



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**ANÁLISIS DE DATOS DE FIABILIDAD
PARA AJUSTAR UNA DISTRIBUCIÓN CONOCIDA**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

A C T U A R I O

P R E S E N T A:

ALBERTO ALBERTO ALFARO



**DIRECTOR DE TESIS:
M. EN C. JOSÉ ANTONIO FLORES DÍAZ
2013**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno

Alberto

Alfaro

Alberto

17 18 04 96

Universidad Nacional Autónoma de
México

Facultad de Ciencias

Actuaría

303000448

2. Datos del tutor

M en C

José Antonio

Flores

Díaz

3. Datos del sinodal 1

Dra

María del Pilar

Alonso

Reyes

4. Datos del sinodal 2

Act

Jaime

Vázquez

Alamilla

5. Datos del sinodal 3

M en I de O

María Isabel

Escalante

Membrillo

6. Datos del sinodal 4

Act

Rosa Daniela

Chávez

Aguilar

7. Datos del trabajo escrito

Análisis de datos de fiabilidad
para ajustar una distribución conocida

78 p

2013

Índice

Introducción.....	1
1. Conceptos de fiabilidad	3
Introducción	3
1.1 Algunos conceptos de fiabilidad.....	4
1.2 Clasificación para funciones de distribución o fiabilidad	8
1.3 El papel de la función de riesgo $r(t)$	13
1.4 Combinaciones de distribuciones	15
1.5 Ordenamientos parciales	16
1.6 Envejecimiento relativo	17
1.7 Distribuciones de vida con tasa de falla en forma de bañera	18
1.8 El papel de la vida residual media	24
1.9 Modelos con tiempos de falla discretos	26
1.10 Envejecimiento bivariado y multivariado	31
2. Aproximación de las funciones de supervivencia y	35
ajuste de algunas distribuciones a datos observados	
Introducción	35
2.1 Censura de datos	35
2.2 Aproximación de las funciones de supervivencia	37
2.2.1 Métodos no paramétricos	38
2.2.1.1 Método actuarial	38
2.2.1.2 Producto límite de Kaplan Meier	40
2.2.2 Métodos paramétricos	43
2.2.2.1 Algunas distribuciones relacionadas con el	43
análisis de supervivencia	
2.2.2.2 Cálculo de parámetros	47
3. Aplicación a un caso práctico	51
Introducción	51
3.1 Datos a utilizar	51

3.2 Estudio del envejecimiento estocástico de la población	53
por medio de un método no paramétrico	
3.3 Estudio del envejecimiento estocástico de la población	60
por medio de un método paramétrico	
3.4 El estadístico Kolmogorov-Smirnov.....	66
3.5 Posibles aplicaciones de los conceptos de fiabilidad	68
4. Dato histórico	72
Introducción	72
4.1 Contenido del artículo	72
Conclusiones	75
Bibliografía	77

Introducción

La Real Academia Española define la fiabilidad como la probabilidad de buen funcionamiento de algo¹, y podrá notarse que en la vida diaria se asigna de manera automática dicha probabilidad en muchas situaciones, por ejemplo, cuando se da por cierto que una tarea será cumplida en tiempo y forma o que un dispositivo podrá ser utilizado en cualquier instante y operará de manera adecuada o incluso que un ser vivo tendrá cierto número de horas, días o años de vida a partir de un momento, sin embargo, podría decirse que nunca se tiene la certeza de que estas suposiciones se cumplirán y la postura que se adopta al respecto depende de la importancia del asunto o de las consecuencias que se tendrían en caso de no cumplirse las cosas como se espera.

El análisis de fiabilidad es la rama de la estadística que estudia los procesos que involucran fallas de dispositivos o muertes de organismos y las situaciones del párrafo anterior son ejemplos básicos donde se pueden aplicar sus resultados para realizar un estudio con las herramientas adecuadas y poder tomar decisiones con el conocimiento de qué tan probable es que se complete una tarea, o bien se cumpla un tiempo de vida. De este modo el análisis de fiabilidad cobra importancia en la industria y en las ciencias de la salud al disminuir la incertidumbre y estimar resultados futuros.

En muchos casos se hace referencia al análisis de fiabilidad como análisis de supervivencia, pudiendo considerarse la misma rama de la estadística, esto debido a que, aunque la palabra supervivencia haga referencia de manera directa a muertes de organismos, estadísticamente pueden tratarse del mismo modo las muertes de organismos que las fallas de dispositivos. En el presente trabajo se utilizarán ambas expresiones de manera indistinta.

El presente trabajo plantea algunas cuestiones propias del análisis de fiabilidad, como funciones, clasificaciones y resultados teóricos. Una

¹ Real Academia Española [en línea]. Diccionario de la Lengua Española. Vigésima segunda edición, [fecha de consulta: 3 de junio de 2013]. Disponible desde: <<http://www.rae.es/rae.html>>.

herramienta propia de la teoría de probabilidad que es muy utilizada en fiabilidad es el desarrollo de funciones de distribución, las cuales son capaces de describir la probabilidad asociada a una variable aleatoria, que generalmente representará tiempos de muerte o falla.

Después de revisar conceptos se plantean métodos para ajustar una distribución conocida a un conjunto de datos, con esta idea se desarrolla un ejemplo donde se utiliza la información de los tiempos de falla para doce aeronaves.

Una vez aplicados los conceptos de fiabilidad se expone un dato histórico sobre un estudio publicado en el año de 1693 por la Royal Society referente a mortalidad, siendo notorio el interés que ha despertado esta rama de estudio desde hace mucho tiempo.

Finalmente, se exponen conclusiones resultantes de la revisión del presente texto y de la conjunción de las fuentes consultadas para su realización.

1. Conceptos de fiabilidad

Introducción

En el presente capítulo se exponen las definiciones necesarias para desarrollar la teoría relacionada con el análisis de supervivencia partiendo de la concepción del tiempo de vida de un sujeto o la vida útil de un artículo como una variable aleatoria, lo cual da lugar a conceptos como función de densidad, de distribución o fiabilidad, también se plantea la tasa de riesgo y la vida residual media.

A partir de lo anterior se procede a clasificar las funciones de distribución de acuerdo a sus características y, dependiendo de la clase a la que pertenezcan, podrán notarse distintas propiedades.

Las distribuciones que describen la fiabilidad de algunas poblaciones pueden tratarse de una combinación de funciones debido a la existencia de subpoblaciones o a la heterogeneidad², al respecto se expone una posible función de densidad, de supervivencia y de tasa de falla.

Para comparar dos variables aleatorias que representan tiempos de vida, se plantean los conceptos de ordenamiento parcial y envejecimiento relativo, los cuales pueden ser utilizados para tomar decisiones cuando debe elegirse entre distintos artículos.

La tasa de falla brinda información importante respecto a la fiabilidad de un artículo y uno de los casos más versátiles es cuando se presenta con forma de bañera, es decir, siendo una función inicialmente decreciente que a partir de un punto se vuelve creciente³; se presentarán las definiciones relacionadas con la

² Heterogeneidad: mezcla de partes de diversa naturaleza en un todo. Real Academia Española [en línea]. Diccionario de la Lengua Española. Vigésima segunda edición, [fecha de consulta: 21 de agosto de 2013]. Disponible desde: <<http://www.rae.es/rae.html>>.

³ Algunos autores consideran la forma de bañera como una función inicialmente decreciente con un periodo intermedio constante para ser finalmente creciente, por simplicidad no se considerará en algunos desarrollos el periodo constante, además de que sería sencillo modificar las definiciones y resultados para incluirlo.

tasa de falla, así como aproximaciones de los puntos de cambio a los que da lugar la forma de bañera.

Se comentan las relaciones entre la tasa de falla y la vida residual media, así como algunos ejemplos de la forma que toma esta última para distintas distribuciones.

Debido a la forma en que se registran las observaciones en algunos experimentos o la naturaleza misma de distintos fenómenos, es necesario plantear distribuciones para tiempos de falla discretos, lo cual se revisa en el tema 1.9.

Finalmente, es importante comentar la forma en que pueden extenderse los conceptos del escenario univariado al multivariado, esto se revisa en el último tema de este capítulo.

1.1 Algunos conceptos de fiabilidad

Para aprovechar las herramientas que brinda la teoría de la probabilidad y comenzar con los conceptos partícipes en el análisis de fiabilidad, puede considerarse el tiempo de vida de un sujeto o vida útil de un artículo que es objeto de estudio, como una variable aleatoria T .

Por ser una variable aleatoria puede asociársele una función de distribución de probabilidad denotada por $F(t)$ y una de densidad denominada $f(t)$, la relación que guardan es:

$$F(t) = \Pr(T \leq t) = \int_0^t f(u) du$$

En numerosos estudios es más utilizada la llamada función de fiabilidad o supervivencia, la cual expresa la probabilidad de que el tiempo de vida T sea mayor a un valor t , es decir:

$$S(t) = \Pr(T > t) = 1 - \Pr(T \leq t) = 1 - F(t)$$

Un valor representativo de la supervivencia de un artículo puede ser la esperanza de vida (esperanza de la variable aleatoria T), denotada por $E(T) = \mu$.

Además, cuando interesa el comportamiento de la mortalidad en un instante durante un periodo de observación puede recurrirse a la tasa de falla, también llamada función de riesgo o fuerza de mortalidad, que se denota por $r(t)$ y puede calcularse por medio de la siguiente fórmula:

$$r(t) = \frac{f(t)}{S(t)} = -d \ln[S(t)]$$

Una función relacionada con la función de riesgo es la función acumulativa de riesgo, definida por:

$$R(t) = \int_0^t r(u) du = -\ln[S(t)]$$

Los valores de las funciones referidas hasta el momento dependen de t , es decir, de la edad del individuo o del tiempo en operación del artículo en cuestión en el instante de interés, lo cual es de esperarse, ya que en muchas situaciones, la edad de un individuo tiene algún efecto en el tiempo de vida restante, llamado vida residual. Un concepto que facilita la asimilación de estas ideas es el de envejecimiento estocástico, este término se refiere al efecto que tiene la edad de un artículo u organismo en su vida residual, descrito en términos de probabilidades. Se dice que existe una situación de envejecimiento positivo (o envejecimiento adverso) cuando la vida residual tiende a decrecer, en algún sentido probabilístico, al aumentar la edad del componente. También se presentan situaciones donde la vida residual se incrementa, en un sentido probabilístico, al aumentar la edad del individuo o componente, en estos casos se habla de un envejecimiento negativo (o envejecimiento benéfico).

Finalmente, cuando la edad de un artículo no tenga efecto en la distribución de su vida residual se dice que el artículo en cuestión no envejece.

A partir de las consideraciones anteriores surge de manera natural la función de supervivencia condicional $S(x|t)$, que puede calcularse por medio de la fórmula:

$$S(x|t) = \frac{P(X > x | X > t \cap X > t)}{P(X > t)} = \frac{P(X > x + t \cap X > t)}{P(X > t)} = \frac{P(X > x + t)}{P(X > t)}$$

$$= \frac{S(t + x)}{S(t)}$$

y puede entenderse como la probabilidad de que un individuo de edad t sobreviva x unidades adicionales de tiempo.

Además, la vida residual media (MRL) en la edad t es $\mu(t) = E(X - t | X > t)$. Nótese que $\mu(0) = \mu$.

Por medio de una serie de relaciones puede mostrarse que una distribución de tiempo de vida puede ser caracterizada por $S(t)$, $S(x|t)$, $r(t)$, $\mu(t)$ e incluso por el segundo momento de la vida residual $E[(X - t)^2 | X > t]$. Algunas de las igualdades que pueden ejemplificar lo anterior son:

La función de supervivencia en términos de la función de densidad:

$$S(x) = P(X > x) = \int_x^{\infty} f(t) dt$$

Por lo cual, la función de densidad en términos de la función de supervivencia es:

$$f(x) = -\frac{dS(x)}{dx}$$

La función de supervivencia en términos de la función de riesgo es:

$$S(x) = \exp \left[- \int_0^x r(u) du \right]$$

La vida residual media en términos de la función de supervivencia es:

$$\mu(x) = \frac{\int_x^\infty (t-x)f(t)dt}{S(x)} = \frac{\int_x^\infty S(t)dt}{S(x)}$$

El proceso inverso es:

$$S(x) = \int_x^\infty f(t)dt = \exp \left[- \int_0^x r(u) du \right] = \exp[-H(x)] = \frac{\mu(0)}{\mu(x)} \exp \left[- \int_0^x \frac{du}{\mu(u)} \right]$$

En particular, la esperanza de la variable aleatoria que representa el tiempo de vida es:

$$E(X) = \mu = \mu(0) = \int_0^\infty tf(t)dt = \int_0^\infty S(t)dt$$

La función de densidad en términos de la función de vida residual media es:

$$f(x) = -\frac{d}{dx}S(x) = r(x)S(x) = \left(\frac{d}{dx}\mu(x) + 1 \right) \left(\frac{\mu(0)}{\mu(x)^2} \right) \exp \left[- \int_0^x \frac{du}{\mu(u)} \right]$$

La función de riesgo en términos de la vida residual media es:

$$r(x) = -\frac{d}{dx} \ln[S(x)] = \frac{f(x)}{S(x)} = \frac{\left(\frac{d}{dx}\mu(x) + 1 \right)}{\mu(x)}$$

Y la varianza en términos de la función de supervivencia es:

$$Var(X) = 2 \int_0^\infty tS(t)dt - \left[\int_0^\infty S(t)dt \right]^2$$

1.2 Clasificación para funciones de distribución o fiabilidad

La función de distribución F puede ser clasificada por ser o no creciente, también por la forma de su función de riesgo asociada e incluso por su vida residual o su vida residual media. A continuación se presentan algunas definiciones⁴ que servirán para agrupar distintas funciones de distribución en clases de fiabilidad.

Los significados de los acrónimos utilizados son los siguientes:

IFR: Increasing failure rate (tasa de falla creciente)

DFR: Decreasing failure rate (tasa de falla decreciente)

DMRL: Decreasing mean residual life (vida residual media decreciente)

IMRL: Increasing mean residual life (vida residual media creciente)

NBU: New better than used (nuevo mejor que utilizado)

NWU: New worse than used (nuevo peor que utilizado)

IFRA: Increasing failure rate average (tasa de falla creciente en promedio)

DFRA: Decreasing failure rate average (tasa de falla decreciente en promedio)

NBUFR: New better than used in failure rate (nuevo mejor que utilizado en tasa de falla)

NWUFR: New worse than used in failure rate (nuevo peor que utilizado en tasa de falla)

NBUFRA: New better than used in failure rate average (nuevo mejor que utilizado en tasa de falla promedio)

NWUFRA: New worse than used in failure rate average (nuevo peor que utilizado en tasa de falla promedio)

NBUE: New better than used in expectation (nuevo mejor que utilizado en esperanza)

NWUE: New worse than used in expectation (nuevo peor que utilizado en esperanza)

⁴ Las siglas utilizadas para hacer referencia a las definiciones a lo largo de todo el texto provienen de sus respectivas expresiones en inglés, se adopta esta notación debido a que la mayoría de la bibliografía relacionada con el tema se encuentra en este idioma y se facilita así la consulta de otras fuentes al lector.

Y los conceptos son:

Definición 1.1: F es tasa de falla creciente (IFR) si $S(x|t)$ es decreciente en $0 \leq t < \infty$ para toda $x \geq 0$.

Definición 1.2: F es tasa de falla decreciente (DFR) si $S(x|t)$ es creciente en $0 \leq t < \infty$ para toda $x \geq 0$.

Definición 1.3: F es vida residual media decreciente (DMRL) si la función de vida residual media $\mu(t) = \int_0^\infty S(x|t)dx$ es decreciente en t , es decir, $\mu(s) \geq \mu(t)$ para $0 \leq s \leq t$. En otras palabras, cuanto más antiguo es un dispositivo, menor es su vida residual media.

Definición 1.4: F es vida residual media creciente (IMRL) si la función de vida residual media $\mu(t) = \int_0^\infty S(x|t)dx$ es creciente en t , es decir, $\mu(s) \leq \mu(t)$ para $0 \leq s \leq t$.

Definición 1.5: F es nuevo mejor que utilizado (NBU) si $S(x|t) \leq S(x)$, esto es, $S(x+t) \leq S(x)S(t)$ para toda $x, t \geq 0$. Esto significa que un dispositivo de cualquier edad tiene un tiempo de vida remanente estocásticamente más pequeño que el de un dispositivo nuevo.

Definición 1.6: F es tasa de falla creciente en promedio (IFRA) si $-\left(\frac{1}{t}\right)\log S(t)$ es creciente en $t > 0$. Esto es equivalente a $S(\alpha t) \geq S^\alpha(t)$, $0 < \alpha < 1$.

Definición 1.7: F es tasa de falla decreciente en promedio (DFRA) si $-\left(\frac{1}{t}\right)\log S(t)$ es decreciente en $t > 0$.

Definición 1.8: F es nuevo peor que utilizado (NWU) si $S(x+t) \geq S(x)S(t)$ para toda $x, t \geq 0$.

Definición 1.9: F es nuevo mejor que utilizado en tasa de falla (NBUFR) si $r(t) > r(0)$ para toda $t > 0$.

Definición 1.10: F es nuevo peor que utilizado en tasa de falla (NWUFR) si $r(t) < r(0)$ para toda $t > 0$.

Definición 1.11: F es nuevo mejor que utilizado en tasa de falla promedio (NBUFRA) si $r(0) \leq \frac{1}{t} \int_0^t r(x) dx$ para toda $t > 0$. La integral $\int_0^t r(x) dx$ acumula la tasa de falla de 0 a t y al dividir entre la longitud del intervalo $(0, t]$ se obtiene una tasa de falla promedio por unidad de tiempo.

Definición 1.12: F es nuevo peor que utilizado en tasa de falla promedio (NWUFRA) si $r(0) \geq \frac{-\log S(t)}{t}$ para $t > 0$.

Definición 1.13: F es nuevo mejor que utilizado en esperanza (NBUE) si $\int_0^\infty S(x|t) dx \leq \mu$ para $t \geq 0$. Esto se puede interpretar como que un artículo de cualquier edad en particular tiene un tiempo de vida remanente medio más pequeño que el de un artículo nuevo.

Definición 1.14: F es nuevo peor que utilizado en esperanza (NWUE) si $\int_0^\infty S(x|t) dx \geq \mu$ para toda $t \geq 0$.

Aunque parezca inesperado, la característica DFR se presenta frecuentemente; en particular, dos comportamientos de poblaciones que suelen dar lugar a una tasa de falla decreciente son el endurecimiento por deformación (en ingeniería mecánica) y la inmunidad adquirida (en organismos biológicos). También se presenta el fenómeno DFR en la parte inicial de la vida del ser humano, llamada “fase de mortalidad infantil”, donde la edad es en realidad benéfica para el organismo. Una mejora en la fiabilidad puede ocurrir por cambios físicos que causen una auto-mejora o por la heterogeneidad que adquiere la población.

