

**TEORÍA DE LA CREDIBILIDAD:  
DESARROLLO Y APLICACIONES EN PRIMAS DE SEGUROS  
Y RIESGOS OPERACIONALES**

FUNDACIÓN MAPFRE no se hace responsable del contenido de esta obra, ni el hecho de publicarla implica conformidad o identificación con la opinión del autor o autores.

Prohibida la reproducción total o parcial de esta obra sin el permiso escrito del autor o del editor.

© 2008, FUNDACIÓN MAPFRE  
Carretera de Pozuelo 52  
28220 Majadahonda. Madrid

[www.fundacionmapfre.com/cienciasdelseguro](http://www.fundacionmapfre.com/cienciasdelseguro)  
[publicaciones.ics@mapfre.com](mailto:publicaciones.ics@mapfre.com)

ISBN: 978-84-9844-105-5  
Depósito Legal: SE-5537-2008

TEORÍA DE LA CREDIBILIDAD:  
DESARROLLO Y APLICACIONES EN  
PRIMAS DE SEGUROS  
Y RIESGOS OPERACIONALES

Emilio Gómez Déniz  
José María Sarabia Alegría

IV PREMIO INTERNACIONAL DE SEGUROS  
JULIO CASTELO MATRÁN



# Presentación

En julio de 2001 el Consejo de Administración de MAPFRE MUTUALIDAD, como homenaje a D. Julio Castelo Matrán, anterior presidente del SISTEMA MAPFRE y de MAPFRE MUTUALIDAD y gran impulsor de su desarrollo -actualmente es presidente de honor de FUNDACIÓN MAPFRE- acordó la creación, con carácter bienal, del Premio Internacional de Seguros Julio Castelo Matrán, destinado a premiar trabajos científicos sobre materias relacionadas con el seguro en España, Portugal y los países de Iberoamérica.

Con este motivo, FUNDACIÓN MAPFRE, a través del Instituto de Ciencias del Seguro, convocó en junio de 2007 la cuarta convocatoria del Premio y designó a un jurado compuesto por relevantes personalidades en los ámbitos profesional, científico y empresarial del mundo del seguro.

En la reunión celebrada el día 25 septiembre de 2008, tras valorar los diversos trabajos que con una alta calidad se habían presentado desde Argentina, Brasil, Cuba, España y Perú, el jurado acordó conceder el Premio, dotado con 35.000 euros y un diploma acreditativo de la obtención del mismo, a D. Emilio Gómez Déniz, Profesor Titular de Métodos Cuantitativos para la Economía y la Empresa de la Universidad de Las Palmas de Gran Canaria, y a D. José María Sarabia, Catedrático de Métodos Cuantitativos para la Economía y la Empresa de la Universidad de Cantabria, por su obra *Teoría de la Credibilidad: Desarrollo y Aplicaciones en Primas de Seguros y Riesgos Operacionales*, cuya publicación promovida por FUNDACIÓN MAPFRE me complace especialmente presentar.

El premio se entrega a los autores en un acto público de acuerdo con lo establecido en las bases de la convocatoria.

FUNDACIÓN MAPFRE, entre otras actividades, fomenta la investigación y la divulgación de conocimientos en relación con el Seguro y la Gerencia de Riesgos y para ello promueve la concesión de ayudas y premios a la investigación en estas áreas potenciando la realización de trabajos científicos como el que se recoge en esta publicación que esperamos sea de interés y utilidad tanto al mundo académico como al profesional relacionado con la actividad aseguradora.

José Manuel Martínez Martínez  
Presidente de FUNDACIÓN MAPFRE

JURADO DEL  
IV PREMIO INTERNACIONAL DE SEGUROS  
JULIO CASTELO MATRÁN

Dña. María José Albert Pérez  
Decana de la Facultad de Ciencias del Seguro, Jurídicas y de la Empresa  
Universidad Pontificia de Salamanca

D. Álvaro Cuervo García  
Catedrático de Economía de la Empresa de la Universidad Complutense de Madrid

D. Antonio Heras Martínez  
Catedrático de Economía Financiera y Actuarial de la Universidad Complutense de Madrid

D. Rafael Illescas Ortiz  
Catedrático de Derecho Mercantil de la Universidad Carlos III de Madrid

D. Filomeno Mira Candel  
Presidente del Instituto de Ciencias del Seguro de FUNDACIÓN MAPFRE

D. Luís Fernando Reglero Campos (†)  
Catedrático de Derecho Civil de la Universidad de Castilla-La Mancha

Dña. Alicia Sanmartín Ruiz  
Presidenta del Instituto de Actuarios Españoles

D. Fernando Suárez González  
Catedrático emérito de Derecho del Trabajo de la Universidad Nacional de Educación a Distancia y Miembro de la Real Academia de Ciencias Morales y Políticas

D. Vicente Tardío Barutel  
Presidente Consejero Delegado de Allianz, Compañía de Seguros y Reaseguros, S.A.

Dña. Mercedes Sanz Septién  
Directora General del Instituto de Ciencias del Seguro de FUNDACIÓN MAPFRE  
Secretaria del Jurado

PREMIO INTERNACIONAL DE SEGUROS  
JULIO CASTELO MATRÁN

Obras galardonados en ediciones anteriores:

2002

*Las Peripicias del Asegurador de Automóviles en el Proceso Penal*

D. Mariano Yzquierdo Tolsada

Catedrático de Derecho Civil de la Universidad de Comillas

2004

*El Seguro de Automóviles: Estado Actual y Perspectiva de la Técnica Actuarial*

Dña. Montserrat Guillén Estany

Catedrática de Economía Aplicada de la Universidad de Barcelona

Dña. Mercedes Ayuso Gutiérrez

Profesora Titular de Economía Aplicada de la Universidad de Barcelona

Dña. Catalina Bolancé Losilla

Profesora Titular de Economía Aplicada de la Universidad de Barcelona

D. Lluís Bermúdez Morata

Profesor Titular de Economía Financiera de la Universidad de Barcelona

Dña. Isabel Morillo López

Profesora Titular de Economía Financiera de la Universidad de Barcelona

Dña. Irene Albarrán Lozano

Profesora Titular de Economía Aplicada de la Universidad de Extremadura

2006

*Las Cargas del Acreedor en el Seguro de Responsabilidad Civil*

D. Osvaldo Lagos Villarreal

Profesor de la Universidad de Los Andes de Santiago de Chile

# Índice general

<b>Introducción</b>	<b>1</b>
<b>1. Teoría de la Credibilidad: Concepto y Perspectiva Histórica</b>	<b>5</b>
<b>2. Principios de Primas y Medidas de Riesgo</b>	<b>15</b>
2.1. Introducción . . . . .	15
2.2. Primas . . . . .	16
2.3. Métodos para la obtención de primas . . . . .	17
2.3.1. Propiedades de los principios de cálculo de primas . . . . .	18
2.3.2. Método ad-hoc: catálogo de principios de prima . . . . .	20
2.3.3. Método basado en caracterizaciones . . . . .	22
2.3.4. Método económico . . . . .	23
2.3.5. Funciones de pérdida y prima de riesgo . . . . .	23
2.4. Propiedades deseables de una medida de riesgo . . . . .	28
2.5. Medidas de riesgo y primas basadas en la cola de la distribución . . . . .	29
2.6. Una clase de medidas de riesgos ponderadas . . . . .	32
2.7. Primas basadas en estadísticos de orden . . . . .	33
<b>3. Modelización del Riesgo en Familias Paramétricas</b>	<b>35</b>
3.1. Introducción . . . . .	35
3.2. Modelos paramétricos . . . . .	36
3.2.1. Distribuciones de Pareto . . . . .	36
3.2.2. Distribuciones exponencial y gamma . . . . .	38
3.2.3. Distribución gamma invertida . . . . .	40
3.2.4. Distribuciones normal y lognormal . . . . .	43
3.2.5. Distribución de Weibull . . . . .	44
3.2.6. Distribución inversa Gaussiana . . . . .	47
3.2.7. Comparación de las colas . . . . .	47
3.3. Caso de estudio: datos normalizados sobre daños de huracanes . . . . .	50
3.3.1. Modelos Ajustados . . . . .	51
3.3.2. Regresión inversa-Gaussiana . . . . .	52

3.3.3. Cálculo de primas VaR y TVaR . . . . .	53
<b>4. Instrumentos de Estadística Bayesiana y Primas Bayes</b>	<b>61</b>
4.1. Introducción . . . . .	61
4.2. Definiciones básicas y teorema de Bayes . . . . .	62
4.3. Inferencia y predicción Bayesianas . . . . .	64
4.4. Aplicaciones Bayesianas en seguros . . . . .	68
4.4.1. Estimación de una cuota de mercado . . . . .	68
4.4.2. Estimación del fraude . . . . .	69
4.5. Análisis Bayesiano de la distribución de pérdidas Pareto . . . . .	72
4.5.1. Intervalos de credibilidad y distribución predictiva . . . . .	74
4.5.2. Caso de estudio: datos de pérdidas por vientos . . . . .	75
4.5.3. Otras distribuciones a priori . . . . .	76
4.6. Estimación Bayesiana de una clase de distribuciones de pérdida . . . . .	76
4.7. Estimación de la prima de riesgo . . . . .	81
4.7.1. Prima colectiva o a priori . . . . .	83
4.7.2. Prima Bayes o a posteriori . . . . .	85
4.8. Sistemas bonus-malus . . . . .	88
4.8.1. Cálculo de primas bonus-malus. Método Bayesiano . . . . .	89
4.8.2. Penalización de las sobrecargas . . . . .	94
4.9. Análisis Bayesiano de la familia exponencial natural . . . . .	99
4.10. Análisis Bayesiano de la familia exponencial de dispersión . . . . .	102
<b>5. Modelos de Credibilidad</b>	<b>105</b>
5.1. Introducción . . . . .	105
5.2. Credibilidad total y parcial . . . . .	106
5.2.1. Credibilidad total . . . . .	106
5.2.2. Credibilidad parcial . . . . .	109
5.3. Modelos de distribución libre . . . . .	110
5.3.1. Modelo de Bühlmann . . . . .	111
5.3.2. Modelo de Bühlmann-Straub . . . . .	117
5.4. Modelos Bayesianos . . . . .	121
5.4.1. Modelo de Jewell . . . . .	123
5.4.2. Modelo de Landsman y Makov . . . . .	128
5.4.3. Modelos Bayesianos jerárquicos . . . . .	130
5.4.4. Un modelo basado en una función general de pérdida . . . . .	151
5.4.5. Modelos de credibilidad basados en robustez . . . . .	158
<b>6. Aplicaciones en Riesgos Operacionales</b>	<b>171</b>
6.1. Introducción . . . . .	171
6.2. Metodología Bayesiana en riesgos operacionales . . . . .	172
6.3. Riesgos operacionales: modelización de la frecuencia . . . . .	172
6.3.1. Modelo Poisson-gamma . . . . .	173

6.3.2. Modelo binomial-beta . . . . .	175
6.4. Riesgos Operacionales: modelización de la severidad . . . . .	180
6.4.1. Modelo lognormal-normal . . . . .	181
6.4.2. Modelo Pareto generalizada-gamma . . . . .	182
6.4.3. Modelo exponencial-gamma invertida . . . . .	184
6.4.4. Modelo Pareto clásica-gamma . . . . .	186
6.5. Modelo de distribución de pérdidas agregadas LDA . . . . .	189
6.5.1. Propiedades generales del proceso LDA . . . . .	194
6.5.2. Procesos LDA con distribución explícita . . . . .	196
6.5.3. Otros Procesos LDA . . . . .	197
6.6. Aplicaciones . . . . .	198
6.6.1. Tipología en procesos LDA . . . . .	199
6.6.2. Aplicación del modelo exponencial-gamma invertida . . . . .	199
<b>7. Conclusiones</b>	<b>207</b>
<b>A. Apéndice: Datos de Klugman (1991)</b>	<b>213</b>
<b>Bibliografía</b>	<b>225</b>
<b>Índice</b>	<b>236</b>

# Introducción

La *teoría de la credibilidad* es un conjunto de técnicas que permiten al asegurador ajustar de modo sistemático las primas de los seguros en función de la experiencia de siniestralidad.

En teoría de la credibilidad entran en juego los dos conceptos clásicos de *riesgo individual* y *riesgo colectivo*, y se resuelve de modo riguroso el problema de cómo analizar la información obtenida de estas dos fuentes para llegar a la prima de seguros y obtener una tarifa justa.

La teoría de la credibilidad como disciplina matemática, toma sus métodos de diversos campos de las matemáticas: la estadística Bayesiana, el análisis funcional, las técnicas de mínimos cuadrados, la modelización sobre el espacio de estados etc. Partiendo de las ideas originales de Whitney y Mowbray autores contemporáneos como Bailey, Longley-Cook, Mayerson, Bühlmann, Straub y Jewell se han encargado de proporcionar a esta teoría una fundamentación matemática rigurosa, convirtiéndola en una de las ramas más atractivas y estudiadas de la ciencia actuarial. Uno de sus principales usos se presenta en el seguro del automóvil, en el que la prima inicial se va transformando sucesivamente a medida que se incorpora la información de la siniestralidad. Son los denominados sistemas de tarificación *bonus-malus*.

Existe en castellano un vacío importante de información en esta área de conocimiento, además de no existir ninguna obra suficientemente actualizada que incluya los avances recientes y sus aplicaciones. La desconexión entre los avances científicos más recientes y la práctica aseguradora, justifica una investigación como la presente, que engloba una gran parte del conocimiento existente sobre credibilidad.

Los objetivos de la investigación consisten en el desarrollo, análisis y aplicación de modelos de credibilidad, tanto en el cálculo de primas como en riesgos operacionales. Se presenta de una manera rigurosa y al mismo tiempo aplicada, tanto los modelos clásicos como los de desarrollo más reciente, además de diversas propuestas novedosas. La exposición y desarrollo de los diferentes modelos viene acompañada con su aplicación en problemas reales que aparecen en la literatura actuarial y de riesgos operacionales.

En las últimas décadas el sector financiero ha experimentado una evolución

sin precedentes caracterizada principalmente por la globalización de los mercados, la incorporación de las nuevas tecnologías al sector, el desarrollo de complejos productos derivados y la incorporación a los mercados de países del tercer mundo como China, India, etc.

En este nuevo entorno se consolida como pieza fundamental del éxito en la gestión financiera el control y la medición de riesgos. De ahí que en la última década haya nacido un término que se está haciendo frecuente en los sistemas financieros. Este término es el de *riesgo operacional*. Un riesgo adscrito a sucesos que no incluyen los de mercado ni los de créditos, y que pueden tener una procedencia interna y/o externa. Debido a las nuevas indicaciones regulatorias conocidas como Basilea II (BIS, 2005) en el sector bancario, se está realizando un considerable esfuerzo en la implementación de procedimientos tanto cualitativos como cuantitativos para modelizar el riesgo operacional.

La investigación se ha estructurado en seis capítulos más uno de conclusiones. Se incluye además un apéndice con datos, las referencias bibliográficas y un índice analítico.

Los contenidos de cada capítulo se exponen a continuación de manera breve.

- El capítulo 1 presenta una revisión de la literatura sobre los modelos existentes de credibilidad. Se comienza con una introducción al tema, para continuar con una descripción de los modelos de credibilidad más recientes.
- Los principios de cálculo de primas, así como la metodología asociada a su obtención, constituye un elemento imprescindible en la tarificación mediante credibilidad. En el capítulo 2 se presentan las propiedades deseables que debe satisfacer un principio de cálculo de prima, así como los diferentes métodos para su obtención. Se estudia el concepto de medida de riesgo, junto con las propiedades deseables que debe satisfacer. Las medidas de riesgo y primas basadas en la cola de la distribución (es decir valor en riesgo VaR y valor en riesgo en la cola TVaR) ocupan un lugar importante en la investigación actual en seguros y serán igualmente estudiadas en este capítulo.
- La modelización del riesgo por medio de familias paramétricas de distribuciones constituye la metodología clásica de establecer la población de referencia tanto en seguros como en teoría de la credibilidad. En el capítulo 3 estudiaremos modelos de distribuciones paramétricas que pueden ser utilizadas para describir el riesgo de un colectivo. Los modelos aquí introducidos serán posteriormente utilizados en los modelos de credibilidad, tanto para modelizar el colectivo de una cartera, como la distribución a priori de algún parámetro de interés. Como parte práctica, presentamos un caso de estudio aplicado a datos normalizados sobre daños de huracanes, donde además de trabajar con diferentes familias paramétricas, se introduce un modelo de regresión basado en la distribución inversa Gaussiana, para finalmente obtener diversas primas basadas en los modelos seleccionados.

- En el capítulo 4 se presentan los instrumentos de estadística Bayesiana y las primas Bayes. La estadística Bayesiana constituye una herramienta básica en credibilidad cuya metodología se ajusta perfectamente a los objetivos de la credibilidad. Por medio de la inferencia Bayesiana, se estimará la prima de riesgo, primero a través de la prima colectiva y posteriormente mediante la prima Bayes.
- El capítulo 5 está dedicado a los diferentes modelos de credibilidad. Se presentarán los modelos clásicos y Bayesianos de credibilidad, así como algunos modelos más recientes desarrollados en los últimos años. Entre estos modelos más recientes incluimos los modelos bayesianos jerárquicos y los modelos basados en robustez bayesiana, la aproximación *gamma-minimax* y la aproximación *posterior regret-gamma minimax*. Se incluyen nuevas aportaciones y se presentan diversas aplicaciones numéricas con datos reales.
- El capítulo 6 está dedicado a la modelización de riesgos operacionales por medio de la teoría de la credibilidad. Se plantean modelos para la cuantificación de la frecuencia y la severidad de pérdidas debidas a riesgos operacionales, además de la modelización de la distribución de pérdidas agregadas (LDA). Se presentan diversas aplicaciones con datos reales y simulados.
- Finalmente, el capítulo 7 incluye las conclusiones más importantes de la investigación.

