

CIMAT

---

---

Centro de Investigación en Matemáticas, A.C.

# Contribuciones del Enfoque Bayesiano en Confiabilidad

Tesis

que para obtener el Grado de

Doctor en Ciencias

con Orientación en

Probabilidad y Estadística

Presenta

Humberto Gutiérrez Pulido

Director de Tesis

Dr. Víctor Aguirre Torres

Guanajuato, Gto., Abril del 2003

## AGRADECIMIENTOS

A Irma, mi esposa, por su apoyo para llevar a cabo este sueño.

Al Cimat, por ser el espacio propicio para aprender.

Al Dr. Víctor Aguirre, por dirigir la realización de este trabajo.

A Román de la Vara, por su amistad y apoyo.

A mis sinodales y profesores: Dr. Andrés Christen, Dr. Víctor Pérez-Abreu, Dr. Enrique Villa y Dr. Rogelio Ramos.

A mis colegas y amigos de la Universidad de Guadalajara: Osvaldo, Porfirio, y Agustín.

A Conacyt, por su apoyo (registro 56932).

A Dios, por todo.

## CONTENIDO

AGRADECIMIENTOS .....	ii
CAPITULO 1    PREFACIO .....	1
CAPITULO 2    ELEMENTOS GENERALES DE CONFIABILIDAD Y DE IN- FERENCIA BAYESIANA .....	3
2.1    Introducción .....	3
2.2    Distribución predictiva .....	4
2.3    Teoría de decisiones .....	5
2.4    Función de Verosimilitud Bajo Censura .....	6
CAPITULO 3    ANÁLISIS BAYESIANO DE ESPECIFICACIÓN DE MODE- LOS EN CONFIABILIDAD .....	9
3.1    Introducción .....	9
3.2    Factor de Bayes .....	13
3.3    Probabilidad Posterior para un Modelo .....	14
3.4    Métodos para Calcular la Verosimilitud Integrada .....	15
CAPITULO 4    VEROSIMILITUD INTEGRADA Y DENSIDAD POSTERIOR PREDICTIVA .....	17
4.1    Modelos Considerados .....	17
4.2    Muestras Completas .....	20
4.3    Muestras Censuradas .....	22
CAPITULO 5    ESPECIFICACIÓN DE LA DISTRIBUCIÓN <i>A PRIORI</i> PARA MODELOS EN CONFIABILIDAD .....	28
5.1    Introducción .....	28
5.2    Intervalos iniciales para la media y la desviación estándar .....	32
5.3    Modelo Normal .....	35
5.4    Modelo Lognormal .....	37
5.5    Modelo Valor Extremo .....	39

5.6	Modelo Weibull . . . . .	41
5.7	Modelo Exponencial . . . . .	44
5.8	Resumen . . . . .	45
5.9	Método general . . . . .	47
5.10	Evaluación de la metodología propuesta . . . . .	48
CAPITULO 6 APLICACIONES DE SELECCIÓN DE MODELOS . . . . .		55
6.1	Tabla de probabilidades posteriores . . . . .	55
6.2	Experimento sobre amortiguadores . . . . .	56
6.3	Experimento sobre aire acondicionado . . . . .	59
6.4	Ejemplos simulados . . . . .	62
CAPITULO 7 ESTUDIO MONTE CARLO PARA EVALUAR ESPECIFICACIÓN DE MODELOS . . . . .		66
7.1	Introducción . . . . .	66
7.2	A priori normal-gama . . . . .	67
7.3	A priori uniforme . . . . .	68
7.4	Conclusiones del Estudio . . . . .	69
CAPITULO 8 ELECCIÓN DE TIEMPOS DE GARANTÍA . . . . .		71
8.1	Introducción . . . . .	71
8.2	Función de Utilidad . . . . .	73
8.3	Cálculo de Utilidad Esperada . . . . .	78
8.4	Ejemplo . . . . .	81
CAPITULO 9 CONCLUSIONES GENERALES Y ESTUDIOS FUTUROS . . . . .		92
9.1	Especificación de Modelos y distribuciones <i>a priori</i> . . . . .	92
9.2	Tiempo de Garantía . . . . .	95

APENDICE . DEMOSTRACIONES .....	97
1.1 Obtención Analítica de la Verosimilitud Integrada .....	97
1.2 Densidad Posterior Predictiva .....	100
1.3 Resultados para la Aproximación de Laplace .....	104
1.4 Utilidad Esperada para el Modelo Exponencial .....	110
APENDICE . ALGORÍTMOS COMPUTACIONALES .....	112
2.1 Generación de una distribución normal-gama en S-Plus .....	112
2.2 Generación de una distribución uniforme en S-Plus .....	112
2.3 Generación de una muestra aleatoria con censura .....	112
2.4 Descripción de los programas para calcular la probabilidad posterior de un modelo y algunas experiencias .....	113
2.5 Descripción del programa para decidir el tiempo de garantía .....	114
REFERENCIAS .....	116

## CAPITULO 1

### PREFACIO

Uno de los campos que actualmente recibe más atención en la comunidad estadística es la confiabilidad, ya que en un contexto de mercados globalizado de productos, es necesario garantizar la lealtad de los clientes actuales y ganar nuevos, y eso se puede lograr a través de productos de calidad que tengan buen desempeño durante un tiempo suficientemente largo, es decir, mediante productos confiables. La confiabilidad como campo de estudio para la estadística tiene sus particularidades: variables aleatorias (tiempos de falla) que toman valores no negativos que suelen tener un comportamiento asimétrico, con sesgo positivo; muestreo censurado y pruebas de vida acelerada para acortar la duración de experimentos; tamaños muestrales pequeños; sistemas complejos en los que actúan muchos componentes con diferentes mecanismos y tiempos de falla, etcétera. Estas características distintivas de la confiabilidad hacen que mucha de la teoría y métodos de la estadística no sean aplicables de manera directa, por lo que se hace necesario adaptarlos o generar nuevos métodos que respondan a las particularidades del área.

En este contexto abordamos, desde la perspectiva Bayesiana, dos problemas que consideramos son vitales en confiabilidad: la especificación del modelo para los datos de tiempo de falla y la selección óptima del tiempo de garantía. Para el primer problema, como lo documentaremos en el capítulo 3, tanto desde la perspectiva clásica como Bayesiana se siguen generando propuestas de solución. En particular desarrollamos un método que ayuda a especificar modelos en confiabilidad que ofrece algunas ventajas respecto a otros trabajos que sobre la materia existen.

Respecto al otro problema, aunque se reconoce que la finalidad de muchos estudios de confiabilidad es decidir el tiempo de garantía, es común que se den recomendaciones ambiguas para su selección: muchas veces sólo se señala, que debe ser elegido con base en percentiles "bajos" de la distribución del tiempo de falla. En este trabajo abordamos este problema, y proponemos una metodología para su selección, que toma en cuenta la confiabilidad del producto, la competitividad del esquema de garantía a los ojos del consumidor (mercadotecnia), efecto sobre la imagen de la empresa cuando el producto falla dentro del período de garantía y los costos en los que incurre el fabricante para cumplir con la garantía prometida.

Para ambas contribuciones hemos considerado cinco de los modelos más usuales en confiabilidad: normal, lognormal, valor extremo, Weibull y exponencial, y contemplamos muestreo completo y con censura: tipo I, II y aleatoria. Veremos que los resultados obtenidos se pueden ampliar hacia otros modelos y tipos de censura.

Para completar las propuestas para la especificación de modelos y la elección del tiempo de garantía fue necesario contar con una forma de definir los hiperparámetros de las distribuciones *a priori* para los parámetros de los modelos considerados. Sin embargo, como los trabajos existentes sobre el particular no cumplían con el enfoque del trabajo, fue necesario desarrollar una propuesta. Para generar esta metodología partimos de que el usuario estaba en posibilidades de proporcionar información sobre intervalos para los cuantiles de las observables, es decir, sobre el tiempo de vida  $X$ , más que de los no observables -los parámetros de los modelos de tiempo de vida-, y a partir de esto se propone una metodología para definir los parámetros de dos tipos

de distribución *a priori*. (uniforme y normal-gama) para los parámetros de los cinco modelos considerados en el trabajo.

En suma en el trabajo hacemos tres contribuciones, una para especificación de modelos, otra para selección del tiempo de garantía y una tercera sobre especificación de las distribución *a priori* para los modelos de confiabilidad estudiados en este trabajo. Esta última, aunque complementa a las dos primeras, tiene su relevancia por sí misma para propósitos de inferencia Bayesiana en confiabilidad. Es hasta los capítulos en los que se desarrolla la parte central de cada una de las contribuciones, donde se planteará con detalle cada problema, se hará una revisión bibliográfica, se generarán y evaluarán las propuestas.

El trabajo se ha organizado de la siguiente manera: en el capítulo 2, con el propósito de hacer más autocontenido el trabajo, vemos elementos de inferencia Bayesiana y generalidades de confiabilidad. En el capítulo 3 abordamos el problema de especificación de modelos, y partiendo de una revisión bibliográfica sobre el tema, establecemos los elementos centrales de la inferencia Bayesiana en cuanto a este tópico. En el cuarto capítulo, desarrollamos los resultados principales que permiten calcular la verosimilitud integrada, que será la base para calcular la probabilidad posterior para un modelo, dados los datos. También se obtendremos para algunos de los modelos considerados la densidad posterior predictiva que será uno de los elementos claves para la elección del tiempo de garantía. En el capítulo cinco proponemos una metodología, en la que a partir de información inicial sobre el intervalo en el que se espera que ocurran algunos cuantiles del tiempo de vida, se calculan los parámetros de dos tipos de distribuciones *a priori* (normal-gama y uniforme) para los parámetros de los cinco modelos (normal, lognormal, valor extremo, Weibull y exponencial). Como se verá este método ofrece algunas ventajas sobre otras propuestas que para el mismo propósito existen, y será aplicado tanto en especificación de modelos como en el tiempo de garantía. En el capítulo 6 se aplicamos la metodología de selección de modelos a ejemplos que han aparecido en la literatura, y veremos que la metodología funciona satisfactoriamente en ellos.

En el séptimo capítulo hacemos un estudio Monte Carlo para considerar el efecto del muestreo en la metodología propuesta de selección de modelos. En el octavo capítulo se concentramos los principales aspectos referentes al tiempo de garantía: revisión bibliográfica del problema, se propone una función de utilidad que consideran los diferentes aspectos de confiabilidad, costos, mercadotecnia e imagen, que están relacionados con la decisión de la garantía. También obtenemos algunos resultados y un par de algoritmos para el cálculo y maximización de la utilidad esperada, que será la base para la decisión. Al final del capítulo 8 se presenta un ejemplo completo en el que se decidimos el tiempo óptimo de garantía. Uno de los algoritmos para la obtención de la utilidad esperada se basa en el método que se conoce como *sampling-importance-resampling*. En el capítulo 9 presentamos las conclusiones del trabajo y algunas investigaciones futuras que se consideramos de interés. Adicionalmente escribimos un apéndice en el que están las demostraciones de algunos resultados que obtuvimos en el trabajo. También escribimos un segundo apéndice en el que se describen algunos algoritmos computacionales y algunas de las experiencias que obtuvimos sobre el particular.

## CAPITULO 2

# ELEMENTOS GENERALES DE CONFIABILIDAD Y DE INFERENCIA BAYESIANA

En este capítulo se presentan los aspectos principales de la inferencia Bayesiana y algunos conceptos generales sobre la función de verosimilitud. No es objetivo del capítulo presentar una versión exhaustiva, sino sólo elementos generales que permiten que el trabajo sea más autocontenido. El lector interesado en profundizar en estas áreas puede consultar algunas de los libros en las que nos hemos basado para escribir este capítulo. Para lo referente a inferencia Bayesiana se puede consultar Bernardo y Smith (1994), DeGroot (1970) y Berger(1985). Sobre Confiabilidad Lawless (1982) y Meeker y Escobar (1998). Ambos temas se tratan en libros como Ibrahim et al. (2001), Martz (1982) y Spizzichino(2001).

### 2.1 Introducción

La inferencia Bayesiana está basado en especificar un modelo de probabilidad  $f(x|\theta)$  para los datos observados,  $\mathbf{X}$ ; dado un vector de parámetros  $\theta$  de valor desconocido. Generándose así la función de verosimilitud,  $L(\mathbf{X}|\theta)$ , que es la distribución conjunta condicional de  $\mathbf{X}$  para un valor dado de  $\theta$ . Entonces se supone que  $\theta$  es una variable aleatoria, que tiene una distribución *a priori* o inicial denotada por  $\pi(\theta)$ . Es precisamente a través de  $\pi(\theta)$  como el investigador incorpora formalmente el conocimiento inicial que tiene sobre el fenómeno bajo estudio. De esta manera las inferencias respecto a  $\theta$  se basan en la distribución *a posteriori* o posterior, la cual es obtenida por el teorema de Bayes:

Si  $H$  denota una hipótesis y  $\mathbf{X}$  los datos, entonces el teorema de Bayes, en su forma más simple (Bernardo y Smith, 1994), establece que

$$P(H|\mathbf{X}) = \frac{P(\mathbf{X}|H)P(H)}{P(\mathbf{X})} \quad (2.1)$$

donde  $P(H)$  es una afirmación probabilística de certidumbre acerca de  $H$  antes de obtener los datos  $\mathbf{X}$ , y  $P(H|\mathbf{X})$  se convierte en una afirmación probabilística de creencia acerca de  $H$  después de haber obtenido los datos. Habiendo especificado  $P(H)$  y  $P(\mathbf{X}|H)$ , el teorema de Bayes proporciona una solución al problema de qué tanto se aprende de los datos.

**Teorema 2.1** *La distribución posterior  $\pi(\theta|\mathbf{X})$  es la distribución condicional de  $\theta$  dados los datos, y está dada por*

$$\pi(\theta|\mathbf{X}) = \frac{L(\mathbf{X}|\theta)\pi(\theta)}{\int_{\Theta} L(\mathbf{X}|\theta)\pi(\theta)d\theta} \quad (2.2)$$

donde  $\Theta$  es el espacio paramétrico de  $\theta$ .

Este resultado se desprende directamente del teorema de Bayes. De (2.2) es claro que  $\pi(\theta|\mathbf{X})$  es proporcional a la verosimilitud multiplicada por la distribución *a priori*,

$$\pi(\boldsymbol{\theta} | \mathbf{X}) \propto L(\mathbf{X} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) \quad (2.3)$$

De aquí que el denominador de (2.2), que es la distribución conjunta marginal de los datos,

$$f(\mathbf{X}) = \int_{\Theta} L(\mathbf{X} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (2.4)$$

que también recibe el nombre de constante normalizadora de  $\pi(\boldsymbol{\theta} | \mathbf{X})$ . Otro nombre que recibe  $f(\mathbf{X})$  es el de Verosimilitud Integrada (VI), que es el que usaremos nosotros.

Habíamos dicho que en estadística Bayesiana las inferencias respecto a  $\boldsymbol{\theta}$  están basados en  $\pi(\boldsymbol{\theta} | \mathbf{X})$ , que como es evidente a partir de (2.3), ésta resulta de la contribución de los datos a través de  $L(\mathbf{X} | \boldsymbol{\theta})$  y del aporte de la información o conocimiento *a priori* cuantificado por  $\pi(\boldsymbol{\theta})$ .

## 2.2 Distribución predictiva

En muchas aplicaciones, después del muestreo, en lugar de hacer inferencias sobre  $\boldsymbol{\theta}$ , es de interés hacer inferencias sobre un valor futuro de la variable  $X$ . En este sentido, supongamos que  $y$  con distribución  $f(y | \boldsymbol{\theta})$  representa un valor futuro de la variable aleatoria  $X$ . Advertimos que para simplificar notación al referirnos al vector de parámetros  $\boldsymbol{\theta}$  en esta sección, nos estamos refiriendo a su estado después del muestreo, es decir, a  $(\boldsymbol{\theta} | \mathbf{X})$ .

**Teorema 2.2** *Bajo el supuesto de que hay independencia entre los datos,  $\mathbf{X}$ , y la observación futura,  $y$ , la distribución posterior predictiva está dada por*

$$f(y | \mathbf{X}) = \int_{\Theta} f(y | \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | \mathbf{X}) d\boldsymbol{\theta} \quad (2.5)$$

La demostración de este resultado básico de la inferencia Bayesiana se puede consultar en Bernardo y Smith (1994, pag. 242). Notemos que en realidad la distribución posterior predictiva  $f(y | \mathbf{X})$  es el valor esperado de  $f(y | \boldsymbol{\theta})$  respecto a  $\boldsymbol{\theta}$  dado  $\mathbf{X}$ , es decir

$$E_{\boldsymbol{\theta} | \mathbf{X}} [f(y | \boldsymbol{\theta})] = \int_{\Theta} f(y | \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | \mathbf{X}) d\boldsymbol{\theta} \equiv f(y | \mathbf{X}).$$

En la mayoría de los modelos y aplicaciones la distribución conjunta marginal  $f(\mathbf{X})$  no tiene una forma analítica cerrada, por lo tanto  $\pi(\boldsymbol{\theta} | \mathbf{X})$  tampoco la tiene, mucho menos  $f(y | \mathbf{X})$ . Esto genera una gran cantidad de literatura que plantean diferentes alternativas para estimar  $f(\mathbf{X})$  y  $f(y | \mathbf{X})$ , o para obtener muestras de  $\pi(\boldsymbol{\theta} | \mathbf{X})$  sin conocer su forma cerrada.

En el capítulo 3 de este trabajo se verá que en estadística Bayesiana es necesaria la estimación de  $f(\mathbf{X})$  para poder calcular la probabilidad posterior de un modelo, dados los datos. Esta probabilidad posterior será la base para especificar modelos en confiabilidad.

En el capítulo 8 se verá que para la mayoría de los modelos considerados en este trabajo será necesario obtener por simulación  $f(y|\mathbf{X})$ , a fin de poder seleccionar el tiempo óptimo de garantía para un producto.

Para algunas aplicaciones es necesario conocer la *densidad a priori predictiva*,  $f(x)$ , que de acuerdo al teorema anterior, está dada por

$$f(x) = \int_{\Theta} f(x|\theta)\pi(\theta)d\theta \quad (2.6)$$

y representa el modelo del tiempo de vida, *a priori* (usando la distribución *a priori* de  $\theta$  en lugar de su densidad posterior).

### 2.3 Teoría de decisiones

Por lo general el objetivo último de un análisis estadístico es tomar una decisión, como por ejemplo fijar un tiempo de garantía, establecer políticas de mantenimiento preventivo en un proceso, determinar los puntos más débiles de un sistema para mejorar la confiabilidad de un producto o un proceso, etcétera. Obviamente es deseable que estas decisiones estén fundamentadas en un buen conocimiento sobre la realidad.

A continuación veremos brevemente los elementos y procedimientos necesarios para un esquema coherente de toma de decisiones de tipo cuantitativo. Los elementos de una decisión en el contexto de inferencia son:

- i)  $a \in A$ , las posibles "respuestas" o "decisiones".
- ii)  $\omega \in \Omega$ , estados desconocido de la realidad.
- iii)  $u : A \times \Omega \rightarrow R$ , una función que vincula la utilidad o ganancia de cada  $(a, \omega)$ , es decir, que evalúa las consecuencias de una "respuesta"  $a$  y una situación resultante de la realidad  $\omega$ .
- iv)  $p(\omega)$ , una especificación, en la forma de una distribución de probabilidad, de los conocimientos actuales acerca de los posibles estados de la realidad.

La elección óptima de la respuesta para un problema de inferencia es aquella  $a \in A$  que *maximiza la utilidad esperada*,

$$u^*(a) = \int_{\Omega} u(a, \omega)p(\omega)d\omega. \quad (2.7)$$

Alternativamente, si en lugar de trabajar con  $u(a, \omega)$ , se trabaja con la llamada función de pérdida

$$l(a, \omega) = h(\omega) - u(a, \omega),$$

donde  $h(\cdot)$  es una función fija arbitraria, la elección óptima de la respuesta es aquella  $a \in A$  que *minimiza la pérdida esperada*,

$$l^*(a) = \int_{\Omega} l(a, \omega)p(\omega)d\omega. \quad (2.8)$$

En el contexto de inferencia paramétrica o predictiva, el estado desconocido de la realidad son parámetros o valores futuros de una variable (observables). Por lo tanto, las creencias actuales,  $p(\omega)$ , se reducen a alguna de las densidades de probabilidad que hemos visto antes:

- $\pi(\theta)$  conocimiento inicial sobre un vector de parámetros,  $\theta$ ;
- $\pi(\theta|\mathbf{X})$  conocimiento acerca de  $\theta$ , dados los datos  $\mathbf{X}$ ;
- $\pi(\Psi|\mathbf{X})$  conocimiento acerca de  $\Psi = g(\theta)$ , dados los datos  $\mathbf{X}$ ;
- $f(y|\mathbf{X})$  conocimiento sobre un dato futuro  $y$ , dados los datos  $\mathbf{X}$ .

En el capítulo 8 se verá que para decidir el tiempo óptimo de garantía,  $t_g$ , además de especificar una adecuada función de utilidad,  $u(t_g, t)$ , que nos relacione la decisión  $t_g$  con el tiempo de vida de un producto  $t$ , se va a requerir la densidad posterior predictiva,  $f(y|\mathbf{X})$ . Ya que ésta va reflejar el conocimiento actual sobre el tiempo de vida del producto.

#### 2.4 Función de Verosimilitud Bajo Censura

Una de las características principales de los datos de confiabilidad es que por lo general son censurados. Y esto ocurre cuando en un experimento de confiabilidad el tiempo de falla es conocido sólo para una parte de los especímenes, mientras que para el resto sólo se sabe que su tiempo de falla es menor o mayor que cierto valor. Formalmente se dice que una observación  $x_i$  es censurada por la derecha en  $q$  si el valor exacto de la observación no se conoce, y sólo se sabe que es mayor que  $q$  ( $x_i > q$ ). Similarmente, se dice que una observación  $x_i$  es censurada por la izquierda en  $q$  si el valor exacto de la observación no se conoce, y sólo se sabe que es menor que  $q$  ( $x_i < q$ ).

Censura tipo I. Este es un tipo de censura muy usual. Una muestra con censura múltiple tipo I surge cuando los especímenes 1, 2, ...,  $n$  de un experimento de confiabilidad están sujetos a períodos límite de observación  $q_1, q_2, \dots, q_n$ ; de tal forma que los tiempos individuales de falla,  $x_i$ , son observados sólo si  $x_i \leq q_i$ . Cuando los límites de censura son iguales,  $q_1 = q_2 = \dots = q_n = q$ , se habla de censura simple tipo I. Si  $\mathbf{X} = (x_1, \dots, x_n)$  es una muestra aleatoria de  $f(x|\theta)$ ,  $x_i$  se observará sólo si  $x_i \leq q_i$ . Supongamos que se observan exactamente  $r$  fallas  $0 \leq r \leq n$ , entonces sea  $x_{(1)}, \dots, x_{(r)}$  los  $r$  tiempos de falla ordenados en forma creciente, y también ordenemos los  $n - r$  tiempos de censura en forma ascendente:  $q_{(1)}, q_{(2)}, \dots, q_{(n-r)}$ . A través de función de verosimilitud se recoge toda la información que sobre  $\theta$  aportan los datos. Se sabe que la verosimilitud es proporcional a la probabilidad de observar lo que se observó (Sprott, 2000, pag. 8). Por ello la verosimilitud aportada por los tiempos de falla observados se recoge a través de  $f_j(x_{(i)}|\theta)$ , y para las observaciones censuradas por  $\Pr(X > q_{(k)}) = 1 - F_j(q_{(k)}|\theta)$ . De aquí que la función de verosimilitud, bajo el modelo  $M_j$ , para una muestra con censura múltiple tipo I está dada por

$$L(\mathbf{X} | \theta, M_j) = \prod_{i=1}^r f_j(x_{(i)} | \theta) \prod_{k=1}^{n-r} (1 - F_j(q_{(k)} | \theta)) \quad (2.9)$$

Detalles adicionales para la obtención de esta expresión pueden consultarse en Lawless (1982, pag. 34). En caso que se tenga un censuramiento simple: ( $q_1 = q_2 = \dots = q_n =$

q), entonces (2.9) toma la siguiente forma

$$L(\mathbf{X} | \theta, M_j) = (1 - F_j(q | \theta))^{n-r} \prod_{i=1}^r f_j(x_{(i)} | \theta) \quad (2.10)$$

Censura tipo II. Esta es una censura por la derecha en el que el experimento corre hasta que una cierta cantidad fija,  $r$ , de unidades falla de  $n$  disponibles ( $r \leq n$ ). Sea  $x_{(1)}, \dots, x_{(r)}$  los  $r$  tiempos de falla ordenados en forma creciente, entonces la verosimilitud estará dada por

$$L(\mathbf{X} | \theta, M_j) = C(1 - F_j(x_{(r)} | \theta))^{n-r} \prod_{i=1}^r f_j(x_{(i)} | \theta) \quad (2.11)$$

donde la constante  $C = \frac{n!}{(n-r)!}$ ,  $r \leq n$  (ver Lawless, 1982, pag. 32). En este tipo de censura la duración del experimento,  $T_e$ , es una variable aleatoria y es igual a la  $r$ -ésima estadística de orden,  $T_e = x_{(r)}$ . Nótese que esta verosimilitud es similar, excepto por la constante  $C$ , a la verosimilitud para censura simple tipo I (expresión 2.10); haciendo en ambos casos  $T_e = x_{(r)} = q$ .

Censura aleatoria. En algunos estudios los tiempos de censura son aleatorios. Por ejemplo, si las unidades experimentales entran al estudio de acuerdo al tiempo en que se vendieron, y el estudio se termina en cierta fecha preestablecida; entonces los tiempos de censura son aleatorios; esto es el tiempo transcurrido desde que un artículo entra al estudio hasta que finaliza el experimento, es aleatorio. Otra situación en la que aparece la censura aleatoria, es cuando aunque todas las unidades experimentales entran al mismo tiempo al estudio por alguna razón externa (independiente de  $\theta$ ) puede ocurrir que ya no es posible darle seguimiento a algunas unidades experimentales. Para propósitos de inferencia uno puede trabajar condicionalmente sobre los tiempos de censura, procediendo como si el censuramiento fuera del tipo I, pero teniendo presente el mecanismo por el que se generaron los datos a fin de justificar tal proceder (ver Lawless, 1982). Entonces bajo el tipo de censuramiento aleatorio considerado antes, la verosimilitud estaría dada por (2.9).

Además, por el principio de verosimilitud, desde el punto de vista la inferencia Bayesiana, la censura múltiple tipo I es igual a la censura aleatoria, ya que la distribución posterior  $\pi(\theta | \mathbf{X})$  depende de los datos sólo a través de la verosimilitud  $L(\mathbf{X} | \theta)$ , que en ambos casos es igual.

Censura por intervalo. En algunos estudios de confiabilidad las fallas sólo son detectadas en los tiempos de inspección, de tal forma que por lo general no se detecta con precisión el tiempo de falla, sólo se sabe que ocurrió en un cierto intervalo. Esto genera datos censurados por intervalo, o también se conocen como datos de inspección o datos agrupados. Este tipo de censura es una combinación de censura por la izquierda y censura por la derecha. Por ejemplo, si los tiempos de inspección están dados por  $t_1, t_2, \dots, t_p$ , y en la primera inspección en el tiempo  $t_1$  se observa que  $d_1$  unidades ya fallaron, lo que se sabrá es que su tiempo de falla ocurrió entre 0 y  $t_1$ , y como  $\Pr(0 < X \leq t_1) = F(t_1) - F(0)$ , entonces su correspondiente verosimilitud estará dada

por  $[F(t_1) - F(0)]^{d_1}$ . Si en general en el intervalo  $i$ -ésimo,  $(t_{i-1}, t_i]$ , se detectaron  $d_i$  unidades que fallaron, entonces la verosimilitud estará dada por:

$$L(\mathbf{X} | \theta, M_j) = (1 - F(t_p | \theta))^{n - \sum_{j=1}^p d_j} \prod_{i=1}^p (F(t_i | \theta) - F(t_{i-1} | \theta))^{d_i} \quad (2.12)$$

Note que el primer factor de la expresión anterior corresponde a la verosimilitud de las unidades que no fallaron hasta el último tiempo de inspección.

## CAPITULO 3

### ANÁLISIS BAYESIANO DE ESPECIFICACIÓN DE MODELOS EN CONFIABILIDAD

En este capítulo se presentan los resultados generales referentes a especificaciones de modelos en confiabilidad, desde una perspectiva Bayesiana.

#### 3.1 Introducción

Una revisión de la literatura clásica de confiabilidad refleja que es práctica común utilizar gráficas de probabilidad con bandas de confianza para evaluar la especificación de la distribución del tiempo de vida, ver por ejemplo Meeker y Escobar (1998, cap. 6). Este tipo de gráficas son útiles para detectar discrepancias fuertes respecto al modelo de los datos, presencia de mezclas de distribuciones (Doganaksoy et. al. 2002), pero son subjetivas y pueden tener poco poder para detectar errores en la especificación. Veamos el siguiente ejemplo.

Muestra	Datos														
1	0.01	0.11	0.13	0.15	0.17	0.32	0.33	0.41	0.42	0.44	0.45	0.48	0.56	0.60	0.80
2	0.04	0.12	0.17	0.20	0.20	0.24	0.29	0.43	0.44	0.45	0.60	0.64	0.68	0.72	1.03
3	0.03	0.09	0.10	0.13	0.14	0.17	0.29	0.31	0.33	0.43	0.53	0.60	0.78	0.78	0.88
4	0.05	0.06	0.19	0.25	0.26	0.30	0.35	0.42	0.70	0.73	0.78	0.90	1.03	1.41	1.45

Tabla 3.1. Muestras generadas de una distribución exponencial ( $n = 15$ ).

Ejemplo 3.1. En la tabla 3.1 se muestran datos simulados de una distribución exponencial y en la figura 3.1 se aprecian las correspondientes gráficas de probabilidad normal. Con base en estas gráficas es difícil rechazar la idea de que estos datos provengan de una distribución normal. Remarcando el hecho de que esto ocurre con una distribución como la exponencial, que tiene una forma radicalmente diferente a la normal.

Ejemplo 3.2. Una de las características particulares de los datos en confiabilidad es que por lo general son muestras censuradas. Por ejemplo los datos de la tabla 3.2 representan la distancia de falla en kilómetros de 38 amortiguadores de vehículos (O'Connor, 1985; pag. 85). Los datos que tienen como exponente un signo +, son tiempos censurados por la derecha, es decir, a esa distancia el amortiguador todavía estaba en buen estado. Con el esquema de censuramiento múltiple por la derecha, de los 38 amortiguadores, sólo se observaron 11 fallas. Utilizando el método Kaplan-Meier para definir las posiciones de probabilidad (Meeker y Escobar, 1998, secc. 3.5), en la figura 3.2 se muestra las correspondientes gráficas de probabilidad con bandas de confianza generadas por el procedimiento *probplot.censor* del sistema S-Plus. Teniendo como alternativa cuatro posibles modelos: (a) normal, (b) lognormal, (c) valor extremo y (d) Weibull. De acuerdo a estas gráficas prácticamente cualquiera de estas distribuciones es factible, quizás la valor extremo pudiera descartarse por el punto fuera de las bandas.

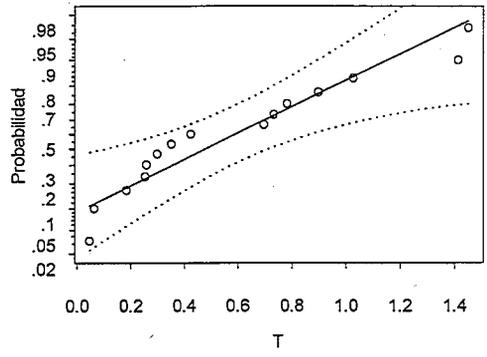
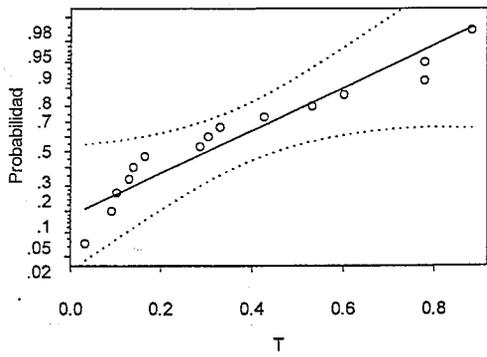
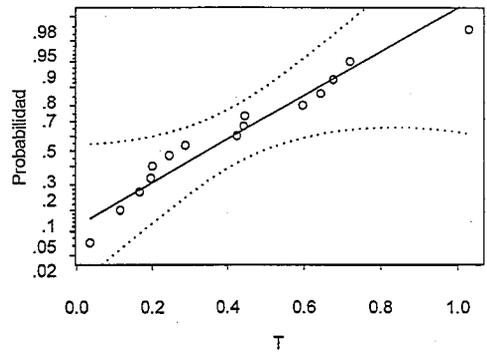
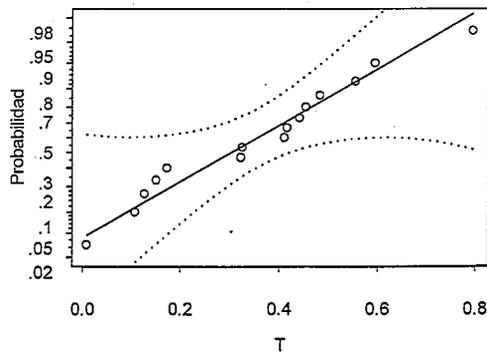


Figura 3.1: Gráfica de probabilidad normal para datos de la tabla 3.1.

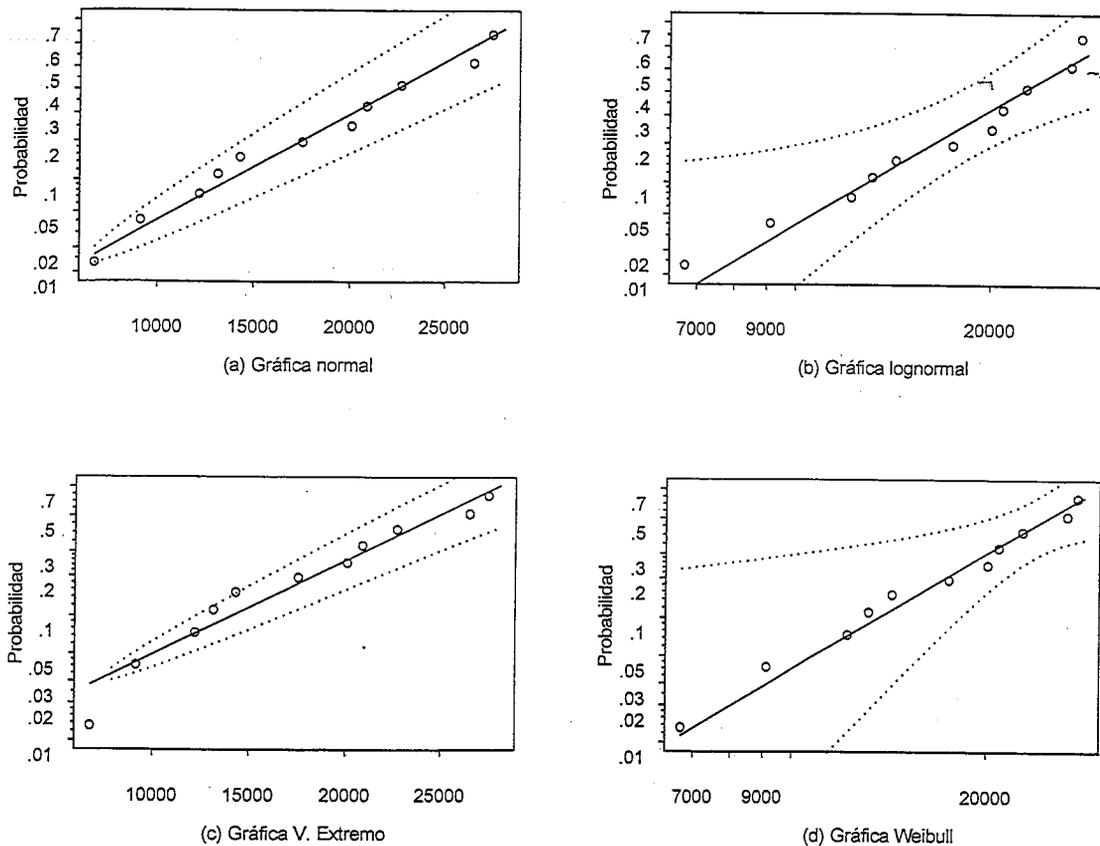


Figura 3.2: Gráficas de probabilidad para datos tabla 3.2.

6700, 6950 <sup>+</sup> , 7820 <sup>+</sup> , 8790 <sup>+</sup> , 9120, 9660 <sup>+</sup> , 9820 <sup>+</sup> , 11310 <sup>+</sup> , 11690 <sup>+</sup> , 11850 <sup>+</sup> , 11880 <sup>+</sup> , 12140 <sup>+</sup> , 12200, 12870 <sup>+</sup> , 13150, 13330 <sup>+</sup> , 13470 <sup>+</sup> , 14040 <sup>+</sup> , 14300, 17520, 17540 <sup>+</sup> , 17890 <sup>+</sup> , 18420 <sup>+</sup> , 18960 <sup>+</sup> , 18980 <sup>+</sup> , 19410 <sup>+</sup> , 20100, 20100 <sup>+</sup> , 20150 <sup>+</sup> , 20320 <sup>+</sup> , 20900, 22700, 23490 <sup>+</sup> , 26510, 27410 <sup>+</sup> , 27490, 27890 <sup>+</sup> , 28100 <sup>+</sup>
---

Tabla 3.2. Distancia de falla para 38 amortiguadores de vehículos.

Con base en los dos ejemplos anteriores, vemos que se requiere contar con herramientas más formales que ayuden en la especificación del modelo para datos de confiabilidad. Desde la perspectiva del enfoque clásico de bondad de ajuste para muestras completas hay muchas alternativas, por ejemplo en Barrio et al. (2000) se hace una amplia revisión sobre la evolución de la teoría asintótica para pruebas de bondad de ajuste y su paralelismo con el desarrollo de la teoría de procesos empíricos. Se aprecia que estas pruebas parten de especificar como hipótesis nula a un modelo  $F_0$ , y se construyen distintos estadísticos que ponen diferente énfasis en la detección de desviaciones de  $F_0$  respecto a la distribución empírica  $F_n$  (desviaciones largas en cualquier punto, desviaciones promedios, mayor peso en las colas, etcétera). De aquí se desprende que para cada modelo  $F_0$  es necesario encontrar la distribución asintótica de cada estadístico, que para pequeñas muestras puede haber casos con baja potencia y para muestras grandes

demasiada potencia. Esto hace que en ocasiones todos los modelos  $F_0$  propuestos se acepten, o bien en el otro extremo, que todos se rechacen. Esto obviamente genera un problema de decisión al usuario de la estadística. Por otro lado, para el caso de muestras censuradas (que no se aborda en el trabajo citado antes), existen trabajos para algunas distribuciones, pero con limitaciones importantes en cuanto a los tamaños de muestra considerados, el nivel y tipo de censuramiento considerado. Por ejemplo, en Anaya y O'Reilly(2001) se aborda el problema de bondad de ajuste para las distribuciones Gaussiana inversa y gama, bajo una muestra con censura tipo I, donde obtienen que la distribución asintótica de los estadísticos de prueba (Anderson-Darling y Cramér-von Mises) depende de la relación de magnitud entre los dos parámetros de esta distribución. Por lo tanto los percentiles de los estadísticos de prueba, obtenidos por simulación, dependen de la significancia, del tamaño de muestra, del nivel de censuramiento y de la relación entre los parámetros. En Pettit(1976) se proponen generalizaciones para los estadísticos Anderson-Darling y Cramér-von Mises para probar normalidad cuando se tienen censura tipo I y II, obteniendo percentiles asintóticos para los mismos, sin embargo la velocidad de convergencia a esos resultados asintóticos son lentos (Lawless, 1982, pag. 438). Otras pruebas de bondad de ajuste para algunos modelos, suponiendo datos censurados, se han hecho teniendo como hipótesis nula a una distribución particular contra un modelo más general que lo incluye, por ejemplo: exponencial contra Weibull, valor extremo contra. log-gama de tres parámetros; este tipo de pruebas tienen la limitante de que no consideran otros modelos como alternativa. En Mann, et al. (1973) desarrollan un estadístico de prueba para la distribución valor extremo considerando censura tipo II; los percentiles del estadístico dependen de la significancia, del nivel de censuramiento y del tamaño de muestra (consideran tamaños de muestra de 3 a 25). Los trabajos que se han desarrollado desde el enfoque de agrupamiento de datos (prueba  $\chi^2$ ) tienen las limitantes de la cantidad de datos y que contemplan esquemas de censuramiento limitado (ver por ejemplo Mihalko y Moore-1980). En suma, en el enfoque clásico cada distribución es un problema diferente, y los resultados que se tienen sólo contemplan opciones limitadas en cuanto al tipo de censuramiento, el nivel de censuramiento y los tamaños muestrales. Además puede darse el caso que todos los modelos propuestos se acepten o por el contrario que todos se rechacen.

Adicionalmente uno de los enfoques propuestos para selección de modelos no anidados, desde la perspectiva clásica, ha sido la generalización del enfoque de razón de verosimilitudes. Sin embargo, además de fundamentarse en resultados asintóticos, el hecho de que uno de los dos modelos tenga que designarse como el "verdadero" y el otro como el "alternativo", influye dramáticamente en los resultados (Rossi, 1989). Esto se debe a la asimetría de la prueba Neyman-Pearson.

En este trabajo se estudia un enfoque Bayesiano aplicado al problema de selección de modelos como herramienta para evaluar la especificación del modelo de confiabilidad. Sobre el particular, existen varios trabajos con diferentes enfoques; por ejemplo Geisser y Eddy (1979), Gelfand y Dey (1994), Berger y Pericchi (1996), Sinha et al. (1999), Hoeting et al. (1999), Gutiérrez-Peña y Walker (2001), Pérez y Berger (2002). Las razones de los diferentes enfoques son (ver Pérez y Berger, 2002): la dificultad para incorporar la información *a priori* para los diferentes parámetros de los modelos, las complicaciones para calcular la probabilidad posterior para los modelos, la dificultad

de utilizar distribuciones iniciales impropias, si se utiliza distribuciones *a priori* propias pero vagas el poder discriminatorio de la metodología es bajo. Lo anterior hace que se sigan proponiendo y explorando alternativas.

En este trabajo se aborda la problemática de especificación de modelos en confiabilidad desde la perspectiva Bayesiana (siguiente sección), se propone una forma de incorporar la información inicial (capítulo 5), se estudian los modelos normal, lognormal, valor extremo, Weibull y exponencial y se verá que la probabilidad posterior para algunos de ellos tiene forma cerrada, en otros casos se obtienen aproximaciones (capítulo 4). También se evalúa la metodología analizando datos de problemas de confiabilidad (capítulo 6) y haciendo un estudio Monte Carlo (capítulo 7) .

### 3.2 Factor de Bayes

El enfoque Bayesiano para probar hipótesis fue desarrollado por Jeffreys como parte fundamental de su programa de inferencia científica (Kass y Raftery, 1995); Jeffreys estaba interesado en comparar las predicciones que se obtenían con dos teorías científicas en competencia. En su enfoque, se introducen modelos estadísticos para representar la probabilidad de los datos de acuerdo a cada una de las dos teorías, y el teorema de Bayes es utilizado para calcular la probabilidad posterior de que una de las dos teorías sea la correcta. Una pieza central en esto fue el llamado factor de Bayes. Veamos.

Sea  $\mathbf{X}$  los datos y supongamos que éstos han sido obtenidos de una de dos hipótesis  $H_1$  y  $H_2$ , de acuerdo con la densidad de probabilidad  $\Pr(X | H_1)$  o  $\Pr(X | H_2)$ . Dadas las probabilidades *a priori*  $\Pr(H_1)$  y  $\Pr(H_2) = 1 - \Pr(H_1)$ , entonces los datos generan probabilidades posteriores  $\Pr(H_1 | X)$  y  $\Pr(H_2 | X) = 1 - \Pr(H_1 | X)$ . Puesto que cualquier opinión *a priori* es transformada en una opinión *a posteriori* a través de los datos; la transformación en sí misma representa la evidencia aportada por los datos. En efecto, tal transformación es utilizada para obtener la probabilidad posterior, sin tomar en cuenta la probabilidad *a priori*. Cuando generamos la escala de momios (momios = probabilidad / (1 - probabilidad)), la transformación toma una forma sencilla. Del teorema de Bayes (ver capítulo 2), obtenemos

$$P(H_k | \mathbf{X}) = \frac{P(\mathbf{X} | H_k)P(H_k)}{P(\mathbf{X} | H_1)P(H_1) + P(\mathbf{X} | H_2)P(H_2)} \quad \text{con } k = 1, 2 \quad (3.1)$$

por lo que

$$\frac{\Pr(H_1 | \mathbf{X})}{\Pr(H_2 | \mathbf{X})} = \frac{P(\mathbf{X} | H_1) P(H_1)}{P(\mathbf{X} | H_2) P(H_2)}$$

y la transformación es simplemente la multiplicación por

$$B_{12} = \frac{P(\mathbf{X} | H_1)}{P(\mathbf{X} | H_2)} \quad (3.2)$$

que es precisamente lo que se conoce como el factor de Bayes. Así, en palabras,

$$\text{Momios posteriores} = \text{Factor de Bayes} \times \text{momios } a \text{ priori}$$

y en forma equivalente el factor de Bayes es la razón de los momios posteriores de  $H_1$  y sus momios *a priori*; sin importar el valor de los momios *a priori*. Notemos que en el enfoque Bayesiano es totalmente simétrico quién es  $H_1$  y quién  $H_2$ , es decir, es irrelevante en cuanto al procedimiento seguido. Por lo tanto, la comparación entre estas hipótesis es a nivel operativo y de presentación de resultados.

Intuitivamente,  $B_{12}$  proporciona una medida de si los datos tienen momios crecientes o decrecientes sobre  $H_1$  en relación a  $H_2$ . Así  $B_{12} > 1$  significa que  $H_1$  es relativamente más plausible a la luz de los datos;  $B_{12} < 1$  significa que la plausibilidad relativa de  $H_2$  es mayor. El factor de Bayes puede utilizarse directamente como criterio de selección de modelos (Rossi, 1989). Sin embargo, como lo veremos en la siguiente sección, nosotros utilizaremos un enfoque diferente.