Existen diversas relaciones entre las clases de fiabilidad presentadas, entre ellas, una serie de implicaciones que ha sido notada desde hace muchos años es la siguiente:

$$IFR \Rightarrow IFRA \Rightarrow NBU \Rightarrow NBUE$$

También existe una relación similar para otro grupo de clases de fiabilidad:

$$DFR \Rightarrow DFRA \Rightarrow NWU \Rightarrow NWUE$$

Las definiciones de clases de fiabilidad han dado lugar a muchos resultados teóricos con importantes aplicaciones, un ejemplo es el siguiente teorema.

Teorema: Considérese un sistema operando durante un periodo de tiempo indefinido. En caso de falla, una reparación o reemplazamiento es llevado a cabo, tomando un tiempo despreciable. Los intervalos sucesivos entre fallas son variables aleatorias independientes e idénticamente distribuidas X_1, X_2, \dots de un proceso de renovación. Sea $N(t)$ que denota el número de renovaciones (reemplazamientos) en $(0, t]$ y $M(t)$ el número esperado de renovaciones en $(0, t]$, esto es, $E(N(t)) = M(t)$.

Sea $E(X) = \mu < \infty$ el tiempo de vida medio de un componente. Entonces, bajo el planteamiento anterior:

- a) Si F es NBUE, entonces $\frac{t}{\mu} - 1 \leq M(t) \leq \frac{t}{\mu}$.
- b) Si F es NWUE, entonces $M(t) \geq \frac{t}{\mu}$.

La demostración puede consultarse en el libro *Stochastic Ageing and Dependence for Reliability* de Chin-Diew Lai y Min Xie (2006).

Puede notarse que este teorema brinda una cota para el número esperado de reemplazamientos en términos del tiempo que se espera que opere el sistema y el tiempo de vida promedio de los repuestos utilizados, la aplicación de este

teorema resulta directa al aproximar el número de refacciones con que se debe contar en una operación donde se requiera no detener la producción.

Entre los distintos enfoques para funciones de fiabilidad se encuentra uno que ha resultado útil para estudiar propiedades de la distribución subyacente y también puede ser aplicado para resolver geoméricamente algunos problemas de mantenimiento estocástico, haciendo referencia con mantenimiento estocástico a los procesos de la industria donde la necesidad de una reparación se presenta de manera no determinística e involucra en muchas ocasiones planes de mantenimiento preventivo diseñados con base en estudios estadísticos de los tiempos de falla en el sistema. Este planteamiento se basa en el concepto de tiempo total a prueba (TTT) y el desarrollo es el siguiente:

Sea F una distribución de tiempo de vida, se define

$$F^{-1}(p) = \inf\{x: F(x) \geq p\}, \quad p \in [0,1]$$

Esto es, una especie de función inversa para $F(x)$, donde el valor de interés es un ínfimo, ya que $F(x)$ es una función acumulativa.

Defínase también

$$H_{F^{-1}}(p) = \int_0^{F^{-1}(p)} S(x)dx, \quad p \in [0,1]$$

Si el tiempo de vida medio μ es finito, entonces

$$H_{F^{-1}}(1) = \int_0^{F^{-1}(1)} S(x)dx = \mu$$

La transformación de tiempo total a prueba escalada está definida por

$$\phi(p) = \frac{H_{F^{-1}}(p)}{H_{F^{-1}}(1)} = \frac{H_{F^{-1}}(p)}{\mu}$$

La curva $\phi(p)$ contra $p \in [0,1]$ es llamada la curva TTT escalada.

Finalmente, si F es una distribución de tiempo de vida continua, la clasificación que se puede obtener de este planteamiento es:

- i) F es IFR (DFR) si y sólo si $\phi(p)$ es cóncava (convexa) en $p \in [0,1]$.
- ii) F es IFRA (DFRA) si y sólo si $\frac{\phi(p)}{p}$ es decreciente (creciente) en $p \in [0,1]$.
- iii) F es NBUE (NWUE) si y sólo si $\phi(p) \geq p$ ($\phi(p) \leq p$) para $p \in [0,1]$.
- iv) F es DMRL (IMRL) si y sólo si $\frac{1-\phi(p)}{1-p}$ es decreciente (creciente) en $p \in [0,1]$.

Cabe considerar que algunos autores han definido más clases de envejecimiento que también pueden ser clasificadas con este planteamiento.

1.3 El papel de la función de riesgo $r(t)$

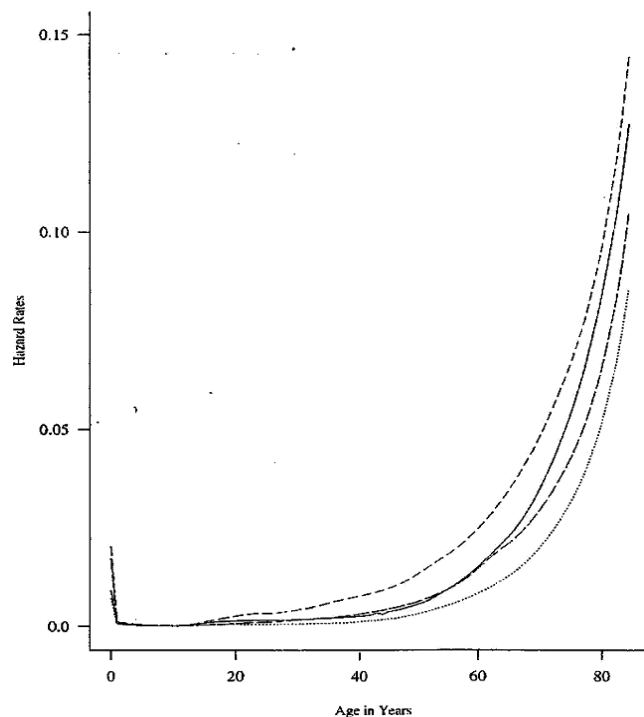
La gráfica de la función de riesgo es muy útil para entender el comportamiento del envejecimiento de un organismo o artículo a lo largo de su vida, particularmente cuando interesa saber en qué momento se tiene un máximo o un mínimo riesgo (probabilidad) de que el organismo muera o el artículo falle. Una caracterización simple de la función de riesgo se presenta a continuación, debe notarse que algunos de los incisos se relacionan de manera directa con las clasificaciones para la función de distribución que se presentaron en la sección 1.2.

Asumiendo que la función de riesgo $r(t)$ es diferenciable, $r(t): R^+ \rightarrow R^+$ y que por creciente se entiende no decreciente y decreciente se refiere a no creciente, se dice que $r(t)$ es:

- Estrictamente creciente si $r'(t) > 0$ para toda t y se denota por I.
- Estrictamente decreciente si $r'(t) < 0$ para toda t y se denota por D.
- Con forma de bañera si $r'(t) < 0$ para $t \in (0, t_0)$, $r'(t_0) = 0$ y $r'(t) > 0$ para $t > t_0$, se denota por BT.
- Con forma de bañera invertida si $r'(t) > 0$ para $t \in (0, t_0)$, $r'(t_0) = 0$ y $r'(t) < 0$ para $t > t_0$, se denota por UBT.
- Con forma de bañera modificada si $r(t)$ es al principio creciente y después tiene forma de bañera, se denota por MBT.

Debe observarse que algunos autores incluyen que $r(t)$ sea constante en un intervalo para las definiciones BT y UBT, se prescindirá de esa característica, aunque para adoptarla sería sencillo modificar la clasificación anterior.

De manera ilustrativa se presenta la siguiente imagen del libro *Survival Analysis Techniques for Censored and Truncated Data* de John P. Klein y Melvin L. Moeschberger (1997), donde se muestra la función de riesgo por todas las causas de mortalidad en la población de Estados Unidos separada por color de piel.



Hazard functions for all cause mortality for the US population in 1989. White males (—); white females (·····); black males (- - - -); black females (— — — —).

Se puede notar en la gráfica una etapa inicial decreciente, seguida de una sección con función de riesgo aproximadamente constante para terminar con una función de riesgo creciente, es decir, dichas funciones de riesgo tienen forma de bañera con intervalo constante.

Los tiempos de supervivencia y falla son a menudo modelados por distribuciones con tasas de falla crecientes o decrecientes. Aunque esto es apropiado para muchos casos, resulta inapropiado si el comportamiento del fenómeno es tal que la mortalidad alcanza un máximo después de un tiempo finito y después decrece lentamente. Por ejemplo, en un estudio sobre la curabilidad del cáncer de mama llevado a cabo por Allan O Langlands, Stuart J Pocock y otros investigadores en 1979, se encontró que el punto más alto de la mortalidad ocurría después de aproximadamente tres años, lo cual sirve para ilustrar que fenómenos con tasa de falla con forma de bañera invertida pueden presentarse en este tipo de estudios.

1.4 Combinaciones de distribuciones

Suelen surgir en poblaciones heterogéneas. Un caso típico es cuando una población consiste de dos subpoblaciones. El comportamiento de las distribuciones tiene algunas particularidades, por ejemplo, las combinaciones de distribuciones DFR siempre son DFR, sin embargo, algunas combinaciones de distribuciones IFR pueden resultar DFR.

Una función de densidad que se puede asociar a una combinación de dos subpoblaciones con funciones de densidad f_1 y f_2 , respectivamente, es:

$$f(t) = pf_1(t) + (1 - p)f_2(t), \quad t \geq 0, \quad 0 \leq p \leq 1$$

A partir de esta función de densidad puede deducirse la función de fiabilidad como:

$$S(t) = pS_1(t) + (1 - p)S_2(t)$$

La tasa de falla puede ser expresada como:

$$r(t) = \frac{pf_1(t) + (1-p)f_2(t)}{pS_1(t) + (1-p)S_2(t)}$$

donde $f_i(t)$ y $S_i(t)$ son las funciones de densidad y de supervivencia de las distribuciones con tasas de falla $r_i(t)$, $i = 1, 2$.

1.5 Ordenamientos parciales

Pueden compararse dos variables que representan tiempos de vida X y Y de varias formas, en términos de las tasas de falla $r_F(t)$ y $r_G(t)$, de las funciones de densidad $f(t)$ y $g(t)$, de las funciones de supervivencia $S_F(t)$ y $S_G(t)$, de las vidas residuales medias $\mu_F(t)$ y $\mu_G(t)$ o a partir de otras características de envejecimiento.

Se presenta a continuación una serie de ordenamientos con definiciones sencillas y una interpretación intuitiva, se utiliza la notación del párrafo anterior para hacer referencia a las tasas de falla, funciones de densidad y supervivencia, así como a la vida residual media de cada una de las funciones involucradas.

Definición 1.15: X es mayor que Y en el ordenamiento de razón de verosimilitud ($X \geq_{LR} Y$) si $\frac{f(t)}{g(t)}$ es decreciente en $t \geq 0$.

Definición 1.16: X es mayor que Y en el ordenamiento de tasa de falla ($X \geq_{FR} Y$) si $r_F(t) \leq r_G(t)$ para toda $t \geq 0$ o $\frac{S_F(t)}{S_G(t)}$ es creciente en $t \geq 0$.

Definición 1.17: X es mayor que Y en el ordenamiento estocástico ($X \geq_{ST} Y$) si $S_F(t) \geq S_G(t)$ para toda $t \geq 0$.

Definición 1.18: X es mayor que Y en el ordenamiento de vida residual media ($X \geq_{MR} Y$) si $\mu_F(t) \geq \mu_G(t)$ para toda $t \geq 0$. Se ha deducido que $X \geq_{MR} Y$ si y sólo si $\frac{\int_t^\infty S_F(x)dx}{\int_t^\infty S_G(x)dx}$ es creciente en $t \geq 0$.

Algunas relaciones entre los ordenamientos anteriores son las siguientes:

- $r_F(t) \leq r_G(t) \Rightarrow \mu_G(t) \leq \mu_F(t)$, lo inverso en general no es cierto.
- Suponiendo que $\mu_G(t) \leq \mu_F(t)$ y $\frac{\mu_G(t)}{\mu_F(t)}$ es una función no decreciente para toda $x \geq 0$, entonces $r_F(t) \leq r_G(t)$.

Existen ejemplos donde se muestra que la comparación en términos de la tasa de falla, es aquello a lo que en realidad se refiere la gente cuando compara el comportamiento de dos productos, esto en términos de envejecimiento estocástico hace referencia al envejecimiento relativo.

1.6 Envejecimiento relativo

Definición 1.19: Sean F y G las funciones de distribución de las variables aleatorias X e Y , respectivamente, se dice que X envejece más rápido que Y , lo cual se denota por $X <_c Y$, si la variable aleatoria $\Lambda_G(X) = -\log S_G(X)$ tiene una tasa de falla creciente, esto es, si las tasas de falla $r_F(t)$ y $r_G(t)$ existen, con $r_F(t) = \frac{f(t)}{1-F(t)}$ y $r_G(t) = \frac{g(t)}{1-G(t)}$, entonces X envejece más rápido que Y si $\frac{r_F(t)}{r_G(t)}$ es una función creciente de t .

Existen algunos resultados basados en el concepto de envejecimiento relativo, los cuales pueden ser utilizados para asignar repuestos e identificar fallas cuando dos componentes (o sistemas) tienen la misma media. Uno de estos

resultados dice que si X envejece más rápido que Y y ambas tienen la misma media, entonces $\text{var}(X) \leq \text{var}(Y)$.

1.7 Distribuciones de vida con tasa de falla en forma de bañera

Los conceptos de envejecimiento monótono han resultado populares entre muchos ingenieros encargados de estudiar la fiabilidad. Sin embargo, existen aplicaciones prácticas donde el efecto del envejecimiento es inicialmente benéfico (una fase donde existe envejecimiento negativo) pero después de cierto periodo inicia una fase donde el envejecimiento es positivo. Conjuntos de datos y la experiencia de distintas industrias indican que muchos productos, en particular los artículos electrónicos, muestran una tasa de falla en forma de bañera.

Se dice que F tiene tasa de falla en forma de bañera (es BT) si su función de riesgo decrece al principio, después se mantiene constante por un periodo para finalmente ser creciente. En otras palabras, $r(t)$ tiene la forma de una bañera. La forma de bañera también suele asignarse a las funciones de riesgo inicialmente decrecientes que después de un punto son crecientes. En general, es sencillo adoptar la teoría de acuerdo a la forma que se designe como bañera.

Durante el periodo inicial de vida las fallas suelen ser causadas por defectos de fabricación o de nacimiento en el caso de seres humanos. Después de este periodo las fallas se pueden considerar aleatorias y al final de la vida la función de riesgo crece debido a que al ser más viejo el dispositivo es más propenso a fallar.

Cabe notar que no existe un consenso sobre quién utilizó por primera vez el término “bañera” para esta clase de distribuciones.

Definición 1.20: Sea F una función de distribución acumulativa con función de tasa de falla continua $r(t)$. Entonces F es BT si existe un t_0 tal que $r(t)$ es

decreciente para $t < t_0$, $r(t)$ es creciente para $t > t_0$, es decir, $r'(t) < 0$ para $t < t_0$, $r'(t_0) = 0$ y $r'(t) > 0$ para $t > t_0$.

En la definición anterior se supone continuidad y dado que este requisito no siempre se cumple puede definirse el concepto BT a través de la función de fiabilidad condicional:

$$S(x|t) = \frac{S(t+x)}{S(t)}, \quad S(t) = 1 - F(t) > 0.$$

Definición 1.21: F es BT si existe un t_0 tal que

- a) $S(x|t)$ es estrictamente creciente en t para $0 \leq t < t_0, 0 \leq x \leq t_0 - t$,
- b) $S(x|t)$ es estrictamente decreciente en t para $t_0 \leq t < \infty, x \geq 0$.

Los puntos de cambio de $r(t)$ son importantes en las distribuciones con tasa de falla no-monótona, la estimación de los mismos en las distribuciones BT es particularmente relevante en el contexto de políticas de mantenimiento, dado que no se quisiera reemplazar un componente antes de que su edad haya pasado el punto de cambio de su función de tasa de falla; si hubiera dos puntos de cambio se buscaría no reemplazar el componente hasta que haya pasado el segundo. En general, la aproximación del punto de cambio es una función de los parámetros del modelo, por lo tanto, un enfoque paramétrico para su aproximación es lo más adecuado.

Otro punto de cambio que puede ser de interés es el de $\mu(t)$, la función de vida residual media, y al respecto varios autores han notado que dicho punto se presenta antes que el de la función de tasa de falla. En otras palabras, el tiempo que minimiza la función de riesgo en forma de bañera no maximiza $\mu(t)$. La función de vida residual media en el intervalo constante (cuando este existe) de la tasa de falla en forma de bañera no es constante, sino decreciente.

Una aplicación directa de estos conceptos surge en las fases de prueba a que son sometidos algunos artículos antes de estar a la venta, en estos periodos se procura simular las condiciones en que estarán operando dichos artículos y

aquellos componentes que fallen durante el periodo de prueba pueden ser desechados o reparados, mientras que los que sobrevivan a este procedimiento serán considerados de buena calidad, sólo entonces son puestos al alcance de los consumidores. Un periodo de prueba insuficiente genera altas tasas de falla iniciales que dan lugar a altos costos de reparación. Un periodo de prueba excesivo puede brindar una baja función de riesgo pero hará que el productor incurra en mayores gastos. Así que un problema importante es encontrar el tiempo de prueba óptimo.