## Capítulo 1

# Teoría de la Credibilidad: Concepto y Perspectiva Histórica

La teoría de la credibilidad constituye uno de los temas centrales de investigación y práctica en seguros. Dicha teoría incluye numerosos modelos que han aparecido desde comienzos del siglo XX hasta la actualidad, elaborándose en buena medida gracias a los avances que se iban produciendo en las ciencias matemáticas, en especial en su apartado de estadística, y más concretamente en estadística Bayesiana. De este modo, diferentes modelos que se desarrollan desde puntos de partida completamente diferentes convergen finalmente a un mismo resultado. Esta particularidad es uno de los atractivos de esta teoría. Su utilidad en las ciencias actuariales ha trascendido a otros campos científicos, algunos no relacionados con la disciplina actuarial. Es posible encontrar modelos de credibilidad en biología (Linacre et al., 2004), economía financiera (Bühlmann et al., 2007), estadística en deportes (Mahler, 1991), etc.

El término *credibilidad* se introdujo por vez primera en seguros en Estados Unidos antes de la primera guerra mundial, en relación con los sistemas de ajuste de primas en seguros de compensación obrera o seguros de accidentes. Por aquél entonces, numerosas empresas ejercieron una fuerte presión a las empresas aseguradoras dada la baja siniestralidad laboral y la elevada tasa de actividad, para que se les reconociera este hecho en los importes de las primas a satisfacer.

Se debe a Whitney (1918) los primeros trabajos en la materia con la publicación de su trabajo en los *Proceedings* de la *Casualty Actuarial Society*. La idea inicial había sido propuesta por Mowbray (1914). De una forma simple, a través de una matemática elemental, propone que la prima que se debe pagar por parte del asegurado incluya tanto la experiencia individual (del asegurado) como la del

colectivo (la cartera de seguros). De este modo el estimador de la prima se calcularía atendiendo a la siguiente expresión:

$$P = Z \cdot X + (1 - Z) \cdot C, \quad (1.1)$$

donde  $X$  es la experiencia individual,  $C$  la información que se dispone sobre el colectivo y  $Z$  un factor de ponderación denominado factor de credibilidad. Obsérvese que la expresión (1.1) es una combinación lineal convexa entre la experiencia individual  $X$  y la del colectivo  $C$ . Es decir, un punto perteneciente al segmento que conecta  $X$  y  $C$ , como se aprecia en la figura 1.1. La expresión (1.1)

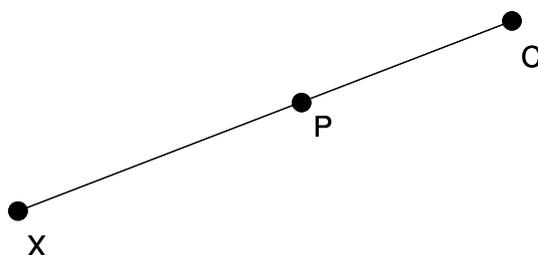


Figura 1.1: Prima basada en la experiencia individual  $X$ , y el colectivo,  $C$ .

dio respuesta formal a la idea que estaba en mente de muchos actuarios de la época: encontrar un mecanismo que permitiese asignar a cada uno de los dos tipos de información, la individual y la colectiva, un peso o ponderación que se complementasen en la determinación de la prima a cobrar.

Intuitivamente el factor de credibilidad  $Z$  debería de satisfacer lo siguiente:

- Ser una función del tiempo de vigencia de la póliza,  $n$ , por tanto  $Z \equiv Z(n)$ ,
- Ser una función creciente en  $n$ , de modo que se aproxime a 1 cuando  $n$  tienda a infinito, mientras que tienda a cero cuando  $n$  tienda a cero. Por tanto, el valor de  $n = 0$  supondría que no se dispone de experiencia para el asegurado (se trataría de un contrato nuevo), y la prima a cobrar en este caso sería  $C$ , esto es, la prima basada en la información acerca del colectivo. En la medida en que aumentase  $n$ , y por tanto se dispusiese de más datos, la información de que se dispone del asegurado tendría mayor peso. Es decir la experiencia de siniestralidad del asegurado tendría más creencia, sería más creíble. El caso extremo,  $n$  tendiendo a infinito, debería proporcionar como valor de la prima  $X$ , esto es que la prima estuviera basada exclusivamente en la experiencia individual.
- El factor de credibilidad  $Z$  debería ser también una función creciente de la varianza de las primas teóricas, con límite 1 cuando aquélla tiende a infinito

y cero cuando tiende a cero. Esto es lógico, puesto que si la cartera no es heterogénea, entonces la prima basada en el colectivo sería el mejor estimador de la prima individual. Por otro lado, mayor heterogeneidad de la cartera debería dar lugar a un mayor peso a la información del asegurado.

A efectos pedagógicos, consideremos el siguiente ejemplo que se propone para una cartera de seguros ficticia.

**Ejemplo 1.1** Supongamos que se dispone de una cartera de seguros compuesta por dos pólizas. Las cantidades reclamadas durante los años 2007 y 2008 aparecen en la tabla 1.1. Atendiendo al criterio de la media de la cantidad reclamada, vamos a estimar la prima de credibilidad basada en la expresión (1.1) para cada uno de los contratos.

Tabla 1.1: Cartera de seguros.

Años	Póliza 1	Póliza 2
2007	20	30
2008	10	20

La cantidad media anual reclamada para cada uno de las dos pólizas es:

$$\text{Póliza 1: } \quad : \quad \bar{x}_1 = 15,$$

$$\text{Póliza 2: } \quad : \quad \bar{x}_2 = 25.$$

Por otro lado, la cantidad media reclamada para la cartera es

$$\text{Cartera de seguros: } \bar{x} = \frac{50 + 30}{2} = 40 = \bar{x}_1 + \bar{x}_2.$$

Basado en la expresión (1.1), la prima a cobrar para cada una de las pólizas será:

$$\text{Póliza 1} \quad : \quad P = Z \cdot 15 + (1 - Z) \cdot 40,$$

$$\text{Póliza 2} \quad : \quad P = Z \cdot 25 + (1 - Z) \cdot 40.$$

Así, para un valor de  $Z = 1$ , la prima a cobrar para las pólizas 1 y 2 coincidiría con la media de reclamaciones de cada uno de ellas, respectivamente. Su propia experiencia de siniestralidad se dice entonces que es creíble al 100 %.

Un valor de  $Z = 0$  provoca que la prima a cobrar sea igual a la cantidad media reclamada del colectivo. Ahora se dirá que la experiencia de cada una de las pólizas no es creíble en absoluto.

Evidentemente cualquier otro valor de  $Z$  conduce a un valor de la prima que pertenece al intervalo  $(X, C)$ .

En la figura 1.2 se muestran los casos posibles según la distancia entre  $P$ ,  $C$  y  $X$ . Así el gráfico (a) muestra el caso en que  $Z < \frac{1}{2}$ . Aquí, la información del colectivo es más creíble que la experiencia del asegurado (póliza), y por tanto  $P$  está más próxima a  $C$ . La situación contraria ocurre en el gráfico (b), en el que  $Z > \frac{1}{2}$  y la experiencia individual es más creíble que la del colectivo. Por supuesto que el caso  $Z = \frac{1}{2}$  nos lleva a un punto situado justo en la mitad del segmento que une  $X$  y  $C$ . Los casos extremos,  $Z = 1$  y  $Z = 0$ , comentados en el ejemplo anterior, se muestran en las figuras (c) y (d), respectivamente.

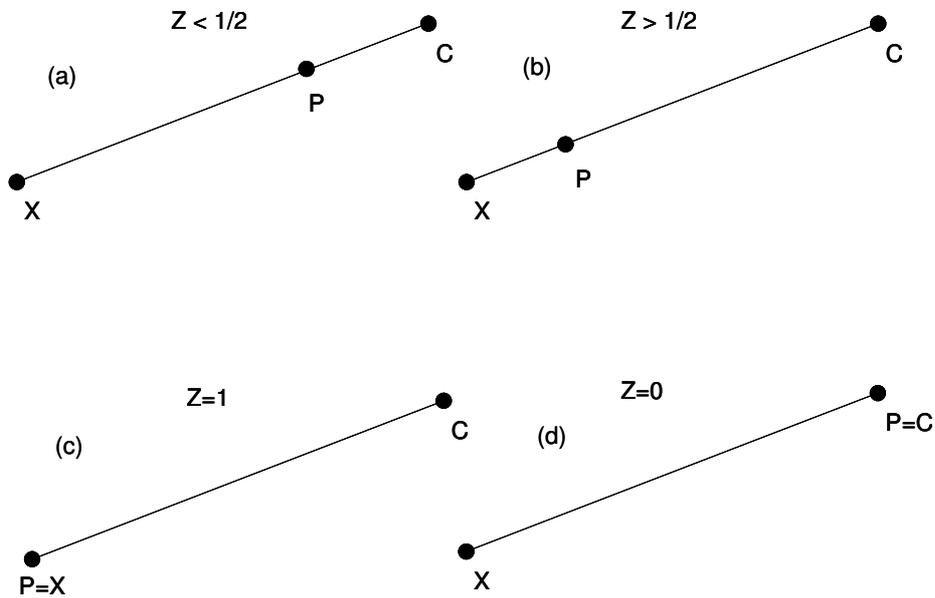


Figura 1.2: Prima basada en la experiencia individual  $X$ , y el colectivo,  $C$ . Diversas posibilidades.

Los distintos modelos de credibilidad desarrollados hasta la fecha, y basados en la expresión (1.1), se han ocupado mayoritariamente de estimar el valor del factor de credibilidad  $Z$ .

La siguiente definición de teoría de la credibilidad fue propuesta por Hickmann (1975).

**Definición 1.1** *La teoría de la credibilidad es el mecanismo que permite el ajuste sistemático de las primas de seguros conforme se obtiene la experiencia de siniestralidad.*

En un sentido más práctico, Hewitt (1970) define credibilidad como un estimador lineal de la prima teórica esperada que sea una combinación convexa entre la hipótesis y la observación.

A partir del trabajo de Whitney (1918) y durante las siguientes décadas, los investigadores que tenían su campo de actuación en la teoría de la credibilidad se dedicaron a dar un carácter riguroso a esta disciplina, desarrollándose así los modelos clásicos (basados en el uso de la estadística clásica) de credibilidad parcial y total. Estos modelos tienen un carácter muy descriptivo.

Posteriormente, en torno a mediados del siglo XX empezaba a tomar cuerpo una nueva visión de la estadística, la Bayesiana, no tardándose mucho en comprobar que muchos estimadores Bayes, obtenidos para ciertas verosimilitudes y su distribución a priori natural conjugada, respondían a la propuesta de Mowbray (1914) y la expresión (1.1) de Whitney (1918), así como a lo formulado en el trabajo de Bailey (1945). De hecho Whitney (1918) ya señalaba que el problema de la credibilidad es un caso de probabilidad inversa (teorema de Bayes). En este aspecto destacamos el trabajo de Mayerson (1964) en el que por vez primera se utilizan los términos credibilidad y estadística Bayesiana.

Bajo esta metodología, la fórmula de credibilidad (1.1) puede interpretarse también de la siguiente manera. Puede considerarse a  $C$  como la información a priori (basada por ejemplo en contratos similares) y  $X$  la nueva información obtenida mediante la observación de la siniestralidad en los últimos años. Finalmente,  $P$  es el resultado de combinar la información a priori junto con la información adquirida para obtener un estimador revisado de la prima. Visto pues de esta manera, la teoría de la credibilidad es un proceso Bayesiano en el que se da entrada a la información a priori junto con la información muestral para obtener un estimador actualizado de la prima. Este proceso de conocimientos a priori y conocimientos actuales se presenta en la figura 1.3.

En 1975 se publica el texto *Credibility. Theory and Applications* que contiene los trabajos, así como discusiones de los mismos, presentados en el *Actuarial Research Conference on Credibility* del 19 al 21 de septiembre de 1974 en la Universidad de California en Berkeley. Este texto incluye todo el conocimiento existente hasta la fecha sobre la teoría de la credibilidad, tanto en seguros vida como no vida, incluyendo reaseguros e IBNR (*Incurred But Not Reported*), es decir la estimación de reservas para reclamaciones pendientes.

Lo que hoy en día se entiende por teoría de la credibilidad moderna surge a finales de la década de los 60 del siglo pasado con la publicación del trabajo de Bühlmann (1967), en los que desarrolla el llamado modelo de distribución libre. El objetivo que se plantea el autor en dichos trabajos es estimar la prima correspondiente a un asegurado o grupo de asegurados que conforman una póliza en una cartera de seguros, restringiéndose a las primas lineales y utilizando el método de los mínimos cuadrados. La solución que obtuvo es de una elegancia aplastante, fundamentalmente por dos motivos. En primer lugar, no asigna ninguna distribu-

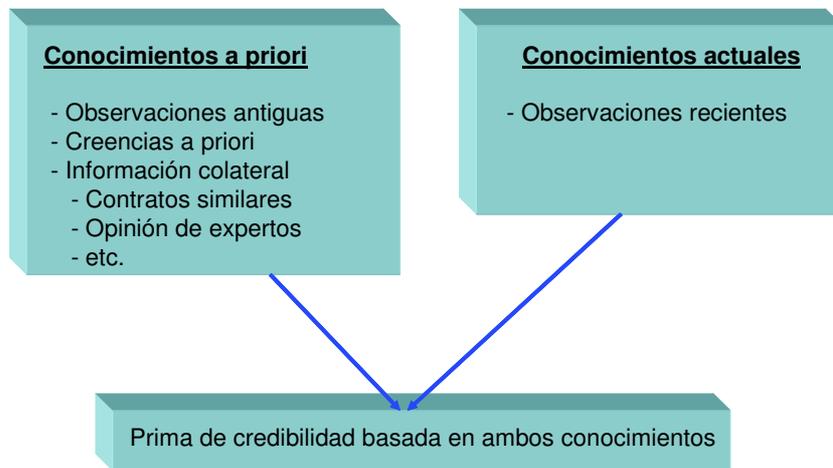


Figura 1.3: Prima de credibilidad: basada en conocimientos a priori y actuales.

ción de probabilidad ni a los riesgos que conforman la cartera de seguros ni a los parámetros de los que dependen dichos riesgos. En segundo lugar, proporcionó la expresión que debía tener el factor de credibilidad. Dicha expresión, a su vez, admite interpretaciones que son bastante plausibles.

Posteriormente en Bühlmann y Straub (1972) se generaliza el modelo clásico de Bühlmann (1967) a una cartera de seguros que proporciona información sobre la cantidad reclamada en cada una de las pólizas así como del número de asegurados. La solución obtenida es más general que la del modelo de Bühlmann (1967), además de incluirla como un caso particular.

Pocos años más tarde, Jewell (1974) se planteó el problema de considerar como verosimilitud asociada al número de reclamaciones, la familia exponencial de distribuciones (FED), que contiene a las distribuciones clásicas de Poisson, binomial, binomial negativa, gamma, normal y otras. Para esta familia existe una distribución a priori natural conjugada (esto es, la distribución a posteriori pertenece a la FED pero con los parámetros actualizados con las observaciones muestrales). Concluyó que el estimador de la prima neta (basada en el criterio de la media) admitía siempre una expresión lineal, con el factor de credibilidad igual al propuesto por Bühlmann (1967).

Recientemente Landsman y Makov (1998) han generalizado el resultado de Jewell (1974) al caso de la familia de distribuciones exponencial de dispersión. Esta familia tiene la ventaja con respecto a la familia exponencial de distribuciones, de incorporar un parámetro que posibilita la presencia de infradispersión, equidispersión o sobredispersión en los datos.

Bühlmann (1975) ha interpretado el problema de decisión del actuario para elegir una prima de seguro a cobrar como un juego entre dos jugadores, el actuario y la naturaleza utilizando el criterio de decisión minimax. De esta manera el problema de la teoría de la credibilidad se ha llevado también al terreno de la teoría de juegos.

Basados en ideas tomadas de la teoría de la decisión y la teoría de juegos, aparecen los modelos gamma-minimax (Eichenauer et al., 1988) y posterior regret gamma-minimax (Gómez-Déniz et al., 2006a y Gómez-Déniz, 2008b). Estos modelos surgen en el entramado de la estadística Bayesiana robusta, en la que se asume que el investigador no tiene conocimiento exacto sobre la distribución a priori del parámetro de riesgo. En todos estos modelos, minimax, gamma-minimax y posterior regret gamma-minimax se derivan fórmulas de credibilidad con el factor de credibilidad igual al de Bühlmann (1967).

A finales de la década de los 80 del siglo pasado Heilmann (1989) extendió los modelos de credibilidad, hasta el momento basados en el criterio de la media o pérdida cuadrática, a otras funciones de pérdidas. Obtuvo entonces nuevas fórmulas de credibilidad bajo distintos pares de verosimilitudes y distribuciones a priori y las pérdidas de varianza y Esscher. El factor de credibilidad que obtuvo era de nuevo igual al de Bühlmann (1967).

Siguiendo también esta línea de trabajo, Gómez-Déniz (2008a) también ha generalizado los distintos modelos de credibilidad existentes hasta la fecha, al utilizar una función de pérdida más general que la cuadrática ponderada, generadora de los principales principios de cálculo de primas, neta, exponencial, Esscher, varianza y otros. La función utilizada es la pérdida cuadrática ponderada y equilibrada, propuesta en Zellner (1994).

Puesto que la distribución a priori del parámetro de riesgo depende, a su vez, de ciertos parámetros (hiperparámetros) es posible considerar alguno de ellos como no conocido y aleatorio. El análisis Bayesiano permite atribuirle a su vez una distribución a priori (hiperpriori). Se entra así en un modelo jerárquico, que ya había sido estudiado, desde una perspectiva clásica por Bühlmann-Straub (1972) y que Klugman (1992) y Gómez-Déniz et al. (2008c) han elaborado siguiendo la metodología bayesiana jerárquica. En ciertos casos es posible obtener de nuevo fórmulas de credibilidad, aunque en esta ocasión el factor de credibilidad no es igual que el de Bühlmann (1967).

