

UNIVERSIDAD CENTRAL DE LAS VILLAS
FACULTAD DE MATEMATICA, FISICA Y COMPUTACION
CARRERA DE CIENCIA DE LA COMPUTACION



TRABAJO DE DIPLOMA

**SISTEMA ESTADÍSTICO PARA EL ESTUDIO DE LA
FIABILIDAD**

AUTOR:

Danilo Valdés Ramírez

TUTOR:

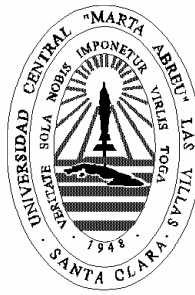
Tirso Reyes Carvajal

CONSULTANTES:

Gladys Casas Cardoso

María del Carmen Chávez Cárdenas

Santa Clara, 3 de julio de 2006
"Año de la Revolución Energética en Cuba"



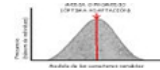
Hago constar que el presente trabajo fue realizado en la Universidad Central Marta Abreu de Las Villas como parte de la culminación de los estudios de la especialidad de Ciencias de la Computación, autorizando a que el mismo sea utilizado por la institución, para los fines que estime conveniente, tanto de forma parcial como total y que además no podrá ser presentado en eventos ni publicado sin la autorización de la Universidad.

Firma del autor

Los abajo firmantes, certificamos que el presente trabajo ha sido realizado según acuerdos de la dirección de nuestro centro y el mismo cumple con los requisitos que debe tener un trabajo de esta envergadura referido a la temática señalada.

Firma del tutor

Firma del jefe del Seminario



A mis padres

El presente trabajo es el fruto del esfuerzo de un grupo de personas a las que deseo agradecer eternamente el apoyo que me han brindado. A mis padres que me han convertido en lo que soy. A mi familia, especialmente a mis hermanas. A mi novia que siempre ha estado a mi lado. A mis compañeros y amigos de cuarto, de aula, de año. A mis profesores, especialmente los doctores Tirso Reyes y Gladys Casas. A mis amigos de la FEU. A la Universidad Central de Las Villas por lo que ella ha significado en mi preparación. Y muy especialmente a todo el que ha luchado por construir esta Revolución.

Resumen

En este trabajo se describe un sistema estadístico capaz de realizar estudios sobre una variable continua, específicamente sobre la variable tiempo. Con el nombre de Sistema Estadístico para el estudio de la Fiabilidad, ofrece los gráficos de las curvas de fiabilidad con los intervalos de confianza y la posibilidad de evaluar sobre ella cualquier intervalo de tiempo después de realizado el estudio. Ofrece también las curvas teóricas y prácticas de probabilidad y un resumen de los datos estadísticos de la muestra. Permite obtener los parámetros estimados y en caso de ajuste a la distribución los valores ideales para la nueva planificación de mantenimiento o tiempo de garantía, según sea el problema. Es capaz también de generar nuevos valores que completen la muestra de ser necesario. Todo esto a partir de una muestra de intervalos de tiempos hasta o entre fallos. En él se han trabajado dos pruebas de bondad del ajuste y tres distribuciones con tres niveles de significación, para esta primera versión.

Contiene además un resumen sobre la historia de las estadísticas y la teoría fiabilidad como ciencias, desde su aparición hasta su desarrollo y sus creadores, es el contenido del primer capítulo. Junto a esto aparecen relacionadas las herramientas computacionales que existen hoy en el mundo que permiten realizar estudios similares con sus deficiencias y ventajas y las condiciones que nos llevaron a realizar este trabajo.

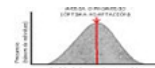
Se presentan los diagramas que describen la estructura interna del sistema, los algoritmos matemáticos y estadísticos utilizados y las instrucciones a manera de manual para su correcto empleo.

Abstract

The present work describes a statistical system able to make studies on a continuous variable, specifically on the variable time. With the name of Statistical System for the study of the Reliability, given a sample of time intervals until or between failures, offers the graphs of the curves of reliability with the confidence intervals and the possibility of evaluating on it, any time interval after made the study, the theoretical and practical curves of probability; a summary of the statistical parameters of the sample and, in case of adjustment to the distribution, the ideal values for the new planning of maintenance or time of guarantee according to is the problem.

It is able also to generate new values that complete the sample to be necessary. Two tests of goodness of fit and three distributions with three levels of meaning for this have worked first version. A summary on the history of the statistics and the theory reliability like sciences, from its appearance and development, to its creators, is the content of the first chapter.

Next to this, appear a summary of the computer operations tools that exist today in the world that allow to make similar studies with their deficiencies and advantages and the conditions that took to us to make this work. The diagrams also appear that describe the internal structure of the system, the mathematical and statistical algorithms used as well as the instructions to way of manual for their correct use.



Índice

Resumen.....	5
Abstract.....	6
Índice	7
Introducción.....	12
1. Capítulo I: VISIÓN GENERAL ACERCA DE LA UTILIZACIÓN DE LA ESTADÍSTICA EN APLICACIONES COMPUTACIONALES. ESTADO DEL ARTE.....	15
1.1 Generalidades.....	15
1.1.1 Concepto de fiabilidad.....	15
1.1.2 Concepto de estadística y estudio estadístico.....	16
1.1.3 Concepto de variable aleatoria continua.....	17
1.1.4 Concepto de muestra y población.....	18
1.1.5 Breve reseña de la teoría de la fiabilidad.....	19
1.2 Aplicaciones generales de la estadística en la ingeniería y la medicina.....	21
1.2.1 Utilización en la teoría de la fiabilidad.....	21
1.2.2 Utilización en la medicina.....	22
1.2.3 Características que potencian su utilización.....	24
1.3 Herramientas computacionales existentes.....	25
1.3.1 Usos y fines.....	26
1.3.2 Deficiencias.....	28
1.4 Soluciones que da el Sistema Estadístico para el estudio de la Fiabilidad a las deficiencias existentes.....	30
1.5 Conclusiones parciales.....	32
2. Capítulo II: DIAGRAMAS Y ESTRUCTURA DEL SISTEMA ESTADÍSTICO PARA EL ESTUDIO DE LA FIABILIDAD.....	35
2.1 Diagrama del modelo del proceso del negocio.....	35

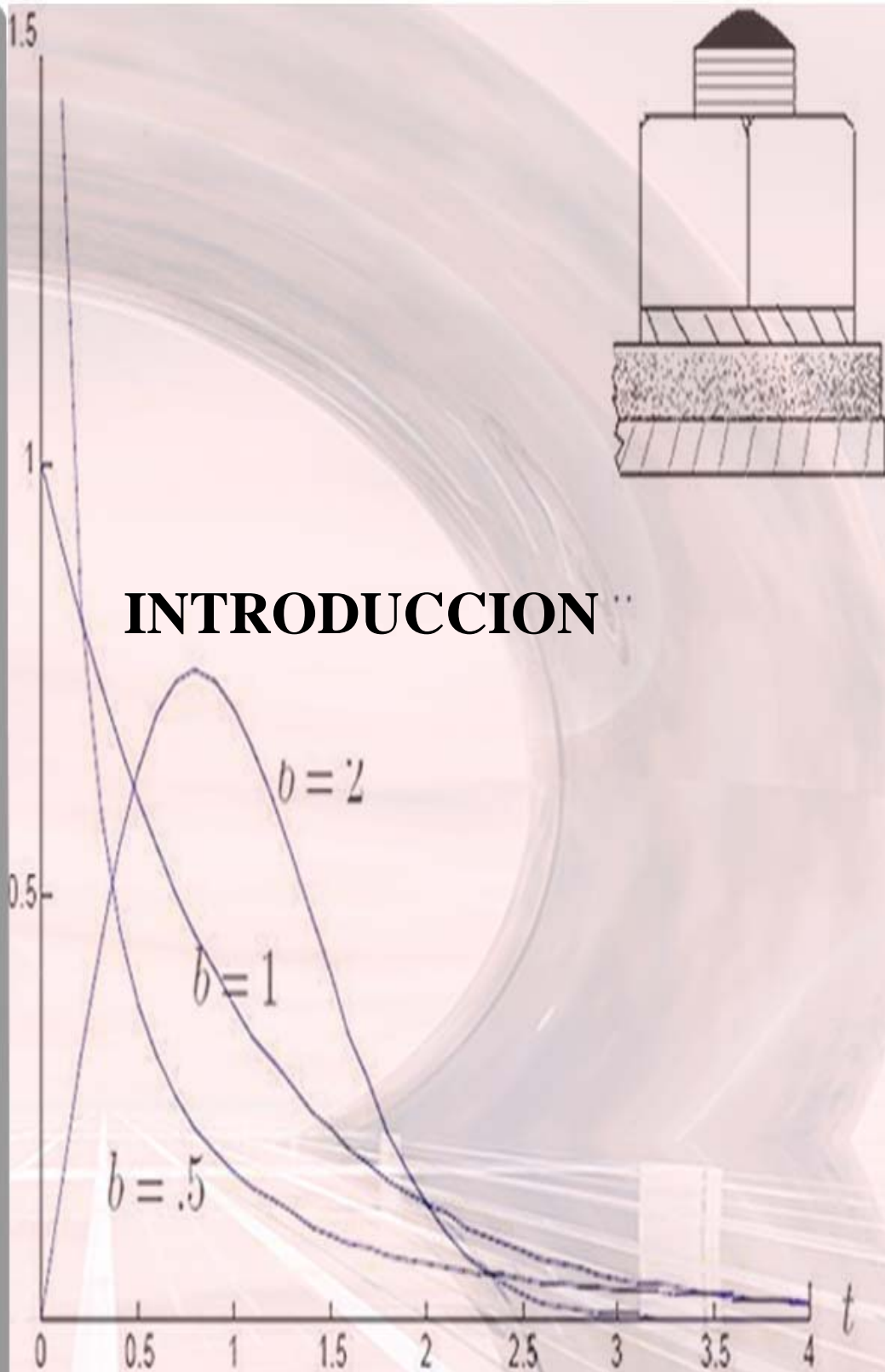
2.1.1 Actividad fundamental.....	36
2.1.2 Diagrama de estados general.	37
2.1.3 Diagramas de secuencias.	37
2.2 Casos de uso.	38
2.2.1 Casos de uso del negocio.	38
2.2.2 Casos de uso del sistema.....	39
2.2.2.1 Especificidad de actividades.....	40
2.3 Diagrama de flujo de datos.	41
2.4 Diagrama de la estructura informativa (entidades).....	43
2.5 Diagrama de clases.	45
2.6 Diseño de ventanas.	46
2.6.1 Tabla de Eventos.....	46
2.7 Conclusiones parciales.....	48
3. Capítulo III: FORMULACIÓN MATEMÁTICA, ESTADÍSTICA Y FIABILÍSTICA.....	50
3.1 Test de bondad del ajuste.....	50
3.1.1 Distribuciones utilizadas. Funciones de probabilidad y probabilidad acumulada.	50
3.1.2 Test Chi cuadrado.	52
3.1.3 Test de Kolmogorov Smirnov.....	53
3.1.3.1 Estadígrafo utilizado.	53
3.2 Estimación de parámetros.....	54
3.2.1 Estimación de la media.	54
3.2.2 Estimación de la varianza y la desviación estándar.....	55
3.2.3 Estimación del parámetro lambda.....	56
3.2.4 Estimación de los parámetros alfa y beta.....	56
3.3 Cálculo de los intervalos de confianza.	57
3.3.1 Intervalos de confianza en la distribución normal.	57
3.3.2 Intervalos de confianza en la distribución exponencial.	58
3.3.2.1 Generalización para muestras con o sin reemplazo.	58



3.3.3 Intervalos de confianza en la distribución de weibull.....	59
3.4 Generación de tiempos para la muestra.	60
3.4.1 Generación asistida por el usuario.	61
3.4.2 Generación empírica.	62
3.5 Conclusiones parciales.....	62
4. Capítulo IV: FUNCIONALIDADES DEL SISTEMA ESTADÍSTICO PARA EL ESTUDIO DE LA FIABILIDAD.....	64
4.1 Abrir un problema.....	64
4.1.1 Desde un fichero.	64
4.1.2 Desde la base de datos.	65
4.1.3 Crear un problema nuevo.....	65
4.2 Entrar datos de la muestra.....	66
4.2.1 Clasificar el problema.	66
4.2.2 Definir una variable.	67
4.3 Ejecutar estudio estadístico.....	67
4.4 Interpretación gráfica de los resultados.	68
4.4.1 Interpretación de la gráfica Q (t).....	69
4.4.2 Interpretación de la gráfica P (t).	70
4.4.3 Interpretación de las gráficas f (t).	71
4.5 Interpretación de los resultados estadísticos.....	71
4.5.1 Significados de los parámetros.	72
4.5.2 Usos de los intervalos de confianza.	72
4.6 Guardar un problema.	73
4.6.1 Hacia un fichero.....	74
4.6.2 Hacia la base de datos.....	74
4.7 Facilidades para el almacenamiento de información.....	75
4.7.1 Portabilidad de los datos.	75
4.7.2 Seguridad y consistencia de las muestras.	76
4.8 Generar nuevos valores.....	76
4.9 Análisis de la velocidad de procesamiento.....	77



4.9.1 Comparación con otras herramientas.....	78
4.10 Requerimientos.....	78
4.11 Conclusiones parciales.....	78
Conclusiones.....	81
Recomendaciones	82
Bibliografía	83
Anexos.....	86



INTRODUCCION

Sistema Estadístico para el estudio de la Fiabilidad

Introducción

La teoría de la fiabilidad nace prácticamente junto al desarrollo que tuvieron la ciencia y la técnica a partir de la Revolución Industrial. La aplicación de métodos estadísticos para la determinación y predicción del comportamiento de los equipos mecánicos que empezaban a aparecer fue el comienzo de esta. Sin embargo durante las dos Guerras Mundiales fue cuando realmente se desarrolló y alcanzó un gran auge. La durabilidad, potencia y garantía de los complejos equipos militares que se desarrollaban debía ser analizada con métodos estadísticos que confirmaran si en verdad “un tanque era mejor que otro” o “si un fusil duraba más tiempo que otro”.

En la actualidad, a pesar de su poca divulgación, esta teoría es aplicada por todas las empresas de diversas formas y sus métodos se han difundido y enriquecido con el desarrollo de las computadoras. Un estudio que puede tomar varias horas, incluso días, en realizarse por la complejidad de los métodos y algoritmos, ahora puede ser realizado en pocos minutos y en ocasiones segundos.

Ante la necesidad de herramientas para realizar estudios de la teoría de la fiabilidad en nuestro país, surge la idea de este proyecto, un sistema creado en Cuba que permita analizar una muestra de tiempos de funcionamiento entre fallos o hasta el primer fallo en equipos, piezas, componentes o individuos. Que sea capaz también de generar una información visual para especialistas o conocedores de estadísticas y fiabilidad. Además, que asista a cualquier persona o industria en la modificación o planificación del ciclo de mantenimientos de sus equipos o componentes. Y que, a la vez, ayude a determinar la garantía de un producto a partir del comportamiento de una muestra de ellos.

El sistema que se presenta tiene como objetivo buscar el ahorro y la eficiencia en el aprovechamiento de los equipos, los componentes, las piezas, las construcciones, en fin, en todo lo que el hombre es capaz de crear, a través de una buena planificación del ciclo de sus mantenimientos; esto convierte al Sistema Estadístico para el estudio de la Fiabilidad en un fuerte aliado de cualquier empresa que desee preservar sus equipos y su producción. También se construye con el objetivo de ser utilizado en la

docencia en asignaturas del pregrado y postgrado de algunas ingenierías fundamentalmente la Ingeniería Mecánica, Industrial y Civil.

El presente trabajo persigue cuatro grandes objetivos que son:

- Crear un sistema computacional que permita realizar estudios de fiabilidad sobre equipos, componentes o individuos de una población o muestra.
- Presentar una herramienta que pueda ser utilizada en la docencia en algunas asignaturas del pregrado de varias ingenierías.
- Mostrar las acciones que se deben realizar para trabajar con este sistema.
- Presentar un primer intento de software estadístico que constituya el punto de arranque para una familia de estos.

En este trabajo se muestra, en su primer capítulo, un resumen de las herramientas que existen hoy en el mundo para llevar a cabo estudios estadísticos o de fiabilidad, sus ventajas y deficiencias, las condiciones que nos llevaron a producir este software y las soluciones que brinda a las deficiencias encontradas. Él mismo tiene el objetivo de introducir la situación actual de las herramientas computacionales existentes para estas actividades. En el segundo capítulo se muestran un grupo de diagramas del sistema que ayudan a su entendimiento y constituyen parte de la documentación del software, necesaria para realizar cualquier mantenimiento o para próximas versiones. El capítulo tercero tiene el objetivo de presentar la justificación matemática del sistema, en él aparece un compendio de los algoritmos matemáticos que son utilizados en el sistema para determinar la bondad del ajuste, estimar los parámetros o calcular los intervalos de confianza, las funciones de probabilidad y probabilidad acumulada de las distribuciones utilizadas así como las ecuaciones para la estimación de parámetros. El cuarto capítulo muestra las características y requerimientos del producto en forma de un pequeño manual para su fácil utilización. Por último se relacionan un grupo de imágenes y diagramas que complementan las explicaciones que se brindan. Entre ellas se encuentran diferentes vistas del software en ejecución y la forma de las funciones de probabilidad y probabilidad acumulada de las distribuciones que se trabajan en esta primera versión, normal, exponencial y de weibull.



1. Capítulo I: VISIÓN GENERAL ACERCA DE LA UTILIZACIÓN DE LA ESTADÍSTICA EN APLICACIONES COMPUTACIONALES. ESTADO DEL ARTE.

1.1 Generalidades

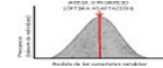
El presente trabajo nos obliga a definir un grupo de conceptos que serán utilizados indistintamente en cada uno de los capítulos y son imprescindibles para el entendimiento de los algoritmos que se utilizan. Además se hará un pequeño recorrido por la historia de la estadística que es la madre de la teoría de la fiabilidad. Se abordará sobre la existencia de las herramientas computacionales en el mundo, sus fines específicos, sus bondades y sus deficiencias. Por qué a pesar de la gran variedad de softwares que ya se han desarrollado, se ha decidido aventurarse en la producción del Sistema Estadístico para el estudio de la Fiabilidad (SEF). También se ha incluido un epígrafe que trata sobre cómo el trabajo intenta solucionar las necesidades de cálculo estadístico y aplicaciones de la teoría de la fiabilidad que aún existen en nuestro país. En este capítulo se hará referencia a varios sitios de Internet, libros y formularios que fueron consultados, las referencias se muestran en el pie de página y no coinciden, en el orden, con la bibliografía que se encuentra ordenada alfabéticamente.

1.1.1 Concepto de fiabilidad.

Según la definición de la CEE (1988), es "el cuerpo de conocimientos que se refieren a la predicción, análisis y reducción del error humano, enfocándose sobre el papel de la persona en las operaciones de diseño, de mantenimiento, uso y gestión de un sistema socio técnico".¹

Entre los objetivos que persigue el estudio de la teoría de la fiabilidad se encuentra la verificación de expresiones subjetivas como: "este producto es más duradero", "la compañía X produce los mejores componentes para computadoras", "los puentes de

¹ Leplat J y G Terssac. (1990). Les facteurs humains de la fiabilité Marseille (Francia), Ed. Octares.



hierro son más fuertes que los de concreto”, “los chinos duran más que los africanos”... Además justificar matemáticamente la toma de decisiones de cualquier empresa antes de producir cierto producto en serie. Determinar la supervivencia de individuos de una población sometidos a características particulares de un medio.

En su forma aplicativa a la ingeniería, la fiabilidad también es conocida como fiabilidad industrial. La definición en la ingeniería de esta teoría es la de “probabilidad de que un equipo realice sus funciones bajo determinadas condiciones en un período de tiempo determinado.” La teoría de la fiabilidad industrial realiza estudios en los productos que obligan a seguir los equipos tanto en su etapa de desarrollo como después de la recepción del producto por el cliente.

Una buena planificación de mantenimientos para un equipo, componente o construcción puede evitar un desastre, una rotura irreparable o el gasto de materiales de mantenimiento por una planificación fuera de tiempo. Precisamente este es uno de los problemas que se trata de resolver a través del estudio de la teoría de la fiabilidad. La fiabilidad en la ingeniería está orientada a los fallos. El problema reside en predecir si puede ocurrir un fallo al utilizar un dispositivo y predecir cuando ocurrirá. Esta información es útil para determinar los plazos de garantías de los productos. Todas las empresas dedicadas a la producción necesitan conocer la eficiencia de sus productos antes de decidir fabricarlo en serie y esta decisión puede ser fundamentada realizando un estudio estadístico de las muestras de tiempos de funcionamiento hasta el primer fallo que se recojan de un grupo de estos productos en fase de prueba.

1.1.2 Concepto de estadística y estudio estadístico.

La estadística está relacionada con los métodos científicos para la recolección, organización, tabulación, presentación y análisis de los datos así como para sacar conclusiones válidas y tomar decisiones razonables a partir de este análisis. De forma resumida se utiliza para denotar datos por sí mismos o los números derivados de estos.² Según el diccionario estadístico online es una Ciencia derivada de la Matemática que estudia los métodos científicos para recoger, organizar, resumir y

² Spiegel M. (Octubre 1961). *Teoría y problemas de Estadística. Estadística*. La Habana 1979.

analizar datos, así como para sacar conclusiones válidas y tomar decisiones razonables basadas en tal análisis.³

Muy relacionado a este concepto de estadística se encuentra el de fiabilidad, visto anteriormente. De aquí puede surgir la idea de realizar un estudio estadístico sobre la fiabilidad de un producto o de una población de individuos pero antes se necesita definir empíricamente qué es un estudio estadístico. Partiendo de las definiciones anteriores se puede decir que un estudio estadístico es “el análisis del comportamiento de una muestra tomada de una o varias poblaciones, a partir de su recogida y organización, para arribar a conclusiones específicas a esa muestra y que, en igualdad de condiciones, pueden ser extendidas a otras poblaciones”.

1.1.3 Concepto de variable aleatoria continua.

Una variable también es definida como un símbolo que puede tomar cualquiera de un conjunto de valores prescritos, denominado el dominio de la variable. Si la variable solo puede tomar un valor es denominada una constante. Una variable que puede tomar cualquier valor entre dos valores dados se denomina variable continua de lo contrario se denomina variable discreta.⁴

Se llaman variables aleatoria, y se representan por X , a una cantidad que se puede determinar cuantitativamente y que tiene un carácter aleatorio, en el sentido de que en distintas observaciones de la misma categoría puede tomar valores diferentes, no determinísticamente determinados, la caracterización de una variable aleatoria supone por tanto, la determinación del conjunto de valores que puede tomar efectivamente la variable y además la probabilidad con que ellos pueden ser tomados.

Parece natural definir que una variable aleatoria es continua si su función de distribución $F(x)$ es una función continua en todos los puntos de \mathbf{R} . En realidad, para caracterizar una variable aleatoria continua, se exige algo más fuerte sobre $F(x)$: que ella sea continua diferenciable, esto es que exista:

³ Data Mining Institute, S.L. (1997). “estadístico.com” en *Diccionario estadístico*. [En línea]. España, disponible en: <http://www.estadistico.com/dic.html?p=1267> [Accedido el día 10 de noviembre de 2005].

⁴ Spiegel M. (Octubre 1961). *Teoría y problemas de Estadística. Variables discretas y continuas*. La Habana 1979.

$$f(x) = F'(x) = \frac{dF(x)}{dx}$$

en casi todos los puntos y sea diferente de cero (la frase “en casi todos los puntos” puede entenderse, por los no matemáticos, en el sentido más simple de “en todos los puntos excepto a lo sumo en un número finito”).

Desde el punto de vista puramente matemático, esta restricción puede ser esencial. Hay ejemplos de funciones continuas, crecientes de 0 a 1, diferenciables “en casi todos los puntos” con derivada nula; pero en la práctica tales funciones no se presentan como funciones de distribución de variables aleatorias conocidas.

Una variable aleatoria discreta tiene un nivel de medición “ordinal” si sus valores posibles han sido establecidos de manera que expresan cierto orden en las categorías del fenómeno aleatorio que caracterizan, aunque la magnitud absoluta de estos valores no tenga un significado específico.⁵

Algunos ejemplos de variables aleatorias continuas son el peso de un grupo de personas, la estatura, el tiempo de funcionamiento sin fallos de un equipo, el tiempo asignado a cada tarea en un procesador de computadora. De manera general todas las muestras relacionadas a tiempos sin restricciones especiales describen variables aleatorias continuas.

1.1.4 Concepto de muestra y población.

Al recolectar datos concernientes a las características de un grupo de individuos u objetos, tales como las estaturas, los pesos o el número de tornillos defectuosos en la producción de una fábrica en un día dado, es casi imposible observar al grupo completo, especialmente si es grande. Por lo que generalmente en vez de tomar el grupo completo, llamado población o universo, solo se recoge una pequeña parte de ellos denominada muestra.

⁵ Grau R, (2004), “Conferencia 1. Introducción a la estadística.” En la UCLV. *Conferencias de Estadísticas*. Facultad MFC, Universidad Central de Las Villas, Santa Clara, 2004.

Una población puede ser denominada finita o infinita. Por ejemplo, la población que consiste en todos los tornillos antes mencionados es finita, mientras que la población referida a los tiempos de vida entre fallos de un equipo es infinita.

Si se supone que se quiere estudiar un conjunto de objetos homogéneos respecto a cierto índice cualitativo o cuantitativo que caracteriza esos objetos. Por ejemplo, si se tiene un grupo de estudiantes, el índice cualitativo puede ser el sexo y el índice cuantitativo puede ser el índice académico.

Cuando en ocasiones es posible y se realiza una investigación total, es decir, se examina cada uno de los objetos del conjunto respecto al índice que interesa, se está analizando la población en su conjunto. Sin embargo en la mayoría de los casos esto no es posible, porque el conjunto es muy grande o infinito o por ejemplo, porque el estudio lleva implícito un conjunto considerable de gastos materiales o la destrucción del objeto de estudio. En estos casos se escogen fortuitamente del total un número limitado de objetos y se someten éstos al estudio. A este subconjunto se le llama conjunto muestral o simplemente muestra y se asocia con el conjunto general o población del cual se extrae.