Existen distintos enfoques para responder a esta cuestión, uno de ellos es maximizar la fiabilidad para un tiempo de tarea dado. Supóngase que T es el tiempo de vida de un componente con tasa de falla $r(t)$ en forma de bañera y se espera que se cumpla una tarea que dura por un tiempo τ , la fiabilidad de completarla será entonces $S(\tau)$. Si se somete el componente a un periodo de prueba por un tiempo b y el mismo sobrevive, entonces la fiabilidad condicional de completar la tarea original está dada por:

$$\frac{S(b + \tau)}{S(b)} = \exp\left(-\int_b^{b+\tau} r(t)dt\right)$$

Para determinar entonces el tiempo de prueba óptimo se busca el valor que maximice la función anterior. Para esto puede definirse el conjunto de tiempos de prueba posibles como:

$$B^* = \left\{ b \geq 0: \frac{S(b + \tau)}{S(b)} = \max_{t \geq 0} \frac{S(t + \tau)}{S(t)} \right\}$$

O

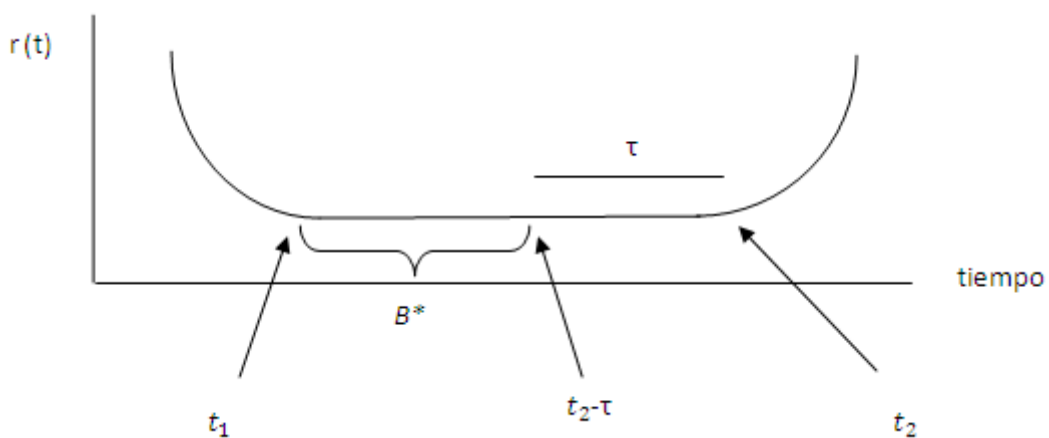
$$B^* = \left\{ b \geq 0: \int_b^{b+\tau} r(s)ds = \min_{t \geq 0} \int_t^{t+\tau} r(s)ds \right\}$$

El siguiente teorema brinda información respecto al tiempo de prueba óptimo.

Teorema: Sea una función de tasa de falla continua $r(t)$ con forma de bañera y puntos de cambio t_1 y t_2 , y sea $\tau > 0$ el tiempo requerido para cumplir una tarea, entonces:

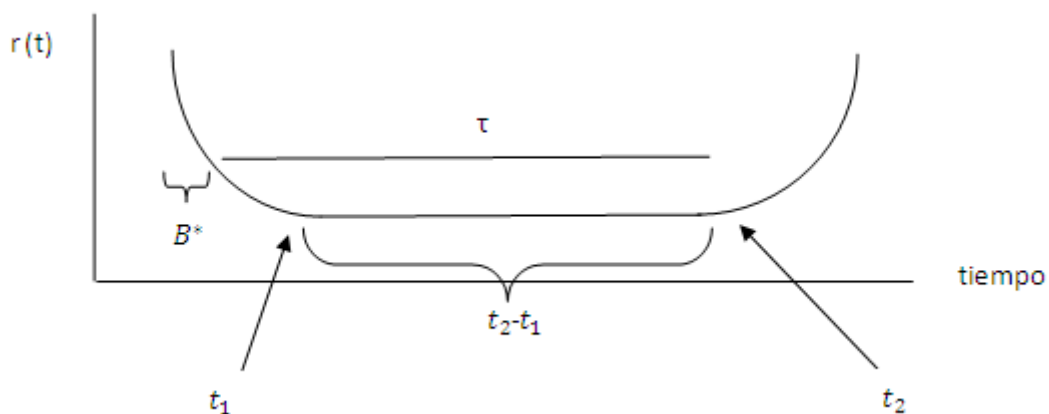
- a) Si $\tau \leq t_2 - t_1$, el tiempo de prueba óptimo ocurre en cada punto del intervalo $[t_1, t_2 - \tau]$, es decir, $B^* = [t_1, t_2 - \tau]$.
- b) Si $\tau > t_2 - t_1$, el tiempo de prueba óptimo ocurre en o antes del primer punto de cambio t_1 , esto es, $B^* = \{b^*\}$ y $b^* \in [0, t_1]$.

El inciso a) se puede visualizar por medio del siguiente diagrama:



Donde se puede notar que en t_1 inicia el intervalo donde la tasa de falla ya no puede ser menor y además es constante, aquí inicia B^* , y termina en el punto $t_2 - \tau$, porque en cualquier punto t posterior a $t_2 - \tau$ sucede que al evaluar la función de riesgo en $t + \tau$, $r(t)$ es creciente.

Para el inciso b) puede observarse el siguiente diagrama:



Donde, al ser τ mayor que $t_2 - t_1$, la longitud del intervalo donde la función de riesgo es constante y mínima, el tiempo de prueba óptimo debe ubicarse en o antes de t_1 para que al sumar el tiempo de prueba óptimo y τ , la función de riesgo se evalúe en todo el intervalo donde es constante, más algún intervalo donde sea decreciente o creciente.

Para más referencias sobre la demostración puede consultarse el libro *Stochastic Ageing and Dependence for Reliability* de Chin-Diew Lai y Min Xie (2006).

Otro enfoque para elegir el periodo de prueba óptimo es tomando como referencia el tiempo de vida útil del producto en cuestión. Supóngase que el artículo es puesto en operación por un periodo fijo llamado b y, de no considerarse el costo, puede tomarse como meta el maximizar el tiempo de vida medio. Es decir, se requiere determinar el valor de b de tal manera que la vida residual media es maximizada, y sólo aquellos artículos que sobrevivan a este tiempo de prueba serán puestos en operación. En otras palabras, se busca una b^* tal que

$$\mu(b^*) = \max_{b \geq 0} \{\mu(b)\}$$

Un resultado útil basado en este enfoque se enuncia en el siguiente teorema.

Teorema: Supóngase que $r(t)$ tiene forma de bañera, es derivable y tiene dos puntos de cambio t_1 y t_2 . Entonces:

- Si $t_1 = 0$, no se necesita un periodo de prueba, ya que $b^* = 0$.
- Si $t_2 = \infty$ y $t_1 > 0$, siempre puede tomarse $b^* = t_1$.
- Si $t_1 = t_2 = \infty$, es decir, F es en realidad DFR, el costo del tiempo de prueba debe ser considerado.
- Si $0 < t_1 \leq t_2 < \infty$, entonces b^* es igual al único punto de cambio t^* de $\mu(t)$.

Puede notarse que nunca es necesario probar los artículos más allá del primer punto de cambio, salvo en el caso de que F sea DFR (tasa de falla decreciente).

Para referencias sobre la demostración puede consultarse el libro *Stochastic Ageing and Dependence for Reliability* de Chin-Diew Lai y Min Xie (2006).

Una cuestión similar al tiempo de prueba óptimo surge con el momento en que se debe reemplazar un componente. En este caso, uno de los enfoques que se pueden utilizar es sustituirlo cuando su edad ha alcanzado un valor previamente decidido. La pregunta a resolver es cuál es la edad óptima para llevar a cabo el reemplazamiento, y un criterio puede ser que el artículo sea cambiado cuando la tasa de falla rebase cierto valor h_r . Con estas consideraciones, la edad de reemplazamiento óptima puede ser determinada al resolver la ecuación $r(t) = r_b$, y puede encontrarse la solución por métodos numéricos.

Además de estos criterios para elegir tiempos de prueba óptimos pueden considerarse otros que han sido planteados por distintos autores, por ejemplo:

- La tasa de falla ha sido utilizada tradicionalmente al tomar decisiones sobre tiempos de prueba debido a su interpretación como fuerza de mortalidad.
- Puede usarse como criterio la maximización de la probabilidad de supervivencia para un tiempo de tarea dado.
- Se puede determinar el tiempo de prueba óptimo al minimizar los costos asociados con las políticas de garantía y con los periodos de prueba.
- Otro criterio posible es balancear la vida residual media y la varianza residual.

Se ha señalado previamente que el instante donde la tasa de falla es mínima no suele coincidir con el instante donde la MRL es máxima y obtener información de la vida residual a partir de la tasa de falla o viceversa no siempre es directo, sin embargo, la función de distribución gaussiana inversa

brinda un ejemplo donde el instante en que $r(t)$ alcanza su máximo coincide con el momento en que la vida residual media $\mu(t)$ logra su mínimo.

La función de distribución gaussiana inversa tiene la siguiente forma:

$$F(t) = \Phi \left\{ \sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right\} + e^{2\lambda/\mu} \Phi \left\{ -\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right\}$$

Esta distribución es una de las más utilizadas para ajustar datos relativos al coste de los siniestros.⁵

1.8 El papel de la vida residual media

Recordando que la vida residual media se define formalmente como $\mu(t) = E(X - t | X > t)$, puede obtenerse la siguiente igualdad:

$$\mu(t) = E(X - t | X > t) = \frac{\int_t^{\infty} S(x) dx}{S(t)}$$

En estudios de fiabilidad industrial o estrategias de reemplazamiento o reparación la función de vida residual media puede proporcionar información más relevante que $r(t)$, esto se debe a que la primera resume toda la distribución de vida residual, mientras que la segunda describe solamente el riesgo de falla inmediata.

Es fácil mostrar que la función de vida residual media determina la distribución de manera única, puede comenzarse con la siguiente igualdad:

$$S(t) = \frac{\mu}{\mu(t)} \exp \left\{ -\int_0^t \mu(x)^{-1} dx \right\}, t \geq 0$$

⁵ Gómez Déniz Emilio. Notas para el curso de doctorado en economía financiera cuantitativa y actuarial. Modelización de riesgo e incertidumbre en seguros y auditoría contable. [fecha de consulta: 12 de agosto de 2013]. Disponible desde: <www.ulpgc.es>

Derivando la ecuación respecto a t se obtiene la siguiente igualdad:

$$f(t) = \frac{\mu(\mu'(t) + 1)}{\mu(t)^2} \exp\left\{-\int_0^t \mu(x)^{-1} dx\right\}, t \geq 0$$

De este modo es posible determinar una función de manera única a través de su vida residual media. En particular, F es exponencial si y sólo si su MRL es constante, esto es, $\mu(t) = \mu$ para toda $t \geq 0$.

Puede entonces recordarse que la función de supervivencia $S(t)$, la tasa de falla $r(t)$ y la vida residual media $\mu(t)$ son equivalentes, en el sentido de que conociendo cualquiera de ellas las otras pueden ser determinadas cuando existen.

Las funciones de vida residual media pueden ser muy simples o demasiado complicadas, al respecto tómesese el ejemplo de la distribución exponencial, la cual, debido a la propiedad de pérdida de memoria, tiene una función de vida residual igual a su media, es decir,

$$\mu(t) = \frac{1}{\lambda}$$

La distribución Pareto ha sido ampliamente usada para modelar datos de tiempos de vida, una de las razones es que su función de vida residual media tiene una forma lineal simple. La función de supervivencia de la distribución Pareto está dada por:

$$S(t) = \left[\frac{A}{A + Bt} \right]^{\frac{1}{B} + 1}$$

y su función de vida residual media tiene la siguiente forma:

$$\mu(t) = A + Bt$$

Sin embargo, también existe el ejemplo de la distribución gaussiana inversa, la cual tiene la siguiente función de distribución:

$$F(t) = \Phi \left\{ \sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right\} + e^{2\lambda/\mu} \Phi \left\{ -\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right\}$$

y su MRL está dada por:

$$\mu(t) = \frac{(\mu - t)\Phi \left(\sqrt{\frac{\lambda}{t}} \left(1 - \frac{t}{\mu} \right) \right) + (\mu - t)^2 e^{\frac{2\lambda}{\mu}} \Phi \left(-\sqrt{\frac{\lambda}{t}} \left(1 + \frac{t}{\mu} \right) \right)}{\Phi \left(\sqrt{\frac{\lambda}{t}} \left(1 - \frac{t}{\mu} \right) \right) + e^{\frac{2\lambda}{\mu}} \Phi \left(-\sqrt{\frac{\lambda}{t}} \left(1 + \frac{t}{\mu} \right) \right)}$$

lo cual muestra lo complicada que puede ser una función de vida residual media.

Un ejemplo de la aplicación del concepto de vida residual media es cuando se estudia la permanencia de empleados en una empresa, al respecto puede decirse que la MRL de un empleado con cuatro años en la empresa es mayor que la correspondiente a un empleado con dos meses de antigüedad, ya que el primero ha invertido más tiempo y se ha especializado en las actividades que desempeña; nótese también que después de un periodo inicial de IMRL (vida residual media creciente), la permanencia de los empleados entra en una etapa de DMRL (vida residual media decreciente) al estar próximos a las edades de retiro.

1.9 Modelos con tiempos de falla discretos

Muchos de los tiempos de falla son continuos en la naturaleza y distintas distribuciones han sido propuestas en la literatura, sin embargo, los datos de falla discretos también surgen en muchas situaciones comunes, por ejemplo:

- Los datos en algunos experimentos son reportados de manera semanal o mensual y las observaciones son el número de fallas en cada periodo sin importar los tiempos de ocurrencia.
- Algunos sistemas operan de manera cíclica y las observaciones consisten en el número de ciclos completados exitosamente antes de presentarse una falla.

Sea la variable aleatoria X con soporte $N^+ = \{1, 2, \dots\}$ que representa el tiempo de vida discreto de un componente, se denota entonces por $f(k)$ la probabilidad de que una falla ocurra en el instante k , es decir,

$$f(k) = \Pr \{X = k\}, \quad k = 1, 2, \dots$$

La función de fiabilidad que corresponde a X está dada por

$$R(k) = \Pr\{X > k\} = \sum_{j=k+1}^{\infty} f(j), \quad k = 1, 2, \dots$$

La función de tasa de falla $r(k)$ está definida como

$$r(k) = P(X = k | T \geq k) = \frac{P(X=k)}{P(X \geq k)} = \frac{f(k)}{R(k-1)}, \quad \text{siempre que } \Pr(T \geq i) > 0.$$

La ecuación anterior puede ser también expresada como

$$r(k) = \frac{R(k-1) - R(k)}{R(k-1)}$$

Es importante señalar que algunos autores han definido la función $R(k)$ como $P(X \geq k)$ en lugar de $P(X > k)$ como se ha hecho previamente.

Puede decirse que las clases de envejecimiento discretas se definen de una manera análoga a sus contrapartes continuas, por ejemplo:

Definición 1.22: F es IFR si $r(k)$ es creciente en $k = 1, 2, \dots$

Definición 1.23: F es DFR si $r(k)$ es decreciente en $k = 1, 2, \dots$

Las distribuciones de tiempos de falla discretos tienen muchas aplicaciones, por ejemplo, ha sido notado que las distribuciones DFR pueden modelar el número de temporadas que un programa de televisión es transmitido antes de ser cancelado. Otro ejemplo se encuentra en los estudios de fiabilidad sobre armas de fuego, donde el número de detonaciones antes de fallar es más importante que la edad del arma al momento de la falla.

Existen algunas diferencias entre los conceptos aplicables para tiempos de falla continuos y discretos, una de las principales se encuentra en la definición y características de la función de tasa de falla.

Por ejemplo $r(k)$ en el caso discreto es la probabilidad condicional de fallar al momento $X = k$ dado que el artículo no ha fallado hasta el instante $k - 1$ y, en consecuencia, $r(k) \leq 1$, sin embargo, en el caso continuo, $r(t)$ puede ser no acotada en algunas situaciones. Más aún, algunos autores han comentado que $r(k)$ para el caso discreto no puede crecer de manera exponencial, lo cual es común para muchos componentes al final de su vida útil.

Curiosamente, la tasa de falla para el caso continuo es comúnmente aproximada como la proporción de dispositivos que fallan en un intervalo por unidad de tiempo dado que han sobrevivido hasta el inicio del intervalo.

Otra característica de la definición de $r(k)$ considerada se encuentra al notar que no es aditiva para sistemas en serie⁶ de componentes independientes. Para señalar esto tómesese $r_i(k)$ como la tasa de falla del i -ésimo componente, así que la función de riesgo del sistema será:

$$r(k) = \frac{R(k-1) - R(k)}{R(k-1)} = \frac{\prod_{i=1}^n R_i(k-1) - \prod_{i=1}^n R_i(k)}{\prod_{i=1}^n R_i(k-1)}$$

⁶ Se toman por sistemas en serie aquellos arreglos de dispositivos que trabajan en conjunto donde se presenta la falla de todo el sistema al fallar cualquiera de sus componentes.

$$= 1 - \prod_{i=1}^n \frac{R_i(k)}{R_i(k-1)} = 1 - \prod_{i=1}^n [1 - r_i(k)] \neq \sum_{i=1}^n r_i(k)$$

Y una diferencia más que puede notarse es que la función acumulativa de riesgo $H(k) = \sum_{i=1}^k r(i)$ no es equivalente a $-\log R(k)$, lo cual sí se cumple en el caso continuo.

Dadas estas diferencias en las características de la función de tasa de falla como se ha definido, algunos autores han planteado una definición alternativa que se denotará por $r^*(k)$.

Definición 1.24: Para una distribución discreta con función de fiabilidad $R(k)$ se define la función de tasa de falla alternativa $r^*(k)$ como:

$$r^*(k) = \log \frac{R(k-1)}{R(k)}, \quad k = 1, 2, \dots$$

La razón de esta nueva definición es que la función de tasa de falla para el caso continuo se puede definir como:

$$r(t) = \frac{f(t)}{R(t)} = -\frac{d}{dt} \log R(t)$$

Y en lugar de usar $R(k-1) - R(k)$ para $f(k)$ como se ha hecho, puede utilizarse $\log R(k-1) - \log R(k)$ para $-d[\log R(t)]/dt$ y definir la tasa de falla como:

$$r^*(k) = -[\log R(k) - \log R(k-1)] = -\log \frac{R(k)}{R(k-1)} = \log \frac{R(k-1)}{R(k)}$$

Puede notarse que $r^*(k)$ no es acotada.

Estudiando la relación que guardan $r(k)$ y $r^*(k)$ puede notarse lo siguiente:

$$r^*(k) = -\log \frac{R(k)}{R(k-1)} = -\log \frac{R(k)-R(k-1)}{R(k-1)} = \log[1 - r(k)], \text{ o también}$$

$$r(k) = 1 - e^{-r^*(k)}$$

Y dada esta relación puede decirse que $r^*(k)$ es creciente (decreciente) si y sólo si $r(k)$ es creciente (decreciente), de este modo pueden definirse las mismas clases de envejecimiento IFR (DFR) discretas a través de $r^*(k)$.

Puede continuarse el análisis de esta nueva definición de la función de tasa de falla al ver el efecto que tiene en la función acumulativa de riesgo.

$$H^*(k) = r^*(1) + r^*(2) + r^*(3) + \dots + r^*(k), \text{ lo cual resulta:}$$

$$H^*(k) = \log \frac{R(0)}{R(1)} + \log \frac{R(1)}{R(2)} + \dots + \log \frac{R(k-1)}{R(k)} = \log \frac{R(0)}{R(k)} = -\log R(k).$$

Donde se puede notar que la propiedad que se cumplía para el caso continuo ahora se mantiene en un enfoque discreto ($H^*(k) = -\log R(k)$).