Recientemente han aparecido nuevos tipos de modelos de credibilidad que se apartan de la ortodoxia. Algunos de estos modelos permiten incorporar experiencia de siniestralidad para más de una aseguradora (Guillén et al., 2008). Otro tipo de modelos tienen en cuenta el hecho empíricamente contrastado de que las distribuciones del número de siniestros en el ramo de automóviles están muy cargadas de ceros, incorporando las llamadas distribuciones infladas de ceros (Boucher y Denuit, 2007). Se han propuesto además modelos de credibilidad tipo panel, basados en varios períodos de tiempo (Gajek et al., 2007).

En conclusión: podemos señalar que la teoría de la credibilidad se ocupa básicamente de medir la importancia que ha de tener la información de un individuo o contrato frente a la información del colectivo (la cartera de seguros) en relación a la prima que debe satisfacer. De ahí que con esta teoría se pretende estimar las primas de seguros a medida que se va obteniendo la experiencia de siniestralidad. Partiendo de las ideas iniciales de Whitney y Mowbray, basadas en argumentos heurísticos con desarrollos matemáticos parciales, Bailey, Mayerson, Bühlmann, Straub y Jewell, entre otros, se encargaron posteriormente de proporcionar a esta teoría un fundamento matemático que ha hecho de la misma una de las ramas más atractivas y estudiadas de la ciencia actuarial.

Una de las principales aplicaciones de dicha teoría se presenta en el seguro de automóviles, en el que la prima inicial se va transformando sucesivamente a medida que se incorpora la información de la siniestralidad (Guillén et al, 2005). Son los denominados sistemas de tarificación bonus-malus. Se parte de un nivel  $x$  neutro, de modo que para niveles superiores a  $x$  el asegurado entra en la escala *malus* y para niveles inferiores a  $x$  el asegurado entra en la escala *bonus*. Con este sistema el asegurado puede ver bonificada o penalizada su prima particular dependiendo de su propia experiencia de reclamaciones.

En una visión más amplia del término, el concepto credibilidad no incluye únicamente expresiones del tipo (1.1), sino que incluye además otras expresiones de riesgo que incorporen tanto la experiencia reciente del riesgo como la experiencia de un colectivo. La expresión (1.1) presenta la ventaja de distribuir el peso que sobre la prima total tiene la experiencia individual.

En definitiva, y sin ser exhaustivos, presentamos a continuación de forma resumida la evolución temporal de los distintos modelos desarrollados en teoría de la credibilidad.

ALGUNOS MODELOS DE CREDIBILIDAD

1. Modelo inicial de Whitney (1918)
2. Modelos clásicos: credibilidad parcial y credibilidad total
3. Modelos Bayesianos (Bailey, Mayerson, 1945)
4. Modelo de Bühlmann de distribución libre, 1967
5. Modelo de Bühlmann-Straub, 1972
6. Modelo de Jewell (FED), 1974
7. Modelo de credibilidad con regresores (Hachemeister, 1975)
8. Modelo minimax (Bühlmann, 1975)
9. Modelo Bayesiano general (Heilmann, 1989)
10. Modelos jerárquicos (Klugman, 1992, Gómez et al., 2008c)
11. Modelo de Landsman y Makov (FDE), 1998
12. Modelo basado en funciones de pérdida generales (Gómez-Déniz, 2008a)
13. Modelos basados en robustez Bayesiana
  - a) Aproximación gamma-minimax (Eichenauer et al., 1988)
  - b) Aproximación posterior regret-gamma-minimax (Gómez-Déniz et al., 2006a)
14. Modelo basado en distribuciones infladas de ceros (Boucher and Denuit, 2007)
15. Modelo de credibilidad basado en múltiples períodos de tiempo (Gajek et al., 2007)
16. Modelos de credibilidad múltiples (Guillén et al., 2008)

En esta investigación se estudiarán algunos de estos modelos de credibilidad, destacando los modelos clásicos de credibilidad parcial y total, los modelos Bayesianos, el modelo de Bühlmann de distribución libre, el modelo de Bühlmann-Straub, el modelo de Jewell, el de Landsman y Makov, el modelo general bayesiano, los modelos jerárquicos, el modelo basado en funciones de pérdida generales, así como algunos modelos basados en el uso de la estadística Bayesiana robusta, gamma-minimax y posterior regret gamma-minimax. El lector interesado en el resto de modelos puede consultar la bibliografía señalada.

## Capítulo 2

# Principios de Primas y Medidas de Riesgo

### 2.1. Introducción

Los principios de cálculo de primas, así como la metodología utilizada para su obtención, constituye un elemento imprescindible en la tarificación mediante credibilidad. Como es bien conocido, la cobertura de un riesgo por parte de una compañía de seguros se establece con la garantía de un contrato, la *póliza* y exige al asegurado pagar un precio, la *prima*. Ajustar una prima como una combinación lineal convexa entre la prima debida únicamente a su experiencia personal y la que debería pagar por el hecho de pertenecer a un colectivo con unas características peculiares (la cartera), forma parte del escenario de la *teoría de la credibilidad*.

En este capítulo estudiaremos detalladamente los principios de primas y las medidas de riesgos. En la sección 2.2 comenzaremos estudiando generalidades sobre las primas. Una prima requiere una serie de propiedades deseables que debe satisfacer. Los diferentes métodos para la obtención de primas y las propiedades que se deben verificar se estudiarán en la sección 2.3. Los métodos para la obtención de primas incluyen los métodos ad-hoc, los basados en caracterizaciones, los métodos basados en modelos económicos y los métodos de las funciones de pérdida. Las propiedades deseables de una medida de riesgo serán estudiadas en la sección 2.4.

Las medidas de riesgo y primas basadas en la cola de la distribución (es decir el VaR y el TVaR) ocupan un lugar importante en la investigación actual en seguros. Dichas medidas se estudiarán en la sección 2.5. Finalmente, las secciones 2.6 y 2.7 se dedican al estudio de algunas medidas de riesgo y primas especiales: las primas ponderadas y las basadas en los estadísticos ordenados.

## 2.2. Primas

El cálculo de la prima se basa en la idea simple y coherente de que la experiencia de reclamaciones ha de estar compensada por pagos fijos, las primas (Bühlmann, 1995). Por lo tanto, ésta no es más que el precio para el seguro (o reaseguro) vendido por la compañía aseguradora. Puesto que lo que la compañía vende es la cobertura de un riesgo, podemos especificar más la definición anterior como sigue.

**Definición 2.1** *La prima es el pago que un asegurado hace a un asegurador por la cobertura total o parcial contra un riesgo.*

De forma reducida, una prima mínima técnica está compuesta de los siguientes elementos:

- Prima pura de riesgo.
- Recargo de seguridad.
- Costo adicional para el beneficio.

Por tanto, la prima o precio del servicio es el coste que para la empresa suponen los siniestros más el margen de beneficio. En este texto nos centraremos en las dos primeras componentes y no haremos mención a la tercera, que en principio no parece tener un componente estocástico.

Para su cálculo se requiere algún estimador del número de reclamaciones esperado y de la cantidad asociada a cada una de dichas reclamaciones. En la práctica es habitual considerar que estos dos componentes, el número de reclamaciones y la cuantía asociada a cada una de las mismas, son independientes. La razón, a parte de conveniencia matemática, obedece a lo siguiente. La mayoría de los factores de riesgo afectan de manera separada a cada una de las componentes. Por ejemplo, una legislación que obligue a llevar el cinturón de seguridad abrochado en el vehículo afecta a la cuantía del daño pero no a la frecuencia del suceso. Por el contrario, una medida más restrictiva sobre la tasa de alcohol a la que un conductor puede pilotar un vehículo afectará al número de reclamaciones pero apenas tendrá efecto sobre la cuantía.

Por otro lado, la forma de recogida de datos por parte de las compañías aseguradoras determinará la metodología de trabajo, y son las siguientes:

1. En algunas ocasiones las compañías recogen datos solamente de los costes totales en unidades monetarias generadas por cada póliza cada año. En este caso la única vía para trabajar es utilizar estos costes totales y todos los modelos y/o estimadores se referirán a la distribución del coste total.
2. En otras ocasiones los datos se recogen separadamente para número de siniestros y coste de cada uno de ellos. La vía para trabajar es, en este caso, componer los dos modelos, del número de siniestros y de la cuantía de los mismos, para obtener la distribución del coste total o cantidad total reclamada.

3. Por último, en algunas ocasiones se cree (algunas compañías lo hacen para seguros de automóviles) que el número de siniestros es la única componente que está bajo el control del asegurador. Esto se piensa así porque una vez ocurrido un siniestro, la cuantía del mismo está fuera de control del asegurador. En este caso la vía para trabajar es hacerlo solamente con el número de siniestros.

En la práctica, si el cálculo de la prima se lleva a cabo teniendo en cuenta sólo el número de reclamaciones, la prima resultante se obtendrá multiplicando ésta por la cantidad media de reclamación, siempre que se haya asumido independencia entre las dos componentes.

El precio correcto, que es llamado *rating*, es vital, pues si es demasiado bajo representa una pérdida para la compañía aseguradora y si es demasiado alto se pierde competitividad frente a otras. Por tanto, una de las labores del actuario consiste en encontrar métodos de cálculo de primas, generalmente llamados en la literatura actuarial *principios de cálculo de primas*.

## 2.3. Métodos para la obtención de primas

En el contexto de la estadística actuarial y el análisis de riesgos, un riesgo equivale a una variable aleatoria. En lo que sigue denotamos por  $X$  la variable aleatoria número de reclamaciones, cantidad reclamada (severidad) o bien una combinación de ambas. Tenemos la siguiente definición de principio de cálculo de prima.

**Definición 2.2** *Un principio de cálculo de prima es una función  $H[X]$  que asigna a un riesgo  $X$  un número real, que es la prima.*

En la práctica, el principio de cálculo de prima dependerá de la función de distribución  $F_X(x)$  que siga la variable aleatoria  $X$ , por lo que en vez de hablar de  $H[X]$  como una función debería hablarse en términos más precisos del funcional  $H[F_X]$  como aparece en Gerber (1979).

En las siguientes secciones describiremos cuatro métodos para la obtención de primas:

- Método ad-hoc.
- Método basado en caracterizaciones.
- Método económico.
- Método de las funciones de pérdida.

### 2.3.1. Propiedades de los principios de cálculo de primas

Como paso previo al estudio de los métodos para la obtención de primas, pasamos a establecer una serie de propiedades deseables o axiomas que debería de satisfacer un principio de prima.

No existe en la literatura actuarial un sistema axiomático comúnmente aceptado de propiedades que un principio de cálculo de prima debería satisfacer. Sin ser exhaustivos, contribuciones importantes en esta materia pueden encontrarse en Gerber (1979), Heilmann (1989), Hürlimann (1994) y Young (2004). Destacamos las siguientes propiedades que un principio de cálculo de prima  $H[X]$  debería satisfacer:

1. *Independencia.*  $H[X]$  depende únicamente de la función de distribución de  $X$ .

Esta propiedad establece que la prima depende sólo de la pérdida monetaria debida al riesgo (descrita a través de la función de distribución) y no de la causa de esta pérdida.

2. *Recargo de seguridad no negativo.*

$$H[X] \geq E[X],$$

Esto significa que para evitar la ruina técnica la ganancia esperada  $H[X] - E[X]$  será no negativa.

3. *Riesgos constantes:*  $H[c] = c$ , para toda constante  $c \geq 0$ .

Esto significa que para un riesgo no aleatorio  $X = c$ , con  $\Pr(X = c) = 1$ , la prima a cobrar será  $c$ .

4. *No estafa.* La prima no excederá a la reclamación máxima posible:  $\sup[X]$ :

$$H[X] \leq \sup[X].$$

5. *Consistencia.* Para cada riesgo  $X$  y cada constante  $c$ ,

$$H[X + c] = H[X] + c.$$

Esto significa que si el beneficio se incrementa en una constante esta constante tiene que ser añadida a la prima.

6. *Homogeneidad positiva.*

$$H[cX] = cH[X], \text{ para todo } c \geq 0,$$

que resulta conveniente para corregir efectos inflacionarios.

7. *Aditividad*: Para todo  $X_1, X_2 \in \mathcal{X}$ :

$$H[X_1 + X_2] = H[X_1] + H[X_2].$$

8. *Subaditividad*: Para todo  $X_1, X_2 \in \mathcal{X}$ :

$$H[X_1 + X_2] \leq H[X_1] + H[X_2].$$

9. *Superaditividad*: Para todo  $X_1, X_2 \in \mathcal{X}$ :

$$H[X_1 + X_2] \geq H[X_1] + H[X_2].$$

10. *Aditividad de riesgos independientes* Si  $X_1$  y  $X_2$  son riesgos independientes entonces:

$$H[X_1 + X_2] = H[X_1] + H[X_2].$$

Esto quiere decir que la incorporación de riesgos independientes afecta a la prima total de modo aditivo.

11. *Aditividad para riesgos comonótonos*:

$$H[X_1 + X_2] = H[X_1] + H[X_2],$$

para todo  $X_1, X_2 \in \mathcal{X}$  riesgos comonótonos.

12. *Monotonicidad*. Si  $X_1(w) \leq X_2(w)$  para todo  $w \in \Omega$ , entonces  $H[X_1] \leq H[X_2]$ .

13. *La prima preserva la ordenación estocástica de primer orden*: Si  $S_{X_1}(t) \leq S_{X_2}(t)$ , para todo  $t \geq 0$ , entonces:  $H[X_1] \leq H[X_2]$ .

14. *La prima preserva la ordenación estocástica stop-loss*: Si  $E[X_1 - d]_+ \leq E[X_2 - d]_+$ , para todo  $d \geq 0$ , entonces:  $H[X_1] \leq H[X_2]$ .

15. *Continuidad*. Si  $X \in \mathcal{X}$ , entonces:

$$\lim_{a \rightarrow 0^+} H[\text{máx}(X - a, 0)] = H[X],$$

y

$$\lim_{a \rightarrow \infty} H[\text{mín}(X, a)] = H[X].$$

16. *Iteratividad*. Si  $X$  y  $\Theta$  son dos riesgos arbitrarios dependientes, entonces:

$$H[X] = H[H[X|\Theta]].$$

Esto significa que la prima para  $X$  puede calcularse en dos pasos. Primero calcular la prima condicionada  $H[X|\Theta]$ , aplicando  $H$  a la distribución condicional de  $X$ . Esta prima condicional es una función de  $\Theta$  y por lo tanto una variable aleatoria en sí misma. Entonces se aplica  $H$  a la distribución de  $H[X|\Theta]$ , para obtener  $H[H[X|\Theta]]$ .

### 2.3.2. Método ad-hoc: catálogo de principios de prima

En esta sección incluimos un catálogo de los principales principios de prima y a continuación incluiremos las propiedades que satisfacen cada uno de ellos. El método ad-hoc para elegir una prima consiste en seleccionar las propiedades adecuadas que debe de satisfacer (de acuerdo con la sección anterior) para a continuación elegir la correspondiente prima.

1. Principio de prima neta:

$$H[X] = E[X]. \quad (2.1)$$

2. Principio del valor esperado:

$$H[X] = (1 + \lambda)E[X], \quad \lambda > 0. \quad (2.2)$$

En este caso el recargo de seguridad es  $\lambda E(X)$ .

El principio de prima neta, así como el de valor esperado raramente se utilizan en seguros de propiedad y daños a terceros. La razón es sencilla, y es que asigna la misma prima a todos los riesgos con la misma media, cuando el colectivo de asegurados es heterogéneo. Además, el primero, no incluye el recargo de seguridad.

3. Principio de varianza:

$$H[X] = E[X] + \lambda \text{Var}[X], \quad \lambda > 0. \quad (2.3)$$

La ventaja de este principio es que no sólo estima la siniestralidad media del riesgo, sino que proporciona además el recargo de seguridad que debe llevar la prima para atender a las desviaciones aleatorias de la siniestralidad. Mediante la expresión anterior se dice que el recargo de seguridad es proporcional a la varianza.

4. Principio de desviación típica:

$$H[X] = E[X] + \lambda \sqrt{\text{Var}[X]}, \quad \lambda > 0. \quad (2.4)$$

Este principio ha sido utilizado por Balbás, Gil y Heras (1990) como medida de riesgo en un problema de reaseguro óptimo.

El principio de desviación típica y el de varianza son los más utilizados en estos los tipos de seguros mencionados anteriormente.

5. Principio de utilidad exponencial:

$$H[X] = \frac{1}{\alpha} \log E[e^{\alpha X}], \quad \alpha > 0. \quad (2.5)$$

Este principio de cálculo de prima viene dado en términos del logaritmo de la función generatriz de momentos de la variable aleatoria  $X$ . En teoría de la decisión a  $\alpha$  se le denomina *constante de aversión al riesgo* (también llamada medida de Arrow-Pratt) asociada al decisor. Aunque posteriormente se obtendrá este principio utilizando una función de pérdida, el nombre de principio de utilidad esperada obedece a que puede obtenerse a partir de la función de utilidad

$$U(z) = \frac{1}{\alpha} (1 - e^{-\alpha z}), \quad (2.6)$$

donde  $z$  mide la riqueza de la aseguradora. Para calcular la prima en este caso se elige  $P$  que satisfaga la ecuación

$$U(z) = \int U(z + H[X] - x) dx, \quad (2.7)$$

donde se iguala la utilidad antes de la operación de aseguramiento con la utilidad esperada después de dicha operación. Sustituyendo (2.7) en (2.6) se obtiene, después de algunos cálculos, la expresión (2.5).

Puede probarse (Gerber, 1979) que la prima exponencial crece conforme lo hace la constante de aversión al riesgo  $\alpha$  y decrece con ella. Además cuando  $\alpha \rightarrow 0$  se obtiene el principio de prima neta y cuando  $\alpha \rightarrow \infty$  la prima resultante es el máximo valor de  $X$  (Kaas et al., 2001).