Es importante señalar que para que los resultados que aporta el estudio puedan tener una repercusión en la población de la que fue extraída la muestra esta última tiene que ser representativa de la población.

1.1.5 Breve reseña de la teoría de la fiabilidad.

La fiabilidad comienza a usarse y a analizarse como una rama de la estadística primeramente por las compañías de seguros. Fue utilizada para calcular la mortalidad derivada de las epidemias y de otros sucesos naturales o creados por el hombre y de esta forma hallar el riesgo de sus pólizas ante un suceso de este tipo. En esta época se usaban como herramientas las tablas de vida, de las cuales se conoce desde 1693 gracias a Edmund Halley, astrónomo inglés.

Ya a principios de 1900 se utilizaban los métodos actuariales para determinar la supervivencia en pacientes sometidos a tratamientos y en el cálculo de la fiabilidad de diferentes equipos principalmente de los ferrocarriles. Por este motivo se puede hacer alusión al término fiabilidad industrial a partir del siglo XIX lo que clasifica a la

teoría de la fiabilidad como una rama joven de la ciencia. Precisamente el desarrollo vertiginoso de la tecnología, de la mecánica y de complejos equipos militares en el siglo XX fue la razón para estudiar, en los centros de producción y mantenimiento, los métodos estadísticos que permitieran la predicción del funcionamiento de cualquier producto.

Entre las distribuciones con una enorme diversidad de aplicaciones en la estadística y en la fiabilidad en particular, se encuentra la normal. También se llama distribución de Gauss o distribución de Laplace-Gauss. Ello se debe a que el matemático francés Pierre Simon de Laplace fue el primero que demostró la siguiente relación, muy

importante en el estudio de la distribución normal $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$.

Sin embargo, muchos autores consideran como auténtico descubridor de la distribución normal a Abraham De Moivre, quien publicó en 1733 un folleto con el título de Approximatio ad summan terminorum binomii, en el que aparece por primera vez la curva de la distribución de errores $(a + b)^n$, que pasando el tiempo, y con no cierta injusticia, se conoce como distribución de Gauss.⁶

En 1939 Walodje Weibull, profesor del Royal Institute of Technology de Suiza propuso una distribución para describir la duración de materiales, que más tarde llevaría su nombre.⁷ La distribución de Weibull es precisamente el centro de las investigaciones en la teoría de la fiabilidad hoy y es la distribución que más se ha estudiado en el Sistema Estadístico para el estudio de la Fiabilidad.

En 1951 Epstein y Sobel comenzaron a aplicar las características de la distribución exponencial como modelo probabilística para estudiar el tiempo de vida en dispositivos. Actualmente este modelo es utilizado en la mayoría de los estudios realizados a equipos electrónicos que poseen la propiedad de ser en su mayoría no reparables y generalmente la muestras de funcionamiento sin fallos recogidas de elementos como estos se ajustan a una distribución de este tipo con un parámetro λ

⁶ Colectivo de autores. Anecdótico matemático. España, disponible en: <http://www.etsi2.ugr.es/profesores/anecdótico/anecdótico.htm>. [Accedido el día 10 de mayo de 2006]

⁷ Colectivo de autores (2003). "edicionsupc.es". Introducción a la Teoría de la Fiabilidad. España. disponible en: <http://www.edicionsupc.es/ftppublic/pdfmostra/ME03301M.pdf> [Accedido el día 15 de mayo de 2006].

determinado. En el **Anexo 1.1** se muestra una tabla que resume algunos de los sucesos más importantes en el desarrollo de las estadísticas como una ciencia.

1.2 Aplicaciones generales de la estadística en la ingeniería y la medicina.

1.2.1 Utilización en la teoría de la fiabilidad.

Durante el siglo XVIII comienza el auge de la estadística descriptiva en asuntos sociales y económicos y es a finales de este siglo y comienzos del XIX cuando empiezan a sentarse verdaderamente las bases teóricas de la teoría de probabilidades con los trabajos de Joseph Louis Lagrange (1736-1813) y Pierre Simon de Laplace (1749-1827), el astrónomo alemán Carl Friedrich Gauss (1777-1855), y de Siméon-Denis Poisson (1781-1840). Previamente cabe destacar el descubrimiento de la distribución normal por Abraham de Moivre (1667-1754), distribución que será posteriormente "redescubierta" por Gauss.

Una vez sentadas las bases de la teoría de probabilidades, el nacimiento de la estadística moderna y su empleo en el análisis de experimentos, se puede situarlo en los trabajos de Francis Galton (1822-1911) y Karl Pearson (1857-1936). Este último publicó en 1892 el libro The Grammar of Science, un clásico en la filosofía de la ciencia y fue él quien ideó el archiconocido test del Chi cuadrado⁸.

Pero es Ronald Arnold Fisher (1890-1962) sin lugar a dudas la figura más influyente de la estadística, situándola como una poderosa herramienta para la planificación y análisis de experimentos. Contemporáneo de Pearson, desarrolló el análisis de la varianza, fue pionero en el desarrollo de numerosas técnicas de análisis multivariante, y en la introducción del método de máxima verosimilitud para la estimación de parámetros. Su libro Statistical Methods for Research Workers publicado en 1925 ha sido probablemente el libro de estadística más utilizado durante mucho tiempo.

A partir del fuerte desarrollo armamentista que sucedió en el mundo tras la primera y segunda Guerra Mundial la necesidad de medir la probabilidad de funcionamiento sin fallos de los aparatos militares provocó una mayor aplicación de los métodos

⁸ Luis M. Molinero. (Julio 2004). Asociación de la Sociedad Española de Hipertensión. Historia del razonamiento estadístico, disponible en: <http://www.seh-lelha.org/historiastat.htm>. España. [Accedido el 29 de mayo de 2006].

estadísticos en estos estudios. Esto unido al desarrollo de la mecánica y la tecnología en el siglo XIX y XX, fue la fuente de la aparición de fiabilidad industrial como toda una teoría.

Desde su surgimiento ha sido aplicada a la determinación de la probabilidad de funcionamiento sin fallos de cualquier equipo, ya sea un componente electrónico, una pieza mecánica, un equipo, una construcción, un individuo o cualquier materia para la que se recoja una muestra de tiempos como variable aleatoria.

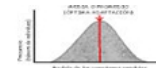
La investigación de sistemas de fiabilidad en general y en particular las funciones de sistemas coherentes se inició en 1961 a partir del artículo publicado por Bimbaum, Esary y Saunders.

Hoy en día los métodos estadísticos posibilitan el análisis de tiempos de vidas entre fallos. Esta propiedad auxilia la planificación de mantenimientos o la sustitución de un equipo, pieza o construcción. Además permite el análisis de una muestra de tiempos de vidas hasta el primer fallo, en el caso de equipos no reparables o si se desea determinar la garantía de producción. Esta aplicación es fundamental a la hora de fabricar un producto a gran escala. Para ello se deben probar algunas muestras del producto y realizando el análisis de la fiabilidad se decide si es rentable o no su producción.

Finalmente se puede asegurar que las ramas de la producción que hoy aplican elementos de la teoría de la fiabilidad para sus producciones o reparaciones son disímiles y abundantes, van desde la electrónica hasta la astronáutica, pasando por la construcción, la agricultura, la mecánica, la computación, en general por casi todos los renglones que rigen la economía de un país.

1.2.2 Utilización en la medicina.

En la actualidad la estadística ocupa un lugar de gran importancia tanto en la investigación como en la práctica médica; en los estudios de medicina de cualquier país se incluyen varias asignaturas dedicadas a la estadística; es difícil, por no decir imposible, que un trabajo de investigación sea aceptado por una revista médica sin la utilización de técnicas y conceptos estadísticos en su planteamiento y en el análisis de los datos. Y sin embargo esta situación es bastante reciente, baste señalar que el gran



auge de la utilización del método estadístico, tanto para la planificación de experimentos como para el análisis de los datos obtenidos, puede ser situado en los trabajos de quien sin lugar a dudas se considera como el padre de la estadística moderna, Ronald A. Fisher (1890-1962), aunque su aplicación generalizada en la medicina tardó más en llegar.

El comienzo de los ensayos clínicos realizados aleatoriamente en medicina, fundamentados en esos trabajos de Fisher, empiezan a despegar con Sir Austin Bradford Hill (1897-1991). Precisamente los editores de Lancet, comprendiendo la necesidad de difundir y explicar las técnicas estadísticas a los médicos, encargaron a Hill escribir una serie de artículos destinados a explicar el uso correcto de la estadística en medicina; artículos que posteriormente darían lugar a un libro del género, Principles of Medical Statistics, cuya primera edición corresponde a 1937, y la última a 1991.

Se citan los trabajos de un eminente médico francés, Pierre-Charles-Alexandre Louis (1787-1872), como uno de los primeros en aplicar el razonamiento científico en un trabajo de investigación sobre la fiebre tifoidea, en el que estudió la mortalidad en relación con la edad de los pacientes, el tiempo de estancia en París y la eficacia de las sangrías.⁹

Otro de los ejemplos de la utilización de la estadística en la medicina es en América, en 1954, cuando se llevó a cabo el mayor y más costoso estudio médico jamás realizado, para evaluar la eficacia de la vacuna de Salk como protección contra la poliomelitis.

En los días de hoy en todos los estudios realizados por las Organizaciones de Naciones Unidas tienen un fuerte basamento estadístico y las publicaciones que se realizan sobre los índices de salud de cada país están analizadas estadísticamente.

Tan importante se considera la ciencia en nuestro país que además es impartida como una asignatura del pregrado en una gran cantidad de carreras de nuestro país, en especial, en las Ciencias Médicas impartida con el nombre de Bioestadística.

⁹ Luis M. Molinero. Julio 2004. Asociación de la Sociedad Española de Hipertensión. Historia del razonamiento estadístico, disponible en: <http://www.seh-lelha.org/historiastat.htm>. España. [Accedido el 29 de mayo de 2006].

1.2.3 Características que potencian su utilización.

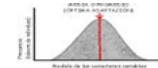
La estadística como todas las matemáticas ha seguido su desarrollo continuo desde su aparición hasta la actualidad. Quizás esta fue una de las más lentas en este proceso, sin embargo a medida que se han incorporado métodos y algoritmos, rápidamente se ha visto su aplicación, muchos han aparecido por necesidades de la vida práctica. De esta forma se puede decir que cada una de las distribuciones que existen describe algún proceso de la vida, la producción o es la base de uno de estos.

En los deportes en particular son ampliamente utilizadas las estadísticas para analizar el comportamiento de cada uno de los equipos y de los jugadores en cada temporada del año, campeonato, territorio o frente a algún oponente y a partir de ahí tomar las decisiones de qué deportistas utilizar en cada partido o cómo jugarle a cada contrincante incluso, si es preferible sacrificar algún partido.

Su gran vínculo con la teoría de las probabilidades y debido a que se dedica a estudiar procesos aleatorios también ha servido para realizar estudios sobre la aleatoriedad de algunos juegos como los naipes, los dados, entre otros.

Sus propiedades de cálculos de promedios permiten generalizar, con un fundamento matemático, las proposiciones verbales que a menudo son planteadas. Mediante los tests de bondad del ajuste se puede llegar a un conocimiento previo de que debe suceder con una muestra determinada, analizando su comportamiento y la distribución a que se ajusta. Así esta propiedad ha sido utilizada en herramientas computacionales creadas para ayudar el manejo de los negocios a las empresas actuales o hasta en la predicción de enfermedades o brotes de infecciones en el mundo con un alto nivel de exactitud.

Además existen métodos estadísticos que permiten la estimación de parámetros para varias distribuciones, estos parámetros describen la distribución en el tiempo o espacio, la forma y la dispersión de un grupo de datos. En el software que con este trabajo se culmina, se estiman los parámetros de las distribuciones normal, exponencial y de weibull a través de fórmulas obtenidas a partir de los métodos de máxima verosimilitud.



Permiten el cálculo de valores que se utilizan como base para la teoría de la fiabilidad y por tanto la planificación de mantenimientos. Algunos de los valores que son calculados a partir de los parámetros son los intervalos de confianza que, siendo otra de las potencialidades del cálculo estadístico, delimitan el riesgo que se puede correr a la hora de tomar una decisión.

De manera general en cualquier histograma o resumen histórico que se desee realizar sobre una población de individuos o elementos necesita de la aplicación de los métodos estadísticos de cálculo. Por lo que actualmente es difícil encontrar una guía de béisbol, un resumen médico de una campaña de vacunación o de higienización, un análisis de encuestas realizadas a la población, los resultados de una votación, la planificación de mantenimientos de un equipo o algún proceso social, político o económico donde no se utilice la estadística.

1.3 Herramientas computacionales existentes.

Con la aparición de la computación, las disímiles aplicaciones de la estadística rápidamente invadieron esta nueva ciencia. De esta forma se comenzó a recoger y analizar el comportamiento estadístico de las computadoras y de sus componentes. Pero la amplia gama de problemas que pueden ser resueltos mediante una computadora abarcaron también a los problemas que necesitaban de la estadística para su resolución y así comenzaron a aparecer herramientas como el STATA (México), el SAS (E.U.A), el SPSS (E.U.A.), el Weibull++ (E.U.A.) entre las más conocidas, todas desarrolladas para el análisis numérico o gráfico de muestras continuas o discretas de datos.

Actualmente sería prácticamente imposible resumir todas las herramientas computacionales que existen en el mundo que realizan cálculos estadísticos. Uno de los intentos realizados para relacionar una gran cantidad de softwares, documentos y sitios web es la página Statistics and Statistical Graphics Resources accesible en: <http://www.math.yorku.ca/SCS/StatResource.html> y que posee cerca de 300 elementos entre sitios, softwares y documentos sobre las estadísticas. Aquí aparecen relacionados, además de los ya mencionados, otros softwares como son, Mathematica, MATLAB, Statistic y StatGraphics.

También la difusión de la Internet por todo el mundo propició la aparición de sitios web dedicados a realizar cálculos estadísticos online entre estos pueden ser mencionados SASCOM online además en <http://statpages.org/> puede ser encontrada una lista de aplicaciones web y de softwares dedicados a realizar análisis estadísticos de datos.

1.3.1 Usos y fines.

Entre los softwares más difundidos en la red que permiten el análisis de datos se encuentra el Weibull++ de ReliaSoft que fue desarrollado para realizar el análisis de datos de tiempos de vida o de funcionamiento sin fallos o entre fallos (Análisis de Weibull) utilizando distintas distribuciones de vida incluyendo todas las formas de distribución Weibull. Él posee varias herramientas de análisis permitiendo diversos tipos de cálculos e informes gráficos.¹⁰ Puede ser utilizado ampliamente en la teoría de la fiabilidad ya sea con fines docentes o en la vida práctica.

SPSS Inc. es un proveedor mundial principal de softwares de análisis predictivo y soluciones. Entre los varios productos que hoy comercializa esta compañía está el SPSS for Windows. Fundado en 1968, hoy SPSS tiene más de 250,000 clientes mundiales y es servido por más de 1,200 empleados en 60 países. SPSS se ha vuelto un líder en las tecnologías del análisis predictivo a través de una combinación de compromiso a la innovación y dedicación a los clientes.¹¹

Se encuentran clientes de SPSS en virtualmente todas las industrias, incluso las telecomunicaciones, la banca, las finanzas, el seguro, la salud, la fabricación, los estudios socioculturales, las investigaciones psicológicas, la educación más alta, el gobierno e investigaciones del mercado. Los clientes del mercado usan el software SPSS de análisis predictivo para anticiparse el cambio, manejan funcionamientos diarios y las iniciativas especiales más eficazmente, y comprenden los beneficios positivos. Incorporando el análisis del predictivo en sus funcionamientos diarios, se

¹⁰ ReliaSoft. (2006). Weibull++ 7. *Análisis de Datos de Vida (Análisis del Weibull)*. EE.UU., disponible en: <http://www.reliasoft.com.br>. [Accedido el 9 de abril del 2006].

¹¹ SPSS Inc. (2006). SPSS. *Softwares and Solutions*. EE.UU., disponible en: http://www.spss.com/products/index.cfm?source=homepage&hpzone=nav_bar. [Accedido el 30 de mayo de 2006]

vuelven grandes empresarios habilitados para dirigir y automatizar las decisiones para ganarse las metas comerciales y lograr la ventaja competitiva.

Como ya se había mencionado otro de los softwares que cuenta con mayor distribución es el SAS, posee una amplia campaña publicitaria dirigida fundamentalmente a su utilización en los negocios capitalistas para hacer más rentable sus producciones o para tomar decisiones más correctas a partir del análisis estadístico de los resultados de su compañía o de algunas similares. Este es un software que entrega tecnologías y especializadas en:

Inteligencia comercial que permite integrar los datos de su empresa, y entregar el acceso rápido a la información de autoservicio y el análisis a todos los niveles de usuario. Fácil de usar las interfaces, lo hacen simple para que fabricantes de decisiones puedan conseguir las respuestas que ellos necesitan ahorrándoles tiempo y manejando buenas decisiones.¹²

La Integración de los datos, ganando un acercamiento integrado para manejar sus volúmenes crecientes de datos de la empresa. Simplificando el proceso de extracción de los datos, transformación y carga mientras mejora la calidad de los datos y la verdadera integración de estos.

Las Soluciones de industria. Dado que una solución no encaja en todos los clientes SAS mantiene las soluciones innovadoras individuales casi para cada industria. Estas soluciones combinan las fuerzas de sus tecnologías con su especialización del dominio para ayudar rápidamente a alcanzar los objetivos del usuario.

La Plataforma de Inteligencia de la empresa de SAS se ocupa del proceso entero de almacenamiento y análisis de datos de una empresa para encontrarse sus objetivos comerciales.

Otro de los conocidos en el mundo es el STATA producido por México y que pone cientos de herramientas estadísticas en las manos del usuario, desde técnicas avanzadas, como modelos de supervivencia con fragilidad, regresiones de panel de datos dinámicos (DPD) , ecuaciones de estimación generalizadas (GEE) modelado

¹² SAS. (2006). Superior Software that give you the power to know. SAS. EE.UU, disponible en: <http://www.sas.com/index.html>. [Accedido 7 de mayo de 2006].

con selección de muestra, (ARCH) y estimación con muestras de encuestas complejas; hasta modelos lineales y generalizados (GLM), regresiones con conteo o resultados binarios, ANOVA/MANOVA, ARIMA, análisis de conglomerados, estandarización de proporciones y análisis de casos controlados; y a las tabulaciones y estadísticas básicas. STATA también tiene herramientas avanzadas para manejar datos especializados tales como datos de supervivencia /duración, datos de series de tiempo, datos de panel/longitudinales, datos categóricos, y reporte de datos. También se pueden leer un conjunto de datos a través de plataformas (Windows, Mac, Linux y Unix) sin necesidad de conversión e importar conjuntos de datos de otros programas estadísticos, hojas de cálculo y bases de datos rápida y fácilmente.¹³

1.3.2 Deficiencias.

A pesar de las potencialidades de los softwares que ya existen en el mundo para realizar estudios estadísticos existen un grupo de deficiencias que nos llevaron a la realización del “Sistema Estadístico para el estudio de la Fiabilidad”.

En el caso del Weibull++ no es un software libre que puede ser utilizado por toda la comunidad sino que tiene que ser pagada su licencia. Sin embargo aún si se desea pagar dicha licencia en el caso particular de nosotros, los cubanos, no se puede por el bloqueo que nos tiene impuesto el gobierno de los Estados Unidos. Muestra de ello es el texto que a continuación se relaciona enviado por una empresaria de la compañía ReliaSoft con oficinas en los Estados Unidos de América ante la solicitud de información sobre el Weibull++ 7:

Dear Sir,

Thank you for expressing an interest in our reliability statistical analysis software. Due to the U.S. trade sanctions imposed on your country, we are unable to supply you with software at this time.

We do, however, offer some free reliability analysis tools on our reference website. Please visit the following link for access to these tools:

¹³ STATA. (2006). STATA. México., disponible en: <http://www.STATA.com.mx> [Accedido 25 de mayo de 2006].

<http://www.weibull.com/itools/index.htm>.

Sincerely,

*Jenna*¹⁴

Estimado Señor,

Gracias por expresar un interés en nuestro software del análisis estadístico de la fiabilidad. Debido a las sanciones de comercio americanas impuestas a su país, nosotros somos incapaces de proporcionarle el software en este momento.

Sin embargo, le ofrecemos algunas herramientas de análisis de fiabilidad libre en nuestro website. Por favor visite el hipervínculo siguiente para el acceso a estas herramientas: <http://www.weibull.com/itools/index.htm>.

Sinceramente,

Jenna

A continuación se dan los datos de la empresa ReliaSoft y de la empresaria Jenna Thompson para que pueda ser verificado el hecho aquí expuesto que define claramente la mayor deficiencia que puede tener el software Weibull++.

Jenna L. Thompson

ReliaSoft Corporation 115 South Sherwood Village Drive Tucson, AZ 85710 USA

Jenna.Thompson@ReliaSoft.com

Toll free (U.S. & Canada): 1-888-886-0410 Phone: +1-520-886-0366 x2239

Fax: +1-520-886-0399 Corporate Web Site: <http://www.Reliasoft.com/>

Reliability Resources Web Site: <http://www.Weibull.com>

Abordando el SPSS como uno de las herramientas computacionales más competentes en la realización de estudios estadísticos, se puede analizar, sin embargo, que su enfoque es fundamentalmente hacia estudios sociales y rara vez es utilizado para estudios de fiabilidad, aunque en sus últimas versiones se han incluido módulos para esta teoría.

Con este software se puede realizar un estudio a cualquier variable aleatoria, sea continua o discreta, esto a pesar de ser conveniente para la estadística, no lo es en el

¹⁴ Thompson, J. (2006) *Your Request*. Correo electrónico a D. Valdés. (dvramirez@gmail.com) 18 de mayo de 2006 [Accedido el 18 de mayo de 2006].



caso de los estudios sobre la teoría de la fiabilidad debido a que las muestras son tiempos de vida o de funcionamientos entre fallos o hasta el primer fallo las que constituyen variables continuas. De esta forma un paquete estadístico para el estudio de la fiabilidad que además analice cualquier variable discreta puede ser excesivamente grande sin necesidad lo que provocaría un mayor consumo de recursos o que incluso no pueda ser utilizado en una computadora con prestaciones limitadas como puede suceder. Por otra parte tanto el SPSS como el SAS no han explotado ampliamente las potencialidades de la distribución de Weibull que logra el ajuste de cualquier muestra de tiempos.

Nuevamente se convierte en la principal deficiencia el costo de los softwares y el país de fabricación en el caso del SAS (E.U.A.) y el STATA (México). Debido ambos a que son producidos en los Estados Unidos de América o en un país firmante de un TLC con el gobierno norteamericano. Pero además se debe señalar, que en los dos casos, los softwares tienen un enfoque mayoritario a los estudios estadísticos de la macroeconomía de las empresas capitalistas diseñadas para superproducciones.

En el Sistema Estadístico para el estudio de la Fiabilidad (SEF) se ha potenciado la realización de los gráficos de los datos debido a que el análisis gráfico, en muchos casos, en la teoría de la fiabilidad brinda una mejor visión del problema y de su solución. Sin embargo los gráficos generados con los softwares antes mencionados no relacionan todas las funciones que se realizan con el SEF.

1.4 Soluciones que da el Sistema Estadístico para el estudio de la Fiabilidad a las deficiencias existentes.

Tomando en cuenta las limitaciones que se han planteado en el epígrafe anterior y conociendo las que posee el “Sistema estadístico para el estudio de la Fiabilidad” se puede establecer una pequeña comparación entre este software y los existentes en el mundo de la estadística y la fiabilidad.

La primera y gran ventaja que ofrece este nuevo producto es que es realizado en Cuba por lo que ninguna empresa, centro de enseñanza o de investigación, institución social o de salud que desee utilizarlo tienen restricciones o prohibiciones surgidas de un bloqueo económico para utilizarlo.

Su fácil interfaz visual le garantiza al usuario un rápido aprendizaje en su utilización. No se necesita de grandes conocimientos de estadísticas o de teoría de fiabilidad para aplicarlo a un problema en específico.

Sus cálculos son basados en tres distribuciones, normal, exponencial y de Weibull, con las que se logra abarcar cualquier problema de la fiabilidad industrial y una gran cantidad de problemas de bioestadísticas.

La interfaz gráfica ha sido detallada para que el usuario con medianos o profundos conocimientos de estadística y de fiabilidad pueda analizar las soluciones brindadas por el software y correr ciertos riesgos en la toma de decisiones, que el producto en su primera versión no implementa, según el tiempo de explotación del equipo o elemento cuya muestra de tiempos de vida se está analizando.

Un resumen textual, se puede ver en el SEF, al realizar el estudio independiente para cada muestra que incluye una propuesta de planificación de mantenimientos o de intervalos de tiempos que son vitales para la toma de decisiones de quienes lo utilizan. Esta propiedad hace que cualquier usuario con pequeños conocimientos de computación pueda acudir al Sistema Estadístico para el estudio de la Fiabilidad como apoyo en la toma de sus decisiones, la determinación de la garantía de producción de un producto o la planificación de mantenimientos de sus equipos.

Almacena sus datos en ficheros que pueden ser leídos con muchos otros softwares convencionales lo que le da gran portabilidad y facilidades para recoger las muestras de datos durante un período de tiempo. Los datos pueden ser entrados al software a partir de un fichero texto con la estructura correcta.

Además le brinda al usuario seguridad e integridad en los datos. Si se utiliza el SEF para recoger y almacenar los datos recogidos durante un período de tiempo se pueden almacenar en una base de datos que posee el software manipulada completamente desde él. Esta base de datos solo utiliza el motor de Microsoft Access que se encuentra presente en casi todas las computadoras del mundo con sistema operativo de la familia Windows.

Posee también opciones para la generación de valores que sigan una distribución especificada por el usuario o similar a la que tienen los datos que ya aparecen para la variable en cuestión. Así no es necesario tener la muestra completa para realizar el

estudio de la fiabilidad del equipo sino que, puede correr cierto riesgo y utilizar valores generados por el SEF para aumentar el tamaño de su muestra y que los resultados obtenidos de la aplicación de los test sean más exactos.