Es importante señalar que también se cumplirá ahora la propiedad aditiva para sistemas en serie de componentes independientes. Esto se muestra a través del siguiente razonamiento.

Supóngase un sistema en serie de n componentes independientes con tasa de falla $r_j^*(k), j = 1, 2, \dots, n$, entonces la fiabilidad del sistema está dada por:

$$\begin{aligned} R(k) &= \prod_{j=1}^n R_j(k) = \prod_{j=1}^n \exp\{-\sum_{i=1}^k r_j^*(i)\} = \exp\{-\sum_{j=1}^n \sum_{i=1}^k r_j^*(i)\} \\ &= \exp\{-\sum_{i=1}^k \sum_{j=1}^n r_j^*(i)\} = \exp\{-\sum_{i=1}^k r^*(i)\} \end{aligned}$$

Donde $r^*(i)$ es la función de tasa de falla para el sistema. Ahora puede notarse que

$$r^*(i) = \sum_{j=1}^n r_j^*(i)$$

Y así con la nueva definición de función de tasa de falla se cumple la propiedad aditiva para sistemas en serie de componentes independientes.

Por último, se presenta un ejemplo del cálculo de la función de tasa de falla con la definición alternativa.

Considérese la distribución Pareto con función de fiabilidad

$$R(k) = \left(\frac{d}{k+d}\right)^c; \quad c, d > 0, \quad k \geq 1$$

La función de tasa de falla con la definición alternativa resulta:

$$r^{*(k)} = \log \frac{R(k-1)}{R(k)} = \log \frac{\left(\frac{d}{k-1+d}\right)^c}{\left(\frac{d}{k+d}\right)^c} = c \log \frac{k+d}{k-1+d} = c \log \left[1 + \frac{1}{k-1+d}\right]$$

1.10 Envejecimiento bivariado y multivariado

Es útil generalizar los conceptos de envejecimiento utilizados hasta el momento a un enfoque bivariado o multivariado, ya que muchos sistemas están compuestos de varios componentes con tiempos de vida dependientes.

Puede comenzarse definiendo algunos conceptos de envejecimiento bivariado, los cuales se extienden de manera natural para más variables.

Se denota por $S(x_1, x_2)$ la probabilidad de que el artículo 1 sobreviva más allá del tiempo x_1 y el artículo 2 sobreviva después de x_2 . Nótese que $S(x_1, x_2) \neq 1 - F(x_1, x_2)$, sin embargo, la función de supervivencia conjunta

$S(x_1, x_2)$ está relacionada con la distribución $F(x_1, x_2)$ por medio de la siguiente ecuación:

$$S(x_1, x_2) = \Pr(X_1 > x_1, X_2 > x_2) = 1 - F_{X_1}(x_1) - F_{X_2}(x_2) + F(x_1, x_2)$$

Algunas características que pueden ser consideradas al extender las definiciones del caso univariado al bivariado o multivariado son las siguientes:

- Deberían basarse en condiciones impuestas a la función de supervivencia conjunta, más que a las variables aleatorias.
- Tendrían que coincidir con aquellas aceptadas para una variable.
- Buscarían dar lugar a una serie de implicaciones y contenciones como la que relaciona las definiciones en el caso univariado.
- Los argumentos del planteamiento debieran ser extensiones naturales de aquellos utilizados en el caso univariado.

A continuación se presentan algunas definiciones de envejecimiento multivariado, cabe señalar que existen variantes para cada concepto dependiendo del enfoque utilizado por distintos autores.

Definición 1.25: Considérese la cantidad $\frac{S(x_1+t_1, x_2+t_2)}{S(x_1, x_2)}$ la cual puede ser interpretada como la probabilidad conjunta de sobrevivir t_i unidades adicionales de tiempo ($i = 1, 2$) dado que el componente i ha sobrevivido hasta el instante x_i . Entonces se dice que F es una distribución bivariada con tasa de falla creciente (IFR) si $S(x_1 + t_1, x_2 + t_2)/S(x_1, x_2)$ es decreciente en $x_1 \geq 0, x_2 \geq 0$ para todo $t_1 \geq 0, t_2 \geq 0$.

Definición 1.26: Sea $T(X, Y)$ un vector no negativo con función de supervivencia $S(t_1, t_2) = \Pr(X > t_1, Y > t_2)$ entonces el vector T es IFRA (tasa de falla creciente en promedio) si:

$$E^\alpha[h(X, Y)] \leq E[h^\alpha(X, Y)/\alpha]$$

para todas las funciones h crecientes, no negativas y continuas y para toda α tal que $0 < \alpha < 1$.

Definición 1.27: Se dice que la distribución $F(X, Y)$ es DMRL (vida residual media decreciente) si $\frac{\int_0^\infty S(x_1+t, x_2+t) dt}{S(x_1, x_2)}$ es decreciente en $x_1, x_2 \geq 0$.

Existen más definiciones para otras clases de envejecimiento, e incluso distintos planteamientos para cada concepto, sin embargo, en la mayoría de los casos, la extensión del concepto univariado al multivariado es directa.

A continuación se presentan tres ejemplos de distribuciones bivariadas:

- Distribución exponencial bivariada de Gumbel tipo I.

Su función de supervivencia conjunta es:

$$S(x, y) = \Pr(X > x, Y > y) = e^{-x-y-\theta xy}; \quad x, y \geq 0, \quad 0 \leq \theta \leq 1$$

- Distribución Farlie-Gumbel-Morgenstern (F-G-M) con marginales exponenciales.

Su densidad conjunta es:

$$f(x, y) = \lambda^2 e^{-\lambda(x+y)} [1 + \alpha(2e^{-\lambda x} - 1)(2e^{-\lambda y} - 1)]; \quad x, y \geq 0$$

- Distribución exponencial bivariada de Marshall y Olkin.

Su función de supervivencia es:

$$S(x, y) = \exp[-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x, y)] \quad \text{con } x, y, \lambda_1, \lambda_2 \text{ y } \lambda_{12} \geq 0.$$

Puede notarse que X y Y son independientes si y sólo si $\lambda_{12} = 0$.

También son de interés algunos modelos desarrollados para describir el comportamiento de sistemas donde existe la posibilidad de restaurar los artículos que fallan aunque la reparación sea imperfecta. Dos de estos modelos posibles son los siguientes:

- Dos artículos, con tiempos de falla X y Y son puestos en operación y cuando uno de ellos presenta una avería es sometido a reparación. Con probabilidad p la compostura no sirve y el artículo es desechado, con probabilidad $1 - p$ es exitosa y el artículo sigue funcionando. Si ambos artículos fallan al mismo tiempo entonces son arreglados de manera independiente y cada uno tiene una probabilidad p de ser desechado y probabilidad $1 - p$ de ser reparado exitosamente.
- En un segundo modelo el planteamiento es el mismo, sólo que si ambos artículos fallan al mismo tiempo, con probabilidad $1 - p$ ambos son restaurados de manera exitosa y con probabilidad p ambos son desechados.

Finalmente, cabe comentar que un enfoque bayesiano puede ser de mucha utilidad cuando existe alguna dependencia del comportamiento del sistema o artículo de interés hacia una cantidad no observable, en caso de no presentarse esta situación, los conceptos definidos en este capítulo bastan para un buen análisis del envejecimiento observado.

2. Aproximación de las funciones de supervivencia y ajuste de algunas distribuciones a datos observados

Introducción

Para aproximar los valores de las funciones de supervivencia existen distintos métodos, los cuales se clasifican en paramétricos y no paramétricos, estos serán revisados en el presente capítulo.

Un concepto importante que será considerado al plantear los procedimientos, es el de observación censurada, lo cual se presenta cuando el fenómeno de interés sucede fuera del periodo de estudio establecido o por algún motivo se desconoce el momento preciso de su ocurrencia.

Finalmente, se exponen diversas distribuciones que han sido utilizadas para ajustarse a datos de fiabilidad y se presenta un método para calcular los parámetros involucrados en las mismas, lo que permitirá en conjunto obtener valores teóricos para las funciones de supervivencia y otras relacionadas.

2.1 Censura de datos

Se dice que una observación está censurada cuando el sujeto o artículo en estudio no es visto hasta la presentación del evento de interés o el seguimiento de su comportamiento comienza después de la presentación de tal evento. Por ejemplo, un grupo de pacientes es sometido a un tratamiento en una clínica y se espera observar cierto resultado a partir de la novena semana, así que hasta entonces inicia el seguimiento por parte de los médicos, sin embargo, algunos pacientes presentan la reacción esperada poco tiempo después de iniciado el tratamiento y en la novena semana no es claro en qué momento debería registrarse la observación, en ese caso se tiene una censura al inicio; por otro lado, se puede establecer por cuestiones administrativas que el periodo de estudio máximo sea de un año, ya que en ese momento todos deberían haber mostrado su reacción al tratamiento, pero cumplido ese plazo

algunos pacientes siguen sin presentar respuesta, así que se les asigna como registro de la observación el máximo, es decir, un año, en ese caso se tiene una censura por la derecha.

También hay censuras por intervalos contenidos en el periodo de observación y, en general, se pueden presentar particularidades en cualquier fenómeno que es objeto de estudio respecto a las observaciones censuradas.

A los datos que se presentan en condiciones donde todos los individuos o dispositivos mueren o fallan dentro del periodo de observación, se les llama datos de mortalidad completa, en el caso contrario se les llama datos de mortalidad incompleta.

Al elegir el tiempo que se utilizará para realizar un estudio y registrar datos se puede truncar o censurar la información, dos de los esquemas más comunes son:

1. Terminar el estudio en un punto predeterminado de tiempo, los datos resultantes se denominan truncados.
2. Terminar el estudio cuando un número predeterminado de muertes (o fallas) ha ocurrido, los datos resultantes se denominan censurados.

Algunos autores llaman censura tipo 1 al primer esquema y censura tipo 2 al segundo. Puede notarse que en el esquema 1, el número de muertes es una variable aleatoria, mientras que en el esquema 2, el tiempo de terminación del estudio es una variable aleatoria; los tiempos individuales de muerte (o falla) son variables aleatorias en ambos casos.

En algunos casos, como en estudios clínicos, los individuos entran en observación en diferentes momentos, cuando esto se combina con la censura tipo 1 mencionada previamente, se habla de censura de tiempo progresiva; si el ingreso en observación en diferentes momentos se combina con la censura tipo 2 se habla de censura de falla progresiva.

Cuando los tiempos de entrada a un estudio son variables puede distinguirse entre el tiempo cronológico y tiempo a partir de que un individuo entró en observación, también llamado tiempo de seguimiento.

El tiempo cronológico de estudio es usualmente el tiempo calendario, en esta escala cada individuo tiene su propia fecha de inicio; los individuos presentes al inicio del estudio se llaman principiantes, y aquellos que entran en observación en un momento posterior se llaman nuevos participantes; durante el estudio las observaciones de algunos individuos pueden ser discontinuas, estas observaciones son llamadas retiros; los individuos presentes al final del estudio pueden ser llamados finalistas.

En el tiempo de seguimiento se asigna el valor $t = 0$ a cada individuo al momento de su entrada al estudio y se mantiene en observación hasta que alguna de las siguientes situaciones se presente: muerte o falla, abandono del estudio por razones desconocidas al momento de entrada al estudio o, fin del estudio. En este esquema no se distingue entre principiantes y nuevos participantes, ya que todos los participantes son principiantes.

Existen además algunas variantes al esquema de tiempo de estudio predeterminado, dos de las más comunes son:

1. Cada individuo se mantiene en observación por el mismo periodo de tiempo a partir de su ingreso al estudio, este método de reunir datos es llamado método de aniversario y es muy utilizado en estudios a largo plazo; el momento para finalizar el estudio es determinado por la última entrada.
2. Combinar un tiempo de término fijo para el estudio con el método de aniversario, el que suceda después, este método es llamado método del último registro y también es bastante utilizado.

2.2 Aproximación de las funciones de supervivencia

Una vez que se cuenta con observaciones de un fenómeno donde se desea analizar la fiabilidad de dispositivos u organismos será útil plantear métodos

que aproximen los valores que toman las funciones relacionadas para contar con una descripción gráfica y cuantitativa del comportamiento de la supervivencia o mortalidad de interés.

Los métodos que brindan esta descripción se pueden clasificar en paramétricos y no paramétricos. Un método es no paramétrico (o de distribución libre) si el mismo no requiere de un supuesto de distribución, en caso contrario se le llama paramétrico.

La preferencia por algún tipo de método dependerá de la finalidad del estudio, de la cantidad de datos con que se cuenta y las características de los mismos, de las herramientas tecnológicas y de la experiencia de quien lleva a cabo la tarea, así como las condiciones particulares de trabajo.

2.2.1 Métodos no paramétricos

Se revisará el producto límite de Kaplan Meier y el actuarial; se trata de procedimientos sencillos y similares aunque existe una diferencia radical en cuanto al establecimiento de intervalos de tiempo para la descripción del fenómeno.

2.2.1.1 Método actuarial

Es adecuado cuando el número de observaciones es grande. Se considera que el tiempo de estudio va de $t = 0$ a $t = A$, el cual se divide en r segmentos previamente definidos, si bien el método no especifica alguna longitud para los r intervalos, la precisión de la aproximación dependerá de la cantidad y extensión de los mismos. De este modo puede considerarse el periodo de observación como la siguiente sucesión de intervalos:

$$[0, a_1), [a_1, a_2), \dots, [a_{j-1}, a_j), \dots, [a_{r-1}, A)$$

Se inicia aproximando la probabilidad de sobrevivir al intervalo $[a_{j-1}, a_j)$ suponiendo que se estaba vivo al comienzo del mismo, la cual se denota por p_j :

$$p_j = \Pr (T \geq a_j | T \geq a_{j-1})$$

donde T es la variable aleatoria que indica el tiempo de vida del organismo o artículo en observación.

La aproximación de p_j dependerá del número de unidades expuestas al riesgo de fallecer o fallar durante $[a_{j-1}, a_j)$, denotado por n_j , y de la cantidad de aquellas que presentan el suceso de interés en este mismo intervalo, llamada d_j .

$$p_j = \frac{n_j - d_j}{n_j} = 1 - \frac{d_j}{n_j}$$

Para obtener el número de individuos expuestos al riesgo durante $[a_{j-1}, a_j)$, es decir, n_j , es importante conocer la cantidad de vivos al comienzo del intervalo n'_j , de fallecidos d_j , de aquellos cuya observación es censurada c_j y de individuos vivos al comienzo del siguiente intervalo $n'_{j+1} = n'_j - d_j - c_j$.

Si en $[a_{j-1}, a_j)$ no hay censuras se tiene que $n_j = n'_j$, en caso de haberlas existirán individuos que no están expuestos al riesgo durante todo el intervalo, sin embargo, pueden aportar información para la aproximación de la función de supervivencia, por ejemplo, si una observación es censurada en el cuarto día de diez, contribuye con $\frac{4}{10} = 0.4$ al número de individuos expuestos al riesgo. Este método asume que si c_j es el número de censuras en $[a_{j-1}, a_j)$, la contribución de las c_j al número n_j será $\frac{c_j}{2}$, esto supone que el momento en

que se presenta la censura de una observación se distribuye de manera uniforme a lo largo del intervalo, de este modo

$$n_j = n'_j - \frac{c_j}{2}$$

En consecuencia

$$\hat{p}_j = 1 - \frac{d_j}{n'_j - \frac{c_j}{2}}$$

Se denota por $\hat{S}(a_i)$ la estimación de la función de supervivencia en el instante a_i y se calcula como el producto de las probabilidades de sobrevivir o no fallar en el intervalo $[a_{i-1}, a_i)$ y todos los intervalos anteriores, siendo a_i el momento donde comienza el intervalo $[a_i, a_{i+1})$, es decir,

$$\hat{S}(a_i) = \hat{p}_1 \hat{p}_2 \hat{p}_3 \dots \hat{p}_i = \prod_{j|a_j \leq a_i} \left[1 - \frac{d_j}{n'_j - \frac{c_j}{2}} \right] = \prod_{j=1}^i \left[1 - \frac{d_j}{n'_j - \frac{c_j}{2}} \right]$$

Para todos los valores t comprendidos en $[a_{j-1}, a_j)$ se puede obtener $\hat{S}(t)$ por interpolación lineal:

$$\hat{S}(t) = \hat{S}(a_{j-1}) + \frac{\hat{S}(a_j) - \hat{S}(a_{j-1})}{a_j - a_{j-1}} (t - a_{j-1})$$

2.2.1.2 Producto límite de Kaplan Meier

Este procedimiento brinda una aproximación de las funciones de supervivencia y otras relacionadas. La principal diferencia con el método actuarial es que en este planteamiento no se considera el previo establecimiento de intervalos de tiempo, sino que los periodos para el cálculo de las probabilidades de supervivencia quedarán determinados por el momento en que muere cada uno

de los organismos o fallan los dispositivos en observación, siendo el primer intervalo el que va del instante cero hasta la primera muerte o falla, el segundo se considera a partir del final del primer periodo y hasta la siguiente falla o muerte y así sucesivamente.

Se establece la siguiente notación:

P_1 la probabilidad de sobrevivir el primer periodo

P_2 la probabilidad de sobrevivir el segundo periodo dado que se sobrevivió al primero

P_t la probabilidad de sobrevivir al periodo t dado que se sobrevivió al periodo $t - 1$.

Por este método, la probabilidad de sobrevivir hasta el periodo t , denotada por $S(t)$, se obtiene al multiplicar las probabilidades P_1, P_2, \dots, P_t , así que

$$S(t) = P_1, P_2, \dots, P_t = \prod_{i=1}^t P_i$$

Esto supone independencia entre las probabilidades de supervivencia a los distintos periodos, ya que la probabilidad de sobrevivir a todos los periodos (intersección) es el producto de todas las probabilidades.

Para estimar las probabilidades P_1, P_2, \dots, P_t se considera la proporción del número de sujetos (o artículos) que sobreviven al final del periodo entre aquellos con los que se contaba al inicio del mismo.

Sea n_t la cantidad de vivos al inicio del periodo t y d_t los sujetos que mueren (o artículos que fallan) durante dicho intervalo, de este modo, $n_t - d_t$ es el número de sobrevivientes al final de t .

Entonces, la estimación de la probabilidad de sobrevivir al periodo t es

$$\hat{p}_t = \frac{n_t - d_t}{n_t} = 1 - \frac{d_t}{n_t}$$

En caso de presentarse observaciones censuradas, n_t adopta la forma $n_t = n_{t-1} - d_{t-1} - c_{t-1}$, donde c_{t-1} es el número de observaciones censuradas durante el periodo $t - 1$.

