,05$ | $\alpha = 0,01$ |
|----------|----------------|----------------|-----------------|-----------------|
| 1 | 1,6424 | 2,7055 | 3,8415 | 6,6349 |
| 2 | 3,2189 | 4,6052 | 5,9915 | 9,2104 |
| 3 | 4,6416 | 6,2514 | 7,8147 | 11,3449 |
| 4 | 5,9886 | 7,6794 | 9,4877 | 13,2767 |
| 5 | 7,2893 | 9,2363 | 11,0705 | 15,0863 |
| 6 | 8,5581 | 10,6446 | 12,5916 | 16,8119 |
| 7 | 9,8032 | 12,017 | 14,0671 | 18,4753 |
| 8 | 11,0301 | 13,3616 | 15,5073 | 20,0902 |
| 9 | 12,2421 | 14,6837 | 16,919 | 21,666 |
| 10 | 13,442 | 15,9852 | 18,307 | 23,2093 |
| 11 | 14,6314 | 17,275 | 19,6752 | 24,725 |
| 12 | 15,812 | 18,5493 | 21,0261 | 26,217 |
| 13 | 16,9848 | 19,8119 | 22,362 | 27,6882 |
| 14 | 18,1508 | 21,0641 | 23,6848 | 29,1412 |
| 15 | 19,3107 | 22,3071 | 24,9958 | 30,578 |
| 16 | 20,4651 | 23,5418 | 26,2962 | 31,9999 |
| 17 | 21,6146 | 24,769 | 27,5871 | 33,4087 |
| 18 | 22,7595 | 25,9894 | 28,8593 | 34,8052 |
| 19 | 23,9004 | 27,2036 | 30,1435 | 36,1908 |
| 20 | 25,0375 | 28,412 | 31,3404 | 37,5663 |
| 21 | 26,1711 | 29,6151 | 32,6706 | 38,9322 |
| 22 | 27,3015 | 30,8133 | 33,9245 | 40,2894 |
| 23 | 28,4288 | 32,0069 | 35,1725 | 41,6383 |
| 24 | 29,5533 | 33,1962 | 36,415 | 42,9798 |
| 25 | 30,6752 | 34,3816 | 37,6525 | 44,314 |
| 26 | 31,7146 | 35,5632 | 38,8851 | 45,6416 |
| 27 | 32,9117 | 36,7412 | 40,1133 | 46,9628 |
| 28 | 34,0266 | 37,9159 | 40,3372 | 48,2782 |
| 29 | 35,1394 | 39,0875 | 42,5569 | 49,5878 |

Anexo 2. Tabla de distribución Chi-Cuadrado

Anexo 3. Código fuente realizado y ejecutado en Matlab para el entorno programado

```
% Organización de los datos
% S1 contiene todos los datos, S2 elimina datos repetidos y S3
agrupa en
% intervalos

SU=1;
S1=1;
S2=2;
S3=3;
NS=4

if SU == S1,
    NUMERIC = xlsread('Base1.xlsx');
end;

if selecciondatos == 2,
    Numeric1 = xlsread('Base1.xlsx');
    Numeric2 = sort(Numeric1);
    ii = size (Numeric1,1);
    iii=1;
    n=1;
    NUMERIC=zeros(n,2);
    NUMERIC (n,1) = n;
    if iii==1,
        NUMERIC (n,2) = Numeric2 (iii,2);
        iii=iii+1;
        n=n+1;
        NUMERIC (n,1) = n;
    end;
    while iii<=ii,
        if Numeric2 (iii,2) == Numeric2 (iii-1,2);

            else
                NUMERIC (n,2) = Numeric2 (iii,2);
                NUMERIC (n,1) = n;
                n=n+1;
            end;
        iii=iii+1;
    end
end;

if selecciondatos == 3,
    Numeric3 = xlsread('Base1.xlsx');
    Numeric4 = sort (Numeric3)
    iv = size (Numeric4,1);
```

```

Nintervalo= round(sqrt(iv));
DatoP= Numeric4(1,2)
DatoF= Numeric4(iv,2)
Tamano= DatoF - DatoP
Intervalo = Tamano / Nintervalo
FrecAbs = (hist(Numeric4,Nintervalo))';
nn=1;
v=1;
T = zeros (Nintervalo,1);
while v <= Nintervalo,
    T(v,1)= nn;
    T(v,2)= (DatoP + (DatoP +Intervalo))/2;
    DatoP = DatoP + Intervalo;
    v=v+1;
    nn= nn+1;
end
NUMERIC = T
end;

%Rango de la mediana
NUMERIC
i = size(NUMERIC,1);
x = (NUMERIC(1,1));
v = 0.0611; %Parametro de ajuste
Ks = zeros(i,1);
xx=1;

while xx <= i,
    y = (NUMERIC(xx,2));
    Ks(xx,1)= y;
    KS = sort((Ks(:))); %Ordena de menor a mayor
    xx = xx + 1;
end

% Linealización y regresión para datos alfa y beta

G=zeros(i,1);
G1=zeros(i,1);
G2=zeros(i,1);

while x <= i,
    Rx = ((x - 0.3)/(i + 0.4));
    G(x,1)=Rx;
    M = log (abs(KS(x,1)-v));
    G1(x,1)=M;
    yi = log(log(1/(1-Rx)));
    G2(x,1)=yi;
    x = x+1;