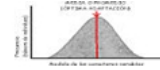
Todas estas bondades que se mencionan vienen a resolver deficiencias de los softwares que permiten la realización de estudios estadísticos o de la fiabilidad de productos o individuos, más conocidos y distribuidos por el mundo. En la mayoría de estos se necesita de algún conocimiento estadístico para su utilización, hay que pagar grandes sumas de dinero para comprarlos algunos no permiten que sus ficheros sean vistos con otro editor por lo que hay que tener el software desde que se va a comenzar a recoger la muestra para el estudio y esto puede demorar un tiempo considerable como para que ya se hayan publicado otros programas más eficientes o menos costosos.

1.5 Conclusiones parciales.

A partir del análisis realizado en este capítulo, basándonos en la importancia que tiene la estadística en los procesos actuales de la sociedad en su conjunto, en lo significativo que ha sido su aporte al desarrollo de la humanidad mediante el análisis matemático de los errores y experiencias pasadas recolectadas en la historia no hay lugar a dudas de que la estadística es una rama de la matemática que se mantendrá en estudio y desarrollo mientras la humanidad exista.

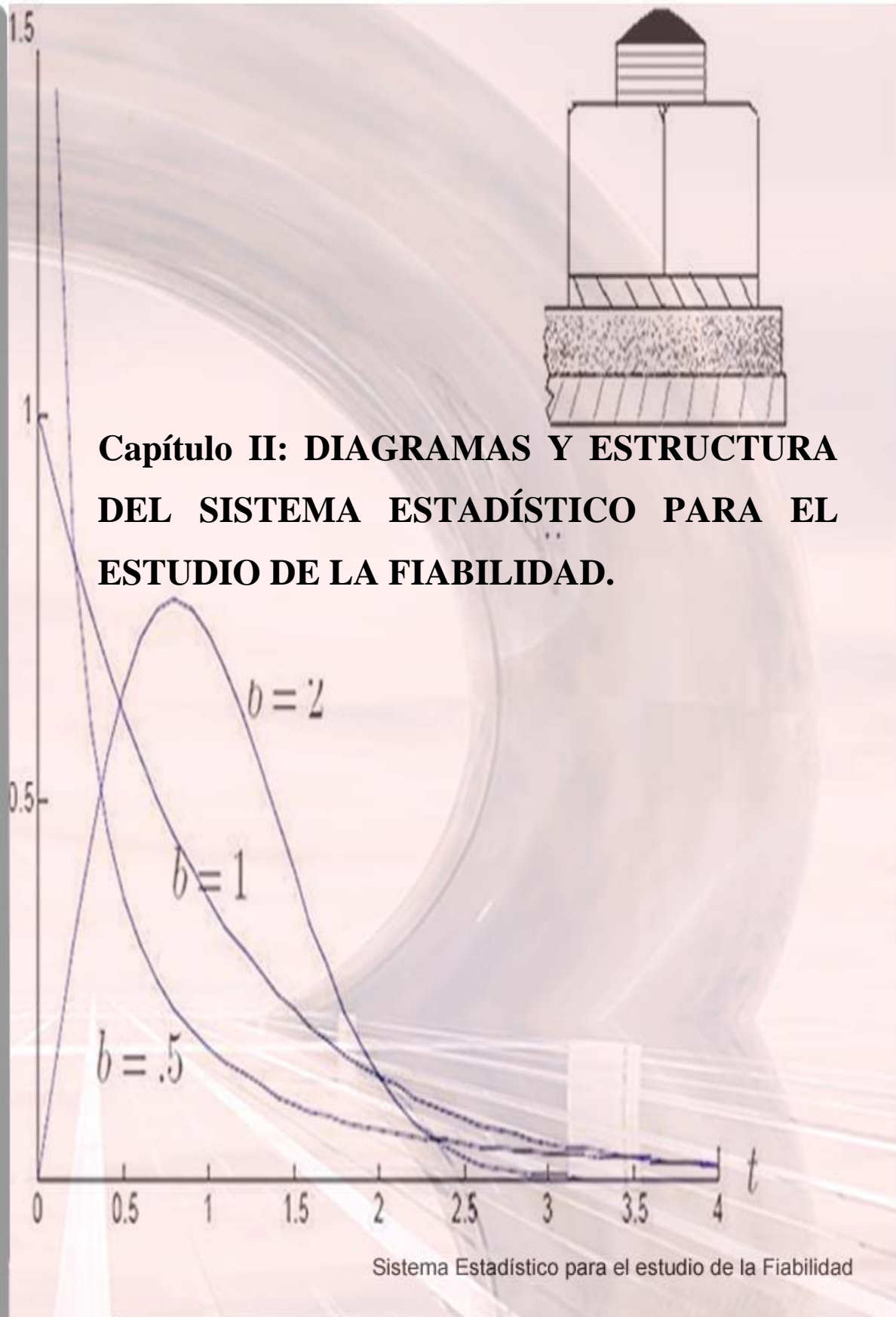
Cada vez que se desarrolle una nueva rama de cualquier ciencia o una nueva ciencia en sí, se necesitará casi seguro de un análisis estadístico de sus componentes, aplicaciones, beneficios, etc.

Por otra parte en el mundo, la mayoría de las transnacionales de la informatización son las que se han dedicado a manipular la información y para ello han desarrollado softwares como los que aquí se han mencionado para recolectar y analizar los datos que luego se convierten en información. Una gran cantidad de estas transnacionales son de los Estados Unidos de América o tienen negocios con ellos, lo que dificulta la compra a Cuba de cualquiera de estas herramientas computacionales.



Estas razones son las que nos han llevado a desarrollar un primer intento de software que permita llevar adelante el análisis estadístico de datos, en este caso tiempos de vida relacionados a la teoría de la fiabilidad. Precisamente ahora cuando estamos en una lucha constante por la obtención de ganancias a partir del ahorro y el incremento de los índices de producción y calidad de nuestros productos, el “Sistema Estadístico para el estudio de la Fiabilidad” se convierte en una herramienta poderosa, utilizable en cualquier lugar del país, para estos fines y viene a aportar un grano de arena a esta batalla.

Capítulo II: DIAGRAMAS Y ESTRUCTURA DEL SISTEMA ESTADÍSTICO PARA EL ESTUDIO DE LA FIABILIDAD.



2. Capítulo II: DIAGRAMAS Y ESTRUCTURA DEL SISTEMA ESTADÍSTICO PARA EL ESTUDIO DE LA FIABILIDAD.

El segundo capítulo está dedicado a acercar al usuario especialista en computación la estructura interna del Sistema Estadístico para el estudio de la Fiabilidad. En siete acápites se presentan los diagramas de flujo de datos, de clases, de la estructura informativa, de secuencias, entre otros. Constituye parte vital de la documentación del SEF y es de obligatoria lectura para próximas versiones.

Para algunos de los diagramas, por su extensión, se han utilizado solo las partes fundamentales del sistema, tal es el caso del diagrama de clases en el cual se han relacionado solo las clases fundamentales.

Cada acápite contiene una explicación acerca del significado del diagrama o de la estructura que tiene en el software.

Una vez concluido el estudio del presente capítulo el lector quedará listo para adentrarse en el manual de usuario o, si lo desea, en el estudio de los métodos matemáticos y estadísticos que se implementan para la obtención de los resultados.

2.1 Diagrama del modelo del proceso del negocio.

Actualmente para realizar un estudio de la fiabilidad de un equipo en una empresa sin ningún sistema computacional, un estadístico se encarga de recoger las muestras de los datos de uno o varios equipos en un período de tiempo; lleva las planillas al archivo y las recoge si necesitan ser modificadas, esta una persona que se encarga de administrar la Base de Datos y manipular las planillas haciendo los cambios cuando un fallo ocurre. Es función del secretario archivar las planillas o entregarlas si el estadístico o el computador las solicita. El jefe de taller ordena realizar el estudio y se encarga de recoger los resultados. Un computador realiza los cálculos estadísticos y a partir de los resultados obtenidos de los parámetros de la distribución ajustada, determinar el tiempo de vida de un equipo de acuerdo con la teoría de la fiabilidad para poder planificar los mantenimientos.

El negocio se representa en el siguiente modelo:

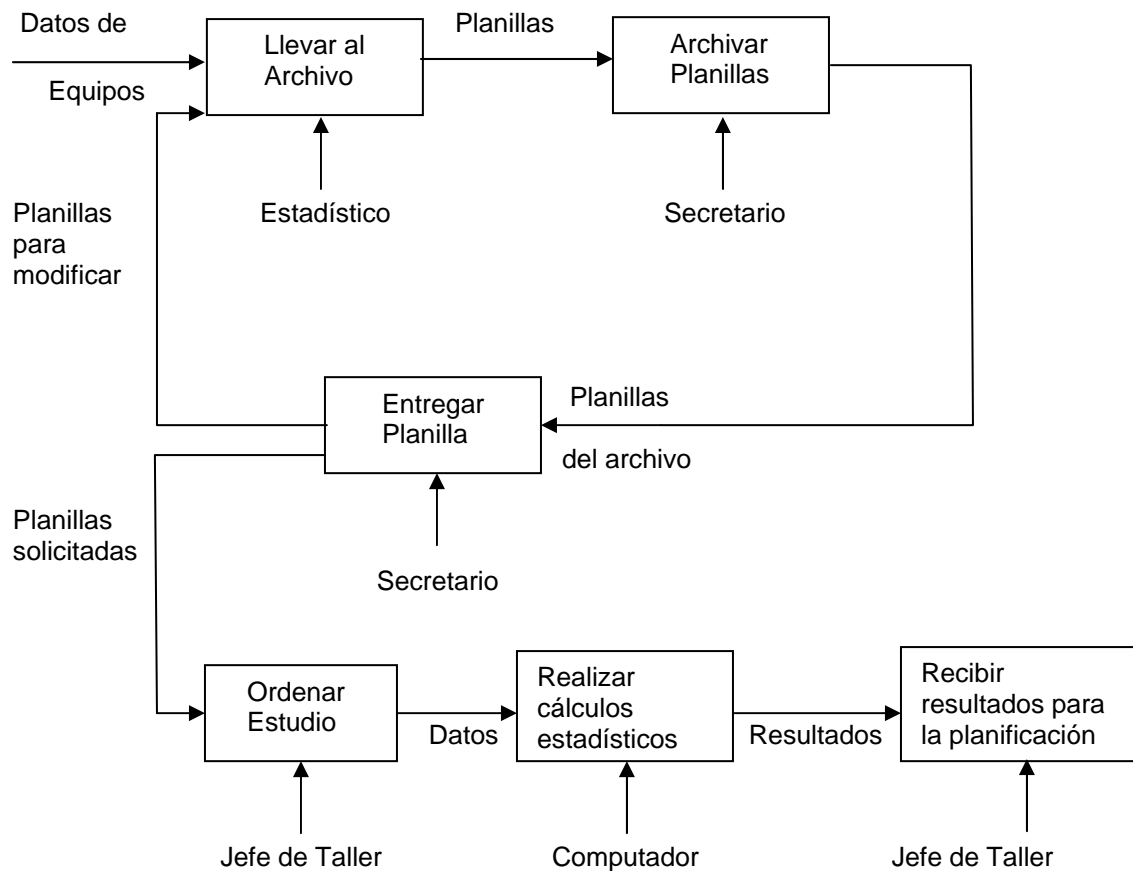


Fig. 2.1 Modelo del negocio.

2.1.1 Actividad fundamental.

La actividad fundamental del Sistema Estadístico para el estudio de la Fiabilidad es la realización de estudios estadísticos a muestras de tiempos de funcionamientos entre fallos o hasta el primero de ellos. A continuación se muestra un diagrama de estados que corresponden a la transición de un estado a otro para cumplir con esta actividad fundamental. En él se pueden observar las acciones que puede realizar cada uno de los actores del sistema y el orden en que se deben completar cada una de ellas.

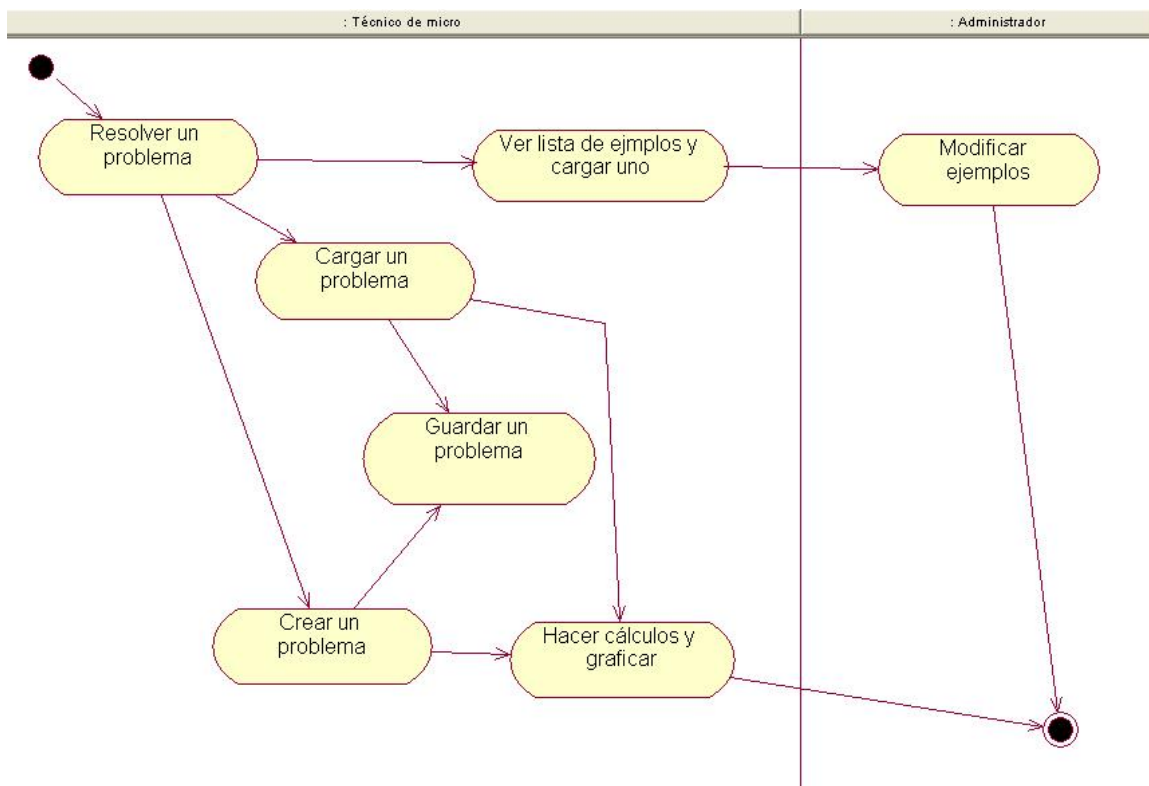


Fig. 2.2 Diagrama de estados de la actividad fundamental.

2.1.2 Diagrama de estados general.

Además del diagrama de la actividad fundamental existe otro que se muestra en el **Anexo 2.1** en el que se relacionan todas las actividades que se realizan en el sistema. Inicialmente al ejecutarlo se encuentra en espera de una orden que puede ser: abrir un problema desde la base de datos, desde un fichero, crear uno nuevo o ver un ejemplo. Al ejecutar una de estas acciones el sistema pasa a un nuevo estado que depende de la acción que se haya seleccionado. A partir de aquí va a ir transitando de estado en estado hasta completar una acción, cuando regresa al estado inicial.

2.1.3 Diagramas de secuencias.

Para cada actividad de las mencionadas en el diagrama de estados, se puede desarrollar un diagrama de secuencias que consiste en la secuencia de las acciones y permite observar sobre qué clase se realiza cada operación, paso a paso, hasta completar una acción. En el siguiente ejemplo se muestran el diagrama de secuencias para la acción: Añadir un problema.

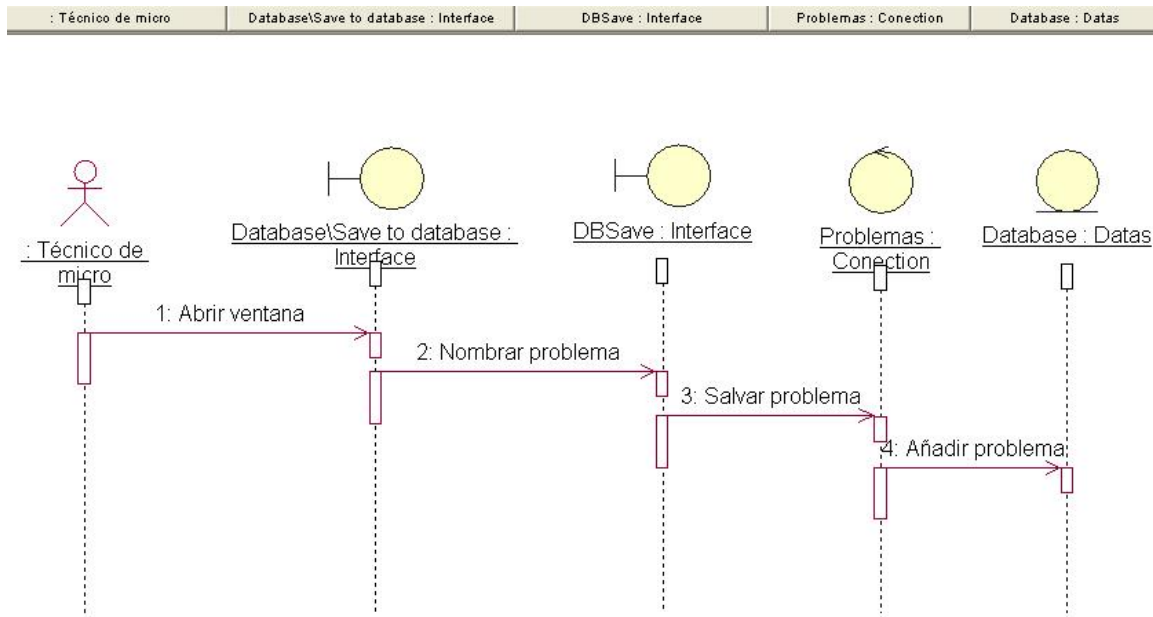


Fig. 2.3 Diagrama de secuencias.

Aquí se puede observar cómo uno de los actores del sistema, el técnico de micro, realiza una acción, abrir una ventana y esta acción desencadena un grupo de operaciones sobre las interfaces de la base de datos y hasta llegar a esta.

2.2 Casos de uso.

El Sistema Estadístico para el estudio de la Fiabilidad invita a su utilización debido a que logra una gran reducción de actores en la realización de la misma operación y con más eficiencia y rapidez. Así de esta forma en el negocio se necesitan varias personas para realizar el estudio estadístico a un equipo, pieza o individuo, dependiendo del tiempo que se quiera estar recogiendo la muestra y de lo experto que sean los individuos que participan en el estudio. Sin embargo, con el SEF con conocimientos medios de estadísticas y de computación el usuario puede rápidamente realizar este estudio y llegar a conclusiones válidas.

2.2.1 Casos de uso del negocio.

En el negocio aparecen como actores el estadístico encargado de recoger las muestras, el secretario que archiva el comportamiento histórico de los equipos, el jefe de taller que decide a que equipos se realizará el estudio y recibe los resultados y el

computador que es el encargado de, con las muestras, realizar la planificación de mantenimiento a partir del resultado.

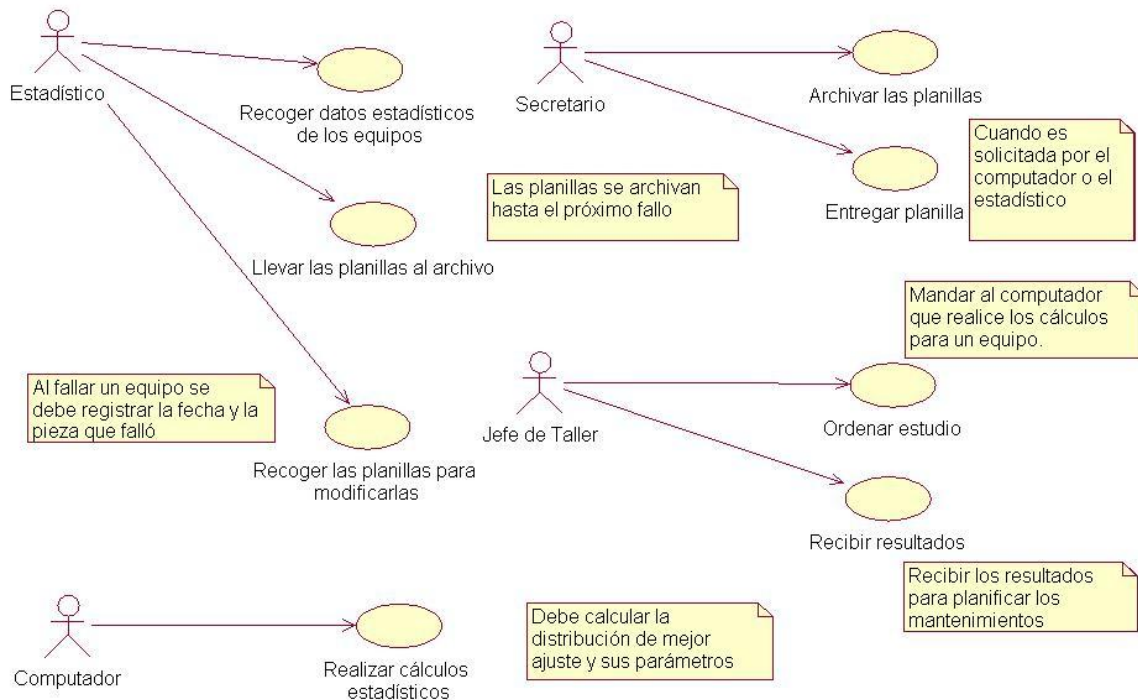


Fig. 2.4 Casos de uso del negocio.

2.2.2 Casos de uso del sistema.

En el sistema además de no necesitar el archivo ni el secretario pues la base de datos o los ficheros pueden almacenar toda esta información con mayor seguridad y menos gastos, el computador, encargado de obtener los resultados, no tiene que ser un experto en la teoría de la fiabilidad ni en la estadística sino que con conocimientos medios de computación puede llevar a cabo un estudio completo, sin embargo si posee algún conocimiento sobre fiabilidad podrá interpretar de mejor forma los resultados que aquí se muestran..

En el sistema aparece además un administrador del sistema que no es necesario pero que puede ser muy útil en ocasiones, si se desea manipular la base de datos. El administrador además será el encargado de dar mantenimiento al sistema, actualizando las versiones del mismo.

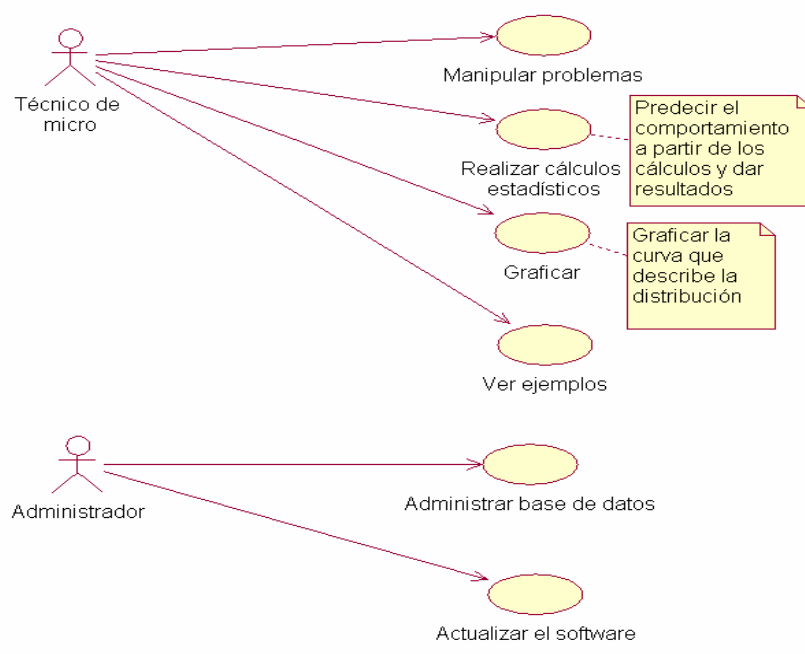


Fig. 2.5 Casos de uso del sistema.

2.2.2.1 Especificidad de actividades.

Para esclarecer las actividades que aparecen en los casos de uso del sistema se muestran estos dos diagramas correspondientes a las actividades fundamentales.

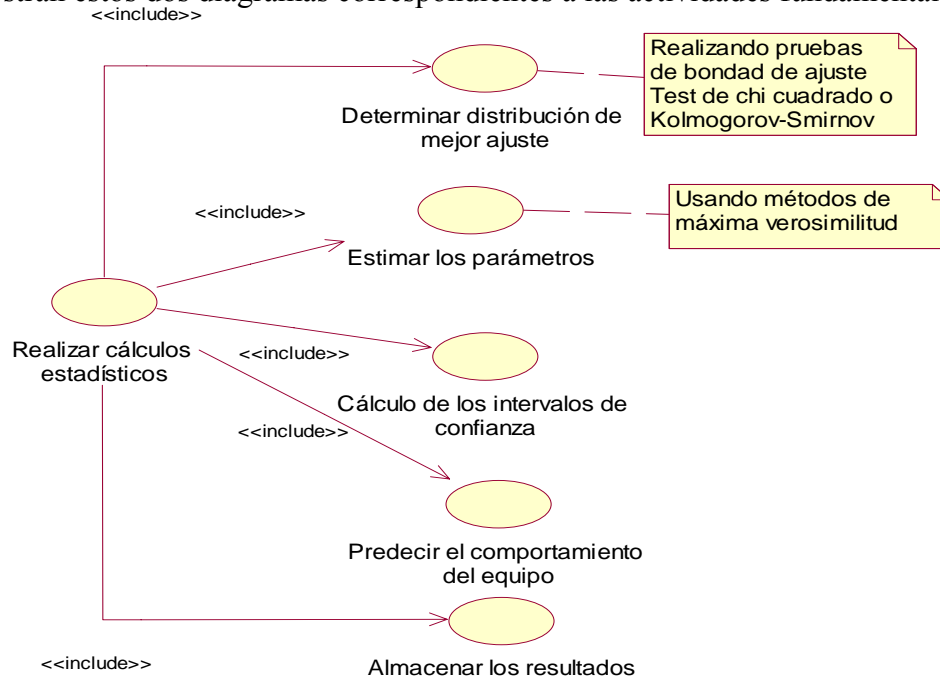


Fig. 2.6 Realizar cálculos estadísticos.