6. Principio de Esscher:

$$H[X] = \frac{E[Xe^{\alpha X}]}{E[e^{\alpha X}]}, \quad \alpha > 0. \quad (2.8)$$

El principio Esscher admite también la siguiente interpretación. Si  $f(x)$  es la función de densidad de probabilidad asociada al riesgo  $X$ , es obvio que

$$g(x) = \frac{e^{\alpha x} f(x)}{\int e^{\alpha x} f(x) dx} \quad (2.9)$$

es una función de densidad de probabilidad genuina. Así, la prima Esscher no es más que la prima neta asociada al riesgo  $X$  bajo la densidad (2.9). Si la variable aleatoria  $Y$  sigue la función de densidad (2.9), su función de distribución es

$$G_X(x) = \frac{\int_0^x e^{\alpha y} f(y) dy}{M_Y(\alpha)}, \quad (2.10)$$

con  $M_Y(\alpha) = \int_0^\infty e^{\alpha y} f(y) dy$ . (2.10) recibe también el nombre de transformada de Esscher de  $F(x) = \int_0^x f(t) dt$ , de ahí que el principio al que hacemos referencia se le conozca como principio de prima Esscher.

7. Principio de prima de azar proporcional:

$$H[X] = \int_0^{\infty} [S_X(t)]^c dt, \quad 0 < c < 1, \quad (2.11)$$

donde  $S_X(t) = \Pr(X > t)$  es la función de supervivencia de  $X$ .

8. Principio de utilidad equivalente. La prima  $H[X]$  debe satisfacer,

$$u(w) = E[u(w - X + H)], \quad (2.12)$$

donde  $u(\cdot)$  es la función de utilidad del asegurado, que es creciente y cóncava, y donde  $w$  representa la riqueza inicial.

9. Principio de prima de Wang:

$$H[X] = \int_0^{\infty} g[S_X(t)] dt, \quad (2.13)$$

donde  $g : [0, 1] \rightarrow [0, 1]$  es una función creciente, cóncava, llamada función de distorsión de la cola. Este principio general da lugar a una amplia clase de primas que han sido detalladamente estudiadas en la literatura (Denneberg (1994) y Wang (1995, 1996)), donde se han propuesto diversas formas funcionales para  $g$ .

10. Principio de prima Suizo. La prima  $H$  es la solución de la ecuación:

$$E[v(X - zH)] = v((1 - z)H), \quad (2.14)$$

en la que  $v(\cdot)$  es una función peso verificando  $v'(z) > 0$ ,  $v''(z) \geq 0$ , si  $z > 0$ .

11. Principio de Prima Holandés. La prima es:

$$H[X] = E[X] + \lambda E[(X - \alpha E(X))_+], \quad \alpha \geq 1, \quad 0 < \lambda \leq 1, \quad (2.15)$$

donde  $X_+ = \max\{X, 0\}$ .

Las propiedades que satisfacen los diferentes principios de cálculo de primas se recogen en la tabla 2.1.

### 2.3.3. Método basado en caracterizaciones

El método basado en caracterizaciones consiste en elegir aquellos principios de primas que vienen caracterizados por un determinado conjunto de propiedades deseables. Por ejemplo, Furman y Zitikis (2008) han caracterizado primas ponderadas en términos de propiedades de invarianza frente a transformaciones de escala, traslaciones y aditividad.

<b>Principio de Prima</b>	<b>Neta</b>	<b>V. Esp</b>	<b>Var</b>	<b>Desv. T</b>	<b>Expo</b>	<b>Esscher</b>	<b>Azar</b>	<b>Util</b>	<b>Wang</b>	<b>Suizo</b>	<b>Holandes</b>
<b>Propiedad</b>											
1	+	+	+	+	+	+	+	+	+	+	+
2	+	+	+	+	+	-	+	+	+	+	+
3	+	-	+	+	+	+	+	+	+	+	+
4	+	-	-	-	+	+	+	+	+	+	+
5	+	-	+	+	+	+	+	+	+	-	-
6	+	+	-	+	-	-	+	-	+	-	+
7	+	+	-	-	-	-	-	-	-	-	-
8	+	+	-	-	-	-	+	-	+	-	+
9	+	+	-	-	-	-	-	-	-	-	-
10	+	+	+	-	+	-	-	-	-	-	-
11	+	+	-	-	-	-	+	-	+	-	-
12	+	+	-	-	+	-	+	+	+	+	+
13	+	+	-	-	+	-	+	+	+	+	+
14	+	+	-	-	+	-	+	+	+	+	+
15	+	+	+	+	+	+	+	+	+	+	+
16	+	+	-	+	+	-	-	-	-	-	-

Figura 2.1: Propiedades de los principios de cálculo de primas (Kass et al. (2001), Young (2004) y Sarabia et al. (2006)).

### 2.3.4. Método económico

La teoría económica de la utilidad esperada puede utilizarse para la obtención de primas. Por ejemplo, el principio de prima de Esscher puede obtenerse a partir de la teoría de la utilidad esperada. Supongamos que un asegurador  $j$  se enfrenta a un riesgo  $X_j$ ,  $j = 1, 2, \dots, n$ . Supongamos que el asegurador  $j$  tiene una función de utilidad exponencial  $u_j(w) = 1 - e^{-\alpha_j w}$ . Bühlmann (1995) define un equilibrio de modo que cada agente maximiza su utilidad esperada. Este equilibrio coincide con el óptimo de Pareto.

**Teorema 2.1** *Con las hipótesis anteriores, el precio para el riesgo  $X$  en el equilibrio viene dado por,*

$$H[X] = \frac{E[Xe^{\alpha Z}]}{E[e^{\alpha Z}]},$$

donde  $Z = X_1 + \dots + X_n$  es el riesgo agregado y  $\alpha = \{\sum_{i=1}^n \alpha_i^{-1}\}^{-1}$ .

### 2.3.5. Funciones de pérdida y prima de riesgo

Consideremos ahora una función de pérdida  $L : \mathbb{R}^2 \rightarrow \mathbb{R}$  que atribuya a algún  $(x, P) \in \mathbb{R}^2$  la pérdida soportada por un decisor que toma la acción  $P$  y se

encuentra con el resultado  $x$  de algún experimento aleatorio. La *prima de riesgo* se define de la siguiente manera.

**Definición 2.3** *Dados un riesgo  $X$  con función de distribución  $F_X(x)$  y una función de pérdida  $L : \mathbb{R}^2 \rightarrow \mathbb{R}$ , la prima de riesgo es el valor de  $H$  que minimiza la pérdida esperada*

$$\int_{\mathcal{X}} L(x, H) dF_X(x) = E[L(X, H)], \quad (2.16)$$

donde  $x$  es el resultado del experimento aleatorio  $X$  y  $H$  la prima cobrada por tomar  $x$ .

Obviamente, si  $X$  es discreta  $H$  deberá minimizar la pérdida esperada

$$\sum_{x=0}^{\infty} L(x, H) \Pr(X = x),$$

donde  $\Pr(X = x) = f_X(x)$  es la función de densidad discreta o función de cuantía.

Esta metodología de cálculo de prima utilizando funciones de pérdidas, fue propuesta por Heilmann (1989), obteniendo de esta manera muchos de los principios de cálculo de prima que ya se utilizaban así como otros nuevos.

Para obtener las diversas primas de riesgo consideramos las funciones de pérdida:

- Cuadrática
- Exponencial
- Cuadrática ponderada

Tenemos los siguientes resultados:

**Teorema 2.2** *Si consideramos la función de pérdida cuadrática, dada por*

$$L(x, H) = (x - H)^2,$$

*resulta,*

$$H[X] = E[X], \quad (2.17)$$

*el principio de prima neta o de equivalencia.*

**Demostración:** Consideremos el funcional,

$$\int L(x, H) dF_X(x) = \int (x - H)^2 dF_X(x).$$

Derivando con respecto a  $H$  e igualando a cero obtenemos:

$$-2 \int (x - H) dF_X(x) = 0,$$

de donde se obtiene el principio de prima neta. ■

**Teorema 2.3** Si consideramos la función de pérdida exponencial, dada por

$$L(x, H) = \frac{1}{\alpha}(e^{\alpha x} - e^{\alpha H})^2,$$

con  $\alpha > 0$ , resulta,

$$H[X] = \frac{1}{\alpha} \log E[e^{\alpha X}], \quad (2.18)$$

que es el principio de utilidad exponencial

**Teorema 2.4** Si consideramos la función de pérdida cuadrática ponderada, con peso  $g(x) = e^{\alpha x}$ , dada por

$$L(x, H) = e^{\alpha x}(x - H)^2,$$

con  $\alpha > 0$ , entonces

$$H[X] = \frac{E[Xe^{\alpha X}]}{E[e^{\alpha X}]}, \quad (2.19)$$

que es el principio de prima Esscher.

**Teorema 2.5** Si consideramos la función de pérdida cuadrática ponderada, con peso  $g(x) = x$ , dada por

$$L(x, H) = x(x - H)^2,$$

entonces:

$$H[X] = \frac{E[X^2]}{E[X]} = E[X] + \frac{Var[X]}{E[X]}, \quad (2.20)$$

que es el principio de varianza.

El siguiente resultado se muestra útil para probar la propiedad de recargo de seguridad no negativo de los principios de cálculo de prima antes presentados.

**Teorema 2.6** Si  $h(x) = x$ , es decir  $L(x, H) = g(x)(x - H)^2$  y  $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , entonces  $g(x)$  es creciente (decreciente) si y sólo si para todo riesgo  $X$  con  $\Pr(X > 0) = 1$  se verifica:

$$H[X] = \frac{E[g(X)X]}{E[g(X)]} \geq E[X], \quad (2.21)$$

o bien,

$$H[X] = \frac{E[g(X)X]}{E[g(X)]} \leq E[X], \quad (2.22)$$

si  $g(x)$  es decreciente.

En general, para obtener las diversas primas de riesgo consideramos funciones de pérdidas de la forma

$$L(x, H) = g(x) [h(x) - h(H)]^2, \quad (2.23)$$

donde  $g(x)$  y  $h(x)$  son funciones apropiadas cuyas esperanzas bajo  $f(\cdot)$  existen. Ahora utilizando (2.16) y el siguiente resultado podemos obtener la prima de riesgo.

**Teorema 2.7** *Si  $h(x)$  es estrictamente creciente y diferenciable (y por tanto invertible) y  $g(x)$  es no negativa, entonces para todo riesgo  $X$  con  $E[g(X)] < \infty$  y  $E[g(X)h(X)] < \infty$ , se tiene que*

$$H[X] = h^{-1} \left\{ \frac{E[g(X)h(X)]}{E[g(X)]} \right\}. \quad (2.24)$$

**Demostración:** Derivando con respecto  $H$  el funcional

$$\int L(x, H) dF_X(x) = \int g(x) [h(x) - h(H)]^2 dF_X(x),$$

resulta

$$-2 \int g(x) [h(x) - h(H)] h'(H) dF_X(x) = 0,$$

de donde

$$h'(H) \int g(x) h(x) dF_X(x) - h(H) h'(H) \int g(x) dF_X(x) = 0.$$

Ahora, puesto que  $h'(H) > 0$  resulta:

$$h(H) = \frac{\int g(x) h(x) dF_X(x)}{\int g(x) dF_X(x)},$$

y por tanto:

$$H[X] = h^{-1} \left\{ \frac{\int g(x) h(x) dF_X(x)}{\int g(x) dF_X(x)} \right\} = h^{-1} \left\{ \frac{E[g(X)h(X)]}{E[g(X)]} \right\}.$$

Si aplicamos el teorema 2.7 con las pérdidas cuadrática, exponencial y cuadrática ponderada obtenemos los principios antes vistos. ■

Por otro lado, cabe también la posibilidad de generar principios de cálculo de primas considerando otras funciones de pérdidas. De este modo, bajo la función de pérdida:

$$L(x, H) = \frac{(x - H)^2}{x}, \quad x > 0, \quad \Pr(X > 0) = 1,$$

se obtiene la prima de riesgo,

$$H[X] = \{E[X^{-1}]\}^{-1}.$$

Si se considera la función de pérdida

$$L(x, H) = \frac{(x - H)^2}{x(x + 1)}, \quad x > 0, \quad \Pr(X > 0) = 1,$$

se obtiene la prima de riesgo

$$H[X] = \frac{E[(X + 1)^{-1}]}{E[(X(X + 1))^{-1}]}.$$

Finalmente, bajo la función de pérdida

$$L(x, H) = (\log x - \log H)^2, \quad x > 0, \quad H > 0, \quad \Pr(X > 0) = 1,$$

la prima de riesgo resultante es:

$$H[X] = \exp\{E[\log(X)]\},$$

que coincide con el riesgo media geométrica.

**Ejemplo 2.1** Vamos a probar que la prima Esscher cumple las propiedades 2, 3, 4, 5 y 7 de las propiedades deseables para un principio de cálculo de prima.

- *Recargo de seguridad no negativo.* El principio Esscher se obtiene a partir de la función de pérdida  $L(x, H) = e^{\alpha x}(x - H)^2$ , que es de la forma (2.23), por tanto haciendo uso del resultado (2.21) se obtiene

$$H[X] = \frac{E[Xe^{\alpha X}]}{E[e^{\alpha X}]} \geq E[X].$$

- *Riesgos constantes:*  $H[c] = \frac{E[ce^{\alpha c}]}{E[e^{\alpha c}]} = \frac{ce^{\alpha c}}{e^{\alpha c}} = c.$
- *No estafa.* Se verifica que  $X \leq \sup[X]$  y por tanto  $Xe^{\alpha X} \leq \sup[X]e^{\alpha X}$  de donde:

$$E[Xe^{\alpha X}] \leq \sup[X]E[e^{\alpha X}],$$

$$\text{y por tanto } H[X] = \frac{E[Xe^{\alpha X}]}{E[e^{\alpha X}]} \leq \sup[X].$$

- *Consistencia:*

$$\begin{aligned} H[X + c] &= \frac{E[(X + c)e^{\alpha(X+c)}]}{E[e^{\alpha(X+c)}]} \\ &= \frac{E[Xe^{\alpha(X+c)}]}{E[e^{\alpha(X+c)}]} + c \\ &= H[X] + c. \end{aligned}$$

- *Aditividad:* Si  $X_1$  y  $X_2$  son dos riesgos que se suponen independientes, entonces:

$$\begin{aligned} H[X_1 + X_2] &= \frac{E[(X_1 + X_2)e^{\alpha(X_1+X_2)}]}{E[e^{\alpha(X_1+X_2)}]} \\ &= \frac{E[X_1e^{\alpha X_1}]E[e^{\alpha X_2}] + E[X_2e^{\alpha X_2}]E[e^{\alpha X_1}]}{E[e^{\alpha X_1}]E[e^{\alpha X_2}]} \\ &= \frac{E[X_1e^{\alpha X_1}]}{E[e^{\alpha X_1}]} + \frac{E[X_2e^{\alpha X_2}]}{E[e^{\alpha X_2}]} \\ &= H[X_1] + H[X_2]. \end{aligned}$$

## 2.4. Propiedades deseables de una medida de riesgo

La medición del riesgo ha dado lugar a un importante número de publicaciones recientes (ver por ejemplo Artzner (1999), Dhaene et al. (2003), Landsman y Sherris (2001), Wang (2000)). La idea es que una medida de riesgo debe permitir obtener las correspondientes primas que reflejen adecuadamente la incertidumbre inherente en la distribución de  $X$ .

**Definición 2.4** *Una medida de riesgo es un funcional  $\rho : X \rightarrow [0, \infty)$  que hace corresponder a un riesgo  $X$  un número real no negativo  $\rho[X]$  (que puede ser infinito), que representa la cantidad adicional que se debe añadir a  $X$  para hacerlo aceptable.*

Varios autores han seleccionado algunos principios para formar un conjunto de requerimientos que debería satisfacer una medida de riesgo. La siguiente definición ha sido propuesta por Artzner et al. (1999).

**Definición 2.5 (Medidas coherentes de riesgo)** *Una medida coherente de riesgo  $\rho[X]$  se define como aquella medida que satisface las siguientes cuatro propiedades:*

1. *Subaditividad:*  $\rho[X_1 + X_2] \leq \rho[X_1] + \rho[X_2]$ .

2. *Monotonicidad:* Si  $X_1(w) \leq X_2(w)$ :  $\rho[X_1] \leq \rho[X_2]$ .
3. *Homogeneidad positiva:* Si  $c > 0$ ,  $\rho[cX] = c\rho[X]$ .
4. *Invarianza frente a traslaciones:* Si  $c > 0$ ,  $\rho[X + c] = \rho[X] + c$ .

Es importante tener en cuenta que *coherente* se define con respecto a un determinado conjunto de axiomas, y no a un conjunto de axiomas universalmente aceptado. Modificando el conjunto de axiomas, se obtiene otra medida *coherente* de riesgo.

## 2.5. Medidas de riesgo y primas basadas en la cola de la distribución

En la última década se ha producido un creciente interés en la práctica financiera y actuarial para la obtención de los cuantiles (o percentiles) de probabilidad de la distribución de un riesgo. Puesto que la interpretación de los cuantiles se realiza en términos probabilísticos y es bastante clara, dichos valores han encontrado un lugar destacado en la gestión de riesgos, por medio del llamada valor en riesgo (VaR). Este concepto responde a la siguiente pregunta, ¿cuánto podemos esperar perder en un día, semana, año, . . . con una probabilidad dada?

En el mundo financiero y actuarial el valor en riesgo se ha convertido en una de las principales herramientas cuantitativas.

**Definición 2.6** Dado un riesgo  $X$  con función de distribución  $F_X$  y un nivel de probabilidad  $q \in (0, 1)$ , se define el valor en riesgo, que lo denotaremos por  $VaR[X; q]$  como,

$$VaR[X; q] = F_X^{-1}(q) = \inf\{x : F_X(x) \geq q\}, \quad 0 < q < 1, \quad (2.25)$$

donde  $F_X^{-1}(\cdot)$  denota la función de distribución inversa.