```

```

end

% plot (G1,G2, '.');
c = polyfit(G1,G2,1);
c1 = G1(1):0.1:G1(length(G1));
c2 = polyval(c,c1);

%plot(c1,c2);
%hold on
%plot(G1,G2, 'x');
xlabel('x')
ylabel('y')

%Parámetros por Máxima Verosimilitud

parmhat = wblfit(Ks);
AlfaMV = parmhat (1,1);
BetaMV = parmhat (1,2);

[M,V] = wblstat(AlfaMV,BetaMV);

AlfaMV % Parámetro con MV
BetaMV % Parámetro con MV
M; % Media
V; %Varianza

% Parámetros por linealización y regresión

Beta = c(1,1)
Intercepto = c(1,2);
Alfa = exp(-(Intercepto/Beta))
Aprox = 0;

% Kolmogorov - Smirnov

while Aprox == 0 || Aprox == y,

%KS = sort((Ks(:))); %Ordena de menor a mayor
n = 1;
FTW=zeros(i,1);
FCUM=zeros(i,1);
FTWEXP=zeros(i,1);
FCUMEXP=zeros(i,1);
PROEMMAX=zeros(i,1);

```

```

PROEMMIN=zeros(i,1);
KS;

while n<=i,
    kss = (NUMERIC(n,1));
    t = (KS(n,1));
    Ftw = ((BetaMV*(t^(BetaMV-1)))/(AlfaMV^BetaMV))*(exp(-
(t/AlfaMV)^(BetaMV))); % Fdp
    FTW(n,1)=Ftw;
    Fcum = (1 - (exp(-(t/AlfaMV)^(BetaMV)))); % Fcum
    FCUM(n,1) = Fcum;
    Proemmin = ((kss-1)/(i));
    Proemmax = (kss/i);
    PROEMMAX(n,1) = Proemmax;
    PROEMMIN(n,1) = Proemmin;
    n = n+1;

end

if Aprox == y && 0.9<=BetaMV<=1.1,
    while n <= i,
        ksss = (NUMERIC(n,1));
        texp = (KS(n,1));
        Ftwexp = (AlfaMV*(exp(-(texp*AlfaMV))));
        FTWEXP(n,1)=Ftwexp;
        Fcumexp = (1 - (exp(-(texp*AlfaMV))));
        FCUMEXP(n,1) = Fcumexp;
        Proemmin = ((ksss-1)/(i));
        Proemmax = (ksss/i);
        PROEMMAX(n,1) = Proemmax;
        PROEMMIN(n,1) = Proemmin;
        n = n+1;
    end
end;

subplot(2,1,1), plot (KS,FTW,'x') %Fdp
title('Fdp')
subplot(2,1,2), plot (KS,FCUM,'x') %Fcumulativa
title('Fcum')

D1 = abs(FCUM - PROEMMAX);
D2 = abs(FCUM - PROEMMIN);
Dmayor1 = max(max(D1));
Dmayor2 = max(max(D2));

if Dmayor1 >= Dmayor2,
    Dmayormax = Dmayor1
else