Cuando las hipótesis son igualmente probables *a priori* —  $P(H_1) = P(H_2) = 0.5$  —, entonces el factor de Bayes es igual a los momios posteriores de  $H_1$ . Sin embargo, no siempre dos hipótesis son igualmente probables *a priori*.

En el caso más simple cuando dos hipótesis son distribuciones simples con parámetros fijos (el caso de hipótesis “simple versus simple”),  $B_{12}$  es la razón de verosimilitud. En otros casos, cuando hay parámetros desconocidos bajo una o ambas hipótesis, el factor de Bayes es todavía dado por (3.2), y en este sentido, éste continua teniendo la forma de una razón de verosimilitud. Sin embargo, las densidades  $P(\mathbf{X} | H_k)$  ( $k = 1, 2$ ) son obtenidas integrando (no maximizando) sobre el espacio paramétrico, así que en la ecuación (3.2),

$$P(\mathbf{X} | H_k) = \int_{\Theta} L(\mathbf{X} | \theta_k, H_k) \pi(\theta_k | H_k) d\theta_k \quad (3.3)$$

donde  $\theta_k$  es el parámetro bajo  $H_k$ ,  $\pi(\theta_k | H_k)$  es la densidad *a priori*, y  $L(\mathbf{X} | \theta_k, H_k)$  es la densidad de probabilidad de  $\mathbf{X}$  dado el valor de  $\theta_k$ , o la función de verosimilitud de  $\theta_k$ . En (Kass y Raftery, 1995) se puede consultar un análisis amplio sobre la interpretación del factor de Bayes y sus diferentes aplicaciones.

### 3.3 Probabilidad Posterior para un Modelo

Sean  $m$  modelos en competencia  $M_1, \dots, M_m$ , y  $\mathbf{X} = (x_1, \dots, x_n)$  una muestra aleatoria de tamaño  $n$  de los datos (la muestra puede ser completa o censurada), entonces por el teorema de Bayes la probabilidad posterior para el modelo  $M_j$ , está dada por:

$$\Pr(M_j | \mathbf{X}) = \frac{f(\mathbf{X} | M_j) \Pr(M_j)}{\sum_{i=1}^m f(\mathbf{X} | M_i) \Pr(M_i)} \quad j = 1, 2, \dots, m \quad (3.4)$$

donde

$$f(\mathbf{X} | M_j) = \int L(\mathbf{X} | \theta_j, M_j) \pi(\theta_j | M_j) d\theta_j. \quad (3.5)$$

Como ya habíamos visto en el capítulo 2, la función dada por (3.5) es la distribución conjunta marginal de los datos bajo el modelo  $M_j$ , y también se le conoce como la constante normalizadora o como la verosimilitud integrada (VI). De esta última forma

nos referiremos a ella en el resto del trabajo. Además  $\pi(\theta_j|M_j)$  es la densidad *a priori* para el vector de parámetros  $\theta_j$ ,  $L(\mathbf{X}|\theta_j, M_j)$  es la verosimilitud bajo el modelo  $M_j$ , y  $\Pr(M_j)$  es la probabilidad o ponderación *a priori* que se le asigna a la certidumbre del modelo  $M_j$ .

De acuerdo a la sección anterior, dedicada al factor de Bayes, la probabilidad posterior para un modelo dada por (3.4), se puede ver como una generalización de la fórmula dada en (3.1), y (3.5) sería equivalente a (3.3). De esta manera con (3.4) se puede evaluar la evidencia que aportan los datos en favor de una distribución de probabilidad como modelo de datos de confiabilidad al compararla contra otras distribuciones.

La problemática para calcular (3.4) se localiza en poder calcular (3.5), la cual depende de la verosimilitud y de la densidad *a priori*. Si (3.5) no tiene solución analítica, habrá que emplear métodos numéricos, lo cual tiene su problemática especial debido a que conforme  $n$  crece el integrando de (3.5) se pone "puntiagudo". Por ello a continuación describimos dos métodos para calcular la VI.

### 3.4 Métodos para Calcular la Verosimilitud Integrada

Como cometamos antes, uno de los problemas para obtener la probabilidad posterior para un modelo, es poder calcular la verosimilitud integrada. Sin embargo, como veremos más adelante, sólo en unos casos tiene solución analítica. Por ello habrá que recurrir a métodos numéricos para obtenerla. Un problema para obtener esta integral por métodos numéricos de integración convencionales es que para tamaños muestrales moderados y grandes el integrando se torna "puntiagudo" alrededor de su máximo, es decir, se tiene una función unimodal cuya masa tiende a concentrarse en su máximo. Pero además como generalmente esta masa es muy pequeña, la mayoría de métodos numéricos tienen problemas para encontrar la región donde la masa del integrando se acumula. En Kass y Raftery (1995) y en Hoeting et al. (1999) se discuten diferentes métodos para calcular la verosimilitud integrada. Nosotros veremos dos métodos: Laplace y Monte Carlo.

Laplace. Se supone que el integrando,  $L(\mathbf{X}|\theta_j, M_j)\pi(\theta_j|M_j)$  es muy puntiagudo alrededor de su máximo  $\tilde{\theta}$ , lo que es razonable a partir de tamaños muestrales relativamente bajos (ver Kass y Raftery, 1995; y Bernardo y Smith, 1994- pag 340). Sea

$$l_j(\theta) = \log [L(\mathbf{X}|\theta, M_j)\pi(\theta|M_j)] \quad (3.6)$$

y  $\tilde{\theta}$  el valor de  $\theta$  que maximiza  $l(\theta)$ .

Expandiendo  $l(\theta)$  hasta los términos cuadráticos alrededor de  $\tilde{\theta}$ , se obtiene una aproximación de  $L(\mathbf{X}|\theta_j, M_j)\pi(\theta_j|M_j)$  que tiene la forma de una densidad normal con media  $\tilde{\theta}$  y matriz de covarianza  $\Sigma = (-D^2l(\tilde{\theta}))^{-1}$ , donde  $D^2l_j(\tilde{\theta})$  es la matriz de segundas derivadas con respecto a  $\theta$ .

Integrando esta aproximación se obtiene que

$$\int_{\Theta} L(\mathbf{X}|\theta, M_j)\pi(\theta|M_j)d\theta \cong (2\pi)^{d/2} |\Sigma|^{-1/2} L(\mathbf{X}|\tilde{\theta}, M_j) \Pr(\tilde{\theta}|M_j) \quad (3.7)$$

donde  $d$  es la dimensión de  $\theta$ .

Así para aplicar esta aproximación se lleva a cabo lo siguiente:

- i) Calcular  $l_j(\theta)$  para cada modelo
- ii) Estimar numéricamente el vector posterior de modas  $(\tilde{\theta})$ ;
- iii) Obtener la matriz  $D^2 l_j(\tilde{\theta})$ .
- iv) Calcular la matriz  $\Sigma = (-D^2 l_j(\tilde{\theta}))^{-1}$
- v) Finalmente evaluar (3.7).

La aproximación del método de Laplace ha sido evaluada en varios trabajos, en Kass y Raftery (1995) se citan varios de ellos. En general hay consenso en que la aproximación es adecuada en aquellos casos en los que la dimensión de  $\theta$  es pequeña y cuando las funciones de verosimilitud no tienen una forma muy diferente a la normal. Por ejemplo, bajo las condiciones concretas descritas en Kass et al. (1990), se demuestra que conforme  $n \rightarrow \infty$ ,  $\int_{\Theta} L(\mathbf{X}|\theta, M_j)\pi(\theta|M_j)d\theta = \hat{I}(1 + O(n^{-1}))$ , donde  $\hat{I}(\cdot)$  es la aproximación de Laplace para la VI. Es decir, el error relativo es de  $O(n^{-1})$ . De esta manera, cuando el método de Laplace es aplicado al numerador y denominador, ya sea para calcular el factor de Bayes (sección anterior) o la probabilidad posterior para un modelo, la aproximación resultante tiene también un error del orden  $O(n^{-1})$  (Kass y Raftery, 1995).

Monte Carlo. Otro método para calcular en forma aproximada la verosimilitud integrada es a través de simulación Monte Carlo. Efectivamente observando la VI:

$$f(\mathbf{X}|M_j) = \int L(\mathbf{X}|\theta_j, M_j)\pi(\theta_j|M_j)d\theta_j$$

ésta es el valor esperado de  $L(\mathbf{X}|\theta_j, M_j)$  con respecto a  $\pi(\theta_j|M_j)$ ; por lo tanto

$$\hat{f}(\mathbf{X}|M_j) = \frac{1}{K} \sum_{i=1}^K L(\mathbf{X}|\theta_j^{(i)}, M_j)$$

donde  $\{\theta_j^{(i)} : i = 1, \dots, K\}$  es una muestra de la distribución *a priori*  $\pi(\theta_j|M_j)$ . La estimación entonces no es otra cosa que el promedio de las verosimilitudes de una muestra de valores del vector de parámetros  $\theta_j$ . La elección de  $K$  debe ser suficientemente grande para garantizar la convergencia (Robert y Casella, 1999).

De esta manera dada la verosimilitud para los diferentes modelos, lo único que se requerirá es generar muestras aleatorias de  $\pi(\theta_j|M_j)$ . Como veremos más adelante, usaremos dos tipos de distribuciones *a priori*: normal-gama y uniforme. Entonces en ambos casos es muy sencillo implementar un algoritmo en S-plus para generar muestras aleatorias de estas distribuciones. Estos algoritmos son el B1 y B2 del apéndice de cómputo, y es muy fácil justificarlos con los elementos que desarrollamos en el siguiente capítulo.

## CAPITULO 4

### VEROSIMILITUD INTEGRADA Y DENSIDAD POSTERIOR PREDICTIVA

En este capítulo se obtienen las expresiones para la verosimilitud integrada y la densidad posterior predictiva para las distribuciones normal ( $M_1$ ), lognormal ( $M_2$ ), valor extremo ( $M_3$ ), Weibull ( $M_4$ ) y exponencial ( $M_5$ ).

#### 4.1 Modelos Considerados

En esta sección estableceremos los modelos que consideraremos en este trabajo, junto con sus generalidades

Modelo normal. La densidad para una distribución normal  $(\mu, \tau)$ , está dada por

$$f_1(x | \mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left[-\frac{\tau}{2}(x - \mu)^2\right] \quad (4.1)$$

con  $-\infty < x < +\infty, -\infty < \mu < +\infty, \tau > 0$ . El parámetro de precisión  $\tau$  es igual al inverso de la varianza:  $\tau = 1/\sigma^2$ . De aquí que si se hace un muestreo sin censura la verosimilitud para el modelo normal está dada por

$$\begin{aligned} L(\mathbf{X} | \mu, \tau, M_1) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left[-\frac{\tau}{2} \sum_{i=1}^n (x_i - \mu)^2\right] \\ &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left[-\frac{\tau}{2} n(\mu - \bar{x}_n)^2\right] \exp\left[-\frac{\tau s_n}{2}\right] \end{aligned} \quad (4.2)$$

con  $\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$ ,  $s_n = \sum_{i=1}^n (x_i - \bar{x}_n)^2$ . La igualdad anterior se obtiene observando que  $\sum_{i=1}^n (x_i - \mu)^2 = n(\mu - \bar{x})^2 + \sum_{i=1}^n (x_i - \bar{x})^2$ .

Considerando las censuras tipo I, tipo II y aleatoria, la verosimilitud para el modelo normal está dada por

$$L(\mathbf{X} | \mu, \tau, M_1) = C \left(\frac{\tau}{2\pi}\right)^{\frac{n}{2}} \exp\left[-\frac{\tau}{2} \sum_{i=1}^r (x_{(i)} - \mu)^2\right] \prod_{j=1}^{n-r} (1 - \Phi(\tau^{1/2}(q_{(j)} - \mu))) \quad (4.3)$$

donde  $\Phi(\cdot)$  es la distribución acumulada normal estándar. Además en el caso de censura tipo I y aleatoria la constante  $C = 1$ . En el caso de censura tipo II,  $C = \frac{n!}{(n-r)!}$  y el tiempo de censura está dado por la  $r$ -ésima estadística de orden,  $q_{(j)} = x_{(r)} = T_e$ . En el caso de tener una censura simple de tipo I, entonces la verosimilitud anterior toma la forma siguiente

$$L(\mathbf{X} | \mu, \tau, M_1) = C \left(\frac{\tau}{2\pi}\right)^{\frac{n}{2}} \exp\left[-\frac{r\tau}{2}(\mu - \bar{x}_r)^2\right] \exp\left[-\frac{\tau s_r}{2}\right] (1 - \Phi(\tau^{1/2}(T_e - \mu)))^{n-r}$$

con  $\bar{x}_r$  igual a la media de los  $r$  tiempos de falla observados y  $s_r = \sum_{i=1}^r (x_{(i)} - \bar{x}_r)^2$ .

Se suele utilizar como distribuciones *a priori* para los parámetros del modelo normal  $(\mu, \tau)$ , una distribución normal-gama  $(m_1, k_1, \alpha_1, \beta_1)$ , en donde la distribución

condicional de  $\mu$  dado  $\tau = k_1$  es normal  $(m_1, k_1\tau)$  y la distribución marginal de  $\tau$  es una gama  $(\alpha_1, \beta_1)$ , por lo que

$$\pi(\mu, \tau | \alpha_1, \beta_1, m_1, k_1) = \left(\frac{k_1\tau}{2\pi}\right)^{1/2} \exp\left[-\frac{k_1\tau}{2}(\mu - m_1)^2\right] \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \exp(-\beta_1\tau)\tau^{\alpha_1-1} \quad (4.4)$$

con  $\alpha_1, \beta_1 > 0$ . Se puede demostrar (ver DeGroot, 1970, pag. 169) que la distribución marginal de  $\mu$  está dada por una distribución *t*-Student con  $2\alpha_1$  grados de libertad, parámetro de localización  $m_1$  y precisión  $\alpha_1 k_1 / \beta$ .

Si en lugar de utilizar una distribución normal-gama como *a priori* para los parámetros de la distribución normal  $(\mu, \tau)$ , utilizamos una distribución uniforme y suponiendo independencia entre éstos, entonces:

$$\pi(\mu, \tau | a_1, b_1, d_1, e_1) = \pi(\mu | a_1, b_1)\pi(\tau | d_1, e_1) = \frac{1}{(b_1 - a_1)} \frac{1}{(e_1 - d_1)} = u_1 \quad (4.5)$$

En el capítulo 5 discutimos la razón de utilizar dos diferentes tipos de distribuciones *a priori*.

Modelo Lognormal. La densidad de la distribución lognormal  $(\mu, \tau)$  está dada por

$$f_2(x | \mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{1/2} x^{-1} \exp\left[-\frac{\tau}{2}(\log(x) - \mu)^2\right] \quad (4.6)$$

con  $0 < x, -\infty < \mu < +\infty, \tau > 0$ . Los parámetros logmedia,  $\mu$ , y logprecisión,  $\tau$ , tienen diferente interpretación respecto a los parámetros equivalentes del modelo normal. Siguiendo un procedimiento al caso normal, se puede ver que la verosimilitud para datos completos para el modelo lognormal está dada por

$$L(\mathbf{X} | \mu, \tau, M_2) = \frac{1}{\prod_{i=1}^n x_i} \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left[-\frac{\tau}{2}n(\mu - \bar{w}_n)^2\right] \exp\left[-\frac{\tau v_n}{2}\right] \quad (4.7)$$

donde  $v_n = \sum_{i=1}^n (\log(x_i) - \bar{w})^2$  y  $\bar{w}_n = \frac{1}{n} \sum_{i=1}^n \log(x_i)$ .

En el caso de censura tipo I, II y aleatoria, la verosimilitud para el modelo lognormal está dada por

$$L(\mathbf{X} | \mu, \tau, M_2) = \frac{1}{\prod_{i=1}^r x_i} C \left(\frac{\tau}{2\pi}\right)^{\frac{r}{2}} e^{\left[-\frac{\tau}{2}(\mu - \bar{w}_r)^2 - \frac{\tau v_r}{2}\right]} \prod_{j=1}^{n-r} \left( \int_{\tau^{1/2}(\log(q_{(j)})-\mu)}^{\infty} (2\pi)^{-\frac{1}{2}} e^{-\frac{w^2}{2}} dw \right)$$

donde  $v_r = \sum_{i=1}^r (\log(x_{(i)}) - \bar{w}_r)^2$  y  $\bar{w}_r = \frac{1}{r} \sum_{i=1}^r \log(x_{(i)})$ . Además la constante  $C$  y los tiempos de censura  $q_{(j)}$ , dependen del tipo de censura (ver Capítulo 2).

Como distribución inicial para los parámetros del modelo lognormal vamos a usar las mismas distribuciones que en el caso normal, sólo que con valores diferentes para los correspondientes hiperparámetros. Así la distribución inicial para  $(\mu, \tau)$  podrá ser normal-gama  $(m_2, k_2, \alpha_2, \beta_2)$ , ver (4.4); o bien una uniforme  $(a_2, b_2, d_2, e_2)$ , como (4.5).

Distribución valor extremo. La función de densidad para la distribución de valores extremos está dada por

$$f_3(x | \mu, \tau) = \tau \exp [\tau(x - \mu) - \exp(\tau(x - \mu))] \quad (4.8)$$

con  $-\infty < x < \infty$ ,  $\tau > 0$ ,  $-\infty < \mu < \infty$ . El parámetro  $\mu$  es de localización y  $\tau$  es de escala. Considerando datos sin censura, la verosimilitud para este modelo está dada por

$$L(\mathbf{X} | M_3, \mu, \tau) = (\tau)^n \exp \left[ \tau \left( -n\mu + \sum_{i=1}^n x_i \right) - \left( \sum_{i=1}^n \exp(\tau(x_i - \mu)) \right) \right] \quad (4.9)$$

y la verosimilitud con muestras censuradas toma la forma siguiente

$$L(\mathbf{X} | M_3, \mu, \tau) = C(\tau)^r \exp \left[ \tau \left( -r\mu + \sum_{i=1}^r x_{(i)} \right) - \left( \sum_{i=1}^r \exp(\tau(x_{(i)} - \mu)) \right) \right] \\ \exp \left[ - \sum_{j=1}^{n-r} \exp(\tau(q_{(j)} - \mu)) \right] \quad (4.10)$$

Respecto a la distribución inicial para los parámetros del modelo valor extremo  $(\mu, \tau)$ , también utilizaremos una distribución normal-gama  $(\alpha_3, \beta_3, m_3, k_3)$ :

$$\pi(\mu, \tau | \alpha_3, \beta_3, m_3, k_3) = \left( \frac{k_3 \tau}{2\pi} \right)^{1/2} \exp \left[ -\frac{k_3 \tau}{2} (\mu - m_3)^2 \right] \frac{\beta_3^{\alpha_3}}{\Gamma(\alpha_3)} \exp(-\beta_3 \tau) \tau^{\alpha_3 - 1} \quad (4.11)$$

o una distribución uniforme  $(a_3, b_3, d_3, e_3)$ , como (4.5).

Distribución Weibull. La densidad para el modelo Weibull  $(\eta, \theta)$  está dada por

$$f_4(x | \eta, \theta) = \frac{\theta}{\eta} \left( \frac{x}{\eta} \right)^{\theta-1} \exp \left[ - \left( \frac{x}{\eta} \right)^\theta \right] \quad \text{con } x > 0, \eta > 0, \theta > 0 \quad (4.12)$$

De aquí que la verosimilitud para datos censurados para el modelo Weibull está dada por

$$L(\mathbf{X} | M_4, \eta, \theta) = C \left( \frac{\theta}{\eta^\theta} \right)^r \left( \prod_{i=1}^r x_{(i)} \right)^{\theta-1} \exp \left[ - \sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta - \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \right] \quad (4.13)$$

Nótese que para datos completos  $n = r$ , por lo que la última sumatoria en el exponente del último término de (4.13), desaparece.

Como distribuciones *a priori* para  $(\eta, \theta)$  utilizaremos las dos que hemos estado utilizando: distribución normal-gama  $(m_4, k_4, \alpha_4, \beta_4)$  :

$$p(\eta, \theta | m_4, k_4, \alpha_4, \beta_4) = \left( \frac{k_4 \theta}{2\pi} \right)^{1/2} \exp \left[ -\frac{k_4 \theta}{2} (\mu - m_4)^2 \right] \frac{\beta_4^{\alpha_4}}{\Gamma(\alpha_4)} \exp(-\beta_4 \theta) \theta^{\alpha_4 - 1}$$

y distribución uniforme  $(a_4, b_4, d_4, e_4)$ .

Distribución Exponencial. La densidad para el modelo exponencial está dada por

$$f_5(x | \theta, \gamma) = \theta \exp[-\theta(x - \gamma)] \quad \theta > 0, \quad 0 \leq \gamma < x \quad (4.14)$$

Vamos a suponer conocido  $\gamma$ . Si se hace un muestreo sin censura la verosimilitud para el modelo exponencial está dada por

$$L(\mathbf{X} | M_5, \theta, \gamma) = \theta^n \exp \left[ -\theta \left( \sum_{i=1}^n x_i - n\gamma \right) \right] \quad (4.15)$$

y bajo las censuras que hemos estado considerando, la verosimilitud está dada por

$$L(\mathbf{X} | \theta) = C\theta^r \exp \left[ -\theta \left( \sum_{i=1}^r x_{(i)} + \sum_{j=1}^{n-r} q_{(j)} - n\gamma \right) \right] \quad (4.16)$$

Como distribución *a priori* para  $\theta$  utilizaremos una distribución gama  $(\alpha_5, \beta_5)$  :

$$\pi(\theta | \alpha_5, \beta_5) = \frac{\beta_5^{\alpha_5}}{\Gamma(\alpha_5)} \exp(-\beta_5 \theta) \theta^{\alpha_5 - 1} \quad (4.17)$$

y una distribución uniforme  $(a_5, b_5)$  :

$$\pi(\theta | a_5, b_5) = \frac{1}{(b_5 - a_5)} = u_5. \quad (4.18)$$

## 4.2 Muestras Completas

En esta sección sólo se presenta la verosimilitud integrada y la densidad posterior predictiva para los modelos que con muestra completas tienen solución analítica. En el caso de utilizar un distribución *a priori* normal-gama, son los modelos normal, lognormal y exponencial; y en con *a priori* uniforme sólo este último. Para detalles de la notación referirse a la sección anterior.

**Teorema 4.1** *Bajo muestras aleatorias sin censura y suponiendo una densidad a priori normal-gama, la verosimilitud integrada está dada por:*

a) Para el modelo normal  $M_1$  :

$$f(\mathbf{X} | M_1, \alpha_1, \beta_1, m_1, k_1) = \frac{\beta_1^{\alpha_1} \Gamma(\frac{n}{2} + \alpha_1) (k_1)^{1/2}}{(k_1 + n)^{1/2} \Gamma(\alpha_1) (2\pi)^{n/2} h_1^{\frac{n}{2} + \alpha_1}} \quad (4.19)$$

con  $h_1 = \frac{1}{2}k_1n \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} + \beta_1 + \frac{s_n}{2}$ ,  $s_n = \sum_{i=1}^n (x_i - \bar{x})^2$  y  $\bar{x}_n$  es la media aritmética.  
 b) En el caso del modelo lognormal  $M_2$ :

$$f(\mathbf{X}|M_2, \alpha_2, \beta_2, m_2, k_2) = \frac{\beta_2^{\alpha_2} \Gamma(\frac{n}{2} + \alpha_2) (k_2)^{1/2}}{(k_2 + n)^{1/2} \Gamma(\alpha_2) (2\pi)^{n/2} h_2^{\frac{n}{2} + \alpha_2} \prod_{i=1}^n x_i} \quad (4.20)$$

con  $h_2 = \frac{1}{2}k_1n \frac{(m_2 - \bar{w}_n)^2}{k_1 + n} + \beta_2 + \frac{u_n}{2}$ .

**Teorema 4.2** Bajo muestreo aleatorio completo de un modelo exponencial  $(\theta, \gamma)$ , y suponiendo conocido  $\gamma$ ; la verosimilitud integrada está dada por:

a) Utilizando una densidad a priori gama  $(\alpha_5, \beta_5)$  para  $\theta$  (4.17), entonces:

$$f(\mathbf{X}|M_5, \gamma) = \frac{\beta_5^{\alpha_5} \Gamma(n + \alpha_5)}{\Gamma(\alpha_5) (s_5)^{n + \alpha_5}} \quad (4.21)$$

donde  $s_5 = \sum_{i=1}^n x_i + \beta_5 - n\gamma$ .

b) Suponiendo una distribución a priori uniforme  $(a_5, b_5)$  para  $\theta$  (4.18):

$$f(\mathbf{X} | M_5, \gamma) = \frac{u_5 \Gamma(n + 1)}{v^{n+1}} [G(b_5; n + 1, v) - G(a_5; n + 1, v)] \quad (4.22)$$

donde  $v = \sum_{i=1}^n x_i - n\gamma$ , y  $G(\cdot; n + 1, v)$  es la distribución gama acumulada con parámetros  $(n + 1, v)$ .

**Teorema 4.3** Bajo muestreo aleatorio completo y suponiendo una densidad a priori normal-gama  $(\alpha_1, \beta_1, m_1, k_1)$ , la densidad posterior predictiva (2.5) está dada por:

a) Para el modelo normal  $M_1$ :

$$f_1(y|\mathbf{X}) = \frac{\Gamma(\frac{2\alpha_1 + n + 1}{2})}{\Gamma(\frac{2\alpha_1 + n}{2}) \Gamma(\frac{1}{2})} \left( \frac{l_t}{2\alpha_1 + n} \right)^{1/2} \left[ \frac{l_t}{(2\alpha_1 + n)} (y - m_t)^2 + 1 \right]^{-\frac{(n + 2\alpha_1 + 1)}{2}} \quad (4.23)$$

donde  $m_t = \frac{k_1 m_1 + n \bar{x}_n}{n + k_1}$  y  $l_t = \left[ \frac{1}{2} n k_1 \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} + \beta_1 + \frac{s_n}{2} \right]^{-1} \frac{(\alpha_1 + n/2)(n + k_1)}{(k_1 + n + 1)}$ . Esta densidad es del tipo  $t$ -Student (Bernardo y Smith, 1994, pag. 122), con parámetros: localización (media)  $m_t$ , precisión  $l_t$ , y  $2\alpha_1 + n$  grados de libertad.

b) Para el modelo lognormal  $M_2$ :

$$f_2(y|\mathbf{X}) = \frac{\Gamma(\frac{2\alpha_2 + n + 1}{2})}{y \Gamma(\frac{2\alpha_2 + n}{2}) \Gamma(\frac{1}{2})} \left( \frac{l}{2\alpha_2 + n} \right)^{1/2} \left[ \frac{l}{(2\alpha_2 + n)} (\log(y) - m_l)^2 + 1 \right]^{-\frac{(n + 2\alpha_2 + 1)}{2}} \quad (4.24)$$

donde  $m_l = \frac{k_2 m_2 + n \bar{w}_n}{n + k_2}$  y  $l = \left[ \frac{1}{2} n k_2 \frac{(m_2 - \bar{w}_n)^2}{k_2 + n} + \beta_2 + \frac{u_n}{2} \right]^{-1} \frac{(\alpha_2 + n/2)(n + k_2)}{(k_2 + n + 1)}$ . Esta densidad es del tipo log $t$ -Student (Percy, 2002), con los siguientes parámetros: localización (media)  $m_l$ , precisión  $l$ , y  $2\alpha_2 + n$  grados de libertad.

**Teorema 4.4** Bajo muestreo aleatorio la densidad posterior predictiva (2.5) para el modelo exponencial ( $M_5$ ), está dada por

a) Utilizando una *a priori* Gama( $\alpha_5, \beta_5$ ) para  $\theta$  (4.17):

$$f_5(y|\mathbf{X}) = \frac{(v_5)^{n+\alpha_5} \Gamma(n + \alpha_5 + 1)}{\Gamma(n + \alpha_5) (v_5 + y - \gamma)^{n+\alpha_5+1}} \quad (4.25)$$

donde  $v_5 = \sum x_i + \beta_5 - n\gamma$ . Esta es una densidad gama-gama( $n + \alpha_5, v_5, 1$ ) (Ver Bernardo y Smith, 1994, pag. 120). La distribución posterior acumulada está dada por:

$$\begin{aligned} F_5(y|\mathbf{X}) &= \int_0^y f_5(t|\mathbf{X}) dt \\ &= \left( 1 - \frac{(v_5 - \gamma)^{n+\alpha_5}}{(v_5 + y - \gamma)^{n+\alpha_5}} \right) \end{aligned} \quad (4.26)$$

b) Utilizando una *a priori* uniforme( $a_5, b_5$ ) (4.18):

$$f_5(y|\mathbf{X}) = \frac{u_5 \Gamma(n+2)}{(t_5 + y)^{n+2}} [G(b_5; n+2, t_5 + y) - G(a_5; n+2, t_5 + y)] \quad (4.27)$$

donde  $G(\cdot; n+2, t_5 + y)$  es la distribución gama acumulada con parámetros ( $n+2, t_5 + y$ ); con  $t_5 = \sum_{i=1}^n x_i - (n+1)\gamma$ .

### 4.3 Muestras Censuradas

En esta sección presentamos resultados necesario para realizar aproximaciones por el método de Laplace o para la inferencia de tiempo de garantía.

**Teorema 4.5** Bajo muestreo censurado (I, II o aleatorio) y suponiendo una densidad *a priori* normal-gama, el hessiano de la fórmula (3.6) está dada por:

a) Para el modelo normal:

$$\begin{aligned} &\frac{d^2}{d\mu^2} l_1(\mu, \tau) \\ &= -(k_1 + r)\tau - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau(\mu - q_{(j)})^2\right) \tau}{(\sqrt{\pi})^3 (\Omega(\mu, \tau, q_{(j)}))^2} \right. \\ &\quad \left. \left( \Omega(\mu, \tau, q_{(j)}) \left[ \pi\sqrt{2\tau}(\mu - q_{(j)}) \right] + 2\sqrt{\pi} \exp\left(-\frac{\tau}{2}(\mu - q_{(j)})^2\right) \right) \right\} \end{aligned} \quad (4.28)$$

donde

$$\Omega(\mu, \tau, q_{(j)}) = 1 + \operatorname{erf}\left(\sqrt{\frac{\tau}{2}}(\mu - q_{(j)})\right) \quad (4.29)$$

con

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (4.30)$$

$$\begin{aligned} & \frac{d^2}{d\tau^2} l_1(\mu, \tau) \\ = & \frac{1}{2} \frac{r-1+2\alpha_1}{\tau^2} - \sum_{j=1}^{n-r} \left\{ \frac{(\mu - q_{(j)}) \exp\left(-\frac{\tau}{2} (\mu - q_{(j)})^2\right)}{4\pi\tau^2 (\Omega(\mu, \tau, q_{(j)}))^2} \right. \\ & \left. \left[ \sqrt{2\pi\tau} \left( \tau (\mu - q_{(j)})^2 + 1 \right) \Omega(\mu, \tau, q_{(j)}) + 2\tau(\mu - q_{(j)}) \exp\left(-\frac{\tau}{2} (\mu - q_{(j)})^2\right) \right] \right\} \end{aligned} \quad (4.31)$$

$$\begin{aligned} & \frac{d^2}{d\tau d\mu} l_1(\mu, \tau) \\ = & -k_1(\mu - m_1) - r(\mu - \bar{x}_r) - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau (\mu - q_{(j)})^2\right)}{(\Omega(\mu, \tau, q_{(j)}))^2 \pi\sqrt{2\tau}} \right. \\ & \left. \left[ \sqrt{\pi} \left( \tau (\mu - q_{(j)})^2 - 1 \right) \Omega(\mu, \tau, q_{(j)}) + (\mu - q_{(j)}) \sqrt{2\tau} \exp\left(-\frac{1}{2}\tau (\mu - q_{(j)})^2\right) \right] \right\} \end{aligned} \quad (4.32)$$

b) En el caso del modelo Lognormal:

$$\begin{aligned} & \frac{d^2}{d\mu^2} l_2(\mu, \tau) \\ = & -(k_2 + r)\tau - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau (\mu - \log(q_{(j)}))^2\right) \tau}{(\sqrt{\pi})^3 (\Omega(\mu, \tau, \log[q_{(j)}]))^2} \right. \\ & \left. \left( \Omega(\mu, \tau, \log[q_{(j)}]) \left[ \pi\sqrt{2\tau}(\mu - \log(q_{(j)})) \right] + 2\sqrt{\pi} \exp\left(-\frac{\tau}{2} (\mu - \log(q_{(j)}))^2\right) \right) \right\} \end{aligned} \quad (4.33)$$

$$\begin{aligned} & \frac{d^2}{d\tau^2} l_2(\mu, \tau) \\ = & \frac{1}{2} \frac{r-1+2\alpha_2}{\tau^2} - \sum_{j=1}^{n-r} \left\{ \frac{(\mu - \log(q_{(j)})) \exp\left(-\frac{\tau}{2} (\mu - \log(q_{(j)}))^2\right)}{4\pi\tau^2 (\Omega(\mu, \tau, \log[q_{(j)}]))^2} \right. \\ & \left. \left[ \sqrt{2\pi\tau} \left( \tau (\mu - \log(q_{(j)}))^2 + 1 \right) \Omega(\mu, \tau, \log[q_{(j)}]) + 2\tau(\mu - \log(q_{(j)})) e^{-\frac{\tau}{2} (\mu - \log(q_{(j)}))^2} \right] \right\} \end{aligned} \quad (4.34)$$

$$\begin{aligned} & \frac{d^2}{d\tau d\mu} l_2(\mu, \tau) \\ &= -k_2(\mu - m_2) - r(\mu - \bar{w}_r) - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau(\mu - \log(q_{(j)}))^2\right)}{(\Omega(\mu, \tau, \log[q_{(j)}]))^2 \pi \sqrt{2\tau}} \right. \\ & \quad \left. \left[ \sqrt{\pi} \left( \tau(\mu - \log(q_{(j)}))^2 - 1 \right) \Omega(\mu, \tau, \log[q_{(j)}]) + (\mu - \log(q_{(j)})) \sqrt{2\tau} e^{-\frac{1}{2}\tau(\mu - \log(q_{(j)}))^2} \right] \right\} \end{aligned} \quad (4.35)$$

donde la función  $\Omega(\mu, \tau, \log[q_{(j)}])$  ya se había definido en (4.29).

c) Para el modelo valor extremo:

$$\frac{d^2 l_3(\mu, \tau)}{d^2 \mu} = -\tau^2 \sum_{i=1}^r \exp(\tau(x_{(i)} - \mu)) - k_3 \tau - \tau^2 \sum_{j=1}^{n-r} \exp(\tau[q_{(j)} - \mu]) \quad (4.36)$$

$$\frac{d}{d\tau^2} l_3(\mu, \tau) = -\frac{(r + \alpha_3 - 0.5)}{\tau^2} - \sum_{i=1}^r (x_{(i)} - \mu)^2 e^{\tau(x_{(i)} - \mu)} - \sum_{j=1}^{n-r} (q_{(j)} - \mu)^2 e^{\tau[q_{(j)} - \mu]} \quad (4.37)$$

$$\begin{aligned} \frac{d^2 l_3(\mu, \tau)}{d\mu d\tau} &= -r + \tau \sum_{i=1}^r (x_{(i)} - \mu) \exp(\tau(x_{(i)} - \mu)) + \sum_{i=1}^r \exp(\tau(x_{(i)} - \mu)) - k_3 \mu \\ & \quad + \tau \sum_{j=1}^{n-r} [q_{(j)} - \mu] \exp(\tau[q_{(j)} - \mu]) + \sum_{j=1}^{n-r} \exp(\tau[q_{(j)} - \mu]) \end{aligned} \quad (4.38)$$

d) En el caso del modelo Weibull:

$$\frac{\partial^2 l_4}{\partial \eta^2} = r \frac{\theta}{\eta^2} - \left( \sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta + \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \right) \frac{\theta(\theta + 1)}{\eta^2} - k_4 \theta \quad (4.39)$$

$$\frac{\partial^2 l_4}{\partial \theta^2} = -\frac{\alpha_4 + r - 0.5}{\theta^2} - \sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta \ln^2 \frac{x_{(i)}}{\eta} - \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \left( \ln^2 \frac{q_{(j)}}{\eta} \right) \quad (4.40)$$

$$\frac{\partial^2 l_4}{\partial \theta \partial \eta} = -\frac{r}{\eta} + \frac{\sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta \left( \theta \ln \frac{x_{(i)}}{\eta} + 1 \right) + \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \left( \theta \ln \frac{q_{(j)}}{\eta} + 1 \right)}{\eta} - k_4 (\eta - m_4) \quad (4.41)$$

**Teorema 4.6** *Bajo un muestreo con censura del modelo Exponencial( $\theta, \gamma$ ), suponiendo conocido  $\gamma$ , entonces la verosimilitud integrada está dada por:*

a) Utilizando como distribución *a priori* para  $\theta$  una distribución gama  $(\alpha_5, \beta_5)$ :

$$f(\mathbf{X}|M_5, \gamma) = \frac{C\beta_5^{\alpha_5}\Gamma(r + \alpha_5)}{\Gamma(\alpha_5) (h)^{r+\alpha_5}} \quad (4.42)$$

con  $h = \sum_{i=1}^r x_{(i)} + \sum_{j=1}^{n-r} q_{(j)} + \beta_5 - n\gamma$ , y  $C$  depende del tipo de censura (ver cap.2)

b) Utilizando una distribución uniforme  $(a_5, b_5)$  como *a priori* para  $\theta$ :

$$f(\mathbf{X}|M_5, \gamma) = \frac{u_5 C \Gamma(r+1)}{(h)^{r+1}} [G(a_{25}; r+1, h) - G(a_{15}; r+1, h)] \quad (4.43)$$

donde  $h = \sum_{i=1}^r x_{(i)} + \sum_{j=1}^{n-r} q_{(j)} - n\gamma$ ,  $G(\cdot; r+1, h)$  es la distribución gama acumulada con parámetros  $(r+1, h)$  y  $u_5$  como en (4.18).

**Teorema 4.7** *Bajo muestreo aleatorio censurado la densidad posterior predictiva (2.5) para el modelo exponencial  $(M_5)$ , está dada por*

a) Utilizando una *a priori* Gama  $(\alpha_5, \beta_5)$  para  $\theta$  (4.17) :

$$f_5(y|\mathbf{X}) = \frac{(p_5)^{r+\alpha_5}\Gamma(r + \alpha_5 + 1)}{\Gamma(r + \alpha_5)(p_5 + y - \gamma)^{r+\alpha_5+1}} \quad (4.44)$$

donde  $p_5 = \sum_{i=1}^r x_{(i)} + (n-r)T_e + \beta_5 - n\gamma$ . Esta es una densidad gama-gama  $(r + \alpha_5, p_5, 1)$  (Ver Bernardo y Smith, 1994, pag. 120).

b) Utilizando una *a priori* uniforme  $(a_5, b_5)$ :

$$f_5(y|\mathbf{X}) = \frac{u_5 \Gamma(r+2)}{(s_5 + y)^{n+2}} [G(b_5; r+2, s_5 + y) - G(a_5; r+2, s_5 + y)] \quad (4.45)$$

con  $s_5 = \sum_{i=1}^r x_{(i)} + (n-r)T_e - (n+1)\gamma$  y  $G(\cdot; r+2, s_5 + y)$  es la distribución gama acumulada con parámetros  $(r+2, s_5 + y)$ .

**Teorema 4.8** *Bajo muestreo censurado (I, II o aleatorio) y suponiendo una densidad *a priori* uniforme, el hessiano de la fórmula (3.6) está dada por:*

a) Para el modelo normal:

$$\begin{aligned} & \frac{d^2}{d\mu^2} l_1(\mu, \tau) \\ &= -r\tau - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau(\mu - q_{(j)})^2\right)\tau}{(\sqrt{\pi})^3 (\Omega(\mu, \tau, q_{(j)}))^2} \right. \\ & \quad \left. \left( \Omega(\mu, \tau, q_{(j)}) \left[ \pi\sqrt{2\tau}(\mu - q_{(j)}) \right] + 2\sqrt{\pi} \exp\left(-\frac{\tau}{2}(\mu - q_{(j)})^2\right) \right) \right\} \quad (4.46) \end{aligned}$$

$$\begin{aligned}
& \frac{d^2}{d\tau^2} l_1(\mu, \tau) \\
&= -\frac{r}{2\tau^2} - \sum_{j=1}^{n-r} \left\{ \frac{(\mu - q_{(j)}) \exp\left(-\frac{\tau}{2}(\mu - q_{(j)})^2\right)}{4\pi\tau^2 (\Omega(\mu, \tau, q_{(j)}))^2} \right. \\
& \quad \left. \left[ \sqrt{2\pi\tau} \left( \tau (\mu - q_{(j)})^2 + 1 \right) \Omega(\mu, \tau, q_{(j)}) + 2\tau(\mu - q_{(j)}) \exp\left(-\frac{\tau}{2}(\mu - q_{(j)})^2\right) \right] \right\} \quad (4.47)
\end{aligned}$$

$$\begin{aligned}
& \frac{d^2}{d\tau d\mu} l_1(\mu, \tau) \\
&= -r(\mu - \bar{x}_r) - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau(\mu - q_{(j)})^2\right)}{(\Omega(\mu, \tau, q_{(j)}))^2 \pi\sqrt{2\tau}} \right. \\
& \quad \left. \left[ \sqrt{\pi} \left( \tau (\mu - q_{(j)})^2 - 1 \right) \Omega(\mu, \tau, q_{(j)}) + (\mu - q_{(j)}) \sqrt{2\tau} \exp\left(-\frac{1}{2}\tau(\mu - q_{(j)})^2\right) \right] \right\} \quad (4.48)
\end{aligned}$$

b) Para el modelo lognormal

$$\begin{aligned}
& \frac{d^2}{d\mu^2} l_2(\mu, \tau) \\
&= -r\tau - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau(\mu - \log(q_{(j)}))^2\right) \tau}{(\sqrt{\pi})^3 (1 + \operatorname{erf}\left(\sqrt{\frac{\tau}{2}}(\mu - \log(q_{(j)}))\right))^2} \right. \\
& \quad \left. \left( \Omega(\mu, \tau, \log[q_{(j)}]) \left[ \pi\sqrt{2\tau}(\mu - \log(q_{(j)})) \right] + 2\sqrt{\pi} \exp\left(-\frac{\tau}{2}(\mu - \log(q_{(j)}))^2\right) \right) \right\} \quad (4.49)
\end{aligned}$$

$$\begin{aligned}
& \frac{d^2}{d\tau^2} l_2(\mu, \tau) \\
&= -\frac{1}{2} \frac{r}{\tau^2} - \sum_{j=1}^{n-r} \left\{ \frac{(n-r)(\mu - \log(q_{(j)})) \exp\left(-\frac{\tau}{2}(\mu - \log(q_{(j)}))^2\right)}{4\pi\tau^2 (\Omega(\mu, \tau, \log[q_{(j)}]))^2} \right. \\
& \quad \left. \left[ \sqrt{2\pi\tau} \left( \tau (\mu - \log(q_{(j)}))^2 + 1 \right) \Omega(\mu, \tau, \log[q_{(j)}]) + 2\tau(\mu - \log(q_{(j)})) e^{\left(-\frac{\tau}{2}(\mu - \log(q_{(j)}))^2\right)} \right] \right\} \quad (4.50)
\end{aligned}$$

$$\begin{aligned}
& \frac{d^2}{d\tau d\mu} l_2(\mu, \tau) \\
&= -r(\mu - \bar{w}_r) - \sum_{j=1}^{n-r} \left\{ \frac{\exp\left(-\frac{1}{2}\tau(\mu - \log(q_{(j)}))^2\right)}{(1 + \operatorname{erf}\left(\sqrt{\frac{\tau}{2}}(\mu - \log(q_{(j)}))\right))^2 \pi\sqrt{2\tau}} \right. \\
& \quad \left. \left[ \sqrt{\pi} \left( \tau (\mu - \log(q_{(j)}))^2 - 1 \right) \Omega(\mu, \tau, \log[q_{(j)}]) + (\mu - \log(q_{(j)})) \sqrt{2\tau} e^{\left(-\frac{1}{2}\tau(\mu - \log(q_{(j)}))^2\right)} \right] \right\} \quad (4.51)
\end{aligned}$$

c) En el caso del modelo valor extremo

$$\frac{d^2 l_3(\mu, \tau)}{d^2 \mu} = -\tau^2 \sum_{i=1}^r \exp(\tau (x_{(i)} - \mu)) - \tau^2 \sum_{j=1}^{n-r} e^{\tau [q_{(j)} - \mu]} \quad (4.52)$$

$$\frac{d^2 l_3(\mu, \tau)}{d^2 \tau} = -\frac{r}{\tau^2} - \sum_{i=1}^r (x_{(i)} - \mu)^2 \exp(\tau (x_{(i)} - \mu)) - \sum_{j=1}^{n-r} [q_{(j)} - \mu]^2 \exp(\tau [q_{(j)} - \mu]) \quad (4.53)$$

$$\begin{aligned} \frac{d^2 l_3(\mu, \tau)}{d\mu d\tau} &= -r + \tau \sum_{i=1}^r (x_{(i)} - \mu) \exp(\tau (x_{(i)} - \mu)) + \sum_{i=1}^r \exp(\tau (x_{(i)} - \mu)) \\ &+ \tau \sum_{j=1}^{n-r} [q_{(j)} - \mu] \exp(\tau [q_{(j)} - \mu]) + \sum_{j=1}^{n-r} \exp(\tau [q_{(j)} - \mu]) \end{aligned} \quad (4.54)$$

d) Para el modelo Weibull

$$\frac{\partial^2 l_4}{\partial \eta^2} = r \frac{\theta}{\eta^2} - \left( \sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta + \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \right) \frac{\theta(\theta + 1)}{\eta^2} \quad (4.55)$$

$$\frac{\partial^2 l_4}{\partial \theta^2} = -\frac{r}{\theta^2} - \sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta \ln^2 \frac{x_{(i)}}{\eta} - \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \left( \ln^2 \frac{q_{(j)}}{\eta} \right) \quad (4.56)$$

$$\frac{\partial^2 l_4}{\partial \theta \partial \eta} = -\frac{r}{\eta} + \frac{\sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta \left( \theta \ln \frac{x_{(i)}}{\eta} + 1 \right) + \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \left( \theta \ln \frac{q_{(j)}}{\eta} + 1 \right)}{\eta} \quad (4.57)$$

## CAPITULO 5

### ESPECIFICACIÓN DE LA DISTRIBUCIÓN *A PRIORI* PARA MODELOS EN CONFIABILIDAD

En este capítulo proponemos cómo especificar los parámetros de las distribuciones *a priori* para los parámetros de cada uno de los modelos de confiabilidad que hemos considerado, para de esa manera completar el procedimiento de especificación de modelos y en general poder hacer inferencia Bayesiana con los modelos de confiabilidad que describimos en el capítulo 4.