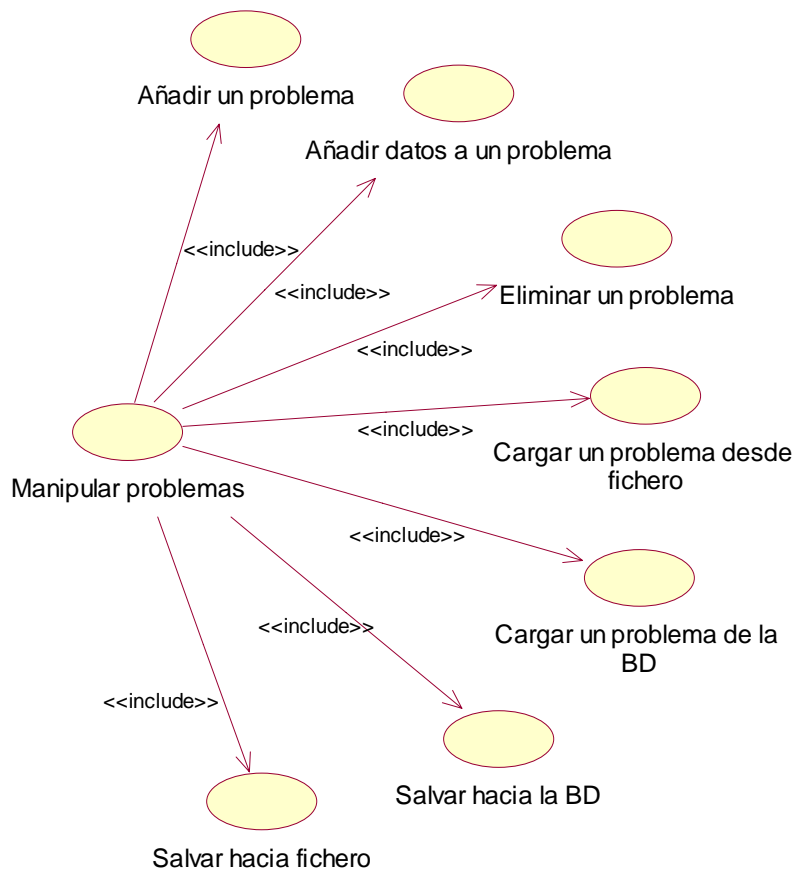


Fig. 2.7 Manipular un problema.

2.3 Diagrama de flujo de datos.

Otro de los diagramas utilizados para modelar o representar un sistema son los DFD. Los diagramas de flujo de datos son herramientas de modelado que permiten representar un sistema como una red de procesos funcionales conectados unos con otros mediante el paso de datos. También son conocidos por: Diagramas de burbujas, Modelo de procesos, Diagramas de flujo de trabajo, etc. Para construir un buen DFD es necesario: - Elegir nombres representativos para los procesos, flujos de datos, almacenes de datos y entidades externas. - Enumerar los procesos. - Volver a dibujar el DFD tantas veces como sea necesario para ganar en claridad y estética. No entrecruzar líneas. - Asegurarse de que el DFD es consistente¹⁵.

¹⁵ Desconocido. 2004. Otros modelos del análisis.ppt, *Diagramas de Flujos de datos*. Facultad MFC, UCLV, Cuba, 2006.

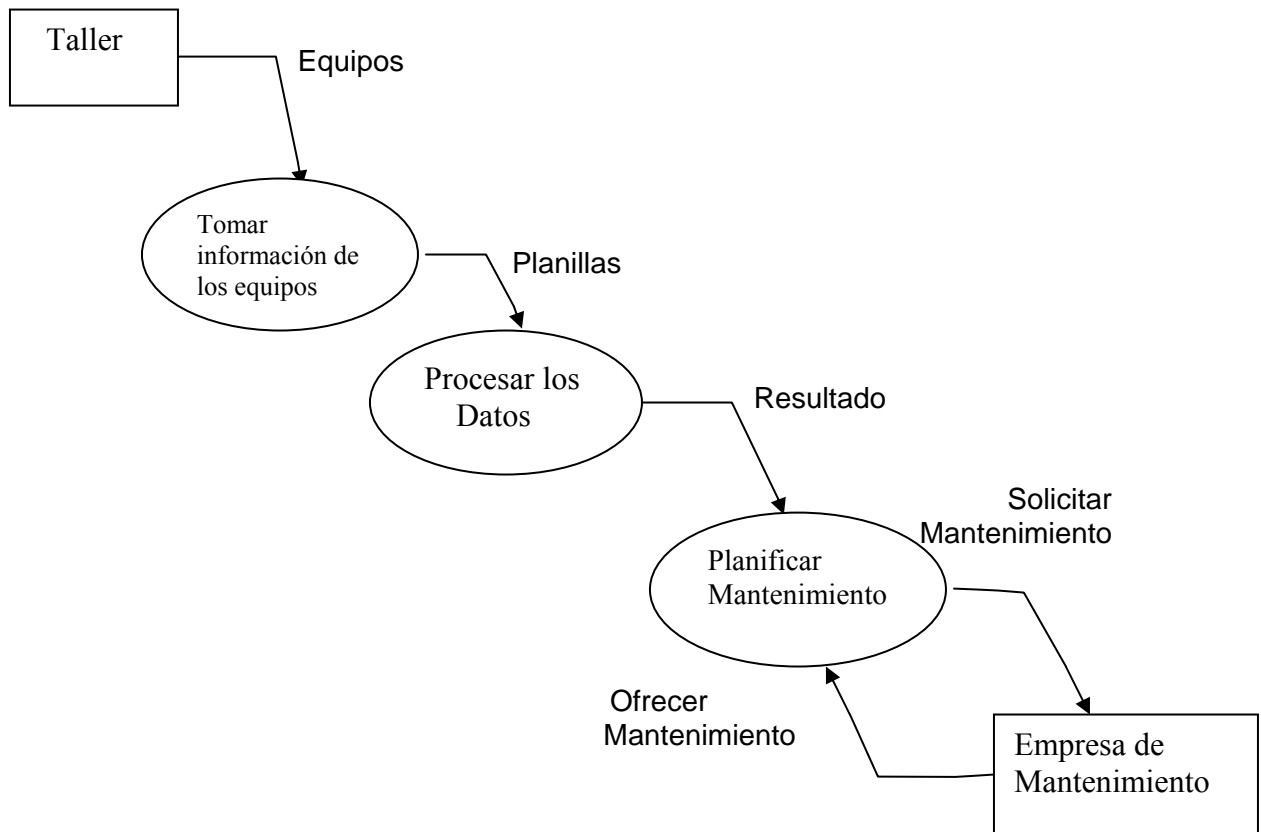
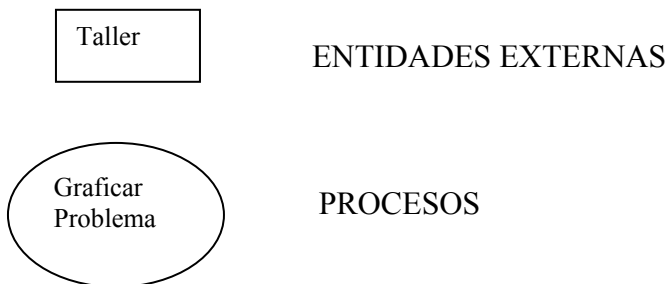


Fig. 2.8 Diagrama de Flujo de Datos



PROCESOS: burbuja, función, transformación. Muestra una parte del sistema que transforma entradas en salidas. Se describe por una frase o sentencia simple que representa qué hace el proceso (verbo - objeto)

ENTIDADES EXTERNAS: usados para modelar entidades externas con las que el sistema se comunica (personas, grupos de personas, organizaciones, otros sistemas).

Están fuera del sistema que va a ser modelado. Están fuera del dominio de cambio de los diseñadores, analistas, desarrolladores. Las relaciones que puedan existir entre ellas no se representan en el DFD.

Refinar los procesos *Tomar Información de los Equipos* y *Procesar los Datos*:

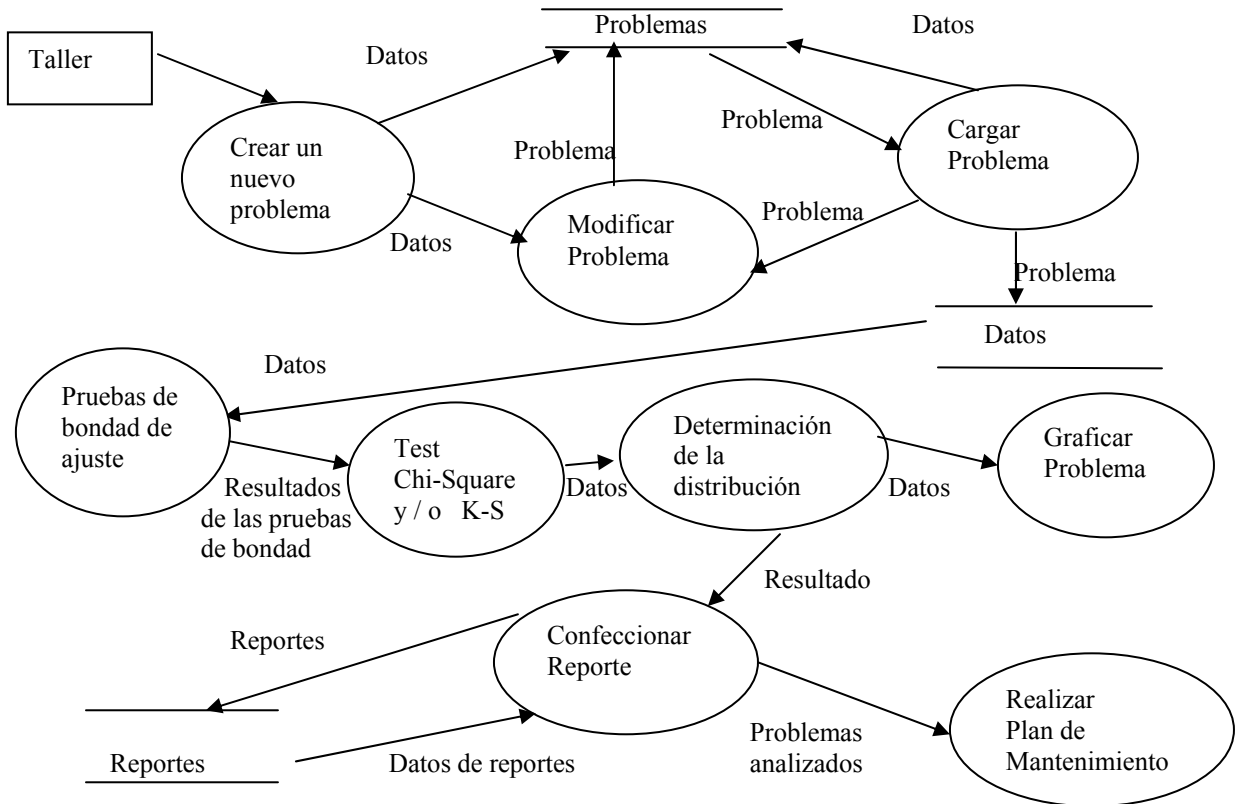


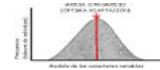
Fig. 2.9 Procesos refinados.

Problemas	ALMACENES DE DATOS
-----------	--------------------

ALMACENES DE DATOS: usados para modelar colecciones de datos que no se encuentran en movimiento. El nombre que se elige para representar el almacén es el plural de los paquetes de datos que entran o salen de él. El almacén no necesariamente expresa datos computarizados.

2.4 Diagrama de la estructura informativa (entidades).

El diagrama de la estructura informativa está conformado por el diagrama de la base de datos **Anexo 2.2** y un diccionario de datos. Los diccionarios de datos son listados organizados de todos los elementos de datos del sistema con definiciones rigurosas y



precisas para que tanto el cliente como el analista tengan un entendimiento común de las entradas, salidas almacenes de datos y cálculos intermedios del sistema¹⁶. Para el SEF el diccionario utilizado se muestra a continuación:

MainTable: {Problem_Name_id + Information + Example}

Sintaxis: Esta tabla es abstracta y de ella van a 'heredar' las tablas que corresponden a los problemas con distribución específica.

Pareto: {Example_id + Distribution (FK)_id + Alfa + X0}

Sintaxis: Corresponde a los problemas con distribución de Pareto.

NormalTruncate: {Example_id+ Distribution (FK)_id + Moda + Media + Varianza}

Sintaxis: Corresponde a la distribución Normal Truncate.

Ganma: {Example_id + Distribution (FK)_id + Value K + Value Lamda}

Sintaxis: Distribución Ganma.

Uniform: {Example_id + Distribution (FK)_id + Value A + Value B}

Sintaxis: Distribución Uniforme.

Triangular: {Example_id + Distribution (FK)_id + Value Min + Value Max+ Moda}

Sintaxis: Distribución Triangular.

Exponential: {Example_id + Distribution (FK)_id + Lamda}

Sintaxis: Distribución Exponencial.

Weibull: {Example_id + Distribution (FK)_id + Alfa + Beta}

Sintaxis: Distribución de Weibull.

Log-Normal: {Example_id + Distribution (FK)_id + Moda + Media + Varianza}

Sintaxis: Distribución Logaritmo Normal.

Rayleigh: {Example_id + Distribution (FK)_id + Value Cit }

Sintaxis: Distribución de Rayleigh.

Normal: {Example_id + Distribution (FK)_id + Media + Varianza}

Sintaxis: Distribución Normal.

Erlang: {Example_id + Distribution (FK)_id + Value K + Value Cita}

Sintaxis: Distribución de Erlang.

¹⁶ Pressman. Ingeniería del software: un enfoque práctico. 4ta edición. Capítulo 12, pág 207

Values: {Variable_id + Row_id + Problem_Name (FK)_id + Value}

Sintaxis: Esta tabla contiene los valores de las variables de cada problema.

Variables: {Var_Name_id + Problem_Name (FK)_id + Column}

Sintaxis: Aquí se guardan las variables de cada problema.

2.5 Diagrama de clases.

En este diagrama se relacionan algunas de las clases que se utilizan en el sistema. Para la notación se ha utilizado la UML (Grady Booch), con el nombre de la clase, los atributos de esa clase y los métodos que implementa. En particular se puede observar que el sistema, debido a que posee una amplia interfaz de intercambio con el usuario, tiene como clases fundamentales Main y ChildWin que son clases que representan la ventana principal de la aplicación y cada una de las ventanas hijas, hojas de cálculo que equivalen a un problema diferente. Alrededor de estas dos clases fundamentales se implementan otras que son necesarias para otras operaciones como la aplicación de las pruebas de bondad del ajuste, UFormTest, la generación manual de valores USimulationForm o la generación automática de valores UAutoSimulation. La clase UClassified se encarga de clasificar inicialmente los problemas según su tipo, de planificación de mantenimientos, de determinación de garantía de producción o de determinación de supervivencia de individuos de una población. Esta clase además guarda la unidad de medida con que se trabajará en ese problema.

A pesar de que la interfaz visual tiene gran importancia en un sistema de este tipo, el módulo matemático se convierte en el centro de todo el problema. La clase principal del Sistema Estadístico para el estudio de la Fiabilidad es UMathematic donde se implementan los métodos para aplicar las pruebas de bondad del ajuste y para el cálculo de los parámetros y de los intervalos de confianza.

Alrededor de esta clase se necesitaron otro grupo de ellas que cumplen con determinadas funciones específicas. La clase TCDF se encarga de calcular los valores de la función de probabilidad acumulada para la distribución y los parámetros calculados. La clase TPDF de forma similar pero calcula los valores de la función de probabilidad.

Las clases TRandom, TVector y TSampleVector se utilizan en la generación de números unidas a las clases de la interfaz visual USimulationForm y UAutoSimulation.

Además existe una conexión con una base de datos en la que se almacenan problemas y ejemplos de muestras. Para establecer esta conexión y manipular la base de datos con el SEF se utilizan las clases que se muestran en el diagrama. UDataBase que contiene entre sus campos el camino de la base de datos y los tipos de consultas que se realizarán sobre la base de datos. UOpen y UDBSave se van a encargar de abrir un problema o un ejemplo desde la base de datos o guardarlo hacia ella respectivamente. El diagrama de clases en notación UML se puede ver en el **Anexo 2.3**.

2.6 Diseño de ventanas.

La interfaz visual, en todo sistema que intercambie información directamente con un humano, tiene mucho valor. Elementos aparentemente sin importancia como los colores que se utilizan, el diseño o las palabras, pueden determinar en el triunfo de un producto sobre otro. Debido a que los fines iniciales del Sistema Estadístico para el estudio de la Fiabilidad son los docentes, se han utilizado ventanas similares a los softwares que se utilizan durante el pregrado en casi todas las carreras que reciben estadísticas como una asignatura. Esta propiedad facilita su utilización. Todas las ventanas se han diseñado utilizando el idioma español y con los colores clásicos del sistema operativo para el que se ha generado.

El sistema comienza con una ventana principal al tamaño de la pantalla, a partir de la cual se puede abrir un problema desde un fichero, desde la base de datos, borrar un problema de la base de datos, ver un ejemplo o crear un nuevo problema con los diálogos clásicos de apertura y otros propios de la aplicación.

Imágenes de las ventanas utilizadas en la aplicación se pueden observar a partir del **Anexo 4.1** y hasta el **Anexo 4.9**.

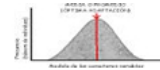
2.6.1 Tabla de Eventos.

En la tabla 2.1 se resumen los eventos que ocurren al ejecutar algunas de las acciones que se pueden realizar a partir de la ventana principal del sistema.

Ventana	¿Qué hace el actor?	¿Qué hace el sistema?	Función
Mostrar Ejemplos	Oprimir en el menú Ayuda, el submenú Ejemplos	El sistema le muestra al actor una nueva ventana donde se puede seleccionar un ejemplo de la lista de ejemplos del sistema para estudiar su distribución.	Brindar información al usuario sobre determinados tipos de problemas.
Salvar en la Base de Datos	En el menú Base de datos, oprimir el submenú Guardar hacia base de datos	Se le muestra al usuario una ventana donde podrá nombrar al problema e incluso asociarle una descripción antes de salvarlo hacia la Base de Datos.	Salvar en la Base de Datos un nuevo problema o uno existente con modificaciones.
Nuevo Problema	Oprimir Nuevo...	Aparece una ventana 'vacía' para que el usuario pueda introducir los datos de un nuevo problema.	Procesar Datos

Tabla 2.1 Tabla de eventos.

Con estas tres acciones, se resumen en la tabla de eventos lo que hace el actor, el sistema, qué ventana se muestra y qué función cumple esta. Similar a como se muestra en la tabla ocurre con todas las demás acciones que se realizan en el SEF, todas están acompañadas de una o varias operaciones en el sistema y la aparición de



nuevas ventanas que solicitan o muestran datos en un constante flujo de información con el usuario.

2.7 Conclusiones parciales.

Con la culminación del segundo capítulo se le presenta al usuario la estructura computacional que tiene el Sistema Estadístico para el estudio de la Fiabilidad. Se han planteado los casos de uso de cada uno de los actores del sistema, las clases fundamentales, la estructura informativa, entre otros aspectos que abordan la estructura interna del SEF. Es de vital importancia conocer los aspectos que aquí se relacionan antes de aventurarse en una nueva versión del software o incluso a darle mantenimiento.



3. Capítulo III: FORMULACIÓN MATEMÁTICA, ESTADÍSTICA Y FIABILÍSTICA.

Todo este tercer capítulo intenta esbozar los métodos matemáticos y en especial estadísticos que se han utilizado para la realización del Sistema Estadístico para el estudio de la Fiabilidad. En cinco acápites se presentan las ecuaciones de las distribuciones utilizadas, los algoritmos de las pruebas de bondad del ajuste así como otras ecuaciones que se necesitaron. La notación matemática utilizada ha sido fiel a la bibliografía intentando mantener consistencia en los términos utilizados en cada una de las ecuaciones o funciones. En este capítulo se hará referencia a varios sitios de Internet, libros y formularios que fueron consultados, las referencias se muestran en el pie de página y no coinciden, en el orden, con la bibliografía que se encuentra ordenada alfabéticamente.

3.1 Test de bondad del ajuste.

En el Sistema Estadístico para el estudio de la Fiabilidad se pueden aplicar dos test matemáticos para demostrar la bondad del ajuste de la muestra a una distribución con parámetros estimados. Los que se han utilizado en este caso por la exactitud de sus resultados son el test de Chi cuadrado y el test de Kolmogorov-Smirnov. Ambos permiten demostrar la bondad del ajuste a las distribuciones normal, exponencial y de Weibull que son las usadas en este software.

3.1.1 Distribuciones utilizadas. Funciones de probabilidad y probabilidad acumulada.

Ya se mencionó que en el Sistema Estadístico para el estudio de la Fiabilidad (SEF) utilizan tres distribuciones a las cuales se pueden ajustar variables continuas, como lo es en todo los casos la variable tiempo.

Por el momento han sido seleccionadas para el trabajo de la fiabilidad las distribuciones normal, exponencial y de weibull. Esto no quiere decir que en una próxima versión del SEF no se puedan incluir otras distribuciones que también se

trabajan en la teoría de la fiabilidad como son las distribuciones de pareto, erlang, log-normal, rayleigh, entre otras.

La distribución normal ha sido seleccionada debido a que en el caso de los equipos mecánicos, en varias ocasiones, los datos de tiempos de vida entre fallos recogidos se ajustan a una distribución de este tipo. Además de lo fácil que resulta el cálculo de los intervalos de confianza en esta distribución y la estimación de los parámetros. La distribución normal es trabajada por casi todos los softwares que existen dedicados a esta temática, fundamentalmente debido a la cantidad de procesos de la vida real que se pueden predecir porque se ajustan a ella. Las funciones de densidad:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

y de probabilidad acumulada o de distribución:

$$F(x) = \int_{-\infty}^x f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} dx$$

también tienen una ecuación sencilla en el caso de esta distribución por lo que la hacen ideal para una primera versión.

Otra de las seleccionadas para el trabajo es la distribución exponencial. En este caso la mayoría de los componentes electrónicos de un equipo son productos no reparables e históricamente se ha demostrado que siguen una distribución exponencial. Esta razón y que es en la que más fácil se pueden estimar los parámetros también influyeron en que se seleccionara. Para la distribución exponencial la función de densidad está dada por la ecuación:

$$f(x) = \lambda e^{-\lambda x}$$

y la de probabilidad acumulada:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Por último fue seleccionada la distribución de weibull debido a que esta curva tiene la característica de que cualquier muestra de tiempos se ajusta a ella y que además con

diferentes valores de sus parámetros se comporta de igual forma que la normal y la exponencial. La complejidad del cálculo de los intervalos de confianza y de la estimación de los parámetros de esta distribución es mucho mayor, sin embargo el aporte desde el punto de vista de la teoría de la fiabilidad que esta distribución brinda nos hizo seleccionarla. En este caso la función de densidad es:

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

Y la de probabilidad acumulada:

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha} \right)^{\beta}} \quad \text{si } t \geq 0, \text{ como es el caso.}$$

En el **Anexo 3.1** se muestran imágenes de las curvas $F(t)$ y $f(t)$ para cada una de las distribuciones aquí mencionadas.

3.1.2 Test Chi cuadrado.

Esta prueba de bondad del ajuste utiliza como base la distribución Chi cuadrado de ahí su nombre y consiste, de manera general, en dividir la muestra en intervalos o clases y comparar las diferencias entre las frecuencias esperadas y las observadas, en cada clase, construyendo el estadígrafo Chi cuadrado mediante la siguiente ecuación:

$$\chi^2 = \sum_{i=1}^k (o_i - e_i)^2 / e_i, \text{ donde } o_i \text{ son las frecuencias observadas y } e_i \text{ las esperadas.}$$

Para el trabajo en el SEF se llevó a un lenguaje de alto nivel el algoritmo matemático que a continuación se muestra:

Paso 1: Construcción de una distribución empírica de frecuencia, para la determinación de las frecuencias de clases.

Paso 2: Sustitución de los parámetros desconocidos por sus estimaciones máximas verosímiles y determinación de la probabilidad teórica de cada intervalo de clases.

Paso 3: Determinación de la frecuencia esperada de cada clase.

Paso 4: Cálculo del estadígrafo χ^2 , según la ecuación antes mencionada.

Paso 5: La decisión se tomará basándose en las siguientes dójimas.¹⁷

¹⁷ Cué, J. (1987). Estadística. Primera Parte. Dócima para la bondad del ajuste. Universidad de la Habana. Facultad de Matemática Cibernética. Ciudad de la Habana.

$$\varphi(o_i) = \begin{cases} 0 & \text{si } x^2 \leq x^2_{1-\alpha}(k-1-s) \\ 1 & \text{en otro caso} \end{cases}$$

donde s es el número de parámetros que fueron sustituidos por sus estimadores máximos verosímiles, x^2 corresponde al valor de la función Chi cuadrado y α corresponde al nivel de significación permitido.

3.1.3 Test de Kolmogorov Smirnov.

Es otro de los que se utilizan para demostrar la bondad del ajuste a una distribución. Tiene una gran ventaja sobre la prueba de Chi cuadrado y es que no es necesario clasificar los datos en intervalos sino que se tratan uno a uno y cada dato equivale a un intervalo, por lo que no hay pérdida de información en la clasificación como si ocurre en la prueba anterior. Para el SEF fue llevado a un lenguaje de alto nivel el algoritmo matemático que a continuación se muestra:

Paso 1: Se estiman los parámetros a partir de la muestra.