Además del no cumplimiento de la propiedad de subaditividad el valor en riesgo presenta algunos inconvenientes. Un valor aislado del VaR para un nivel de probabilidad  $q$  no proporciona información sobre el tamaño de la cola superior de la distribución. Esto es un considerable inconveniente puesto que un actuario además de la frecuencia, está interesado en la severidad de un suceso. Nos interesa por tanto una medida de riesgo que nos indique cuanto de malo es un suceso malo.

En este sentido se define el valor en riesgo en la cola, como la media de todos los valores en riesgo en la cola de la distribución.

**Definición 2.7** Dado un riesgo  $X$  con función de distribución  $F_X$  y un nivel de probabilidad  $q \in (0, 1)$ , se define el valor en riesgo en la cola, y lo representaremos

por  $TVaR[X; q]$  como,

$$TVaR[X; q] = \frac{1}{1-q} \int_q^1 VaR[X; q] dq, \quad 0 < q < 1. \quad (2.26)$$

En el caso de riesgos con función de distribución  $F_X$  de tipo continuo, una importante propiedad del valor en riesgo en la cola (2.26) es que coincide con la esperanza condicionada en la cola, definida por,

$$TCE[X; q] = E[X|X > VaR[X; q]], \quad (2.27)$$

que mide el valor medio de los percentiles en la cola de la distribución. Furman y Landsman (2007) han argumentado que el TCE (o equivalentemente el TVaR en el caso continuo) puede ser considerado como un principio de prima y puede obtenerse como una prima ponderada, como veremos en la siguiente sección.

**Ejemplo 2.2** Las distribuciones probabilísticas de extremos se han convertido en elemento indispensable en el análisis de datos, tanto en finanzas como en seguros. Consideremos una distribución de Gumbel con función de distribución,

$$F(x; \mu, \sigma) = \exp \left\{ - \exp \left\{ - \frac{x - \mu}{\sigma} \right\} \right\}, \quad -\infty < x < \infty, \quad (2.28)$$

donde  $\sigma > 0$  y  $\mu \in \mathbb{R}$ . Mediante la función inversa de (2.28) obtenemos el valor en riesgo,

$$VaR[X; q] = \mu - \sigma \log[-\log(q)], \quad 0 < q < 1. \quad (2.29)$$

El valor en riesgo en la cola viene dado por la media de (2.29) de acuerdo con la definición (2.26). Se puede demostrar que,

$$TVaR[X, q] = \frac{1}{1-q} [\mu(1-q) - \sigma \{-\gamma + Ei[\log q] - q \log[-\log q]\}], \quad (2.30)$$

donde  $\gamma = 0.57722\dots$ , es la constante de Euler y  $Ei[\cdot]$  es la función exponencial integral, definida por,

$$Ei[z] = - \int_{-z}^{\infty} \frac{e^{-t}}{t} dt. \quad (2.31)$$

Veamos un ejemplo de cálculo de estas dos medidas.

**Ejemplo 2.3** Un analista de riesgos tiene que estimar temperaturas extremas de una determinada zona. La tabla 2.1 muestra 15 observaciones de estas temperaturas en grados centígrados. Los estimadores de máxima verosimilitud de  $\mu$  y  $\sigma$  son:

$$\hat{\mu} = 33,5, \quad \hat{\sigma} = 2,241$$

Los valores en riesgo y en riesgo en la cola para un nivel de probabilidad  $q$  se obtienen por medio de (2.29) y (2.30) sustituyendo  $\mu$  y  $\sigma$  por sus estimadores de máxima verosimilitud. La figura 2.2 representa las dos funciones. De este modo:

$$\text{VaR}[X; 0,99] = 43,8089; \quad \text{TVaR}[X; 0,99] = 46,0556$$

y

$$\text{VaR}[X; 0,999] = 48,9792; \quad \text{TVaR}[X; 0,999] = 51,2207.$$

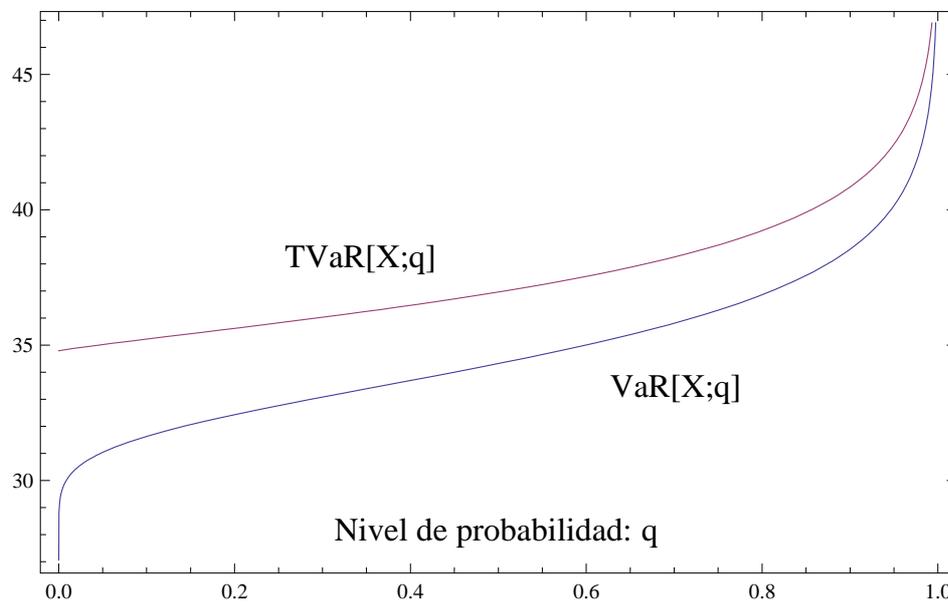


Figura 2.2: Valor en riesgo (VaR) y valor en riesgo en la cola (TVaR) de la distribución de Gumbel con los datos del Ejemplo 2.3.

Tabla 2.1: Datos de temperaturas extremas (Bury, 1999).

32.7	30.4	31.8	33.2	33.8	35.3	34.6	33.0
32.0	35.7	35.5	36.8	40.8	38.7	36.7	

## 2.6. Una clase de medidas de riesgos ponderadas

A partir del Teorema general 2.7 se pueden obtener una gran variedad de principios de primas. Recientemente, Furman y Zitikis (2008) han propuesto una importante subclase de (2.24). Sea  $w : [0, \infty) \rightarrow [0, \infty]$  una función peso tal que  $E[w(X)] < \infty$ . Consideremos la clase de distribuciones ponderadas considerada por Patil y Rao (1978):

$$F_{X_w}(x) = \frac{E[1(X \leq x)w(X)]}{E[w(X)]},$$

donde  $1(S)$  es la función indicador, que vale 1 si  $S$  es cierto y 0 en otro caso. Por extensión de (2.13) podemos considerar la clase de primas ponderadas:

$$H[X_w] = \int_0^\infty \bar{F}_{X_w}(x) dx. \quad (2.32)$$

Un número importante de primas son casos especiales de  $H[X_w]$ . Para probar esto, escribimos  $H[X_w]$  como,

$$H[X_w] = \frac{E[Xw(X)]}{E[w(X)]}. \quad (2.33)$$

Entonces, la prima  $H[X_w]$  incluye como casos particulares (Heilman (1989) y Kamps (1998)):

- Prima neta:  $E[X]$ , cuando  $w(x)$  es constante,
- Prima varianza modificada:  $E[X] + \text{Var}[X]/E[X]$  cuando  $w(x) = x$ ,
- Prima de Esscher:  $E[Xe^{\lambda X}]/E[e^{\lambda X}]$ , eligiendo  $w(x) = e^{\lambda x}$ ,
- Prima de Kamps: (1998)  $E[X(1-e^{-\lambda X})]/E[1-e^{-\lambda X}]$  cuando  $w(x) = 1-e^{-\lambda x}$ ,

así como las primas basadas en la cola de la distribución:

- Valor en riesgo en la cola:  $\text{TVaR}[X; q]$ , eligiendo  $w(x) = 1\{X > \text{VaR}[X; q]\}$ ,
- Prima varianza en la cola modificada:  $\text{TVaR}[X; q] + \frac{\text{Var}[X|X > \text{VaR}[X; q]]}{\text{TVaR}[X; q]}$  cuando  $w(x) = x \cdot 1\{X > \text{VaR}[X; q]\}$ .

Otra clase de prima pueden obtenerse considerando la función peso,

$$w(x) = x^c,$$

para algún valor fijo  $c \in (0, 1]$ . Esta función da lugar a,

$$H[X_c] = \frac{E[X^{c+1}]}{E[X^c]}. \quad (2.34)$$

Puesto que (2.34) es el ratio de dos momentos, puede ser fácilmente calculado para un buen número de distribuciones.

**Ejemplo 2.4** Consideremos un riesgo  $X$  con distribución lognormal, definido por medio de la función de densidad,

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right\}, \quad x > 0.$$

Puesto que  $E[X^r] = \exp(\mu r + r^2\sigma^2/2)$  si  $r > 0$ , la expresión (2.34) se convierte en:

$$H[X_c] = \exp\{\mu + (c + 1/2)\sigma^2\}.$$

## 2.7. Primas basadas en estadísticos de orden

Además de los principios descritos anteriormente, existen otros muchos principios que se han descrito en la literatura (Tsanakas y Desli (2003), Heilpern (2003), Young (2004), Denuit et al. (2006), entre otros).

En esta sección describiremos brevemente principios de prima basados en los estadísticos de orden del riesgo, descrito tanto en el contexto del modelo de riesgo individual o colectivo. Sea  $X_1, \dots, X_N$  una muestra de riesgos donde  $N$  una variable aleatoria de conteo con  $\Pr(N \geq 1) = 1$ . Consideremos la muestra de riesgos ordenada:

$$X_{1:N} \leq X_{2:N} \leq \dots \leq X_{N:N},$$

donde  $X_{1:N}$  representa el menor valor,  $X_{2:N}$  el segundo menor etc. Definimos la variable aleatoria,

$$R_N = \sum_{k=1}^N f_k(X_{k:N})$$

donde  $f_i$  son funciones reales con  $f_i(x) \leq x$  y  $\sum_{i=1}^n f_i(x_i) \geq 0$  si  $x_1 \leq \dots \leq x_n$ . A partir de  $R_N$ , Kremer (1985) define la prima,

$$H[R_N] = E[R_N] + \lambda\sqrt{\text{Var}(R_N)}. \quad (2.35)$$

con  $\lambda \geq 0$ . Si  $f_i = c_i = 1$  son constantes y  $\lambda = 0$  obtenemos la prima neta del modelo de riesgo clásico. La variable  $R_N$  permite además modelizar determinados reaseguros. Si disponemos de una muestra fija de  $n$  reclamaciones podemos definir la prima basada en el principio de varianza del máximo:

$$H[X_{n:n}] = E[X_{n:n}] + \lambda\sqrt{\text{Var}[X_{n:n}]}, \quad (2.36)$$

válida en el contexto del modelo de riesgo individual.

## Capítulo 3

# Modelización del Riesgo en Familias Paramétricas

### 3.1. Introducción

En este capítulo estudiaremos modelos de distribuciones paramétricas que pueden ser utilizadas para describir el riesgo de un colectivo, tanto en seguros como en riesgos operacionales.

Los modelos que aquí introducimos serán también utilizados en los modelos de credibilidad, tanto para modelizar el colectivo de una cartera, como las distribuciones a priori de algún parámetro de interés.

Entre los textos que incluyen teoría y aplicaciones sobre distribuciones probabilísticas podemos destacar: Arnold et al. (1992 y 1999), Balakrishnan y Nevzorov (2003), Kleiber y Kotz (2003), Castillo et al. (2005), Sarabia et al. (2006) y los clásicos Johnson, Kotz y Balakrishnan (1994 y 1995). En este capítulo estudiaremos medidas de riesgo que pueden ser utilizadas en credibilidad para modelizar datos de pérdidas. Más concretamente, obtendremos expresiones para las dos medidas de riesgo en la cola que actualmente se están utilizando en la práctica como son el VaR y el TVaR. El cálculo de estas medidas es muy reciente en la moderna práctica actuarial.

Los contenidos de este capítulo son los siguientes. En la sección 3.2 presentaremos los siguientes modelos paramétricos de uso habitual en la ciencia actuarial: las distribuciones de Pareto, las distribuciones exponencial, gamma y gamma invertida, las distribución normal y lognormal, la distribución de Weibull y la distribución inversa Gaussiana. Obtendremos los VaR y TVaR para los diferentes modelos. El análisis de las colas de una distribución constituye para el analista de riesgos uno de los aspectos claves a la hora de elegir una distribución de pérdidas. Este aspecto será estudiado en la sección 3.2.7. En la sección 3.3 presentamos un caso de

estudio aplicado a datos normalizados sobre daños de huracanes. Se propondrán diversos modelos paramétricos, además de un modelo de regresión basado en la distribución inversa Gaussiana. Finalmente se obtendrán diversas primas basadas en los modelos seleccionados.

## 3.2. Modelos paramétricos

En esta sección estudiaremos diversos modelos paramétricos, y obtendremos algunas medidas de riesgo importantes. Concretamente, estudiaremos los siguientes grupos de distribuciones paramétricas:

- Distribuciones de Pareto (Pareto I y II).
- Distribuciones exponencial, gamma y gamma invertida.
- Distribuciones normal y lognormal.
- Distribución de Weibull.
- Distribución inversa Gaussiana.

### 3.2.1. Distribuciones de Pareto

Las distribuciones de Pareto han sido ampliamente utilizadas en la literatura actuarial y de riesgos por sus buenas propiedades. El texto de Arnold (1983) es una de las referencias clásicas sobre las distribuciones de Pareto en sus diferentes aspectos. Desde el punto de vista del análisis de riesgos, el ajuste de la cola de una distribución es un aspecto clave: interesa que los sucesos poco probables pero con gran severidad tengan una probabilidad no despreciable. La función de densidad de la distribución de Pareto converge a cero mucho más lentamente que la distribución normal y lognormal. Por tanto, resulta mucho más seguro utilizarla para determinar las primas de un reaseguro en los tramos de grandes siniestros (sección 3.2.7).

Como es bien conocido, existen diversos tipos de distribuciones de Pareto. Una de las más utilizadas es la distribución clásica de Pareto, que forma parte de la jerarquía de distribuciones propuesta por Arnold (1983), donde se comienza por la distribución más sencilla y se van enriqueciendo las distribuciones, hasta tener un total de cuatro modelos anidados. En esta investigación nos interesaremos en la distribución de Pareto clásica y la distribución de Pareto tipo II a veces llamada Pareto generalizada.

Una variable aleatoria  $X$  sigue una distribución clásica de Pareto, si la función de densidad es:

$$f(x; \alpha, \sigma) = \frac{\alpha \sigma^\alpha}{x^{\alpha+1}}, \quad x > \sigma > 0, \quad (3.1)$$

con  $\alpha > 0$ , y  $f(x) = 0$  si  $x < \sigma$ . Por otro lado, una variable aleatoria  $X$  sigue una distribución de Pareto tipo II (o generalizada) si la función de densidad viene dada por:

$$f(x; \alpha, \sigma) = \frac{\alpha/\sigma}{(1 + x/\sigma)^{\alpha+1}}, \quad x > 0 \quad (3.2)$$

y  $f(x; \alpha, \sigma) = 0$  si  $x < 0$ , donde  $\alpha > 0$  es un parámetro de forma y  $\sigma$  un parámetro de escala. Para la distribución clásica de Pareto, la tabla 3.1 muestra algunos valores aproximados del parámetros  $\alpha$  dependiendo del tipo de suceso con que trabajemos (Antal (2001)). Estos valores pueden resultar de utilidad a la hora de asignar los parámetros de la distribución a priori de  $\alpha$  en inferencia Bayesiana.

Tabla 3.1: Distribución clásica de Pareto: valores aproximados del parámetro de forma (Antal (2001)).

Tipo de suceso	Valor aproximado de $\alpha$
Terremotos y tormentas	$\approx 1.0$
Incendios	$\approx 2.0$
Incendios en industrias	$\approx 1.5$
Responsabilidad general	$\approx 1.8$
Responsabilidad general automóviles	$\approx 2.5$
Daños en el trabajo	$\approx 2.0$

Las cantidades VaR y TVaR de las distribuciones de Pareto se pueden obtener de forma explícita. Para el caso de la distribución clásica de Pareto con densidad (3.1), el valor en riesgo viene dado por,

$$\text{VaR}[X; q] = \sigma(1 - q)^{-1/\alpha}, \quad 0 < q < 1. \quad (3.3)$$

invirtiendo la función de distribución asociada a (3.1).

Para la existencia del TVaR se necesita una media finita, en este caso  $\alpha > 1$ . Usando (2.26) se obtiene el valor en riesgo en la cola,

$$\text{TVaR}[X; q] = \frac{\alpha\sigma}{\alpha - 1}(1 - q)^{-1/\alpha}, \quad \alpha > 1, \quad 0 < q < 1. \quad (3.4)$$

La figura 3.1 muestra las expresiones (3.3) y (3.4) según un nivel de probabilidad  $q$ . En el caso que  $0 < \alpha \leq 1$  (donde la media no existe,  $E[X] = \infty$ ), el VaR se hace máximo, que suele corresponder a riesgos asociados a terremotos y tormentas. El riesgo del suceso disminuye a medida que aumenta el valor de  $\alpha$  y por tanto van existiendo los sucesivos momentos. El TVaR disminuye igualmente a medida que aumenta el valor de  $\alpha$ .

Para la distribución de Pareto tipo II, invirtiendo la función de distribución, obtenemos el valor en riesgo que viene dado por,

$$\text{VaR}[X; q] = \sigma[(1 - q)^{-1/\alpha} - 1], \quad 0 < q < 1, \quad (3.5)$$

cuya expresión siempre existe. Si promediamos la expresión (3.5) por medio de la fórmula (2.26) obtenemos el valor en riesgo en la cola, que viene dado por,

$$\text{TVaR}[X; q] = \frac{\sigma}{1 - q} \left[ \frac{\alpha}{\alpha - 1} (1 - q)^{1-1/\alpha} - (1 - q) \right], \quad \alpha > 1, \quad 0 < q < 1, \quad (3.6)$$

que requiere la condición  $\alpha > 1$  para su cálculo.