#### 5.1 Introducción

En la sección 3.1 señalamos que una de las dificultades principales para aplicar la metodología de selección de modelos descrita en el capítulo 3 o para calcular el factor de Bayes, es la problemática de definir la distribución *a priori*  $\pi(\theta)$  para los parámetros de los modelos. Ya que de entrada, es muy difícil que el usuario de la estadística posea el conocimiento suficiente para que pueda establecer un rango de variación, por ejemplo, para cada uno de los parámetros de los modelos considerados. Esto en particular para el caso de los modelos que hemos estado considerando en los capítulos previos, equivaldría a que el usuario tendría que poseer una **información similar** para todos los parámetros y especificar por ejemplo rangos de variación equivalentes lo mismo para la media en el caso normal, que para la logmedia de la distribución lognormal o el parámetro de forma del modelo Weibull. Algo que francamente parece poco realista. Esta problemática ha llevado a diferentes enfoques Bayesianos de solución por ejemplo muestras de ensayo, donde parte de la información de la muestra es utilizada para estimar la distribución inicial, y con el resto tomar la decisión (Berger y Pericchi, 1996; Pérez y Berger, 2002)), suponer que se dispone de datos históricos previos, similares a los que se quieren obtener, con base en los cuales se estiman los parámetros de la distribución *a priori* (Ibrahim, et al. 2001, sec. 3.3 y 6.1), etcétera. También varias aproximaciones desde la perspectiva Bayesiana semi-paramétrica o no paramétrica (ver por ejemplo, Sinha et. al 1999; Gutiérrez-Peña y Walker, 2001; y referencias citadas en estos trabajos).

En selección de modelo no es posible utilizar distribuciones *a priori* impropias (Bernardo y Smith, 1994, secc. 5.4), ya que al final el factor de Bayes o la probabilidad posterior para un modelo contienen constantes indefinidas (Kass y Raftery, 1995; Berger y Pericchi, 1996). Para superar esta problemática de las constantes indefinidas, existen algunos trabajos, ya sea para estimarlas por simulación (Chen y Shao, 1997), o para en el contexto de prueba de hipótesis utilizar "muestras imaginarias de ensayo" (Spiegelhalter y Smith, 1982). Donde se parte de imaginar que se dispone de un pequeño conjunto de datos, a través de los cuales se obtiene un valor para la razón de las constantes indefinidas. Como evolución de este enfoque, y en parte para superar sus limitaciones, se ha propuesto usar parte de la información para definir el valor de estas constantes (Berger y Pericchi, 1996).

Lo anterior muestra que el problema de especificar la distribución *a priori* en el contexto de especificación de modelo es un problema abierto. Esto en realidad es parte de las propuestas que se siguen generando para mejorar y enriquecer el análisis

Bayesiano. Por ejemplo en Wilson (1994), Wolfson (1995) y Kadane y Wolfson (1998) se discuten diferentes aspectos cognitivos y psicológicos del proceso mediante el cual el conocimiento del experto es transformado en afirmaciones probabilísticas. Es también reconocido que a pesar de la importancia del problema de especificar la distribución *a priori*, éste ha recibido poca atención en la literatura Bayesiana (O'Hagan, 1998).

En las siguientes secciones proponemos un procedimiento para definir las distribuciones iniciales para los parámetros de los diferentes modelos de confiabilidad que aquí hemos estudiado, que puede ser ampliado hacia otros modelos no analizados en este trabajo. Consideramos que este procedimiento es una buena alternativa para la especificación de modelos en confiabilidad y para tomar decisiones con base en ellos, como es el caso de decidir el tiempo de garantía.

Para nuestra propuesta hemos partido de una especie de consenso entre los expertos Bayesianos, acerca de que la especificación de la *a priori* debe partir de información disponible sobre las observables, es decir, sobre el tiempo de vida  $X$ , más que de los no observables -los parámetros de los modelos de tiempo de vida-, ver por ejemplo Winkler (1981), Singpurwalla (1988), Wolfson (1995), Kadane y Wolfson (1998), O'Hagan(1998), Garthwaite y O'Hagan(2000), Percy(2002). Antes de describir nuestra propuesta, vamos a señalar algunos elementos de las contribuciones que se han citado antes.

En Kadane y Wolfson (1998) se propone un procedimiento para la especificación de los hiperparámetros para el modelo lineal normal, donde inicialmente el usuario especifica un rango de variación para cada uno de los  $p$  variables del modelo,  $x_j$ . A partir de esto se genera mediante un algoritmo realizaciones posibles del vector de variables,  $\mathbf{x}$ . De aquí se generan por lo menos  $p + 3$  realizaciones de  $\mathbf{x}$ , y por cada realización el usuario debe especificar los percentiles 50, 75 y 90 de la distribución posterior predictiva. Este enfoque supone demasiado conocimiento inicial del usuario.

En O'Hagan(1998) y Garthwaite y O'Hagan(2000) se propone un método para elegir la media y la desviación estándar iniciales para un modelo de pronóstico. En estos trabajos, cuando la distribución *a priori* predictiva  $f(x)$  se supone que es simétrica y unimodal, se propone que la media *a priori* sea dada por el experto en forma directa dando un número  $M$ , y para la desviación estándar *a priori*, inspirados en las propiedades de la distribución normal, se obtenga a partir de que el experto dé un par de cuantiles de la distribución *a priori* predictiva. Sean  $t_{p_1}$  y  $t_{p_2}$  estos cuantiles ( $t_{p_1} < t_{p_2}$ ). Se sabe que la función cuantil de la distribución normal para  $\mu$  y  $\sigma$  dados, está definida por  $t_p = \mu + \phi^{-1}(p)\sigma$ , donde  $\phi^{-1}(p) = z_p$  es el cuantil  $p$  de la distribución normal estándar. De aquí que es razonable como regla práctica proponer el valor de  $\sigma$  inicial a partir de la diferencia entre estos cuantiles, ya que:  $t_{p_2} - t_{p_1} = \mu + \phi^{-1}(p_2)\sigma - \mu - \phi^{-1}(p_1)\sigma$ ; de aquí que

$$\sigma = \frac{t_{p_2} - t_{p_1}}{z_p} \quad (5.1)$$

donde  $z_p = \phi^{-1}(p_2) - \phi^{-1}(p_1)$  es el rango intercuantílico de la distribución normal estándar. En O'Hagan(1998) se sugiere que se elijan los cuantiles de tal forma que el rango intercuantílico cubra el 33.3%, 50% ó 66% de los valores de  $x$ ; lo que se puede obtener, respectivamente, con la siguientes elección de cuantiles: 0.33 – 0.67,

0.25 - 0.75, 0.17 - 0.83; en estos casos  $z_p$  tomaría los valores 0.8614, 1.3490 y 1.9348, respectivamente. La decisión de que pareja de cuantiles elegir, se toma con base en el conocimiento *a priori* que tenga el experto de  $f(x)$ .

Por otro lado si desde el inicio se cree que  $f(x)$  tiene un sesgo positivo, como es frecuente en distribuciones del tiempo de vida, entonces en O'Hagan(1998) se propone que la elección *a priori* de la media y la desviación estándar se haga aplicando la misma regla de "dedo" anterior, pero ahora inspirados en una distribución con sesgo positivo como lo es la distribución lognormal( $\mu_l, \sigma_l$ ). Para ello recordemos que la función cuantil de esta distribución, para  $\mu_l$  y  $\sigma_l$  dados, está definida por  $t_p = \exp[\mu_l + \phi^{-1}(p)\sigma_l]$ . Además la media y desviación estándar para una variable con distribución lognormal están dadas por  $E(X | \mu_l, \sigma_l) = \exp(\mu_l + 0.5\sigma_l^2)$  y  $\sqrt{V(X | \mu_l, \sigma_l)} = \{\exp(2\mu_l + \sigma_l^2)[\exp(\sigma_l^2) - 1]\}^{1/2}$ . Así si el usuario da *a priori* dos cuantiles de  $f(x)$ ,  $t_{p_1}$  y  $t_{p_2}$ ; entonces  $\log(t_{p_2}) - \log(t_{p_1}) = \sigma_l z_p$ , y  $\sigma_l = [\log(t_{p_2}) - \log(t_{p_1})] / z_p$ . Definiendo

$$w = \exp(\sigma_l^2/2) \equiv \exp\left\{\frac{[\log(t_{p_2}) - \log(t_{p_1})]^2}{2z_p^2}\right\} \quad (5.2)$$

y suponiendo que el experto también da *a priori* el valor de la mediana de  $f(x)$ ,  $D$ , y como de acuerdo a la función cuantil  $D = \exp(\mu_l)$ , entonces los valores *a priori* para la media y la desviación estándar de  $f(x)$ , estarían dados por

$$\mu = \exp(\mu_l + 0.5\sigma_l^2) = wD \quad (5.3)$$

$$\sigma = \{\exp(2\mu_l + \sigma_l^2)[\exp(\sigma_l^2) - 1]\}^{1/2} = wD(w^2 - 1)^{1/2} \quad (5.4)$$

Otro procedimiento, utilizado por Wolfson(1995) y Percy(2002), para definir los hiperparámetros de las distribuciones *a priori* para los parámetros de algunos modelos, consiste en obtener la forma analítica de la distribución *a priori* predictiva acumulada  $F(x)$ , e igualar ésta con los cuantiles dados por el experto; con lo que se obtienen ecuaciones simultáneas que tienen como incógnita a los hiperparámetros. El único caso que en estas obras se trabajan es el del modelo exponencial ( $\theta$ ), suponiendo una distribución *a priori* gama( $\alpha, \beta$ ) para  $\theta$ . Con esto la distribución acumulada *a priori* predictiva está dada por  $F(x) = 1 - [\beta/(\beta + x)]^\alpha$ . Por lo tanto si el usuario da *a priori* un par de cuantiles,  $t_{p_1}$  y  $t_{p_2}$ , entonces se obtienen las siguientes ecuaciones simultáneas no lineales:  $F(t_{p_1}) = p_1$  y  $F(t_{p_2}) = p_2$ . Al resolverlas se obtienen los valores de  $\alpha$  y  $\beta$ . Adicionalmente Wolfson (1995) señala sobre lo anterior que es deseable que el experto dé más de dos cuantiles, y bajo esta circunstancia propone estimar por mínimos cuadrados los parámetros  $\alpha$  y  $\beta$ . Así si el usuario da el valor de cinco cuantiles,  $t_{p_1}, \dots, t_{p_5}$ , que pueden ser los correspondientes a  $p_i = 0.10, 0.25, 0.50, 0.75, 0.90$ ; por ejemplo, entonces los hiperparámetros  $\alpha$  y  $\beta$  son aquellos que minimizan la siguiente expresión:

$$\sum_{i=1}^5 [t_{p_i} - F^{-1}(t_{p_i})]^2$$

donde  $F^{-1}(t_{p_i}) = \beta \left[ \left( \frac{1}{1-p_i} \right)^{1/\alpha} - 1 \right]$  es la función cuantil de la *a priori* predictiva para el modelo exponencial.

Sin embargo, este procedimiento para definir los hiperparámetros para el modelo exponencial ya sea con base en dos o más cantiles, es difícil extenderlo a otros modelos. Al respecto Percy (2002) reconoce la dificultad de extender este enfoque a otros modelos más complicados como el Weibull, ya que entre otras cosas, éste modelo no tiene distribución conjugada y no se ha encontrado alguna distribución *a priori* para sus dos parámetros que permita obtener una distribución *a priori* predictiva y una distribución posterior con expresiones analíticas. Algo similar a lo anterior puede decirse respecto a la distribución valor extremo. De hecho es típico que para estos dos modelos se suponga independencia entre sus parámetros y que para el parámetro de escala se utilice una distribución normal y para el de forma una distribución gama (ver Ibrahim et al. 2001, pag. 35).

En Singpurwalla (1988) se describe un algoritmo para especificar los hiperparámetros de la distribuciones *a priori* para los parámetros de un modelo Weibull( $\eta, \theta$ ). Dado que en el caso de esta distribución la mediana para valores dados de ( $\eta, \theta$ ), está dada por  $D = \eta(\log(2))^{1/\theta}$ ; en el trabajo en cuestión se supone que el experto es capaz de trasladar su incertidumbre sobre  $D$  en términos de alguna distribución con media  $m$  y desviación estándar  $s$ . Suponiendo que tal distribución es normal, se propone una distribución normal truncada como distribución *a priori* para  $D$ . Respecto al otro parámetro del modelo Weibull,  $\theta$ , consideran una distribución *a priori* gama( $\alpha, \beta$ ), y dado que la media y varianza de esta distribución están dadas por  $\alpha/\beta$  y  $\alpha/\beta^2$ , respectivamente; entonces la incertidumbre del experto sobre  $\theta$  se puede expresar nuevamente en términos de una media y una desviación estándar, que al ser igualadas con  $\alpha/\beta$  y  $\sqrt{\alpha/\beta^2}$ , se obtienen dos ecuaciones simultáneas de donde se despejan los hiperparámetros  $\alpha$  y  $\beta$ . Para integrar estas dos distribuciones *a priori* al algoritmo que proponen para obtener la función de confiabilidad posterior, suponen independencia entre  $D$  y  $\theta$ . Nótese que en esta propuesta  $D$  es la mediana del modelo, que no necesariamente es igual a la mediana de la distribución *a priori* predictiva.

Para los modelo normal y lognormal, al usar como distribuciones *a priori* para sus dos parámetros una distribución normal-gama, sus distribuciones *a priori* predictivas son una *t-Student* y *logt-Student*, respectivamente (ver Bernardo y Smith, 1994; y Percy, 2002). En forma específica estas distribuciones *a priori* predictivas, utilizando la misma notación establecida en la sección 4.1, están dadas por una *t-Student*( $m_1, \lambda_1, 2\alpha_1$ ) y *logt-Student*( $m_2, \lambda_2, 2\alpha_2$ ); cuyas expresiones analíticas son, respectivamente:

$$f_1(x) = \frac{\Gamma[(2\alpha_1 + 1)/2]}{\Gamma[\alpha_1] \Gamma[1/2]} \left( \frac{\lambda_1}{2\alpha_1} \right)^{1/2} \left( \frac{\lambda_1}{2\alpha_1} (x - m_1)^2 + 1 \right)^{-\frac{2\alpha_1+1}{2}}$$

$$f_2(x) = \frac{\Gamma[(2\alpha_2 + 1)/2]}{\Gamma[2\alpha_2] \Gamma[1/2] x} \left( \frac{\lambda_2}{2\alpha_2} \right)^{1/2} \left( \frac{\lambda_2}{2\alpha_2} (\log(x) - m_2)^2 + 1 \right)^{-\frac{2\alpha_2+1}{2}}$$

con  $\lambda_1 = \beta_1^{-1} (k_1 + 1)^{-1} k_1 \alpha_1$  y  $\lambda_2 = \beta_2^{-1} (k_2 + 1)^{-1} k_2 \alpha_2$ . Por tanto, extendiendo el método propuesto en Wolfson (1995) y Percy (2002) para el modelo exponencial, el

usuario tendría que dar *a priori* cuatro cuantiles para la distribución *a priori* predictiva,  $t_{p_i}$  ( $i = 1, \dots, 4$ ), y luego se tendría que resolver cuatro ecuaciones en forma simultánea para cada modelo. Estas ecuaciones tienen la siguiente forma para cada modelo  $M_j$

$$\int_{-\infty}^{t_{p_i}} f_j(x) dx = p_i \quad \text{con } i = 1, \dots, 4$$

que deben ser resueltas en forma simultánea en términos de los cuatro hiperparámetros  $\alpha_j, \beta_j, m_j$  y  $k_j$ . Pero resolver estas ecuaciones no lineales de manera simultánea no parece sencillo, para empezar porque las integrales ahí indicadas no tienen solución analítica, es decir, porque no existe forma analítica para la distribución acumulada *a priori* predictiva. En suma para resolver estas ecuaciones habría que generar el algoritmo numérico correspondiente, que no parece ser un trabajo sencillo.

Aunque se contara con tal algoritmo no sería suficiente en el caso de selección de modelos, ya que como lo señalamos antes, en el caso de los modelos Weibull y valor extremo no existe una expresión analítica ni siquiera para la densidad *a priori* predictiva. Además no se trata de tener diferentes procedimientos para cada uno de los modelos que se están considerando en un momento dado. Más bien requerimos de un procedimiento único para definir los hiperparámetros de todos los modelos considerados, para de esa forma no favorecer a ninguno de los modelos. Además pensando en otras aplicaciones un procedimiento único para la especificaciones de los hiperparámetros en modelos de confiabilidad, tendría la ventaja de disminuir la necesidad de metodologías que en un momento dado tendría que utilizar el usuario de la estadística.

Con base en lo anterior podemos concluir que no existe un método que permita especificar los hiperparámetros de las distribuciones *a priori* de los cinco parámetros de los modelos de confiabilidad que en este trabajo se están considerando, que además responda a las necesidad de imparcialidad que se requieren en el contexto de selección de modelos. Por ello a continuación proponemos un método que cumple con las características requeridas.

## 5.2 Intervalos iniciales para la media y la desviación estándar

Nosotros consideramos que es factible que el usuario de la estadística tenga información *a priori* sobre la distribución del tiempo de vida que le permita establecer rangos donde espera ocurra la media o la mediana del tiempo de vida, y también rangos donde espera estén algunos cuantiles, con lo cuales se pueda establecer un rango de ocurrencia para la desviación estándar del tiempo de vida. Es decir, consideramos que es más realista que el usuario pueda aportar tales intervalos, a que aporte el valor exacto de la mediana o la media como Garthwaite y O'Hagan(2000) y O'Hagan (1998) y/o a que el experto dé un valor para varios cuantiles del tiempo de vida, como se supone en los dos últimos trabajos citados antes y en Kadane y Wolfson (1998) y Percy (2002). Nosotros generaremos un método único para especificar el valor de los hiperparámetros de la distribución *a priori* para los parámetros de los diferentes modelos que se han considerado en este trabajo, suponiendo que el usuario es capaz de aportar los intervalos en los que espera ocurran la media (o mediana) y un par de cuantiles. El método estará basado en la información que tiene el usuario sobre los observables (tiempo de vida),

pero no estará fundamentado en la densidad *a priori* predictiva  $f(x)$  por las dificultades que ya se comentaron antes (con excepción del modelo exponencial no existen formas analíticas para la distribución acumulada *a priori* predictiva). Más bien partiremos directamente del modelo de tiempo de vida,  $f(x|\theta)$ , en cierto sentido de forma similar a como lo hace Singpurwalla (1988) para especificar la mediana del tiempo de vida.

Supongamos que el experto es capaz de proponer intervalos donde espera que ocurra un par de cuantiles del tiempo de vida del producto. Sean  $t_{p_1}$  y  $t_{p_2}$  estos cuantiles ( $t_{p_1} < t_{p_2}$ ) y sus intervalos no traslapados  $[it_{p_1}, st_{p_1}]$  y  $[it_{p_2}, st_{p_2}]$ , respectivamente. A partir de lo anterior, e inspirados en la forma en que en Garthwaite y O'Hagan(2000) y O'Hagan (1998) se estima la desviación estándar *a priori*  $S$ , nosotros proponemos dos formas de estimar el rango en el que se espera ocurra  $S$ , dependiendo de si se supone una distribución del tiempo de vida simétrica o no. Para el caso en que se suponga simétrica, nosotros proponemos que el intervalo de ocurrencia para la desviación estándar,  $S$ , esté dado por  $[LI_s, LS_s]$ , donde

$$LI_s = \frac{it_{p_2} - st_{p_1}}{z_p} \quad \text{y} \quad LS_s = \frac{st_{p_2} - it_{p_1}}{z_p} \quad (5.5)$$

donde  $z_p = \phi^{-1}(p_2) - \phi^{-1}(p_1)$  es el rango intercuantílico de la distribución normal estándar. Nótese que estas expresiones son similares a (5.1), sólo que en nuestro caso se ha considerado la distancia más corta y más larga entre los límites iniciales para los intervalos de los cuantiles.

Para la media  $M$ , como se supone que la distribución es simétrica y unimodal, entonces proponemos que el usuario dé directamente un intervalo donde espera que ocurra el tiempo medio de vida, sea este intervalo  $[LI_m, LS_m]$ .

Estos intervalos incorporan algo que siempre va a estar presente: la incertidumbre del usuario. Esto es una ventaja sobre las propuestas descritas antes. Además estos intervalos protegen de una subestimación o una sobrestimación sobre la percepción de la distribución del tiempo de vida, que son señaladas en O'Hagan (1998) y en las referencias citadas en ese mismo trabajo.

Para el caso más común en confiabilidad en que se suponga una distribución asimétrica (con sesgo positivo), inspirados en Garthwaite y O'Hagan(2000) y O'Hagan (1998), en particular en las expresiones (5.2), (5.3) y (5.4), nosotros proponemos que los intervalos para la media  $M$  y la desviación estándar del tiempo de vida  $S$  se obtenga como se describe enseguida.

Igual que antes el usuario da intervalos no traslapados  $[it_{p_1}, st_{p_1}]$  y  $[it_{p_2}, st_{p_2}]$  para los cuantiles  $t_{p_1}$  y  $t_{p_2}$  del tiempo de vida, respectivamente. Ahora por la falta de simetría el usuario da un intervalo de ocurrencia para la mediana, sea  $[D_i, D_s]$  este intervalo. De forma similar a (5.2), obtengamos

$$w_i = \exp \left\{ \frac{[\log(it_{p_2}) - \log(st_{p_1})]^2}{2z_p^2} \right\}$$

$$w_s = \exp \left\{ \frac{[\log(st_{p_2}) - \log(it_{p_1})]^2}{2z_p^2} \right\}$$

además sea  $\bar{w} = (w_i + w_s)/2$  y  $\bar{D} = (D_i + D_s)/2$ . Respaldados en (5.3), proponemos que el intervalo  $[LI_m, LS_m]$  para la media se obtenga con

$$LI_m = \bar{w}D_i \text{ y } LS_m = \bar{w}D_s \quad (5.6)$$

y tomando en cuenta (5.4), proponemos que el intervalo  $[LI_s, LS_s]$  para la desviación estándar se obtendrá de acuerdo a

$$LI_s = w_i\bar{D}(w_i^2 - 1)^{1/2} \text{ y } LS_s = w_s\bar{D}(w_s^2 - 1)^{1/2}. \quad (5.7)$$

Sobre los cuantiles específicos sobre los que el usuario debe generar los intervalos, aunque el procedimiento formalmente no requiere algún para especial de cuantiles, nos unimos a la sugerencia de O'Hagan(1998), de que éstos no sean percentiles muy pequeños o muy grandes. De aquí que en la media de lo posible se sugiere que se elijan los cuantiles de tal forma que el rango intercuantílico cubra el 33.3%, 50% ó 66% de los valores de  $x$ ; lo que se puede obtener, respectivamente, con la siguientes elección de cuantiles: 0.33 – 0.67, 0.25 – 0.75, 0.17 – 0.83; en estos casos  $z_p$  tomaría los valores 0.8614, 1.3490 y 1.9348, respectivamente. La decisión de que pareja de cuantiles elegir, se toma con base en el conocimiento *a priori* que tenga el experto sobre la distribución del tiempo de vida.

Dados los intervalos para la media,  $M$ , y la desviación estándar,  $S$ , del tiempo de vida. Es posible reexpresar esta incertidumbre del usuario de otra forma. Para ello, dado que  $M$  y  $S$  son variables aleatorias y si recurrimos a las propiedades de la distribución normal y a la cota de Camp y Meidel para la desigualdad de Chebyshev para distribuciones unimodales (ver Duncan, 1989, pag. 103 ), es razonable como regla de “dedo” traducir la información que aporta el usuario, de la siguiente manera:

$$\begin{aligned} E(M) &= \frac{(LI_m + LS_m)}{2} \\ V(M) &= \left[ \frac{(LS_m - LI_m)}{6} \right]^2 \\ E(S) &= \left[ \frac{(LI_s + LS_s)}{2} \right] \\ V(S) &= \left[ \frac{(LS_s - LI_s)}{6} \right]^2 \end{aligned} \quad (5.8)$$

Esta información a su vez vamos a ligarla con las expresiones para la media y la varianza para cada uno de los modelos, que son función de sus parámetros, y será la base para proponer un procedimiento para definir los hiperparámetros. Esto lo explicaremos para cada modelo en el las secciones siguientes. Antes de ello, a continuación señalamos porque hemos utilizado en el trabajo dos tipos de distribuciones *a priori*.

Es bien conocido que el factor de Bayes y por extensión la probabilidad posterior para un modelo es sensible a la distribución *a priori* utilizada (Sinharay y Stern, 2002). Por ello, pensando en una distribución informativa y con la idea de que la información inicial para todos los modelos sea similar, utilizamos una distribución normal-gama( $\alpha_i, \beta_i, \mu_i, k_i$ ) como *a priori* para los parámetros de los modelos normal, lognormal,

valor extremo y Weibull (ver 4.4). Se ha hecho así, porque esta distribución es conjugada para los modelos normal y lognormal (Bernardo y Smith, 1994, sec. 5.2); y dado que los modelos valor extremo y Weibull no tienen una distribución conjugada (Ibrahim, et al. 2001), entonces se ha dejado para estos modelos también la distribución normal-gama, de esta manera se deja en igualdad de circunstancias a los cuatro modelos que tienen dos parámetros. En el caso del modelo exponencial (4.14), se propone como distribución informativa inicial a una gama, debido a que ésta es su conjugada.

Adicionalmente, se ha evaluado como distribución *a priori* para los parámetros de todos los modelos considerados a una uniforme  $(a_i, b_i, d_i, e_i)$ , ver (4.5), dado que esta distribución es propia pero poco informativa (vaga). Esto con la intención de disminuir el riesgo de que la información inicial influya en la especificación del modelo, y esto coincidiría con lo que sugieren varios autores, en el sentido de que la *a priori* sea relativamente plana sobre el rango de los posibles valores de los parámetros (ver por ejemplo George, 1999; y Kass y Raftery, 1995). En suma, se ha propuesto dos tipos de distribuciones *a priori*: normal-gama y uniforme, una informativa y otra vaga; con la intención de evaluar ambas opciones respecto a su posible influencia en la especificación del modelo y en general en la inferencia.

### 5.3 Modelo Normal

Si  $X$  sigue una distribución normal, y dados valores de sus parámetros  $\mu$  y  $\tau$ , entonces

$$\begin{aligned} E(X \mid \mu, \tau) &= \mu \\ V(X \mid \mu, \tau) &= \tau^{-1}. \end{aligned} \quad (5.9)$$

La información para la media,  $M$ , y la desviación estándar,  $S$ , dados por (5.8) y obtenidas directamente de los intervalos aportados por el usuario; puede verse como información para  $E(X \mid \mu, \tau)$  y  $\sqrt{V(X \mid \mu, \tau)}$ ; y en esto a su vez tiene un equivalente en términos de los valores de los parámetros  $\mu$  y  $\tau$ , como lo veremos enseguida para cada tipo de distribución *a priori*.

A priori normal-gama. Suponiendo una distribución inicial normal-gama  $(\alpha_1, \beta_1, m_1, k_1)$  para los parámetros  $(\mu, \tau)$  del modelo normal. Analizando la información que aporta el usuario sintetizada en (5.8), el significado de los parámetros  $(\mu, \tau)$  y las expresiones para la media y la varianza del modelo normal (5.9), es natural suponer que

$$\begin{aligned} E(\mu) &= E(M) \\ V(\mu) &= V(M) \end{aligned}$$

Como en general  $\tau = 1/S^2$ , define la relación entre el parámetro de precisión,  $\tau$ , y la desviación estándar,  $S$ . Entonces sin presuponer ninguna distribución, es razonable que

$$E(\tau) = \frac{1}{E^2(S)}$$

Además como un valor extremo de  $S$  es  $E(S) + 3\sqrt{V(S)} = LS_s$ , entonces  $1/[LS_s]^2$  se puede ver como un valor extremo superior para la distribución de  $\tau$ . Luego, inspirados en las propiedades de la distribución normal, al dividir entre tres la distancia entre  $E(\tau)$  y  $1/[LS_s]^2$  se obtiene la desviación estándar de  $\tau$ . De esta forma como regla de "dedo" el valor de  $V(\tau)$  se obtiene de la siguiente manera:

$$V(\tau) = \left[ \frac{E(\tau) - \frac{1}{[LS_s]^2}}{3} \right]^2$$

Así bajo la suposición inicial de que  $(\mu, \tau)$  siguen la distribución normal-gama  $(\alpha_1, \beta_1, m_1, k_1)$ , ver (4.4), entonces  $\tau$  sigue una distribución gama  $(\alpha_1, \beta_1)$  y como la esperanza y la varianza de una distribución gama está dada por  $\alpha_1/\beta_1$  y  $\alpha_1/\beta_1^2$ , respectivamente; luego

$$\begin{aligned} \frac{\alpha_1}{\beta_1} &= E(\tau) \\ \frac{\alpha_1}{\beta_1^2} &= V(\tau) \end{aligned}$$

resolviendo para  $\alpha_1$  y  $\beta_1$ , obtenemos:

$$\begin{aligned} \beta_1 &= \frac{E(\tau)}{V(\tau)} \\ \alpha_1 &= \frac{E^2(\tau)}{V(\tau)} \end{aligned}$$

Además, como la distribución marginal de  $\mu$  es  $t$  con parámetro de localización  $m_1$ ,  $2\alpha_1$  grados de libertad y precisión  $\alpha_1 k_1 / \beta_1$  (DeGroot, 1970, pag. 171). De aquí que

$$\begin{aligned} E(\mu) &= m_1 \\ &= E(M) \end{aligned}$$

$$\begin{aligned} V(\mu) &= \frac{\beta_1}{k_1(\alpha_1 - 1)} \\ &= V(M) \end{aligned}$$

luego

$$k_1 = \frac{\beta_1}{V(M)(\alpha_1 - 1)}$$

Con esto tenemos completamente especificados los valores de los hiperparámetros de la distribución *a priori* Normal-Gama para los parámetros  $(\mu, \tau)$  del modelo normal, con base a la información inicial que aporta el usuario.

Este procedimiento lo vamos a generalizar al resto de los modelos considerados en este trabajo, y estamos conscientes que quizás no sea un procedimiento muy ortodoxo, ya que la información que aporta el usuario en realidad es sobre la distribución *a priori* predictiva. Sin embargo, como ya lo hemos dicho la distribución acumulada *a priori* predictiva de la mayoría de los modelos utilizados en confiabilidad no tienen forma analítica y esto genera complicaciones que hasta donde sabemos no han sido resueltas. Por ello estamos proponiendo este otro enfoque, que tiene parte de similitud a lo hecho por Singpurwalla (1988) para tener información sobre el parámetro de escala de una distribución Weibull mediante la especificación por parte del usuario de la mediana del modelo para el tiempo de vida. En cualquier caso, en la sección final de este capítulo demostraremos empíricamente que el método funciona al menos para los modelos que hemos considerado en este trabajo.

A priori uniforme. Si como distribución inicial para  $(\mu, \tau)$  se utiliza una distribución uniforme  $(a_1, b_1, d_1, e_1)$ , ver (4.5), entonces dado que deseamos que los parámetros de cada distribución siga una distribución uniforme, eso implica que cuando el usuario da los intervalos para la media y la desviación estándar, no esté dando ninguna preferencia por alguna zona del intervalo donde haya más posibilidades de que se encuentre tal media y desviación estándar. Con los argumentos anteriores, dados los intervalos iniciales que aporta el usuario, y analizando las ecuaciones de (5.9), entonces es justificable que los parámetros de la distribución uniforme para  $\mu$ , estén dados por

$$a_1 = LI_m \quad y \quad b_1 = LS_m.$$

De la misma manera los parámetros de la distribución uniforme para  $\tau$ , proponemos estén dados por

$$d_1 = \frac{1}{(LS_s)^2} \quad y \quad e_1 = \frac{1}{(LI_s)^2}.$$

#### 5.4 Modelo Lognormal

Si  $X$  tiene la distribución lognormal dada en (4.6), entonces tenemos que :

$$E(X \mid \mu, \tau) = \exp\left(\mu + \frac{0.5}{\tau}\right)$$

$$V(X \mid \mu, \tau) = \exp(2\mu + \tau^{-1})[\exp(\tau^{-1}) - 1]$$

Despejando  $\mu$  y a  $\tau$ , y simplificando notación, obtenemos:

$$\mu = \log(E(X)) - 0.5 \log\left(\frac{V(X)}{E^2(X)} + 1\right)$$

$$\tau = 1/\log\left(\frac{V(X)}{E^2(X)} + 1\right)$$

De aquí y siguiendo un razonamiento similar a lo hecho para el modelo normal, se obtiene

$$E(\mu) = \log(E(M)) - 0.5 \log\left(\frac{E^2(S)}{E^2(M)} + 1\right) \quad (5.10)$$

$$E(\tau) = \frac{1}{\log\left(\frac{E^2(S)}{E^2(M)} + 1\right)} \quad (5.11)$$

A priori normal-gama. Utilizando como distribución inicial para  $(\mu, \tau)$  una distribución normal gama  $(\alpha_2, \beta_2, m_2, k_2)$ . Entonces como  $\mu$  es el parámetro con mayor impacto sobre la localización de la distribución, entonces para especificar  $V(\mu)$ , proponemos derivarla a partir del valor más pequeño que el usuario esperaría para la media,  $LI_m$ , de la siguiente manera:

$$\begin{aligned} V(\mu) &= \left( \frac{E(\mu) - \left[ \log(LI_m) - 0.5 \log\left(\frac{E^2(S)}{E^2(M)} + 1\right) \right]}{3} \right)^2 \\ &= \left[ \frac{\log\left(\frac{E(M)}{LI_m}\right)}{3} \right]^2 \end{aligned}$$

De forma similar para especificar  $V(\tau)$  se parte del valor más grande que el usuario esperaría para la desviación estándar,  $LS_s$ , y sustituyendo  $(LS_s)^2$  por  $E^2(S)$  en (5.11), y procediendo de manera similar al caso normal, tenemos que:

$$V(\tau) = \left[ \frac{E(\tau) - \frac{1}{\log\left(\frac{LS_s^2}{E^2(M)} + 1\right)}}{3} \right]^2$$

De esta manera a partir de la información proporcionada por el usuario hemos completado los valores esperados y las varianzas de los parámetros de la distribución lognormal  $(E(\mu), V(\mu), E(\tau)$  y  $V(\tau))$ . Por lo que utilizando los mismos argumentos que en el caso de la distribución normal, obtenemos que los valores de los hiperparámetros de la correspondiente distribución *a priori* normal-gama, están dados por:

$$\beta_2 = \frac{E(\tau)}{V(\tau)}$$

$$\alpha_2 = \frac{E^2(\tau)}{V(\tau)}$$

$$m_2 = E(\mu)$$

$$k_2 = \frac{\beta_2}{V(\mu)(\alpha_2 - 1)}$$

A priori uniforme. En el caso de que como distribución inicial para  $(\mu, \tau)$  se utilice una distribución uniforme  $(a_2, b_2, d_2, e_2)$ , ver (4.5). Entonces dada (5.10), el rango en el que se esperaría ocurra  $\mu$  y en consecuencia los parámetros de su correspondiente distribución uniforme estarían dados por

$$a_2 = \log(LI_m) - 0.5 \log \left( \frac{E^2(S)}{E^2(M)} + 1 \right) \quad \text{y} \quad b_2 = \log(LS_m) - 0.5 \log \left( \frac{E^2(S)}{E^2(M)} + 1 \right).$$

Respecto a  $\tau$ , observando (5.11) proponemos que los valores extremos de  $\tau$  y en consecuencia los parámetros de su distribución uniforme estén dados por

$$d_2 = \frac{1}{\log \left( \frac{LS_m^2}{E^2(M)} + 1 \right)} \quad \text{y} \quad e_2 = \frac{1}{\log \left( \frac{LI_m^2}{E^2(M)} + 1 \right)}.$$

### 5.5 Modelo Valor Extremo

Si  $X$  sigue una distribución valor extremo dada en (4.8), entonces

$$E(X \mid \mu, \tau) = \mu - \frac{0.5772}{\tau}$$

$$V(X \mid \mu, \tau) = \frac{\pi^2}{6\tau^2}$$

donde 0.5772 es la constante de Euler. Despejando de estas ecuaciones a  $\mu$  y a  $\tau$ , y simplificando notación, obtenemos:

$$\mu = E(X) + \frac{0.5772\sqrt{6}}{\pi} \sqrt{V(X)}$$

$$\tau = \frac{\pi}{\sqrt{6}\sqrt{V(X)}} \quad (5.12)$$

A priori normal-gama. Si se utiliza la distribución normal-gama  $(\alpha_3, \beta_3, m_3, k_3)$  de (4.11), como *a priori* para  $(\mu, \tau)$ . Entonces de (5.12) y utilizando razonamientos similares a los hechos para las distribuciones que ya se han analizado, se obtiene que

$$E(\mu) = E(M) + 0.45004E(S)$$

$$E(\tau) = \frac{\pi}{\sqrt{6}E(S)}$$

Nuevamente al ser  $\mu$  un parámetro de localización, entonces para especificar  $V(\mu)$  se toma en cuenta el valor más pequeño que el usuario esperaría para la media,  $LI_m$ , de la siguiente manera:

$$\begin{aligned}
 V(\mu) &= \left( \frac{E(\mu) - (LI_m + 0.45004E(S))}{3} \right)^2 \\
 &= \left( \frac{E(M) - LI_m}{3} \right)^2
 \end{aligned}$$

Para especificar  $V(\tau)$  se parte de (5.12) y del valor más grande que el usuario esperaría para la desviación estándar,  $LS_s$ , de la siguiente manera:

$$V(\tau) = \left[ \frac{E(\tau) - \frac{\pi}{\sqrt{6}LS_s}}{3} \right]^2$$

De esta forma con  $E(\mu)$ ,  $V(\mu)$ ,  $E(\tau)$  y  $V(\tau)$  en términos de la información proporcionada por el usuario, y dado que la distribución *a priori* para  $(\mu, \tau)$  es una normal-gama, entonces utilizando los mismos argumentos que en el caso de la distribución normal, obtenemos que los valores de los hiperparámetros de la distribución *a priori* están dados por:

$$\begin{aligned}
 \beta_3 &= \frac{E(\tau)}{V(\tau)} \\
 \alpha_3 &= \frac{E^2(\tau)}{V(\tau)}
 \end{aligned}$$

$$m_3 = E(\mu)$$

$$k_3 = \frac{\beta_3}{V(\mu)(\alpha_3 - 1)}$$

A priori uniforme. En caso de usar una distribución uniforme  $(a_3, b_3, d_3, e_3)$  como priori para los parámetros  $(\mu, \tau)$  del modelo valor extremo, entonces analizando (5.12) y procediendo de una manera similar a lo hecho para el modelo normal, el rango en el que se espera ocurra  $\mu$  está dado por

$$a_3 = LI_m + 0.45004E(S) \quad \text{y} \quad b_3 = LS_m + 0.45004E(S).$$

De la misma manera a partir de (5.12), el rango en el que se espera ocurra  $\tau$  y los correspondientes parámetros de su distribución uniforme estarían dados por

$$d_3 = \frac{\pi}{\sqrt{6}LS_s} \quad \text{y} \quad e_3 = \frac{\pi}{\sqrt{6}LI_s}.$$

## 5.6 Modelo Weibull

Si  $X$  tiene una distribución Weibull( $\eta, \theta$ ) dada por (4.12), entonces se tiene que

$$\begin{aligned} E(X \mid \eta, \theta) &= \eta \Gamma\left(1 + \frac{1}{\theta}\right) \\ V(X \mid \eta, \theta) &= \eta^2 \left\{ \Gamma\left(1 + \frac{2}{\theta}\right) - \Gamma^2\left(1 + \frac{1}{\theta}\right) \right\} \end{aligned} \quad (5.13)$$

Es necesario resolver estas ecuaciones en términos de  $\eta$  y  $\theta$ , como hemos venido haciendo en los casos anteriores. Al ser estas ecuaciones no lineales es necesario recurrir a un método numérico para su solución. Una forma de simplificar la solución es observando lo siguiente:

$$\frac{E(X)}{\sqrt{V(X)}} = \frac{\Gamma\left(1 + \frac{1}{\theta}\right)}{\left\{ \Gamma\left(1 + \frac{2}{\theta}\right) - \Gamma^2\left(1 + \frac{1}{\theta}\right) \right\}^{1/2}} \equiv g(\theta) \quad (5.14)$$

De esta manera tenemos la función no lineal  $g(\theta)$  sólo en términos de  $\theta$ . Para resolver en términos de  $\theta$  es necesario recurrir a métodos numéricos. Si se opta por utilizar el método de Newton-Rapson dado un valor inicial para el parámetro,  $\theta_0$ , el siguiente valor de  $\theta_1$  será dado por:

$$\theta_1 = \theta_0 - \frac{g(\theta_0)}{g'(\theta_0)}$$

donde la derivada para la función  $g(\theta)$  está dada por

$$\begin{aligned} g'(\theta) &= \frac{1}{2} \Gamma\left(1 + \frac{1}{\theta}\right) \times \\ & \frac{-2\Gamma\left(1 + \frac{2}{\theta}\right)\Psi\left(1 + \frac{1}{\theta}\right) + 2\Gamma^2\left(1 + \frac{1}{\theta}\right)\Psi\left(1 + \frac{1}{\theta}\right) + 2 \left[ \Gamma\left(1 + \frac{2}{\theta}\right)\Psi\left(1 + \frac{2}{\theta}\right) - \Gamma^2\left(1 + \frac{1}{\theta}\right)\Psi\left(1 + \frac{1}{\theta}\right) \right]}{\theta^2 \left( \sqrt{\Gamma\left(1 + \frac{2}{\theta}\right) - \Gamma^2\left(1 + \frac{1}{\theta}\right)} \right)^3} \end{aligned}$$

Para llegar a esta expresión se recurrió al siguiente resultado (Abramowitz y Stegun, 1965):

$$\frac{d\Gamma(v)}{dv} = \Psi(v)\Gamma(v)$$

donde  $\Psi(v)$  es la función digama, que está definida por:

$$\begin{aligned} \Psi(v) &= \frac{d \log \Gamma(v)}{dv} \\ &= \log(v) - \frac{1}{2v} - \frac{1}{12v^2} + \frac{1}{120v^4} - \frac{1}{252v^6} + \dots \end{aligned}$$

De esta forma para encontrar el valor de  $\theta$  aplicando el método de Newton-Rapson, faltaría proponer un valor inicial. Con la idea de generar un algoritmo que con la información inicial dado por el usuario se genere tal valor inicial de  $\theta$ , vamos a recurrir a propiedades del modelo Weibull. De acuerdo a esto, la función cuantil del modelo Weibull para valores dados de  $(\eta, \theta)$  está definida por:

$$X_p = \eta[-\log(1-p)]^{1/\theta}$$

de aquí sea  $p = 0.63 = (1 - e^{-1})$ , por lo tanto el cuantil  $X_{0.63} = \eta$ . Entonces hacemos  $\eta_0 = E(M)$ , ya que no deben ser tan diferentes entre sí. Por otro lado si pensamos en el cuantil 0.7, entonces la distancia entre éste y el cuantil 0.63 debe ser una fracción de la desviación estándar  $E(S)$ , por ejemplo:

$$X_{0.7} = E(M) + 0.2E(S),$$

pero de acuerdo a la función cuantil:

$$X_{0.7} = \eta[-\log(1-0.7)]^{1/\theta}$$

igualando estas dos últimas ecuaciones, sustituyendo en la primera a  $\eta$  por  $E(M)$ , y despejando  $\theta$  obtenemos como valor inicial de  $\theta$  a:

$$\theta_0 = \frac{\log(-\log(0.3))}{\log\left(1 + \frac{0.2E(S)}{E(M)}\right)}$$

Una vez encontrado el valor de  $\theta$ , es fácil encontrar el valor de  $\eta$ , ya que de la ecuación (5.13), se ve que:

$$\eta = \frac{E(X)}{\Gamma\left(1 + \frac{1}{\theta}\right)} \quad (5.15)$$

A priori normal-gama. Si se utiliza una distribución normal-gama  $(\alpha_4, \beta_4, m_4, k_4)$  como distribución inicial para  $(\eta, \theta)$ . Entonces de (5.14), (5.15) y utilizando razonamientos similares a los hechos para las distribuciones que ya se han analizado, estamos en posibilidades de explicar cómo proponemos asignar los valores de los hiperparámetros para tal distribución *a priori*. Para ello a continuación planteamos cómo calcular  $E(\eta)$ ,  $E(\theta)$ ,  $V(\eta)$  y  $V(\theta)$ . Proponemos definir a  $E(\theta)$  igual al  $\theta_h$  tal que:

$$E(\theta) = \left\{ \theta_h \mid g(\theta_h) - \frac{E(M)}{E(S)} = 0 \right\} \quad (5.16)$$

donde  $g(\theta)$  esta definida por (5.14). La solución se encuentra aplicando las recomendaciones que dimos antes para encontrar el cero de la función  $g(\theta)$ .