Paso 2: Se calcula el estadígrafo S_n . (epígrafe 3.1.3.1)

Paso 3: Se generan los valores de la función de probabilidad acumulada con los parámetros estimados para los valores de la muestra.

Paso 4: Se calcula la máxima diferencia entre el estadígrafo S_n y la función de probabilidad acumulada.

Paso 5: Se compara la máxima diferencia encontrada con el valor de D_n con $1-k$ grados de libertad. Donde k es el nivel de significación permitido y D_n son los valores críticos para la prueba de la bondad del ajuste de la muestra de tamaño n . Se puede aceptar el ajuste si la máxima diferencia encontrada es menor o igual que el valor crítico D_n .¹⁸

3.1.3.1 Estadígrafo utilizado.

El estadígrafo utilizado para la prueba de bondad del ajuste de Kolmogorov-Smirnov ya se había presentado como S_n . La ecuación que describe a este estadígrafo es:

$$S_n = \begin{cases} 0 & \text{si } x < x'_1 \\ k/n & \text{si } x'_k \leq x < x'_{k+1} \\ 1 & \text{si } x \geq x'_n \end{cases} \quad k = 1, 2, \dots, n-1$$

Donde k es la posición del valor x , n es el tamaño de la muestra y x'_1, x'_2, \dots, x'_n son los valores x_1, x_2, \dots, x_n ordenados ascendentemente.

La función definida de esta forma cumple propiedades interesantes y una de las más importantes es que a medida que n crece, $S_n(x)$ se aproxima cada vez más a la función $F(x)$ que sigue la muestra.

3.2 Estimación de parámetros.

En el cálculo de la bondad del ajuste (epígrafes 3.1.2 y 3.1.3) se necesitan los valores numéricos de los parámetros de las distribuciones a las cuales se va a demostrar el ajuste. En otros softwares, se le pregunta al usuario que parámetros desea utilizar, en el SEF se estiman a partir de la muestra.

Para la tres distribuciones que se han trabajado en esta versión se estiman lo parámetros alfa (α), beta (β), desviación estándar (σ), varianza (σ^2), media (μ) y lambda (λ) según sea la distribución.

3.2.1 Estimación de la media.

Este parámetro es estimado en las tres distribuciones referidas, normal, exponencial y de weibull; aunque tiene diferentes formas de estimación.

De esta forma para la distribución normal la media que se utiliza es la media aritmética, calculada a partir de la muestra sumando todos los valores de tiempo y dividiéndolo entre la cantidad de valores, lo que da lugar a la siguiente ecuación¹⁹:

¹⁸ Cué, J. (1987). Estadística. Segunda Parte. Prueba de Kolmogorov-Smirnov para la bondad del ajuste. Universidad de la Habana. Facultad de Matemática Cibernética. Ciudad de la Habana.

¹⁹ Cué, J. (1987). Estadística. Primera Parte. La media aritmética. Universidad de la Habana. Facultad de Matemática Cibernética. Ciudad de la Habana.

$$\bar{m} = \bar{\mu} = \frac{\sum_{i=1}^n t_i}{n}$$

donde \bar{m} sería la media poblacional y $\bar{\mu}$ la media muestral, pero como ya se había dicho que para el SEF se toma como media poblacional la media muestral, partiendo de que la muestra tomada sea representativa de la población, se dice que $\bar{m} = \bar{\mu}$.

Mediante esta misma ecuación se calcula la media para la distribución exponencial, tomando la media aritmética como media muestral.

Sin embargo para la estimación de la media en la distribución de weibull se utiliza una ecuación²⁰ un tanto más compleja, basada en los valores de la distribución gamma y de los parámetros alfa y beta (epígrafe 3.2.4).

$$\bar{m} = \bar{\mu} = \alpha \Gamma \left(1 + \frac{1}{\beta} \right) \text{ donde } \Gamma(x) = \frac{x^{x-1}}{e^x} \left(\frac{0.2119184}{x^{1.00357}} + 2.5066283 \right)$$

Para la función gamma ha sido utilizada esta ecuación partiendo de la teoría de que cumple con los requisitos de exactitud para la teoría de la fiabilidad y es más sencilla de evaluar computacionalmente.

3.2.2 Estimación de la varianza y la desviación estándar.

De igual forma que la media, la varianza de estas distribuciones tiene su forma independiente de cálculo. Para la distribución normal ha sido utilizada la ecuación²¹:

$$V(x) = \sigma^2 = \frac{1}{n} \left(\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right)$$

De la cual aparece la demostración en la bibliografía referenciada. La desviación estándar es calculada a partir de la varianza, como su raíz cuadrada $\sigma = \sqrt{V(x)}$.

Para la distribución exponencial no ha sido necesario estimar la varianza ni la desviación estándar sino el parámetro lambda (epígrafe 2.2.3).

²⁰ Reyes T, 2003, Formulario sobre las distribuciones utilizadas en la teoría de la fiabilidad. Sin publicar. Facultad de Ingeniería Mecánica, UCLV.

²¹ Cué, J. (1987). Estadística. Primera Parte. La varianza. Universidad de la Habana. Facultad de Matemática Cibernética. Ciudad de la Habana.

En la distribución de weibull la varianza está dada por²²:

$$V(\mathbf{x}) = \alpha^2 \left[\Gamma \left(1 + \frac{2}{\beta} \right) - \Gamma^2 \left(1 + \frac{1}{\beta} \right) \right]$$

nuevamente utilizando la función gamma.

3.2.3 Estimación del parámetro lambda.

El parámetro lambda (λ) estimado para las distribuciones exponencial y de weibull tiene dos ecuaciones para el cálculo respectivo. En el caso de la distribución exponencial este parámetro es quien decide que forma van a seguir las funciones de densidad y de probabilidad acumulada y está dado por el inverso de la media muestral

$$\lambda = \frac{1}{\mu} \text{ donde } \mu \text{ es la media muestral}$$

En la distribución de weibull la ecuación que define el parámetro lambda depende de los estimadores de alfa y beta:

$$\lambda = \alpha^{-\beta} = \frac{1}{\alpha^{\beta}}$$

3.2.4 Estimación de los parámetros alfa y beta.

Estos parámetros son característicos de la distribución de weibull, en ella definen el resto de los parámetros, la forma que tendrán las funciones de densidad y probabilidad acumulada y los valores de los intervalos de confianza (epígrafe 3.3).

El parámetro alfa es calculado a partir de beta y de los valores de la muestra mediante una ecuación sencilla²⁰:

$$\alpha = \left(\frac{\sum_{i=1}^n t_i^{\beta}}{n} \right)^{\frac{1}{\beta}}$$

²² BioMates, 2005, Distribuciones de probabilidad. *Distribución de Weibull*. [En línea], España, disponible en: http://www.telefonica.net/web2/biomates/tabl/tabl_weibull.htm [Accedido el 20/09/2005]

Esto nos hace pensar que la mayor complejidad está en la estimación del parámetro beta y de hecho este debe ser estimado por un método de máxima verisimilitud escogiendo el valor de beta que “mejor” cumpla con la ecuación²⁰:

$$\left(\frac{n}{\beta} + \sum_{i=1}^n \ln t_i \right) \sum_{i=1}^n t_i^{\beta} - n \sum_{i=1}^n (t_i^{\beta} \ln t_i) = 0$$

3.3 Cálculo de los intervalos de confianza.

Una de las potencialidades del SEF es que permite calcular los intervalos de confianza para cualquier muestra y todas las distribuciones que en él se trabajan, además se muestran en el gráfico de la función de probabilidad acumulada alrededor de la media muestral. Cada distribución tiene una forma particular de calcular estos intervalos de confianza, basándose siempre en los parámetros estimados anteriormente. Es por ello que aquí se presentan tres epígrafes, cada uno referente a una distribución.

3.3.1 Intervalos de confianza en la distribución normal.

Después de haber estimado la media y la desviación típica se pueden calcular los intervalos de confianza para esta distribución. En el dominio de problemas que el SEF puede resolver no se incluyen los que poseen muestras infinitas, esto quiere decir que siempre se va a trabajar con una muestra finita de datos por lo que se puede asegurar que los intervalos de confianza están dados por una simple ecuación²³:

$$\text{inferior} = \mu - z_c \frac{\sigma}{\sqrt{n}} \quad \text{superior} = \mu + z_c \frac{\sigma}{\sqrt{n}}$$

donde z_c son los valores críticos, es decir, el valor obtenido en una distribución normal $N(0, 1)$ para una probabilidad de $\{\alpha/2\}$ grados de libertad, para c nivel de confianza.

En ocasiones se comete el error de generalizar el método de cálculo de los intervalos de confianza de la distribución normal a todas las distribuciones, esto se debe

²³ Cué, J. (1987). Estadística. Primera Parte. Intervalo de confianza para la media. Página 152
Universidad de la Habana. Facultad de Matemática Cibernética. Ciudad de la Habana.

fundamentalmente a que no existe una teoría general para el cálculo en todas las distribuciones y por otro lado que es mucho más sencillo que en el resto.

3.3.2 Intervalos de confianza en la distribución exponencial.

Ya se había mencionado anteriormente que la distribución exponencial generalmente logra buenos ajustes a muestra de tiempo de funcionamiento sin fallos o entre fallos a los componentes electrónicos o equipos de este tipo. Sin embargo puede que otros con otras características también se ajusten y entonces en la estadística se establecen diferencias entre los que no son reparables y lo que si. Sobre esto existe toda una teoría que podrá ser abordada por el SEF en próximas versiones, aunque ahora se están analizando muestras que equivalen a un período de tiempo, lo que quiere decir que no se ha fijado un número de fallos sino un intervalo de tiempo para recoger las muestras. Esto nos permite generalizar todos los problemas que resuelve el SEF. Para muestras sin reemplazo, que es el caso que se está analizando en este trabajo los intervalos de confianza para la distribución exponencial se muestran a continuación²⁰:

$$\frac{2n\mu}{\chi^2_{(1-\frac{\alpha}{2}; 2n)}} < \mu < \frac{2n\mu}{\chi^2_{(\frac{\alpha}{2}; 2n)}} \text{ donde } \chi^2 \text{ es el valor de chi cuadrado}$$

3.3.2.1 Generalización para muestras con o sin reemplazo.

En la fiabilidad para la distribución exponencial se diferencian los métodos según sea el tipo de problema para el que se recogió la muestra. Ya se había hecho alusión a esto en el epígrafe anterior. Un ejemplo de esta diferencia es en el cálculo de los intervalos de confianza en los que para muestreo con reemplazo, o sea, para equipos no reparables que son repuestos por otros, el intervalo inferior es calculado evaluando la función Chi cuadrado en $2n+2$ y no en $2n$ como ocurre en las muestras tomadas del muestreo sin reemplazo. Esto es fundamentado en estudios realizados con varios ejemplos los que han abordado que los fallos que no se tienen en cuenta después de la última muestra tomada en los equipos no reparables equivale a evaluar la función Chi cuadrado en el doble del tamaño de la muestra más dos.

En el caso de los estudios que se pueden realizar con el SEF lo que se ha fijado es el intervalo de tiempo y no la cantidad de fallos por lo que se generalizan los problemas a equipos reparables, o sea muestreo sin reemplazo.

3.3.3 Intervalos de confianza en la distribución de weibull.

Aún cuando la distribución de weibull es la más potente en la teoría de la fiabilidad para realizar estudios estadísticos a muestras de tiempos de funcionamientos entre o hasta el fallo, no existen muchos documentos donde se refleje el cálculo de los intervalos de confianza para esta distribución. En la investigación no se encontró ningún software capaz de calcular estos intervalos. A partir de la bibliografía consultada se llevaron a un lenguaje de alto nivel las ecuaciones para encontrar estos intervalos. De estos estudios se obtuvo.

Límite de confianza inferior ²⁰:

$$t_{ib} = \alpha_{ib} \min \left(\Gamma \left(1 + \frac{1}{\beta_{ib}} \right); \Gamma \left(1 + \frac{1}{\beta_{sb}} \right) \right)$$

Límite de confianza superior ²⁰:

$$t_{ib} = \alpha_{sb} \max \left(\Gamma \left(1 + \frac{1}{\beta_{ib}} \right); \Gamma \left(1 + \frac{1}{\beta_{sb}} \right) \right)$$

Donde

$$\alpha_{ib} = \lambda_{sb}^{-\frac{1}{\beta}} = \frac{1}{\lambda_{sb}^{\frac{1}{\beta}}}$$

$$\alpha_{sb} = \lambda_{ib}^{-\frac{1}{\beta}} = \frac{1}{\lambda_{ib}^{\frac{1}{\beta}}}$$

$$\lambda_{ib} = \lambda - U \left(1 - \frac{\alpha}{2} \right) \sigma (\lambda)$$

$$\lambda_{sb} = \lambda + U \left(1 - \frac{\alpha}{2} \right) \sigma (\lambda)$$

$$\beta_{ib} = \beta - U \left(1 - \frac{\alpha}{2} \right) \sigma (\beta)$$

$$\beta_{sb} = \beta + U \left(1 - \frac{\alpha}{2} \right) \sigma (\beta)$$

$$\sigma_{(\lambda)}^2 = \frac{\frac{n}{\beta^2} + \lambda \sum_{i=1}^n (t_i^\beta \ln t_i)}{\frac{n}{\lambda^2} \left(\frac{n}{\beta^2} + \lambda \sum_{i=1}^n (t_i^\beta \ln^2 t_i) \right) - \left(\sum_{i=1}^n (t_i^\beta \ln t_i) \right)^2}$$

$$\sigma_{(\beta)}^2 = \frac{\frac{n}{\lambda^2}}{\frac{n}{\lambda^2} \left(\frac{n}{\beta^2} + \lambda \sum_{i=1}^n (t_i^\beta \ln^2 t_i) \right) - \left(\sum_{i=1}^n (t_i^\beta \ln t_i) \right)^2}$$

$U(x)$ es el percentil de la función normal normalizada que para los niveles de fiabilidad que se utilizan en la primera versión del SEF se muestran en la tabla 3.1:

x	U(x)
0.95	1.645
0.90	1.282
0.85	1.0360

Tabla 3.1 Percentiles de la función normal normalizada.

Si se incluyen nuevos niveles de fiabilidad sería necesario incluir los nuevos valores del percentil de la función normal normalizada sin embargo para una nueva versión estos valores pueden ser calculados.

3.4 Generación de tiempos para la muestra.

Para la mayoría de los estudios estadísticos que se realizan tanto manualmente como asistido por herramientas computacionales la mayor limitante es conocer hasta qué punto la muestra tomada es representativa de la población o si el tamaño de la muestra es suficiente para los cálculos que se realizan. Así por ejemplo, se plantea que para que la prueba de Chi cuadrado tenga buenos resultados es necesario no solo que las frecuencias esperadas de los intervalos o clases sea superior 5 sino que el tamaño de la muestra sea superior a cien. La prueba de Kolmogorov Smirnov no pone tantos requisitos, sin embargo el estadígrafo utilizado en el SEF para comparar, S_n (epígrafe 3.1.3.1) la propiedad fundamental que lo caracteriza es que a medida que n crece se acerca más a la función $F(x)$ a la cual se quiere demostrar el ajuste. Por esta

razón en ocasiones es necesario correr el riesgo de utilizar métodos de simulación para generar un grupo de valores que se acerquen a lo que sucedería si se continúa el estudio, de lo contrario se requiere de mucho tiempo más. En el trabajo el usuario debe ser capaz de decidir cuando puede usar un método de simulación para generar más valores.

Los valores se generan usando una semilla que siempre es diferente. Esta se calcula a partir de la fecha actual de la siguiente forma:

Resultado = MilliSecondOf (DateTime) +
SecondOf (DateTime) * 1000 +
MinuteOf (DateTime) * 60000 +
HourOf (DateTime) * 360000 +
DayOf (DateTime) * 8640000 +
MonthOf (DateTime) * 103680000

Aquí la función DateTime obtiene la fecha y el tiempo que tiene el sistema en el momento en que se invoca la función hasta el nivel de milisegundo. Luego se espera un milisegundo en el software para garantizar que nunca se repita la misma semilla.

3.4.1 Generación asistida por el usuario.

Para esta opción del SEF el usuario debe conocer qué distribución van a seguir los datos que él desea sean generados. Para ella se usa una implementación de la generación de números aleatorios con los parámetros calculados a partir de la pequeña muestra que ya ha sido dada. En caso de que se deseen generar valores sin una muestra inicial se van a crear con valores por defecto, por ejemplo $\mu = 0$ $\sigma^2 = 1$. Los números son generados usando una biblioteca basada en el resumen "Computer Generation of Statistical Distributions" de Richard Saucier - ARMY RESEARCH LABORATORY - March 2000.

La opción se puede ejecutar por el menú en la opción Simulación \ Simulación Manual. Donde aparecerá un diálogo pidiendo los datos necesarios para la generación.

3.4.2 Generación empírica.

El uso de la generación empírica en el SEF no debe ser motivo de abuso por parte del usuario y solo debe ser utilizada en caso realmente necesario y que se esté seguro de que los datos generados no contaminan la muestra. Para esta opción se calculan los parámetros de la muestra, necesarios para poder ejecutarla y a partir de ella se generan empíricamente los próximos valores de tiempo según el comportamiento que han tenido los intervalos hasta el momento.

Para el cálculo computacional ha sido utilizado el mismo resumen que para la generación manual. La diferencia entre la generación manual y la asistida por el usuario es que, en este caso, no es necesario que el usuario conozca la distribución que siguen los datos sino que se toma como vector inicial la muestra que ya existe.

La opción se puede ejecutar por el menú en la opción Simulación \ Autosimulación. Donde aparecerá un diálogo pidiendo los datos necesarios para la generación.

3.5 Conclusiones parciales

Cerrando este tercer capítulo, dedicado a los algoritmos, la ecuaciones y los métodos matemáticos utilizados en el Sistema Estadístico para el estudio de la Fiabilidad, se puede resumir que existe un gran campo de aplicaciones de la matemática y de la estadística en la fiabilidad que pueden ser muy usados en la ingeniería en general. Por otro lado se ha llegado a un grupo de resultados basados en estudios realizados a equipos en funcionamiento y en propiedades de las distribuciones abordadas. Puntualmente se hace énfasis, en los métodos expuestos aquí para determinar los intervalos de confianza según la distribución de mejor ajuste y la forma en que han sido llevados a la computación usando algoritmos avanzados de aproximación.

Capítulo IV: FUNCIONALIDADES DEL SISTEMA ESTADÍSTICO PARA EL ESTUDIO DE LA FIABILIDAD.



Sistema Estadístico para el estudio de la Fiabilidad

4. Capítulo IV: FUNCIONALIDADES DEL SISTEMA ESTADÍSTICO PARA EL ESTUDIO DE LA FIABILIDAD.

Este capítulo presenta un compendio de las acciones que se deben realizar para completar un estudio estadístico con el SEF y cómo observar y almacenar los resultados obtenidos. Con la conclusión del mismo se puede llevar a cabo cualquier estudio de fiabilidad con el software. Dividido en once acápites presenta además las características y los requerimientos del sistema.

4.1 Abrir un problema.

Por lo general en todas las aplicaciones que intercambian información con el usuario aparece esta opción para permitirle escoger los datos con lo que va a trabajar. En el caso del SEF puede comenzar un estudio de cualquiera de las tres formas que aparecen en los siguientes acápites.

4.1.1 Desde un fichero.

Para abrir un fichero en el SEF puede hacerlo de dos formas. Similar a como sucede en otros sistemas posee un menú *archivo* en el que aparece la opción *abrir* que al ser ejecutada muestra un diálogo de apertura para que el usuario escoja el fichero y la extensión. También existe una barra de herramientas en la que se muestra un botón *abrir* que realiza la misma función que la opción anterior.

En el Sistema Estadístico para el estudio de la Fiabilidad se utilizan dos tipos de ficheros para almacenar la información. Los ficheros textos con extensión (*.txt) guardan la información en claro y puede ser vista con cualquier otro software capaz de leer textos. La estructura en que se guardan los datos es sencilla, aparecen en la cabeza de cada columna el nombre de la variable y a continuación, hacia abajo, los valores de esa variable. Un fichero creado con esta estructura en cualquier otro software puede ser utilizado por el SEF.

La información, además, puede ser guardada en forma estructurada si se desea garantizar que los datos no sean variados con otro sistema que no sea el SEF. Para este tipo de fichero se utiliza la extensión (*.sst), particular del sistema. El esquema

que se utiliza es simple y en cada localización se guarda la cadena que contiene el escaque con su respectivo valor de fila y columna. De esta forma los valores de las variables o sus nombres solo pueden ser modificados con el SEF.

Las dos formas de abrir un problema desde un fichero pueden ser vistas en el **Anexo 4.1**.

4.1.2 Desde la base de datos.

De manera similar a como se abre un problema desde un fichero, se puede abrir uno desde la base de datos. El SEF posee una base de datos con una estructura que aparece en el **Anexo 2.2** en la que se guardan un grupo de problemas y ejemplos muestras que se ajustan a varias distribuciones para que le sirva de ayuda al usuario. Si el usuario desea abrir un ejemplo de la base de datos, debe pulsar la opción *ejemplos* en el menú *ayuda*. A continuación le aparece un diálogo con una lista de los ejemplos de distribuciones y un comentario opcional sobre esta muestra.

Si por el contrario desea abrir un problema y no un ejemplo, en el menú *base de datos* debe pulsar el submenú *abrir desde base de datos*, luego aparece un diálogo con la lista de problemas y los comentarios sobre ellos. Para seleccionar un problema debe marcarlo con el cursor y luego presionar *ver*, o simplemente presionar dos veces clic sobre el problema deseado.

Las ventanas para abrir un problema desde la base de datos se muestran en el **Anexo 4.2**.

El menú *base de datos* posee dos opciones más, *borrar problema* y *guardar problema*. Para borrar un problema al seleccionar esta opción le aparece un diálogo similar a si presiona *abrir*, para confirmar que lo desea eliminar debe seleccionar el problema y presionar dos veces el botón izquierdo del mouse sobre él o seleccionarlo y presionar sobre el botón *borrar*. Una vez borrado no puede ser recuperado. La opción *guardar problema* se explica detalladamente en el acápite 4.6.2.

4.1.3 Crear un problema nuevo.

Para crear un nuevo problema y comenzar a recoger las muestras de tiempo, el usuario lo puede hacer de dos formas. En ambos casos se trata de la manera

tradicional en que lo hacen la mayoría de las aplicaciones que intercambian información con el usuario.

En el menú *archivo* del Sistema Estadístico para el estudio de la Fiabilidad aparece como primera opción, *nuevo*; presionándola aparece un diálogo para clasificar el problema y luego se crea una nueva ventana hija en la que pueden ser recogidos los valores de tiempo que forman la muestra estadística.

Otra forma de crear un nuevo problema es pulsando, en la barra de herramientas, el botón *nuevo* señalado, igual que en casi todas las aplicaciones, por una pequeña hoja en blanco. Este botón cumple exactamente las mismas funciones que el submenú *nuevo*. Estas opciones se muestran en el **Anexo 4.3**.

4.2 Entrar datos de la muestra.

Para recoger los datos de una muestra es necesario primeramente crear un nuevo problema, clasificarlo, definir la variable y luego comenzar a recoger los valores de tiempo. Los problemas en el SEF se muestran al usuario como una hoja de cálculo en la que cada columna puede almacenar los valores de tiempo de una variable, todas las variables constituyen un problema y no necesariamente todas las variables tienen la misma cantidad de valores.

Si el usuario intenta entrar un valor que no sea numérico en un escaque de una variable el SEF lo sustituye por un cero. Los valores de tiempo recogidos pueden ser enteros o reales en cuyo caso se separan las cifras enteras de las decimales por un punto.

4.2.1 Clasificar el problema.

El sistema en esta primera versión puede ser utilizado con tres objetivos:

- Planificar o modificar los ciclos de mantenimiento de un equipo.
- Determinar la garantía de producción de un producto.
- Determinar la supervivencia de individuos de una población.

La variable tiempo puede ser trabajada con cinco unidades minutos, horas, días, meses y años. Por esta razón al abrir un problema o crear uno nuevo el SEF pide una clasificación para el problema. Para auxiliar la selección, aparece un texto con una

pequeña información sobre ese tipo de problema al pasar el mouse sobre la clasificación correspondiente. Una vez seleccionado el tipo de problema y la unidad de tiempo en que se va a trabajar pulsando el botón *aceptar* aparece la hoja para comenzar a entrar los valores.

Para trabajar con otra unidad de tiempo que no aparezca en el SEF puede seleccionar cualquiera de las que aparecen y sustituirla en el reporte final. Una vista de la ventana utilizada para clasificar los problemas se muestra en el **Anexo 4.4**.

4.2.2 Definir una variable.

Una variable puede ser definida en cualquier momento durante la entrada de datos de la muestra, sin embargo el estudio estadístico no puede ser ordenado hasta que no se encuentren definidas todas las variables. No se pierden los datos el usuario comienza entrándolos y luego define la variable.

Para definir una variable solo es necesario pulsar el botón derecho del mouse sobre el escaque correspondiente a la fila encabezada con *variable* y la columna donde se colocarán los valores de tiempo, o sea, el encabezado de cada columna. Una vez realizada esta acción aparece una caja de texto en la que se debe teclear el identificador de la variable. Para culminar la definición se puede presionar la tecla enter o presionar el botón izquierdo del mouse fuera de la caja de texto.

Importante resaltar que el nombre de la variable debe ser un identificador, en caso de que se intente utilizar un carácter no válido en el nombre de la variable este será sustituido por un guión bajo. Ver **Anexo 4.5**.

En todo momento las acciones a realizar o los mensajes de SEF se muestran en la barra de estado.

4.3 Ejecutar estudio estadístico.

Esta es la opción con la que se obtienen los datos estadísticos y fiabilísticos de la muestra que se está estudiando. Para comenzar a realizar este estudio se puede hacer de tres formas. A través del menú *estadística* que tiene entre sus opciones una llamada *encontrar distribución*, otra forma es en la barra de herramientas presionando

el botón *encontrar distribución* señalado con un lente, y la tercera simplemente presionando la tecla *F9*.