### 3.2.2. Distribuciones exponencial y gamma

La distribución exponencial viene definida por medio de la función de densidad,

$$f(x; \sigma) = \frac{1}{\sigma} \exp(-x/\sigma), \quad x > 0, \quad (3.7)$$

donde  $\sigma > 0$ . Se utiliza para modelizar procesos con tasa de fallo constante y viene caracterizada por el hecho de ser la única distribución continua que verifica la propiedad de pérdida de memoria.

Invirtiendo la función de distribución asociada a (3.7) se obtiene valor en riesgo:

$$\text{VaR}[X; q] = -\sigma \log(1 - q), \quad 0 < q < 1. \quad (3.8)$$

Por otro lado, el valor en riesgo en la cola es,

$$\text{TVaR}[X; q] = \sigma[1 - \log(1 - q)], \quad 0 < q < 1, \quad (3.9)$$

puesto que (fórmula (2.26)):

$$\begin{aligned} \text{TVaR}[X; q] &= \frac{1}{1 - q} \int_q^1 [-\sigma \log(1 - q)] dq \\ &= \sigma[1 - \log(1 - q)]. \end{aligned} \quad (3.10)$$

La distribución gamma es un de los modelos de distribuciones paramétricas más utilizadas en riesgos cuando se dispone de datos positivos, unimodales y con asimetría positiva. Se dice que una variable aleatoria  $X$  sigue una distribución gamma de parámetros  $\alpha$  y  $\sigma$  (representaremos  $X \sim \mathcal{G}(\alpha, \sigma)$ ) si su función de densidad se puede escribir como:

$$f(x; \alpha, \sigma) = \frac{x^{\alpha-1} \exp(-x/\sigma)}{\sigma^\alpha \Gamma(\alpha)}, \quad \text{si } x \geq 0 \quad (3.11)$$

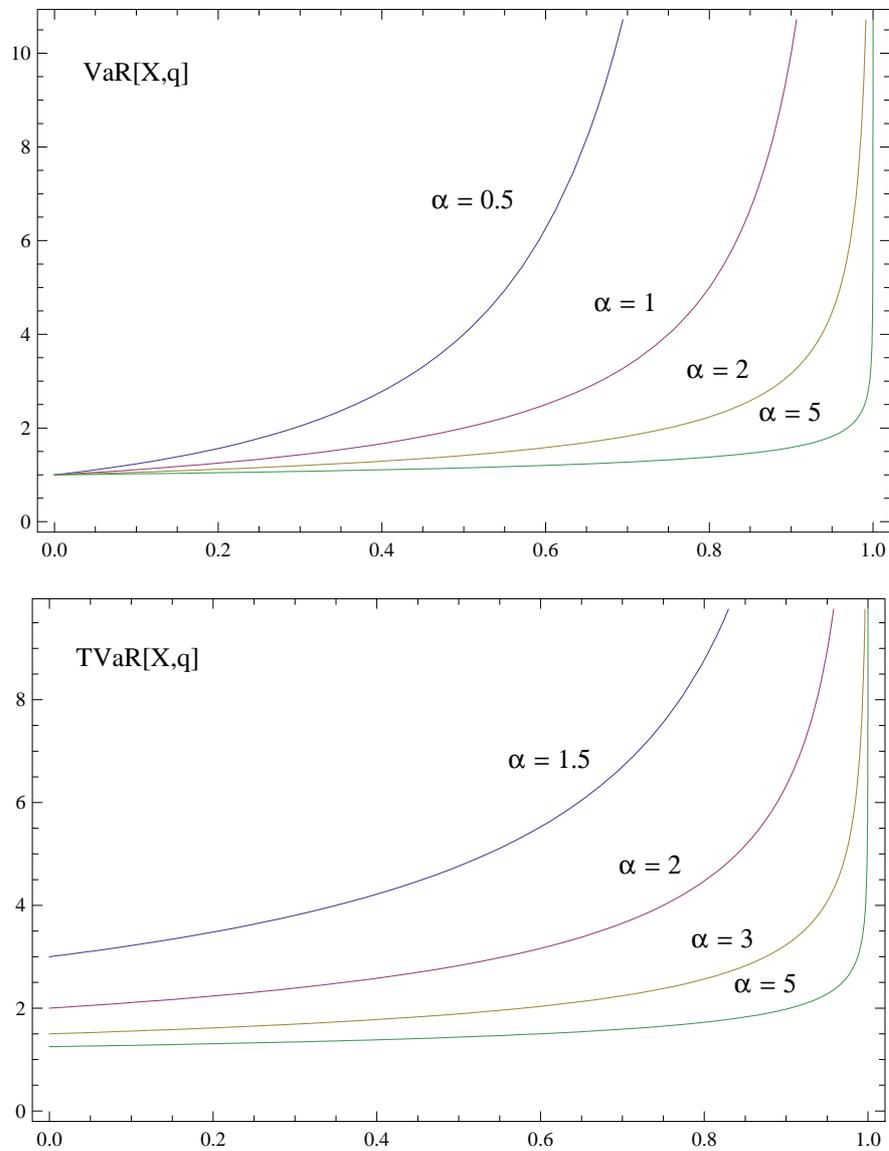


Figura 3.1: Valor en riesgo VaR (gráfico superior) y valor en riesgo en la cola TVaR (gráfico inferior) de la distribución clásica de Pareto con  $\sigma = 1$ , según diversos valores del parámetro de forma  $\alpha$ .

y  $f(x; \alpha, \sigma) = 0$  si  $x < 0$ , donde  $\alpha, \sigma$  son números reales positivos. El parámetro  $\alpha$  es un parámetro de forma y  $\sigma$  un parámetro de escala. La distribución gamma contiene como casos particulares las distribuciones de Erlang ( $\alpha \in \mathbf{N}$ ), exponencial ( $\alpha = 1$ ), chi-cuadrado ( $\sigma = 2$  y  $\alpha = \nu/2$ ), y la normal si  $\alpha \rightarrow \infty$ . La función de distribución de (3.11) se puede escribir como,

$$F(x; \alpha, \sigma) = IG[x/\sigma; \alpha], \quad x > 0, \quad (3.12)$$

donde  $IG[X; a]$  representa la función razón gamma incompleta definida como:

$$IG[x; a] = \frac{1}{\Gamma(a)} \int_0^x t^{a-1} \exp(-t) dt, \quad x > 0, \quad (3.13)$$

con  $a > 0$ . A partir de esta definiciones, el valor en riesgo de la distribución gamma viene dado por,

$$\text{VaR}[X; q] = \sigma \cdot IG^{-1}[q; \alpha], \quad 0 < q < 1, \quad (3.14)$$

donde  $IG^{-1}[z; \alpha]$  es la función inversa de (3.13). Notar que esta función está bien definida puesto que (3.13) es monótona creciente. El valor del TVaR se puede obtener como,

$$\text{TVaR}[X; q] = \alpha \sigma \cdot \frac{1 - IG[\text{VaR}[X; q]/\sigma; \alpha + 1]}{1 - IG[\text{VaR}[X; q]/\sigma; \alpha]}. \quad (3.15)$$

Las figuras 3.2 y 3.3 muestran las expresiones del VaR y TVaR para las distribuciones exponencial y gamma, respectivamente, según un nivel de probabilidad  $q \in (0, 1)$ .

### 3.2.3. Distribución gamma invertida

La distribución gamma invertida es un modelo de distribución probabilística de uso en estadística Bayesiana, como distribución a priori para el parámetro de una exponencial (capítulo 6). Se trata de la transformada inversa de la distribución gamma y tiene como función de densidad:

$$f(x; \alpha, \theta) = \frac{\theta^\alpha x^{-(\alpha+1)} \exp(-\theta/x)}{\Gamma(\alpha)}, \quad x > 0, \quad (3.16)$$

donde  $\alpha, \theta > 0$ . A diferencia de la distribución gamma, los momentos positivos no siempre existen. Se verifica que:

$$E[X] = \frac{\theta}{\alpha - 1}, \quad \text{si } \alpha > 1 \quad (3.17)$$

y

$$\text{Var}[X] = \frac{\theta^2}{(\alpha - 1)^2(\alpha - 2)}, \quad \text{si } \alpha > 2. \quad (3.18)$$

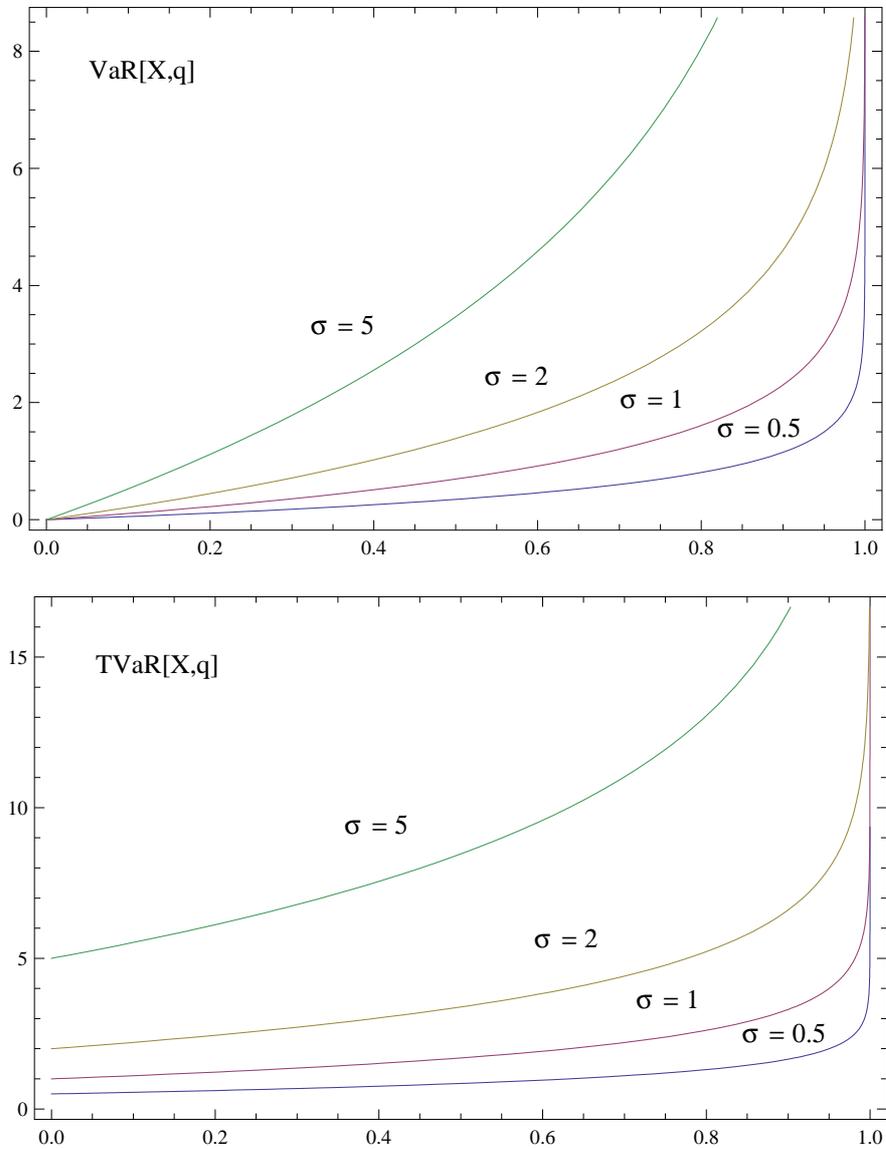


Figura 3.2: Valor en riesgo VaR (gráfico superior) y valor en riesgo en la cola TVaR (gráfico inferior) de la distribución exponencial según diversos valores del parámetro  $\sigma$ .

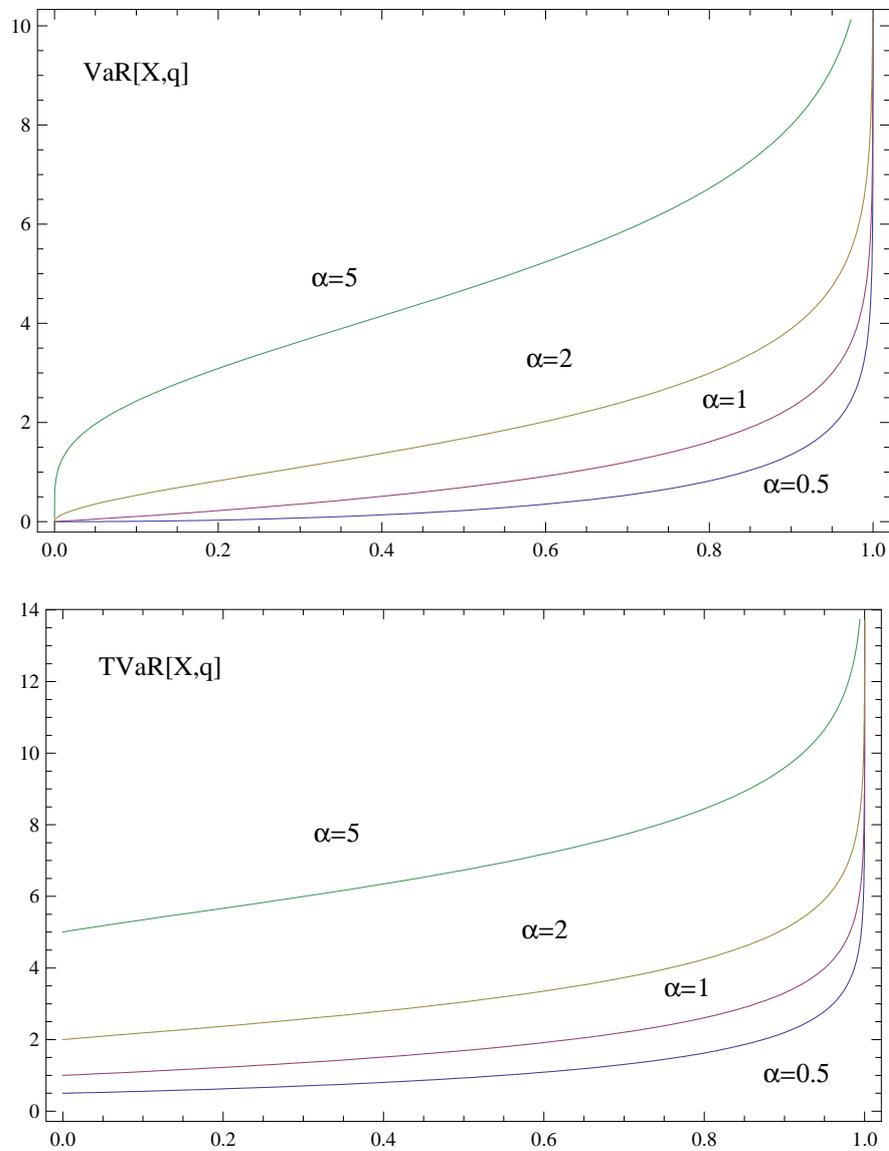


Figura 3.3: Valor en riesgo VaR (gráfico superior) y valor en riesgo en la cola TVaR (gráfico inferior) de la distribución gamma con  $\sigma = 1$  según diversos valores del parámetro de forma  $\alpha$ .

### 3.2.4. Distribuciones normal y lognormal

La distribución normal juega un papel central en la modelización estadística, y se utiliza frecuentemente en credibilidad. Si disponemos de un gran número de datos de pérdidas, la aproximación normal resulta plausible en los diversos modelos clásicos de riesgo, tanto individual como colectivo (sección 6.5). La función de densidad de una normal  $\mathcal{N}(\mu, \sigma^2)$  viene dada por,

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, \quad -\infty < x < \infty, \quad (3.19)$$

donde  $\mu \in \mathbb{R}$  y  $\sigma > 0$ . La función de distribución es,

$$F(x; \mu, \sigma) = \Phi\left(\frac{x-\mu}{\sigma}\right), \quad (3.20)$$

donde  $\Phi(z)$  es la función de distribución de una normal estándar  $\mathcal{N}(0, 1)$ . Usando (3.20) obtenemos el valor en riesgo, que viene dado por:

$$\text{VaR}[X; q] = \mu + \sigma\Phi^{-1}(q), \quad 0 < q < 1, \quad (3.21)$$

donde  $\Phi^{-1}(z)$  es la función de distribución inversa o función de percentiles. Por medio de la expresión (2.26) del capítulo anterior se obtiene el TVaR, que viene dado por:

$$\text{TVaR}[X; q] = \mu + \sigma \frac{\phi[\Phi^{-1}(q)]}{1-q}, \quad 0 < q < 1, \quad (3.22)$$

donde  $\phi(z)$  representa la función de densidad de una variable normal estandarizada.

La distribución lognormal es uno de los modelos paramétricos más utilizados en riesgos. Se dice que la variable aleatoria  $X$  sigue una distribución lognormal si,

$$\log(X) \sim \mathcal{N}(\mu, \sigma^2). \quad (3.23)$$

Se suele representar  $X \sim \mathcal{LN}(\mu, \sigma^2)$ . Las funciones de distribución y de densidad de una distribución lognormal se obtienen mediante cambio de variable, y vienen dadas por,

$$F_{LN}(x; \mu, \sigma) = \Phi\left(\frac{\log x - \mu}{\sigma}\right), \quad x > 0, \quad (3.24)$$

$$f_{LN}(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right\}, \quad x > 0, \quad (3.25)$$

respectivamente.

El valor en riesgo de la distribución lognormal se obtiene a través de (3.24) y viene dado por,

$$\text{VaR}[X; q] = \exp[\mu + \sigma\Phi^{-1}(q)], \quad 0 < q < 1, \quad (3.26)$$

donde  $\Phi^{-1}(z)$  representa nuevamente la función de cuantiles de la distribución normal estándar. El valor en riesgo en la cola viene dado por,

$$\text{TVaR}[X; q] = \exp(\mu + \sigma^2/2) \frac{\Phi(\sigma - \Phi^{-1}(q))}{1 - q}, \quad 0 < q < 1, \quad (3.27)$$

que siempre existe, y donde no es necesaria ninguna condición adicional sobre los parámetros. En las figuras 3.4 y 3.5 aparecen representadas el VaR y TVaR de las distribuciones normal y lognormal, respectivamente, según un nivel de probabilidad  $q$ .