A partir de (5.15) proponemos calcular a  $E(\eta)$  de la siguiente manera:

$$E(\eta) = \frac{E(M)}{\Gamma\left(1 + \frac{1}{E(\theta)}\right)}$$

Por otro lado, analizando el comportamiento de las soluciones de

$$g(\theta) - \frac{E(M)}{E(S)} = 0$$

en función de diferentes valores de  $E(M)$  y  $E(S)$ ; y sus consecuencias en la forma de estimar  $\eta$  :

$$\eta = \frac{E(M)}{\Gamma(1 + \frac{1}{\theta})}$$

se aprecia que los cambios en  $E(M)$  afectan principalmente los valores de  $\eta$ ; mientras que los cambios en  $E(S)$  afectan predominantemente los valores de  $\theta$ . Por ello, dada la forma en que antes hemos propuesto calcular  $E(\eta)$  y  $E(\theta)$ , y siguiendo razonamientos similares a los hechos con las otras distribuciones, proponemos calcular a  $V(\eta)$  y  $V(\theta)$  de la siguiente manera. Sea  $\theta_s$  un valor extremo de los posibles valores de  $\theta$ , que se obtiene de la solución de la siguiente ecuación

$$\theta_s = \left\{ \theta \mid g(\theta) - \frac{E(M)}{LI_s} = 0 \right\} \quad (5.17)$$

con este valor  $\theta_s$  obtenemos  $V(\theta)$  de la siguiente manera

$$V(\theta) = \left( \frac{|E(\theta) - \theta_s|}{3} \right)^2$$

Ahora para obtener  $V(\eta)$ , se obtiene  $\theta_m$  como solución de la siguiente ecuación

$$\theta_m = \left\{ \theta \mid g(\theta) - \frac{LI_m}{E(S)} = 0 \right\} \quad (5.18)$$

Con este valor  $\theta_m$  obtenemos un  $\eta_m$  extremo, de la siguiente manera

$$\eta_m = \frac{LI_m}{\Gamma(1 + \frac{1}{\theta_m})} \quad (5.19)$$

y con este valor  $\eta_m$ , se obtiene  $V(\eta)$  de la siguiente manera

$$V(\eta) = \left( \frac{E(\eta) - \eta_m}{3} \right)^2$$

De esta forma con  $E(\eta)$ ,  $V(\eta)$ ,  $E(\theta)$  y  $V(\theta)$  en términos de la información proporcionada por el usuario, y dado que la distribución *a priori* para  $(\eta, \theta)$  es una normal-gama, entonces utilizando los mismos argumentos que en el caso de la distribución normal, se obtiene que los valores de los hiperparámetros de la distribución *a priori* están dados por:

$$\beta_4 = \frac{E(\theta)}{V(\theta)}$$

$$\alpha_4 = \frac{E^2(\theta)}{V(\theta)}$$

$$m_4 = E(\eta)$$

$$k_4 = \frac{\beta_4}{V(\eta)(\alpha_4 - 1)}$$

A priori uniforme. En caso de usar una distribución uniforme  $(a_4, b_4, d_4, e_4)$  como *a priori* para  $(\eta, \theta)$ , entonces para proceder de una manera similar a lo hecho para el caso normal, primero para determinar el rango en el que se espera ocurra  $\eta$  observemos (5.15), (5.14) y (5.16), entonces proponemos como rangos de ocurrencia para  $\eta$  a:

$$a_4 = \frac{LI_m}{\Gamma(1 + \frac{1}{E(\theta)})} \quad y \quad b_4 = \frac{LS_m}{\Gamma(1 + \frac{1}{E(\theta)})}$$

donde  $E(\theta)$  se obtiene de acuerdo a (5.16). Ahora respecto a  $\theta$ , observando (5.14) y analizando la forma que se obtuvo un extremo para  $\theta$  mediante (5.17); entonces de forma similar proponemos obtener valores extremos de  $\theta$  de la siguiente manera

$$\theta_1 = \left\{ \theta \mid g(\theta) - \frac{E(M)}{LI_s} = 0 \right\} \quad (5.20)$$

y

$$\theta_2 = \left\{ \theta \mid g(\theta) - \frac{E(M)}{LS_s} = 0 \right\} \quad (5.21)$$

Con estos valores proponemos que los parámetros de la distribución uniforme para  $\theta$  estén dados por

$$d_4 = \min(\theta_1, \theta_2) \quad y \quad e_4 = \max(\theta_1, \theta_2).$$

### 5.7 Modelo Exponencial

Si  $X$  tiene la distribución exponencial dada en (4.14), entonces tenemos que:

$$E(X \mid \theta, \gamma) = \gamma + (1/\theta)$$

por lo que despejando  $\theta$  y simplificando notación

$$\theta = \frac{1}{E(X) - \gamma} \quad (5.22)$$

A priori gama. Suponiendo conocido el parámetro  $\gamma$  y si la distribución *a priori* para  $\theta$  es una gama( $\alpha_5, \beta_5$ ), entonces procediendo de igual manera que en las subsecciones previas y observando (5.22), proponemos que

$$E(\theta) = \frac{1}{E(M) - \gamma}$$

y la varianza *a priori* para  $\theta$ :

$$V(\theta) = \left( \frac{\frac{1}{LI_m - \gamma} - \frac{1}{LS_m - \gamma}}{6} \right)^2$$

Como  $\theta$  sigue una distribución gama ( $\alpha_5, \beta_5$ ), y la esperanza y la varianza de tal distribución está dada por  $\alpha_5/\beta_5$  y  $\alpha_5/\beta_5^2$ , respectivamente, entonces

$$E(\theta) = \alpha_5/\beta_5 \quad \text{y} \quad V(\theta) = \alpha_5/\beta_5^2$$

resolviendo estas expresiones para  $\alpha_5$  y  $\beta_5$ , obtenemos que

$$\alpha_5 = \frac{E^2(\theta)}{V(\theta)}$$

$$\beta_5 = \frac{E(\theta)}{V(\theta)}$$

A priori uniforme. Suponiendo conocido el parámetro  $\gamma$  y si la distribución *a priori* para  $\theta$  es una uniforme( $a_5, b_5$ ), entonces procediendo de igual manera que en las subsecciones previas y observando (5.22), proponemos que el rango de ocurrencia para  $\theta$  y en consecuencia los parámetros para su distribución inicial estén dados por

$$a_5 = \frac{1}{ES_m - \gamma} \quad \text{y} \quad b_5 = \frac{1}{EI_m - \gamma}$$

## 5.8 Resumen

Hasta aquí hemos visto como obtener los parámetros de la distribución *a priori* normal-gama para los parámetros de los diferentes modelos, aunque hemos dejado expresiones implícitas. Por ejemplo, vimos que en el caso del modelo lognormal los parámetros de la distribución normal-gama( $\alpha_2, \beta_2, m_2, k_2$ ) están dados por

$$\beta_2 = \frac{E(\tau)}{V(\tau)}$$

$$\alpha_2 = \frac{E^2(\tau)}{V(\tau)}$$

$$m_2 = E(\mu)$$

$$k_2 = \frac{\beta_2}{V(\mu)(\alpha_2 - 1)}$$

donde  $E(\mu)$ ,  $E(\tau)$ ,  $V(\mu)$  y  $V(\tau)$  a su vez se obtienen a partir de la información inicial que aporta el usuario. Por esta razón, y con el ánimo de resumir lo hecho para cada modelo, en la tabla 5.1 se sintetiza la forma explícita de calcular los parámetros de la distribución *a priori* normal-gama para los parámetros de cada modelo. Algo similar se hace en la tabla 5.2 pero para el caso de una distribución *a priori* uniforme. Para la construcción de estas tablas hemos supuesto que con base en el conocimiento inicial del usuario sobre intervalos para los cuantiles de la distribución del tiempo de vida, y aplicando el método descrito en la sección 5.2, se han calculado intervalos en los que se espera ocurra la media,  $M$ , y la desviación estándar,  $S$ , del tiempo de vida. Recordemos que estos intervalos los denotamos con  $[LI_m, LS_m]$  y  $[LI_s, LS_s]$ , respectivamente. Además se debe considerar las expresiones (5.8), que son obtenidas de a partir de éstos intervalos. Sólo cabe enfatizar algunas expresiones para el modelo Weibull, donde en la tabla 5.1,  $E(\theta)$  se obtiene de (5.16),  $\theta_s$  de (5.17);  $\eta_m$  de (5.19) y (5.18); y en la tabla 5.2,  $E(\theta)$  se obtiene de (5.16),  $\theta_1$  de (5.20), y  $\theta_2$  de (5.21).

Modelo	$\beta_i$	$\alpha_i$	$m_i$	$k_i$
$M_1$	$\frac{1}{E^2(S) \left[ \frac{1}{E^2(S)} - \frac{1}{[LS_s]^2} \right]^2}$	$\frac{\beta_1}{E^2(S)}$	$E(M)$	$\frac{\beta_1}{V(M)(\alpha_1 - 1)}$
$M_2$	$\frac{\frac{9}{\log\left(\frac{E^2(S)}{E^2(M)} + 1\right)}}{\log^2\left[\left(\frac{LS_s^2}{E^2(M)} + 1\right)\left(\frac{E^2(S)}{E^2(M)} + 1\right)^{-1}\right]}$	$\frac{\beta_2}{\log\left(\frac{E^2(S)}{E^2(M)} + 1\right)}$	$\log\left(\frac{E^2(M)}{(E^2(S) + E^2(M))^{1/2}}\right)$	$\frac{9\beta_2}{(\alpha_2 - 1)\log^2\left(\frac{E(M)}{LI_m}\right)}$
$M_3$	$\frac{9\pi}{\sqrt{6}E(S)\left[E(\tau) - \frac{\pi}{\sqrt{6}LS_s}\right]^2}$	$\frac{\pi\beta_3}{\sqrt{6}E(S)}$	$E(M) + 0.45E(S)$	$\frac{9\beta_3}{(\alpha_3 - 1)[E(M) - LI_m]^2}$
$M_4$	$\frac{9E(\theta)}{ E(\theta) - \theta_s ^2}$	$E(\theta)\beta_4$	$\frac{E(M)}{\Gamma\left(1 + \frac{1}{E(\theta)}\right)}$	$\frac{9\beta_4\Gamma^2\left(1 + \frac{1}{E(\theta)}\right)(\alpha_4 - 1)^{-1}}{(E(M) - \eta_m\Gamma\left(1 + \frac{1}{E(\theta)}\right))^2}$
$M_5$	$\frac{\frac{36}{E(M) - \gamma}}{\left(\frac{1}{LI_m - \gamma} - \frac{1}{LS_m - \gamma}\right)^2}$	$\frac{\beta_5}{E(M) - \gamma}$		

Tabla 5.1: Cálculo de los hiperparámetros de la distribución normal-gama  $(\alpha_i, \beta_i, m_i, k_i)$ .

Modelo	$a_i$	$b_i$	$d_i$	$e_i$
Normal	$LI_m$	$LS_m$	$(LS_s)^{-2}$	$(LI_s)^{-2}$
Lognor	$\log\left(\frac{LI_m E(M)}{\sqrt{E^2(S)+E^2(M)}}\right)$	$\log\left(\frac{LS_m E(M)}{\sqrt{E^2(S)+E^2(M)}}\right)$	$\frac{1}{\log\left(\frac{LS_s^2}{E^2(M)}+1\right)}$	$\frac{1}{\log\left(\frac{LI_s^2}{E^2(M)}+1\right)}$
V.Extr.	$LI_m + 0.45E(S)$	$LS_m + 0.45E(S)$	$\frac{\pi}{\sqrt{6}LS_s}$	$\frac{\pi}{\sqrt{6}LI_s}$
Weib	$\frac{LI_m}{\Gamma(1+\frac{1}{E(\theta)})}$	$\frac{LS_m}{\Gamma(1+\frac{1}{E(\theta)})}$	$\min(\theta_1, \theta_2)$	$\max(\theta_1, \theta_2)$
Expon	$\frac{1}{ES_m-\gamma}$	$\frac{1}{EI_m-\gamma}$		

Tabla 5.2: Cálculo de los parámetros de la distribución uniforme( $a_i, b_i, d_i, e_i$ ).

### 5.9 Método general

En las secciones previas hemos aplicado el método que estamos proponiendo para especificar los valores de los hiperparámetros de las distribuciones *a priori* para los parámetros de los modelos que se han considerado en este estudio. Si en general consideramos modelos con dos parámetros  $\theta = (\theta_1, \theta_2)$ , el primero con mayor influencia en la media y el segundo en la varianza de la distribución del modelo considerado. Además dada la información dada por el usuario expresada en los intervalos iniciales para la media y la desviación estándar,  $[LI_m, LS_m]$  y  $[LI_s, LS_s]$ , respectivamente y a través de (5.8). De acuerdo a lo que hemos explicado en secciones previas, el método, con pequeñas variantes para cada modelo, consiste en general en los siguientes pasos:

1. Obtener  $E(X|\theta) = h_1(\theta_1, \theta_2)$  y  $V(X|\theta) = h_2(\theta_1, \theta_2)$
2. Haciendo  $E(X|\theta) = E(M)$  y  $V(X|\theta) = E^2(S)$ , resolver en términos de  $\theta_1$  y  $\theta_2$  las dos ecuaciones simultáneas que se generan:  $h_1(\theta_1, \theta_2) = E(M)$  y  $h_2(\theta_1, \theta_2) = E^2(S)$ .
3. Sea  $\theta_1^*$  y  $\theta_2^*$  la solución de estas ecuaciones, que se pueden expresar o al menos ver como funciones de  $E(M)$  y  $E^2(S)$ . Denotemos esto por las funciones  $r_1$  y  $r_2$  de la siguiente manera:  $\theta_j^* = r_j(E(M), E^2(S))$
4. En caso de utilizar una distribución normal-gama( $\alpha_i, \beta_i, m_i, k_i$ ), como para  $\theta_2$  se está suponiendo una distribución gama( $\alpha_i, \beta_i$ ), entonces hacer  $E(\theta_2) = \theta_2^* = \alpha_i/\beta_i$  y calcular  $V(\theta_2) = \{|r_2[E(M), LS_s^2] - r_2[E(M), LI_s^2]|/6\}^2 = \alpha_i/\beta_i^2$ . Resolver estas dos ecuaciones y obtener los valores de  $\alpha_i$  y  $\beta_i$ . Respecto a  $\theta_1$ , hacer  $E(\theta_1) = \theta_1^*$  y calcular  $V(\theta_1) = \{|r_1[LI_m, E^2(S^2)] - r_1[LS_m, E^2(S^2)]|/6\}^2 = \beta_i/(k_i(\alpha_i - 1))$ . Resolver esta última ecuación para  $k_i$ .
5. En caso de una distribución uniforme( $a_i, b_i, d_i, e_i$ ),  $a_i$  y  $b_i$  se asignan al valor más pequeño y al más grande, respectivamente, de  $r_1[LI_m, E^2(S^2)]$  y  $r_1[LS_m, E^2(S^2)]$ . De la misma manera,  $d_i$  y  $e_i$  se asignan al valor más pequeño y al más grande, respectivamente, de  $r_2[E(M), LS_s^2]$  y  $r_2[E(M), LI_s^2]$ .

## 5.10 Evaluación de la metodología propuesta

Una forma de evaluar el método propuesto para la elección de los hiperparámetros de las *a priori* normal-gama( $\alpha_i, \beta_i, m_i, k_i$ ) y de la uniforme( $a_i, b_i, d_i, e_i$ ), es obteniendo la densidad *a priori* predictiva para cada uno de los modelos y compararlas. Si estas densidades son similares será un indicativo que la información inicial evaluada en términos de la densidad *a priori* predictiva es similar para todos los modelos. También se podrá comparar la media y desviación estándar de esta densidad con los intervalos iniciales para la media y la desviación estándar. Obviamente se espera que la media y desviación estándar de esta densidad estén dentro de los intervalos correspondientes.

De acuerdo a lo anterior necesitamos generar la densidad *a priori* predictiva para cada modelo y cada tipo de *a priori*. es decir necesitamos obtener (ver 2.6):

$$f(x) = \int_{\Theta} f(x|\theta)\pi(\theta)d\theta \quad (5.23)$$

Como habíamos comentado antes cuando se utiliza una distribución normal-gama como *a priori*,  $f(x)$  tiene forma analítica sólo para el caso exponencial, normal y lognormal. En cambio cuando se utiliza *a priori* uniforme no tiene forma analítica para ningún de los modelos. En los casos que  $f(x)$  no tenga forma analítica la obtendremos por simulación. Para ello de acuerdo a Smith y Roberts (1993), una muestra  $x_1, \dots, x_n$  de  $f(x)$  se obtiene a través de una muestra de  $f(x|\theta^{(i)})$ , con  $\theta^{(i)}$  ( $i = 1, \dots, n$ ) muestra generada a partir de la *a priori*  $\pi(\theta)$ . Por lo anterior, el algoritmo para generar por simulación una muestra de  $f(x)$  es el siguiente.

1. Dar intervalos iniciales para la media y la desviación estándar del tiempo de vida. Estos intervalos se pueden generar a través de los intervalos *a priori* para los cuantiles de la distribución del tiempo de vida, tal como se explicó en la segunda sección de este capítulo.
2. A partir de los intervalos anteriores calcular los parámetros de la distribuciones *a priori*, ya sea *a priori* normal-gama( $\alpha_i, \beta_i, m_i, k_i$ ) o uniforme( $a_i, b_i, d_i, e_i$ ).
3. Con lo hiperparámetros definidos obtener  $\theta^{(i)}$  de la *a priori* (esto se hace aplicando el algoritmo B1 y B2 del apéndice).
4. Generar una muestra del modelo  $f(x|\theta^{(i)})$ . Esto se puede hacer con las rutinas tradicionales de S-plus, dados los modelos que hemos considerado en este trabajo.

A continuación presentaremos tres ejemplos de donde se aplicará el algoritmo anterior. En particular en el paso 1 del algoritmo partiremos directamente de intervalos para la media y la desviación estándar, que finalmente es un procedimiento único para todos los modelos.

Ejemplo 5.1. En este ejemplo utilizamos magnitudes pequeñas. Supongamos que por alguno de los dos métodos descritos en la sección 5.2 se llegó a que los intervalos iniciales para la media y la desviación estándar del tiempo de vida, son  $[7, 10]$  y  $[1, 2.5]$ , respectivamente. Aplicando el algoritmo descrito antes se generó una muestra de tamaño 15000

de las densidades *a priori* predictiva para cada modelo, y se utilizó el comando *density* de *S-Plus* para estimar en forma no paramétrica las densidades correspondientes. En la figura 5.1 se muestran las densidad *a priori* predictivas utilizando los dos tipos de *a priori*.

A partir de estas gráficas se puede apreciar que exceptuando el modelo exponencial, el resto de las densidades *a priori* predictivas son relativamente similares para cada tipo de *a priori*. La razón de la discrepancia del modelo exponencial se puede justificar a partir del hecho que esta sólo tiene un parámetro, con el cual sólo se controla la media. En este sentido su media es relativamente similar al resto. En el caso de los otros modelos, como éstos tienen dos parámetros, esto ayuda a controlar tanto la media como la dispersión. Así de la figura 5.1 se aprecia que para los otros modelos se nota una buena similitud entre las densidades, tanto en el caso de la *a priori* normal-gama como uniforme. Con los datos simulados de las densidades *a priori* predictivas también se calculó la media y la desviación estándar de estos datos (ver tabla 5.3). De ahí se aprecia que la media en los diferentes casos tendió a ubicarse al centro del intervalo inicial correspondiente, [7, 10], lo cual son buenas noticias. En el caso de la dispersión, además del modelo exponencial que ya se había comentado, los valores de la desviación estándar también tiende a ubicarse al centro del intervalo inicial, [1, 2.5]. Por lo tanto, para este ejemplo, el método de definir los hiperparámetros para los dos tipos de *a priori* para los diferentes modelos funcionó adecuadamente, ya que está garantizando una información inicial similar para los diferentes modelos.

Modelo	A priori normal-gama		A priori uniforme	
	Media	D. Estándar	Media	D. Estándar
Normal	8.50	1.85	8.53	1.74
Lognormal	8.53	1.89	8.35	1.69
V. Extremo	8.47	1.84	8.60	1.82
Weibull	8.47	1.98	8.59	1.82
Exponencial	8.50	6.90	8.38	6.74

Tabla 5.4. Estadísticos para las densidades *a priori* predictivas para el ejemplo 5.1.

**Ejemplo 5.2.** Supongamos que por alguno de los dos métodos descritos en la sección 5.2 se llegó a que los intervalos iniciales para la media y la desviación estándar del tiempo de vida, son [50, 75] y [5, 15], respectivamente. Se aplicó el algoritmo descrito antes para generar una muestra de tamaño 15000 de las densidades *a priori* predictiva para cada modelo, y se utilizó el comando *density* de *S-Plus* para estimar en forma no paramétrica las densidades correspondientes. En la figura 5.2 se muestran las densidad *a priori* predictivas obtenidas.

A partir de estas gráficas se puede apreciar que exceptuando el modelo exponencial, que ya se había comentado en el ejemplo anterior, el resto de las densidades *a priori* predictivas son relativamente similares para cada tipo de *a priori*. En la tabla 5.4 se aprecia la media y la desviación estándar de las densidades *a priori* predictivas. De donde se aprecia en cuanto tendencia central una similitud en todos los casos, incluyendo los dos tipos de *a priori*, ya que la media está prácticamente al centro del

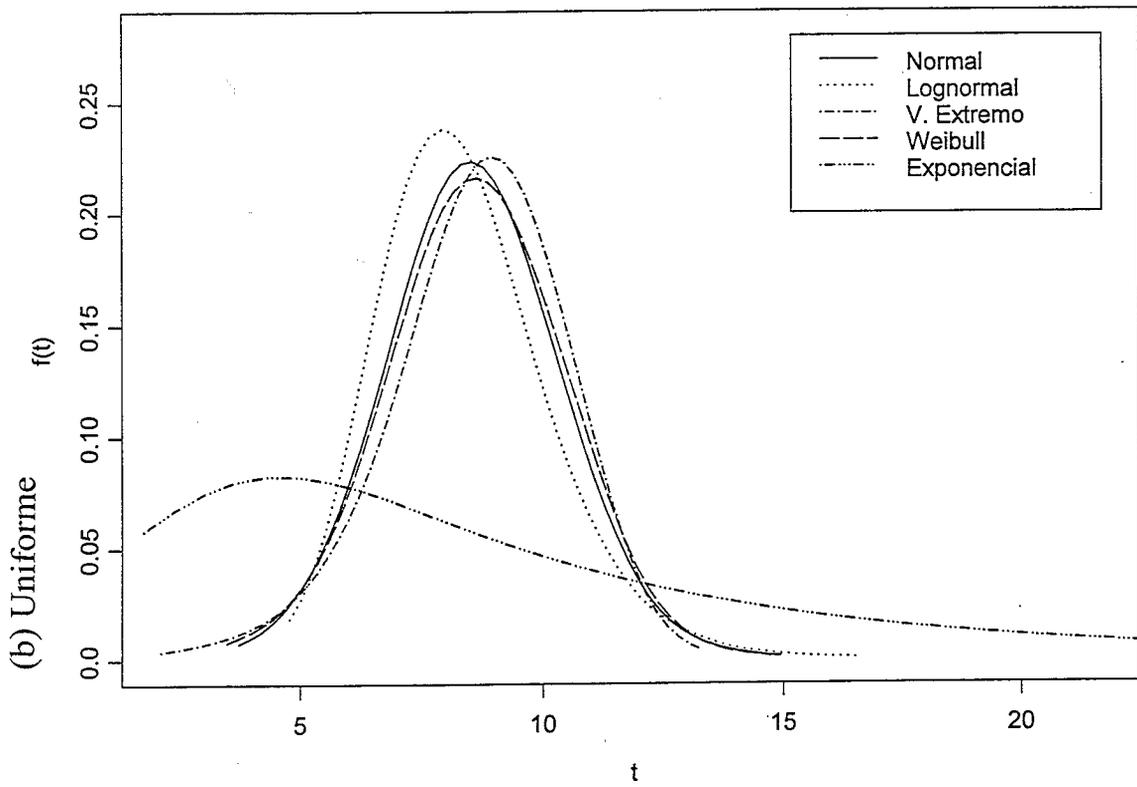
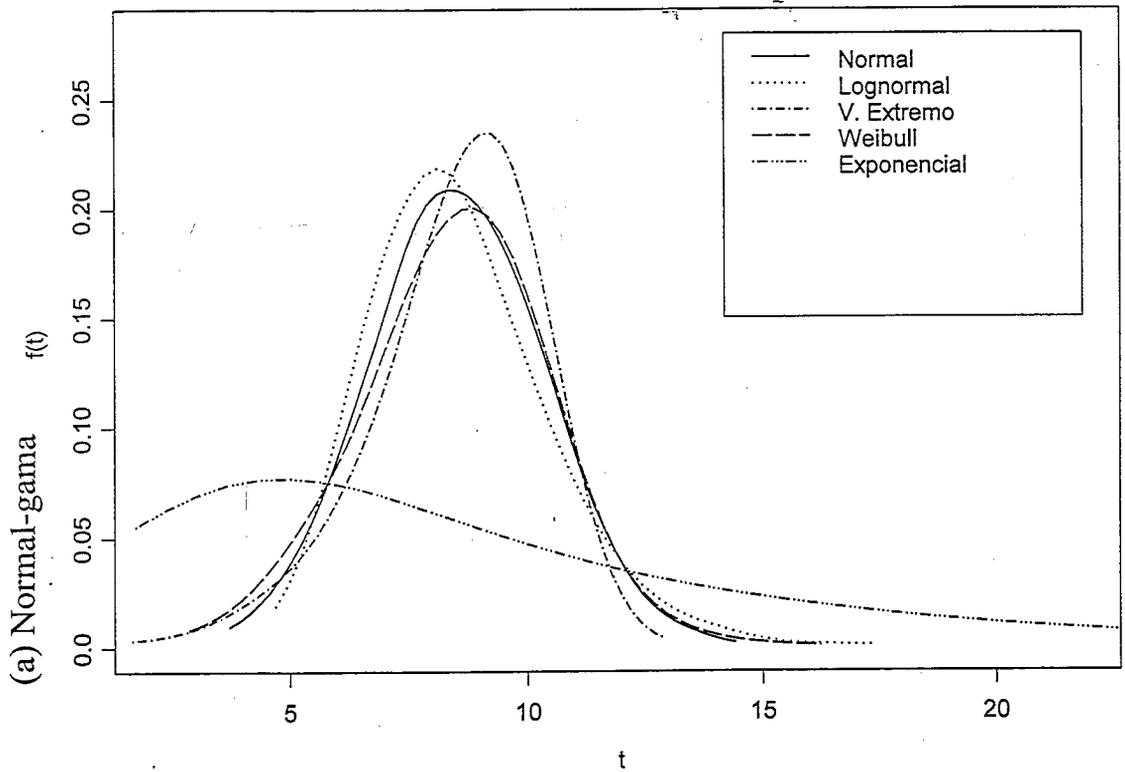


Figura 5.1: Densidades *a priori* predictivas para el ejemplo 5.1.

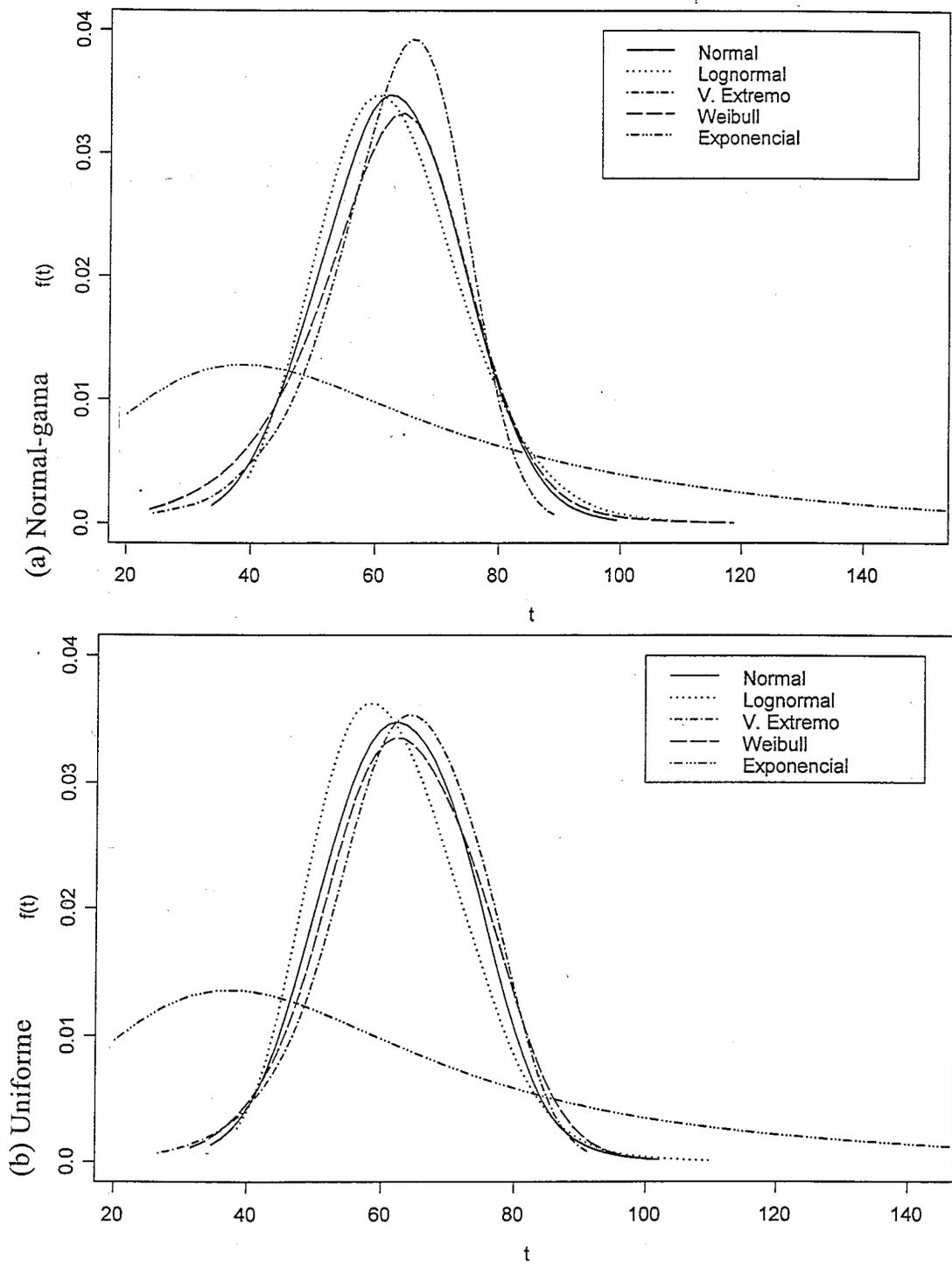


Figura 5.2: Densidades *a priori* predictivas para el ejemplo 5.2.

intervalo inicial para la media, [50, 75]. En el caso de la dispersión, los valores de la desviación estándar también tienden a ubicarse al centro del correspondiente intervalo inicial, [5, 15]. Por lo tanto, en este ejemplo, el método de definir los hiperparámetros para los dos tipos de *a priori* para los diferentes modelos funcionó adecuadamente.

Modelo	A priori normal-gama		A priori uniforme	
	Media	D. Estándar	Media	D. Estándar
Normal	62.55	11.03	62.52	10.66
Lognormal	62.87	11.35	61.36	10.66
V. Extremo	62.60	10.86	63.34	11.08
Weibull	62.20	12.56	63.26	11.37
Exponencial	63.31	43.41	59.95	41.27

Tabla 5.4. Estadísticos para las densidades *a priori* predictivas para el ejemplo 5.2.

**Ejemplo 5.3.** En el ejemplo 3.2 presentamos los datos de tiempo de vida de amortiguadores, y en el siguiente capítulo analizaremos nuevamente estos datos para especificar su distribución. Ahí veremos que [20000, 35000] y [5000, 15000] son los intervalos *a priori* para la media y la desviación estándar del tiempo de vida. Se generó una muestra de tamaño 15000 de la densidad *a priori* predictiva de los diferentes modelos, y se utilizó el comando *density* de *S-Plus* para estimar en forma no paramétrica la densidad correspondiente. En la figura 5.3 se muestran las densidades obtenidas.

A partir de estas gráficas se puede apreciar que exceptuando el modelo exponencial, el resto de las densidades *a priori* predictivas son relativamente similares para cada tipo de *a priori*. Siendo más similares cuando se utiliza una distribución uniforme como *a priori*. En la tabla 5.5 se muestra la media y la desviación estándar de tales densidades. De donde se aprecia en cuanto tendencia central una similitud en todos los casos, incluyendo los dos tipos de *a priori*. En particular las medias caen casi al centro del intervalo inicial dado para la media, [20000, 35000]. En el caso de la desviación estándar también tiende a ubicarse en el centro del correspondiente intervalo inicial, [5000, 15000]. Resalta un poco una mayor dispersión en el caso del modelo Weibull comparativamente con el resto, cuando se utiliza una distribución normal-gama como *a priori*. Por lo anterior, para este ejemplo, podemos concluir que la metodología propuesta para definir los hiperparámetros para los dos tipos de *a priori* para los diferentes modelos, funcionó adecuadamente.

Modelo	A priori normal-gama		A priori uniforme	
	Media	D. Estándar	Media	D. Estándar
Normal	27525.8	10468.3	27360.23	9007.3
Lognormal	27695.26	10602.9	26124.54	8759.8
V. Extremo	27395.13	10435.15	28239.59	9719.9
Weibull	27810.9	13104.3	27892.13	9710.6
Exponencial	27545.07	27804.54	26777.14	27469.25

Tabla 5.5. Estadísticos para las densidades *a priori* predictivas para el ejemplo 5.3.

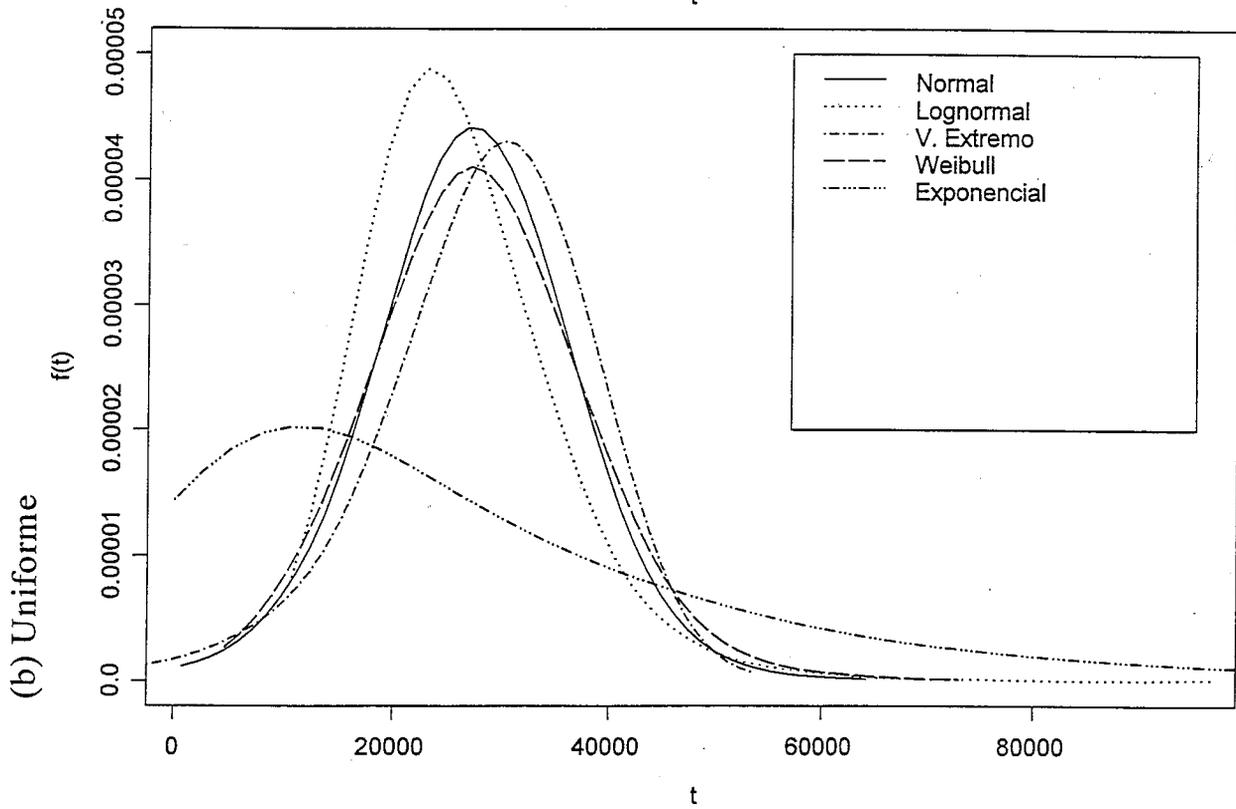
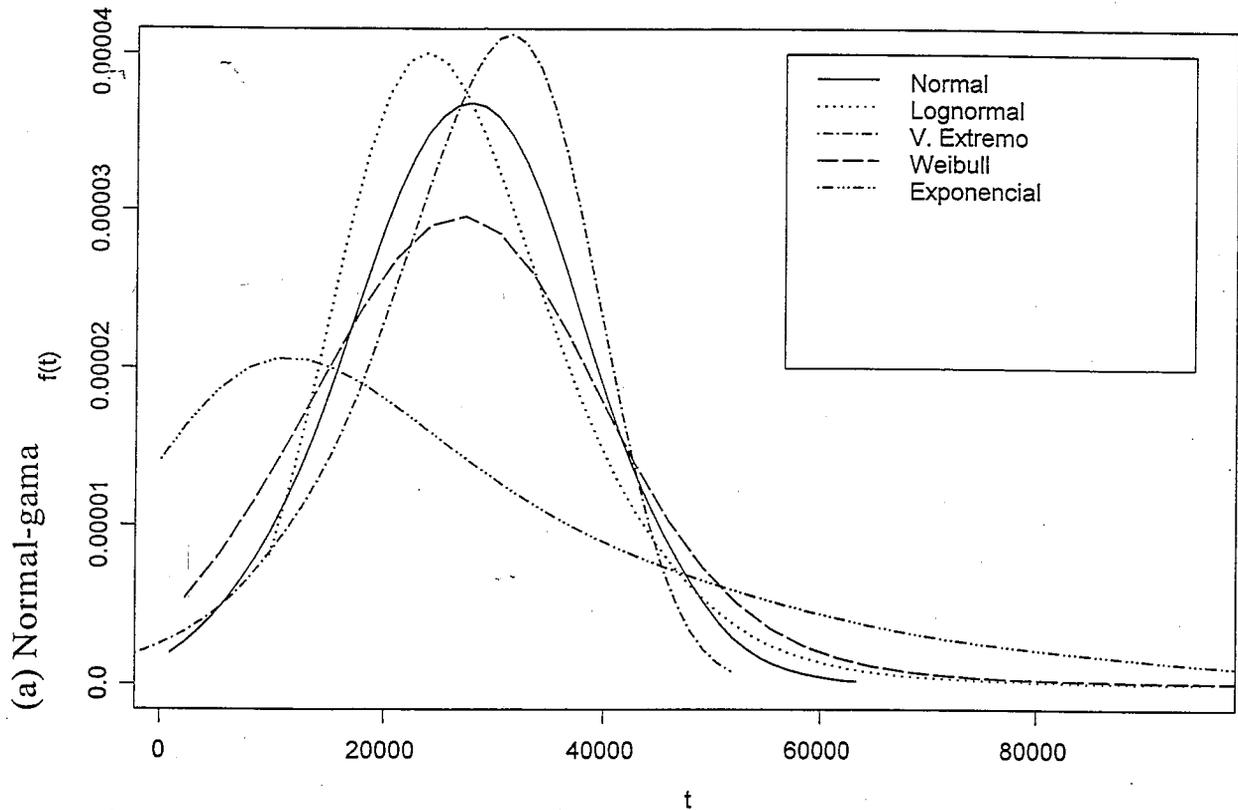


Figura 5.3: Densidades *a priori* predictivas para el ejemplo 5.3.

Conclusión. Hemos propuesto una metodología para que la información inicial del usuario en términos de intervalos para los cuantiles del tiempo de vida del producto, permitan definir intervalos para la media y la desviación estándar del tiempo de vida. Y esto últimos intervalos sirvieron de base para definir los parámetros de las distribuciones *a priori* de los parámetros de los modelos que se han considerado en este trabajo. Considerando dos tipos de *a priori*, en esta sección hemos evaluado esta metodología. Para ello comparamos las distribuciones *a priori* predictivas. Los ejemplos anteriores nos han mostrado que el método está incorporando de una manera consistente y similar para los diferentes modelos, el conocimiento inicial.

## CAPITULO 6

### APLICACIONES DE SELECCIÓN DE MODELOS

En este capítulo se aplica la metodología de selección de modelos que se ha propuesto en los capítulos anteriores, se empleara sobre datos que se han reportado en la literatura y datos simulados.

#### 6.1 Tabla de probabilidades posteriores

Vamos a aplicar la metodología propuesta para evaluar cuál de los cinco modelos (normal- $M_1$ , lognormal- $M_2$ , valor extremo- $M_3$ , Weibull- $M_4$  y exponencial- $M_5$ ) tiene mayores méritos como modelo de distintos conjuntos de datos. Para ello, de acuerdo al capítulo 5, necesitamos proponer intervalos iniciales donde se espera ocurran ciertos cuantiles, para al aplicar la metodología descrita en la sección 5.2 generar intervalos donde se espera que ocurran la media y la desviación estándar del tiempo de vida. Obviamente lo deseable hubiese sido proponer estos intervalos de forma previa a obtener los datos, sin embargo, como lo datos con los que trabajaremos ya existe, vamos a proponer directamente los intervalos para la media y la desviación estándar, ya que estos son comunes a todos los modelos.

A partir de los intervalos para la media y la desviación estándar calcularemos los hiperparámetros de las distribuciones *a priori* para los parámetros de los diferentes modelos considerados, aplicando para ello los resultados resumidos en las tablas 5.1 y 5.2. Utilizaremos los dos tipos de distribuciones *a priori* que se han detallado en los capítulos previos. La comparación de modelos se hace de forma general, calculando la probabilidad posterior para cada modelo al tomar en cuenta todos los demás (3.4). Es decir, calculando

$$\Pr(M_j|\mathbf{X}) = \frac{f(\mathbf{X}|M_j) \Pr(M_j)}{\sum_{i=1}^5 f(\mathbf{X}|M_i) \Pr(M_i)} \quad j = 1, 2, \dots, 5 \quad (6.1)$$

Como ya lo habíamos comentado en la sección 3.3,  $\Pr(M_j|\mathbf{X})$  proporciona una evidencia cuantitativa de qué tan adecuado es cada modelo al ser comparado contra el resto de los modelos considerados. Nótese que  $\sum_{j=1}^5 \Pr(M_j|\mathbf{X}) = 1$ , por lo que los modelos que tengan una mayor probabilidad posterior serán los modelos, de entre los considerados, con mayores méritos para ser considerados como modelos para los datos correspondientes. La probabilidad inicial para cada modelo será la misma ( $\Pr(M_j) = 0.5$ ).

Además con la idea de tener mayores elementos de evaluación de los méritos de cada modelo en relación otro, es posible comparar de dos en dos los modelos. Esto en un momento dado permitirá tener argumentos específicos en favor en contra de un modelo, al ser referido a otro de manera específica. Así, además de la evaluación general proporcionada por  $\Pr(M_j|\mathbf{X})$ , proponemos, a partir de la verosimilitud integrada de cada modelo  $f(\mathbf{X} | M_i)$  construir la tabla de probabilidades pareadas, en la que se comparan en pares los modelos considerados. Para ello si se desea comparar el modelo  $M_i$  con el modelo  $M_j$ , y se considera sólo estos dos modelos como las posibles opciones, entonces de acuerdo a (3.4) la probabilidad posterior del modelo  $M_i$  estará dada por

$$\Pr(M_{i,j}|\mathbf{X}) = \frac{f(\mathbf{X} | M_i) \Pr(M_i)}{f(\mathbf{X} | M_j) \Pr(M_j) + f(\mathbf{X} | M_i) \Pr(M_i)} \quad \text{con } i \neq j. \quad (6.2)$$

Esta expresión se obtiene directamente de (3.4), haciendo  $\Pr(M_i) = 0.5$ . Al ordenar estas probabilidades se obtiene una tabla como la 6.2. De aquí que en la medida que  $\Pr(M_{i,j} | \mathbf{X})$  sea mayor que 0.5, el modelo  $M_i$  será más adecuado que el modelo  $M_j$ . De esta manera, si en total se evalúan  $m$  modelos y se comparan por pares mediante el cálculo de (6.2); estas probabilidades pueden organizarse en una tabla con  $m$  renglones y  $m$  columnas. Si los renglones de esta tabla los denotamos con el subíndice  $i$  y las columnas con el subíndice  $j$ , entonces la interpretación de esta tabla será como sigue: en el renglón  $i$  se observaran las probabilidades posteriores del modelo  $M_i$  al ser comparado contra el resto de los modelos que se indican en cada columna; por lo tanto si estas probabilidades son grandes y mayores que 0.5, entonces de acuerdo a la metodología este modelo resulta más adecuado para los datos. Por el contrario, si todas o casi todas las probabilidades del renglón  $i$  son pequeñas y menores que 0.5, entonces el modelo  $M_i$  no es un buen candidato a ser especificado como el modelo para el tiempo de falla. La interpretación de las columnas es opuesta: valores bajos de las probabilidades en la columna de una distribución indican que esa distribución es una buena alternativa como modelo. Esto se debe a que

$$\Pr(M_{i,j} | \mathbf{X}) + \Pr(M_{j,i} | \mathbf{X}) = 1$$

En los conjuntos de datos que se consideran enseguida se muestran ejemplos de esta tabla de probabilidades posteriores.

## 6.2 Experimento sobre amortiguadores

En el ejemplo 3.2 se presentaron los datos sobre el tiempo de vida (en kilómetros) de amortiguadores de automóviles. Analizando los datos, es claro que estamos ante el caso de censura múltiple tipo I o quizás una censura aleatoria. En la figura 3.2 se presentan las gráficas de probabilidad para tales datos. A partir de estas gráficas es difícil la elección de la distribución, excepto probablemente la valor extremo, que se puede descartar ya que un punto se sale de las correspondientes bandas de confianza.

Vamos a aplicar a estos datos la metodología que hemos descrito para evaluar cuál de los cinco modelos (normal- $M_1$ , lognormal- $M_2$ , valor extremo- $M_3$ , Weibull- $M_4$  y exponencial- $M_5$ ) es más adecuado. Para ello, de acuerdo al capítulo 5, necesitamos proponer intervalos iniciales donde esperamos que ocurra la media y la desviación estándar del tiempo de vida. Obviamente lo deseable hubiese sido proponer estos intervalos de forma previa a obtener los datos, sin embargo, como los datos ya existían, vamos a proponerlos a partir de observar los datos y con suficiente amplitud. Como intervalos de trabajo para la media y la desviación estándar proponemos a [20000, 35000] y [5000, 15000], respectivamente. Más adelante se verá que en este caso esta elección no influye demasiado en las conclusiones.