Al ejecutar una de estas opciones aparece un diálogo (**Anexo 4.6**) en el que se observa, en la parte superior, la lista de variables definidas en el problema, cuatro botones para seleccionar o deseleccionar las variables de una en una o todas de una vez y otra lista, inicialmente vacía, de variable seleccionadas para realizarle el estudio.

En la parte inferior se observan tres bloques, *Test*, *Distribución* y *Significación*. En el bloque *Test* se pueden seleccionar cualquiera de las pruebas de bondad del ajuste que allí aparecen, incluso todas; solo los marcados son las que se aplican a las variables seleccionadas. En el bloque *Distribución* se pueden seleccionar las distribuciones a las que se intentará ajustar la muestra, igualmente pueden ser seleccionadas varias distribuciones a la vez. Por último en el bloque *Significación* se escoge el nivel de confianza o significación con el que se desea que se compruebe el ajuste de la muestra. En este caso solo puede ser marcado un nivel y para esta primera versión aparecen los niveles de 95%, 90% y 85% debido a que son los más utilizados en la teoría de la fiabilidad.

También aparecen los botones *aceptar*, *cancelar* y *defecto* que cumplen con las funciones de ejecutar el estudio, cancelar el estudio o seleccionar los *test*, la distribución y el nivel de significación por defecto respectivamente.

Al completar la acción de ejecutar el estudio estadístico en el SEF se pueden observar los gráficos de funciones y un resumen de los parámetros estadísticos de la muestra con el texto sobre la información de fiabilidad de la muestra.

4.4 Interpretación gráfica de los resultados.

Una de las potencialidades del SEF es la información gráfica que le brinda al usuario sobre la muestra, permitiéndole interactuar con estas gráficas. En uno de ellos se muestran las gráficas de las curvas $F(t) = Q(t)$ conocida como curva de vida, la curva teórica de $F(t)$ para que se pueda comparar con la práctica y $P(t)$ llamada curva de fiabilidad. Ambas son señaladas con colores diferentes y graficadas, por el



eje del tiempo, entre el cero y el mayor valor que posee la muestra y por las ordenadas, entre cero y uno. **Anexo 4.7.**

El otro gráfico mostrado es el de las funciones de densidad de la muestra y teórica. Este gráfico aunque no es interactivo le brinda al usuario la información visual acerca de lo “bueno” que es el ajuste de la muestra a la distribución teórica.

Para ver estos gráficos se puede presionar el botón *gráficos* de la barra de herramientas señalado por una **G** o seleccionar la hoja correspondiente al gráfico en el sistema.

Si se intenta ver cualquiera de los gráficos antes de ordenar el estudio estadístico un mensaje de error será mostrado, indicando que no se han seleccionado variables para graficar y se regresará a la hoja inicial.

En la parte superior derecha del gráfico se muestra una lista de las variables que han sido graficadas. Presionando dos veces el botón izquierdo del mouse sobre otra de las variables que allí aparecen se muestra el gráfico de esta nueva variable.

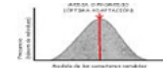
Debajo de esta lista se muestran los botones *intervalo de confianza* y *guardar gráfico* con el primero se pueden visualizar los intervalos de confianza y la media en la curva $P(t)$ y con el segundo se guarda el gráfico hacia un fichero de tipo (*.bmp).

Otra sección que aparece en este gráfico es la de evaluar un valor de tiempo en la gráfica $P(t)$. En la caja de texto se entra el valor de tiempo que se desea evaluar y al presionar el botón *evaluar* se muestra en la gráfica $P(t)$ la probabilidad de funcionamiento sin fallo de ese equipo transcurrido el intervalo de tiempo dado. Con el botón borrar se eliminan todas la evaluaciones que se muestran en la curva y los intervalos de confianza si también se han mostrado.

Por último en la parte inferior derecha se muestra un pequeño gráfico con la distribución espacial de los valores de la muestra, en pequeños puntos negros, como aparecen en la muestra y en puntos grises, ordenados de mayor a menor.

4.4.1 Interpretación de la gráfica $Q(t)$.

Como ya se había planteado anteriormente la curva $Q(t)$ coincide con la curva $F(t)$ en la estadística. Para cada una de las distribuciones que se trabajan en el SEF se muestran las funciones de probabilidad acumulada en el acápite 3.1.1 del capítulo



anterior. A partir del análisis gráfico de las curvas teórica y práctica que se muestran, la primera con una línea fina de color negro y la segunda con una línea de color purple, se puede corroborar el resultado estadístico sobre la bondad del ajuste. **Anexo 4.7.** A medida que coincidan más estas curvas debe ser mejor el ajuste. Sin embargo puede suceder que en un punto las diferencias sean significativas y esto no se observe los gráficos.

Esta curva como toda función de probabilidad acumulada es no decreciente por lo que siempre se va a mostrar como una curva creciente o constante en el eje de las ordenadas.

4.4.2 Interpretación de la gráfica $P(t)$.

La curva $P(t)$ que se muestra en el gráfico con un color azul es conocida como curva de fiabilidad y corresponde a la diferencia $1 - Q(t)$. Sobre ella se calculan los intervalos de confianza, la media y se evalúa cualquier intervalo de tiempo que el usuario desee, después de haber ejecutado el estudio estadístico. Esta función tiene una gran importancia para el estudio de la teoría de la fiabilidad, pues de su interpretación se puede deducir la probabilidad que existe de que un equipo funcione un período de tiempo sin fallar.

Por ejemplo una vez realizado el estudio estadístico a un auto, este refleja que los rodamientos traseros fallaron con una media de 2100 horas de funcionamiento con una probabilidad de 50% y con un intervalo de confianza inferior de 2000 horas con una probabilidad de 68%. El productor de este auto puede, a partir de estos resultados tomar como garantía de producción de su auto 2000 horas y conoce que existe una probabilidad de 68% de que el auto funcione sin fallar. Pero no satisfecho con este valor el productor desea disminuir esta garantía y correr menos riesgo de roturas y evalúa en la curva $P(t)$ la probabilidad de funcionamiento sin fallo de los rodamientos traseros del auto para 1500 horas obteniendo 87% con un nivel de significación del 95%. De esta forma el productor propone que sus autos deben cambiar los rodamientos traseros a las 1500 horas con una seguridad de funcionamiento del 87%.

4.4.3 Interpretación de las gráficas $f(t)$.

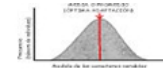
Las funciones de probabilidad para las distribuciones que se trabajan en esta primera versión del SEF se presentan en el acápite 3.1.1 del capítulo anterior. Basados en estas ecuaciones se grafican los valores de $f(t)$ para la muestra en color gris y la forma que tiene la curva teórica en color verde. La observación de estas curvas con un conocimiento mínimo de estadísticas nos puede decir si la muestra se ajusta a la distribución o no. Para una distribución normal si la muestra se ajusta a ella debe describir una campana de Gauss, la amplitud y forma de la campana depende de los parámetros que se hayan estimado. Para una distribución exponencial la curva $f(t)$ se muestra como una función decreciente acercándose cada vez más a cero. Para una distribución de weibull no existe una forma definida. Si el parámetro beta es cercano a tres se parece a una distribución normal, sin embargo si está cercano a uno se parece a una distribución exponencial. Ver **Anexo 4.8**.

4.5 Interpretación de los resultados estadísticos.

Para personas con conocimientos de estadísticas y fiabilidad el SEF provee una información sobre los parámetros que han sido estimados para cada distribución, los valores de las diferencias entre la curva teórica y práctica observados y permitidos, la prueba de bondad del ajuste realizada, el nombre de la variable a la que se le realizó y el nivel de significación con el que se aplicó el test. De esta forma el usuario experto puede corroborar la información que se muestra para la planificación de mantenimientos, determinación de supervivencia o de garantía de producción.

Para la distribución normal se muestran los valores estimados de la media, la varianza, la desviación estándar, los valores máximo y mínimo encontrados en la muestra y las diferencias permitidas y observadas según la prueba que se esté aplicando. Para la distribución exponencial no se muestra el valor de la varianza sino el parámetro lambda estimado.

En el caso de la distribución de weibull, debido a la particular propiedad de que cualquier muestra se ajusta a ella no se muestra máxima diferencia encontrada. Los parámetros que se estiman son alfa y beta. Beta se estima por un método de máxima



verosimilitud y alfa es calculado a partir de beta, todas estas ecuaciones se muestran en el acápite 3.2.4 del capítulo anterior.

Estos resultados estadísticos pueden ser observados de cuatro formas. Seleccionando la hoja *reporte de resultados*. Presionando la opción *mostrar resultados* del menú *estadística*. Presionando el botón *mostrar reporte de resultados* de la barra de herramientas. O simplemente presionando la tecla *F7*. Si intenta visualizar los resultados antes de ejecutar el estudio estadístico no se mostrará nada en la hoja. Ver **Anexo 4.9**

4.5.1 Significados de los parámetros.

Todos los parámetros estimados en este sistema se utilizan para llegar el resultado que se le ofrece al usuario. En particular la media, calculada de una forma u otra según sea la distribución a la que se desea realizar el ajuste, es muy utilizada en la estimación del resto de los parámetros y de los intervalos de confianza. En la distribución normal se calcula, además, la desviación estándar a partir de la varianza, su valor, da la medida de la dispersión de los valores alrededor de la media estimada. En la distribución exponencial el valor de lambda es calculado como el inverso de la media y también constituye una medida de dispersión por lo que da la forma que tiene la muestra.

En la distribución de weibull ya había sido dicho que el parámetro beta aproximado a uno convierte a esta en una distribución exponencial y aproximado a tres en una normal, sin embargo estos parámetros pueden tomar cualquier valor positivo por lo que en el SEF se trabajan esta distribución como ella misma, independientemente del valor de beta. Los valores de alfa y beta determinan los intervalos de confianza superior e inferior.

4.5.2 Usos de los intervalos de confianza.

Son calculados de forma diferente para cada una de las distribuciones que se trabajan en el SEF y poseen una gran importancia en la teoría de la fiabilidad.

Una vez realizado el estudio estadístico, el usuario puede evaluar en la gráfica de $P(t)$ los valores de tiempo calculados como intervalos inferior y superior de confianza así

como la media. Al pulsar el botón *intervalos de confianza* se muestran los intervalos de tiempo y la probabilidad de funcionamiento sin fallo para esa muestra.

En la práctica cuando se intenta modificar o planificar el ciclo de mantenimientos de un equipo se puede escoger un ciclo más cercano al intervalo inferior o superior, según sean las condiciones de explotación del equipo y el nivel de riesgo que se desea correr que este falle antes del mantenimiento. Para la determinación de garantías de producción se utiliza por lo general el valor de tiempo del intervalo inferior para evitar grandes gastos en reparaciones. Para la determinación de supervivencias de individuos se usa un valor más cercano a la media aunque si se desea tener una mayor supervivencia se debe tomar el intervalo inferior.

Por ejemplo, un grupo de puentes construidos por una misma compañía han sido estudiados con el SEF, obteniendo como resultados que la muestra se ajusta a una distribución normal, que tienen una media de vida útil de 15.6 años. Al calcular los intervalos de confianza se obtiene que corresponden a 12 y 19.2 años. El usuario teniendo en cuenta las implicaciones sociales y de pérdidas de vidas que traería la rotura de un puente no puede tomar otro valor para la reconstrucción de los puentes que el intervalo de confianza inferior. Sin embargo en otros ejemplos puede que sea más importante prolongar el tiempo y correr el riesgo por lo que significa en gastos un mantenimiento o reparación. Estas evaluaciones no son realizadas por el SEF en esta primera versión sino que es el usuario quien debe tenerlas en cuenta para la toma de decisiones.

4.6 Guardar un problema.

Recoger todos los intervalos de tiempos para completar una muestra que pueda ser analizada en el SEF y sea representativa de la población puede demorar, por lo que el usuario puede guardar un problema después de haber trabajado con él. De esta forma los datos se encuentran almacenados en el ordenador en una forma segura. Los problemas pueden ser guardados hacia un fichero de texto o estructurado o hacia la base de datos.

4.6.1 Hacia un fichero.

Un problema puede ser guardado en dos tipos de ficheros de los cuales puede ser abierto nuevamente en cualquier ocasión. Uno de ellos son los ficheros textos con extensión (*.txt) que una vez guardado el problema en él se puede observar con cualquier otro software para leer textos. La estructura con la que se guardan los datos del problema es simple, se muestra en el encabezado de la columna el nombre de cada variable y a continuación los valores de intervalos de tiempos.

Si por el contrario los datos se guardan en un fichero estructurado, con extensión (*.sst), su contenido solo puede ser observado con el Sistema Estadístico para el estudio de la Fiabilidad, aunque no es difícil construir una aplicación capaz de leerlos pues la estructura que se guarda en cada campo tiene la siguiente forma:

Estructura

INT columna

INT fila

CADENA valor

Además no existen restricciones de seguridad sobre estos ficheros.

Para guardar un problema el usuario lo puede hacer de cualquiera de las tres formas en que lo realiza en cualquier otra aplicación. La primera mediante el menú *archivo* en el submenú *guardar como*, al presionarla aparece un diálogo para que se escoja la dirección donde se localizará el fichero el nombre y la extensión del mismo. La segunda presionando la opción *guardar* del mismo menú, en este caso si el fichero no fue creado nuevo o ya se habían guardado sus cambios, se reescribe el ya existente, de lo contrario se muestra el mismo diálogo que en la opción *guardar como*. La tercera forma es presionando el botón *guardar* de la barra de herramientas que posee la misma funcionalidad que el menú *guardar*.

4.6.2 Hacia la base de datos.

Para guardar un problema hacia la base de datos el usuario puede hacerlo de dos formas. Una de ellas es en el menú *base de datos* presionar el submenú *guardar hacia base de datos* tras la cual aparece un diálogo preguntando el nombre con el que se guardará el problema y un comentario, opcional, sobre el mismo. Si existe en la base

de datos un problema con ese nombre el SEF lo indica y lo sobrescribe si el usuario confirma, de lo contrario el nombre debe ser cambiado. En la parte inferior del diálogo se muestra una barra de progreso indicando la velocidad con la que se van guardando cada una de las variables, esto puede ser útil, debido a que en ocasiones esta opción puede demorar para muestras muy grandes.

La otra forma en que se puede guardar hacia la base de datos es pulsando el botón *guardar hacia la base de datos* de la barra de herramientas señalado por una unidad de computadora y que ejecuta el mismo diálogo que la opción anterior.

Una vez guardado el problema en la base de datos se conserva en la hoja inicial y puede ser modificado, guardado hacia un fichero si se desea o continuar trabajando con él.

4.7 Facilidades para el almacenamiento de información.

El Sistema Estadístico para el estudio de la Fiabilidad puede almacenar en su base de datos tanta información como su ordenador sea capaz. Además el usuario puede guardar información en ficheros usando unidades externas de almacenamiento. En caso de que la base de datos sufra algún tipo de inconsistencia, por una mala manipulación, es aconsejable que un especialista en el tema analice la estructura de la misma y repare el daño, si no es posible, el sistema puede funcionar perfectamente sin la base de datos, solo se pierden los ejemplos y las opciones del menú *base de datos*, esto es indicado por el SEF en caso extremo de ocurrir.

4.7.1 Portabilidad de los datos.

Un problema guardado en el SEF en un fichero de tipo texto puede ser analizado con cualquier otro sistema estadístico. El usuario puede comparar los resultados que le da el Sistema estadístico para el estudio de la Fiabilidad con otros softwares como el SPSS, el Excel, entre otros. La fácil estructura en que se guardan los datos en estos ficheros permite que estos otros sistemas puedan leerlos y realizar los análisis. Además un fichero estructurado o de texto puede ser transportado de un ordenador a otro por cualquiera de las vías que existen, por correo electrónico, por la red o usando dispositivos de almacenamiento de información. El tamaño de los ficheros generados

por el SEF depende únicamente del tamaño de la muestra y de la cantidad de variables que posea el problema.

En el sistema además se generan dos tipos de ficheros más que de igual forma que los anteriores pueden ser transportados, almacenados u observados con otros softwares. Los gráficos se guardan hacia un fichero, mapa de bits, con extensión (*.bmp) y el reporte de resultados con los valores de la media, los intervalos de confianza y el nivel de significación para cada una de las variables y los test aplicados, se pueden guardar en un fichero texto.

4.7.2 Seguridad y consistencia de las muestras.

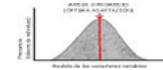
La base de datos del SEF le garantiza al usuario que las muestras recogidas y almacenadas en ellas se mantengan seguras y consistentes. La información con la que se trabaja en la fiabilidad puede ser de mucho valor por lo que este aspecto tiene peculiar importancia. La base de datos no debe ser manipulada con otro software que no sea el SEF para garantizar su consistencia.

Sin embargo ninguna información almacenada en un ordenador es cien por ciento segura por lo que es aconsejable que el usuario almacene sus muestras tanto en la base de datos como en ficheros en dispositivos de memoria.

4.8 Generar nuevos valores.

Para ejecutar esta opción el usuario tiene dos variantes. En el menú principal existe una opción llamada *simulación* dedicada a la generación de valores. En esta versión del sistema se han implementado dos formas de generar los valores. La primera a partir de los valores de tiempo que ya tenga la variable y la segunda el usuario debe indicar la distribución para generar los nuevos intervalos.

La primera se ejecuta presionando el submenú *auto simulación* del menú *simulación* o presionando el botón *auto simulación* de la barra de herramientas señalado con una PC. Al comenzar esta acción aparece un diálogo con una lista de las variables del problema y una caja de texto para que se indique la cantidad de valores a generar y la variable. Presionando el botón aceptar comienza la generación de valores que aparecen en la hoja inicial a continuación del último valor de la muestra. Mientras se



generan los intervalos, una barra de progreso indica el por ciento de generación. Si por el contrario se presiona el botón cancelar solo se culminará la acción.

La segunda forma de simulación se puede ejecutar en el mismo menú pero seleccionando el submenú *simulación manual* o presionando el botón *generación manual* de la barra de herramientas señalado por una mesa de trabajo de un ingeniero. De manera similar a la anterior aparecerá un diálogo, la única diferencia consiste en que en éste se muestra una lista de las distribuciones para que el usuario decida con cual se deben generar los valores.

Para calcular los nuevos valores el SEF estima los parámetros de la muestra. Si se intenta generar valores desde cero, es decir sin haber recogido ningún valor de la muestra, se utilizan los valores por defecto de las medidas de forma y dispersión de la distribución. Por ejemplo para una distribución normal se utiliza la media cero y la varianza uno.

4.9 Análisis de la velocidad de procesamiento.

Este aspecto está muy relacionado a las prestaciones del ordenador en que se esté utilizando el sistema. Por ello para tener una idea, se debe acudir a las comparaciones con otros sistemas para ver los puntos a favor y en contra de esta herramienta.

En un ordenador con prestaciones medias, 2.4 GHz de velocidad y 256 Mb de RAM el Sistema Estadístico para el estudio de la Fiabilidad demora aproximadamente cinco segundos en iniciar. En el mismo ordenador se analiza un caso medio, un problema de cuatro variables de 250 valores cada una, analizando la bondad del ajuste a las tres distribuciones que se utilizan en esta versión y aplicando las dos pruebas de bondad del ajuste para un nivel de significación del 95% se tarda aproximadamente cuatro segundos.

Estos datos que se han dado pueden variar en función de la velocidad del procesador, la cantidad de memoria RAM o la cantidad de aplicaciones que se encuentran en ejecución.

4.9.1 Comparación con otras herramientas.

Las principales ventajas del Sistema Estadístico para el estudio de la Fiabilidad no consisten en la velocidad de procesamiento de los datos sino en la información que brinda. Ya en el primer capítulo se mostró un análisis de las ventajas y desventajas sobre otros sistemas existentes.

Sin embargo teniendo en cuenta que este sistema ha sido desarrollado para el estudio de la fiabilidad y no para la estadística en general. Es por ello que para generar los resultados estadísticos y de fiabilidad tarda menos que otras herramientas como el SPSS que calcula otra cantidad de valores debido a su proyección a estudios sociales. No pudo ser comparado con otras herramientas como el STATA o el SAS debido a que no realiza las mismas operaciones o no pudieron ser consultadas.

4.10 Requerimientos.

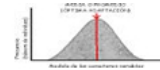
El sistema Estadístico para el estudio de la Fiabilidad puede ser instalado en cualquier ordenador con sistema operativo Windows. Para su correcto funcionamiento es necesario que la resolución de la pantalla sea de 1024 X 768 aunque una resolución menor puede ser utilizada con el inconveniente de que quizás los gráficos no se muestran completamente.

El sistema operativo debe tener como idioma principal el inglés. En caso de no ser así el SEF mostrará un mensaje de error al ejecutar el estudio estadístico.

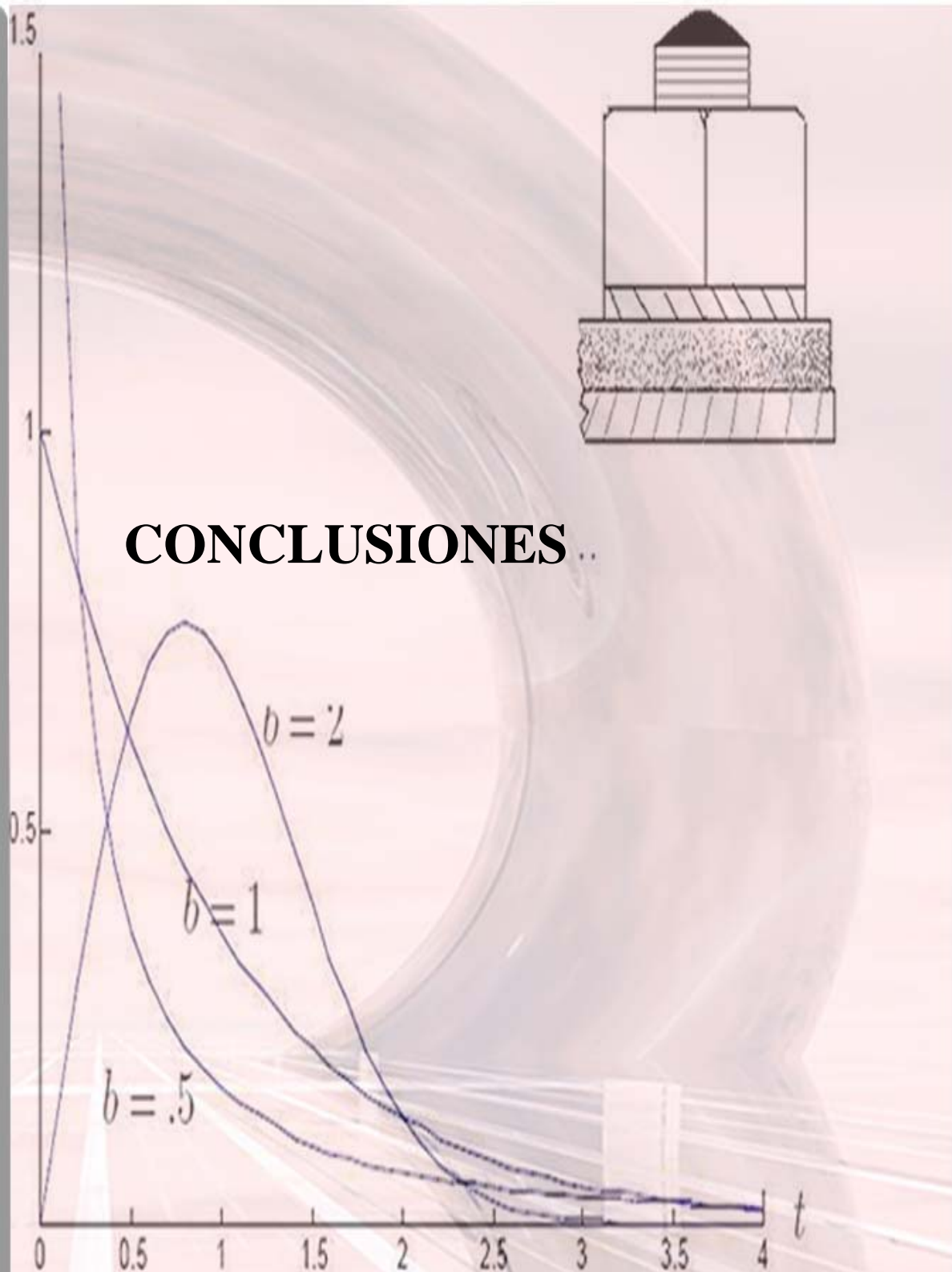
Para poder utilizar los ejemplos o los problemas almacenados en la base de datos es necesario que el sistema operativo cuente con el motor de bases de datos Microsoft Access. Este requisito no es imprescindible, su no cumplimiento solo limita el uso de la base de datos pero no del resto del sistema que puede seguir funcionando y guardando los datos hacia ficheros.

4.11 Conclusiones parciales.

Al concluir este capítulo, a manera de resumen, se enfatiza en que el Sistema Estadístico para el estudio de la Fiabilidad es un software con características visuales similares a las aplicaciones del sistema operativo para el que se ha generado. Se basa en el continuo intercambio de información con el usuario por lo que su interfaz visual



tiene particular importancia. Puede realizar estudios a varias muestras de tiempos de funcionamientos entre o hasta el fallo de un equipo o individuo, dándole al usuario resultados estadísticos y de fiabilidad, numéricos y gráficos. Es capaz de almacenar información en una base de datos o hacia ficheros que pueden ser leídos con otros softwares. Y por último no requiere de grandes prestaciones para realizar estas acciones.



CONCLUSIONES ..

Sistema Estadístico para el estudio de la Fiabilidad

Conclusiones

Al concluir este trabajo es necesario recapitular los principales aspectos que en él se han analizado y que han quedado demostrados a lo largo de los cuatro capítulos con los que cuenta. Luego de obtener algunos conocimientos sobre la aparición de la estadística y la fiabilidad como ciencias:

- Se ha analizado la situación actual de las herramientas computacionales que existen dedicadas a estas temáticas, la pequeña posibilidad de nuestro país de trabajar con ellas y de ahí la necesidad de construir una propia.
- Se ha resumido la estructura del software a través de varios diagramas, las características de sus clases, de la estructura informativa y los detalles de la interfaz de intercambio con el usuario.
- Han sido presentados los algoritmos matemáticos convertidos en algoritmos computacionales y las ecuaciones de las distribuciones trabajadas.
- Se ha logrado brindar al usuario una guía de cómo trabajar con el software y cuales son las formas de obtener la máxima explotación de los resultados que brinda esta herramienta computacional para realizar estudios estadísticos y sobre la teoría de la fiabilidad.
- Se logró construir un Sistema Estadístico para el estudio de la Fiabilidad capaz de realizar análisis estadísticos a cualquier muestra de intervalos tiempos y ofrecer resultados acordes a las características de la muestra.