### 3.2.5. Distribución de Weibull

La distribución de Weibull generaliza a la distribución exponencial, de modo que admite tasas de fallo constantes, crecientes y decreciente. Esta distribución también se utiliza en problemas de datos extremos mínimos. Concretamente, se obtiene como límite del mínimo de muestras independientes procedentes de distribuciones limitadas en la cola izquierda. La podemos definir en términos de la función de distribución como,

$$F(x; \alpha, \sigma, \mu) = 1 - \exp\left[-\left(\frac{x - \mu}{\sigma}\right)^\alpha\right], \quad x \geq \mu > 0, \quad (3.28)$$

donde  $\sigma > 0$  es un parámetro de escala,  $\mu$  un parámetro de localización y  $\alpha > 0$  el parámetro de forma. Si  $\alpha = 1$  se obtiene la distribución exponencial y si  $\alpha = 2$  y  $\mu = 0$  se obtiene la distribución de *Rayleigh*.

Invirtiéndolo la función (3.28) se obtiene el valor en riesgo, que viene dado por,

$$\text{VaR}[X; q] = \mu + \sigma[-\log(1 - q)]^{1/\alpha}, \quad 0 < q < 1. \quad (3.29)$$

El valor en riesgo en la cola se obtiene mediante el siguiente resultado.

**Teorema 3.1** *El valor en riesgo en la cola de distribución de Weibull definida en (3.28) viene dado por,*

$$\text{TVaR}[X; q] = \mu + \frac{\sigma\Gamma(1 + 1/\alpha)}{1 - q} \{1 - IG[-\log(1 - q); 1 + 1/\alpha]\}, \quad (3.30)$$

donde  $IG[x; a]$  representa la función razón gamma incompleta definida en (3.13).

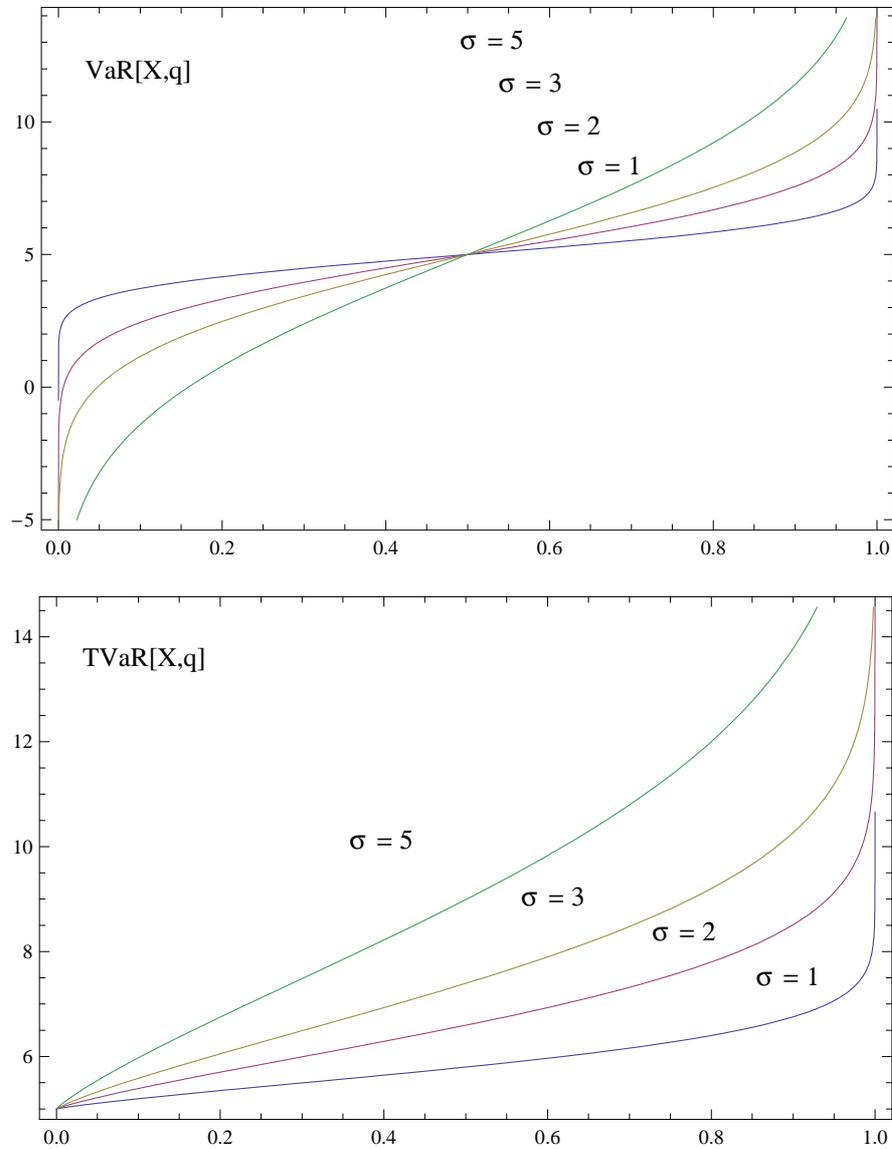


Figura 3.4: Valor en riesgo VaR (gráfico superior) y valor en riesgo en la cola TVaR (gráfico inferior) de la distribución normal con  $\mu = 5$  según diversos valores de la desviación típica  $\sigma$ .

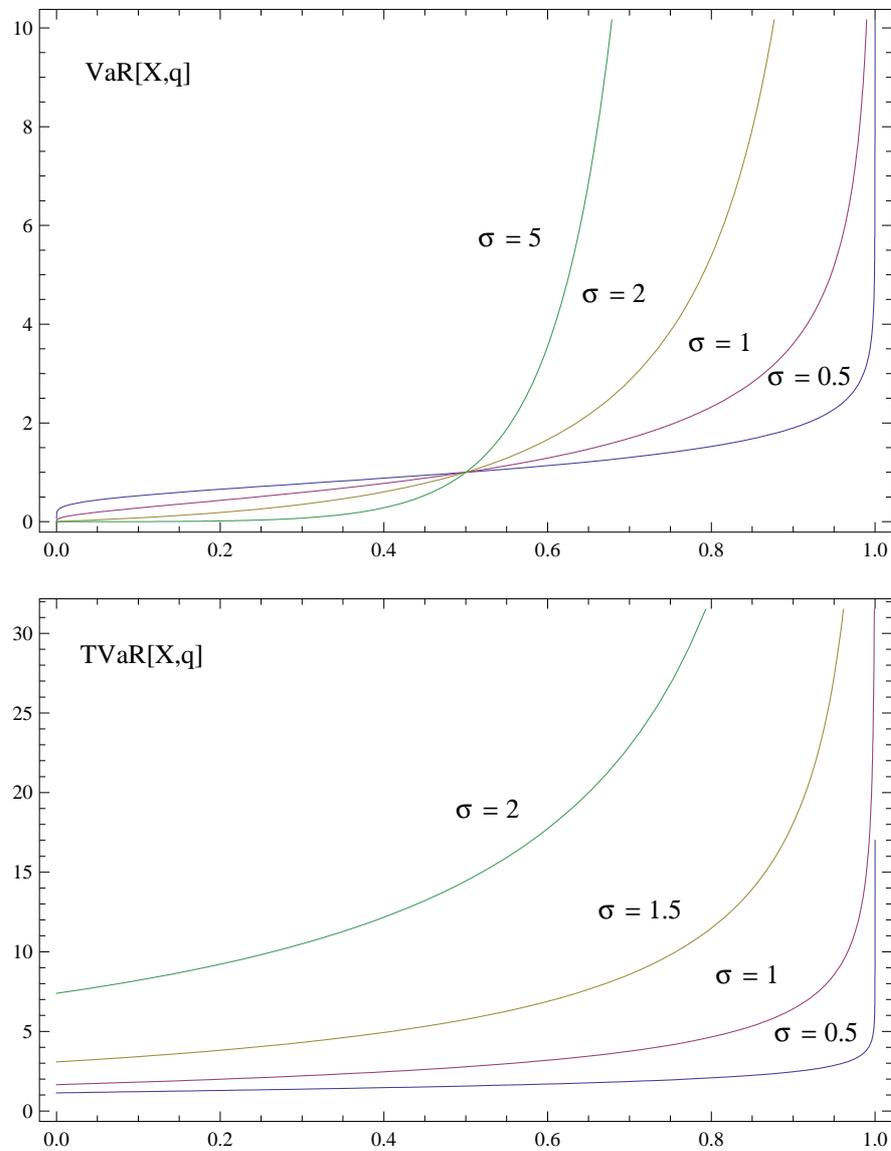


Figura 3.5: Valor en riesgo VaR (gráfico superior) y valor en riesgo en la cola TVaR (gráfico inferior) de la distribución lognormal con  $\mu = 0$  según diversos valores del parámetro  $\sigma$ .

**Demostración:** Por medio de la definición de TVaR y usando (3.29) tenemos:

$$\begin{aligned} \text{TVaR}[X; q] &= \mu + \frac{\sigma}{1-q} \int_q^1 \{[-\log(1-q)]^{1/\alpha}\} dq \\ &= \mu + \frac{\sigma}{1-q} \int_{-\log(1-q)}^{\infty} t^{1/\alpha} \exp(-t) dt, \end{aligned}$$

expresión que conduce a (3.30). ■

La figura 3.6 muestra el VaR y TVaR de una distribución de Weibull según diversos valores del parámetro  $\alpha$ .

### 3.2.6. Distribución inversa Gaussiana

La distribución inversa Gaussiana se ha convertido en una de las distribuciones más utilizadas en estadística actuarial y análisis de riesgos. Esta distribución posee unas propiedades muy similares a las de la distribución normal. La definimos en términos de la función de densidad como:

$$f(x; \lambda, \mu) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}, \quad x > 0, \quad (3.31)$$

donde  $\lambda, \mu > 0$ . El coeficiente de asimétrica es positivo y presenta una moda. La función de distribución viene dada por:

$$F(x; \lambda, \mu) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{2\lambda/\mu} \Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right), \quad x > 0.$$

de donde:

$$\text{VaR}[X; q] = F^{-1}(q; \lambda, \mu).$$

El TVaR se obtiene usando la fórmula 2.26.

### 3.2.7. Comparación de las colas

El análisis de las colas de una distribución constituye para el actuario uno de los aspectos más importantes en la elección de una distribución de pérdidas. Veamos una comparación entre las colas de varias distribuciones.

En el caso de la distribución de Pareto, cuando  $x \rightarrow \infty$ , la función de distribución converge a cero más lentamente que una distribución exponencial (sección 3.2.2), y se dice que la distribución de Pareto presenta una cola más pesada que la distribución exponencial. En la figura 3.7 aparece representada la cola de una distribución de Pareto:

$$F(x) = 1 - \left(\frac{\sigma_0}{x}\right)^\alpha, \quad x > \sigma_0,$$

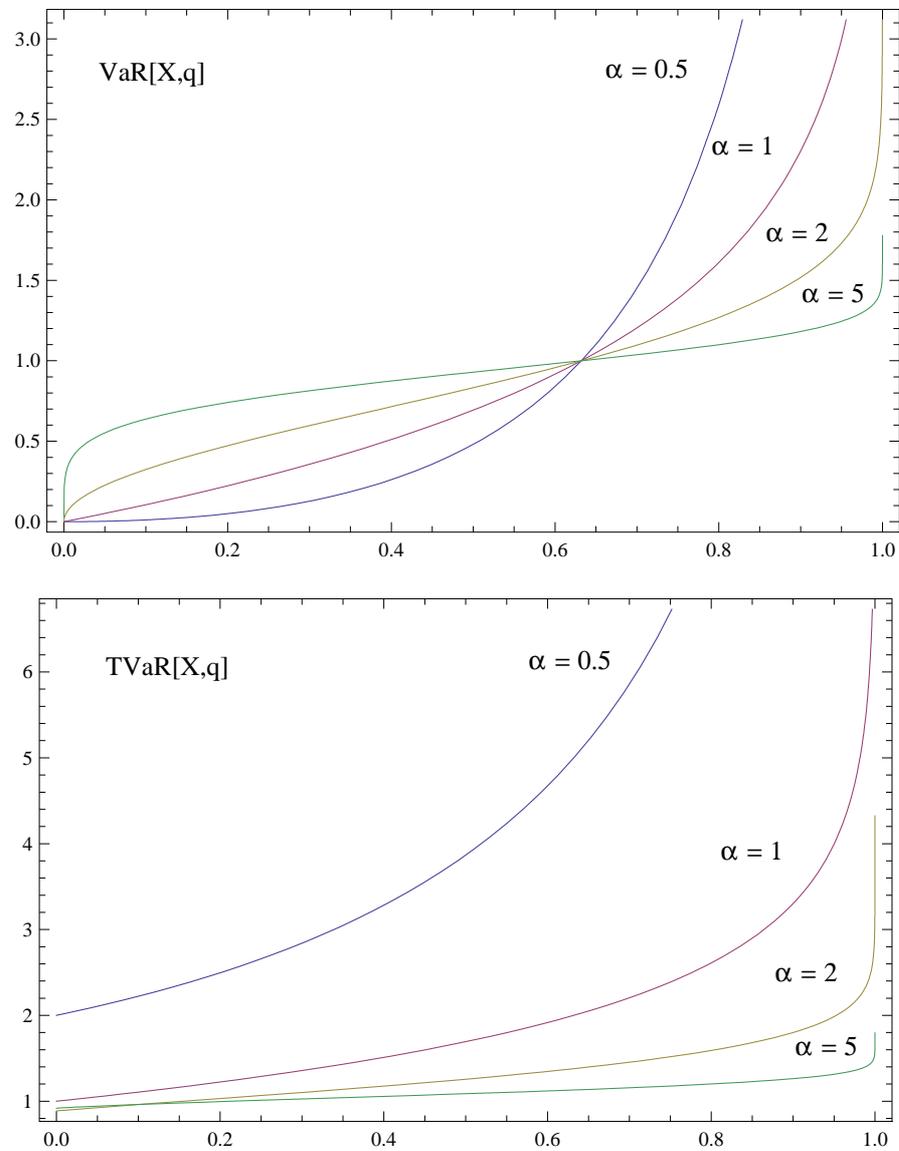


Figura 3.6: Valor en riesgo VaR (gráfico superior) y valor en riesgo en la cola TVaR (gráfico inferior) de la distribución de Weibull con  $\mu = 0$  y  $\sigma = 1$  según diversos valores del parámetro de forma  $\alpha$ .

donde  $\sigma_0 = 0.4$  y  $\alpha = 1.2$ . Para realizar la comparación consideramos distribuciones de tipo exponencial y normal truncadas en  $\sigma_0$ , con funciones de distribución,

$$F_{X|X>\sigma_0}(x) = \frac{F_X(x) - F_X(\sigma_0)}{1 - F_X(\sigma_0)}, \quad x > \sigma_0 > 0$$

y 0 si  $x < \sigma_0$ . A continuación elegimos los valores de los parámetros de modo que las correspondientes funciones de densidad tengan el mismo punto de partida. Por tanto, el parámetro  $\theta$  de la exponencial verifica:  $\frac{1}{\theta} = \frac{\alpha}{\sigma_0}$ , mientras que la desviación típica  $\sigma$  de la distribución normal debe verificar:

$$\frac{\phi(\sigma_0/\sigma)}{\sigma[1 - \Phi(\sigma_0/\sigma)]} = \frac{\alpha}{\sigma_0},$$

que da lugar a  $\sigma = 0,46936$ . Las gráficas de las funciones de supervivencia de las tres distribuciones (Pareto, exponencial y normal) aparecen representadas en la figura 3.7. Notar, de que a pesar de que las tasas de decrecimiento de las tres distribuciones comienzan siendo iguales, la normal tiende a 0 más rápidamente que las distribuciones de Pareto y exponencial. El significado práctico que proporciona este gráfico es claro. Pérdidas grandes son mucho más probables en distribuciones de tipo Pareto que en distribuciones normales. La distribución exponencial se encuentra en medio de las distribuciones normal y de Pareto. Una tipología reducida de colas de distribuciones de pérdidas (de menos pesadas a más pesadas) aparece en la tabla 3.2 (Embrechts et al., 1997). Más detalles acerca de las colas de distribuciones usadas en finanzas y seguros aparece en Heyde y Kou (2004) y Fung y Seneta (2007).

Tabla 3.2: Tipología de colas de distribuciones de pérdidas (de menos peso a más peso). Entre paréntesis algunos ejemplos de distribuciones con ese tipo de cola.

Tipo de cola	Comportamiento de la distribución
<i>Normales</i> (Normal)	$\Pr(X > x) \sim e^{-x^2}$
<i>Exponenciales</i> (Exponencial y Gamma)	$\Pr(X > x) \sim e^{-x}$
<i>Lognormales</i> (Lognormal, Benktander-I)	$\Pr(X > x) \sim e^{-\log^2 x}$
<i>Weibull</i> (Weibull-pesadas, Benktander-II)	$\Pr(X > x) \sim e^{-x^\gamma}, \quad 0 < \gamma < 1$
<i>Potenciales</i> (Pareto, LogGamma, estables*)	$\Pr(X > x) \sim x^{-\alpha}$

\* Distribuciones estables con exponente  $\alpha \in (0, 2)$ . El signo  $\sim$  significa en este contexto equivalencia en el infinito.