A priori normal-gama. Dados los intervalos *a priori* para la media y la desviación estándar, utilizando una distribución normal-gama como *a priori* para los parámetros de los cinco modelos en competencia. Entonces, aplicando los resultados del capítulo 5 sintetizados en la tabla 5.1, se obtienen los valores de los hiperparámetros de la distribución *a priori* que se muestran en la tabla 6.1.

Modelo	$\beta_i$	$\alpha_i$	$m_i$	$k_i$
Normal	$2916 \times 10^6$	29.16	27500	16.568
Lognormal	4.083	32.88	10.1598	11.367
V. extremo	631554	81	32000	0.00126
Weibull	2.2932	6.8757	30796.57	$5.2046 \times 10^{-8}$
Exponencial	2850909	103.67		

Tabla 6.1: Parámetros de la *a priori* normal-gama para datos de amortiguadores.

Con estos parámetros y calculando la verosimilitud integrada por simulación Monte Carlo (20000 replicas), se obtienen los resultados de la tabla 6.2. Analizando la última columna de esta tabla se aprecia que los modelos con mayores méritos son el normal y el Weibull ( $\Pr(M_1|\mathbf{X}) = 0.42$  y  $\Pr(M_4|\mathbf{X}) = 0.41$ ), seguidos de forma lejana por el valor extremo. En Meeker y Escobar (1998, pag. 176) se consideran estos mismos datos, y con base a gráficas de probabilidad con bandas de confianza deciden el modelo Weibull sobre el lognormal, pero no consideran el modelo normal. De la misma tabla 6.2, se aprecia que para estos datos el modelo Weibull es muy superior al lognormal, ya que  $\Pr(M_{4,2}|\mathbf{X}) = 0.9061$ , y  $\Pr(M_{2,4}|\mathbf{X}) = 0.0939$ .

Modelo	(a priori normal-gama)					$\Pr(M_j \mathbf{X})$
	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	
Normal, $M_1$		0.9077	0.7717	0.5048	0.9999	0.4214
Lognormal, $M_2$	0.0923		0.2558	0.0939	0.9992	0.0416
V. extremo, $M_3$	0.2283	0.7442		0.2317	0.9997	0.1234
Weibull, $M_4$	0.4952	0.9061	0.7683		0.9999	0.4136
Exponencial, $M_5$	0.0001	0.0008	0.0003	0.0001		0.0000

Tabla 6.2. Probabilidad posterior para los cinco modelos, datos de amortiguadores.

Para ver qué tanto influye la elección de los intervalos *a priori* para la media y la desviación estándar sobre las probabilidades posteriores de la tabla 6.2, se hicieron cambios a estos intervalos y se observaron los resultados que a continuación comentamos (los cálculos se hicieron tanto con Monte Carlo como con Laplace). Algunos cambios que fueron estudiados en la media fueron: [20000, 40000], [15000, 40000], [25000, 40000], [25000, 35000], [25000, 30000], [15000, 30000], [10000, 30000], en estos casos en general se mantuvo la elección de Weibull, seguida por normal. Respecto a la desviación estándar, algunos cambios estudiados fueron: [3000, 10000], [5000, 10000] con elección similar a la que se da a partir de la tabla 6.2 (Weibull seguida de normal como tendencia). Otros cambios evaluados en la desviación estándar fueron: [5000, 20000], [5000, 25000], con los cuales, el modelo Weibull sigue estando en los primeros planos, pero el modelo normal pasa a un segundo plano, apareciendo en su lugar el modelo lognormal. Con este análisis podemos ver aunque los intervalos iniciales influyen en la elección, en este caso, esta influencia fue relativamente poca, dándose principalmente en el intervalo para la desviación estándar. Entonces para el ejemplo, el método tiene una aceptable robustez en cuanto a la elección de los intervalos iniciales. Nótese que los intervalos evaluados no son radicalmente diferentes entre sí, pero en la medida que si lo sean pueden influir más en la decisión, como sería lógico esperarlo.

A priori uniforme. Aplicando los resultados del capítulo 5, sintetizados en la tabla 5.2, se obtienen los valores de los hiperparámetros de la distribución *a priori* para cada modelo que se muestra en la tabla 6.3.

Modelo	Parámetros de la distribución uniforme( $a_i, b_i, d_i, e_i$ )			
	$a_i$	$b_i$	$d_i$	$e_i$
Normal	20000	35000	$4.4444 \times 10^{-9}$	$4 \times 10^{-8}$
Lognormal	9.8414	10.4010	3.8394	30.7473
V. extremo	24500.4	39500.4	$8.55 \times 10^{-5}$	0.0002565
Weibull	22397.51	39195.64	1.9077	6.4286
Exponencial	$2.8571 \times 10^{-5}$	$5 \times 10^{-5}$		

Tabla 6.3: *A priori* uniforme, datos de amortiguadores.

Con estos parámetros iniciales y calculando la verosimilitud integrada por simulación Monte Carlo para cada modelo (con  $M=20000$  replicas), se obtiene la tabla 6.4 de probabilidades posteriores. De la última columna de esta tabla se aprecia que el modelo Weibull es el que tiene una probabilidad posterior mayor. Lo siguen el modelo normal y valor extremo, en ese orden. Nuevamente al comparar el modelo Weibull contra el lognormal, se aprecia que el primero a parece con más méritos:  $\Pr(M_{4,2}|\mathbf{X}) = 0.8972$ .

Al comparar las tablas 6.2 y 6.4, donde se muestra los dos tipos de *a priori*, se aprecia que en general se llega a las mismas conclusiones, pero cuando se utiliza la distribución normal-gama la discriminación parece ligeramente más clara. Así, dados los resultados de las tablas 6.2 y 6.4, y con base en los cinco modelo considerados, la primera elección del modelo para el tiempo de falla de los amortiguadores sería el Weibull, seguida en forma cercana por la normal.

Modelo	<i>A priori</i> uniforme					$\Pr(M_j \mathbf{X})$
	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	
Normal, $M_1$		0.8453	0.5526	0.3848	0.9998	0.2762
Lognormal, $M_2$	0.1547		0.1844	0.1028	0.9988	0.0461
V. extremo, $M_3$	0.4474	0.8156		0.3362	0.9997	0.2307
Weibull, $M_4$	0.6152	0.8972	0.6638		0.9999	0.4454
Exponencial, $M_5$	0.0002	0.0012	0.0003	0.0001		0.0001

Tabla 6.4. Probabilidad posterior para los cinco modelos, datos de amortiguadores.

Para ver qué tanto influye la elección de los intervalos *a priori* para la media y la desviación estándar, cambiamos estos intervalos y vimos los resultados que a continuación comentamos (los cálculos se hicieron con el método Monte Carlo). Algunos cambios que fueron estudiados en la media fueron: [20000, 40000], [15000, 40000], [25000, 40000], [25000, 35000], [25000, 30000], [15000, 30000], [10000, 30000], [10000, 25000] y [10000, 20000]; en estos casos en general se mantuvo la elección de Weibull, seguida por normal, observando una mejor estabilidad de las probabilidades posteriores al ser comparado con el mismo análisis para la *a priori* normal-gama. Respecto a la desviación estándar, algunos cambios estudiados fueron: [3000, 10000], [5000, 10000], [5000, 20000] y [5000, 25000] con elección similar a la que se da a partir de la tabla 6.4 (Weibull seguida

de normal). De esta manera se observa en este caso respecto a la elección del intervalo para la desviación estándar hay mejor estabilidad si comparamos con lo que pasó en el caso de la *a priori* normal-gama. Con este análisis podemos ver aunque los intervalos iniciales influyen en la elección, en este caso, esta influencia fue relativamente poca.

### 6.3 Experimento sobre aire acondicionado

En Proschan(1963) se presentan tiempos de falla del equipo de aire acondicionado de un aeroplano, sugiriendo que éstos siguen una distribución exponencial. Estos datos no censurados han recibido mucha atención en la literatura; ver por ejemplo Berger y Pericchi (1996) y Gutiérrez-Peña y Walker (2001). Los tiempos de falla se presentan a continuación

1 3 5 7 11 11 11 12 14 14 14 16 16 20 21 23 42 47  
52 62 71 71 87 90 95 120 120 225 246 261

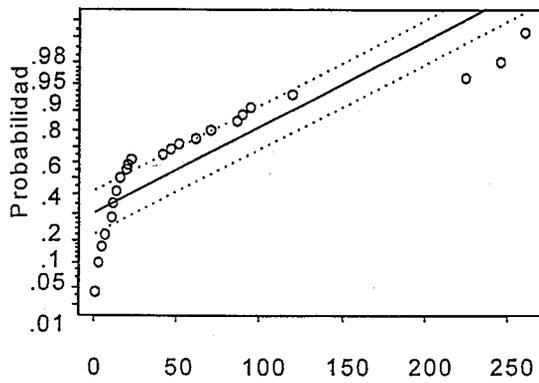
En la figura 6.1 se aprecia la gráfica de probabilidad para estos tiempos de falla para diferentes distribuciones. De aquí se aprecia que la distribución valor extremo estaría descartada, y decidirse entre las distribuciones exponencial, lognormal o Weibull realmente sería difícil. Vamos a ver cómo refleja esto la tabla de probabilidades posteriores. Estos tiempos de falla no están censurados, por lo que para obtener la verosimilitud integrada es posible aplicar las expresiones analíticas que se obtuvieron en el capítulo 4 para los caso normal, lognormal y exponencial. Se utiliza como intervalos de trabajo *a priori* para la media y la desviación estándar a [50, 70] y [50, 90], respectivamente. Más adelante se ve que esta elección no influye demasiado.

A(j normal-gama. Utilizando una distribución normal-gama como *a priori*, y dados los intervalos para la media y la desviación estándar, se obtienen los parámetros de la distribución *a priori* para cada modelo que se muestran en la tabla 6.5.

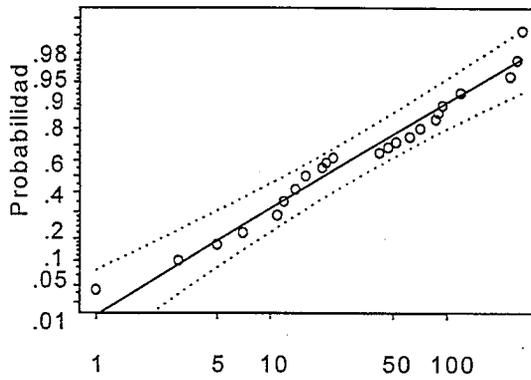
Modelo	$\beta_i$	$\alpha_i$	$m_i$	$k_i$
Normal	282558.7	57.66	60	448.78
Lognormal	105.2	122.47	3.665	234.52
V. extremo	9947	182.25	91.5	4.939
Weibull	64.98	55.9	55.557	0.04967
Exponencial	18375	306.25		

Tabla 6.5: Parámetros de la *a priori* normal-gama para datos de aire acondicionado.

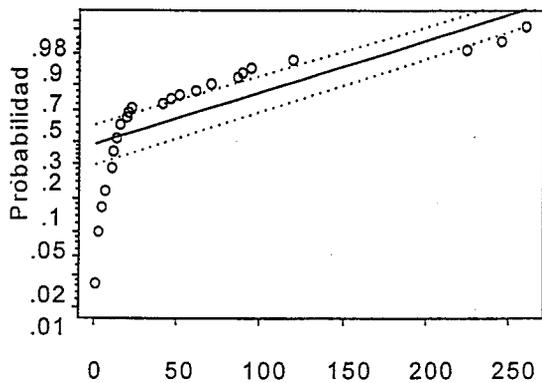
Con los parámetros iniciales se calcula la verosimilitud integrada con la aproximación de Laplace para los modelos Weibull y valor extremos, y para los otros modelos de manera exacta. En la tabla 6.6 se muestran las probabilidades posteriores. De donde se ve que se debe descartar definitivamente a la distribución valor extremo y normal como modelos. También la distribución lognormal tiene poca viabilidad frente a los modelos Weibull y exponencial. Y entre estos últimos la metodología da una ligera preferencia por el modelo Weibull. Esta preferencia se podría acentuar o incluso invertir si en la probabilidad inicial por estos dos modelos se declara una preferencia por



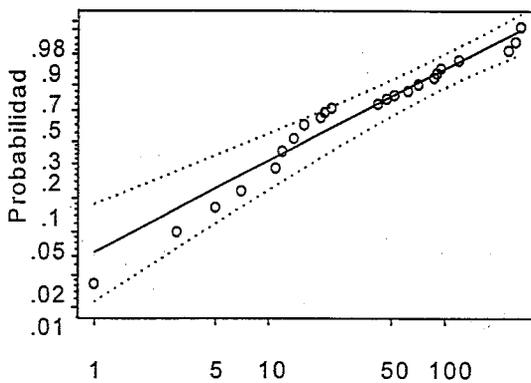
(a) Gráfica normal



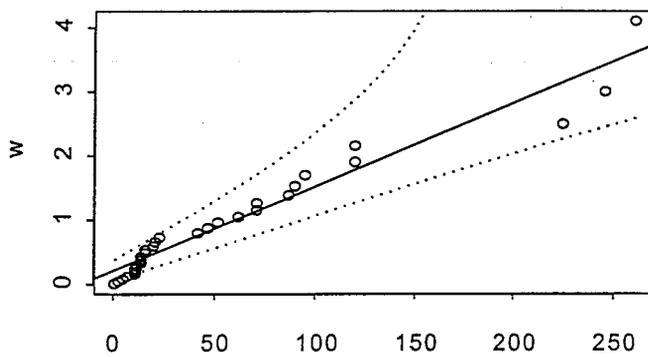
(b) Gráfica lognormal



(c) Gráfica valor extremo



(d) Gráfica Weibull



(e) Gráfica exponencial

Figura 6.1: Gráficas de probabilidad para los datos de equipos de aire acondicionado.

un modelo más sencillo (exponencial) o más flexible (Weibull), según se prefiera. Por ejemplo en el caso que nos ocupa, si la probabilidad inicial para exponencial fuera de 0.7 por 0.3 el Weibull, en la probabilidad posterior pareada  $\Pr(M_{5,4} | \mathbf{X}) = 0.63$  en favor de exponencial.

En otros trabajos se han analizado estos datos. Por ejemplo en Gutiérrez-Peña y Walker (2001), aplicando otro procedimiento Bayesiano, se llega a que las distribuciones Weibull y exponencial son equivalentes. En cambio en Berger y Pericchi (1996), aplicando otro enfoque Bayesiano, el orden de preferencia es exponencial, lognormal y Weibull. Este último trabajo, en relación a sus conclusiones sobre el modelo lognormal, está en clara contradicción tanto con nuestros resultados como los de Gutiérrez-Peña y Walker (2001).

Modelo	A priori normal-gama					$\Pr(M_j   \mathbf{X})$
	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	
$M_1$		0.0000	1	0	0	0
$M_2$	1		1	0.0093	0.0126	0.0054
$M_3$	0.0000	0.0000		0.0000	0.0000	0
$M_4$	1	0.9907	1		0.5778	0.5740
$M_5$	1	0.9874	1	0.4222		0.4206

Tabla 6.6. Probabilidades posteriores, datos de aire acondicionado.

Para ver qué tanto influye la elección de los intervalos *a priori* para la media y la desviación estándar, hicimos cambios a estos intervalos y vimos los resultados que a continuación comentamos (los cálculos se hicieron de manera exacta para los modelos normal, lognormal y exponencial y con Monte Carlo para los restantes). Algunos cambios que fueron estudiados en la media fueron: [50, 60], [50, 80], [50, 90], [40, 70], [30, 70], [40, 60]; en todos estos casos los dos modelos con probabilidades posteriores más altas y relativamente similares, son el exponencial y Weibull; y en un tercer término muy lejano el lognormal. Es decir se mantiene la elección reportada en la tabla 6.6. Respecto a la desviación estándar, algunos cambios estudiados fueron: [50, 70], [50, 80], [50, 100], [40, 90], [40, 70], [30, 60]; ocurre una elección similar a la que se da a partir de la tabla 6.6 (Weibull y exponencial con probabilidades similares; y en tercer término en forma lejana lognormal). Con este análisis podemos ver que aunque los intervalos iniciales influyen en las probabilidades posteriores, en este caso, esta influencia prácticamente no afecta la elección del modelo.

A priori uniforme. Aplicando los resultados de la tabla 5.2, se obtienen los valores de los hiperparámetros de la distribución *a priori* uniforme para cada uno de los modelos que se muestran en la tabla 6.7.

Modelo	$a_i$	$b_i$	$d_i$	$e_i$
Normal	50	70	0.000123	0.0004
Lognormal	3.48246	3.81893	0.848425	1.896256
V. extremo	81.5	101.5	0.01425	0.02565
Weibull	46.2978	64.8170	0.684756	1.2054
Exponencial	0.014286	0.02		

Tabla 6.7: Parámetros de la *a priori* uniforme para datos de aire acondicionado.

Con estos parámetros iniciales se calculó la verosimilitud integrada por simulación Monte Carlo para los modelos  $M_1$  a  $M_4$  (con 20000 replicas) y en forma analítica para el modelo  $M_5$ . En la tabla 6.8 se muestran las probabilidades posteriores, que son muy similares a las expuestas en la tabla 6.6. De la última columna de esta tabla se aprecia que los modelos con mayores probabiidades son Weibull y exponencial, con un empate prácticamente. Descartando el resto de las distribuciones como modelos para estos datos. Esto refuerza la evidencia contra el modelo lognormal, ya que éste frente a Weibull y exponencial, tiene muy baja probabilidad posterior. Por ejemplo si utilizamos las probabilidades pareadas  $\Pr(M_{4,2} | \mathbf{X}) = 0.9738$  y  $\Pr(M_{5,2} | \mathbf{X}) = 0.9721$ ; lo que afirma el error en el que incurre Berger y Pericchi (1996) al poner el modelo lognormal por arriba del Weibull.

Modelo	A priori uniforme					$\Pr(M_j   \mathbf{X})$
	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	
Normal $M_1$		0	1	0	0	0
Lognormal $M_2$	1.0		1	0.0262	0.0279	0.0136
Valor extremo $M_3$	0	0		0	0	0
Weibull $M_4$	1.0	0.9738	1.0		0.5163	0.5090
Exponencial $M_5$	1.0	0.9721	1.0	0.4837		0.4774

Tabla 6.8. Probabilidades pareadas, datos de aire acondicionado.

En suma si tomamos en cuenta los resultados de las tabla 6.6 y 6.8, y con base en los cinco modelo considerados, la elección del modelo para el tiempo de falla del aire acondicionado está entre el Weibull y exponencial.

Para ver qué tanto influye la elección de los intervalos *a priori* para la media y la desviación estándar, hicimos cambios a estos intervalos y vimos los resultados que a continuación comentamos (los cálculos se hicieron con Monte Carlo). Algunos cambios que fueron estudiados en la media fueron: [50, 60], [50, 80], [50, 90], [40, 70], [30, 70], [40, 60]; en todos estos casos los dos primeros modelos, con probabilidades relativamente similares, son el exponencial y Weibull; y en un tercer término muy lejano el lognormal. Respecto a la desviación estándar, algunos cambios estudiados fueron: [50, 70], [50, 80], [50, 100], [40, 90], [40, 70], [30, 60]; ocurre una elección similar a la que se da a partir de la tabla 6.8 (Weibull y exponencial con probabilidades similares; y en tercer término en forma lejana lognormal). Con este análisis podemos ver aunque los intervalos iniciales influyen en las probabilidades posteriores, en este caso, esta influencia fue muy poca respecto a los modelos candidatos a ser elegidos.

#### 6.4 Ejemplos simulados

En el ejemplo 3.1 se presentaron cuatro muestras de datos simulados de una distribución exponencial (ver tabla 3.1). A pesar de ello, de la gráfica de probabilidad normal de la figura 3.1 no se rechaza con claridad la idea de normalidad.

Aunque nos interesa principalmente los modelos exponencial y normal, vamos a evaluar también a los demás, para ver cómo se adecuan a los datos. Observando la magnitud de los datos se propusieron intervalos amplios para la media y la desviación estándar: [0.5, 2.0] y [0.5, 2.0], respectivamente.

Modelo	$\beta_i$	$\alpha_i$	$m_i$	$k_i$
$M_1$	37.87	24.24	1.25	26.08
$M_2$	30.25	43.64	-0.123	7.60
$M_3$	62.37	64	1.81	15.84
$M_4$	3.13	3.13	1.25	12.3
$M_5$	12.8	10.2		

Tabla 6.9: Parámetros de la *a priori* normal-gama para datos del ejemplo 3.1.

A priori normal-gama. Primero aplicando las fórmulas de la tabla 5.1 se obtienen los hiperparámetros de la distribución *a priori* normal-gama para cada modelo que se muestran en la tabla 6.9. Tomando en cuenta éstos y calculando la verosimilitud integrada de manera exacta para los modelos normal, lognormal y exponencial, y de con la aproximación de Laplace para los modelo Weibull y valor extremo, se obtiene la tabla de probabilidades posteriores que se muestra en el cuadro 6.10. De la última columna de esta tabla se aprecia que el modelo exponencial es el que resulta con mayores méritos en todas las muestras, exceptuando la muestra 2, donde el modelo lognormal es le que tiene la mayor probabilidad posterior, aunque no muy superior que la que tiene el modelo exponencial,  $\Pr(M_{2,5} | \mathbf{X}) = 0.5861$ . También se aprecia que el modelo Weibull compite en forma cercana con el modelo exponencial. Respecto al modelo normal, este debe descartarse definitivamente como modelo de estos datos, en contraste con la información que aportó la gráfica de probabilidad de la figura 3.1.

Modelo	Muestra	A priori normal-gama					$\Pr(M_j   \mathbf{X})$
		$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	
Normal $M_1$	1		0	0.9320	0	0	0
	2		0	0.9162	0	0	0
	3		0	0.9231	0	0	0
	4		0.0002	0.8345	0.0001	0	0
Lognormal $M_2$	1	1		1	0.0435	0.0351	0.019
	2	1		1	0.6194	0.5861	0.4309
	3	1		1	0.4785	0.3655	0.2560
	4	0.9998		1	0.3095	0.1852	0.0133
V. extremo $M_3$	1	0.0680	0		0	0.000	0
	2	0.0838	0		0.000	0.0001	0
	3	0.0769	0		0	0	0
	4	0.1655	0		0	0	0
Weibull $M_4$	1	1	0.9564	1		0.4684	0.4595
	2	1	0.3806	1		0.4652	0.2648
	3	1	0.5215	1		0.3857	0.2997
	4	0.9999	0.6905	1		0.3365	0.2839
Exponencial $M_5$	1	1	0.9649	1	0.5316		0.5216
	2	1	0.4139	1	0.5348		0.3043
	3	1	0.6345	1	0.6143		0.4443
	4	1	0.8148	1	0.6635		0.5834

Tabla 6.10. Probabilidades posteriores, datos del ejemplo 3.1.

A(j) uniforme. Los parámetros de la distribución *a priori* se muestran en la tabla 6.11. Con estos y obteniendo la verosimilitud integrada en forma analítica para los casos normal, lognormal y exponencial y por simulación Monte Carlos (20000 replicas) para los modelos Weibull y valor extremo, se calcula la tabla de probabilidades posteriores que se muestra en el cuadro 6.12, donde se observa una clara predominancia del modelo exponencial, seguido por Weibull. En particular en la competencia del modelo normal con el exponencial, la metodología distingue muy bien que los datos proceden de una distribución exponencial. Esto contrasta con lo que se observa en la gráfica de probabilidad de la figura 3.1. Además, analizando los resultados pareados de los modelos exponencial y Weibull, se aprecia que la metodología siempre favorece al exponencial. Este es un hecho relevante, ya que como es bien conocido, la distribución exponencial es un caso particular de la Weibull.

De acuerdo a las tablas 6.10 y 6.12, los resultados de los tipos de *a priori*: normal-gama y uniforme, coinciden en determinar los modelos mejor clasificados (exponencial y Weibull).

Modelo	$a_i$	$b_i$	$d_i$	$e_i$
Normal	0.5	2	0.25	4
Lognormal	0	0.3466	0.7875	6.738
V. extremo	1.06	2.563	0.6413	2.565
Weibull	0.5	2.0	0.6482	2.7
Exponencial	0.5	2.0		

Tabla 6.11: Parámetros de la *a priori* uniforme para datos simulados.

Modelo	Muestra	A priori uniforme					Pr( $M_j X$ )
		$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	
Normal $M_1$	1		0.9999	0.9999	0.0008	0.0003	0.0002
	2		0.9966	0.9996	0.0016	0.0012	0.0007
	3		0.9985	0.9997	0.0011	0.0003	0.0002
	4		0.9964	0.9859	0.0471	0.0130	0.0107
Lognormal $M_2$	1	0.0001		0.4737	0	0	0
	2	0.0034		0.8373	0	0	0
	3	0.0015		0.8914	0	0	0
	4	0.0036		0.1998	0.0002	0	0
V. extremo $M_3$	1	0.0001	0.5263		0	0	0
	2	0.0004	0.1062		0	0	0
	3	0.0003	0.1627		0	0	0
	4	0.0141	0.8002		0.0007	0.0002	0.0001
Weibull $M_4$	1	0.9992	1	1		0.2923	0.2931
	2	0.9984	1	1		0.4435	0.4362
	3	0.9989	1	1		0.1949	0.1947
	4	0.9529	0.9998	0.9993		0.2110	0.2104
Exponencial $M_5$	1	0.9997	1	1	0.7077		0.7067
	2	0.9987	1	1	0.5565		0.5631
	3	0.9997	1	1	0.8051		0.8003
	4	0.9870	1	0.9998	0.7890		0.7788

Tabla 6.12. Probabilidad posterior, datos ejemplo 3.1.

## CAPITULO 7

### ESTUDIO MONTE CARLO PARA EVALUAR ESPECIFICACIÓN DE MODELOS

En este capítulo se presentan los resultados de un estudio Monte Carlo para evaluar la especificación de modelos.

#### 7.1 Introducción

Con la intención de evaluar el impacto de la variación muestral sobre la metodología de selección de modelos que se ha descrito en los capítulos anteriores, se decidió hacer un estudio Monte Carlo. Un trabajo más reducido, pero con el mismo propósito se presenta en Gutiérrez-Pulido y et al. (2002). El estudio que aquí presentamos se realizó de acuerdo a los siguientes puntos:

1. Se propusieron densidades relativamente similares, para valores dados de sus parámetros. Estos modelos se muestran en la figura 7.1, y fueron los modelos de los que se generaron muestras aleatorias, para ver si la metodología es capaz de detectar de qué modelo en particular se ha generado cada muestra..
2. Los intervalos iniciales para la media y la desviación estándar especificados, fueron [560, 570] y [5, 10], respectivamente.
3. Se estableció muestreo aleatorio ( $n=35$ ) y muestreo censurado tipo II, con dos niveles de censuramiento ( $r=25, 15$ ).
4. De cada modelo de la figura 7.1 se generan  $k$  (400) muestras para cada tipo de muestreo (en total  $5 \times k = H$  muestras por cada tipo de muestreo). Para cada muestra,  $\mathbf{X}_h$ , se calculó la verosimilitud integrada para los cinco modelos:

$$f(\mathbf{X}_h | M_i) \quad i = 1, 2, \dots, 5$$

considerando los dos tipos de distribuciones *a priori* que hemos propuesto en el trabajo. El cálculo de  $f(\mathbf{X}_h | M_i)$  se hizo con el método Monte Carlo con 3000 replicas.

5. Suponiendo una probabilidad *a priori* similar para cada modelo,  $\Pr(M_i) = 0.5$ , se calculó la tabla de probabilidades posteriores, tanto totales como pareadas. Es decir, para cada muestra  $\mathbf{X}_h$  se calculó la probabilidad posterior total para cada modelo:

$$\Pr(M_j | \mathbf{X}_h) = \frac{f(\mathbf{X}_h | M_j) \Pr(M_j)}{\sum_{i=1}^5 f(\mathbf{X}_h | M_i) \Pr(M_i)} \quad i, j = 1, 2, \dots, 5; \quad h = 1, \dots, H$$

y también las probabilidades posteriores considerando los modelos por pares:

$$\Pr(M_{i,j} | \mathbf{X}_h) = \frac{f(\mathbf{X}_h | M_i)}{f(\mathbf{X}_h | M_i) + f(\mathbf{X}_h | M_j)} \quad \text{con } i \neq j \text{ y } i, j = 1, 2, \dots, 5; \quad h = 1, \dots, H$$

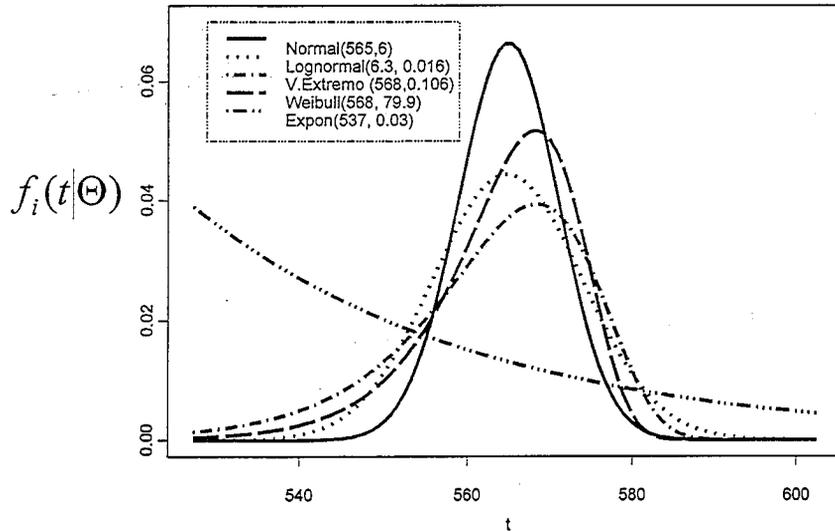


Figura 7.1: Densidades comparadas (en normal y lognormal se indica la desviación estándar).

6. Aunque se analizó el comportamiento de  $\Pr(M_j | \mathbf{X}_h)$  y  $\Pr(M_{i,j} | \mathbf{X}_h)$ . En este trabajo sólo se presenta la media,  $\bar{p}_j$  y  $\bar{p}_{i,j}$  de estas probabilidades. Es decir, sólo reportamos:

$$\bar{p}_j = \frac{1}{H} \sum_{h=1}^H \Pr(M_j | \mathbf{X}_h) \quad \text{y} \quad \bar{p}_{i,j} = \frac{1}{H} \sum_{h=1}^H \Pr(M_{i,j} | \mathbf{X}_h)$$

En seguida presentamos los resultados para los dos tipos de distribuciones *a priori* y para el caso completo y censurado.

## 7.2 A priori normal-gama

En la tabla 7.1 se muestran los resultados del estudio que hemos descrito antes, al usar como *a priori* una distribución normal-gama. Las probabilidades promedio se pueden ver como la probabilidad de una decisión correcta. De un análisis detallado de los resultados de la tabla 7.1 se aprecia que la metodología ve muy similares a los modelos lognormal y normal simulados (como de hecho se puede corroborar en la figura 7.1). Esto pasa también para los modelos Weibull y valor extremo. La prueba tiende a elegir bien el modelo valor extremo al ser comparada contra lognormal y normal;

incluso hay mejor detección conforme aumenta la censura. Algo similar puede decirse de la elección de la normal en relación a Weibull y valor extremo. La discriminación en relación a la exponencial y viceversa, es muy satisfactoria; independientemente de la censura. A lo anterior, hay una excepción que se da en la pobre elección de valor extremo al contrastarla con exponencial, cuando se tiene alto censuramiento.

En los casos que hemos comparado, conforme aumentó el censuramiento, aunque en algunos casos hay una disminución de la probabilidad posterior para llegar a la elección correcta, esta disminución no parece ser fuerte. Incluso se dan casos en los que conforme se disminuyó censuramiento aumentó la probabilidad de una decisión correcta.

### 7.3 A priori uniforme

Utilizando una distribución uniforme como *a priori* para hacer el estudio Monte Carlo que se describió al inicio del presente capítulo, en la tabla 7.2 se muestran los resultados obtenidos. De donde se aprecia que sin censura los resultados son muy similares a cuando se utilizó como *a priori* una normal-gama: es decir los modelos lognormal y normal simulados son muy similares; esto también para los modelos Weibull y valor extremo; además la prueba tiende a elegir bien el modelo valor extremo y Weibull al ser comparadas contra lognormal y normal. Sin embargo, contrariamente al caso normal-gama, conforme aumente el censuramiento la probabilidad de una decisión correcta disminuye, aunque en algunos casos se sigue manteniendo una tendencia satisfactoria a tener una buena decisión. En general podríamos decir que el censuramiento afecta más al poder de discriminación de la metodología cuando se utiliza una *a priori* uniforme.

Datos de:	n=35	(a priori normal-gama)				
	r	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$
Normal $M_1$	35		0.5015	0.6424	0.6526	1.0
	25		0.4920	0.6743	0.6403	1.0
	15		0.4903	0.7363	0.7202	1.0
Lognormal $M_2$	35	0.5011		0.9100	0.8490	1.0
	25	0.4872		0.6148	0.6003	0.9907
	15	0.4889		0.5783	0.5871	0.9909
V.extremo $M_3$	35	0.7217	0.7485		0.1360	0.8606
	25	0.8782	0.8899		0.2948	0.6601
	15	0.9039	0.9139		0.3472	0.4901
Weibull $M_4$	35	0.7615	0.7771	0.5492		1.0
	25	0.7357	0.7478	0.5191		0.9712
	15	0.7251	0.7380	0.5094		0.9320
Exponencial $M_5$	35	1.0	1.0	1.0	1.0	
	25	0.9986	0.9988	0.9981	0.9959	
	15	0.9943	0.9946	0.9898	0.9845	

Tabla 7.1. Probabilidades pareadas promedio,  $\bar{p}_{i,j}$ .

Por otro lado la probabilidad promedio de elección del modelo exponencial se mantiene alta aun en alto censuramiento. Por el contrario en la elección de los mod-

elos lognormal y Weibull frente al modelo exponencial se ve afectado cuando hay alto censuramiento.

Datos de:	n=35	Modelo en competencia ( <i>a priori</i> uniforme)				
	r	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$
Normal ( $M_1$ )	35		0.5002	0.7988	0.7871	1.0
	25		0.4974	0.7415	0.6960	1.0
	15		0.4960	0.6736	0.6683	0.9461
Lognormal ( $M_2$ )	35	0.5022		0.8261	0.8413	1.0
	25	0.4940		0.5759	0.6161	0.9829
	15	0.4904		0.4816	0.4745	0.5961
V.extremo( $M_3$ )	35	0.7870	0.8072		0.5954	0.9906
	25	0.8631	0.8742		0.6419	0.6702
	15	0.6280	0.6675		0.5371	0.9999
Weibull( $M_4$ )	35	0.7687	0.7831	0.4262		1.0
	25	0.7067	0.7188	0.4229		0.9304
	15	0.7645	0.7760	0.4977		0.3946
Exponencial( $M_5$ )	35	1.0	1.0	1.0	1.0	
	25	1.0	1.0	1.0	1.0	
	15	0.9900	0.9910	0.9824	0.9822	

Tabla 7.2. Probabilidad posterior promedio,  $\bar{p}_{i,j}$ .

#### 7.4 Conclusiones del Estudio

El estudio Monte Carlo nos permite sacar las siguientes conclusiones

- En términos generales en el problema analizado al utilizar una distribución normal-gama como *a priori*, el nivel de censuramiento utilizado (más del 50%) no tuvo un efecto drástico en el poder discriminatorio de la metodología. No es el caso de una distribución *a priori* uniforme, ya que en este caso el censuramiento afectó más el poder discriminatorio.
- Cuando ninguno de los modelos propuestos es el correcto, la metodología jerarquiza cuál de ellos está más cerca del modelo correcto. Esto queda en evidencia en las tablas 7.1 y 7.2. Por ejemplo de la tabla 7.1 se aprecia que si con muestreo aleatorio se generan datos del modelo normal, y se propone como modelo correcto a cualquiera de los otros cuatro, entonces la metodología señala que el que más se le acerca es el lognormal con probabilidad  $1 - 0.5015 = 0.4985$ , seguido en forma cercana por valor extremo y Weibull, y en último término y de forma muy lejana el exponencial. Esto garantiza que en la práctica aunque no se tenga el modelo correcto, la metodología seleccionara el que más se aproxime al verdadero. Por lo tanto queda descartado, el que la metodología concluya que ninguno de los considerados es el verdadero, como ocurre en ocasiones con las pruebas de falta de ajuste del enfoque clásico.

- Aunque los modelos estudiados se pueden considerar relativamente similares, como es evidente en la figura 7.1, la metodología en general tendió a hacer una selección adecuada. En los casos de modelos similares, la metodología prácticamente ha declarado un empate, y cuando la diferencia es drástica, prácticamente nunca se equivoca.
- Además la metodología que hemos propuesto podría utilizarse para implementar, en el contexto de confiabilidad, el procedimiento que se conoce como ponderación de modelos (Hoeting et al. 1999), en donde se promedia las distribuciones posteriores de cada uno de los modelos considerados, ponderados por las probabilidades posteriores del modelo:

$$\Pr(q | \mathbf{X}) = \sum_{i=1}^m \Pr(q | M_i, \mathbf{X}) \Pr(M_i | \mathbf{X})$$

donde  $q$  es la cantidad de interés sobre lo que se quiere hacer inferencias. Medigan y Raftery (1994) señalan que ponderar sobre todos los modelos como se indica, proporciona una mejor predicción que cualquiera de los modelos individuales. La desventaja de este enfoque, en todo caso, sería la mayor complejidad de cálculos y el no tener un modelo de referencia específico.

## CAPITULO 8

### ELECCIÓN DE TIEMPOS DE GARANTÍA

Uno de los objetivos básicos de muchos de los estudios de confiabilidad es decidir el tiempo o esquema de garantía para un producto, sin embargo esto prácticamente no se aborda en la literatura clásica de confiabilidad. Es común que sólo se sugiera implícitamente que el tiempo de garantía se fija con base en cuantiles bajos de la distribución del tiempo de vida, y se omite cualquier tipo de consideraciones económicas para esta elección. El objetivo del presente capítulo es abordar el problema de decidir el tiempo de garantía considerando los modelos *normal*, *lognormal*, *valor extremo*, *Weibull* y *exponencial*; y una función de utilidad que incorpore las diferentes consideraciones económicas y de mercadotecnia de la decisión.

#### 8.1 Introducción

Un aspecto que es cada día más relevante en un mercado globalizado de productos es la elección adecuada del tiempo y esquema de garantía (o fecha de caducidad) de un producto, ya que esto por sí mismo habla de la calidad y confiabilidad del producto, y muchas de las veces es una especie de carta de presentación que ayuda a comunicar características del producto y a reforzar las estrategias de mercadotecnia para ampliar la penetración del producto en el mercado (Padmanabhan, 1996; Kelley, 1996). Por ejemplo, en un estudio de The Federal Trade Commission(1985), de EU, encontró que el 75% de los consumidores encuestados estaba dispuesto a pagar una cantidad extra por productos con mejores esquemas de garantía. Por el contrario, si el período de garantía es pequeño esto podrá tener un efecto negativo en la elección del consumidor (Kelley, 1996; Emons, 1989) y además los competidores podrán utilizarlo como una herramienta comparativa. Sin embargo, si la fijación del período de garantía no se basa en la verdadera calidad y confiabilidad del producto, se puede incurrir en serios problemas, ya que si una gran cantidad de productos fallan antes del período de garantía los consumidores van a sentirse insatisfechos y la empresa tendrá que sufragar gastos adicionales para cumplir con tal garantía. En Blischke (1996), aplicando procesos de renovación, se hace un análisis de los costos que implica cumplir con la garantía, y hace estimaciones que pueden ir desde un pequeño porcentaje del precio original del producto hasta porcentajes del orden del 50%. En otras palabras, cuando el productor fije el precio de venta del producto, debe considerar entre los costos del producto el costo por cumplir con la garantía. Por ejemplo, sólo para dar una cifra, en 1995 se "llamó" en Estados Unidos a 17.8 millones de vehículos dentro del período de garantía debido a defectos de tales unidades (Inman y Gonsalvez, 1998), esto evidentemente generó costos de corto plazo para los fabricantes por las reparaciones y reemplazos, pero las dudas e insatisfacción ocasionada a los propietarios de los automóviles podrían ocasionar pérdidas de ventas futuras.

En el caso de productos cuyo tiempo de garantía se fija más bien en términos de una fecha de caducidad o fecha de consumo preferente, también es crítica su elección. Cuando ese período es largo eso ayudará a generar confianza al consumidor y a que al estar el producto más días en los anaqueles de los supermercados se incremente la probabilidad de que un cliente se lleve el producto. El asunto de la fecha de caducidad

o garantía se vuelve crítico en algunas ramas industriales como las alimenticias o farmacéuticas, ya que el consumidor podría demandar una compensación (Crane, 2000). Por lo anterior, la fijación del período de garantía es un problema de confiabilidad, pero también es una decisión que tiene que ver con asuntos de mercadotecnia, costos, leyes, etcétera.

En la literatura clásica de confiabilidad el problema de seleccionar un tiempo de garantía prácticamente no se aborda, ver por ejemplo Lawless(1982), Nelson (1982), Ansell y Phillips (1994) y Meeker y Escobar (1998). A lo que se limitan es señalar, las más de las veces en forma implícita, que el tiempo de garantía se fija con base en cuantiles bajos de la distribución del tiempo de vida, sin embargo nunca se señala cuál cuantil es el adecuado y qué tipo de consideraciones económicas o de mercado se deben hacer. Esto, a pesar de que se establece que el tiempo de garantía juega un papel importante en muchos estudios de confiabilidad. Por ejemplo, en Meeker y Escobar (1998, pag. 2), se señala algunas razones para realizar estudios de confiabilidad, entre las cuales nos interesa destacar dos: "Assesing characteristics of materials over a warranty period..." y "Predicting product warranty cost". De lo anterior, se concluye que existe un vacío importante en la literatura clásica de confiabilidad en cuanto a la forma que se aborda el problema de la elección del período de garantía.

En la literatura estadística, varios trabajos han sido orientado a aprovechar las bases de datos de las empresas en cuanto a reclamaciones por tiempos de garantía para estimar el número esperado de reclamaciones de garantía de las unidades en uso, así como los costos asociados (ver por ejemplo, Kalbfleisch et al. 1991; Rao, 1995; Kim y Rao, 2000). En este tipo de trabajos no se cuestiona la forma en la que se ha seleccionado el período de garantía, mas bien se estudian las consecuencias de esa elección. También se ha propuesto el uso de estos datos junto con la información de las unidades que no han fallado para estimar la distribución del tiempo de vida, ver por ejemplo Suzuki (1985a, b) y Lawless(1998).

En Inman y Gonsalvez (1998) se sustenta desde el punto de vista económico la necesidad de desarrollar un programa de pruebas después de que un nuevo modelo de carro ha sido lanzado, apoyándose en el desempeño de éstas unidades en ciertos clientes mayoristas claves como empresas de mensajería, por ejemplo. Todo esto como una estrategia para mejorar la calidad y confiabilidad de esos diseños, y así reducir los altos costos de garantía. Este artículo, y otros como Allen (2001) reflejan que para ciertos productos ya hay políticas y reglas de mercado que obligan a los productores a fijar ciertos períodos de garantía, lo que ocasiona en el futuro, sin preverlos con precisión, altos costos para cumplirlos. Lo ideal sería decidir en forma óptima la garantía, considerando todos los factores, y prever de antemano los costos asociados a esa decisión, para de esa manera enfocar programas de mejora de la calidad y confiabilidad de los productos para reducir esos costos.

En este sentido el estudio del tiempo garantía ha sido abordado desde muchas perspectivas. Por ejemplo en Blischke y Myrthy-Eds (1996), se analiza la garantía de un producto desde diferentes perspectivas: legales, costos, mercadotecnia, sociales, administración, y de diseño e ingeniería. Por ejemplo Menezes y Currim (1992) proponen determinar el tiempo de garantía y costo de un producto, optimizando una función de demanda (loglineal) que depende de doce variables: precio del producto, el tiempo de

garantía, la publicidad, los puntos de venta, la confiabilidad del producto, y características distintivas del producto; y éstas mismas variables pero medidas en el producto de la competencia. Sin embargo este modelo requiere tanta información previa, que en la práctica difícilmente son aplicables para tomar una decisión oportuna. De hecho el ejemplo que se presenta en el trabajo citado antes, además de dar por conocidos aspectos como la distribución del tiempo de vida del producto, termina decidiendo el tiempo de garantía y precio con base en un modelo que tiene un coeficiente de determinación de 0.57. Por lo tanto la calidad de la estimación deja mucho que desear. Así, no obstante que hay diferentes enfoques para diseñar el esquema de garantía para un producto, no existe una teoría completa, unificada y práctica al respecto (Marcellus y Pirojboot, 1996). Entonces los estudios que sobre el tema existen abordan los esquemas de garantía desde diferentes perspectivas, aportando la información de ese enfoque para evaluar o tomar una decisión.

En este contexto, nuestro objetivo es abordar la elección del tiempo de garantía del producto desde la perspectiva de la estadística Bayesiana, con la idea de decidir el tiempo de garantía a partir de una buena estimación del tiempo de vida del producto y a partir de una función de utilidad que incorpore diferentes consideraciones de costos, mercadotecnia y calidad (satisfacción del cliente). Como lo veremos más adelante, los trabajos de Singpurwalla y Wilson (1993, 1998) están orientados en un sentido relativamente similar al que proponemos, pero utilizan una función de utilidad que considera pocos aspectos relacionados con la decisión, y el modelo de tiempo de vida que usan está indexado por dos escalas (uso y tiempo). Por ello, nuestra idea sería proponer una función de utilidad que incluya consideraciones adicionales para una buena decisión y considerar los modelos de tiempo de vida más usuales en confiabilidad.

En suma, el objetivo del presente capítulo es abordar el problema de decidir el período de garantía para un producto y proponer un esquema de decisión óptima basado en métodos Bayesianos para los modelos normal, lognormal, valor extremo, Weibull y exponencial. Para ello en la sección 8.2 establecemos la función de utilidad, en la 8.3 el cálculo de la utilidad esperada para los diferentes modelos, y en la sección 8.4 se presenta un ejemplo.