Recomendaciones

El Sistema Estadístico para el estudio de la Fiabilidad, en su primera versión, no ha logrado abarcar la gran cantidad de métodos estadísticos que se utilizan en la teoría de la fiabilidad. Este ha sido a penas un primer intento para acercarse a esta ciencia tan importante como usada en el mundo actual. Solo se han tomado tres distribuciones estadísticas para esta versión y se han utilizando los tres niveles de significación más usados en la fiabilidad para la realización de las pruebas de bondad del ajuste.

Por estas razones expuestas se incluyen entre las recomendaciones para próximas versiones del SEF:

- La implementación de otras pruebas de bondad del ajuste.
- La inclusión de otras distribuciones continuas que se utilizan mucho y son de gran importancia para la fiabilidad, entre las que se encuentran: Gamma, Erlang, Pareto, Log-normal y Rayleigh.
- Obtener los valores de los percentiles utilizados o las constantes, tomadas de tablas actualmente, a partir del cálculo computacional.
- Incluir otros niveles de significación, de ser posible, el que el usuario desee.

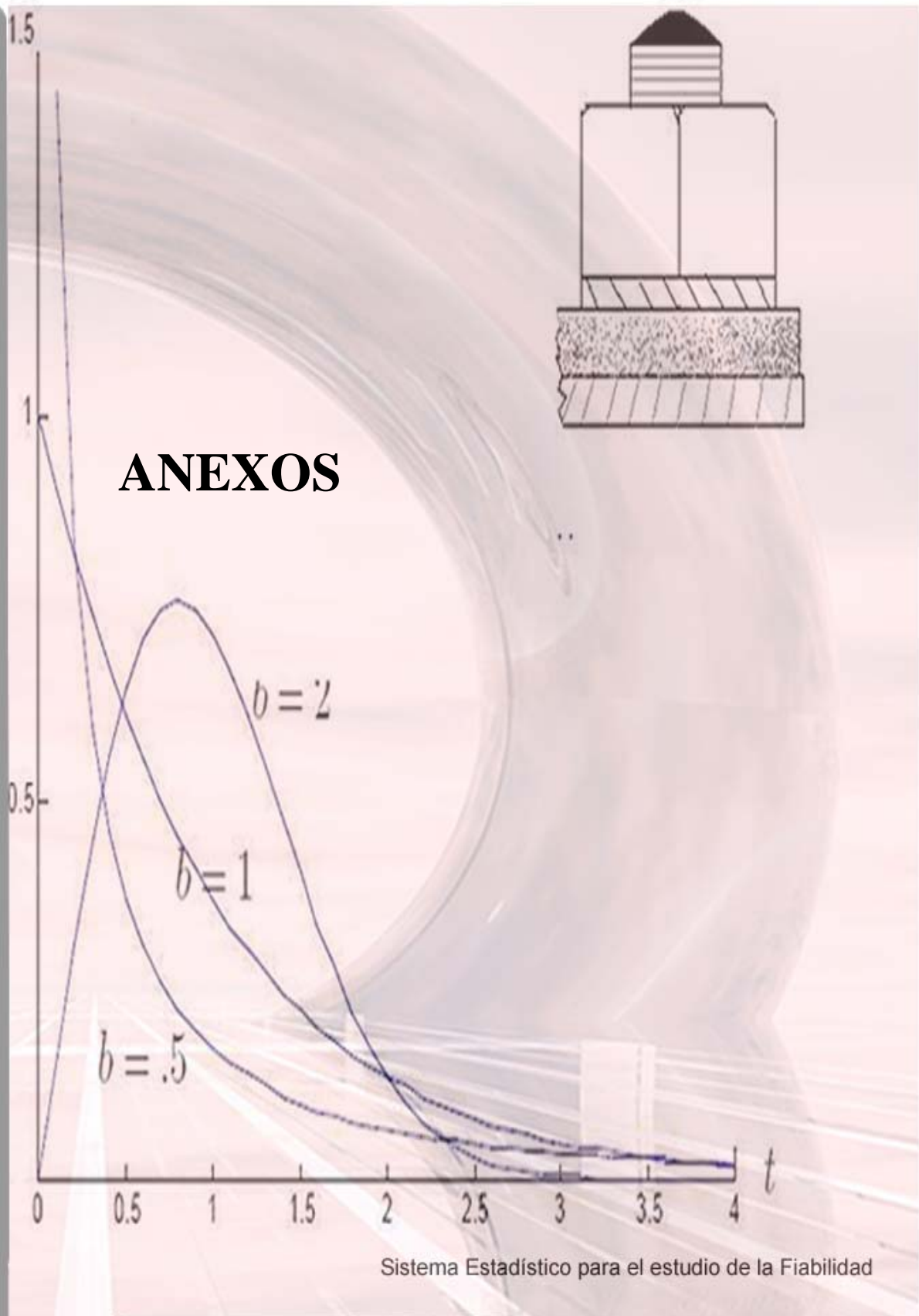
Estas recomendaciones quedan expuestas para próximas versiones del software.

Bibliografía

- [BIO05] BioMates, 2005, Distribuciones de probabilidad. Distribución de Weibull. [En línea], España, disponible en:
http://www.telefonica.net/web2/biomates/tabl/tabl_weibull.htm
[Accedido el 20/09/2005].
- [COL06] Colectivo de autores. Anecdótico matemático. España, disponible en:
<http://www.etsi2.ugr.es/profesores/jmaroza/anecdótico/anecdótico.htm>
[Accedido el día 10 de mayo de 2006]
- [COL03] Colectivo de autores (2003). “edicionsupc.es”. Introducción a la Teoría de la Fiabilidad. España, disponible en:
<http://www.edicionsupc.es/ftppublic/pdfmostra/ME03301M.pdf>
[Accedido el día 15 de mayo de 2006]
- [DAT97] Data Mining Institute, S.L. (1997). “estadístico.com” en Diccionario estadístico. [En línea]. España, disponible en:
<http://www.estadistico.com/dic.html?p=1267> [Accedido 10 de noviembre de 2005].
- [DES04] Desconocido. 2004. Otros modelos del análisis.ppt, Diagramas de Flujos de datos. Facultad MFC, UCLV, Cuba, 2006.
- [GRA04] Grau R, (2004), “Conferencia 1. Introducción a la estadística.” En la UCLV. Conferencias de Estadísticas. Facultad MFC, Universidad Central de Las Villas, Santa Clara, 2004.
- [LEP90] Leplat J y G Terssac. (1990). Les facteurs humains de la fiabilité Marseille (Francia), Ed. Octares.
- [MOL04] Molinero Luis M. (Julio 2004). Asociación de la Sociedad Española de Hipertensión. Historia del razonamiento estadístico, disponible en: <http://www.seh-elha.org/historiastat.htm>. España. [Accedido el 29 de mayo de 2006].
- [PRE] Pressman. Ingeniería del software: un enfoque práctico. 4ta edición. Capítulo 12, pág 207
- [REY03] Reyes T, 2003, Formulario sobre las distribuciones utilizadas en la teoría de la fiabilidad. Sin publicar. Facultad de Ingeniería Mecánica, UCLV.

- [REY97] Reyes T, 1997, Aportes a los métodos y Procedimientos en el cálculo de los índices de Fiabilidad y Mantenimiento. T.D. UCLV.
- [REL06] ReliaSoft. (2006). Weibull++ 7. Análisis de Datos de Vida (Análisis del Weibull). EE.UU., disponible en: <http://www.reliasoft.com.br>. [Accedido el 9 de abril del 2006].
- [SAS06] SAS. (2006). Superior Software that give you the power to know. SAS. EE.UU, disponible en: <http://www.sas.com/index.html>. [Accedido 7 de mayo de 2006].
- [SPI61] Spiegel M. (Octubre 1961). Teoría y problemas de Estadística. La Habana 1979.
- [SPS06] SPSS Inc. (2006). SPSS. Softwares and Solutions. EE.UU, disponible en: http://www.spss.com/products/index.cfm?source=homepage&hpzone=nav_bar. [Accedido el 30 de mayo de 2006]
- [STA06] STATA. (2006). STATA. México., disponible en: <http://www.STATA.com.mx> [Accedido 25 de mayo de 2006].
- [THO06] Thompson, J. (2006) Your Request. Correo electrónico a D. Valdés. (dvramirez@gmail.com) 18 de mayo de 2006 [Accedido el 18 de mayo de 2006].

ANEXOS



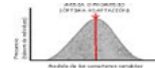
Anexos.

Anexo 1.1: (Tabla cronológica de sucesos importantes en la estadística).

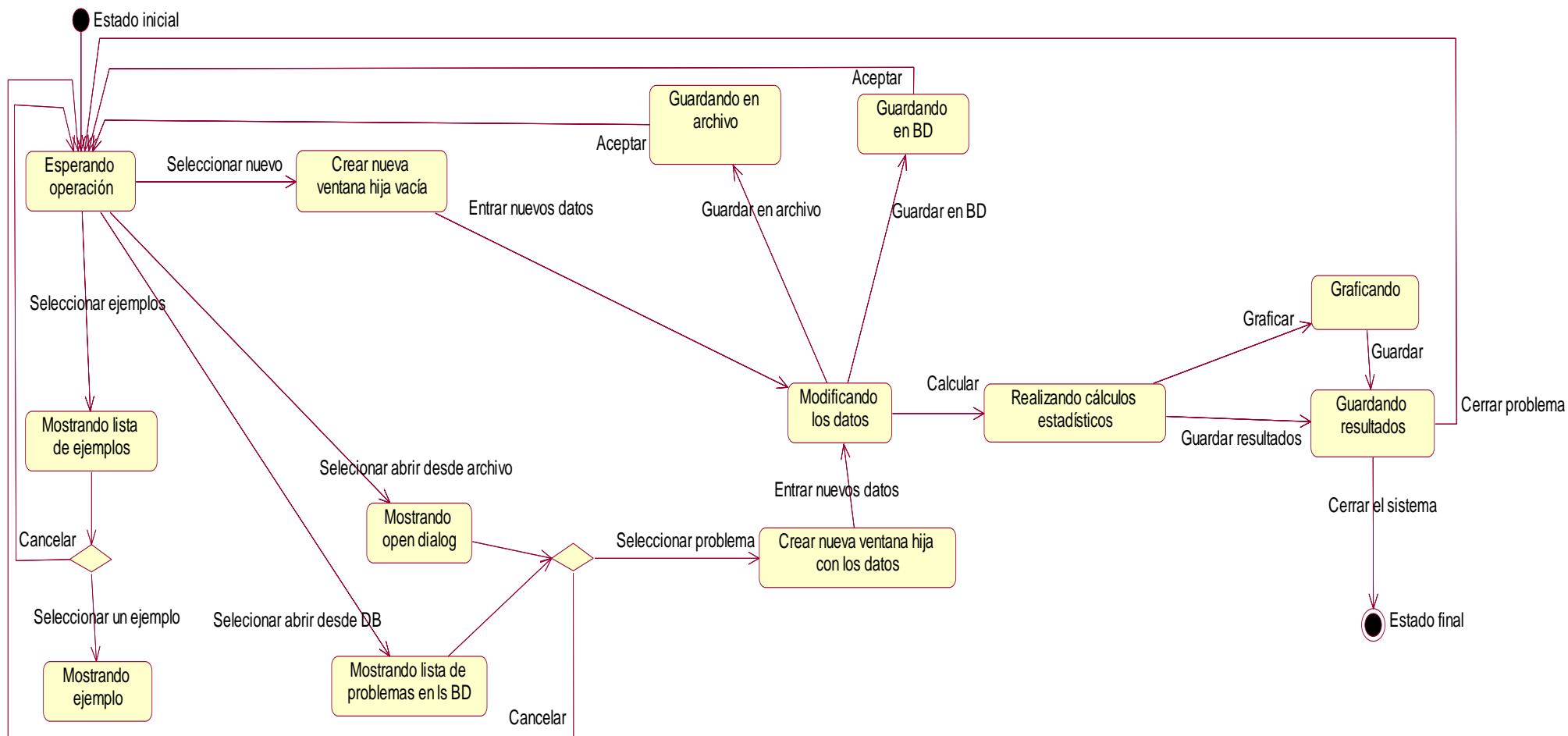
Relación de diferentes sucesos de interés en el desarrollo de la estadística, elegidos quizás algo arbitrariamente. Tomado de <http://www.seh-lelha.org/historiastat.htm>

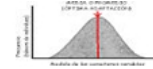
Año	Suceso	Autor
1532	Registro semanal de muertes en Londres	Sir W. Petty
1539	Registro de bautismos, bodas y fallecimientos en Francia	
1654	Correspondencia relativa a los juegos de azar con dados	P. de Fermat B. Pascal
1656	Publicación de <i>Ratiociniis in ludo aleae</i>	Ch. Huygens
1662	Primer estudio demográfico publicado basado en actas de defunción	J. Graunt
1663	Publicación póstuma del libro <i>Liber de ludo aleae</i>	G. Cardano
1713	Publicación póstuma de <i>Ars Conjectandi</i>	J. Bernoulli
1714	Publicación de <i>Teoría de la probabilidad</i>	A. De Moivre
1763	Publicación póstuma de <i>Essay Towards Solving a Problem in the Doctrine of Chances</i> . Contiene el famoso teorema de Bayes	Rev. Bayes
1809	Publicación de <i>Theoria Motus Corporum Coelestium</i>	C.F. Gauss
1812	Publicación de <i>Théorie analytique des probabilités</i>	P.S. Laplace
1833	Planteamiento de los conceptos básicos para la creación de una máquina de calcular	Ch. Babbage
1834	Creación de la <i>Statistical Society of London</i>	
1837	Distribución de Poisson	D.S. Poisson
1839	Creación de la <i>American Statistical Association</i> (Boston)	
1855	Epidemiología del cólera	J. Snow
1889	Publicación de <i>Natural Inheritance</i>	F. Galton
1895	Introducción de las distribuciones asimétricas	K. Pearson
1900	Desarrollo de la prueba del χ^2	K. Pearson
1900	Redescubrimiento de los trabajos de Mendel	W. Bateson
1903	Desarrollo del Análisis de Componentes Principales	K. Pearson
1908	Publicación de <i>The probable error of a mean</i>	W.S. Gosset "Student"

1910	Publicación de <u>An introduction to the theory of statistics</u>	G.U. Yule
1911	Publicación <u>The grammar of science</u>	K. Pearson
1915	Distribución del coeficiente de correlación	R.A. Fisher
1921	Publicación <u>Treatise on probability</u>	J.M. Keynes
1925	Publicación <u>Statistical Methods for Research Workers</u>	R.A. Fisher
1925	Métodos de máxima verosimilitud ML	R.A. Fisher
1928	Pruebas para verificación de hipótesis	J. Neyman, E.S. Pearson
1933	Publicación de <u>On the empirical determination of a distribution</u>	A.N. Kolmogorov
1933	Aparece Probit análisis	C. Bliss
1935	Publicación de <u>The Design of Experiments</u>	R.A. Fisher
1936	Publicación de <u>Relations between two sets of variables</u>	H. Hotelling
1937	Cálculo de las tasas de desempleo en USA mediante muestreo aleatorio	M. Hansen, F. Stephan
1937	Publicación <u>Principles of Medical Statistics</u>	A.B. Hill
1940	Publicación del libro de texto <u>Statistical Methods</u>	G.W. Snedecor
1945	Pruebas no paramétricas	F. Wilcoxon
1947	Métodos secuenciales	A. Wald
1950	Publicación <u>Experimental Design</u>	W.G. Cochran, G.M. Cox
1950	Construcción del primer ordenador moderno con lámparas de vacío UNIVAC 1	
1974	Comercialización del microprocesador de 8 bit y 2 Mhz Intel 8080	
1976	Publicación <u>Time Series Analysis Forecasting And Control</u>	George Box, Gwilym Jenkins, Holden-Day
1977	Publicación <u>Exploratory Data Analysis</u>	J. Tukey
1977	Comercialización del ordenador personal Apple II.	
1981	Comercialización del IBM PC 5150	

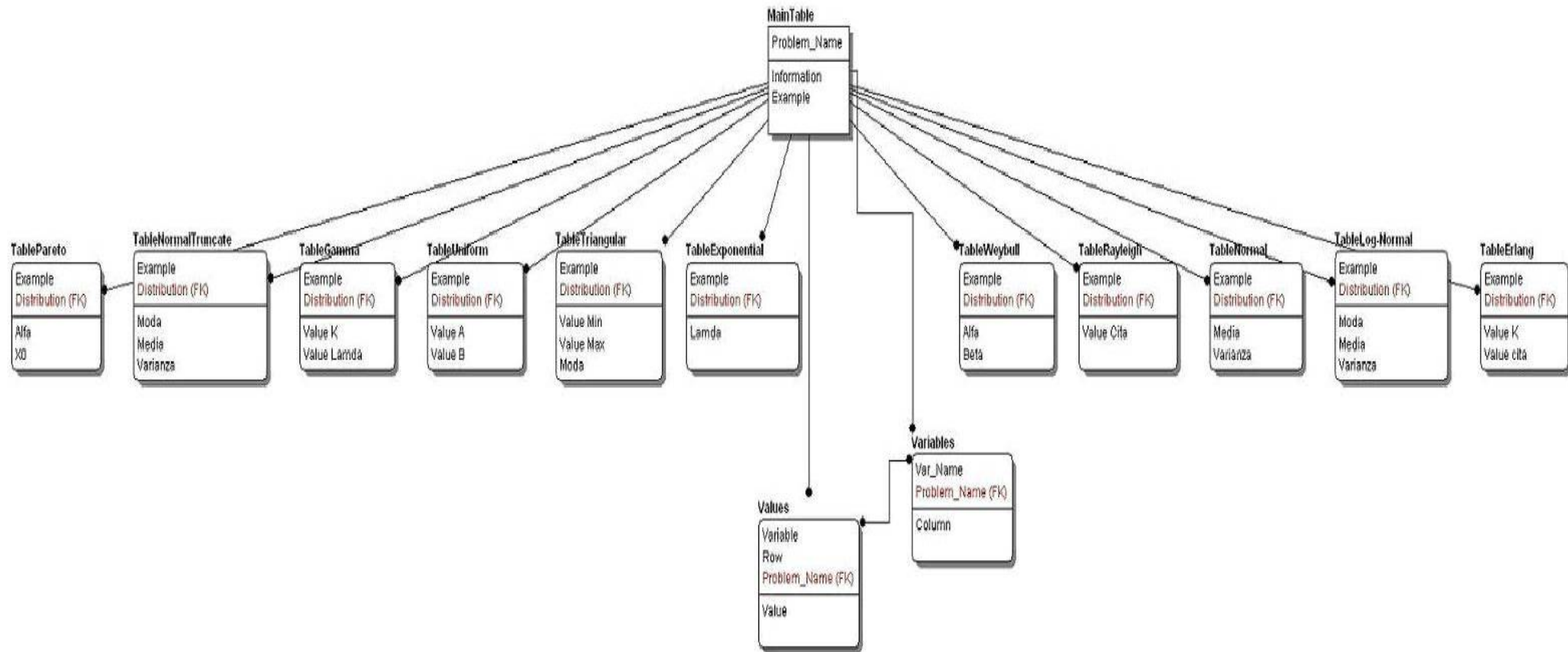


Anexo 2.1: Diagrama de estados general.

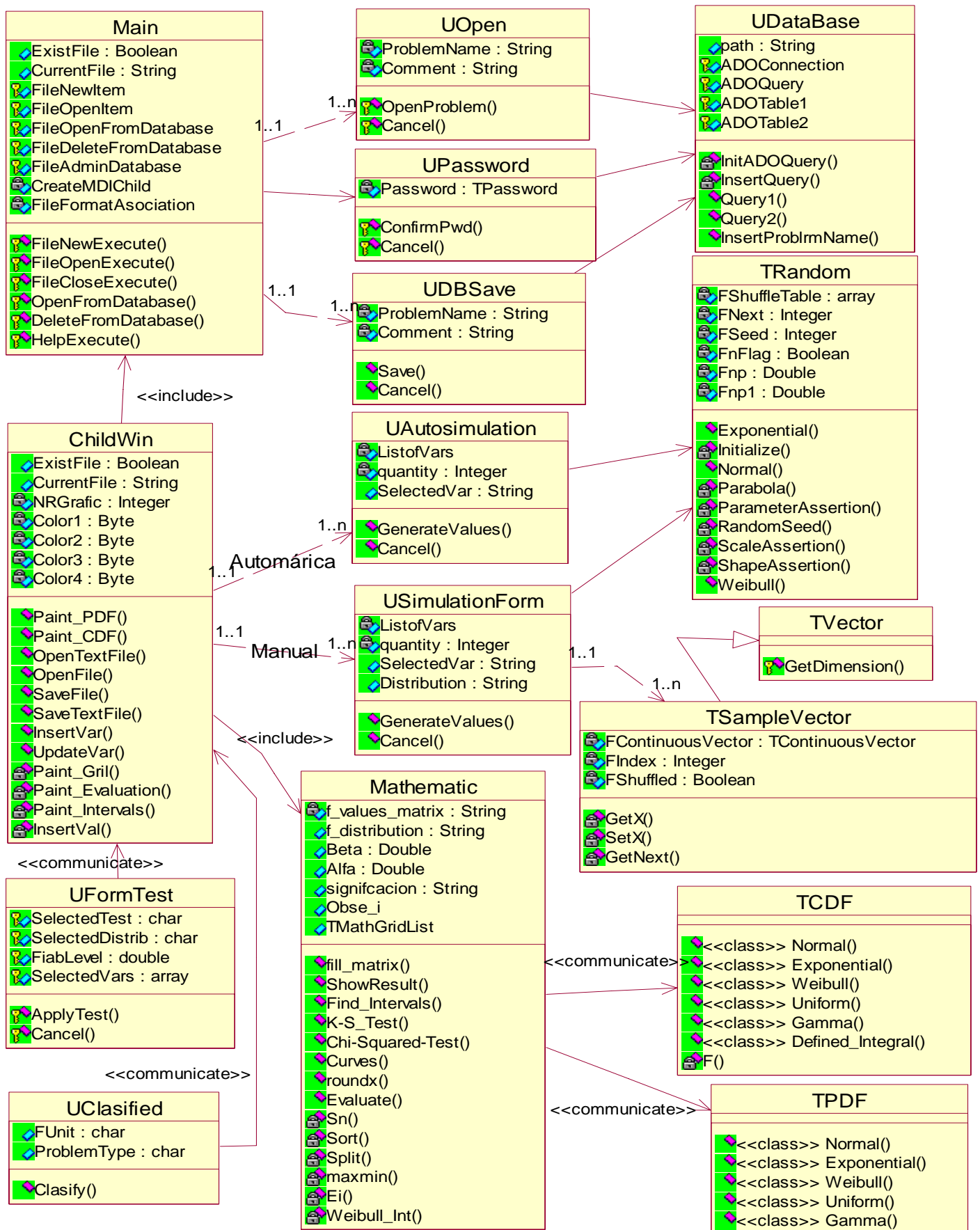




Anexo 2.2: Diagrama de la base de datos.

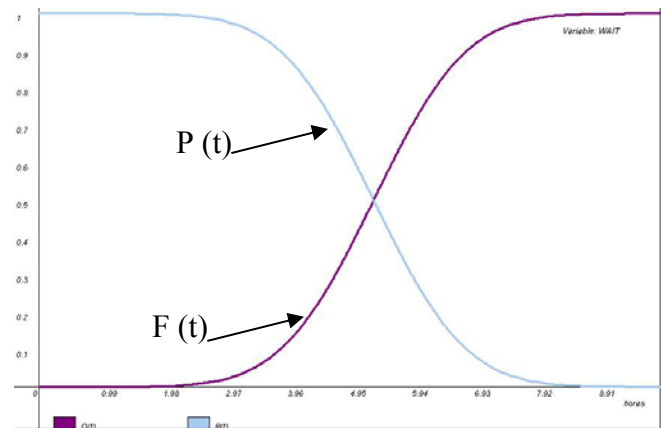
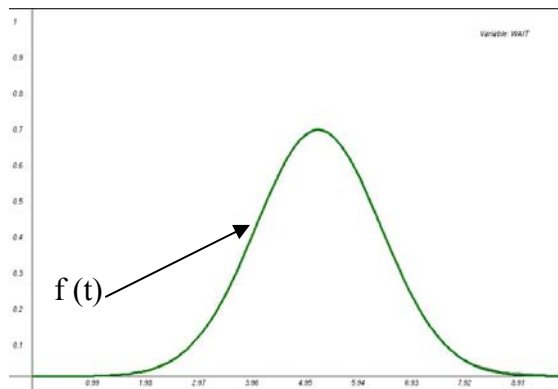


Anexo 2.3: Diagrama de clases (Notación UML).

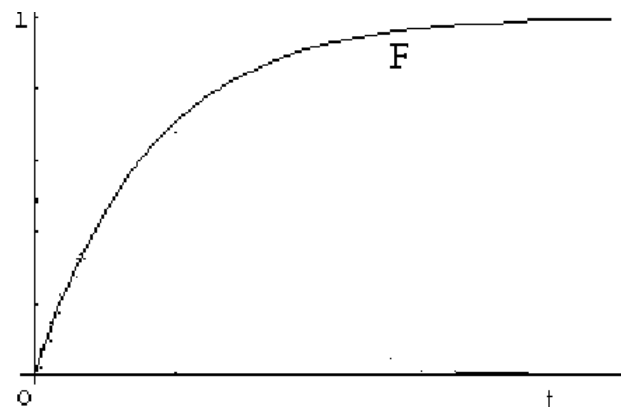
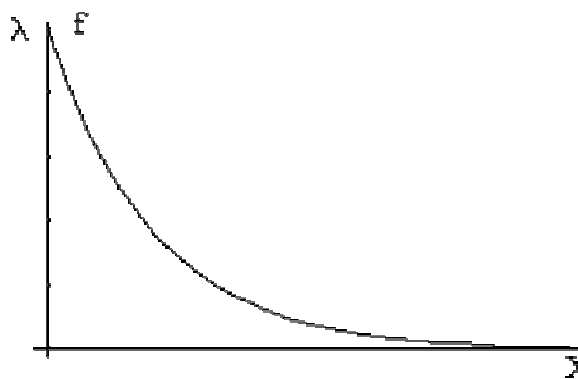


Anexo 3.1: (Imágenes de las curvas $F(t)$ y $f(t)$).