Retomando la estimación de la función de supervivencia establecida previamente se obtiene el estimador conocido como producto límite de Kaplan Meier:

$$\hat{S}(t) = \prod_{i=1}^t \left(1 - \frac{d_i}{n_i}\right)$$

En caso de no existir observaciones censuradas se obtiene la siguiente igualdad a causa de la cancelación de algunos términos en la serie de productos

$$\begin{aligned} \hat{S}(t) &= \prod_{i=1}^t \left(1 - \frac{d_i}{n_i}\right) = \left(1 - \frac{d_1}{n_1}\right) \left(1 - \frac{d_2}{n_2}\right) \dots \left(1 - \frac{d_{t-1}}{n_{t-1}}\right) \left(1 - \frac{d_t}{n_t}\right) \\ &= \left(\frac{n_1 - d_1}{n_1}\right) \left(\frac{n_2 - d_2}{n_2}\right) \dots \left(\frac{n_{t-1} - d_{t-1}}{n_{t-1}}\right) \left(\frac{n_t - d_t}{n_t}\right) \\ &= \left(\frac{n_2}{n_1}\right) \left(\frac{n_3}{n_2}\right) \dots \left(\frac{n_t}{n_{t-1}}\right) \left(\frac{n_t - d_t}{n_t}\right) \dots \text{porque } n_k = n_{k-1} - d_{k-1} \\ &= \frac{n_t - d_t}{n_1} \end{aligned}$$

Continuando con este enfoque y recordando el significado de la función de densidad como la probabilidad de muerte o falla para un tiempo fijo, puede aproximarse $f(t)$ mediante la proporción de muertes o fallas en un periodo o instante entre el total de organismos o dispositivos en el estudio.

$$\hat{f}(t) = \frac{d_t}{n}$$

Otra función de interés, particularmente en el campo de las matemáticas actuariales del seguro de personas es la fuerza de mortalidad, la cual

representa el riesgo de muerte por unidad de tiempo y se aproxima por la proporción de pacientes que mueren en un intervalo dado que han sobrevivido hasta el inicio del mismo, es decir,

$$\hat{r}(t) = \frac{d_t}{n_t}$$

2.2.2 Métodos paramétricos

2.2.2.1 Algunas distribuciones relacionadas con el análisis de supervivencia

Se han desarrollado a lo largo de los años distintos planteamientos que describen la mortalidad de organismos o la falla de dispositivos; estos modelos pueden ser caracterizados por la función de densidad o de distribución que proponen, ya que a partir de ellas se pueden obtener los valores de la probabilidad de supervivencia que asigna dicho modelo a cada instante. Se presentan a continuación algunas de las distribuciones más utilizadas en análisis de fiabilidad junto con las propiedades de las mismas⁷.

Distribución exponencial

Es también llamada exponencial negativa, tiene una función de densidad dada por

$$f(t) = \lambda e^{-\lambda t}, \lambda > 0, t \geq 0$$

y cuenta con una propiedad llamada en algunos textos “pérdida de memoria”, la cual se puede describir por medio de la siguiente igualdad

$$S(x|t) = S(x) \text{ para toda } x, t \geq 0,$$

⁷ En caso de consultar más datos sobre estas distribuciones debe ponerse atención en la parametrización que utiliza cada autor.

la cual significa que la probabilidad de supervivencia para un periodo adicional de duración x es el mismo sin importar la edad actual, en otras palabras, la fiabilidad condicional es igual a la no condicional. Esta propiedad hace que la distribución exponencial sea útil para describir el comportamiento de la mortalidad de organismos que no envejecen, esto sucede cuando la edad de un componente no tiene efecto alguno en su vida residual.

Otra característica de esta distribución es que tiene tasa de falla constante, $r(t) = \lambda$ para toda $t \geq 0$.

Su esperanza es $E(X) = \frac{1}{\lambda}$ y su varianza $var(X) = \frac{1}{\lambda^2}$.

La distribución exponencial es un caso particular de las distribuciones gamma, Weibull y Gompertz, entre otras.

La distribución exponencial es usada ampliamente en el campo de fiabilidad y una de las razones es que suele ser un miembro común de casi todas las clases de supervivencia presentadas en el capítulo 1 (IFR, DFR, etc.).

Distribución gamma

Su función de densidad asociada es

$$f(t) = \frac{\lambda^\alpha t^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda t}; \quad a, \lambda > 0$$

como se señaló anteriormente, la distribución exponencial es un caso especial de la distribución gamma cuando $\alpha = 1$.

Cuando α es un entero positivo, $F(t)$ puede ser calculada por medio de la fórmula

$$F(t) = 1 - \sum_{i=0}^{\alpha-1} \frac{(\lambda t)^i}{i!} e^{-\lambda t}; \quad t \geq 0$$

La esperanza está dada por $E(X) = \frac{\alpha}{\lambda}$ y la varianza por $var(X) = \frac{\alpha}{\lambda^2}$.

Esta distribución ha sido utilizada por Brown y Flood (1947) para describir el tiempo de vida de los vasos de cristal circulando en una cafetería y por Birnbaum y Saunders (1958) para modelar el tiempo de vida de diversos materiales.

Distribución Weibull

El físico sueco Waloddi Weibull utilizó en 1939 esta distribución para modelar la resistencia a la rotura de materiales y en 1951 para otras aplicaciones, por esta razón lleva su nombre.

La función de densidad está dada por

$$f(x) = \frac{\alpha}{\beta} x^{\alpha-1} e^{-\frac{x^\alpha}{\beta}}; \quad 0 < x < \infty, \alpha, \beta > 0$$

La función de distribución

$$F(t) = \int_0^t f(x) dx = \int_0^t \frac{\alpha}{\beta} x^{\alpha-1} e^{-\frac{x^\alpha}{\beta}} dx = 1 - e^{-\frac{t^\alpha}{\beta}}$$

La función de fiabilidad

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

La función de riesgo

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

Además, su esperanza está dada por

$$\mu = \frac{\beta^{\frac{1}{\alpha}}}{\alpha} \Gamma\left(\frac{1}{\alpha}\right)$$

La distribución exponencial puede ser obtenida al tomar $\alpha = 1$ en la distribución Weibull.

Distribución Gompertz

Gompertz desarrolló en 1825 un modelo de probabilidad para mortalidad humana. Comenzó estableciendo la función de riesgo como

$$r(t) = Bc^t; \quad t \geq 0, B > 0, c \geq 0$$

Es útil notar que $r(t)$ es creciente para $c > 1$ y decreciente para $c < 1$. Además, $r(t) = B$ cuando $c = 1$, de este modo la distribución exponencial es un caso especial de la Gompertz, esto puede afirmarse gracias a un resultado que señala que una tasa de falla constante es condición necesaria y suficiente para que la distribución sea exponencial.

Al aplicar algunos procedimientos de ecuaciones diferenciales resulta la siguiente función de supervivencia

$$S(t) = e^{-B(c^t-1)/\log c}; \quad t \geq 0$$

finalmente, la función de densidad tiene la siguiente expresión

$$f(t) = Bc^t e^{-B(c^t-1)/\log c}; \quad t \geq 0, B > 0, c \geq 0$$

Al discutir sobre teoría de fiabilidad para envejecimiento y longevidad, Gavrilov y Gavrilova (2001) comentaron que mientras que la distribución Weibull es más

comúnmente aplicable para tiempos de falla de dispositivos técnicos, la distribución Gompertz es más común para sistemas biológicos.⁸

Distribución Makeham

A partir de una generalización de la distribución Gompertz resulta la distribución Makeham, también conocida como Gompertz-Makeham. Su función de supervivencia está dada por

$$S(t) = \exp\left\{-\alpha t + \left(\frac{\beta}{\lambda}\right)(e^{\lambda t} - 1)\right\}; \quad t \geq 0, \alpha, \beta, \lambda > 0$$

su función de riesgo es

$$r(t) = \alpha + \beta e^{\lambda t}$$

y la función de densidad está dada por

$$f(t) = \exp\left[-\alpha t + \left(\frac{\beta}{\lambda}\right)(e^{\lambda t} - 1)\right] [\alpha - \beta(e^{\lambda t})]$$

Existen otras distribuciones que pueden utilizarse en análisis de fiabilidad, sin embargo, con las anteriores bastará para describir muchos fenómenos debido a su versatilidad y demostrada eficacia.

2.2.2.2 Cálculo de parámetros

Algunos criterios que se pueden tomar en cuenta para elegir una distribución que pueda aproximar un conjunto de datos de fallas o muertes son los siguientes:

⁸ Lai Chin-Diew y Xie Min. Stochastic Ageing and Dependence for Reliability. Ed. Springer. Estados Unidos, 2006.

- 1) Simplicidad de la expresión de las funciones de fiabilidad.
- 2) Flexibilidad o capacidad para describir varias situaciones.
- 3) Bases físicas de la distribución e interpretación de los parámetros.
- 4) Calidad de las estimaciones de los parámetros del modelo.

Una vez elegida la distribución, se requiere aproximar el valor de los parámetros, para lo cual existen distintos métodos y uno de los más utilizados es el de máxima verosimilitud.

Máxima verosimilitud

Supóngase que los datos x son los valores observados de una variable aleatoria X de una familia paramétrica de funciones de densidad, $X \sim f(x; \theta)$, donde, en general, θ es multidimensional, $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$ para alguna $d \geq 1$. Puede considerarse en particular el caso cuando $X = (X_1, \dots, X_n)$ es una muestra de elementos independientes e idénticamente distribuidos a partir de algún modelo paramétrico. Después de observar x , la función de verosimilitud queda definida por

$$L(\theta) \equiv L(\theta; x) = f(x; \theta)$$

vista como una función de θ para una x fija. La estimación máximo verosímil $\hat{\theta}(x)$ queda definida por el valor de θ que maximiza $L(\theta)$, siendo $L(\theta)$ la distribución condicional de θ dados los valores de x , con una muestra x fija; cuando se considera a $\hat{\theta}(X)$, una variable aleatoria, se habla del estimador máximo verosímil.

$\hat{\theta}(X)$ cuenta con características importantes respecto a los conceptos de consistencia, sesgo, eficiencia y suficiencia, las cuales lo hacen un candidato importante para aproximar el valor de un parámetro.

A manera de ejemplo se presenta a continuación el estimador máximo verosímil del parámetro utilizado en una distribución exponencial, se utilizará una parametrización distinta a la presentada previamente aunque muy utilizada en la literatura.

La función de densidad de la distribución exponencial puede ser escrita como:

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$$

Con esta función de densidad, la función de verosimilitud para una muestra de elementos independientes e idénticamente distribuidos es:

$$L(x_1, x_2, \dots, x_n; \theta) = \frac{1}{\theta^n} e^{-\frac{\sum_{i=1}^n x_i}{\theta}}$$

Suele calcularse el logaritmo de la función de verosimilitud antes de continuar con el cálculo del estimador para simplificar el procedimiento.

$$\ln (L(x_1, x_2, \dots, x_n; \theta)) = \ln \left(\frac{1}{\theta^n} e^{-\frac{\sum_{i=1}^n x_i}{\theta}} \right) = \frac{-\sum_{i=1}^n x_i}{\theta} - n \ln \theta$$

Para encontrar los puntos críticos se deriva respecto a θ y se iguala a cero:

$$\frac{d \ln (L(x_1, x_2, \dots, x_n; \theta))}{d\theta} = \frac{\sum_{i=1}^n x_i}{\theta^2} - \frac{n}{\theta} = 0$$

Al despejar θ en la segunda parte de la ecuación se obtiene:

$$\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}$$

Aquí ya se denota el estimador por $\hat{\theta}$.

Para comprobar que el estimador se trata de un máximo se calcula la segunda derivada del logaritmo de la función de verosimilitud y se evalúa en el estimador:

$$\frac{d^2 \text{Ln} (L(x_1, x_2, \dots, x_n; \theta))}{d^2 \theta} = \frac{d \left(\frac{\sum_{i=1}^n x_i}{\theta^2} - \frac{n}{\theta} \right)}{d\theta} = \frac{-2 \sum_{i=1}^n x_i}{\theta^3} + \frac{n}{\theta^2}$$

Evaluando en $\hat{\theta}$ resulta:

$$\frac{-2 \sum_{i=1}^n x_i}{\left(\frac{\sum_{i=1}^n x_i}{n} \right)^3} + \frac{n}{\left(\frac{\sum_{i=1}^n x_i}{n} \right)^2} = \frac{-2n^3}{(\sum_{i=1}^n x_i)^2} + \frac{n^3}{(\sum_{i=1}^n x_i)^2} = \frac{-n^3}{(\sum_{i=1}^n x_i)^2}$$

Como n es un número positivo, el resultado es negativo y se puede decir que la segunda derivada del logaritmo de la función de verosimilitud es negativa al evaluarla en el estimador $\hat{\theta}$, es decir, se trata de un máximo, y así se concluye que $\hat{\theta}$ es el estimador máximo verosímil.

3. Aplicación a un caso práctico

Introducción

Dado que el deterioro a través del tiempo es un factor común en la mayoría de los dispositivos o seres vivos, la aplicación de los conceptos de fiabilidad es frecuente en muchas áreas de la industria, la salud, las ciencias sociales, etc.

En el presente capítulo se desarrolla un análisis del comportamiento del envejecimiento estocástico en un grupo de dispositivos, incluyendo la aproximación por métodos no paramétricos de las funciones relacionadas y el ajuste de una función de densidad conocida por métodos paramétricos.

3.1 Datos a utilizar

Se utilizará una tabla que contiene la información de la duración, en horas operativas, de los intervalos entre fallas sucesivas del equipo de aire acondicionado en 13 aeronaves Boeing 720. Esta tabla puede ser encontrada en el libro “The Statistical Analysis of Series of Events” de D. R. Cox y P. A. W. Lewis. La información fue inicialmente publicada en el artículo “Theoretical explanation of observed decreasing failure rate” de Proschan F., incluido en el volumen 5 de Technometrics.

La tabla es la siguiente:

Tabla 3.1. Intervalos de tiempo entre fallas sucesivas en equipos de aire acondicionado de aeronaves Boeing 720.

Aeronave												
1	2	3	4	5	6	7	8	9	10	11	12	13
Intervalos de tiempo entre fallas sucesivas												
194	413	90	74	55	23	97	50	359	50	130	487	102
15	14	10	57	320	261	51	44	9	254	493	18	209
41	58	60	48	56	87	11	102	12	5		100	14
29	37	186	29	104	7	4	72	270	283		7	57
33	100	61	502	220	120	141	22	603	35		98	54

181	65	49	12	239	14	18	39	3	12		5	32
	9	14	70	47	62	142	3	104			85	67
	169	24	21	246	47	68	15	2			91	59
	447	56	29	176	225	77	197	438			43	134
	184	20	386	182	71	80	188				230	152
	36	79	59	33	246	1	79				3	27
	201	84	27	15	21	16	88				130	14
	118	44	153	104	42	106	46					230
	34	59	26	35	20	206	5					66
	31	29	326		5	82	5					61
	18	118			12	54	36					34
	18	25			120	31	22					
	67	156			11	216	139					
	57	310			3	46	210					
	62	76			14	111	97					
	7	26			71	39	30					
	22	44			11	63	23					
	34	23			14	18	13					
		62			11	191	14					
		130			16	18						
		208			90	163						
		70			1	24						
		101			16							
		208			52							
					95							

Dado que los valores contenidos en cada columna no muestran una tendencia a aumentar o disminuir, puede suponerse que después de repararse un equipo de aire acondicionado su funcionamiento es similar al que tiene después de cualquier otra reparación o antes de la primera.

Una cualidad de este conjunto de datos es que puede tomarse la información de las primeras doce aeronaves para aproximar algunas funciones relacionadas con el envejecimiento estocástico y verificar si el modelo obtenido se ajusta de una manera adecuada a la información de la decimotercera aeronave, pudiendo esto pensarse como una nueva adquisición de la empresa que opera con los Boeing 720 y cuyo análisis y planeación se ven beneficiados con la experiencia previamente registrada.

3.2 Estudio del envejecimiento estocástico de la población por medio de un método no paramétrico

Utilizando el método del producto límite de Kaplan Meier se pueden aproximar los valores de algunas funciones que ayudan a describir el comportamiento del envejecimiento estocástico de la población.

De la tabla 3.1 se tomarán los datos de los primeros 12 sistemas de aire acondicionado, los cuales son reparados en varias ocasiones pero puede suponerse que cada tiempo de falla corresponde a un equipo distinto, teniendo un total de 197, este planteamiento hará más sencilla la explicación de las conclusiones que se obtendrán.

Las fórmulas presentadas en el capítulo 2 y que ahora servirán para realizar las aproximaciones deseadas son las siguientes:

$$\hat{p}_t = \frac{n_t - d_t}{n_t} = 1 - \frac{d_t}{n_t}$$

$$\hat{S}(t) = \prod_{i=1}^t \left(1 - \frac{d_i}{n_i}\right)$$

$$\hat{f}(t) = \frac{d_t}{n}$$

$$\hat{r}(t) = \frac{d_t}{n_t}$$

$$\hat{F}(t) = 1 - \hat{S}(t)$$

Para el valor $t = 1$ puede notarse la existencia de dos fallas en el intervalo de tiempo $(0,1]$ y el número de equipos útiles al inicio de dicho periodo es de 197, con esta información pueden realizarse las aproximaciones de $P_1, S(1), f(1), r(1)$ y $F(1)$:

$$\hat{P}_1 = 1 - \frac{d_1}{n_1} = 1 - \frac{2}{197} = 0.98984772$$

$$\hat{S}(1) = \prod_{i=1}^1 \left(1 - \frac{d_i}{n_i}\right) = \left(1 - \frac{d_1}{n_1}\right) = \left(1 - \frac{2}{197}\right) = 0.98984772$$

$$\hat{f}(1) = \frac{d_1}{n} = \frac{2}{197} = 0.01015228$$

$$\hat{r}(1) = \frac{d_1}{n_1} = \frac{2}{197} = 0.01015228$$

$$\hat{F}(1) = 1 - \hat{S}(1) = 1 - 0.98984772$$

El procedimiento para el tiempo $t = 2$ es similar, notando que el número de equipos útiles al inicio del intervalo es $n_2 = n_1 - d_1$ y en la tabla 3.1 puede observarse que sólo se presenta una falla para $t = 2$. La aproximación de las funciones resulta:

$$\hat{P}_2 = 1 - \frac{d_2}{n_2} = 1 - \frac{1}{195} = 0.99487179$$

$$\begin{aligned} \hat{S}(2) &= \prod_{i=1}^2 \left(1 - \frac{d_i}{n_i}\right) = \left(1 - \frac{d_1}{n_1}\right) \left(1 - \frac{d_2}{n_2}\right) = \left(1 - \frac{2}{197}\right) \left(1 - \frac{1}{195}\right) \\ &= 0.98477157 \end{aligned}$$

$$\hat{f}(2) = \frac{d_2}{n} = \frac{1}{197} = 0.00507614$$

$$\hat{r}(2) = \frac{d_2}{n_2} = \frac{1}{195} = 0.00512821$$

$$\hat{F}(2) = 1 - \hat{S}(2) = 1 - 0.98477157 = 0.01522843$$

Se concentran a continuación las aproximaciones de las funciones relacionadas con fiabilidad para los valores de t contemplados en la tabla 3.1.