## 8.2 Función de Utilidad

Decidir el tiempo o esquema de garantía,  $t_g$ , de un producto es una decisión, que deseamos soportar en un esquema coherente de toma de decisiones de tipo cuantitativo. Por ello de acuerdo a lo expuesto al final del capítulo 2, esta decisión se fundamenta en una función de utilidad  $u : A \times \Omega \rightarrow R$ . La decisión es un tiempo de garantía,  $a \equiv t_g \in A \equiv R^+$ , y la realidad desconocida la representa el tiempo de vida del producto,  $\omega \equiv t \in \Omega \equiv R^+$ . Además las creencias actuales acerca de los posibles estados de la realidad es lógico que lo represente la densidad posterior predictiva, por lo tanto,  $p(\omega) \equiv f(t|\mathbf{X})$ . De esta manera, de acuerdo a la ecuación (2.7), la elección óptima del tiempo de garantía  $t_g$  es el que maximiza la utilidad esperada:

$$u^*(t_g) \equiv E[u(t, t_g)] = \int_0^{\infty} u(t_g, t) f(t|\mathbf{X}) dt. \quad (8.1)$$

De aquí que para poder tomar esta decisión, necesitamos obtener  $f(t|X)$  para los diferentes modelos considerados en este estudio, aspecto que abordaremos en la siguiente sección. Además es necesario definir una función de utilidad  $u(t_g, t)$  que refleje las diferentes consecuencias de tomar la decisión,  $t_g$ . Esto es lo que haremos en esta sección.

De acuerdo a lo dicho en la primera sección de este capítulo, proponemos que la función de utilidad tome en cuenta tres aspectos fundamentales:

- i) Los beneficios económicos de un cierto tiempo de garantía,  $b(t_g)$ . En este rubro consideraremos los beneficios en mercadotecnia, imagen, probabilidad de venta, etcétera.
- ii) El costo directo en la que incurre el fabricante,  $r(t, t_g)$ , cuando el producto falla (al tiempo  $t$ ) dentro del período de garantía; por lo tanto aquí se debe incluir el costo directo por cumplir con tal garantía.
- iii) El costo de insatisfacción del cliente debido a que el producto falla (al tiempo  $t$ ) dentro del período de garantía,  $I(t, t_g)$ . Este es un costo indirecto que debe reflejar el efecto de un cliente insatisfecho sobre la imagen o prestigio de la empresa.

De acuerdo a estos tres puntos, que están en las mismas unidades, la función de utilidad estará dada por

$$u(t_g, t) = b(t_g) - r(t, t_g) - I(t, t_g) \quad (8.2)$$

A continuación propondremos cómo definir cada uno de estos componentes de la función de utilidad

Beneficio,  $b(t_g)$ . Decíamos en la primera sección de este capítulo que entre mayor sea el período de garantía más es el beneficio, sin embargo es poco realista suponer una función de beneficio creciente y no acotada como se propone en Singpurwalla y Wilson (1998), quienes proponen una función logaritmo. Ya que dar un período de garantía mucho mayor al de los competidores es probable que ya no traiga un beneficio real e incluso puede causar duda o suspicacia de parte del cliente. Además a mayor sea el período de garantía mayores serán los costos asociados a su cumplimiento. Por lo anterior resulta más adecuada una función de beneficio por este concepto que sea creciente, pero que esté acotada. Una función que cumple con las características señaladas antes es la siguiente:

$$b(t_g) = A_2[1 - e^{-A_1 t_g}] \quad (8.3)$$

que es una función creciente, cuya rapidez de crecimiento es proporcional a la constante  $A_1$ ; además está acotada asintóticamente por la línea recta  $h(t_g) = A_2$ . Las constantes  $A_1$  y  $A_2$  deben ser especificadas por el productor en función de los beneficios que se esperan por tener ciertos períodos de garantía. Para hacer ello, se puede pensar en las ganancias por vender cierta cantidad de productos, resultado de considerar el precio de ventas menos el costo del mismo (Singpurwalla y Wilson, 1993), o también a

partir de hacer consideraciones para un sólo producto pensando en su precio de venta (Singpurwalla y Wilson, 1998). Nosotros vamos a trabajar por ésta última estrategia.

De esta manera para especificar  $A_1$  y  $A_2$ , proponemos considerar los siguientes aspectos: el precio de venta del producto,  $p_v$  y un tiempo de garantía estándar  $t_e$ , en el que *a priori* se esperaría continuar con el nivel de utilidad actual ( $p_v$ ); este tiempo puede obtenerse a partir de tiempos de garantía para productos similares ya sea del mismo productor o de los competidores. Además especificar un tiempo de garantía mayor al estándar,  $t_c$ , con el cual se esperaría generará un ventaja competitiva, y con ello un cierto incremento proporcional en las ventas,  $I_v$ , y con ello en la utilidad. Entonces la elección de  $A_2$  debe ser igual al beneficio esperado con este incremento en ventas, en otras palabras:

$$A_2 = p_v(1 + I_v) \quad (8.4)$$

La elección de  $A_1$  debe ser tal que  $h(t_e) = p_v$  cuando  $A_2 = p_v(1 + I_v)$ , esto es

$$p_v(1 + I_v)[1 - e^{-A_1 t_e}] = p_v$$

despejando de aquí  $A_1$ , obtenemos

$$A_1 = t_e^{-1} \log \left( \frac{1 + I_v}{I_v} \right) \quad (8.5)$$

Otra función que tiene características similares a las que tiene la función de beneficio propuesta, es la tangente hiperbólica, y bajo ciertas condiciones puede dar una mayor flexibilidad en la especificación de la función beneficio.

Costo por cumplir garantía  $r(t, t_g)$ . En este rubro se debe incluir el costo directo en el que incurre el fabricante cuando el producto falla dentro del período de garantía. Para ello es necesario considerar los tres planes típicos de garantía (Menezes y Currim, 1992):

- *Reemplazo.* Bajo este esquema, sin costo para el consumidor, el fabricante reemplaza el producto que falla dentro del período de garantía, por uno nuevo. El nuevo producto es garantizado sólo por el período original de la garantía. Este esquema de garantía es muy usual para productos pequeños, que son durables, pero que no son demasiado costosos; y generalmente se aplica a partes o componentes de productos. Por ejemplo partes de carros. Este tipo de garantía también se le conoce como “libre de fallas”, “garantía estándar”, “garantía completa”, en Blischke (1996) se hace un análisis más extenso de esta forma de garantía.
- *Reparación.* En este esquema el producto que falla dentro del período de garantía es reparado, sin costo para el consumidor. Esta forma de garantía es usual en productos complejos y durables, como carros, refrigeradores, etc.
- *Inversamente proporcional al uso.* Bajo este esquema, si un producto falla dentro del período de garantía, la empresa le da al consumidor un reembolso o un crédito

para una nueva compra, que es proporcional al valor de la porción no usada de la garantía. En otras palabras, la garantía (reembolso, crédito para una nueva compra) es inversamente proporcional al tiempo de uso. Esta forma de garantía es utilizada para productos que fallan por su desgaste de uso y que no es factible repararlos, por lo que deben ser reemplazados por un producto nuevo. Por ejemplo baterías y llantas para automóviles.

De acuerdo a lo anterior, si el producto falla dentro de la garantía, proponemos que el costo en el que incurre la empresa esté dado por la siguiente función

$$r(t, t_g) = A_4 \left(1 - \frac{A_3 t}{t_g}\right) \quad (8.6)$$

donde  $A_3 t/t_g$  es el factor de cargo al cliente por uso, y que ayudará a definir un esquema de garantía del tipo inversamente proporcional al uso. De esa manera si se tiene un esquema de garantía de reemplazo o reparación, entonces  $A_3 = 0$ . Además la constante  $A_4$  debe ser igual al costo directo en el que incurre el fabricante por reparar o reemplazar el producto. Esta forma de definir este componente de la función de utilidad es más general que el propuesto en Menzefricke (1992) y en Singpurwalla y Wilson (1998), ya que ellos sólo consideran el esquema de garantía inversamente proporcional al uso.

Costo de insatisfacción  $I(t, t_g)$ . En este otro componente de la función de utilidad no considerado en Menzefricke (1992) y Singpurwalla y Wilson (1998), incluiremos el costo para el productor (costo indirecto) debido a que el producto falla (al tiempo  $t$ ) dentro del período de garantía. O sea este es el costo de insatisfacción del cliente debido a que el producto falla dentro del período de garantía (Patankar y Mitra, 1996). El cliente no espera que el producto funcione para siempre, pero sí tiene ciertas expectativas que son reforzadas o incrementadas por un tiempo de garantía largo, por lo que si el producto falla relativamente rápido después de la compra, la insatisfacción del consumidor podría ser significativa (Kelley, 1996). Aunque la garantía minimize o atenúe tal insatisfacción, cualquier reclamo de una garantía genera costo para el consumidor que no son cubiertos por la garantía (tiempo, desplazamientos, la interrupción en el uso del producto, frustración por no cumplimiento de expectativa, etcétera). Por ello es necesario tomar en cuenta esta insatisfacción en la función de utilidad. De tal manera que esta función penalice garantías largas si es que no están respaldadas por la confiabilidad del producto. Porque es ampliamente reconocido ahora que un cliente insatisfecho afecta la imagen o prestigio de la empresa y eso repercute negativamente en las ventas futuras. Por ejemplo Deming (1989, pag. 94) cita un trabajo realizado en la industria automotriz, en el que se afirma que un propietario satisfecho de un carro vale más que 10 folletos publicitarios, ya que éste regresa sin publicidad, es probable que comunique su satisfacción a ocho personas en promedio, e incluso puede que traiga a un amigo. Sin embargo si está insatisfecho e irritado le contará sus problemas a un promedio de dieciséis personas. Obviamente la insatisfacción del cliente será mayor en cuanto más pronto falle el producto, ya que aunque la falla sea cubierta por la garantía, la falla causa molestias. Por ello proponemos que este costo indirecto se cuantifique mediante una función como la siguiente:

$$I(t, t_g) = A_5 \left(1 - \frac{t}{t_g}\right). \quad (8.7)$$

La especificación de  $A_5$  se puede hacer a partir de considerar el costo del máximo nivel de insatisfacción, que se daría si el producto falla en forma muy temprana ( $t$  cercano a cero). Como es difícil cuantificar esto, proponemos que se asigne como una proporción,  $q_i$ , de la utilidad de la venta del producto, por lo tanto

$$A_5 = q_i \times p_v. \quad (8.8)$$

Con base en las consideraciones anteriores, sustituyendo las expresiones (8.3), (8.6) y (8.7) en la función de utilidad (8.2), entonces la función de utilidad que proponemos está definida de la siguiente manera:

$$u(t, t_g) = \begin{cases} A_2[1 - e^{-A_1 t_g}] - A_4\left(1 - \frac{A_3 t}{t_g}\right) - A_5\left(1 - \frac{t}{t_g}\right) & \text{si } t \leq t_g \\ A_2[1 - e^{-A_1 t_g}] & \text{si } t > t_g. \end{cases} \quad (8.9)$$

Hemos sumado los diferentes componentes de la función de utilidad debido a que todos sus componentes están definidos en costos monetarios. Como es evidente si se revisa la forma en que hemos recomendado especificar las constantes  $A_j$  ( $j = 1, \dots, 5$ ). Sustituyendo (8.9) en (8.1), la utilidad esperada está dada por

$$u^*(t_g) \equiv E(u(t, t_g)) = \int_0^{t_g} \left[ A_2(1 - e^{-A_1 t_g}) - A_4\left(1 - \frac{A_3 t}{t_g}\right) - A_5\left(1 - \frac{t}{t_g}\right) \right] f(t|\mathbf{X}) dt + \int_{t_g}^{\infty} A_2[1 - e^{-A_1 t_g}] f(t|\mathbf{X}) dt$$

Como

$$\int_0^{t_g} [A_2(1 - e^{-A_1 t_g})] f(t|\mathbf{X}) dt + \int_{t_g}^{\infty} A_2[1 - e^{-A_1 t_g}] f(t|\mathbf{X}) dt = A_2[1 - e^{-A_1 t_g}]$$

entonces la utilidad esperada toma la forma siguiente

$$u^*(t_g) \equiv E(u(t, t_g)) = A_2[1 - e^{-A_1 t_g}] - \int_0^{t_g} \left[ A_4\left(1 - \frac{A_3 t}{t_g}\right) + A_5\left(1 - \frac{t}{t_g}\right) \right] f(t|\mathbf{X}) dt \quad (8.10)$$

por lo tanto, la decisión óptima para el tiempo de garantía,  $t_g^*$ , está dada por el  $t_g$  que maximice la utilidad esperada, es decir,

$$t_g^* = \arg \max_{t_g} \left[ \left\{ A_2[1 - e^{-A_1 t_g}] - \int_0^{t_g} \left( A_4\left(1 - \frac{A_3 t}{t_g}\right) + A_5\left(1 - \frac{t}{t_g}\right) \right) f(t|\mathbf{X}) dt \right\} \right] \quad (8.11)$$

Como en los más de los casos las expresiones anteriores no tienen solución analítica, entonces será necesario calcularla por métodos numéricos o bien por simulación. Nosotros optaremos por ésta última alternativa.

### 8.3 Cálculo de Utilidad Esperada

En la sección anterior establecimos, que el tiempo óptimo de garantía resultaba de maximizar la utilidad esperada (8.10, 8.11). Para ello es necesario conocer la densidad posterior predictiva. En esta sección veremos la forma de calcular la utilidad esperada para diferentes modelos, y con base en eso obtener el tiempo óptimo de garantía. Lo haremos para los diferentes modelos estadísticos que se consideraron en el capítulo 4, junto con los dos tipos de distribuciones *a priori* (normal-gama y uniforme) que se discutieron en el capítulo 5. Para el único modelo que existen expresiones analíticas para la pérdida esperada es para el exponencial, y para el resto recurriremos a un método por simulación.

Modelo Exponencial (*a priori* gama). En la sección 4.1 establecimos generalidades del modelo exponencial: su densidad (4.14), la verosimilitud para muestreo aleatorio (4.15) y muestreo bajo censura (4.16). También obtuvimos la densidad posterior predictiva al utilizar *a priori* gama (4.25). Con estos elementos estamos en posibilidades de formular el siguiente resultado.

**Teorema 8.1** *La utilidad esperada (8.10) para el modelo exponencial está dada por:*

a) Con *a priori* gama y muestreo sin censura:

$$u^*(t_g) = A_2(1 - e^{-A_1 t_g}) - (A_4 + A_5) \left( 1 - \frac{(v_5 - \gamma)^{n+\alpha_5}}{(v_5 + t_g - \gamma)^{n+\alpha_5}} \right) + \frac{(A_4 A_3 + A_5)}{t_g} \left( \frac{\gamma v_5^{-\alpha_5-n} - t_g (v_5 - \gamma + t_g)^{-\alpha_5-n}}{\alpha_5 + n} + \frac{v_5^{-n-\alpha_5+1} - (v_5 - \gamma + t_g)^{-n-\alpha_5+1}}{(-1 + n + \alpha_5)(n + \alpha_5)} \right) \quad (8.12)$$

con  $v_5 = \sum_{i=1}^n x_i + \beta_5 - n\gamma$ .

b) Con *a priori* gama y muestreo censurado:

$$u^*(t_g) = A_2(1 - e^{-A_1 t_g}) - (A_4 + A_5) \left( 1 - \frac{(p_5 - \gamma)^{r+\alpha_5}}{(p_5 + t_g - \gamma)^{r+\alpha_5}} \right) + \frac{(A_4 A_3 + A_5)}{t_g} \left( \frac{\gamma p_{5g}^{-\alpha_5-n} - t_g (p_5 - \gamma + t_g)^{-\alpha_5-n}}{\alpha_5 + r} + \frac{p_5^{-r-\alpha_5+1} - (p_5 - \gamma + t_g)^{-r-\alpha_5+1}}{(-1 + r + \alpha_5)(r + \alpha_5)} \right) \quad (8.13)$$

con  $p_5 = \sum_{i=1}^r x_{(i)} + (n-r)T_e + \beta_5 - n\gamma$ .

Para encontrar en (8.12) o (8.13), el valor de  $t_g$  que maximiza la utilidad, se aplica algún método numérico convencional, por ejemplo el método de interpolación polinomial de Brent (ver Brent, 1973), y que es en el que se basa el procedimiento *optimize* del sistema computacional *S-Plus 2000*. También puede usarse el método cuasi-Newton (ver Dennis y Mei 1979) que viene implementado en la función *nlmin* de *S-Plus 2000*.

Dada la utilidad esperada (8.1), y de acuerdo a (Robert y Casella, 1999, pag. 75), el cálculo de ésta por simulación Monte Carlo estará dada por

$$u^*(t_g) = \frac{1}{M^*} \sum_{j=1}^{M^*} u(t_j^*, t_g) \quad (8.14)$$

donde  $\{t_j^*; j = 1, \dots, M^*\}$  es una muestra de la densidad posterior predictiva  $f(t|\mathbf{X})$ . De esta manera para decidir el tiempo óptimo de garantía se evaluará (8.14) para un rango amplio de posibles valores del tiempo de garantía  $t_g$ , y se seleccionará el que maximice la utilidad. De acuerdo a lo anterior lo que se requiere es ver la forma de generar una muestra aleatoria de  $t_j^*$  de la densidad  $f(t|\mathbf{X})$ . A continuación vamos a explicar cómo hacer esto para los modelos normal y lognormal, suponiendo muestreo completo y *a priori* normal-gama.

Modelo Normal y Lognormal (muestreo sin censura y *a priori* normal-gama). En el teorema 4.3 vimos que bajo muestreo sin censura y *a priori* normal-gama  $(\alpha_1, \beta_1, m_1, k_1)$ , la densidad posterior predictiva para el modelo normal  $f_1(t|\mathbf{X})$ , (4.23), es una densidad *t*-Student  $(m_t, l_t, 2\alpha_1 + n)$ , con los siguientes parámetros: localización (media)  $m_t = \frac{k_1 m_1 + n \bar{x}_n}{n + k_1}$ , precisión  $l_t = \left[ \frac{1}{2} n k_1 \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} + \beta_1 + \frac{s_n}{2} \right]^{-1} \frac{(\alpha_1 + n/2)(n + k_1)}{(k_1 + n + 1)}$  y  $2\alpha_1 + n$  grados de libertad. Además sabemos, (ver Bernardo y Smith, 1994, pag. 123), que si  $y$  se distribuye *t*-Student  $(\mu, \lambda, \alpha)$ , entonces  $z = \lambda^{1/2}(y - \mu)$  se distribuye *t*-Student estándar  $(0, 1, \alpha)$ . Es precisamente la distribución *t*-Student estándar la que viene incluida en la mayoría de los programas estadísticos computacionales. Por lo anterior para obtener una muestra de  $t$  de  $f_1(t|\mathbf{X})$ , se aplican los siguientes pasos:

1. Obtener  $z$  de una *t*-Student estándar con  $2\alpha_1 + n$  grados de libertad
2. Calcular  $t^* = l_t^{-1/2} z + m_t$ ; entonces de acuerdo a lo que hemos dicho  $t$  se distribuye de acuerdo a la densidad posterior predictiva  $f_1(t|\mathbf{X})$  del modelo normal.
3. Se repite los pasos anterior  $M^*$  veces, y con los  $t_i^*$  obtenidos se evaluará (8.14) para un rango amplio de posibles valores del tiempo de garantía  $t_g$ , y se seleccionará el que maximice la utilidad.

Respecto al modelo lognormal, en el inciso (b) del teorema 4.3, vimos que bajo muestreo completo y *a priori* normal-gama  $(\alpha_2, \beta_2, m_2, k_2)$ , la densidad posterior predictiva para el modelo lognormal  $f_2(t|\mathbf{X})$ , (4.24), es una densidad log*t*-Student  $(m_l, l_l, 2\alpha_2 + n)$ , con los siguientes parámetros: localización (media)  $m_l = \frac{k_2 m_2 + n \bar{w}_n}{n + k_2}$ , precisión  $l_l = \left[ \frac{1}{2} n k_2 \frac{(m_2 - \bar{w}_n)^2}{k_2 + n} + \beta_2 + \frac{u_n}{2} \right]^{-1} \frac{(\alpha_2 + n/2)(n + k_2)}{(k_2 + n + 1)}$  y  $2\alpha_2 + n$  grados de libertad. Además sabemos que si  $w$  se distribuye log*t*-Student  $(\mu_l, \lambda_l, \alpha_l)$ , entonces la variable  $\log(w)$  se distribuye *t*-Student  $(\mu_l, \lambda_l, \alpha_l)$ . De esta manera, y recurriendo a la *t*-Student estándar, para obtener una muestra de  $t$  de  $f_2(t|\mathbf{X})$ , se aplican los siguientes pasos:

1. Obtener  $z$  de una *t*-Student estándar con  $2\alpha_2 + n$  grados de libertad
2. Calcular  $t^* = \exp \left( l_l^{-1/2} z + m_l \right)$ ; entonces de acuerdo a lo que hemos dicho  $t^*$  se distribuye de acuerdo a la densidad posterior predictiva  $f_2(t|\mathbf{X})$  del modelo lognormal.

3. Se repite los pasos anterior  $M^*$  veces, y con los  $t_i^*$  obtenidos se evaluará (8.14) para un rango amplio de posibles valores del tiempo de garantía  $t_g$ , y se seleccionará el que maximice la utilidad.

Sampling-importance-resampling (obtención por simulación de la predictiva). De los modelos que establecimos en el capítulo 4, en las subsecciones anteriores hemos visto cómo obtener la utilidad esperada para el modelo exponencial con *a priori* gama y bajo ambos tipos de muestreo (completo y censurado), y para los modelos normal y lognormal con muestreo completo y *a priori* normal-gama. Para *todos los casos restantes* no existe una forma analítica para la densidad posterior predictiva  $f(t|\mathbf{X})$ , y en consecuencia tampoco para la utilidad esperada. Por ello para decidir el tiempo de garantía será necesario obtener por simulación una muestra de la densidad posterior predictiva, esto lo vamos a hacer con la técnica de simulación conocida como Sampling-importance-resampling (ver Bernardo y Smith, 1994, pag. 350). Esta técnica se basa en el siguiente resultado, conocido como método de Aceptación-Rechazo( Accept-Reject).

**Teorema 8.2** Sean  $g(x)$  y  $h(y)$  densidades, a la primera la llamaremos densidad instrumental y la segunda densidad objetivo. Si existe una constante  $M$  tal que  $h(x) \leq Mg(x)$  en todo el soporte de  $h$ ; entonces se puede generar  $h$  a partir de  $g$  con el siguiente algoritmo:

1. Generar  $X \sim g$  y  $U \sim u(0, 1)$ ;
2. Aceptar  $Y = X$  si  $U \leq \frac{h(X)}{Mg(X)}$ ;
3. Regrese a 1 en otro caso.

Entonces la distribución de  $Y$  es  $h$ . La demostración de este teorema es muy sencilla, parte de ver que  $\Pr(Y \leq y) = \Pr\left(X \leq y \mid U \leq \frac{h(X)}{Mg(X)}\right)$ , y se puede consultar en (Robert y Casella, 1999, pag. 49).

Si en el anterior teorema, hacemos la densidad objetivo  $h$  igual a la densidad posterior:  $h(\theta|\mathbf{X}) = L(\mathbf{X}|\theta)\pi(\theta)$ , a la densidad instrumental  $g$  igual a la distribución *a priori*  $\pi(\theta)$ , y a  $M = L(\mathbf{X}|\hat{\theta})$ , donde  $\hat{\theta}$  maximiza la función de verosimilitud. Entonces se sigue de manera directa que

$$h(\theta|\mathbf{X}) = L(\mathbf{X}|\theta)\pi(\theta) \leq L(\mathbf{X}|\hat{\theta})\pi(\theta) \equiv M\pi(\theta) \quad (8.15)$$

Con esto tenemos una forma de generar la distribución posterior con base a la *a priori*. En particular de (8.15), se aprecia que la forma de aceptar en el paso 2 del algoritmo que es parte del teorema anterior, es con

$$U \leq \frac{h(\theta|\mathbf{X})}{M\pi(\theta)} \equiv \frac{L(\mathbf{X}|\theta)\pi(\theta)}{L(\mathbf{X}|\hat{\theta})\pi(\theta)} = R(\mathbf{X}|\theta) \quad (8.16)$$

donde  $R(\mathbf{X}|\theta)$  es la función de verosimilitud relativa, que en general bajo el modelo  $M_i$  está definida de la siguiente manera

$$R(\mathbf{X}|\theta) = \frac{L(\mathbf{X}|\theta, M_i)}{L(\mathbf{X}|\hat{\theta}, M_i)}$$

donde  $\hat{\theta}$  maximiza la función de verosimilitud.

Con estos elementos, estamos en posibilidades de describir un método muy usual para generar muestras aleatorias de la densidad posterior,  $\pi(\theta|\mathbf{X})$ , o de la densidad posterior predictiva  $f(y|\mathbf{X})$ . Este método de muestreo se conoce como Sampling-importance-resampling. Que aunque es más general, adaptándolo a nuestras necesidades, consiste en los siguientes pasos (ver Bernardo y Smith, 1994, pag. 350):

- a) Generar una muestra aleatoria  $\theta_i, i = 1, \dots, M$ , de la priori  $\pi(\theta)$  (ver algoritmo B1 y B2 del apéndice de cómputo).
- b) Retener el  $i$ -ésimo valor de la muestra,  $\theta_i$ , con probabilidad  $R(\mathbf{X}|\theta_i)$ . Esto se hace generando una realización aleatoria,  $U_i$ , de una distribución uniforme(0, 1), y retener  $\theta_i$  si  $U_i \leq R(\mathbf{X}|\theta_i)$ .
- c) De acuerdo al teorema 8.2, los valores retenidos de la muestra,  $\theta_1^*, \dots, \theta_{M^*}^*$  ( $M^* \leq M$ ), es una muestra aleatoria de la densidad posterior  $\pi(\theta|\mathbf{X})$ .
- d) Para cada  $\theta_j^*$  se genera una  $t_j^*$  directamente del modelo  $f(t|\theta_j^*)$ , entonces  $t_j^*$ , con  $j = 1, \dots, M^*$ , es en realidad una muestra de la densidad posterior predictiva  $f(t|\mathbf{X})$  (ver Smith y Robert, 1993). Nótese que dados valores de los parámetros  $\theta_j^*$  es muy sencillo generar  $t_j^*$  de los modelos que hemos considerado  $f(t|\theta_j^*)$  (normal, lognormal, valor extremo, Weibull y exponencial).

Con los  $t_j^*$  obtenidos se evalúa la utilidad esperada (8.14) para un rango amplio de posibles valores del tiempo de garantía  $t_g$ , y se seleccionará el que maximice la utilidad esperada. Obviamente este proceso de optimización de (8.14), con los  $t_j^*$  obtenidos, puede hacerse directamente con un método numérico de optimización como los que se citaron al final del teorema 8.1.

#### 8.4 Ejemplo

En esta sección presentaremos un ejemplo completo de especificación del tiempo de garantía, suponiendo una función de pérdida como la de (8.9).

Planteamiento del Problema e Información Inicial. A partir de datos parciales proporcionados por un fabricante de un componente automotriz (balatas) se ha completado los datos de este ejemplo. También se van a especificar los valores necesarios para definir los parámetros de la función de pérdida (ver sección 8.2)

El producto consiste en un juego de balatas (un par), y tiene un precio de venta de 900 pesos ( $p_v = 900$ ). Usualmente se ha venido dando una garantía de diez mil kilómetros bajo condiciones de uso normal ( $t_e = 10$ ). Aunque esto es una garantía estándar en el mercado, el fabricante tiene interés en aumentar el tiempo de garantía,

y ha evaluado que si ésta se eleva a 13 mil kilómetros ( $t_c = 13$ ), podría aumentar sus ventas un 10% ( $I_c = 0.1$ ). Con estos elementos, estamos en posibilidades de especificar las constantes  $A_1$  y  $A_2$  de la función de utilidad. De acuerdo a (8.5):

$$\begin{aligned} A_1 &= t_e^{-1} \log \left( \frac{1 + I_v}{I_v} \right) \\ &= (10)^{-1} \log \left( \frac{1 + 0.1}{0.1} \right) \\ &= 0.23979 \end{aligned}$$

Considerando (8.4):

$$\begin{aligned} A_2 &= p_v(1 + I_c) \\ &= 900(1 + 0.1) \\ &= 990.0 \end{aligned}$$

Por otro lado, la forma de garantía que se suele utilizar es el de esquema de reemplazo, es decir, el fabricante reemplaza el producto que falla dentro del período de garantía, por uno nuevo. De aquí que el costo de reemplazo es igual al costo de producción de 750. Por lo tanto,  $A_3 = 0$  y  $A_4 = 750$ .

Respecto al costo de imagen máximo por una falla dentro del período de garantía se decide establecer igual al 12% del precio de venta ( $q_i = 0.12$ ), por ello, de acuerdo a (8.8):

$$\begin{aligned} A_5 &= q_i \times p_v \\ &= 0.12 \times 900 \\ &= 108 \end{aligned}$$

Con información parcial se simularon los datos de la tabla 8.1. La información está en miles de kilómetros, los números con asterisco significa que fueron censurados a ese tiempo. Los datos se simularon de un modelo Weibull, con censura aleatoria independiente de los parámetros Weibull.

<p>9.2, 10.4, 11.2*, 12.1, 12.1*, 12.7*, 12.7*, 13.1*, 13.2, 13.2*,  14.4, 14.5*, 14.7*, 14.9*, 15.1, 15.3, 15.5*, 15.6, 16.1*, 16.4,  16.4*, 16.5*, 17.0*, 17.1, 17.4*, 17.5, 17.6*, 18.0, 18.1*, 18.2,  18.5*, 18.7*, 19.9*, 19.9, 20.1, 20.3, 20.7*, 23.2, 25.1*, 26.1*</p>
--

Tabla 8.1. Datos en miles de kilómetros para duración de balatas.

Partiendo de que el fabricante espera un tiempo medio de vida del producto entre [18, 24], y la desviación estándar entre [2.5, 6.5]. Con base en esto, y aplicando los resultados de la tabla 5.2 se obtienen los valores de los hiperparámetros de las distribuciones *a priori* uniforme para los parámetros de los diferentes modelos. Dando una probabilidad *a priori* similar para los cinco modelos y obteniendo por simulación Monte

Carlo la verosimilitud integrada, obtenemos las siguientes probabilidades posteriores para cada modelo:  $\Pr(M_1 | \mathbf{X}) = 0.3658$ ,  $\Pr(M_2 | \mathbf{X}) = 0.2103$ ,  $\Pr(M_3 | \mathbf{X}) = 0.04991$ ,  $\Pr(M_4 | \mathbf{X}) = 0.3739$  y  $\Pr(M_5 | \mathbf{X}) = 0$ . Por lo tanto los dos modelos con mayores méritos son el Weibull, seguido del normal y lognormal. Si se utiliza como a priori una distribución normal-gama, las probabilidades posteriores son:  $\Pr(M_1 | \mathbf{X}) = 0.5129$ ,  $\Pr(M_2 | \mathbf{X}) = 0.0804$ ,  $\Pr(M_3 | \mathbf{X}) = 0.0507$ ,  $\Pr(M_4 | \mathbf{X}) = 0.3560$  y  $\Pr(M_5 | \mathbf{X}) = 0$ . Por lo tanto con esta opción los dos modelos con mayores probabilidades posteriores son el normal y el Weibull.

Con base a lo anterior, haremos la elección del tiempo de garantía para las balatas considerando dos modelos: Weibull y normal. Esto ayudará a tomar una mejor decisión.

Selección óptima del tiempo de garantía (modelo Weibull). Utilizando un modelo Weibull para el tiempo de vida y a partir de la información de la subsección anterior, estamos en posibilidades de aplicar el método Sampling-importance-resampling, para con base en él obtener una muestra de la distribución posterior predictiva, y así poder calcular la utilidad esperada (8.14).

Percentil	Predicción
0	2.344
1	8.303
5	12.150
10	14.525
15	15.801
20	16.889
25	17.751
30	18.448

Tabla 8.2. Percentiles de la predictiva para modelo Weibull (*a priori* normal-gama).

**A priori normal-gama.** Utilizando una *a priori* normal-gama para los parámetros del modelo Weibull, y aplicando el algoritmo Sampling-importance-resampling con  $M = 6500$  replicas, se obtuvo una muestra de tamaño 1657 de la distribución posterior predictiva que se muestra en el histograma de la figura 8.1. De donde se puede apreciar claramente el tiempo de vida (en miles de kilómetros) del producto. Algunos percentiles claves de esta densidad se muestra en la tabla 8.2, que fueron obtenidos a partir de la muestra y utilizando la función *quantile* del sistema S-Plus.

En la figura 8.2 se aprecia la utilidad esperada, el tiempo de garantía que maximiza ésta función es  $t_g^* = 12.627$  miles de kilómetros, que prácticamente corresponde al percentil 5.5 de la distribución posterior predictiva. La función de utilidad evaluada en este punto es igual a 899.13, que sería la utilidad esperada máxima. Además en la figura 8.2, se puede apreciar que sólo en un rango pequeño alrededor de  $t_g^* = 12.627$ , se esperarían utilidades similares, por lo tanto de acuerdo a esto el tiempo de garantía por el que finalmente se opte debe ser muy cercano a este valor.

**A priori uniforme.** Utilizando una distribución uniforme para los parámetros del modelo Weibull, y aplicando el algoritmo Sampling-importance-resampling con  $M =$

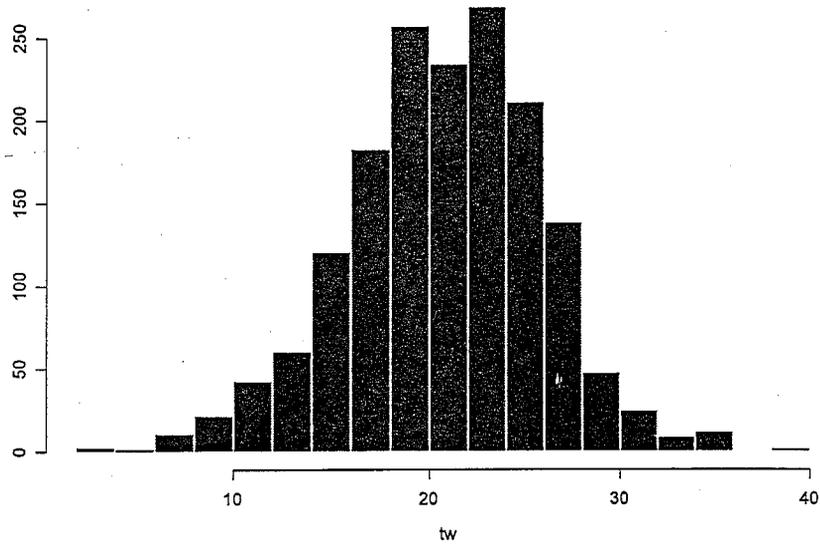


Figura 8.1: Predictiva para balatas, modelo Weibull (*a priori* normal-gama).

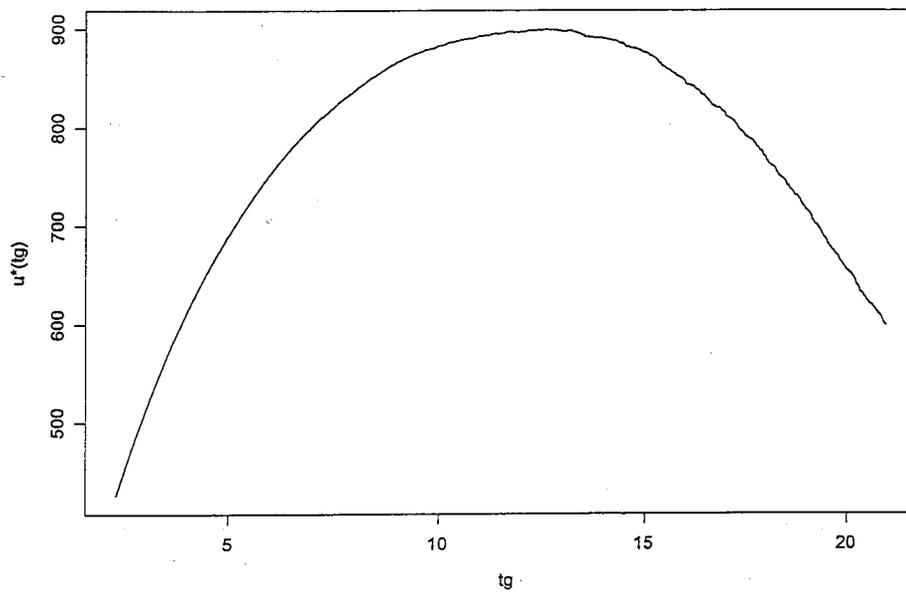


Figura 8.2: Utilidad esperada para balatas, modelo Weibull (*a priori* normal-gama).

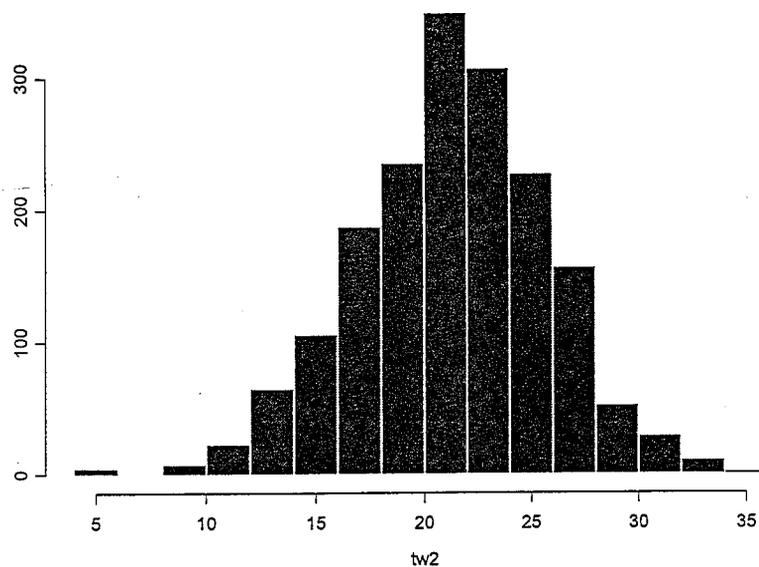


Figura 8.3: Predictiva para balatas, modelo Weibull (*a priori* uniforme).

15000 replicas, con lo que se obtuvo una muestra de tamaño 1769 de la distribución posterior predictiva que se muestra en el histograma de la figura 8.3. De donde se puede apreciar que claramente el tiempo de vida (en miles de kilómetros) del producto. Algunos percentiles claves de esta densidad se muestra en la tabla 8.3.

Percentil	Predicción
0	4.058
1	10.297
5	13.740
10	15.682
15	16.760
20	17.662
25	18.374
30	19.223

Tabla 8.3. Percentiles de la predictiva para modelo Weibull (*a priori* uniforme).

En la figura 8.4 se aprecia la función de utilidad esperada, el tiempo de garantía que maximiza ésta función es  $t_g^* = 12.82$ , que corresponde al percentil 3 de la distribución posterior predictiva. La función de utilidad evaluada en este punto es igual a 922.02, que sería la utilidad esperada máxima. Además en la figura 8.4, se puede apreciar que en un rango de 12 a 14 la función de utilidad es bastante similar, por lo tanto como tiempo de garantía se podría incluso dar un tiempo mayor a 13.00 mil kilómetros.

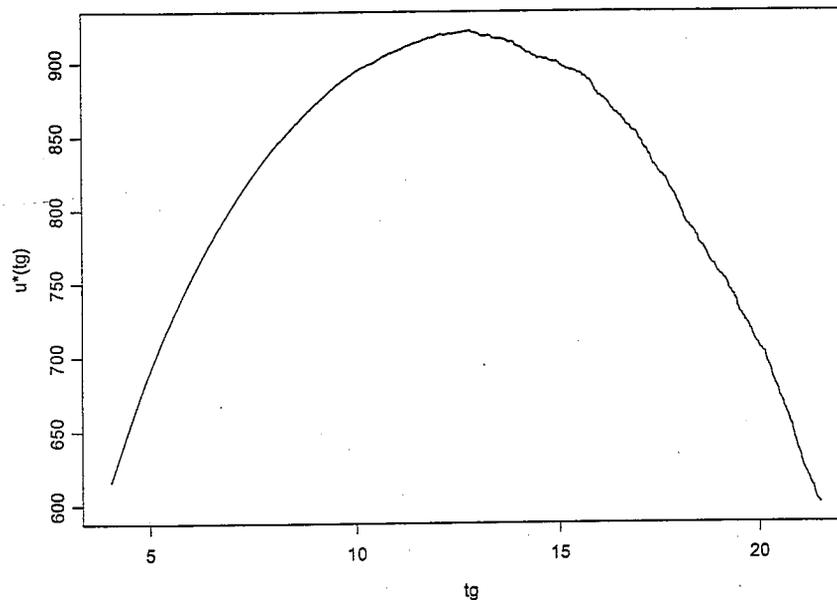


Figura 8.4: Utilidad esperada para balatas, modelo Weibull (*a priori* uniforme).

Selección óptima del tiempo de garantía (modelo normal). Utilizando un modelo normal para el tiempo de vida y a partir de la información de la subsección anterior, estamos en posibilidades de aplicar el método Sampling-importance-resampling, para con base en él obtener una muestra de la distribución posterior predictiva, y así poder calcular la utilidad esperada (8.14).

**A priori normal-gama.** Utilizando una *a priori* normal-gama para los parámetros del modelo normal, y aplicando el algoritmo Sampling-importance-resampling con  $M = 6500$  réplicas, una muestra de tamaño 1845 de la distribución posterior predictiva, esta muestra se aprecia en el histograma de la figura 8.5. De donde se puede apreciar claramente el tiempo de vida (en miles de kilómetros) del producto. Algunos percentiles claves de esta densidad se muestra en la tabla 8.4.

Percentil	Predicción
0	3.926
1	10.124
5	13.121
10	14.902
15	15.900
20	16.945
25	17.762
30	18.540

Tabla 8.4. Percentiles de la predictiva para modelo normal (*a priori* normal-gama).

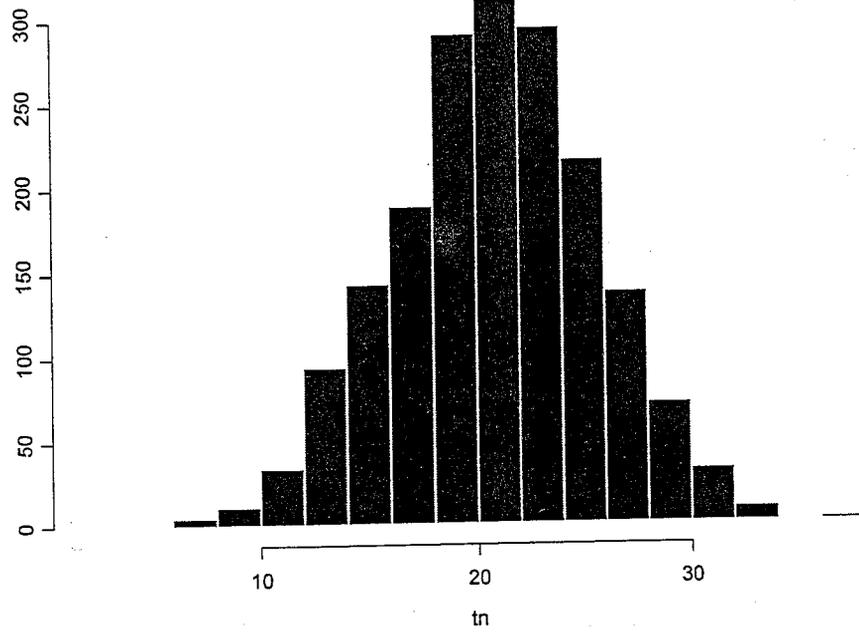


Figura 8.5: Predictiva para balatas, modelo normal (*a priori* normal-gama).

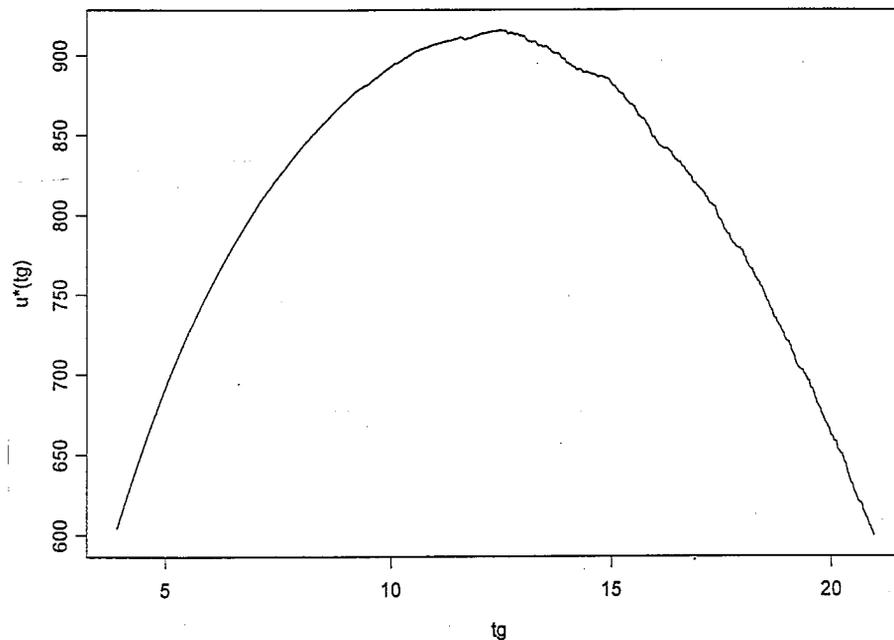


Figura 8.6: Utilidad esperada para balatas, modelo normal (*a priori* normal-gama).

En la figura 8.6 se aprecia la utilidad esperada, el tiempo de garantía que maximiza esta función es  $t_g^* = 12.489$  miles de kilómetros, que corresponde al percentil 3.5 de la distribución posterior predictiva. La función de utilidad evaluada en este punto es igual a 915.88, que sería la utilidad esperada máxima. Además en la figura 8.6, se puede apreciar que sólo en un rango pequeño al rededor de  $t_g^* = 12.489$ , se esperarían utilidades similares, por lo tanto de acuerdo a esto el tiempo de garantía por el que finalmente se opte debe ser muy cercano a este valor.

**A priori uniforme.** Utilizando una distribución uniforme para los parámetros del modelo normal, y aplicando el algoritmo Sampling-importance-resampling con  $M = 15000$  replicas, se obtuvo una muestra de tamaño 1383 de la distribución posterior predictiva, que se aprecia en el histograma de la figura 8.7. De donde se puede apreciar claramente el tiempo de vida (en miles de kilómetros) del producto. Algunos percentiles claves de esta densidad se muestra en la tabla 8.5.