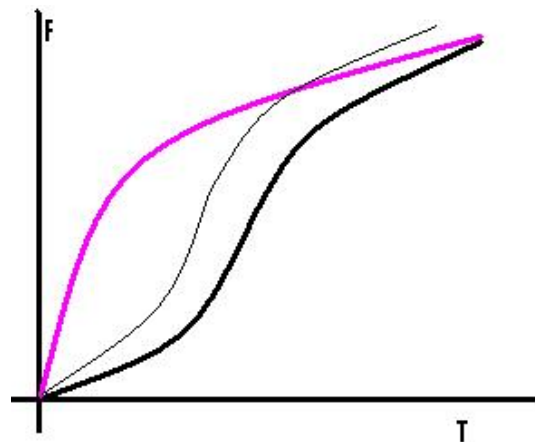
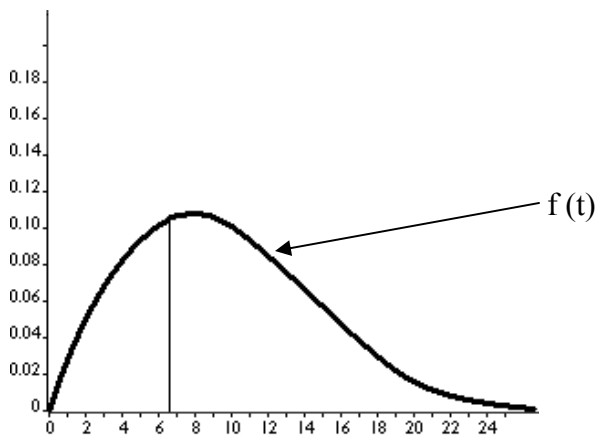
Distribución normal

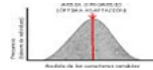


Distribución exponencial

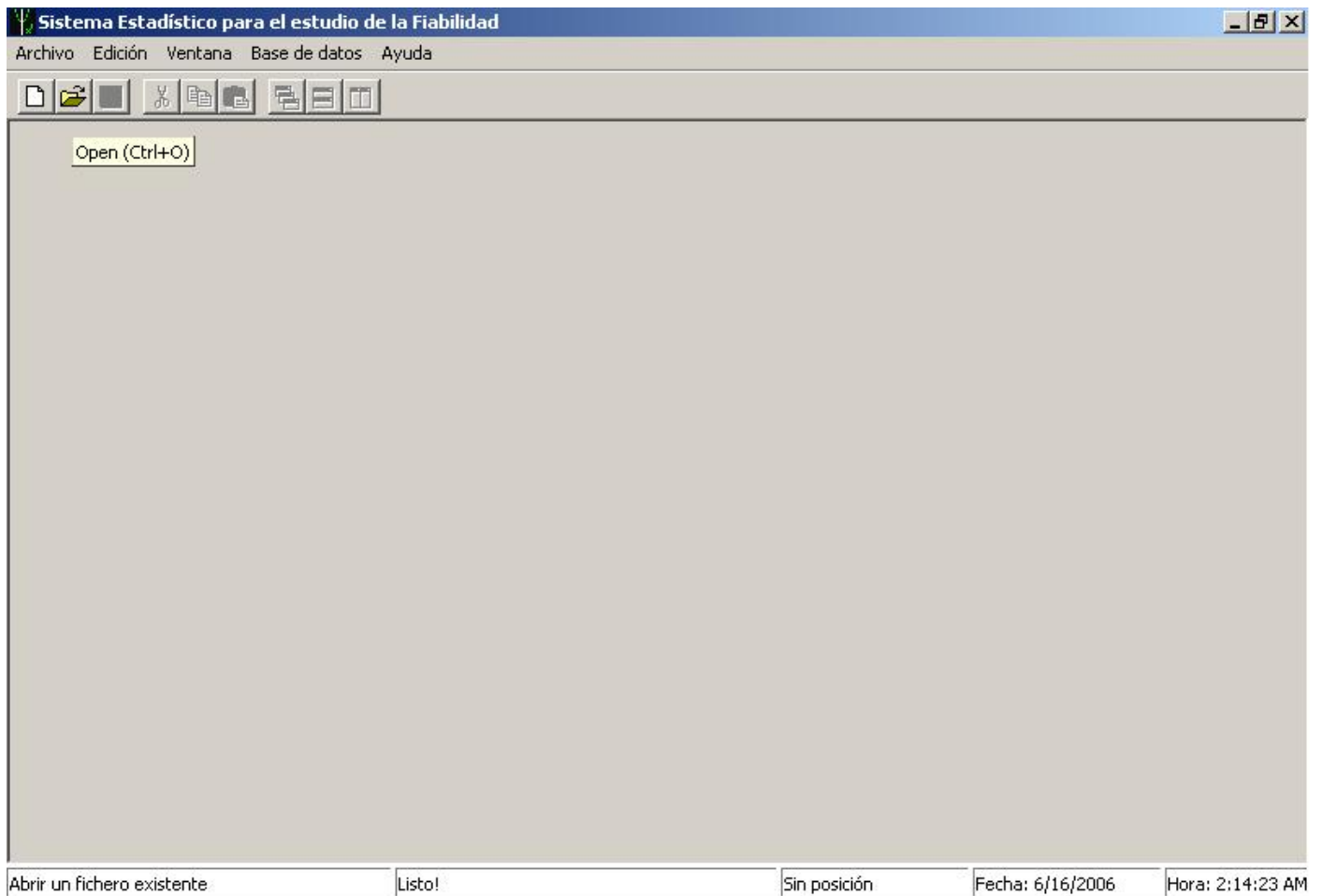
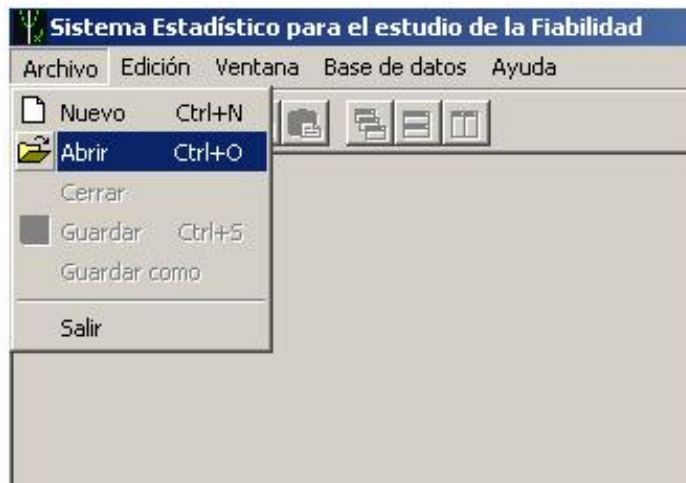


Distribución de weibull. ($F(t)$ y $f(t)$ pueden tener varias formas en esta distribución)

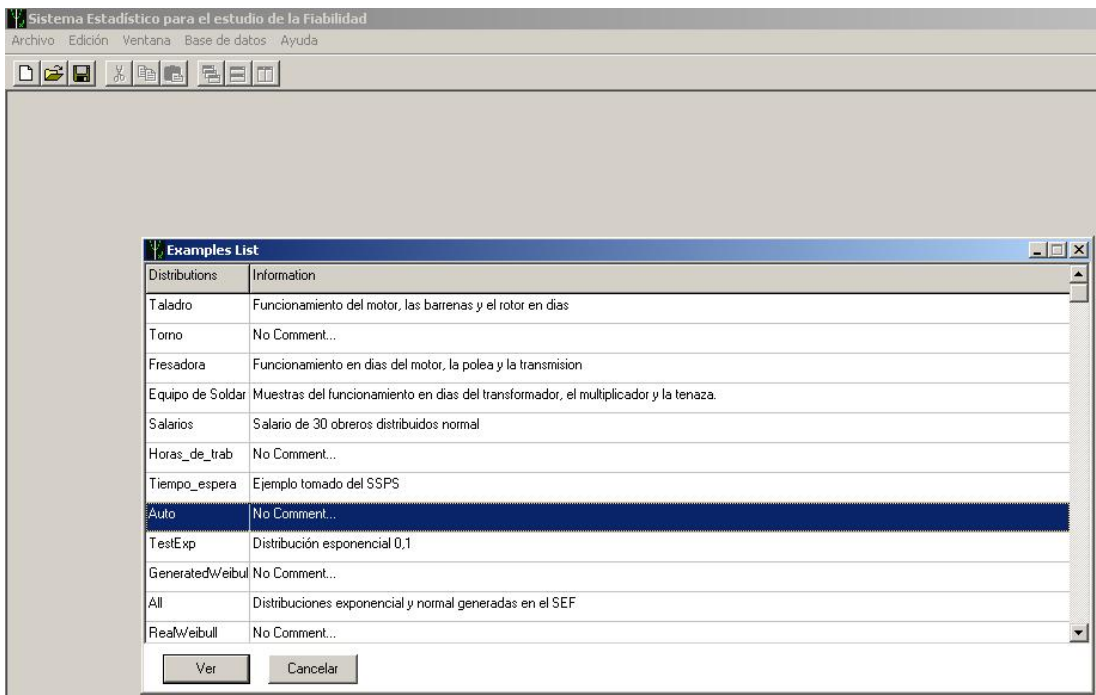
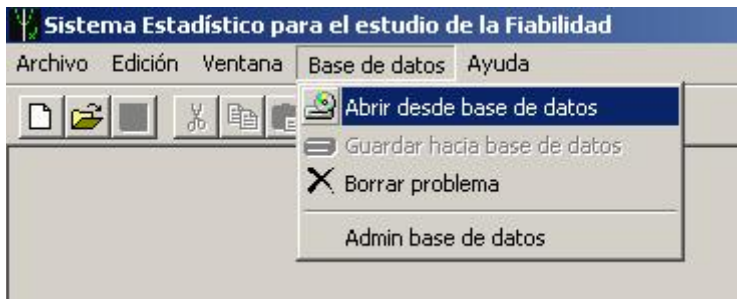




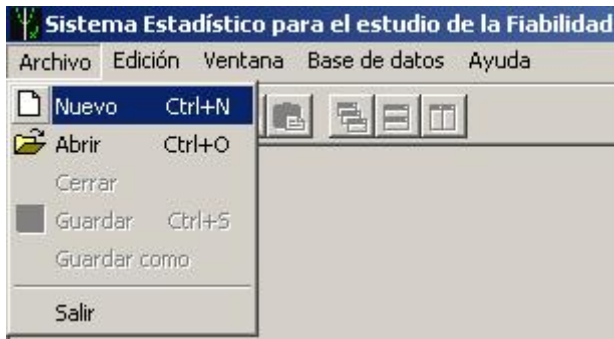
Anexo 4.1: Diseño de Ventanas (Abrir un problema desde un fichero).

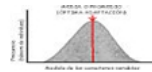


Anexo 4.2: Diseño de Ventanas (Abrir un problema desde la base de datos).



Anexo 4.3: Diseño de Ventanas (Crear un nuevo problema).





Anexo 4.4: Diseño de Ventanas (Clasificar un problema).

Clasifique su problema

☒ Planificación de mantenimientos

☐ Determinación de garantías de producción

☐ Determinación de supervivencias en individuos

Unidad de medida

☐ Minutos

☒ Horas

☐ Días

☐ Meses

☐ Años

Aceptar

Este tipo de problema le permite hacer una planificación del mantenimiento de su equipo o pieza, en el gráfico podrá observar los intervalos de confianza y la media así como evaluar cualquier valor de tiempo en la curva de fiabilidad

SEF versión 1.0.2

Gráfico de fiabilidad que muestra la curva de fiabilidad y la curva de mantenimiento. El gráfico incluye una escala de tiempo (t) y una escala de fiabilidad (F). Se muestran los intervalos de confianza y la media. El gráfico también incluye una imagen de un motor y un diagrama de un sistema de fiabilidad.

Anexo 4.5: Diseño de Ventanas (Definición de variables).

Sistema Estadístico para el estudio de la Fiabilidad - ['Auto']

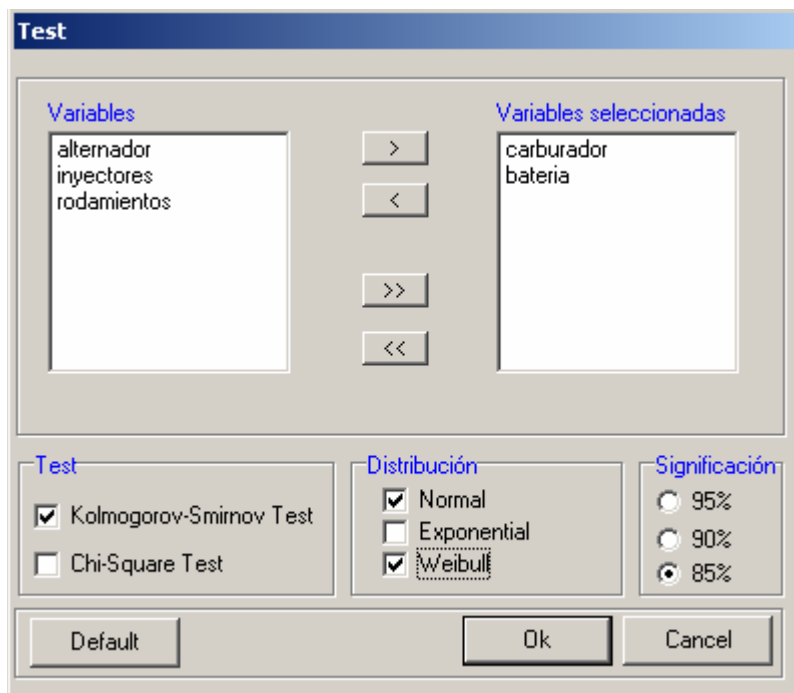
Archivo Edición Ventana Base de Datos Estadística Simulación Help

Hoja de calculo Función de distribución acumulada Función de probabilidad Reporte de Resultados

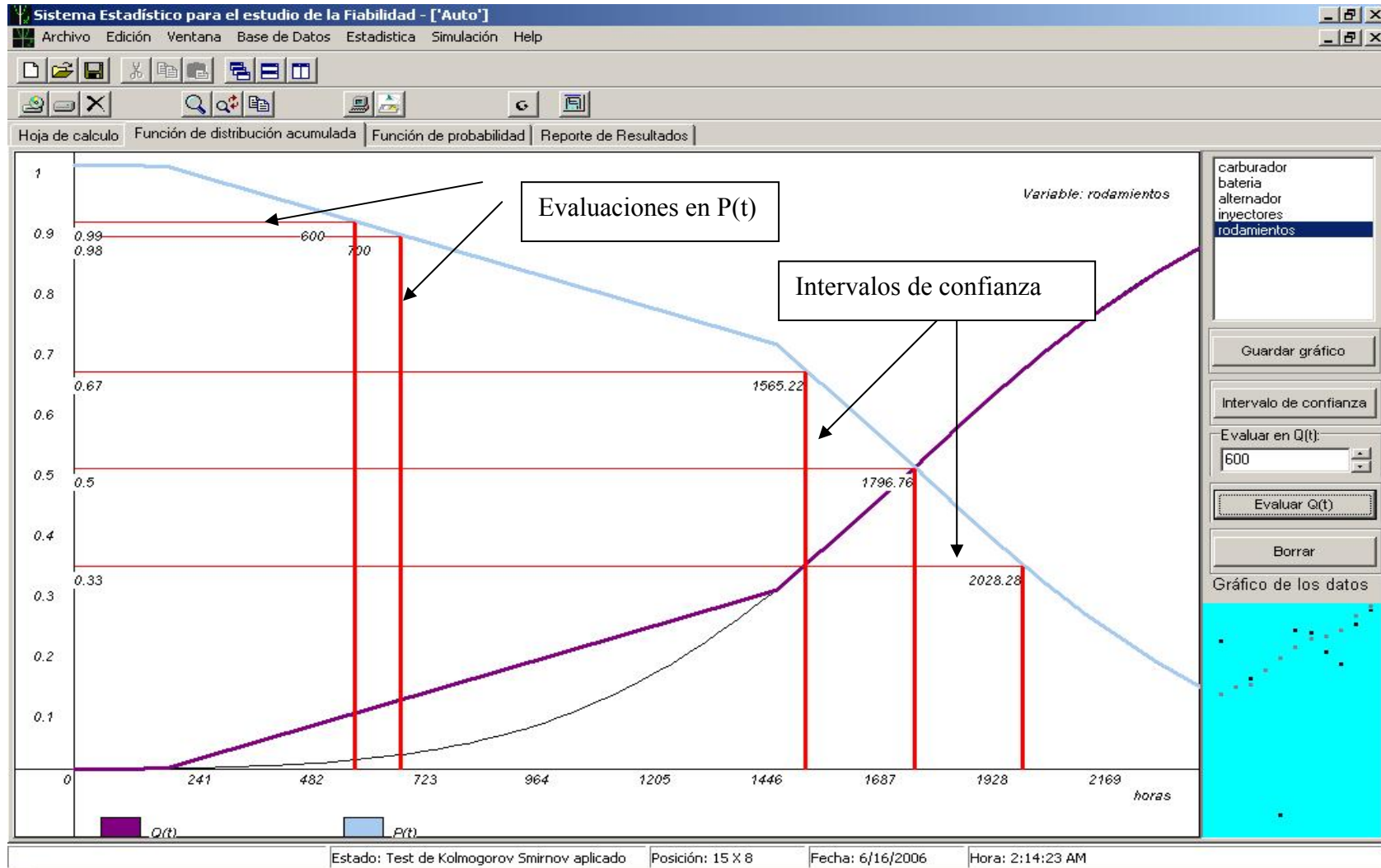
Variables	carburador	bateria	alternador	inyectores	var5	var6	var7	va
1	120	54	347	457	1992	rodamientos		
2	100	32	312	327	1500			
3	90	67	291	654	1600			
4	85	84	300	193	2410			
5	157	15	312	164	200			
6	134	94	481	581	2100			
7	164	72	127	624	2080			
8	191	64	261	931	1873			
9	82	58	324		1754			
10	54	51	258		2164			
11	72	37	245		2314			
12	197	69	192		1574			
13	168	61	194					
14	143	49	347					

Presione enter al teclear la variable Posición: 13 X 5

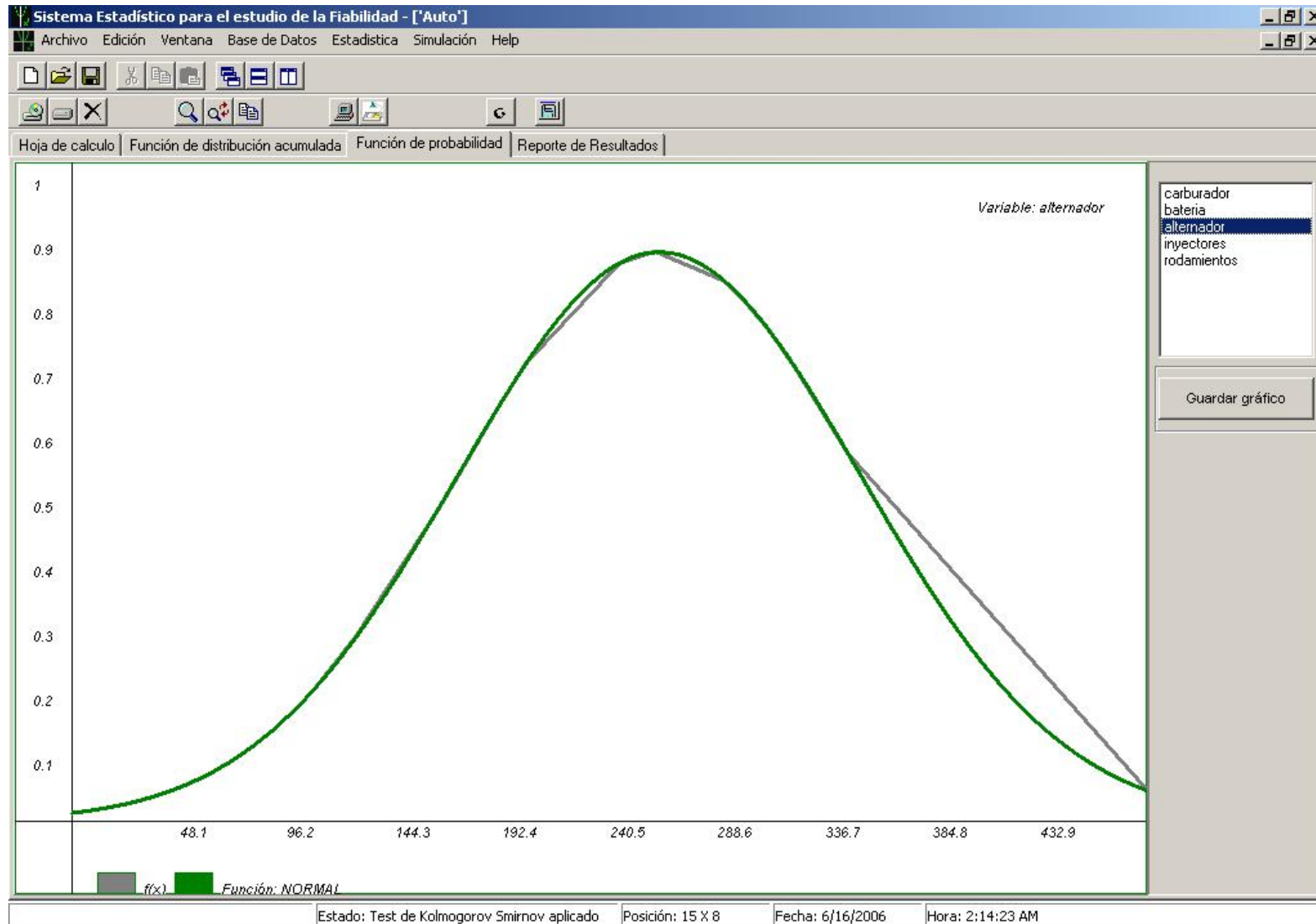
Anexo 4.6: Diseño de Ventanas (Ejecución de estudio estadístico).



Anexo 4.7: Diseño de Ventanas (Curvas P (t) y Q (t), Evaluación).



Anexo 4.8: Diseño de Ventanas (Curvas $f(t)$).





Anexo 4.9: Diseño de Ventanas (Reporte de resultados).

Sistema Estadístico para el estudio de la Fiabilidad - ['Auto']

Archivo Edición Ventana Base de Datos Estadística Simulación Help

Hoja de calculo Función de distribución acumulada Función de probabilidad Reporte de Resultados

Variable	Valor
TEST DE CARBURADOR	KOLMOGOROV-SMIRNOV
Media:	137.55
Por ciento de fiabilidad:	85
Máximo valor encontrado:	234
Mínimo valor encontrado:	54
Desviación estándar:	47.7152753319102
Varianza:	2276.7475
Máxima diferencia encontrada:	0.0924805518367346
Nivel de significación:	0.24600000679493
TEST DE BATERIA	KOLMOGOROV-SMIRNOV
Media:	56.2666666666667
Por ciento de fiabilidad:	85
Máximo valor encontrado:	94
Mínimo valor encontrado:	15
Desviación estándar:	19.8711404358739
Varianza:	394.862222222222
Máxima diferencia encontrada:	0.102858812759062
Nivel de significación:	0.282999992370605
TEST DE ALTERNADOR	KOLMOGOROV-SMIRNOV
Media:	262.647058823529
Por ciento de fiabilidad:	85
Máximo valor encontrado:	481
Mínimo valor encontrado:	107
Desviación estándar:	90.3255166713023
Varianza:	8158.69896193772
Máxima diferencia encontrada:	0.118125250637811

El Sistema Estadístico para el estudio de la fiabilidad ha concluido su estudio para la variable carburador.
 Le da como propuesta, para el establecimiento o reajuste del ciclo de mantenimiento preventivo, el intervalo comprendido entre 122.19 horas y 152.91 horas obtenido a partir del valor de la media muestral 137.56 horas
TEST KOLMOGOROV-SMIRNOV APLICADO
 variable: carburador
 distribución: normal
 nivel de significación permitido: 0.15

El Sistema Estadístico para el estudio de la fiabilidad ha concluido su estudio para la variable bateria.
 Le da como propuesta, para el establecimiento o reajuste del ciclo de mantenimiento preventivo, el intervalo comprendido entre 48.88 horas y 63.66 horas obtenido a partir del valor de la media muestral 56.27 horas
TEST KOLMOGOROV-SMIRNOV APLICADO
 variable: bateria
 distribución: normal
 nivel de significación permitido: 0.15

El Sistema Estadístico para el estudio de la fiabilidad ha concluido su estudio para la variable alternador.
 Le da como propuesta, para el establecimiento o reajuste del ciclo de mantenimiento preventivo, el intervalo comprendido entre 231.1 horas y 294.2 horas obtenido a partir del valor de la media muestral 262.64 horas
TEST KOLMOGOROV-SMIRNOV APLICADO
 variable: alternador
 distribución: normal
 nivel de significación permitido: 0.15

El Sistema Estadístico para el estudio de la fiabilidad ha concluido su estudio para la variable inyectores.
 Le da como propuesta, para el establecimiento o reajuste del ciclo de mantenimiento preventivo, el intervalo comprendido entre 368 horas y 614.74 horas obtenido a partir del valor de la media muestral 491.38 horas
TEST KOLMOGOROV-SMIRNOV APLICADO

Estado: Test de Kolmogorov Smirnov aplicado Posición: 15 X 8 Fecha: 6/16/2006 Hora: 2:14:23 AM