Tabla 3.2. Aproximaciones de las funciones relacionadas con fiabilidad a partir de los datos de la tabla 3.1.

Tiempo	Equipos útiles al inicio del periodo	Equipos que fallan durante el periodo	\hat{P}_t	$\hat{f}(t)$	$\hat{S}(t)$	$\hat{F}(t)$	$\hat{r}(t)$
1	197	2	0.98984772	0.01015228	0.98984772	0.01015228	0.01015228
2	195	1	0.99487179	0.00507614	0.98477157	0.01522843	0.00512821
3	194	4	0.97938144	0.02030457	0.96446701	0.03553299	0.02061856
4	190	1	0.99473684	0.00507614	0.95939086	0.04060914	0.00526316
5	189	5	0.97354497	0.02538071	0.93401015	0.06598985	0.02645503
7	184	3	0.98369565	0.01522843	0.91878173	0.08121827	0.01630435
9	181	2	0.98895028	0.01015228	0.90862944	0.09137056	0.01104972
10	179	1	0.99441341	0.00507614	0.9035533	0.0964467	0.00558659
11	178	4	0.97752809	0.02030457	0.88324873	0.11675127	0.02247191
12	174	4	0.97701149	0.02030457	0.86294416	0.13705584	0.02298851
13	170	1	0.99411765	0.00507614	0.85786802	0.14213198	0.00588235
14	169	6	0.96449704	0.03045685	0.82741117	0.17258883	0.03550296
15	163	3	0.98159509	0.01522843	0.81218274	0.18781726	0.01840491
16	160	3	0.98125	0.01522843	0.79695431	0.20304569	0.01875
18	157	6	0.96178344	0.03045685	0.76649746	0.23350254	0.03821656
20	151	2	0.98675497	0.01015228	0.75634518	0.24365482	0.01324503
21	149	2	0.98657718	0.01015228	0.74619289	0.25380711	0.01342282
22	147	3	0.97959184	0.01522843	0.73096447	0.26903553	0.02040816
23	144	3	0.97916667	0.01522843	0.71573604	0.28426396	0.02083333
24	141	2	0.9858156	0.01015228	0.70558376	0.29441624	0.0141844
25	139	1	0.99280576	0.00507614	0.70050761	0.29949239	0.00719424
26	138	2	0.98550725	0.01015228	0.69035533	0.30964467	0.01449275
27	136	1	0.99264706	0.00507614	0.68527919	0.31472081	0.00735294
29	135	4	0.97037037	0.02030457	0.66497462	0.33502538	0.02962963
30	131	1	0.99236641	0.00507614	0.65989848	0.34010152	0.00763359
31	130	2	0.98461538	0.01015228	0.64974619	0.35025381	0.01538462
33	128	2	0.984375	0.01015228	0.63959391	0.36040609	0.015625
34	126	2	0.98412698	0.01015228	0.62944162	0.37055838	0.01587302
35	124	2	0.98387097	0.01015228	0.61928934	0.38071066	0.01612903
36	122	2	0.98360656	0.01015228	0.60913706	0.39086294	0.01639344
37	120	1	0.99166667	0.00507614	0.60406091	0.39593909	0.00833333
39	119	2	0.98319328	0.01015228	0.59390863	0.40609137	0.01680672

41	117	1	0.99145299	0.00507614	0.58883249	0.41116751	0.00854701
42	116	1	0.99137931	0.00507614	0.58375635	0.41624365	0.00862069
43	115	1	0.99130435	0.00507614	0.5786802	0.4213198	0.00869565
44	114	3	0.97368421	0.01522843	0.56345178	0.43654822	0.02631579
46	111	2	0.98198198	0.01015228	0.55329949	0.44670051	0.01801802
47	109	2	0.98165138	0.01015228	0.54314721	0.45685279	0.01834862
48	107	1	0.99065421	0.00507614	0.53807107	0.46192893	0.00934579
49	106	1	0.99056604	0.00507614	0.53299492	0.46700508	0.00943396
50	105	2	0.98095238	0.01015228	0.52284264	0.47715736	0.01904762
51	103	1	0.99029126	0.00507614	0.5177665	0.4822335	0.00970874
52	102	1	0.99019608	0.00507614	0.51269036	0.48730964	0.00980392
54	101	1	0.99009901	0.00507614	0.50761421	0.49238579	0.00990099
55	100	1	0.99	0.00507614	0.50253807	0.49746193	0.01
56	99	2	0.97979798	0.01015228	0.49238579	0.50761421	0.02020202
57	97	2	0.97938144	0.01015228	0.4822335	0.5177665	0.02061856
58	95	1	0.98947368	0.00507614	0.47715736	0.52284264	0.01052632
59	94	2	0.9787234	0.01015228	0.46700508	0.53299492	0.0212766
60	92	1	0.98913043	0.00507614	0.46192893	0.53807107	0.01086957
61	91	1	0.98901099	0.00507614	0.45685279	0.54314721	0.01098901
62	90	3	0.96666667	0.01522843	0.44162437	0.55837563	0.03333333
63	87	1	0.98850575	0.00507614	0.43654822	0.56345178	0.01149425
65	86	1	0.98837209	0.00507614	0.43147208	0.56852792	0.01162791
67	85	1	0.98823529	0.00507614	0.42639594	0.57360406	0.01176471
68	84	1	0.98809524	0.00507614	0.4213198	0.5786802	0.01190476
70	83	2	0.97590361	0.01015228	0.41116751	0.58883249	0.02409639
71	81	2	0.97530864	0.01015228	0.40101523	0.59898477	0.02469136
72	79	1	0.98734177	0.00507614	0.39593909	0.60406091	0.01265823
74	78	1	0.98717949	0.00507614	0.39086294	0.60913706	0.01282051
76	77	1	0.98701299	0.00507614	0.3857868	0.6142132	0.01298701
77	76	1	0.98684211	0.00507614	0.38071066	0.61928934	0.01315789
79	75	2	0.97333333	0.01015228	0.37055838	0.62944162	0.02666667
80	73	1	0.98630137	0.00507614	0.36548223	0.63451777	0.01369863
82	72	1	0.98611111	0.00507614	0.36040609	0.63959391	0.01388889
84	71	1	0.98591549	0.00507614	0.35532995	0.64467005	0.01408451
85	70	1	0.98571429	0.00507614	0.35025381	0.64974619	0.01428571
87	69	1	0.98550725	0.00507614	0.34517766	0.65482234	0.01449275
88	68	1	0.98529412	0.00507614	0.34010152	0.65989848	0.01470588
90	67	2	0.97014925	0.01015228	0.32994924	0.67005076	0.02985075
91	65	1	0.98461538	0.00507614	0.3248731	0.6751269	0.01538462
95	64	1	0.984375	0.00507614	0.31979695	0.68020305	0.015625
97	63	2	0.96825397	0.01015228	0.30964467	0.69035533	0.03174603
98	61	1	0.98360656	0.00507614	0.30456853	0.69543147	0.01639344
100	60	2	0.96666667	0.01015228	0.29441624	0.70558376	0.03333333
101	58	1	0.98275862	0.00507614	0.2893401	0.7106599	0.01724138
102	57	1	0.98245614	0.00507614	0.28426396	0.71573604	0.01754386

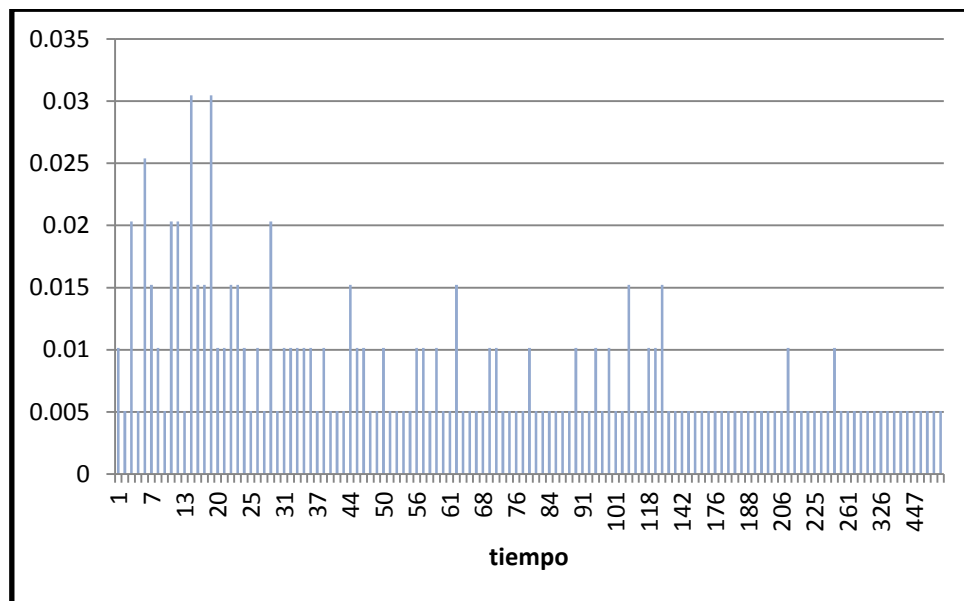
104	56	3	0.94642857	0.01522843	0.26903553	0.73096447	0.05357143
106	53	1	0.98113208	0.00507614	0.26395939	0.73604061	0.01886792
111	52	1	0.98076923	0.00507614	0.25888325	0.74111675	0.01923077
118	51	2	0.96078431	0.01015228	0.24873096	0.75126904	0.03921569
120	49	2	0.95918367	0.01015228	0.23857868	0.76142132	0.04081633
130	47	3	0.93617021	0.01522843	0.22335025	0.77664975	0.06382979
139	44	1	0.97727273	0.00507614	0.21827411	0.78172589	0.02272727
141	43	1	0.97674419	0.00507614	0.21319797	0.78680203	0.02325581
142	42	1	0.97619048	0.00507614	0.20812183	0.79187817	0.02380952
153	41	1	0.97560976	0.00507614	0.20304569	0.79695431	0.02439024
156	40	1	0.975	0.00507614	0.19796954	0.80203046	0.025
163	39	1	0.97435897	0.00507614	0.1928934	0.8071066	0.02564103
169	38	1	0.97368421	0.00507614	0.18781726	0.81218274	0.02631579
176	37	1	0.97297297	0.00507614	0.18274112	0.81725888	0.02702703
181	36	1	0.97222222	0.00507614	0.17766497	0.82233503	0.02777778
182	35	1	0.97142857	0.00507614	0.17258883	0.82741117	0.02857143
184	34	1	0.97058824	0.00507614	0.16751269	0.83248731	0.02941176
186	33	1	0.96969697	0.00507614	0.16243655	0.83756345	0.03030303
188	32	1	0.96875	0.00507614	0.15736041	0.84263959	0.03125
191	31	1	0.96774194	0.00507614	0.15228426	0.84771574	0.03225806
194	30	1	0.96666667	0.00507614	0.14720812	0.85279188	0.03333333
197	29	1	0.96551724	0.00507614	0.14213198	0.85786802	0.03448276
201	28	1	0.96428571	0.00507614	0.13705584	0.86294416	0.03571429
206	27	1	0.96296296	0.00507614	0.1319797	0.8680203	0.03703704
208	26	2	0.92307692	0.01015228	0.12182741	0.87817259	0.07692308
210	24	1	0.95833333	0.00507614	0.11675127	0.88324873	0.04166667
216	23	1	0.95652174	0.00507614	0.11167513	0.88832487	0.04347826
220	22	1	0.95454545	0.00507614	0.10659898	0.89340102	0.04545455
225	21	1	0.95238095	0.00507614	0.10152284	0.89847716	0.04761905
230	20	1	0.95	0.00507614	0.0964467	0.9035533	0.05
239	19	1	0.94736842	0.00507614	0.09137056	0.90862944	0.05263158
246	18	2	0.88888889	0.01015228	0.08121827	0.91878173	0.11111111
254	16	1	0.9375	0.00507614	0.07614213	0.92385787	0.0625
261	15	1	0.93333333	0.00507614	0.07106599	0.92893401	0.06666667
270	14	1	0.92857143	0.00507614	0.06598985	0.93401015	0.07142857
283	13	1	0.92307692	0.00507614	0.06091371	0.93908629	0.07692308
310	12	1	0.91666667	0.00507614	0.05583756	0.94416244	0.08333333
320	11	1	0.90909091	0.00507614	0.05076142	0.94923858	0.09090909
326	10	1	0.9	0.00507614	0.04568528	0.95431472	0.1
359	9	1	0.88888889	0.00507614	0.04060914	0.95939086	0.11111111
386	8	1	0.875	0.00507614	0.03553299	0.96446701	0.125
413	7	1	0.85714286	0.00507614	0.03045685	0.96954315	0.14285714
438	6	1	0.83333333	0.00507614	0.02538071	0.97461929	0.16666667
447	5	1	0.8	0.00507614	0.02030457	0.97969543	0.2
487	4	1	0.75	0.00507614	0.01522843	0.98477157	0.25

493	3	1	0.66666667	0.00507614	0.01015228	0.98984772	0.33333333
502	2	1	0.5	0.00507614	0.00507614	0.99492386	0.5
603	1	1	0	0.00507614	0	1	1

Ahora pueden graficarse los valores de las funciones aproximadas para tener una idea más clara del comportamiento del envejecimiento estocástico de la población. El eje vertical de las gráficas representa la probabilidad de supervivencia o falla de acuerdo a la definición de cada función.

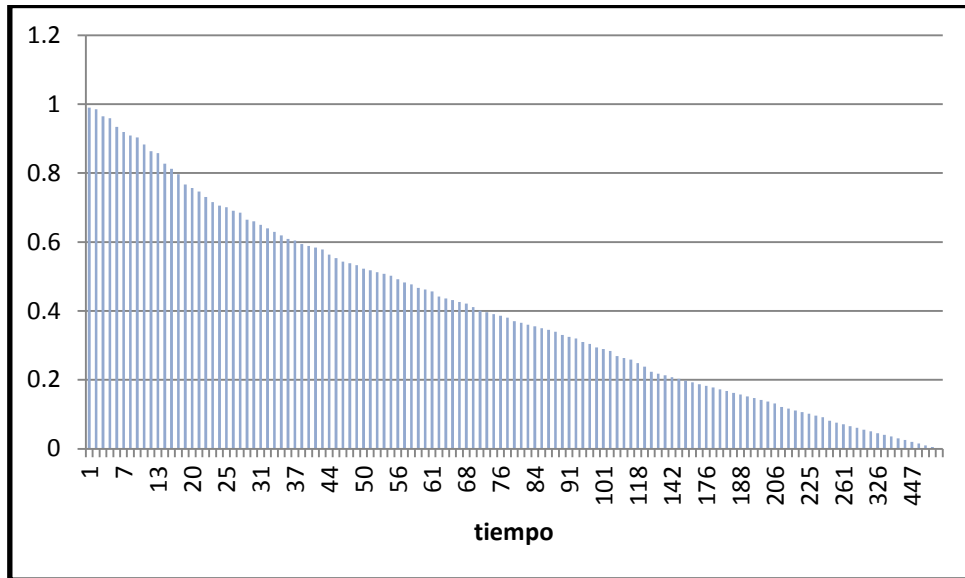
Se puede comenzar con la función de densidad:

Gráfico 3.1. $\hat{f}(t)$: función de densidad



Puede notarse que la función de densidad alcanza algunos valores relativamente altos comparados con el promedio para los tiempos de falla menores a 140, es decir, estos valores de t presentan una mayor frecuencia que el resto y es posible que la edad sea en realidad benéfica para los sistemas de aire acondicionado.

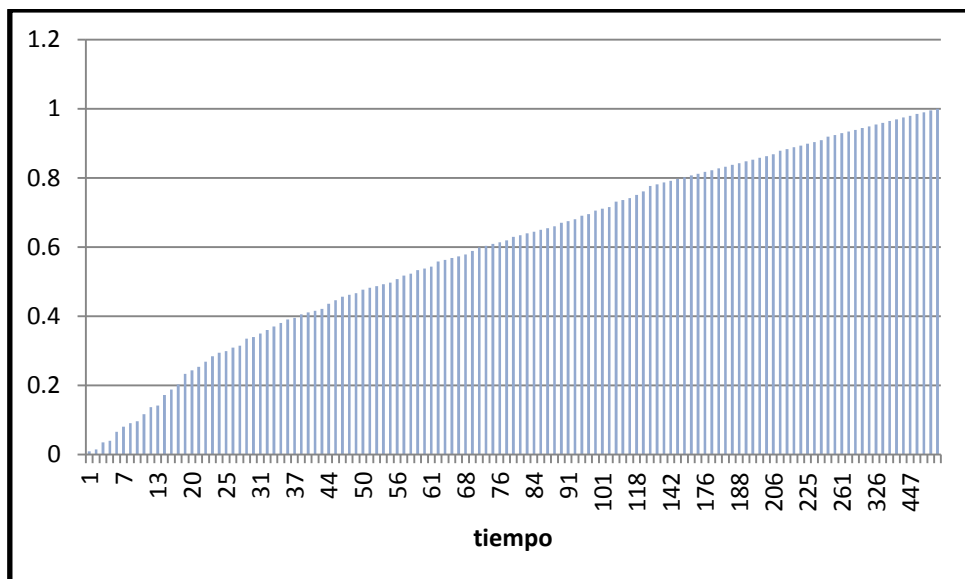
Gráfico 3.2. $\hat{S}(t)$: *función de supervivencia*



Puede decirse que el decaimiento de la función de supervivencia es relativamente uniforme, es decir, no se muestra una edad crítica para los equipos de aire acondicionado.