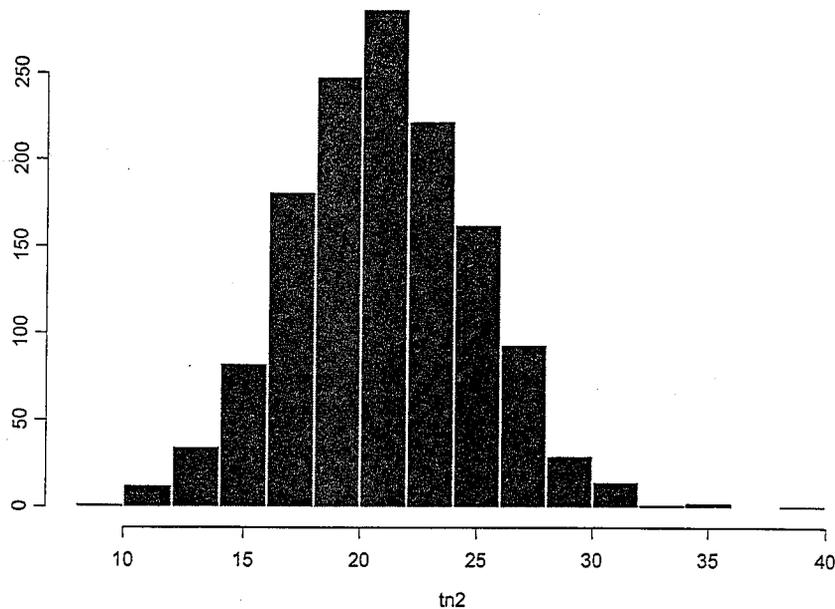


Figura 8.7: Predictiva para balatas con modelo normal (*a priori* uniforme).

Percentil	Predicción
0	9.090
1	11.917
5	14.496
10	16.084
15	17.011
20	17.664
25	18.279
30	18.908

Tabla 8.5. Percentiles de la predictiva para modelo normal (*a priori* uniforme).

En la figura 8.8 se aprecia la función de utilidad esperada, el tiempo de garantía que maximiza ésta función es  $t_g^* = 13.19$  miles de kilómetros, que prácticamente corresponde al percentil 3 de la distribución posterior predictiva. La función de utilidad evaluada en este punto es igual a 932.65, que sería la utilidad esperada máxima. Además en la figura 8.8, se puede apreciar que sólo en un rango pequeño al rededor de  $t_g^* = 13.19$ , se esperarían utilidades similares, por lo tanto de acuerdo a esto el tiempo de garantía por el que finalmente se opte debe ser muy cercano a este valor.

Discusión. En las dos subsecciones previas hemos decidido el tiempo óptimo de garantía para las balatas considerando dos modelos y dos tipos de distribuciones *a priori*. Los aspectos más relevantes de esto se sintetizan en la tabla 8.6. De esta tabla se aprecia que

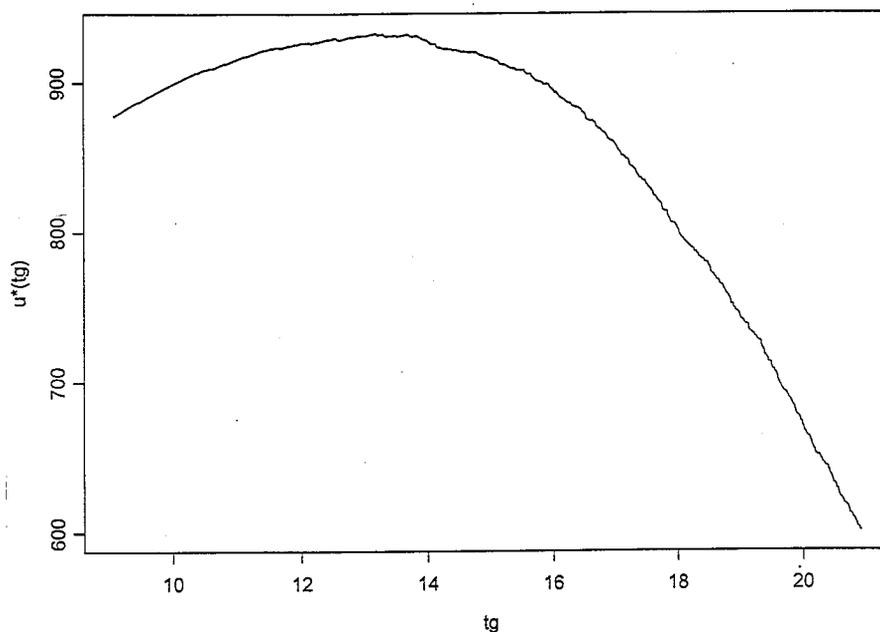


Figura 8.8: Utilidad esperada para balatas modelo normal (*a priori* uniforme).

el tiempo óptimo de garantía que se logra en todos los casos es relativamente similar, una pequeña excepción se logra con el modelo normal y *a priori* uniforme, donde el óptimo es un poco mayor que el resto. Esta excepción se minimiza más si se observan las gráficas para la función de utilidad esperada (figuras 8.2, 8.4, 8.6 y 8.8), donde se aprecia que en torno a sus óptimos estas funciones tienen una caída lenta. Por lo tanto a partir de los cuatro casos se puede concluir un tiempo óptimo de garantía similar entre 12.0 y 14.5. Adicionalmente en todos los casos, los tiempos óptimos de garantía corresponde a percentiles pequeños de la distribución posterior predictiva (entre 3 y 5.5).

<i>a priori</i>	Modelo			
	Weibull		Normal	
	Normal-gama	Uniforme	Normal-gama	Uniforme
Tiempo óptimo de garantía	12.627	12.82	12.489	13.19
Percentil	5.5	3	3.5	3
Utilidad esperada máxima	899.13	922	915.88	932.65

Tabla 8.6. Resumen de resultados para la elección del tiempo óptimo de garantía

Otro aspecto a notar para ambos modelos es que cuando se utiliza una *a priori* uniforme el tiempo de garantía tiende a ser ligeramente más alto; algo similar ocurre respecto a la utilidad esperada, que con esta *a priori* tiende a ser ligeramente mayor.

Ambos aspectos se deben a la influencia de la priori uniforme sobre la forma de la distribución posterior predictiva.

El mismo tipo de percentiles que se obtuvieron (bajos como es de suponerse) y la similitud de los resultados para ambas *a priori* y los dos modelos utilizados, habla en favor de la metodología que se ha propuesto. Por lo tanto, dados los elementos que la función de utilidad tomó en cuenta y los datos sobre la duración del producto, se tienen elementos objetivos para decidir el tiempo de garantía. Los estudios realizados apoyan la idea de aumentar la garantía del producto de 10.00 como es actualmente a un valor cercano a 13.00 mil kilómetros.

Por lo que muestra el ejemplo y por los argumentos que respaldan la propuesta que hemos hecho, consideramos que se tiene una metodología consistente y cuantitativa para decidir el tiempo óptimo de garantía. Esto evitaría la recomendación vaga: seleccionar el tiempo de garantía de acuerdo a un percentil bajo.

## CAPITULO 9

### CONCLUSIONES GENERALES Y ESTUDIOS FUTUROS

A continuación destacamos los aspectos más relevantes del trabajo así como algunos aspectos que consideramos de interés para estudios futuros.

#### 9.1 Especificación de Modelos y distribuciones *a priori*

Con base en lo hecho para especificación de modelos y de las distribuciones *a priori* en los capítulos anteriores, podemos destacar las siguientes conclusiones:

- Apoyarse en el procedimiento Bayesiano que hemos propuesto para especificar el modelo para datos de confiabilidad permite utilizar información previa que posee el usuario y toda la información generada durante el experimento de confiabilidad. Hemos visto a través de los ejemplos analizados y el estudio de simulación, que la metodología que proponemos tiene un buen poder discriminatorio, por lo que complementa adecuadamente la subjetividad de la gráfica de probabilidad.
- Instrumentar la metodología que proponemos considerando dos tipos de distribuciones *a priori*: normal-gama y uniforme, la primera más informativa que la segunda, proporciona doble validación, lo cual ayuda a mejorar la especificación del modelo en confiabilidad. De hecho en los ejemplos que se han presentado, independientemente de cuál distribución *a priori* se utilizó, el modelo con mayor probabilidad posterior fue el mismo. Adicionalmente se vio a través de las aplicaciones de la metodología vistas en el capítulo 6, que la metodología tiene una buena robustez a la elección particular de los intervalos iniciales, para los tamaños de muestra considerados. Siendo un poco menos robusta en el caso de *a priori* normal-gama.
- Dada sólo la información *a priori* para el tiempo de vida en términos de intervalos para los cuantiles del tiempo de vida, y estos convertidos de acuerdo con nuestra propuesta en intervalos para la media y la desviación estándar, generamos una metodología consistente de incorporar esta información para definir dos tipos de *a priori* (normal-gama y uniforme) para los parámetros de cinco modelos (normal, lognormal, valor extremo, Weibull y exponencial). Consideramos que esta forma de incorporar la información inicial tiene ventajas sobre las propuestas Bayesianas similares en el contexto de selección de modelos; ya que no se utiliza la parte de la información proporcionada por los datos en muestras de ensayo para definir las *a priori* (Berger y Pericchi, 1996; Pérez y Berger, 2002), no recurrimos a muestras imaginarias (Spiegelhalter y Smith, 1982) sino a información dada por el usuario, y tampoco es necesario tener datos históricos similares (Ibrahim, 2001, cap. 6). Además esta metodología puede utilizarse para cualquier otro propósito de inferencia Bayesiana para estos modelos de confiabilidad.
- Los intervalos para los cuantiles es la información *a priori* en que se basa la metodología que hemos propuesto en el capítulo 5. Esto coincide con la tendencia Bayesiana de que la especificación de la *a priori* debe partir de información disponible sobre las observables, es decir, sobre el tiempo de vida  $X$ , más que

de los no observables -los parámetros de los modelos de tiempo de vida. Adicionalmente, el supuesto de que el usuario puede aportar información *a priori* sobre *intervalos* para ciertos cuantiles del tiempo de vida, es razonable y menos exigente respecto a las metodologías para definir distribuciones *a priori* que requieren tener datos históricos (Ver Ibrahim, 2001, cap. 6) o las que proponen utilizar parte de la información para utilizarla como muestras de ensayo (Berger y Pericchi, 1996; Pérez y Berger, 2002). También creemos que al generar la información en intervalos se protege contra sub o sobreestimación, cosa que no ocurre cuando se parte de un sólo valor para los cuantiles, como lo hacen para el modelo exponencial Wolfson (1995) y Percy (2002).

- La metodología que hemos propuesta para especificar los parámetros de las distribuciones *a priori* para los parámetros de los diferentes modelos considerados, trabaja directamente con los modelos y fue factible aplicarla para los modelos más usuales en confiabilidad. Este enfoque de partir directamente del modelo, aunque no muy ortodoxo, demostramos que funciona adecuadamente para casos estándar. Esto contrasta con el enfoque, considerado más ortodoxo, de partir de la distribución *a priori* predictiva, que hasta donde se conoce sólo tiene solución para el caso más simple del modelo exponencial, y no así para el resto de los modelos (Percy, 2002).
- Con muestras completas y *a priori* normal-gama obtuvimos formas analíticas para la verosimilitud integrada de tres de los modelos (normal, lognormal y exponencial). Para el modelo exponencial esto también aplica para muestras con censura. Lo cual elimina las complicaciones tradicionales para el cálculo de la verosimilitud integrada. En el resto de los casos, generamos dos algoritmos para obtener la verosimilitud integrada: por simulación Monte Carlo y por la aproximación de Laplace.
- La metodología que hemos desarrollado para especificación de modelos es general: funciona para cualquier tamaño de muestra y tipo de censura, esto contrasta con las metodologías tipo bondad de ajuste, como se señaló en la sección 3.1, que depende de aproximaciones asintóticas, de cada modelo y del tipo de censura; cada variación de lugar a un caso diferente. En este sentido, aunque en el trabajo sólo se han considerado los modelos y las censuras más usuales (I, II y aleatoria), la metodología puede generalizarse fácilmente a otros modelos y tipos de censura, como por ejemplo censura por la izquierda y por intervalo. Esto es una ventaja respecto al enfoque de bondad de ajuste.
- Independientemente de que entre los modelos considerados se tenga el verdadero o no, la metodología propuesta jerarquiza cuál de ellos está más cerca del modelo correcto, por lo que se tiene la alternativa de quedarse con él, o bien proponer otros que sean más similares a éste. Esto contrasta con la especie de indefinición que puede ocurrir en el enfoque clásico de bondad de ajuste, cuando todos los modelos son rechazados o cuando más de uno no es rechazado.

- A través de la probabilidad *a priori* para cada modelo  $-\Pr(M_i)-$ , se puede establecer una preferencia previa por ciertos tipos de modelos, ya sea con base en conocimiento inicial o considerando facilidad o ventajas de uso. Esto podría utilizarse formalmente en el cálculo de la probabilidad posterior y así ayudar en la especificación.
- Como se mencionó al final del capítulo 7, la metodología de especificación de modelos que hemos propuesto podría utilizarse para implementar, en el contexto de confiabilidad, el procedimiento que se conoce como ponderación de modelos (Hoeting et al. 1999), en donde se promedia las distribuciones posteriores de cada uno de los modelos considerados, ponderados por las probabilidades posteriores del modelo:

$$\Pr(q|\mathbf{X}) = \sum_{i=1}^m \Pr(q | M_i, \mathbf{X}) \Pr(M_i|\mathbf{X})$$

donde  $q$  es la cantidad de interés sobre lo que se quiere hacer inferencias. Se sabe este procedimiento proporciona una mejor predicción que cualquiera de los modelos individuales (Medigan y Raftery, 1994). Aunque tiene la desventaja de una mayor complicación en cálculos.

Algunas investigaciones futuras que consideramos de interés en el contexto de especificación de modelos son:

- Extender las metodologías de especificación de modelos y la de definición de distribución *a priori* a otros modelos de uso en confiabilidad (ver Meeker y Escobar, 1998, cap. 5). Esto implicaría establecer sus distribuciones *a priori* y valorar la posibilidad de extender la metodología descrita en el capítulo 5 para éstas. Además resolver el correspondiente problema del cálculo de la verosimilitud integrada.
- Ampliar el trabajo realizado, considerando censuras por la izquierda y por intervalo.
- Evaluar otros métodos para calcular la verosimilitud integrada. Lo hecho para la Aproximación de Laplace serviría de base para aplicar otro tipo de aproximaciones, cuando se esté lejos de los supuestos del método de Laplace, ver por ejemplo (Genz y Kass, 1997).
- Abordar desde la perspectiva Bayesiana el problema de especificación de modelos en confiabilidad en el contexto de pruebas de vida acelerada. Ahí se tendría que generar propuestas para identificar tanto el modelo probabilístico como el modelo físico de falla.

## 9.2 Tiempo de Garantía

- La elección del tiempo de garantía de un producto es una decisión importante que tiene varias repercusiones, y que por lo tanto la estadística no puede limitarse a señalar: "seleccionar el tiempo de garantía de acuerdo a un percentil bajo de la distribución del tiempo de vida del producto".
- Generamos una propuesta consistente y cuantitativa para la elección del tiempo de garantía, que toma en cuenta la confiabilidad del producto, la competitividad del esquema de garantía a los ojos del consumidor (mercadotecnia), imagen de la empresa en caso de que el producto falle dentro del período de garantía y los costos en los que incurre la empresa para cumplir con la garantía prometida.
- Esta propuesta la respaldamos con un ejemplo, en el cual se utilizaron dos modelos y los dos tipos de *a priori*. La similitud en los resultados obtenidos hablan en favor de la metodología.
- La metodología propuesta supone que el usuario es capaz de proponer intervalos iniciales donde espera que ocurran cuantiles de la distribución del tiempo de vida, para a partir de ahí estimar intervalos para la media y la desviación estándar. Esto intervalos permiten especificar dos tipos de distribuciones *a priori*: uniforme y normal-gama. Por lo tanto la decisión podrá tomarse con base en ambas distribuciones *a priori*, o eligiendo aquella que refleje mejor el nivel de conocimiento inicial que el usuario tiene sobre el tiempo de vida del producto.
- La metodología que se propuso para elegir el tiempo de garantía, contempla cinco de los modelos más usuales en confiabilidad: normal, lognormal, valor extremo, Weibull y exponencial; y contempla datos completos y con cesura tipo I, II y aleatoria. Pero además es fácilmente generalizable a otro tipo de censura (intervalo, por ejemplo).
- Para el modelo exponencial derivamos la forma analítica de la función de utilidad esperada, por lo que en este caso encontrar el tiempo de garantía que maximiza la utilidad esperada se logra con rutinas sencillas de optimización de una función univariada. Mientras que para el modelo normal y lognormal, bajo muestreo sin censura y *a priori* normal-gama, proponemos obtener la utilidad esperada por simulación directa de la distribución posterior predictiva (*t*-Student y *logt*-Student). Para el resto de los casos generamos un algoritmo basado en el método *sampling-importance-resampling*.

Algunas investigaciones futuras que consideramos de interés en el contexto de decidir el esquema óptimo de garantía de un producto son:

- Ampliar la metodología, donde se incluyan otros modelos usuales en confiabilidad y otros tipos de censura (por la izquierda y por intervalo).
- Suponer otro tipo de información inicial y otras formas de definir los parámetros de las distribuciones *a priori*.

- Abordar el problema de diseño asociado a la decisión de la garantía: lo que hemos hecho es abordar el problema de decisión respecto a la inferencia sobre el tiempo de garantía. Sin embargo, para llegar a ello primero es necesario diseñar y realizar un experimento para obtener la información para tomar la decisión; por lo tanto es necesario diseñar el experimento que minimice los costos totales, y esto incluye tamaño de muestra y tipo y nivel de censura. Estos costos totales se desprenden de la función de utilidad para decidir el tiempo de garantía y de los costos de la experimentación (ver Berger, 1985, cap. 7)
- Abordar desde la perspectiva Bayesiana el problema de elegir el tiempo de garantía con base en datos de pruebas de vida acelerada.

## APENDICE DEMOSTRACIONES

En este apéndice se dan demostraciones y detalles matemáticos complementarios de algunos de los resultados obtenidos a lo largo del trabajo.

### 1.1 Obtención Analítica de la Verosimilitud Integrada

En esta sección se obtiene la VI para los casos que ésta tiene forma analítica.

DEMOSTRACION del teorema 4.1. Sabemos que la verosimilitud integrada (VI) está dada por

$$f(\mathbf{X}|M_i, \alpha_i, \beta_i, m_i, k_i) = \int_0^\infty \int_{-\infty}^\infty L(\mathbf{X}|\mu, \tau, M_i) \pi(\mu, \tau | \alpha_i, \beta_i, m_i, k_i) d\mu d\tau$$

a) **Caso normal:** sustituyendo las expresiones para la verosimilitud normal (4.2) y la densidad *a priori* normal gama (4.4), y combinando términos obtenemos que la VI está dada por

$$\frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \int_0^\infty \int_{-\infty}^\infty \frac{(k_1)^{1/2} \tau^{\frac{n-1}{2} + \alpha_1}}{(2\pi)^{(n+1)/2}} \exp \left[ -\frac{k_1 \tau}{2} (\mu - m_1)^2 - \frac{\tau}{2} n (\mu - \bar{x}_n)^2 - \beta_1 \tau - \frac{\tau s_n}{2} \right] d\mu d\tau$$

De aquí, desarrollando los términos que involucran a  $\mu$ , combinando y completando un binomio cuadrado obtenemos que la VI es

$$\frac{\beta_1^{\alpha_1} (k_1)^{1/2}}{\Gamma(\alpha_1)} \int_0^\infty \int_{-\infty}^\infty \frac{\tau^{\frac{n-1}{2} + \alpha_1}}{(2\pi)^{(n+1)/2}} \exp \left( -\frac{1}{2} \tau (k_1 + n) \left[ \mu - \frac{(k_1 m_1 + n \bar{x}_n)}{k_1 + n} \right]^2 \right) \exp \left[ -\frac{1}{2} \tau k_1 n \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} \right] \exp \left( -\beta_1 \tau - \frac{\tau s_n}{2} \right) d\mu d\tau$$

Ahora, de la forma de las densidades normal y gama, completando en el caso normal y separando expresiones obtenemos que la VI es

$$\frac{\beta_1^{\alpha_1} (k_1)^{1/2}}{(k_1 + n)^{1/2} \Gamma(\alpha_1) (2\pi)^{n/2}} \int_0^\infty \int_{-\infty}^\infty \tau^{\frac{n-2}{2} + \alpha_1} \exp \left[ - \left( \frac{1}{2} k_1 n \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} + \beta_1 + \frac{s_n}{2} \right) \tau \right] \left( \frac{\tau (k_1 + n)}{2\pi} \right)^{1/2} \exp \left( -\frac{1}{2} \tau (k_1 + n) \left[ \mu - \frac{(k_1 m_1 + n \bar{x}_n)}{k_1 + n} \right]^2 \right) d\mu d\tau$$

Los dos últimos términos de esta expresión corresponden a una densidad normal para  $\mu$ , que al estar siendo integrada sobre todo el espacio paramétrico, esa integral es igual a uno. Por ello, integrando, completando una densidad gama y haciendo  $h_1 = \frac{1}{2} k_1 n \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} + \beta_1 + \frac{s_n}{2}$ , obtenemos que la VI es

$$\frac{\beta_1^{\alpha_1} \Gamma(\frac{n}{2} + \alpha_1) (k_1)^{1/2}}{(k_1 + n)^{1/2} \Gamma(\alpha_1) (2\pi)^{n/2} h_1^{\frac{n}{2} + \alpha_1}} \int_0^\infty \frac{h_1^{\frac{n}{2} + \alpha_1}}{\Gamma(\frac{n}{2} + \alpha_1)} \tau^{\frac{n}{2} + \alpha_1 - 1} \exp[-h_1 \tau] d\tau$$

La expresión dentro del integrando corresponde a una densidad gama para  $\tau$ , que al estar siendo integrada sobre todo el espacio paramétrico es igual a uno. Por ello, finalmente obtenemos que la verosimilitud integrada para el modelo  $M_1$  esta dada por

$$\frac{\beta_1^{\alpha_1} \Gamma(\frac{n}{2} + \alpha_1) (k_1)^{1/2}}{(k_1 + n)^{1/2} \Gamma(\alpha_1) (2\pi)^{n/2} h_1^{\frac{n}{2} + \alpha_1}}$$

que es lo que queríamos demostrar.

b) Para el modelo lognormal. Sustituyendo las expresiones para la verosimilitud lognormal (4.7) y la densidad *a priori* normal gama (4.4), y combinando términos obtenemos que la VI para  $f(\mathbf{X}|M_2, \alpha_2, \beta_2, m_2, k_2)$  está dada por

$$\frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2) \prod_{i=1}^n x_i} \int_0^\infty \int_{-\infty}^\infty \frac{(k_2)^{1/2} \tau^{\frac{n-1}{2} + \alpha_2}}{(2\pi)^{(n+1)/2}} \exp \left[ -\frac{k_2 \tau}{2} (\mu - m_2)^2 - \frac{\tau}{2} n (\mu - \bar{w}_n)^2 \right] \exp \left( -\beta_2 \tau - \frac{\tau u_n}{2} \right) d\mu d\tau$$

De aquí, desarrollando los términos que involucran a  $\mu$ , combinando y completando un binomio cuadrado obtenemos que la VI es

$$\frac{\beta_2^{\alpha_2} (k_2)^{1/2}}{\Gamma(\alpha_2) \prod_{i=1}^n x_i} \int_0^\infty \int_{-\infty}^\infty \frac{\tau^{\frac{n-1}{2} + \alpha_2}}{(2\pi)^{(n+1)/2}} \exp \left( -\frac{1}{2} \tau (k_2 + n) \left[ \mu - \frac{(k_2 m_2 + n \bar{w}_n)}{k_2 + n} \right]^2 \right) \exp \left[ -\frac{1}{2} \tau k_2 n \frac{(m_2 - \bar{w}_n)^2}{k_2 + n} \right] \exp \left( -\beta_2 \tau - \frac{\tau u_n}{2} \right) d\mu d\tau$$

Ahora, reescribiendo el integrando en la forma de las densidades normal y gama, completando en el caso normal y separando expresiones obtenemos que la VI es

$$\frac{(k_2 + n)^{-1/2} \beta_2^{\alpha_2} (k_2)^{1/2}}{\Gamma(\alpha_2) (2\pi)^{n/2} \prod_{i=1}^n x_i} \int_0^\infty \int_{-\infty}^\infty \tau^{\frac{n-2}{2} + \alpha_2} \exp \left[ -\left( \frac{1}{2} k_2 n \frac{(m_2 - \bar{w}_n)^2}{k_2 + n} + \beta_2 + \frac{u_n}{2} \right) \tau \right] \left( \frac{\tau (k_2 + n)}{2\pi} \right)^{1/2} \exp \left( -\frac{1}{2} \tau (k_2 + n) \left[ \mu - \frac{(k_2 m_2 + n \bar{w}_n)}{k_2 + n} \right]^2 \right) d\mu d\tau$$

Los dos últimos términos de esta expresión corresponden a una densidad normal para  $\mu$ , que al estar siendo integrada sobre todo el espacio paramétrico, esa integral es igual a uno. Por ello, integrando, completando una densidad gama y haciendo  $h_2 = \frac{1}{2} k_2 n \frac{(m_2 - \bar{w})^2}{k_2 + n} + \beta_2 + \frac{u_n}{2}$ , obtenemos que la VI es

$$\frac{\beta_2^{\alpha_2} \Gamma(\frac{n}{2} + \alpha_2) (k_2)^{1/2}}{(k_2 + n)^{1/2} \Gamma(\alpha_2) (2\pi)^{n/2} h_2^{\frac{n}{2} + \alpha_2} \prod_{i=1}^n x_i} \int_0^\infty \frac{h_2^{\frac{n}{2} + \alpha_2}}{\Gamma(\frac{n}{2} + \alpha_2)} \tau^{\frac{n}{2} + \alpha_2 - 1} \exp[-h_2 \tau] d\tau$$

Esta expresión dentro del integrando corresponde a una densidad gama para  $\tau$ , que al estar siendo integrada sobre todo el espacio paramétrico es igual a uno. Por ello, finalmente obtenemos que la verosimilitud integrada para el modelo  $M_2$  esta dada por

$$\frac{\beta_2^{\alpha_2} \Gamma(\frac{n}{2} + \alpha_2) (k_2)^{1/2}}{(k_2 + n)^{1/2} \Gamma(\alpha_2) (2\pi)^{n/2} h_2^{\frac{n}{2} + \alpha_2} \prod_{i=1}^n x_i}$$

que es lo que queríamos demostrar. ■

#### DEMOSTRACION del teorema 4.2.

a) **A priori gama.** Dada la expresión (4.15) para la verosimilitud de un modelo exponencial y utilizando una distribución gama( $\alpha_5, \beta_5$ ) como distribución inicial para  $\theta$ , entonces la verosimilitud integrada está dada por

$$f(\mathbf{X}|M_5, \gamma) = \int_0^\infty \theta^n \exp \left[ -\theta \left( \sum_{i=1}^n x_i - n\gamma \right) \right] \frac{\beta_5^{\alpha_5}}{\Gamma(\alpha_5)} \theta^{\alpha_5-1} \exp(-\beta_5 \theta) d\theta$$

Rearreglando términos y completando una densidad gama, incluyendo sus términos constantes, obtenemos que la VI es

$$\frac{\beta_5^{\alpha_5}}{\Gamma(\alpha_5)} \frac{\Gamma(n + \alpha_5)}{s_5^{n+\alpha_5}} \int_0^\infty \frac{b_1^{n+\alpha_5}}{\Gamma(n + \alpha_5)} \theta^{n+\alpha_5-1} \exp(-s_5 \theta) d\theta$$

donde  $s_5 = \sum_{i=1}^n x_i + \beta_5 - n\gamma$ . Como el integrando de esta expresión corresponde a una densidad gama y la integración se está haciendo sobre todo el espacio paramétrico, entonces la integral es igual a uno. De aquí que al integrar se obtenga que la VI es

$$f(\mathbf{X}|M_5, \gamma) = \frac{\beta_5^{\alpha_5}}{\Gamma(\alpha_5)} \frac{\Gamma(n + \alpha_5)}{s_5^{n+\alpha_5}}$$

que es lo que queríamos demostrar.

b) **A priori uniforme.** Dada la expresión (4.15) para la verosimilitud de un modelo exponencial y utilizando una distribución uniforme( $a_5, b_5$ ) como distribución inicial para  $\theta$ , entonces la VI está dada por

$$\int_{a_5}^{b_5} u_5 \theta^n \exp \left[ -\theta \left( \sum_{i=1}^n x_i - n\gamma \right) \right] d\theta$$

Rearreglando términos y completando una densidad gama, incluyendo sus términos constantes, obtenemos que la VI es

$$\frac{u_5 \Gamma(n + 1)}{v^{n+1}} \int_{a_5}^{b_5} \frac{v^{n+1}}{\Gamma(n + 1)} \theta^n \exp[-v\theta] d\theta$$

con  $v = \sum_{i=1}^n x_i - n\gamma$ . Como el integrando de esta expresión corresponde a una densidad gama, entonces integrando se obtiene que la VI está dada por

$$\frac{u_5 \Gamma(n + 1)}{v^{n+1}} [G(b_5; n + 1, v) - G(a_5; n + 1, v)]$$

donde  $G(\cdot; n + 1, v)$  es la distribución gama acumulada con parámetros  $(n + 1, v)$ . Esta expresión es la que queríamos demostrar. ■

## 1.2 Densidad Posterior Predictiva

En esta sección se obtiene, cuando existe, la forma analítica para los modelos considerados en este trabajo.

DEMOSTRACION del teorema 4.3. De la definición de densidad posterior predictiva (2.5), sabemos que ésta se obtiene con

$$f(y|\mathbf{X}) = \int_{\Theta} f(y|\theta)\pi(\theta|\mathbf{X})d\theta$$

por lo tanto para demostrar los dos incisos del teorema será necesario considerar la densidad posterior para  $\theta$ , en cada caso. La cual, en general, es proporcional a

$$\pi(\theta|\mathbf{X}) \propto L(\mathbf{X}|\theta)\pi(\theta)$$

a) Para el modelo normal. Sustituyendo la verosimilitud para un modelo normal (4.2), la densidad *a priori* para sus parámetros (4.4), la expresión para el modelo normal (4.1) y combinando términos, obtenemos que

$$f(y|\mathbf{X}) \propto \int_0^{\infty} \int_{-\infty}^{\infty} \tau^{\frac{n}{2}+\alpha_1} \exp \left[ -\frac{\tau}{2}(y-\mu)^2 - \frac{k_1\tau}{2}(\mu-m_1)^2 - \frac{\tau}{2}n(\mu-\bar{x}_n)^2 \right] \exp \left( -\beta_1\tau - \frac{\tau s_n}{2} \right) d\mu d\tau$$

De aquí, desarrollando los términos que involucran a  $\mu$ , combinando y completando un binomio cuadrado obtenemos

$$\propto \int_0^{\infty} \int_{-\infty}^{\infty} \tau^{\frac{n}{2}+\alpha_1} \exp \left( -\frac{1}{2}\tau(k_1+n+1) \left[ \mu - \frac{(k_1m_1+n\bar{x}_n+y)}{k_1+n+1} \right]^2 \right) \exp \left[ -\frac{1}{2}\tau \frac{n(y-\bar{x}_n)^2 + nk_1(m_1-\bar{x}_n)^2 + k_1(y-m_1)^2}{k_1+n+1} \right] \exp \left( -\beta_1\tau - \frac{\tau s_n}{2} \right) d\mu d\tau$$

Ahora pensando en la forma de las densidades normal y gama, completando en el caso normal y separando expresiones obtenemos

$$\propto \int_0^{\infty} \int_{-\infty}^{\infty} \exp \left[ -\left( \frac{1}{2} \frac{n(y-\bar{x}_n)^2 + nk_1(m_1-\bar{x}_n)^2 + k_1(y-m_1)^2}{k_1+n+1} + \beta_1 + \frac{s_n}{2} \right) \tau \right] \tau^{\frac{n-1}{2}+\alpha_1} \left( \frac{\tau(k_1+n+1)}{2\pi} \right)^{1/2} \exp \left[ -\frac{1}{2}\tau(k_1+n+1) \left[ \mu - \frac{(k_1m_1+n\bar{x}_n+y)}{k_1+n+1} \right]^2 \right] d\mu d\tau$$

Los dos últimos términos tenemos de esta expresión corresponden a una densidad normal para  $\mu$ , que al estar siendo integrada sobre todo el espacio paramétrico, esa integral es igual a uno. De aquí que

$$f(y|\mathbf{X}) \propto \frac{1}{h^{\frac{n+1}{2} + \alpha_1}} \int_0^\infty \frac{h^{\frac{n+1}{2} + \alpha_1}}{\Gamma(\frac{n+1}{2} + \alpha_1)} \tau^{\frac{n+1}{2} + \alpha_1 - 1} \exp[-h\tau] d\tau$$

donde  $h = \frac{\frac{1}{2}n(y - \bar{x}_n)^2 + nk_1(m_1 - \bar{x}_n)^2 + k_1(y - m_1)^2}{k_1 + n + 1} + \beta_1 + \frac{s_n}{2}$ . La expresión dentro del integrando corresponde a una densidad gama para  $\tau$ , que al estar siendo integrada sobre todo el espacio paramétrico es igual a uno. Por ello

$$f(y|\mathbf{X}) \propto h^{-\frac{(n+1+2\alpha_1)}{2}}$$

Ahora lo que sigue es encontrar la forma de esta distribución. Para ello, juntando los términos de  $h_1$  que involucran a  $y$ , y completando el binomio cuadrado, obtenemos que

$$h = \frac{1}{2} \frac{(n + k_1)}{(k_1 + n + 1)} \left( y - \frac{k_1 m_1 + n \bar{x}_n}{n + k_1} \right)^2 + \frac{1}{2} n k_1 \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} + \beta_1 + \frac{s_n}{2}$$

y a su vez toma la siguiente forma

$$h = \frac{l_t}{(2\alpha_1 + n)} (y - m_t)^2 + 1$$

donde  $m_t = \frac{k_1 m_1 + n \bar{x}_n}{n + k_1}$  y  $l_t = \left[ \frac{1}{2} n k_1 \frac{(m_1 - \bar{x}_n)^2}{k_1 + n} + \beta_1 + \frac{s_n}{2} \right]^{-1} \frac{(\alpha_1 + n/2)(n + k_1)}{(k_1 + n + 1)}$ , por lo tanto

$$f(y|\mathbf{X}) \propto \left[ \frac{l_t}{(2\alpha_1 + n)} (y - m_t)^2 + 1 \right]^{-\frac{(n+2\alpha_1+1)}{2}}$$

esta densidad es del tipo  $t$ -Student (ver Bernardo y Smith, 1994, pag. 122), con los siguientes parámetros: localización (media)  $m_t$ , precisión  $l_t$ , y  $2\alpha_1 + n$  grados de libertad. De aquí que la densidad posterior predictiva para el modelo normal utilizando una distribución *a priori* normal-gama( $\alpha_1, \beta_1, m_1, k_1$ ), esté dada por

$$f_1(y|\mathbf{X}) = \frac{\Gamma\left(\frac{2\alpha_1 + n + 1}{2}\right)}{\Gamma\left(\frac{2\alpha_1 + n}{2}\right) \Gamma\left(\frac{1}{2}\right)} \left( \frac{l_t}{2\alpha_1 + n} \right)^{1/2} \left[ \frac{l_t}{(2\alpha_1 + n)} (y - m_t)^2 + 1 \right]^{-\frac{(n+2\alpha_1+1)}{2}}$$

que era lo que queríamos demostrar.

**b) Para el modelo lognormal.** Sustituyendo la verosimilitud para un modelo normal (4.7), la densidad *a priori* para sus parámetros similar a (4.4), la expresión para el modelo lognormal (4.6) y combinando términos, obtenemos que

$$f_2(y|\mathbf{X}) \propto \int_0^\infty \int_{-\infty}^\infty \frac{\tau^{\frac{n}{2} + \alpha_2}}{y} \exp \left[ -\frac{\tau}{2} [\log(y) - \mu]^2 - \frac{k_2 \tau}{2} (\mu - m_2)^2 - \frac{\tau}{2} n (\mu - \bar{w}_n)^2 \right] \exp(-\beta_2 \tau - \frac{\tau u_n}{2}) d\mu d\tau$$

De aquí, desarrollando los términos que involucran a  $\mu$ , combinando y completando un binomio cuadrado obtenemos que  $f_2(y|\mathbf{X})$  es proporcional a

$$\int_0^\infty \int_{-\infty}^\infty \frac{\tau^{\frac{n}{2}+\alpha_1}}{y} \exp \left[ -\frac{1}{2}\tau(k_2+n+1) \left[ \mu - \frac{(k_2 m_2 + n\bar{w}_n + \log(y))}{k_2+n+1} \right]^2 \right] \\ \exp \left[ -\frac{1}{2}\tau \frac{n(\log(y) - \bar{w}_n)^2 + nk_2(m_2 - \bar{w}_n)^2 + k_2(\log(y) - m_2)^2}{k_2+n+1} \right] \exp \left( -\beta_2\tau - \frac{\tau u_n}{2} \right) d\mu d\tau$$

Ahora considerando la forma de las densidades normal y gama, completando en el caso normal y separando expresiones obtenemos que  $f_2(y|\mathbf{X})$  es proporcional a

$$\int_0^\infty \int_{-\infty}^\infty \exp \left[ -\left( \frac{1}{2} \frac{n(\log(y) - \bar{w}_n)^2 + nk_2(m_2 - \bar{w}_n)^2 + k_2(\log(y) - m_2)^2}{k_2+n+1} + \beta_2 + \frac{u_n}{2} \right) \tau \right] \\ \frac{\tau^{\frac{n-1}{2}+\alpha_2}}{y} \left( \frac{\tau(k_2+n+1)}{2\pi} \right)^{1/2} \exp \left[ -\frac{1}{2}\tau(k_2+n+1) \left[ \mu - \frac{(k_2 m_2 + n\bar{w}_n + \log(y))}{k_2+n+1} \right]^2 \right] d\mu d\tau$$

Los dos últimos términos tenemos de esta expresión corresponden a una densidad normal para  $\mu$ , que al estar siendo integrada sobre todo el espacio paramétrico, esa integral es igual a uno. De aquí que

$$f_2(y|\mathbf{X}) \propto \frac{1}{y t^{\frac{n+1}{2}+\alpha_2}} \int_0^\infty \frac{t^{\frac{n+1}{2}+\alpha_2}}{\Gamma(\frac{n+1}{2} + \alpha_2)} \tau^{\frac{n+1}{2}+\alpha_2-1} \exp[-t\tau] d\tau$$

donde  $t = \frac{1}{2} \frac{n(\log(y) - \bar{w}_n)^2 + nk_2(m_2 - \bar{w}_n)^2 + k_2(\log(y) - m_2)^2}{k_2+n+1} + \beta_2 + \frac{u_n}{2}$ . La expresión dentro del integrando corresponde a una densidad gama para  $\tau$ , que al estar siendo integrada sobre todo el espacio paramétrico es igual a uno. Por ello

$$f_2(y|\mathbf{X}) \propto y^{-1} t^{-\frac{(n+1+2\alpha_1)}{2}}$$

Ahora lo que sigue es encontrar la forma de esta distribución. Para ello, juntando los términos de  $t$  que involucran a  $\log(y)$ , y completando el binomio cuadrado, obtenemos que

$$t = \frac{1}{2} \frac{(n+k_2)}{(k_2+n+1)} \left[ \log(y) - \frac{k_2 m_2 + n\bar{w}_n}{n+k_2} \right]^2 + \frac{1}{2} n k_2 \frac{(m_2 - \bar{w}_n)^2}{k_2+n} + \beta_2 + \frac{u_n}{2}$$

que a su vez toma la siguiente forma

$$t = \frac{l}{(2\alpha_2+n)} [\log(y) - m_l]^2 + 1$$

donde  $m_l = \frac{k_2 m_2 + n\bar{w}_n}{n+k_2}$  y  $l = \left[ \frac{1}{2} n k_2 \frac{(m_2 - \bar{w}_n)^2}{k_2+n} + \beta_2 + \frac{u_n}{2} \right]^{-1} \frac{(\alpha_2+n/2)(n+k_2)}{(k_2+n+1)}$ , por lo tanto

$$f_2(y|\mathbf{X}) \propto \frac{1}{y} \left[ \frac{l}{(2\alpha_2 + n)} (\log(y) - m_l)^2 + 1 \right]^{-\frac{(n+2\alpha_2+1)}{2}}$$

esta densidad es del tipo Logt-Student (Percy, 2002), con los siguientes parámetros: localización (media)  $m_l$ , precisión  $l$ , y  $2\alpha_2 + n$  grados de libertad. De aquí que la densidad posterior predictiva para el modelo lognormal utilizando una distribución *a priori* normal-gama( $\alpha_2, \beta_2, m_2, k_2$ ), esté dada por

$$f_2(y|\mathbf{X}) = \frac{\Gamma\left(\frac{2\alpha_2+n+1}{2}\right)}{y\Gamma\left(\frac{2\alpha_2+n}{2}\right)\Gamma\left(\frac{1}{2}\right)} \left(\frac{l}{2\alpha_2+n}\right)^{1/2} \left[\frac{l}{(2\alpha_2+n)} [\log(y) - m_l]^2 + 1\right]^{-\frac{(n+2\alpha_2+1)}{2}}$$

que era lo que queríamos demostrar. ■

DEMOSTRACION del teorema 4.4.

a) Utilizando una *a priori* gama( $\alpha_5, \beta_5$ ). Suponiendo conocido el parámetro  $\gamma$ , sustituyendo la verosimilitud para un modelo exponencial (4.15), y la densidad *a priori* para  $\theta$ ; obtenemos que la distribución posterior para  $\theta$  está dada por

$$\begin{aligned} \pi(\theta|\mathbf{X}) &\propto L(\mathbf{X}|\theta)\pi(\theta|\alpha_5, \beta_5) \\ &\propto \theta^{n+\alpha_5-1} \exp\left[-\theta\left(\sum x_i + \beta_5 - n\gamma\right)\right] \end{aligned}$$

Esta expresión tiene la forma de una densidad gama( $n + \alpha_5, \sum x_i + \beta_5 - n\gamma$ ). Por lo tanto la densidad posterior para  $\theta$  está dada por

$$\pi(\theta|\mathbf{X}) = \frac{v_5^{n+\alpha_5}}{\Gamma(n + \alpha_5)} \theta^{n+\alpha_5-1} \exp[-v_5\theta] \quad (1.21)$$

donde  $v_5 = \sum x_i + \beta_5 - n\gamma$ . De aquí, y dado el modelo exponencial (4.14), la densidad posterior predictiva está dada por:

$$\begin{aligned} f_5(y|\mathbf{X}) &= \int_0^\infty f(y|\theta)\pi(\theta|\mathbf{X})d\theta \\ &= \frac{v_5^{n+\alpha_5}}{\Gamma(n + \alpha_5)} \int_0^\infty \theta^{n+\alpha_5+1-1} \exp[-(v_5 + y - \gamma)\theta] d\theta \end{aligned}$$

completando una densidad gama, obtenemos

$$f_5(y|\mathbf{X}) = \frac{\Gamma(n + \alpha_5 + 1)}{(v_5 + y - \gamma)^{n+\alpha_5+1}} \int_0^\infty \frac{(v_5 + y - \gamma)^{n+\alpha_5+1}}{\Gamma(n + \alpha_5 + 1)} \theta^{n+\alpha_5+1-1} \exp[-(v_5 + y - \gamma)\theta] d\theta$$

la integral es igual a uno, por ser una densidad gama integrada sobre todo el espacio paramétrico, por lo que

$$f_5(y|\mathbf{X}) = \frac{v_5^{n+\alpha_5}}{\Gamma(n + \alpha_5)} \frac{\Gamma(n + \alpha_5 + 1)}{(v_5 + y - \gamma)^{n+\alpha_5+1}}$$

esta expresión tiene la forma de una densidad gama-gama( $n + \alpha_5, v_5, 1$ ) (Ver Bernardo y Smith, 1994, pag. 120). Que es lo que queríamos demostrar.

b) **Utilizando una *a priori* uniforme**( $a_5, b_5$ ). Dada la distribución inicial para  $\theta$ , y dada la expresión (4.15) para la verosimilitud de un modelo exponencial, entonces la densidad posterior para  $\theta$  está dada por

$$\pi(\theta | \mathbf{X}) = u_5 \theta^n \exp \left[ -\theta \left( \sum_{i=1}^n x_i - n\gamma \right) \right]$$

de aquí que la densidad posterior predictiva está dada por

$$\begin{aligned} f_5(y | \mathbf{X}) &= \int_0^\infty f(y|\theta)\pi(\theta|\mathbf{X})d\theta \\ &= u_5 \int_{a_5}^{b_5} \theta^{n+1} \exp \left[ - \left( \sum_{i=1}^n x_i - (n+1)\gamma + y \right) \theta \right] d\theta \end{aligned}$$

Rearreglando términos y completando una densidad gama, incluyendo sus términos constantes, obtenemos

$$f_5(y | \mathbf{X}) = u_5 \frac{\Gamma(n+2)}{(t_5 + y)^{n+1+1}} \int_{a_5}^{b_5} \frac{(t_5 + y)^{n+1+1}}{\Gamma(n+2)} \theta^{n+1} \exp [-(t_5 + y)\theta] d\theta$$

con  $t_5 = \sum_{i=1}^n x_i - (n+1)\gamma$ . De aquí que

$$f_5(y | \mathbf{X}) = \frac{u_5 \Gamma(n+2)}{(t_5 + y)^{n+2}} [G(b_5; n+2, t_5 + y) - G(a_5; n+2, t_5 + y)]$$

donde  $G(\cdot; n+2, t_5 + y)$  es la distribución gama acumulada con parámetros  $(n+2, t_5 + y)$ . Esto es lo que queríamos demostrar. ■

### 1.3 Resultados para la Aproximación de Laplace

En esta sección se documenta la obtención del hessiano para los diferentes modelos, y que es necesario para calcular la aproximación de Laplace para la verosimilitud integrada.