```

```

    Dmayormax = Dmayor2
end;

KOLSM = xlsread('Limites de aceptacion para Kolmogorov-
Smirnov.xlsx');

if i<=20,

    Facep = KOLSM(i,1);

    if Facep == i,
        Facepta = KOLSM(i,NS);
    end;

end;

if i >= 21 && i <=45,
    if i>= 21 && i <=24,
        Facepta = KOLSM(20,NS);
    else
        if i == 25,
            Facepta = KOLSM(21,NS);
        else
            if 26<=i<=29,
                Facepta = KOLSM(21,NS);
            else
                if i == 30,
                    Facepta = KOLSM(22,NS);
                else
                    if 31<=i<=34,
                        Facepta = KOLSM(22,NS);
                    else
                        if i == 35,
                            Facepta = KOLSM(23,NS);
                        else
                            if 36<=i<=39,
                                Facepta = KOLSM(23,NS);
                            else
                                if i == 40,
                                    Facepta = KOLSM(24,NS);
                                else
                                    if 41<=i<=44,
                                        Facepta = KOLSM(24,NS);
                                    else
                                        if i == 45,
                                            Facepta = KOLSM(25,NS);
                                        end;
                                    end;
                                end;
                            end;
                        end;
                    end;
                end;
            end;
        end;
    end;
end;

```

```

end;
end;
end;
end;
end;
end;
end;
end;

if i>45 && NS==2,
    Facepta=(1.07/sqrt(i));
end;

if i>45 && NS==3,
    Facepta=(1.22/sqrt(i));
end;

if i>45 && NS==4,
    Facepta=(1.36/sqrt(i));
end;

if i>45 && NS==5,
    Facepta=(1.63/sqrt(i));
end;

Facepta

if Facepta>=Dmayormax,
    'La Hipótesis se acepta'
else
    'La Hipótesis se rechaza'
end;

% Ji-Cuadrado

CHICUA = xlsread('Lmites de aceptacion para Ji-cuadrado.xlsx');

int = round(sqrt(i)); % Define Usuario
Primero = KS(1,1);
Ultimo = KS(i,1);
longitud = Ultimo - Primero;
Inter = longitud/int;
s=1;
ti = Primero;
ji = zeros(int+1,1);

```

```

Ei = zeros (int,1);
x2 = zeros (int,1);

if int <= 29,

    while s<=int+1;
        Fcumu = (1 - (exp(-(ti/AlfaMV)^(BetaMV))));
        ji(s,1) = Fcumu;
        ti = ti + Inter;
        s=s+1;
    end

ss=1;
    while ss<=int;
        res= (ji(ss+1,1) - ji(ss,1))*i;
        Ei(ss,1) = res ;
        ss = ss + 1;
    end
ni = hist(Ks,int);
Oi=ni';
ji;
Ei;

sss=1;
    while sss<=int;
        Xcua = ((Oi(sss,1) - Ei(sss,1))^2)/(Ei(sss,1));
        x2(sss,1) = Xcua;
        sss = sss + 1;
    end
x2;
Chicadrado = sum (x2)

Limit = size(CHICUA,1);
JJ=1;
FE = zeros (Limit,1);
while JJ <= Limit,
    fe = (CHICUA(JJ,NS));
    FE(JJ,1)= fe;
    JJ = JJ+1;
end

Faceptar = CHICUA(int-1,1);

    if Faceptar == int-1,
        Facepto = CHICUA(int-1,NS)
    end;

if Facepto>=Chicadrado,
    'La Hipótesis se acepta'

```

```

else
    'La Hipótesis se rechaza'
end;

else
    'La prueba Ji-Cuadrado no se puede realizar debido a que el
número de intervalos supera los permitidos'
end;

if BetaMV<=1.1 && BetaMV >=0.9,

    if 0.95<=BetaMV<=1.05,
        Aprox = input('Desea aproximar el Parámetro a una distribución
exponencial: ');
        if Aprox == y;
            BetaMV = 1;

            else
                if Aprox == n,
                    BetaMV = parmhat (1,2);
                end;
            end;
        end;
    end;

if BetaMV<=2.1 && BetaMV >=1.9,

    if 1.9<=BetaMV<=2.1,
        Aprox = input('Desea aproximar el Parámetro a una distribución
Rayleigh: ');
        if Aprox == y,
            BetaMV = 2;

            else
                if Aprox == n,
                    BetaMV = parmhat (1,2);
                end;
            end;
        end;
    end;

    if Aprox == n,
        break
    end;
    if Aprox == 0,
        break
    end;
end;
end;

```

Anexo 3. Código fuente entorno programado