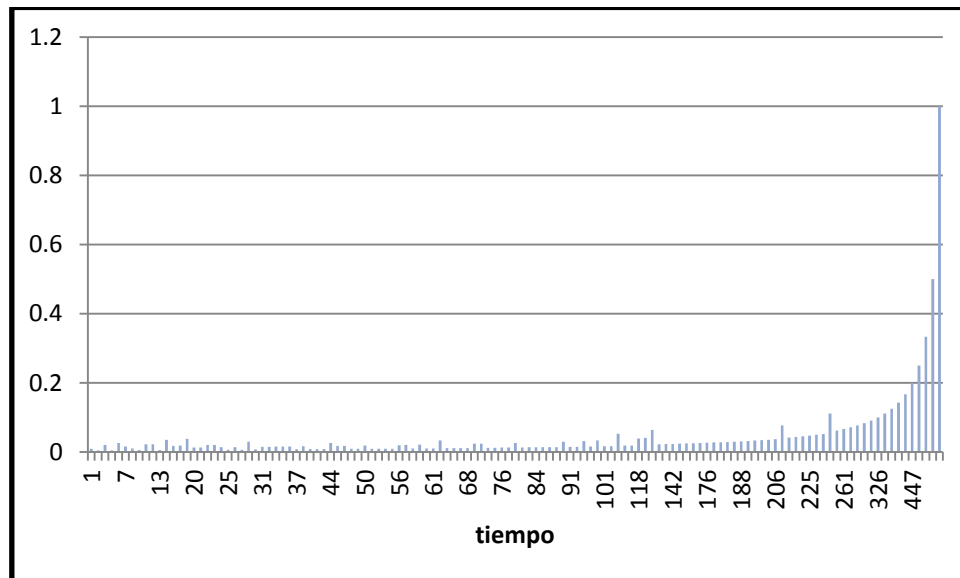
La función de distribución acumulada:

Gráfico 3.3. $\hat{F}(t)$: *función de distribución*



Por último, la función de riesgo:

Gráfico 3.4. $\hat{r}(t)$: *función de riesgo*



A partir de las 438 horas de operación se nota un incremento importante en la tasa de falla, ya que en ese momento son pocos los equipos de aire acondicionado que siguen funcionando y después de dicho tiempo de uso es muy probable que cualquiera de ellos falle.

3.3 Estudio del envejecimiento estocástico de la población por medio de un método paramétrico

En esta sección se aplicará un método paramétrico que también ayudará a describir el envejecimiento estocástico de la población observada.

Como se planteó anteriormente, un método paramétrico supone una distribución relacionada con el fenómeno en cuestión, en este caso puede tomarse en cuenta que una función utilizada en muchas ocasiones para describir el comportamiento de fallas en sistemas mecánicos o eléctricos es la distribución Weibull, la cual ha sido ampliamente estudiada a lo largo de los años y puede ser un buen candidato para modelar el envejecimiento

estocástico de la población en cuestión; sin embargo, si el objeto de estudio fuera un sistema biológico podría considerarse que la distribución Gompertz o tal vez alguna otra fuera un mejor candidato, como lo han mencionado diversos autores.

La función de densidad de la distribución Weibull está dada por:

$$f(x) = \frac{\alpha}{\beta} x^{\alpha-1} e^{-\frac{x^\alpha}{\beta}}; \quad 0 < x < \infty, \alpha, \beta > 0$$

Como fue señalado previamente, uno de los métodos más utilizados para estimar los parámetros de una distribución es el de máxima verosimilitud.

Para estimar por dicho método los parámetros de una distribución Weibull se parte de la función de verosimilitud, que en este caso está dada por:

$$L(\alpha, \beta; x_1, \dots, x_n) = \prod_{i=1}^n \left(\frac{\alpha}{\beta}\right) x_i^{\alpha-1} e^{-\frac{x_i^\alpha}{\beta}} = \left(\frac{\alpha}{\beta}\right)^n \left(\prod_{i=1}^n x_i^{\alpha-1}\right) e^{-\frac{\sum_{i=1}^n x_i^\alpha}{\beta}},$$

para encontrar los estimadores máximo verosímiles de los parámetros deben buscarse los valores de α y β que maximizan la función de verosimilitud, para esto suele ser útil calcular el logaritmo de $L(\alpha, \beta; x_1, \dots, x_n)$, ya que se simplifica la expresión y no se altera la ubicación de los valores críticos.

$$\begin{aligned} \ln[L(\alpha, \beta; x_1, \dots, x_n)] &= \ln \left[\left(\frac{\alpha}{\beta}\right)^n \left(\prod_{i=1}^n x_i^{\alpha-1}\right) e^{-\frac{\sum_{i=1}^n x_i^\alpha}{\beta}} \right] \\ &= n \ln \left(\frac{\alpha}{\beta}\right) + (\alpha - 1) \left(\sum_{i=1}^n \ln x_i\right) - \frac{1}{\beta} \sum_{i=1}^n x_i^\alpha \\ &= n \ln \alpha - n \ln \beta + (\alpha - 1) \left(\sum_{i=1}^n \ln x_i\right) - \frac{1}{\beta} \sum_{i=1}^n x_i^\alpha \end{aligned}$$

Para encontrar un valor crítico puede utilizarse el criterio de las derivadas, para esto se calculan las derivadas parciales respecto a α y β y se igualan a 0.

$$\frac{\partial \ln(L)}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \ln x_i - \frac{1}{\beta} \sum_{i=1}^n x_i^\alpha \ln x_i = 0$$

$$\frac{\partial \ln(L)}{\partial \beta} = -\frac{n}{\beta} + \frac{1}{\beta^2} \sum_{i=1}^n x_i^\alpha = 0$$

Se despeja el parámetro β de ambas ecuaciones:

$$\beta = \frac{\sum_{i=1}^n x_i^\alpha \ln x_i}{\frac{n}{\alpha} + \sum_{i=1}^n \ln x_i}$$

$$\beta = \frac{\sum_{i=1}^n x_i^\alpha}{n}$$

Puede ahora igualarse el lado derecho de las dos últimas ecuaciones para no manejar por el momento el parámetro β .

$$\frac{\sum_{i=1}^n x_i^\alpha}{n} = \frac{\sum_{i=1}^n x_i^\alpha \ln x_i}{\frac{n}{\alpha} + \sum_{i=1}^n \ln x_i}$$

Lo cual suele expresarse de la siguiente forma:

$$\frac{1}{n} \sum_{i=1}^n \ln x_i = \left[\frac{\sum_{i=1}^n x_i^\alpha \ln x_i}{\sum_{i=1}^n x_i^\alpha} - \frac{1}{\alpha} \right]$$

La ecuación anterior se puede resolver por métodos numéricos⁹ para encontrar el valor de $\hat{\alpha}$, es decir, la estimación máximo verosímil de α , la cual se sustituye en la siguiente ecuación para obtener el valor de $\hat{\beta}$:

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i^{\hat{\alpha}}}{n}$$

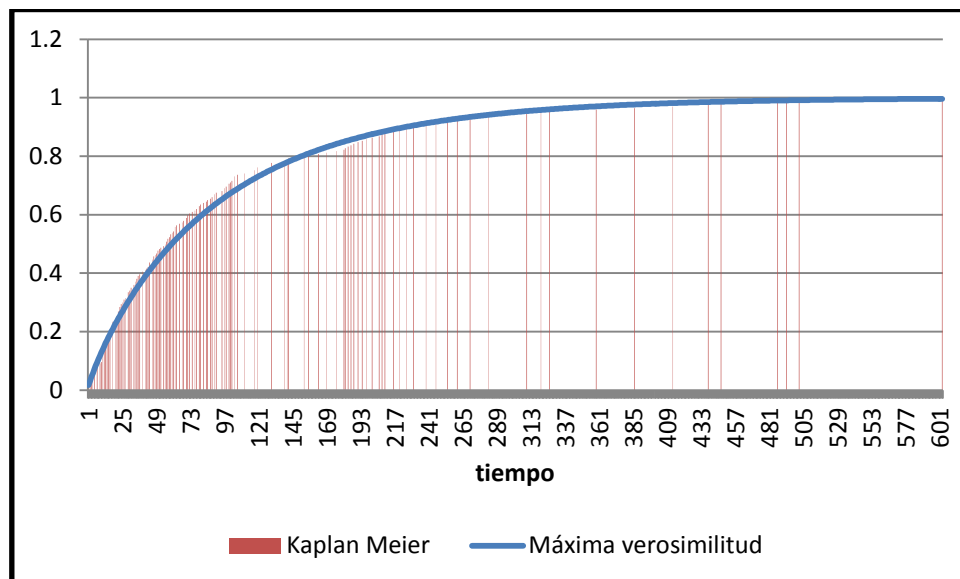
⁹ En el desarrollo del presente texto se utilizó el método de bisección, también conocido como búsqueda binaria y que está basado en el teorema del valor intermedio. Para más información al respecto puede consultarse el libro de la bibliografía referente a Análisis Numérico.

Resolviendo las ecuaciones para $\hat{\alpha}$ y $\hat{\beta}$ con ayuda de algún software¹⁰ se pueden obtener los valores $\hat{\alpha} = 0.9052$ y $\hat{\beta} = 58.3838$, los cuales se utilizarán para ajustar una distribución Weibull a la aproximación que se tiene de las funciones de supervivencia, densidad, distribución y riesgo, hechas por el método de Kaplan Meier. La función de densidad propuesta será entonces:

$$f(x) = \frac{0.9052}{58.3838} x^{0.9052-1} e^{-\frac{x^{0.9052}}{58.3838}}$$

Los siguientes gráficos muestran de manera conjunta la descripción que puede obtenerse del envejecimiento de los dispositivos en cuestión por los métodos de Kaplan Meier y máxima verosimilitud.

Gráfico 3.5. $\hat{F}(t)$: función de distribución



Se presenta en la siguiente página la función de supervivencia.

¹⁰ El software utilizado fue Matlab, sin embargo, puede utilizarse cualquier otro que permita la programación necesaria para implementar el método de bisección.

Gráfico 3.6. $\hat{S}(t)$: función de supervivencia

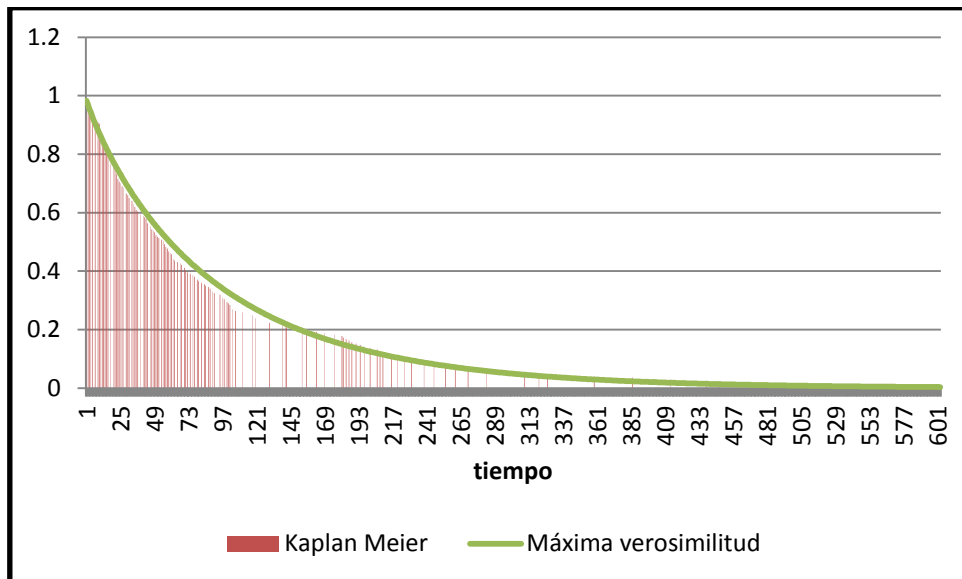


Gráfico 3.7. $\hat{f}(t)$: función de densidad

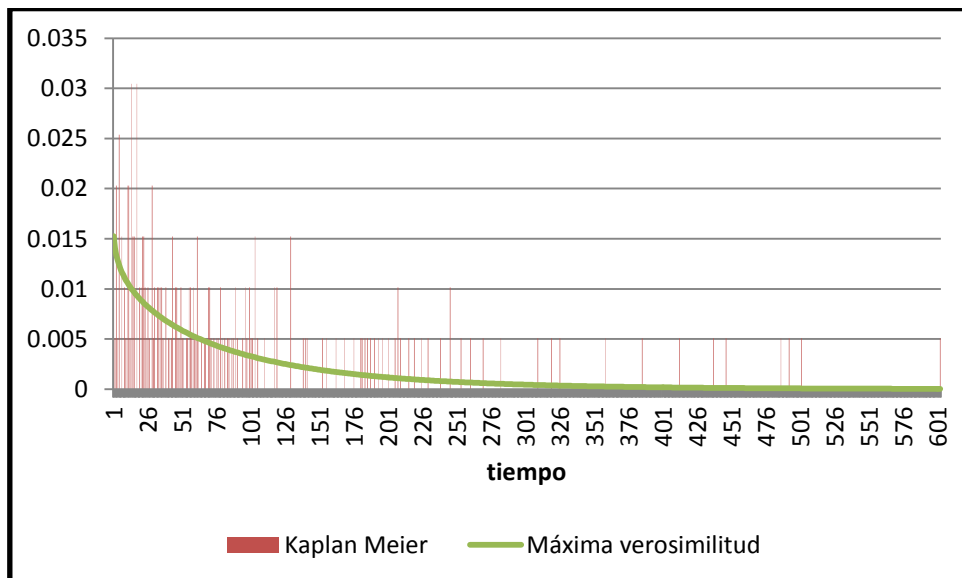
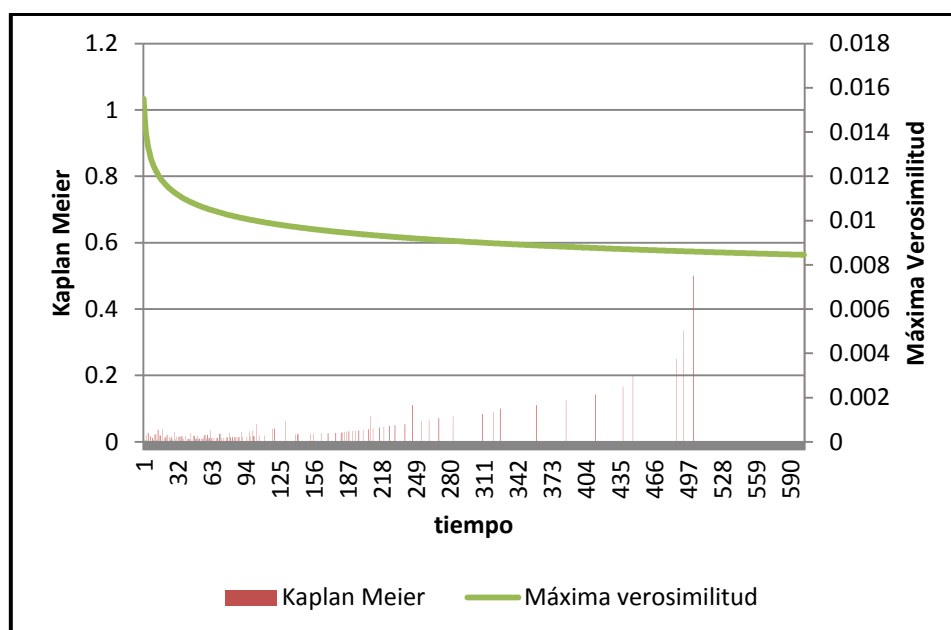


Gráfico 3.8. $\hat{r}(t)$: función de riesgo

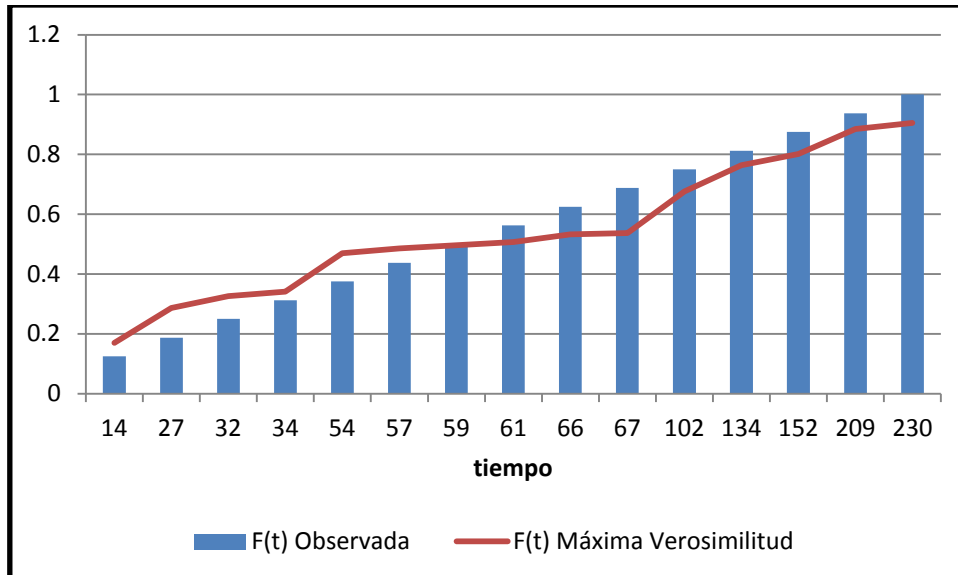


Puede notarse una gran similitud en los gráficos donde se muestran los datos relacionados con la función de distribución y la de supervivencia; en el caso de la densidad la similitud es menos evidente, sin embargo, ambos métodos presentan un comportamiento decreciente con valores parecidos al inicio y después la aproximación por máxima verosimilitud distribuye de manera suave la probabilidad de que se presente alguna falla a lo largo del tiempo; en el caso de la tasa de falla la diferencia es más notoria, esta se puede deber a que contempla para su aproximación por el método de Kaplan Meier sólo los dispositivos útiles en un intervalo de tiempo sin importar lo que haya sucedido previamente y, al paso de las horas de uso, cuando cada vez quedan menos equipos funcionando, el efecto que causa en la función de riesgo la falla de alguno de los dispositivos restantes, es mayor. Es importante recordar que incluso se han planteado definiciones alternativas para la tasa de falla debido a las diferencias entre su comportamiento para una distribución continua y el de su aproximación por medio de un enfoque discreto.

Finalmente, se muestra el ajuste de la distribución estimada por máxima verosimilitud utilizando la información de las primeras 12 aeronaves al

comportamiento de la decimotercera (descrito por el método de Kaplan Meier), notando así la capacidad de previsión que proporciona el modelo desarrollado.

Gráfico 3.9. función de distribución de la decimotercera aeronave



3.4 El estadístico Kolmogorov-Smirnov

Con base en el gráfico 3.9, que muestra la función de distribución acumulada construida con el método de máxima verosimilitud, puede señalarse que la bondad de ajuste del modelo es buena, sin embargo, es importante contar con una evidencia estadística que respalde esta afirmación.

Una de las pruebas de bondad de ajuste más populares es el uso del estadístico Kolmogorov-Smirnov, el cual toma su nombre de los dos matemáticos rusos que lo presentaron por primera vez.

De manera general, el estadístico consiste en crear una función por “escalones” inferiores y superiores que contenga a la función de distribución en cuestión con una probabilidad específica.

Se denota por x_1, x_2, \dots, x_n una muestra ordenada a la cual se le desea ajustar una distribución.