**DEMOSTRACION del teorema 4.5.** En todos los casos obtendremos el hessiano de (3.6), es decir de  $l_i(\theta_i | \alpha_i, \beta_i, m_i, k_i) = \log[L(\mathbf{X}|\theta_i, M_i)\pi(\theta_i | \alpha_i, \beta_i, m_i, k_i)]$ .

a) **Modelo normal.** Sustituyendo la verosimilitud bajo censura para el modelo normal (4.3) y la *a priori* (4.4), entonces  $l_1(\mu, \tau | \alpha_1, \beta_1, m_1, k_1)$  está dada por

$$\begin{aligned} &\log \left[ \frac{C\beta_1^{\alpha_1} (k_1)^{1/2}}{\Gamma(\alpha_1)(2\pi)^{(r+1)/2}} \right] + \left( \frac{r-1}{2} + \alpha_1 \right) \log(\tau) - \frac{k_1\tau}{2} (\mu - m_1)^2 \quad (1.32) \\ &-\frac{r\tau}{2} (\mu - \bar{x}_r)^2 - \beta_1\tau - \frac{\tau s_r}{2} + \sum_{j=1}^{n-r} \log \left( \int_{\tau^{1/2}(q_{(j)} - \mu)}^\infty \left( \frac{1}{2\pi} \right)^{\frac{1}{2}} \exp \left[ -\frac{w^2}{2} \right] dw \right) \end{aligned}$$

Derivando esta expresión respecto a  $\mu$  obtenemos

$$\frac{d}{d\mu} l_1(\mu, \tau) = -k_1 \tau (\mu - m_1) - r \tau (\mu - \bar{x}_r) + \sum_{j=1}^{n-r} \exp\left(-\frac{1}{2} \tau (\mu - q_{(j)})^2\right) \frac{\sqrt{2\tau}}{\sqrt{\pi} (1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}}(\mu - q_{(j)})))}$$

donde la función  $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$  se definió en (4.30). Si esta última expresión se deriva nuevamente respecto a  $\mu$ , se obtiene (4.28); y si se deriva respecto a  $\tau$ , se obtiene (4.32). Si se deriva (1.32) respecto a  $\tau$  se obtiene

$$\frac{d}{d\tau} l_1(\mu, \tau) = -\frac{k_1}{2} (\mu - m_1)^2 - \frac{r}{2} (\mu - \bar{x}_r)^2 - \beta_1 - \frac{s_r}{2} + \frac{1}{2} \frac{r-1+2\alpha_1}{\tau} + \sum_{j=1}^{n-r} \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{1}{2} \tau (\mu - q_{(j)})^2\right) \frac{(\mu - q_{(j)})}{(1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}}(\mu - q_{(j)})))}$$

que al volverse a derivar respecto a  $\tau$ , se obtiene (4.31). Con esto finalizamos la demostración del inciso a) del teorema.

b) **Modelo lognormal.** En este caso,  $l_2(\mu, \tau | \alpha_2, \beta_2, m_2, k_2)$  está dada por

$$\log \left[ \frac{C \beta_2^{\alpha_2} (k_2)^{1/2}}{\Gamma(\alpha_2) (2\pi)^{(\tau+1)/2}} \right] - \log \left( \prod_{i=1}^{\tau} x_{(i)} \right) + \left( \frac{r-1}{2} + \alpha_2 \right) \log(\tau) - \frac{k_2 \tau}{2} (\mu - m_2)^2 - \frac{r\tau}{2} (\mu - \bar{w}_r)^2 - \beta_2 \tau - \frac{\tau u_r}{2} + \sum_{j=1}^{n-r} \log \left( \int_{\tau^{1/2}(\log(q_{(j)}) - \mu)}^{\infty} \left( \frac{1}{2\pi} \right)^{\frac{1}{2}} \exp \left[ -\frac{y^2}{2} \right] dy \right) \quad (1.33)$$

Si derivamos esta expresión respecto a  $\mu$ , obtenemos

$$\frac{d}{d\mu} l_2(\mu, \tau) = -k_2 \tau (\mu - m_2) - r \tau (\mu - \bar{w}_r) + \sum_{j=1}^{n-r} \exp\left(-\frac{1}{2} \tau (\mu - q_{(j)})^2\right) \frac{\sqrt{2\tau}}{\sqrt{\pi} (1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}}(\mu - q_{(j)})))}$$

donde  $\operatorname{erf}(x)$  está definida en (4.30). Al derivar esta expresión nuevamente respecto a  $\mu$  se obtiene (4.33), y si se deriva respecto a  $\tau$  se llega a 4.35. Si derivamos (1.33) respecto a  $\tau$ , entonces se obtiene

$$\frac{d}{d\tau} l_2(\mu, \tau) = -\frac{k_2}{2} (\mu - m_2)^2 - \frac{r}{2} (\mu - \bar{w}_r)^2 - \beta_2 - \frac{u_r}{2} + \frac{1}{2} \frac{r-1+2\alpha_2}{\tau} + \frac{(n-r)}{\sqrt{2\pi\tau}} \exp\left(-\frac{1}{2} \tau [\mu - \log(q_{(j)})]^2\right) \frac{[\mu - \log(q_{(j)})]}{(1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}} [\mu - \log(q_{(j)})]))}$$

y al derivar ésta respecto a  $\tau$  nuevamente se obtiene (4.34). Con esto finalizamos este inciso de la demostración.

c) **Modelo valor extremo.** En este caso,  $l_3(\mu, \tau | \alpha_3, \beta_3, m_3, k_3)$  está dado por

$$\log(ce) + (r + \alpha_3 - 0.5) \log(\tau) + \tau \left( -r\mu + \sum_{i=1}^r x_{(i)} \right) - \beta_3 \tau - \sum_{i=1}^r \exp[\tau(x_{(i)} - \mu)] - \frac{k_3 \tau}{2} (\mu - m_3)^2 - \sum_{j=1}^{n-r} \exp(\tau[q_{(j)} - \mu]) \quad (1.34)$$

donde  $ce$  es una constante que no depende de  $\mu$  y  $\tau$ . Derivando esta expresión respecto a  $\mu$ , se obtiene:

$$\frac{dl_3(\mu, \tau)}{d\mu} = -\tau r + \tau \sum_{i=1}^r \exp(\tau(x_{(i)} - \mu)) - k_3 \tau (\mu - m_3) + \tau \sum_{j=1}^{n-r} \exp(\tau[q_{(j)} - \mu])$$

que al derivarla respecto a  $\mu$  se obtiene (4.36), y al derivarla respecto a  $\tau$  se obtiene (4.38). Si (1.34) se deriva respecto a  $\tau$  se obtiene

$$\frac{d}{d\tau} l_3(\mu, \tau) = \frac{(r + \alpha_3 - 0.5)}{\tau} - r\mu + \sum_{i=1}^r x_{(i)} - \beta_3 - \sum_{i=1}^r (x_{(i)} - \mu) e^{\tau(x_{(i)} - \mu)} - 0.5k_3 (\mu - m_3)^2 - \sum_{j=1}^{n-r} (q_{(j)} - \mu) e^{\tau[q_{(j)} - \mu]}$$

que al volverla a derivar respecto a  $\tau$  se obtiene (4.37). Finalizamos así este inciso de la demostración.

d) **Modelo Weibull.** En este caso  $l_4(\eta, \theta | M_4, \alpha_4, \beta_4, m_4, k_4)$  está dada por

$$\log(cw) + (\alpha_4 + r - 0.5) \log(\theta) - r\theta \log(\eta) + (\theta - 1) \log \left( \prod_{i=1}^r x_{(i)} \right) - \sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta - \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta - \beta_4 \theta - \frac{k_4 \theta}{2} (\eta - m_4)^2 \quad (1.35)$$

donde  $cw$  es una constante que no depende de  $\eta$  y  $\theta$ . Derivando esta expresión respecto a  $\eta$  obtenemos

$$\frac{\partial l_4}{\partial \eta} = -r \frac{\theta}{\eta} + \theta \frac{\sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta + (n-r) \left( \frac{q_{(j)}}{\eta} \right)^\theta}{\eta} - k_4 \theta (\eta - m_4)$$

Al derivar esta expresión respecto a  $\eta$  obtenemos (4.39), y si se deriva respecto a  $\theta$  se obtiene (4.41). Si (1.35) se deriva respecto a  $\theta$  se obtiene

$$\frac{\partial l_4}{\partial \theta} = \frac{\alpha_4 + r - 0.5}{\theta} - r \ln \eta + \log \left( \prod_{i=1}^r x_{(i)} \right) - \left( \sum_{i=1}^r \left( \frac{x_{(i)}}{\eta} \right)^\theta \ln \frac{x_{(i)}}{\eta} \right) - \beta_4 - \sum_{j=1}^{n-r} \left( \frac{q_{(j)}}{\eta} \right)^\theta \ln \frac{q_{(j)}}{\eta} - 0.5k_4 (\eta - m_4)^2$$

que al derivarla respecto a  $\theta$  se obtiene (4.40). De esta manera finalizamos la demostración del último inciso del teorema 4.5. ■

**DEMOSTRACION del teorema 4.6.** En donde se encuentra la VI para el modelo exponencial.

a) **A priori gama.** Dada la expresión (4.16) para la verosimilitud de un modelo exponencial y utilizando una distribución  $\text{gama}(\alpha_5, \beta_5)$  como distribución inicial para  $\theta$ , entonces la VI está dada por

$$\int_0^{\infty} C\theta^r \exp \left[ -\theta \left( \sum_{i=1}^r x_{(i)} + \sum_{j=1}^{n-r} q_{(j)} - n\gamma \right) \right] \frac{\beta_5^{\alpha_5}}{\Gamma(\alpha_5)} \theta^{\alpha_5-1} \exp(-\beta_5\theta) d\theta$$

Reordenado términos, completando una densidad gama e incluyendo sus términos constantes, obtenemos que la VI está dada por

$$\frac{C\beta_5^{\alpha_5}}{\Gamma(\alpha_5)} \frac{\Gamma(r + \alpha_5)}{h^{r+\alpha_5}} \int_0^{\infty} \frac{h^{r+\alpha_5}}{\Gamma(r + \alpha_5)} \theta^{r+\alpha_5-1} \exp[-h\theta] d\theta$$

con  $h = \sum_{i=1}^r x_{(i)} + \sum_{j=1}^{n-r} q_{(j)} + \beta_5 - n\gamma$ . El integrando de esta expresión corresponde a una densidad gama y la integración se está haciendo sobre todo el espacio paramétrico, entonces la integral es igual a uno. De aquí que al integrar se obtenga que la VI sea igual a

$$\frac{C\beta_5^{\alpha_5}}{\Gamma(\alpha_5)} \frac{\Gamma(r + \alpha_5)}{h^{r+\alpha_5}}$$

que es lo que queríamos demostrar.

b) **A priori uniforme.** Dada la expresión (4.16) para la verosimilitud de un modelo exponencial y utilizando una distribución uniforme  $(a_5, b_5)$  como distribución inicial para  $\theta$ , entonces la VI está dada por

$$\int_{a_5}^{b_5} u_5 C\theta^r \exp \left[ -\theta \left( \sum_{i=1}^r x_{(i)} + \sum_{j=1}^{n-r} q_{(j)} - n\gamma \right) \right] d\theta$$

Completando términos para obtener una densidad gama, obtenemos que VI es

$$\frac{u_5 C\Gamma(r + 1)}{(h)^{r+1}} \int_{a_5}^{b_5} \frac{(h)^{r+1}}{\Gamma(r + 1)} \theta^r \exp[-h\theta] d\theta$$

donde  $h = \sum_{i=1}^r x_{(i)} + \sum_{j=1}^{n-r} q_{(j)} - n\gamma$ . Como la integral es sobre una densidad gama, entonces integrando obtenemos que la VI está dada por

$$\frac{u_5 C\Gamma(r + 1)}{(h)^{r+1}} [G(b_5; r + 1, h) - G(a_5; r + 1, h)]$$

con  $G(\cdot; r + 1, h)$  es la distribución gama acumulada con parámetros  $(r + 1, h)$ . Que es lo que queríamos demostrar en este inciso. ■

**DEMOSTRACION del teorema 4.7.** Esta demostración es similar a la del teorema 4.4, sólo cambiando la verosimilitud bajo muestreo aleatorio (4.15), por la verosimilitud bajo censura (4.16). De esta forma se llega a que:

a) Utilizando una *a priori* Gama( $\alpha_5, \beta_5$ ) para  $\theta$  (4.17), la densidad posterior predictiva es:

$$f_5(y|\mathbf{X}) = \frac{(p_5)^{r+\alpha_5} \Gamma(r + \alpha_5 + 1)}{\Gamma(r + \alpha_5) (p_5 + y - \gamma)^{r+\alpha_5+1}}$$

donde  $p_5 = \sum_{i=1}^r x_{(i)} + (n-r)T_e + \beta_5 - n\gamma$ . Esta es una densidad gama-gama( $r + \alpha_5, p_5, 1$ ) (Ver Bernardo y Smith, 1994, pag. 120).

b) Utilizando una *a priori* uniforme( $a_5, b_5$ ), y siguiendo la demostración del inciso (b) del teorema 4.4, obtenemos que

$$f_5(y|\mathbf{X}) = u_5 \frac{\Gamma(r+2)}{(t_5+y)^{r+1+1}} \int_{a_5}^{b_5} \frac{(t_5+y)^{r+1+1}}{\Gamma(r+2)} \theta^{n+1} \exp[-(s_5+y)\theta] d\theta$$

con  $s_5 = \sum_{i=1}^r x_{(i)} + (n-r)T_e - (n+1)\gamma$ . De aquí que

$$f_5(y|\mathbf{X}) = \frac{u_5 \Gamma(r+2)}{(s_5+y)^{n+2}} [G(b_5; r+2, s_5+y) - G(a_5; r+2, s_5+y)]$$

donde  $G(\cdot; r+2, t_5+y)$  es la distribución gama acumulada con parámetros ( $r+2, s_5+y$ ). Esto es lo que queríamos demostrar. ■

**DEMOSTRACION del teorema 4.8.** En todos los casos obtendremos el hessiano de (3.6), es decir de  $l_i(\theta_i | a_i, b_i, d_i, e_i) = \log[L(\mathbf{X}|\theta_i, M_i)\pi(\theta_i | a_i, b_i, d_i, e_i)]$ . Utilizando como *a priori* una uniforme.

a) **Modelo normal.** En este caso  $l_1(\mu, \tau | a_1, b_1, d_1, e_1)$  está dado por

$$\begin{aligned} \log(Cu_1) + \frac{r}{2} [\log(\tau) - \log(2\pi)] - \frac{\tau}{2} [r(\mu - \bar{x}_r)^2 + s_r] + \\ \sum_{j=1}^{n-r} \log \left( \int_{\tau^{1/2}(q_{(j)}-\mu)}^{\infty} \left( \frac{1}{2\pi} \right)^{\frac{1}{2}} \exp \left[ -\frac{w^2}{2} \right] dw \right) \end{aligned} \quad (1.36)$$

derivando esta expresión respecto a  $\mu$ , se obtiene

$$\frac{dl_1}{d\mu} = -r\tau(\mu - \bar{x}_r) + \sum_{j=1}^{n-r} \exp \left( -\frac{1}{2}\tau(\mu - q_{(j)})^2 \right) \frac{\sqrt{2\tau}}{\sqrt{\pi} (1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}}(\mu - q_{(j)})))}$$

si esta expresión se vuelve a derivar respecto a  $\mu$ , entonces se obtiene (4.46), y si se deriva respecto a  $\tau$ , entonces se obtiene (4.48). Si (1.36) se deriva respecto a  $\tau$  se obtiene

$$\begin{aligned} \frac{d}{d\tau} l_1(\mu, \tau) = \frac{r}{2\tau} - \frac{\tau}{2} (\mu - \bar{x}_r)^2 - \frac{s_r}{2} + \\ \sum_{j=1}^{n-r} \frac{1}{\sqrt{2\pi\tau}} \exp \left( -\frac{1}{2}\tau(\mu - q_{(j)})^2 \right) \frac{(\mu - q_{(j)})}{(1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}}(\mu - q_{(j)})))} \end{aligned}$$

que al volverse a derivar respecto a  $\tau$ , se obtiene (4.47). De esta forma finalizamos la demostración de este inciso.

b) Modelo lognormal. En este caso  $l_2(\mu, \tau | a_2, b_2, d_2, e_2, M_2)$  está dado por

$$\log(cu_2) - \log\left(\prod_{i=1}^r x_{(i)}\right) + \log\left(\frac{\tau}{2\pi}\right) \frac{r}{2} - \frac{r\tau}{2}(\mu - \bar{w}_r)^2 - \frac{\tau u_r}{2} + \sum_{j=1}^{n-r} \log(1 - \Phi(\tau^{1/2}(\log(q_{(j)}) - \mu))) \quad (1.37)$$

Al derivar esta expresión respecto a  $\mu$  se obtiene

$$\frac{dl_2}{d\mu} = -r\tau(\mu - \bar{w}_r) + \sum_{j=1}^{n-r} \exp\left(-\frac{1}{2}\tau(\mu - \log(q_{(j)}))^2\right) \frac{\sqrt{2\tau}}{\sqrt{\pi}(1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}}(\mu - \log(q_{(j)}))))}$$

si esta expresión se vuelve a derivar respecto a  $\mu$  se obtiene (4.49), y si se deriva respecto a  $\tau$  se obtiene (4.51). Si (1.37) se deriva respecto a  $\tau$  se obtiene

$$\frac{d}{d\tau} l_2(\mu, \tau) = -\frac{r}{2}(\mu - \bar{w}_r)^2 - \frac{u_r}{2} + \frac{1}{2\tau} + \sum_{j=1}^{n-r} \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{1}{2}\tau(\mu - \log(q_{(j)}))^2\right) \frac{(\mu - \log(q_{(j)}))}{(1 + \operatorname{erf}(\sqrt{\frac{\tau}{2}}(\mu - \log(q_{(j)}))))}$$

que al derivarla respecto a  $\tau$  se obtiene (4.50).

c) Modelo valor extremo. En este caso  $l_3(\mu, \tau | a_3, b_3, d_3, e_3)$  está dada por

$$\log(cu_3) + r \log(\tau) + \tau \left(-r\mu + \sum_{i=1}^r x_{(i)}\right) - \sum_{i=1}^r \exp(\tau(x_{(i)} - \mu)) - \sum_{j=1}^{n-r} \exp(\tau[q_{(j)} - \mu]) \quad (1.38)$$

Derivando esta expresión respecto a  $\mu$ , se obtiene:

$$\frac{dl_3(\mu, \sigma)}{d\mu} = -\tau r + \tau \sum_{i=1}^r \exp(\tau(x_{(i)} - \mu)) + \tau \sum_{j=1}^{n-r} e^{\tau[q_{(j)} - \mu]}$$

al derivar esta expresión respecto a  $\mu$  se obtiene (4.52), y si se deriva respecto a  $\tau$  se obtiene (4.54). Si (1.38) se deriva respecto a  $\tau$  se obtiene

$$\frac{dl_3}{d\tau} = \frac{r}{\tau} - r\mu + \sum_{i=1}^r x_{(i)} - \sum_{i=1}^r (x_{(i)} - \mu) \exp[\tau(x_{(i)} - \mu)] - \sum_{j=1}^{n-r} [q_{(j)} - \mu] \exp(\tau[q_{(j)} - \mu])$$

que al volverla a derivar respecto a  $\tau$  se obtiene (4.53).

d) Modelo Weibull. En este caso  $l_4(\eta, \theta | a_4, b_4, d_4, e_4)$ , está dada por

$$\log(cu_4) + r \log(\theta) - r\theta \log(\eta) + (\theta - 1) \log\left(\prod_{i=1}^r x_{(i)}\right) - \sum_{i=1}^r \left(\frac{x_{(i)}}{\eta}\right)^\theta - \sum_{j=1}^{n-r} \left(\frac{q_{(j)}}{\eta}\right)^\theta \quad (1.39)$$

Si esta expresión se deriva respecto a  $\eta$  se obtiene

$$\frac{\partial l_4}{\partial \eta} = -r \frac{\theta}{\eta} + \theta \frac{\sum_{i=1}^r \left(\frac{x(i)}{\eta}\right)^\theta + \sum_{j=1}^{n-r} \left(\frac{q(j)}{\eta}\right)^\theta}{\eta}$$

que al derivar respecto a  $\eta$  esta expresión se obtiene (4.55), y si se deriva respecto a  $\theta$  se obtiene (4.57). Si (1.39) se deriva respecto a  $\theta$  se obtiene

$$\frac{\partial l_4}{\partial \theta} = \frac{r}{\theta} - r \ln \eta + (\ln p) - \left( \sum_{i=1}^r \left(\frac{x(i)}{\eta}\right)^\theta \ln \frac{x(i)}{\eta} \right) - (n-r) \left(\frac{q(j)}{\eta}\right)^\theta \ln \frac{q(j)}{\eta}$$

y al volverla a derivar respecto a  $\theta$  se obtiene (4.56). Con esto finalizamos la demostración del último inciso del teorema 4.7. ■

#### 1.4 Utilidad Esperada para el Modelo Exponencial

DEMOSTRACION del teorema 8.1. La expresión general para la utilidad esperada está dada por (8.10):

$$E(u(t, t_g)) = A_2(1 - e^{-A_1 t_g}) - \int_0^{t_g} \left[ A_4 \left(1 - \frac{A_3 t}{t_g}\right) + A_5 \left(1 - \frac{t}{t_g}\right) \right] f(t | \mathbf{X}) dt$$

por ello para la demostración del teorema, sustituiremos la expresión correspondiente a la densidad posterior predictiva  $f(t | \mathbf{X})$ , y haremos la integración.

a) **A priori gama y muestreo aleatorio.** Como en este caso  $f(t | \mathbf{X})$  está dada por (4.25), entonces sustituyendo, la utilidad esperada está dada por

$$E(u(t, t_g)) = A_2(1 - e^{-A_1 t_g}) - \int_0^{t_g} \left[ A_4 \left(1 - \frac{A_3 t}{t_g}\right) + A_5 \left(1 - \frac{t}{t_g}\right) \right] \frac{(v_5 - \gamma)^{n+\alpha_5} \Gamma(n + \alpha_5 + 1)}{\Gamma(n + \alpha_5) (v_5 + t - \gamma)^{n+\alpha_5+1}} dt$$

Simplificando,  $E(u(t, t_g))$  es igual a

$$A_2(1 - e^{-A_1 t_g}) - (A_4 + A_5) \int_0^{t_g} f_5(t | \mathbf{X}) dt + \frac{(A_4 A_3 + A_5)}{t_g} \int_\gamma^{t_g} t f_5(t | \mathbf{X}) dt$$

Como de acuerdo a (4.26)

$$\begin{aligned} & \int_0^{t_g} f_5(t | \mathbf{X}) dt = F_5(t_g) \\ & = \left( 1 - \frac{(v_5 - \gamma)^{n+\alpha_5}}{(v_5 + t_g - \gamma)^{n+\alpha_5}} \right) \end{aligned}$$

Además, integrando por partes,

$$\int_{\gamma}^{t_g} t f_5(t|\mathbf{X}) dt = \frac{\gamma v_{5g}^{-\alpha_5-n} - t_g (v_5 - \gamma + t_g)^{-\alpha_5-n}}{\alpha_5 + n} + \frac{v_5^{-n-\alpha_5+1} - (v_5 - \gamma + t_g)^{-n-\alpha_5+1}}{(-1 + n + \alpha_5)(n + \alpha_5)}$$

De esta manera,  $E(u(t, t_g))$  es igual a

$$A_2(1 - e^{-A_1 t_g}) - (A_4 + A_5) \left( 1 - \frac{(v_5 - \gamma)^{n+\alpha_5}}{(v_5 + t_g - \gamma)^{n+\alpha_5}} \right) + \frac{(A_4 A_3 + A_5)}{t_g} \left( \frac{\gamma v_{5g}^{-\alpha_5-n} - t_g (v_5 - \gamma + t_g)^{-\alpha_5-n}}{\alpha_5 + n} + \frac{v_5^{-n-\alpha_5+1} - (v_5 - \gamma + t_g)^{-n-\alpha_5+1}}{(-1 + n + \alpha_5)(n + \alpha_5)} \right)$$

que es lo que queríamos demostrar.

b) **Bajos muestreo censurado y a priori gama.** Dada la similitud entre  $f_5(t|\mathbf{X})$  bajo muestreo aleatorio (4.25) y bajo censura (4.44), la demostración es similar a la realizada para el inciso a). Por ello la utilidad esperada en este caso está dada por

$$E[u(t, t_g)] = A_2(1 - e^{-A_1 t_g}) - (A_4 + A_5) \left( 1 - \frac{(p_5 - \gamma)^{r+\alpha_5}}{(p_5 + t_g - \gamma)^{r+\alpha_5}} \right) + \frac{(A_4 A_3 + A_5)}{t_g} \left( \frac{\gamma p_{5g}^{-\alpha_5-n} - t_g (p_5 - \gamma + t_g)^{-\alpha_5-n}}{\alpha_5 + r} + \frac{p_5^{-r-\alpha_5+1} - (p_5 - \gamma + t_g)^{-r-\alpha_5+1}}{(-1 + r + \alpha_5)(r + \alpha_5)} \right)$$

con  $p_5 = \sum_{i=1}^r x_{(i)} + (n - r)T_e + \beta_5 - n\gamma$ . ■

## APENDICE

### ALGORÍTMOS COMPUTACIONALES

En este apéndice presentamos algunos aspectos de algoritmos computacionales y otros elementos relacionados.

#### 2.1 Generación de una distribución normal-gama en S-Plus

Utilizando las librerías del Sistema S-Plus versión 2000, se genera una muestra aleatoria  $(n, g)$  de una distribución normal-gama  $(m_i, k_i, \alpha_i, \beta_i)$ , expresión (4.4), aplicando los siguientes pasos:

1. Generar un muestra de una densidad gama  $(\alpha_i, \beta_i)$ :  $g <- rgamma(1, \alpha_i, \beta_i)$
2. Generar un número aleatorio de una normal con media  $m_i$  y desviación estándar  $(k_i * g)^{0.5}$ , esto se hace de forma directa con:  $n <- rnorm(1, m_i, (k_i * g)^{0.5})$

La pareja simulada  $(n, g)$  sigue la distribución normal-gama  $(m_i, k_i, \alpha_i, \beta_i)$  dada por (4.4).

#### 2.2 Generación de una distribución uniforme en S-Plus

Para generar un pareja  $(m, t)$  de la distribución uniforme  $(a_i, b_i, d_i, e_i)$ , expresión (4.5), se aplican los siguientes pasos:

1. Generar una muestra de una densidad uniforme  $(a_i, b_i)$ :  $m <- runif(1, a_i, b_i)$
2. Generar un número uniforme  $(d_i, e_i)$ :  $t <- runif(1, d_i, e_i)$

La pareja simulada  $(m, t)$  sigue la distribución uniforme  $(a_i, b_i, d_i, e_i)$  dada por (4.5).

#### 2.3 Generación de una muestra aleatoria con censura

Recordemos que si  $U_1, \dots, U_n$  es una muestra aleatoria de una distribución  $u(0, 1)$ , si además  $t_p = F_T^{-1}(p | \theta)$  es la función cuantil para la distribución de la variable aleatoria  $T$ . Entonces  $T_1 = F_T^{-1}(U_1 | \theta), \dots, T_n = F_T^{-1}(U_n | \theta)$  es una muestra aleatoria de  $F_T | \theta$ .

En el caso de datos censurados, sea  $U_{(i)}$  la  $i$ -ésima estadística de orden de una muestra aleatoria de tamaño  $n$  de una distribución  $Unif(0, 1)$ , aplicando las propiedades de las estadísticas de orden se obtiene que la distribución de  $U_{(i)}$ , dada  $U_{(i-1)}$  está dada por (Meeker y Escobar, 1998):

$$\Pr [U_{(i)} \leq u | U_{(i-1)} = u_{(i-1)}] = 1 - \left[ \frac{1 - u}{1 - u_{(i-1)}} \right]^{(n-i+1)}, \quad u \geq u_{(i-1)}$$

De esta forma obteniendo la función cuantil de esta distribución, se obtiene que una observación aleatoria de  $U_{(i)}$  está dado por

$$U_{(i)} = 1 - [1 - U_{(i-1)}] \times (1 - U_B)^{1/(n-i+1)}, \quad i = 1, \dots, n. \quad (2.31)$$

Muestras Censuradas (tipo II) Aplicando (2.31) el algoritmo para generar muestras aleatorias censuradas por número de fallas, con  $n$  unidades y  $r$  fallas, es el siguiente:

1. Generar  $U_1, \dots, U_r$  observaciones aleatorias de una distribución  $Unif(0, 1)$

2. Calcular la muestra aleatoria de estadísticas de orden aplicando (2.31):

$$\begin{aligned}
 U_{(1)} &= 1 - [1 - U_{(0)}] \times (1 - U_1)^{1/n} \text{ con } U_{(0)} = 0 \\
 U_{(2)} &= 1 - [1 - U_{(1)}] \times (1 - U_2)^{1/(n-1)} \\
 &\vdots \\
 U_{(r)} &= 1 - [1 - U_{(r-1)}] \times (1 - U_r)^{1/(n-r+1)}
 \end{aligned}$$

3. La muestra aleatoria de  $F(t | \theta)$  es

$$T_{(i)} = F^{-1}[U_{(i)} | \theta], \quad i = 1, \dots, r. \quad (2.32)$$

Muestras Censuradas (tipo I) De igual manera aplicando (2.31), el algoritmo para generar muestras aleatorias censuras por tiempo, con  $n$  unidades y tiempo de censura  $t_c$ , es el siguiente:

1. Generar una nueva observación  $U_i$  de una distribución  $Unif(0, 1)$ . Calcular  $U_{(i)} = 1 - [1 - U_{(i-1)}] \times (1 - U_i)^{1/(n-i+1)}$  y  $T_{(i)} = F^{-1}[U_{(i)} | \theta]$ .
2. Si  $T_{(i)} > t_c$ , parar. La muestra estará formada por los tiempos de falla  $T_{(1)}, \dots, T_{(i-1)}$  y  $(n - i + 1)$  observaciones censuradas.
3. Si  $T_{(i)} \leq t_c$ , incrementar  $i$  y regresar al primer paso.

#### 2.4 Descripción de los programas para calcular la probabilidad posterior de un modelo y algunas experiencias

Los programas para calcular la probabilidad posterior de un modelo fueron realizados en S-plus, y tiene la estructura general que se describe con los siguientes puntos.

1. Al inicio del programa se dan los intervalos donde se espera que ocurra el tiempo medio de vida y la desviación estándar. Con esta información y de acuerdo a las expresiones de las tablas 5.1 y 5.2, se calcula los hiperparámetros de los dos tipos de distribuciones *a priori* (normal-gama y uniforme) para los parámetros de los cinco modelos que hemos considerado en este trabajo.
2. Se introduce el vector que contiene los tiempos de vida, y cuando es el caso se introduce el vector que contiene los códigos de censura.
3. Después se procede a calcular la verosimilitud integrada: para ello se tienen programas (o rutinas) para todas las opciones que reflejan los teoremas del capítulo 4. Es decir, cálculo exacto, cuando esto es posible; aproximación de Laplace y simulación Monte Carlo. Por ejemplo en el caso de la aproximación de Laplace se programó el algoritmo correspondiente que se describe en el capítulo 3.

4. Con la VI se calcula la probabilidad posterior para cada uno de los modelos, al ser comparado con los restantes.

Algunas de las experiencias que se lograron en el proceso de calcular la probabilidad posterior, y que consideramos importante comentar, son las siguientes:

1. Al principio desarrollamos programas para calcular la verosimilitud integrada mediante curvatura gaussiana. Sin embargo, el algoritmo y por lo tanto los programas, son muy vulnerables a que los límites de integración sean los adecuados: que no sean muy amplios y que contengan la curva a ser integrada. Por ello, a pesar de varios intentos en esta dirección, los programas nunca lograron una estabilidad aceptable en cuanto a su funcionamiento. Por ello se decidió no incluirlo como una estrategia de cómputo de la VI en este trabajo. Ya en el apartado de investigaciones futuras del capítulo 9, se comentó que entre las opciones que habrá que explorar en el futuro es explorar formas alternativas de cómputo de la VI, en este sentido quizás la curvatura gaussiana pueda seguir siendo una alternativa.
2. Los valores de la VI son en general muy pequeños, y puede ir desde centésimas hasta valores del orden  $10^{-70}$ . A medida que los valores de los datos (variable) son mayores, la VI se vuelve más pequeña. Por ello, de ser posible, es mejor expresar los datos en escalas más generales; por ejemplo kilómetros en lugar de metros; días o años, en lugar horas; etcétera.
3. Por algunos estudios exploratorios, que no hemos reportado formalmente, la calidad de la aproximación de Laplace cuando se utiliza *a priori* uniforme es mala, comparada con la que se logra cuando se utiliza *a priori* normal-gama.

## 2.5 Descripción del programa para decidir el tiempo de garantía

El programa computacional se desarrolló en S-plus 2000, y algunas de sus características son:

1. Primero se especifica los parámetros de la función de utilidad y los intervalos para la media y la desviación estándar.
2. Se introducen los datos del tiempo de vida.
3. Enseguida se calculan los parámetros de dos tipos de distribuciones *a priori* (ver tabla 5.1 y 5.2).
4. Se genera una muestra aleatoria de la distribución *a priori*.
5. Luego se definen las funciones de verosimilitud y verosimilitud relativa.
6. Se genera una muestra aleatoria de la distribución posterior predictiva mediante el algoritmo Sampling-importance-resampling.
7. Para un número amplio de posibles tiempos de garantía, se evalúa la función de utilidad considerando la muestra generada previamente.

8. Se obtiene el tiempo de garantía óptimo, la utilidad esperada en ese punto, una gráfica de la función de utilidad esperada.
9. A la muestra aleatoria de la distribución posterior predictiva, se le obtienen algunos percentiles claves y se hace un histograma.

## REFERENCIAS

1. Abramowitz, M. y Stegun, I.A. Eds. (1965). *Handbook of Mathematical Functions*. Dover: New York.
2. Al-Awadhi, S.A. y Garthwaite, P.H. y (1998). "An elicitation method for multivariate normal distributions". *Commun. Statist.-Theory Meth.*, 27(5), 1123-1142.
3. Allen, R.E. (2001). "Lean: a supplier weapon against warranty cost." *Manufacturing Engineering*, 126, 18-19.
4. Anaya, K. y O'Reilly, F. (2001). "Test of fit for the inverse gaussian and gamma distributions under censoring". *Commun. Statist.-Theory Meth*, 30, 4, 735-773.
5. Ansell, J.I. y M.J. Phillips (1994). *Practical Methods for Reliability Data Analysis*. Clarendon Press: Oxford.
6. Barrio, E., Cuesta-Albertos, J.A. y Matrán, C. (2000). "Contributions of empirical and quantile processes to the asymptotic Theory of Goodness-of-fit Tests". *Test*, 9, 1-96.
7. Berger, J.O. (1993). "An overview of robust bayesian analysis". Technical Report 93-53C, Department of Statistics, Purdue University.
8. Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlang: New York.
9. Berger, J.O. y Pericchi L.R. (1996). "The intrinsic Bayes factor for model selection and prediction". *J. Amer. Statist. Assoc.* 91, 109-121.
10. Bernardo, J.M. y Smith, A.F.M. (1994). *Bayesian Theory*. Wiley: Chichester.
11. Blischke, W.R. y Murthy, D.N.P., Editores (1996). *Product Warranty Handbook*. Marcel.Dekker: New York
12. Blischke (1996). "The basic free\_replacement warranty and related rebate warranties". In *Product Warranty Handbook*, ed. por. Blischke, W.R. y Murthy, D.N.P. Marcel.Dekker: New York, pp. 265-291.
13. Brent, R. (1973). *Algorithms for minimization without derivatives*. Prentice-Hall: Englewood Cliffs, NJ, USA.

14. Chen, M-H. y Shao, Q-M. (1997). "Estimating ratios of normalizing constantes for densities with different dimensions". *Statistical Sinica*, 7, 607-630.
15. Crane, M. (2000). "Supreme court: your office can be searched without a warrant" *Medical Economics*, 77 (11); 43-44.
16. DeGroot, M.H. (1970). *Optimal Statistical Decisions*. McGraw-Hill: New York.
17. Deming., E.W. (1989). *Calidad, Productividad y Competitividad*. Díaz de Santos; Madrid.
18. Doganaksoy, N., Hahn, G.J., y Meeker, W.Q. (2002). "Reliability analysis by failure mode". *Quality Progress*, 35, 6, 47-52.
19. Duncan, A. (1989). *Estadística Industrial*. AlfaOmega: Madrid.
20. Emons, W. (1989). "On the limitation of warranty duration". *Journal of Industrial Economics*, 37, 3, 287-301.
21. Federal Trade Commission (1985). *An Evaluation of the Warranty Rules of the Magnuson-Moss Act*. U.S. Government Printing Office, Washington.
22. Garthwaite, P.H.(1992). "Preposterior expected loss as a scoring rule for prior distributions". *Commun. Statist.-Theory Meth.*, 21(12), 3601-3619.
23. Garthwaite, P.H. y O'Hagan, A.(2000). "Quantifying expert opinion: an experimental study". *The Statistician*, 49, part 4, 455-477.
24. Garthwaite, P.H. y Al-Awadhi, S.A. (2001). "Non-conjugate prior distribution assessment for multivariate normal sampling". *J. Roy. Statist. Soc. Ser. B*, 63,1, 95-110.
25. Geisser, S. y Eddy, W.F. (1979). "A Predictive approach to model selection". *J. Amer. Statist. Assoc.* 74, 153-160.
26. Gelfand, A.E. y Dey, D.K. (1994). "Bayesian model choice: asymptotics and exact calculations". *J. Roy. Statist. Soc. Ser. B* 56, 501-514.
27. Genz, A. y Kass, R.E. (1997). "Subregion-adaptive integration of functions having a dominant peak". *J. of Comput. and Graphical Statistics*, 6, 92-111.
28. George, G.I. (1999). Discusión de "Bayesian model averaging: a tutorial" por Hoeting, et. al. *Statistical Science*, 14, 382-417.
29. Gutiérrez-Peña, E. y Walker, S.G. (2001). "A Bayesian predictive approach to model selection". *J. of Statistical Planning and Inference*, 93, 259-276.
30. Gutiérrez-Pulido, H., Aguirre-Torres, V. y Christen, A. (2002). "Model selection in reliability: a Bayesian approach". *Communications of third Internatioanal Conference of Mathmatical Methods in Reliability*.Throngeim, Noruega.

31. Hoeting, J.A., Madigan, D., Raftery, A.E. y Volinsky, C.T. (1999). "Bayesian model averaging: a tutorial". *Statistical Science*, 14, 382-417.
32. Ibrahim, J.G., Chen, M-H., y Sinha, D. (2001). *Bayesian Survival Analysis*. Springer-Verlag: New York.
33. Inman, R.R. and Gonsalvez, (1998). "A cost-benefit model for production vehicle testing" *IIE Transactions*, 30, 1153-1160.
34. Kadane, J.B. y Wolfson, L.J.(1998). "Experiencies in elicitation". *The Statistician* 47, 3-19.
35. Kalbfleisch, J.D., J.F. Lawless and J.A. Robinson (1991). "Methods for the Analysis and prediction of warranty claims." *Technometrics*, V ol. 33, No. 3, 273-285.
36. Kass, R.E. y Raftery, A. E. (1995). "Bayes factor". *J. Amer. Statist. Assoc.* 90, 773-795.
37. Kass, R.E., Tierney, L. y Kadane, J.B. (1990). "The validity of posterior asymptotic expansions based on Laplace's method" In *Bayesian and Likelihood Methods in Statistics and Econometrics*, eds. S. Geisser, J.S: Hodges, S.J. Press, y A. Zellner. North-Holland: New-York.
38. Kass, R.E. y Vaidyanathan, S. (1992). "Approximate Bayes factor and orthogonal parameters, with applications to testing equality of two binomial proportions". *Journal of the Royal Statistical Society, Ser. B* 54, 129-144.
39. Kelley, C.A. (1996). "Warranty and consumer behavior: product choice" In *Product Warranty Handbook*, ed. por. Blischke, W.R. y Murthy, D.N.P. Marcel.Dekker: New York, pp. 409-419.
40. Kim, H-G y B.M. Rao (2000). "Expected warranty cost of two-attribute free-replacement warranties based on a bivariate exponential distribution." *Computers & Industrial Engineering*, 38, 425-434.
41. Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: Wiley.
42. Lawless, J.F. (1998). "Statistical analysis of product warranty data". *International Statistical Review*, 66, 41-60.
43. Mann, N.R., Schafer, R.E. y Ferting, K.W. (1973). "A new goodness of fit Test for the two parameter Weibull or extreme value distribution". *Commun. Stat*, 2, 383-400.
44. Martz, H.F. (1982). *Bayesian Reliability Analysis*. New York: Wiley.
45. Marcellus, R. y Pirojboot, B.(1996). "Design of warranty policies to balance consumer and producer risks and benefits". In *Product Warranty Handbook*, ed. por. Blischke, W.R. y Murthy, D.N.P. Marcel.Dekker: New York, pp. 483-510.

46. Medigan, D y Raftery, A.E. (1994). "Model selection and accounting for model uncertainty in graphical models using Occam's window". *J. Amer. Statist. Assoc.* 89, 1535-1546.
47. Meeker, W.Q. and Escobar, E. (1998). *Statistical Methods for Reliability Data*. New York: Wiley.
48. Menezes, M.A.J. y Currim, I.S. (1992). "An approach for determination of warranty length". *Intern. J. of Research in Marketing*, 9, 177-195.
49. Menzefricke, U. (1992). "On the variance of total warranty claims". *Commun. Statist.-Theory Meth.* 21 (3), 779-790.
50. Mihalko, D.P. y Moore, D.S. (1980). "Chi-square tests of fit for type II censored data". *The Annals of Statistics* 8,3, 625-644.
51. Mira, A. y Nicholls, G. (2001). "Bridge estimation of the probability density at a point". Preprint.
52. Nelson, W. (1982). *Applied Life Data Analysis*. John Wiley: New York.
53. O'Connor, P.D.T. (1985). *Practical Reliability Engineering* (Second Edition). Wiley: New York.
54. O'Hagan, A. (1998). "Eliciting expert beliefs in substantial practical applications". *The Statistician* 47, 21-35.
55. Padmanabhan, V. (1996). "Marketing and warranty". In *Product Warranty Handbook*, ed. por. Blischke, W.R. y Murthy, D.N.P. Marcel.Dekker: New York, pp. 393-407.
56. Patankar, J.G. y Mitra, A. (1996). "Warranty and consumer behavior: warranty execution" In *Product Warranty Handbook*, ed. por. Blischke, W.R. y Murthy, D.N.P. Marcel.Dekker: New York, pp. 421-438.
57. Percy, D.F. (2002). "Bayesian enhanced strategic decision making for reliability". *European Journal of Operational Research* 139, 133-145.
58. Perez, J.M. y Berger, J.O. (2002). "Expected-posterior prior distributions for model selection". *Biometrika*, 89, 491-511.
59. Pettit, A.N. (1976). "Cramér-von Mises statistics for testing normality with censored samples". *Biometrika*, 63, 475-481.
60. Proschan, F. (1963). "Theoretical explanation of observed failure rate". *Technometrics*, 5, 375-383.
61. Raftery, A.E. (1994). "Approximate Bayes factors and accounting for model uncertainty in generalized linear models". Technical Report No. 255, Department of Statistics, University of Washington.

62. Rao, B.M. (1995). "Algorithms for the free replacement warranty with phase-type lifetime distributions." *IIE Transactions*, 7, 348-358.
63. Ramirez-Beltran, N. (1999). "Statistical analysis of drug stability". *Computers & Industrial Engineering*, 37, 351-355.
64. Robert, C.P. y Casella, G. (1999). *Monte Carlo Statistical Methods*. Springer: New York.
65. Rossi, P.E. (1989). "Comparison of alternative functional forms in production" In *New Approaches to Modeling, Specification Selection and Econometric Inference*, Barnett, W.A. y Gallant, A.R. (Eds). Cambridge Univ. Press: New York.
66. Singpurwalla, N.D. (1988). "An interactive pc-based procedure for reliability assessment incorporating expert opinion and survival data". *Journal of the American Statistical Association*, 83, 43-51.
67. Singpurwalla, N.D. y Wilson, S.P. (1993). "The warranty problem: Its statistical and game theoretic aspects." *SIAM Review*, 35, 17-42.
68. Singpurwalla; N.D. y Wilson, S.P. (1998). "Failure models indexed by two scales." *Adv. Appl. Prob.*, 30, 1058-1072
69. Sinharay, S. y Stern, H.S. (2002). "On the sensitivity of Bayes factors to the prior distributions". *The American Statistician*, 56, 3, 196-201.
70. Sinha, D., Chen, M-H. y Ghosh, S.K. (1999). "Bayesian analysis and model selection for interval-censored survival data". *Biometrics*, 55, 585-590.
71. Smith, A.F.M. y Roberts, G.O. (1993). "Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods". *J. Roy. Statist. Soc. Ser. B*, 55,1, 3-23.
72. Spiegelhalter, D.J y Smith, A.F.M. (1982). "Bayes factor for linear and log-linear models with vague priori information". *Journal of the Royal Statistical Society, Ser. B*, 44, 377-387.
73. Spizzichino, F. (2001). "Subjective probability models for lifetimes". Boca Raton, FL:Chapman & Hall/CRC.
74. Sprott, D.A. (2000). *Statistical Inference in Science*. Springer.
75. Suzuki, K. (1985a). "Estimation method of lifetime based on the record of failure during the warranty period". *Journal of the American Statistical Association*, 80, 66-72.
76. Suzuki, K. (1985b). "Nonparametric estimation of lifetime distribution from incomplete field data". *Technometrics*, 27, 263-271.

77. Weir, T. (2000) "Swell allowances return, drawing flak and favor". *Supermarket Business*, Vol. 55, Iss. 5, 19-21.
78. Wilson, A.G. (1994). "Cognitive factors affecting subjective probability assessment". Discussion Paper 94-02, Institute of Statistics and Decisions Sciences, Duke University.
79. Winkler, R.L. (1981). "Combining probability distributions from dependent information sources" *Management Science*, 27 (5), 479-488.
80. Wolfson, L.J. (1995). "Elicitation of priors and utilities for Bayesian analysis". Tesis Doctoral, Departamento de Estadística, Carnegie Mellon University: Pittsburgh.