A partir de la muestra anterior se crea la siguiente función:

$$S_n(x) = \begin{cases} 0, & x < x_1 \\ \frac{k}{n}, & x_k < x < x_{k+1} \\ 1, & x \geq x_n \end{cases}$$

Y, si $F(x)$ denota la distribución propuesta, el estadístico Kolmogorov-Smirnov se define como:

$$\max_x |F(x) - S_n(x)|$$

Existen, en muchas fuentes, tablas de los valores críticos del estadístico y aproximaciones para cuando se trabaja con muestras grandes, por ejemplo, en *Introduction to Mathematical Statistics* de Paul G. Hoel (1966).

Puede señalarse que al aplicar la prueba de Kolmogorov-Smirnov para bondad de ajuste a la distribución desarrollada en el tema 3.3, se obtiene un valor para el estadístico de prueba de 0.0501, cuando se rechazaría la distribución propuesta si dicho número fuera mayor a 0.0931. Así se cuenta con un indicador que respalda el modelo además de la evidencia gráfica.

Si la información se presentara como una observación censurada o truncada de los datos, existen también estadísticos Kolmogorov-Smirnov para ambos casos, para más información puede consultarse el libro *Survival Models and Data Analysis* de Regina C. Elandt-Johnson y Norman L. Johnson (1980).

Cabe señalar que existe un método gráfico para analizar de manera rápida la bondad de ajuste de una distribución a una serie de datos observada. Si se denota por $F_0(x)$ a una distribución observada, ya sea una tabla de mortalidad o una distribución aceptada para las observaciones en cuestión y por $\hat{F}(x)$ a la

distribución que se propone para ajustarse a las observaciones, entonces, si efectivamente $\hat{F}(x)$ fuera un buen candidato para ajustarse a las observaciones, en una gráfica de $F_0(x)$ contra $\hat{F}(x)$ debería suceder que los puntos con coordenadas $[F_0(x), \hat{F}(x)]$ estuvieran cerca de la diagonal entre los puntos (0,0) y (1,1). Este método es sencillo y puede sugerir continuar con otra clase de análisis, sin embargo, no existe un nivel de significancia en las conclusiones que puedan obtenerse.

3.5 Posibles aplicaciones de los conceptos de fiabilidad

Retomando las definiciones presentadas en el capítulo 1, se está en posibilidad de aplicar alguno de los resultados que existen en la literatura sobre fiabilidad. Puede comenzarse buscando las clases de fiabilidad a las que pertenece la distribución ajustada en la sección anterior.

Retomando, se dice que una distribución F es IFR (tasa de falla creciente) si $S(x|t)$ es decreciente en $0 \leq t < \infty$ para toda $x \geq 0$ y se considera que F es DFR (tasa de falla decreciente) si $S(x|t)$ es creciente en $0 \leq t < \infty$ para toda $x \geq 0$. Estas son, de algún modo, clasificaciones básicas que pueden ser útiles. A continuación se verifica si la distribución ajustada pertenece a alguna de estas clases de fiabilidad.

La función de supervivencia (fiabilidad) de la distribución Weibull está dada por:

$$S(t) = e^{-\frac{t^\alpha}{\beta}}$$

Entonces:

$$S(x|t) = \frac{S(t+x)}{S(t)} = \frac{e^{-\frac{(t+x)^\alpha}{\beta}}}{e^{-\frac{t^\alpha}{\beta}}} = e^{-\frac{(t+x)^\alpha}{\beta} + \frac{t^\alpha}{\beta}}$$

Para verificar si $S(x|t)$ es creciente o decreciente en t para toda $x \geq 0$ puede utilizarse el criterio de la derivada. Para simplificar el procedimiento se calcula el logaritmo natural de $S(x|t)$:

$$\ln(S(x|t)) = \ln\left(e^{\frac{-(t+x)^\alpha}{\beta} + \frac{t^\alpha}{\beta}}\right) = \frac{-(t+x)^\alpha}{\beta} + \frac{t^\alpha}{\beta}$$

Derivando respecto a t el resultado del desarrollo anterior:

$$\frac{d}{dt} \ln(S(x|t)) = -\frac{1}{\beta} (\alpha(t+x)^{\alpha-1}) + \frac{\alpha t^{\alpha-1}}{\beta} = \frac{\alpha}{\beta} [-(t+x)^{\alpha-1} + t^{\alpha-1}]$$

Ahora se verifica si la derivada es positiva o negativa:

$$\frac{\alpha}{\beta} [-(t+x)^{\alpha-1} + t^{\alpha-1}] \geq 0 \Leftrightarrow -(t+x)^{\alpha-1} + t^{\alpha-1} \geq 0 \Leftrightarrow t^{\alpha-1} \geq (t+x)^{\alpha-1}$$

$$\Leftrightarrow (t^{\alpha-1})^{\frac{1}{\alpha-1}} \leq ((t+x)^{\alpha-1})^{\frac{1}{\alpha-1}}$$

$$\Leftrightarrow t \leq t+x \Leftrightarrow 0 \leq x$$

Nótese que se invirtió la desigualdad en el segundo renglón de este último desarrollo porque el valor asignado al parámetro α es de 0.9052 y $\alpha - 1$ resulta negativo.

El último renglón indica que la derivada del logaritmo de la función $S(x|t)$ será positiva y, en consecuencia, la función $S(x|t)$ será creciente siempre y cuando $0 \leq x$, sin embargo, esta era una condición impuesta desde el inicio, así que puede considerarse que la distribución F es DFR.

Ahora, aprovechando la serie de implicaciones $DFR \Rightarrow DFRA \Rightarrow NWU \Rightarrow NWUE$, presente en el capítulo 1, puede señalarse que la distribución F es DFRA, NWU y NWUE sin necesidad de algún procedimiento.

Utilizando entonces que F es NWUE se puede recurrir a un resultado también presente en el capítulo 1 y que se enuncia del siguiente modo:

Teorema: Considérese un sistema operando durante un periodo de tiempo indefinido. En caso de falla, una reparación o reemplazamiento es llevado a cabo, tomando un tiempo despreciable. Los intervalos sucesivos entre fallas son variables aleatorias independientes e idénticamente distribuidas X_1, X_2, \dots de un proceso de renovación. Sea $N(t)$ que denota el número de renovaciones (reemplazamientos) en $(0, t]$ y $M(t)$ el número esperado de renovaciones en $(0, t]$, esto es, $E(N(t)) = M(t)$.

Sea $E(X) = \mu < \infty$ el tiempo de vida medio de un componente. Entonces, bajo el planteamiento anterior:

a) Si F es NBUE (nuevo mejor que utilizado en esperanza), entonces

$$\frac{t}{\mu} - 1 \leq M(t) \leq \frac{t}{\mu}$$

b) Si F es NWUE (nuevo peor que utilizado en esperanza), entonces

$$M(t) \geq \frac{t}{\mu}$$

Para aplicar este teorema a la información de la decimotercera aeronave puede comenzarse por calcular el valor de μ para la distribución Weibull con los parámetros $\alpha = 0.9052$ y $\beta = 58.3838$:

$$\mu = \frac{\beta^{\frac{1}{\alpha}}}{\alpha} \Gamma\left(\frac{1}{\alpha}\right) = \frac{58.3838^{\frac{1}{0.9052}}}{0.9052} \Gamma\left(\frac{1}{0.9052}\right) = 103.5765$$

Como se ha planteado que F es NWUE, entonces $M(t) \geq \frac{t}{103.5765}$.

Es decir, el número esperado de renovaciones (o reparaciones) que habrá sufrido el sistema de aire acondicionado de la aeronave número 13 será mayor o igual a $\frac{t}{103.5765}$, donde t es el número total de horas operativas desde que inició su funcionamiento.

Por ejemplo, los tiempos entre las primeras cinco reparaciones del sistema son: 102, 209, 14, 57 y 54. Así que en el momento $t = 436$ han sido necesarias 5 reparaciones y el teorema anterior señala que el número esperado de reparaciones para el valor $t = 436$ será mayor o igual a $\frac{t}{103.5765} = 4.2094$. De este modo se pueden prever las refacciones y tiempo de trabajo que se requerirán por concepto de reparaciones sabiendo el número de horas operativas que ha cumplido un sistema de aire acondicionado en cualquiera de las 13 aeronaves.

4. Dato Histórico

Introducción

En el año 1693 Mr. E. Halley publicó en el volumen 17 de Philosophical Transactions un artículo donde se utiliza un registro de los nacimientos y funerales de la ciudad de Breslaw, ubicada en la actual provincia de Baja Silesia al suroeste de Polonia para obtener tablas de mortalidad y comprobar los precios de las anualidades en esa época. El nombre del artículo es *“An Estimate of the Degrees of the Mortality of Mankind, drawn from curious Tables of the Births and Funerals at the City of Breslaw; with an Attempt to ascertain the Price of Annuities upon Lives”*.

4.1 Contenido del artículo

Mr. E. Halley retoma dos estudios relacionados con mortalidad referentes a las ciudades de Londres y Dublín, comentando respecto a ellos que tienen tres defectos importantes: primero, el número de personas era erróneo, segundo, no se contaba con las edades de las personas que fallecían y tercero, ambas ciudades recibían una gran cantidad de extranjeros que podían morir en ellas, lo cual se veía reflejado en un exceso de los funerales respecto a los nacimientos. Estas características hacían que la información utilizada fuera poco adecuada para un estudio de mortalidad.

Por otro lado, la ciudad de Breslaw se encuentra relativamente lejos del mar y en aquella época no se trataba de un centro de comercio, por lo cual la confluencia de extranjeros era poca. Adicionalmente, el registro utilizado por Mr. E. Halley contenía la información de las edades y sexos de las personas fallecidas y parecía estar hecho con la mayor exactitud posible. Por estas razones la información parecía ideal para un estudio de mortalidad.

Mr. E. Halley planteó una hipótesis sobre el crecimiento de la población tomando en cuenta el número de nacidos y fallecidos además de la situación de que algunos hombres eran llamados a servir en el ejército y al conjuntar su

análisis con el registro de nacimientos y muertes, obtuvo una tabla con los siguientes datos:

Edad	Muertos	Edad	Muertos	Edad	Muertos	Edad	Muertos
7	11	28	8	55	9	81	3
8	11		7	56	9		4
9	6	35	7		10	84	2
	5 ½	36	8	63	12		1
14	2		9 ½		9 ½	90	1
	3 ½	42	8	70	14	91	1
18	5		9	71	9	98	0
	6	45	7	72	11	99	1/5
21	4 ½		7		9 ½	100	3/5
	6 ½	49	10	77	6		
27	9	54	11		7		

Concentrando la información del número de personas por edad que mueren anualmente.

En los renglones donde no se especifica el dato *Edad* debe entenderse que la columna *Muertos* agrupa el número de decesos anuales para las edades comprendidas entre la edad anterior y la posterior.

Finalmente, se generó la siguiente tabla de mortalidad:

Edad	Vivos	Edad	Vivos	Edad	Vivos
1	1000	36	481	71	131
2	855	37	472	72	120
3	798	38	463	73	109
4	760	39	454	74	98
5	732	40	445	75	88
6	710	41	436	76	78
7	692	42	427	77	68
8	680	43	417	78	58
9	670	44	407	79	49
10	661	45	397	80	41
11	653	46	387	81	34
12	646	47	377	82	28
13	640	48	367	83	23
14	634	49	357	84	20
15	628	50	346		
16	622	51	335		
17	616	52	324		
18	610	53	313		
19	604	54	302		

20	598	55	292
21	592	56	282
22	586	57	272
23	579	58	262
24	572	59	252
25	567	60	242
26	560	61	232
27	553	62	222
28	546	63	212
29	539	64	202
30	531	65	192
31	523	66	182
32	515	67	172
33	507	68	162
34	499	69	152
35	490	70	142

Mr. E. Halley propuso varios usos para esta tabla, entre ellos señaló que se pueden obtener las posibilidades relativas de sobrevivir a uno o más años dependiendo de la edad de una persona, por ejemplo, una persona de 65 años tiene posibilidades de 182 a 10 de llegar a la edad de 66 años, ya que de 192 personas con vida a la edad de 65 años 10 fallecen en el transcurso de un año y 182 sobreviven.

Otro uso para la tabla es encontrar el número de años que se esperaría que una persona viviera dada su edad, el método que propone el autor es encontrar la edad en la cual el número de vivos de acuerdo a la tabla corresponde a la mitad del número de vivos en la edad de estudio, por ejemplo, en la edad de 40 años existen 445 sujetos vivos, dicho número se vería reducido a la mitad entre los 61 y 62 años, por lo cual, sería razonable esperar que una persona de 40 años viviera 21 o 22 años más.

Además de estas aplicaciones, Mr. E. Halley hace observaciones sobre el cálculo de anualidades, lo cual es el principal objetivo de su artículo y para lo cual puede considerarse una base fundamental el análisis de supervivencia.

Conclusiones

Desde hace muchos años el análisis de fiabilidad ha despertado el interés de especialistas en diversas disciplinas y hoy en día juega un papel fundamental en el área de la industria, las finanzas y la salud, entre otras. Las ventajas de su aplicación son plausibles y tienen una repercusión económica importante.

Los conceptos como función de densidad, distribución o supervivencia hacen directa la posibilidad de recurrir a la teoría sobre variables aleatorias existente aún sin pensar en fiabilidad, siendo útil esta abstracción para trabajar con rigurosidad y un campo de acción amplio. De manera particular puede señalarse la gran cantidad de distribuciones desarrolladas a lo largo de la historia, muchas de ellas surgidas a partir de escenarios ajenos a fiabilidad pero ahora utilizables en este campo.

Para los casos donde deba elegirse entre dos opciones descritas por variables aleatorias se ha presentado el concepto de envejecimiento relativo y los ordenamientos parciales, siendo auxiliares en la comparación del comportamiento de dos artículos por medio de criterios bien definidos.

Una de las formas con que cuenta el ser humano para influir en la fiabilidad de los artículos es el periodo de prueba, donde la duración del mismo puede decidirse después de realizar un análisis por medio de herramientas como la vida residual media o la tasa de falla, viéndose recompensado dicho estudio con la reducción de costos de reparación o garantía.

Aún con todo esto el desarrollo del análisis de fiabilidad no ha finalizado y las bases sólidas con las que cuenta dan lugar a una investigación continua obteniendo resultados que lo dotan de un mayor campo de acción.

En el caso práctico desarrollado puede notarse la sencillez con que se aplican los resultados de esta teoría, siendo fáciles de interpretar y debido a la relación que guardan entre ellos, la comprobación de un supuesto para un conjunto de datos da lugar a que se pueda utilizar más de un teorema. También se

comprueba, una vez más, lo adecuada que es la distribución Weibull para modelar tiempos de falla de sistemas electrónicos, además de ser un modelo parsimonioso y con una fácil estimación de sus parámetros. Así mismo se pone de manifiesto la unión o contención intrínseca entre el análisis de supervivencia y el análisis de fiabilidad; al estudiar el mismo tipo de situaciones, en algunas ocasiones con nomenclatura distinta, pero siempre con las mismas bases y resultados comunes.

Una diferencia entre fiabilidad y supervivencia surge al implementar políticas de mantenimiento, ya que la idea de “mantenimiento” pareciera tener poco sentido en un contexto de supervivencia y menos sentido tienen las ideas de reemplazamiento o reparación, sin embargo, en fiabilidad surgen de modo natural. A pesar de esto, cabe mencionar que en un contexto de supervivencia la inmunidad adquirida podría tener efectos equivalentes a un mantenimiento preventivo en términos estadísticos.

BIBLIOGRAFÍA

Burden Richard L. y Faires J. Douglas. **Análisis Numérico**. Séptima edición. Thomson Learning. México, 2004.

Cancela Rivas Gema. **Cáncer de Hígado en Hospitales Públicos de la Comunidad de Madrid: Análisis de Supervivencia**. Tesis doctoral. Facultad de medicina. Universidad Complutense de Madrid. España. 2011

D. R. Cox y P. A. W. Lewis. **The Statistical Analysis of Series of Events**. Methuen & Co. Ltd. Gran Bretaña, 1966.

Elandt-Johnson Regina C. y Johnson Norman L. **Survival Models and Data Analysis**. John Wiley & Sons, Inc. Estados Unidos, 1980.

Halley, E. **An Estimate of the Degrees of the Mortality of Mankind, drawn from curious Tables of the Births and Funerals at the City of Breslaw; with an Attempt to ascertain the Price of Annuities upon Lives**. Philosophical Transactions of the Royal Society, Vol. XVII, Pp. 596-610. Nueva York, 1963.

Hannon Bruce y Ruth Matthias. **Modeling and Simulation in Science and Mathematics Education**. Ed. Springer. Estados Unidos. 1999

Hoel Paul G. **Introduction to Mathematical Statistics**. Tercera edición. Wiley International Edition. Estados Unidos, 1966

Huber Catherine, Limnios Nikolaos, et al. **Mathematical Methods in Survival Analysis, Reliability and Quality of Life**. Wiley International Edition. Gran Bretaña, 2008.

Klein John P. y Moeschberger Melvin L. **Survival Analysis Techniques for Censored and Truncated Data**. Springer –Verlag New York, Inc. Estados Unidos. 1997

Kullback Solomon. **Information Theory and Statistics**. Wiley International Edition. Estados Unidos. 1959

Lai Chin-Diew y Xie Min. **Stochastic Ageing and Dependence for Reliability**. Ed. Springer. Estados Unidos, 2006.

Langlands A. O., Pocock S. J., Kerr G. y Gore S. M. **Long-Term Survival of Patients with Breast Cancer; a Study of Curability of the Disease**. British Medical Journal, 2, 1247-1251. Inglaterra, 1979.

Lee Elisa T. y Wenyu Wang John. **Statistical Methods for Survival Data Analysis**. Tercera edición. John Wiley & Sons, Inc. Estados Unidos. 2003

Machin David, Yin Bun Cheung, et al. **Survival Analysis A Practical Approach**. Segunda Edición. Wiley International Edition. Gran Bretaña, 2006.

Pentz M. J., F. Aprahamian, et al. **El Manejo de Datos Experimentales**. Traducción de Guillermo Fonthal. Ed. Mcgraw-Hill. México. 1974

Rivas López María Jesús. **Análisis de Supervivencia**. Ed. La Muralla. España, 2000.

Verdin Medina Luis Armando. **Estimación de Máxima Verosimilitud en la Distribución Weibull para Muestras Completas, Censuradas y su Aplicación en el Análisis de Tiempos de Vida**. Tesis de licenciatura. División de Ciencias Forestales. Universidad Autónoma Chapingo. México, 2005.

Young G. A. y Smith R. L. **Essentials of Statistical Inference**. Cambridge University Press. Estados Unidos, 2005.