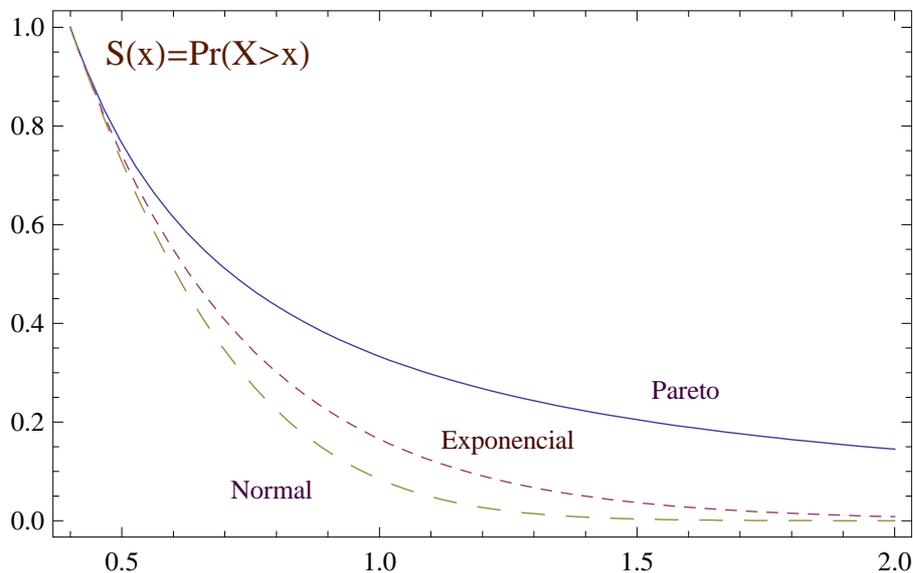


Figura 3.7: Funciones de supervivencia de las distribuciones de Pareto, exponencial y normal.

### 3.3. Caso de estudio: datos normalizados sobre daños de huracanes

Como aplicación de las secciones precedentes, vamos a realizar un ejercicio de tarificación usando diferentes primas basadas en diversos modelos paramétricos, algunos de los cuáles se han descrito previamente. Los datos que utilizaremos son datos reales sobre daños en huracanes.

En una importante contribución, Pielke y Landsea (1998) presentan datos normalizados sobre daños de huracanes en Estados Unidos durante el período de 1925 a 1995. Los datos sobre daños, junto con el nombre del huracán, el año y la categoría aparecen en la tabla 3.3. El concepto de *dato normalizado* permite una correcta comparación entre los datos de daños. Los datos están en dólares constantes del año 1995 y han sido normalizados según el crecimiento de la propiedad personal y los cambios en la población. Detalles sobre el proceso de normalización pueden encontrarse en el citado artículo de Pielke y Landsea (1998).

El histograma de los datos aparece en la figura 3.8. Dicho histograma es similar a muchas de las distribuciones a las que se enfrentan los actuarios, presentando los siguientes hechos estilizados:

- Posee un valor modal,
- Presenta una fuerte asimetría positiva y una curtosis alta,
- Unos pocos valores son mucho mayores que el resto (es decir, presentan una cola larga).

### 3.3.1. Modelos Ajustados

Se han ajustado un total de ocho modelos paramétricos, habituales cuando se trabaja con datos de pérdidas. Los modelos con: gamma, Pareto II, lognormal, inversa Gaussiana, poli-exponencial, distribución de Gumbel de mínimos, distribución de Gumbel de máximos y distribución de Weibull. Las funciones de densidad de estos modelos aparecen en la tabla 3.4.

Los modelos han sido ajustados maximizando la función de verosimilitud de los datos observados:

$$\log L(\underline{\theta}; \mathbf{x}) = \prod_{i=1}^n \log f(x_i; \underline{\theta}),$$

donde  $\mathbf{x} = (x_1, \dots, x_n)$  son los daños observados y  $\underline{\theta} = (\theta_1, \dots, \theta_p)^T$  el vector de parámetros. Como estimadores iniciales de los parámetros se han elegido los estimadores de momentos. Los errores estándar de los parámetros han sido obtenidos mediante la inversa de la matriz de información observada. Si la matriz de información observada es,

$$I(\theta_i, \theta_j) = -\frac{\partial^2 \log L(\underline{\theta}; \mathbf{x})}{\partial \theta_i \partial \theta_j}, \quad i, j = 1, 2, \dots, p,$$

donde los valores de los parámetros se sustituyen por los estimadores de máxima verosimilitud, entonces los errores estándar se obtienen como,

$$\text{ste}(\hat{\theta}_i) = \sqrt{I^{-1}(\theta_i, \theta_i)}, \quad i = 1, 2, \dots, p.$$

Puesto que la estimación de los modelos se ha realizado mediante máxima verosimilitud, como criterios de selección de modelos se ha trabajado con criterios basados en la verosimilitud. En concreto, el valor de la verosimilitud y el del estadístico AIC de Akaike. Mediante el valor de la verosimilitud se elige aquel modelo con mayor verosimilitud. Con objeto de corregir el estadístico anterior se trabaja con el estadístico AIC, que penaliza según el número de parámetros del modelo.

Las estimaciones correspondientes a los ocho modelos ajustados aparecen en la tabla 3.5. Se han incluido los estimadores de máxima verosimilitud y las estimaciones de los errores estándar de los parámetros. La tabla 3.6 recoge el ranking de los modelos según el criterio de máxima verosimilitud (Rango 1) y según el

estadístico AIC (Rango 2). Según los dos criterios los mejores modelos corresponden a (en este orden): distribución inversa Gaussiana, lognormal y gamma. Los puestos intermedio varían ligeramente (distribuciones poli-exponencial, Weibull y Pareto II), mientras que los últimos puestos corresponden a las distribuciones de extremos de Gumbel (mínimos y máximos). Los gráficos de las tres funciones de densidad correspondientes junto con el gráfico conjunto del histograma de los datos y de los modelos ajustados aparecen en las figuras 3.9, 3.10 y 3.11. Los estimadores de los parámetros de estos tres modelos son todos significativos, si realizamos el correspondiente contraste de hipótesis. Para el cálculo de las primas usaremos estos tres modelos.

### 3.3.2. Regresión inversa-Gaussiana

La idea de esta sección es construir un nuevo modelo con variables explicativas o covariables en la distribución ajustada del daño. Como es claro, existen otras variables que explican el daño y que pueden mejorar el modelo ajustado. En este caso disponemos de dos variables explicativas, como son el año en que se produjeron los años y la categoría del huracán. Se pueden considerar otras variables siempre que los datos estén disponibles. En esta caso usaremos la distribución inversa Gaussiana definida en la sección 3.2.6, puesto que es el modelo que mejor se ajusta a los datos, según los criterios antes comentados.

Para especificar el modelo de regresión con la distribución inversa Gaussiana, vamos a considerar un vector  $\mathbf{z}$  de  $m$  covariables conocidas. Existen dos procedimientos de introducir las covariables:

- A través de la media:  $\mu(\mathbf{z}) = \exp(\theta^\top \mathbf{z})$ .
- A través del parámetro de escala:  $\lambda(\mathbf{z}) = \exp(\theta^\top \mathbf{z})$ ,

donde  $\theta = (\theta_1, \dots, \theta_m)^\top$  es el vector de coeficientes de la regresión. Se han ensayado las dos opciones, y se han obtenido mejores resultados con la regresión a través de la media. Por tanto, tenemos el modelo,

$$X|\mathbf{z} \sim \mathcal{IG}(\exp(\theta^\top \mathbf{z}), \lambda), \quad (3.32)$$

donde la función de verosimilitud viene dada por,

$$\log L(\lambda, \underline{\theta}; \mathbf{x}|\mathbf{z}) = \prod_{i=1}^n \log f(x_i; \exp(\theta^\top \mathbf{z}_i), \lambda), \quad (3.33)$$

donde  $f(x; \mu, \lambda)$  es la función de densidad de la inversa Gaussiana definida en (3.31) y  $\mathbf{z} = (1, z_1, z_2)^\top$  son los regresores, donde  $z_1$  corresponde al año y  $z_2$  a la categoría del huracán.

El modelo finalmente ajustado aparece en la tabla 3.7. De las dos covariables consideradas, únicamente el año del huracán ha resultado significativa, por lo que

la otra covariable no ha sido incluida. Nuevamente los valores de los parámetros son significativos. Mediante los dos criterios de selección de modelos, el modelo de regresión inversa Gaussiana mejora a todos los modelos anteriormente considerados. Este modelo puede usarse por tanto con carácter predictivo.

### 3.3.3. Cálculo de primas VaR y TVaR

Como apartado final de esta sección se han obtenido las primas VaR y TVaR para los tres modelos finalmente seleccionados: inversa Gaussian, lognormal y gamma, para niveles de la probabilidad de  $q = 0.9, 0.95, 0.99$  y  $0.999$ . Los resultados obtenidos se encuentran en la tabla 3.8. Las primas obtenidas a partir de los modelos inversa gaussiana y lognormal son similares, si bien en la parte superior de la cola la distribución lognormal da lugar a primas de mayor magnitud. Las primas obtenidas por medio de la distribución gamma son de menor magnitud.

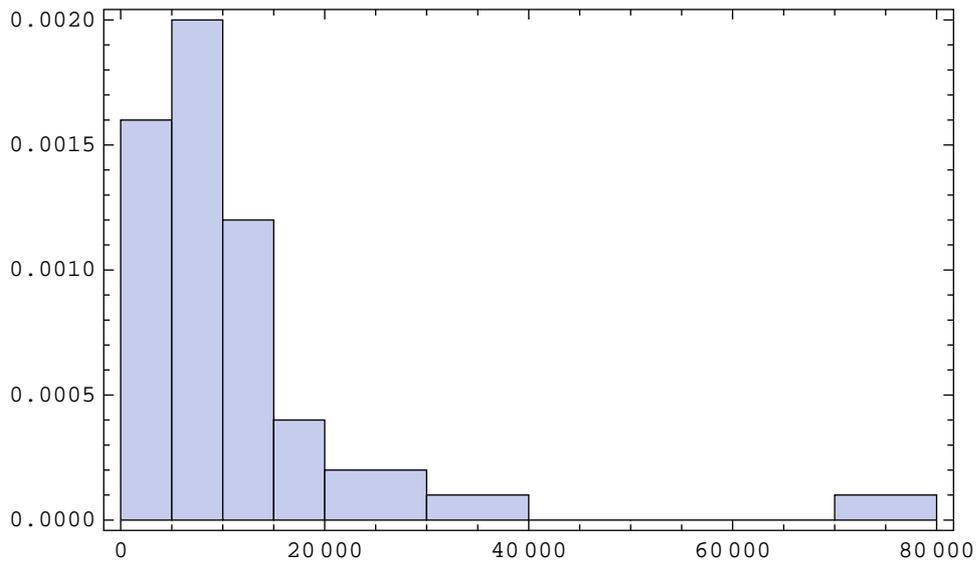


Figura 3.8: Histograma de los datos normalizados de daños debidos a huracanes.

Tabla 3.3: Datos normalizados de daños debidos a huracanes en USA entre 1925-95.  
Fuente: Pielke y Landsea (1998).

Rango	Huracán	Año	Categoría	Daños en billones de dólares US
1	SE FL-AL	1926	4	72.303
2	Andrew (SE FL-LA)	1992	4	33.094
3	N TX (Galvenston)	1900	4	26.619
4	N TX (Galvenston)	1915	4	22.602
5	SW FL	1944	3	16.864
6	New England	1938	3	16.629
7	SE FL-Lake Okeechobee	1928	4	13.795
8	Betsy (SE FL-LA)	1965	3	12.434
9	Donna (FL-E United States)	1960	4	12.048
10	Camille (MS-LA-VA)	1969	5	10.965
11	Agnes (NW FL, NE United States)	1972	1	10.705
12	Diane (NE United States)	1955	1	10.232
13	Hugo (SC)	1989	4	9.380
14	Carol (NE United States)	1954	3	9.066
15	SE FL-LA-AL	1947	4	8.308
16	Carla (NE United States)	1961	4	7.069
17	Hazel (SC-NC)	1954	4	7.039
18	NE United States	1944	3	6.536
19	SE FL	1945	3	6.313
20	Frederic (AL-MS)	1979	3	6.293
21	SE FL	1949	3	5.838
22	S Texas	1919	4	5.368
23	Alicia (N TX)	1983	3	4.056
24	Celia (S TX)	1970	3	3.338
25	Dora (NE FL)	1964	2	3.108
26	Opal (NW FL-AL)	1995	3	3.000
27	Cleo (SE FL)	1964	2	2.435
28	Juan (LA)	1985	1	2.399
29	Audrey (LA-N TX)	1957	4	2.396
30	King (SE FL)	1950	3	2.266

Tabla 3.4: Modelos de distribuciones de daño.

Distribución	Función de densidad
$\mathcal{IG}(\mu, \lambda)$	$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda(x-\mu)^2}{2x\mu^2}\right\}$
$\mathcal{LN}(\mu, \sigma)$	$f(x; \mu, \sigma) = \frac{1}{x\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right\}$
$\mathcal{G}(\alpha, \sigma)$	$f(x) = \frac{x^{\alpha-1} \exp(-x/\sigma)}{\sigma^\alpha \Gamma(\alpha)}$
$\mathcal{PE}(w, \lambda_1, \lambda_2)$	$f(x; w, \lambda_1, \lambda_2) = \frac{1-w}{\lambda_1} \exp(-x/\lambda_1) + \frac{w}{\lambda_2} \exp(-x/\lambda_2)$
$\mathcal{W}(\gamma, \sigma)$	$f(x; \gamma, \sigma) = \frac{\gamma}{\sigma} \left(\frac{x}{\sigma}\right)^{\gamma-1} \exp(-(x/\sigma)^\gamma)$
$\mathcal{PII}(\alpha, \sigma)$	$f(x; \alpha, \sigma) = \frac{\alpha/\sigma}{(1+x/\sigma)^{\alpha+1}}$
$\mathcal{Gmax}(\lambda, \sigma)$	$f(x; \lambda, \sigma) = \frac{1}{\sigma} \exp\left\{-\frac{x-\lambda}{\sigma} - \exp\left(-\frac{x-\lambda}{\sigma}\right)\right\}$
$\mathcal{Gmin}(\lambda, \sigma)$	$f(x; \lambda, \sigma) = \frac{1}{\sigma} \exp\left\{\frac{x-\lambda}{\sigma} - \exp\left(\frac{x-\lambda}{\sigma}\right)\right\}$

Tabla 3.5: Datos de daños en USA 1925-95: modelos ajustados mediante máxima verosimilitud. Entre paréntesis se incluyen los errores estándar.

Rango	Distribución	Parámetros ajustados
1	$\mathcal{IG}(\mu, \lambda)$	$\hat{\mu} = 11749.9$ (2146.15); $\hat{\lambda} = 11739.4$ ; (3031.11)
2	$\mathcal{LN}(\mu, \sigma)$	$\hat{\mu} = 8.98467$ (0.152243); $\hat{\sigma} = 0.8339$ ; (0.107652)
3	$\mathcal{G}(\alpha, \sigma)$	$\hat{\alpha} = 1.4361$ (0.33645); $\hat{\sigma} = 8181.84$ ; (2286.5)
4	$\mathcal{PE}(w, \lambda_1, \lambda_2)$	$\hat{w} = 0.9061$ (0.1793); $\hat{\lambda}_1 = 32202.4$ ; (36848.0) $\hat{\lambda}_2 = 9631.4$ ; (2478.6)
5	$\mathcal{W}(\gamma, \sigma)$	$\hat{\gamma} = 1.10923$ (0.140589); $\hat{\sigma} = 12302.6$ ; (2152.78)
6	$\mathcal{PII}(\alpha, \sigma)$	$\hat{\alpha} = 12.2171$ (23.4233); $\hat{\sigma} = 131538.0$ ; (271581.0)
7	$\mathcal{Gmax}(\lambda, \sigma)$	$\hat{\lambda} = 7236.26$ (1192.02); $\hat{\sigma} = 6333.51$ ; (1007.42)
8	$\mathcal{Gmin}(\lambda, \sigma)$	$\hat{\lambda} = 20040.1$ (4316.07); $\hat{\sigma} = 22127.4$ ; (2429.21)

Tabla 3.6: Datos de daños en USA 1925-95. Valor del logaritmo la verosimilitud y del estadístico AIC (cambiada de signo). Rango 1: basado en el criterio de máxima verosimilitud y Rango 2: basado en el estadístico AIC.

Rango 1	Rango 2	Distribución	$-\log L$	$-AIC$
1	1	Inversa Gaussiana	306.265	308.265
2	2	Lognormal	306.658	308.658
3	3	Gamma	310.650	312.650
4	6	Poli-Exponencial	310.727	313.727
5	4	Weibull	310.836	312.836
6	5	Pareto II	310.982	312.982
7	7	Gumbel máximos	313.988	315.988
8	8	Gumbel mínimos	341.377	343.377

Tabla 3.7: Datos de daños en USA 1925-95; modelo ajustado de regresión inversa Gaussiana mediante máxima verosimilitud. Se incluyen los errores estándar y estadísticos de máxima verosimilitud y AIC.

Parámetros / Estadísticos de selección	Valores ajustados	Errores estándar
$\theta_0$	9.86363	0.372305
$\theta_1$	-0.0101	0.00549
$\lambda$	13103.1	3383.20
$-\log L$	304.669	
$-AIC$	307.669	

Tabla 3.8: Datos de daños en USA 1925-95: Cálculo de primas valor en riesgo  $\text{VaR}[X; q]$  y valor en riesgo en la cola  $\text{TVaR}[X; q]$  para los tres modelos seleccionados.

Primas	Inversa Gaussiana	Lognormal	Gamma
$\text{VaR}[X; 0,90]$	25.184	23.233	24.746
$\text{VaR}[X; 0,95]$	34.343	31.453	31.058
$\text{VaR}[X; 0,99]$	58.588	55.522	45.352
$\text{VaR}[X; 0,999]$	98.225	104.979	65.332
$\text{TVaR}[X; 0,90]$	39.365	36.965	33.739
$\text{TVaR}[X; 0,95]$	49.566	47.153	39.927
$\text{TVaR}[X; 0,99]$	75.625	76.582	54.045
$\text{TVaR}[X; 0,999]$	116.901	135.840	73.892

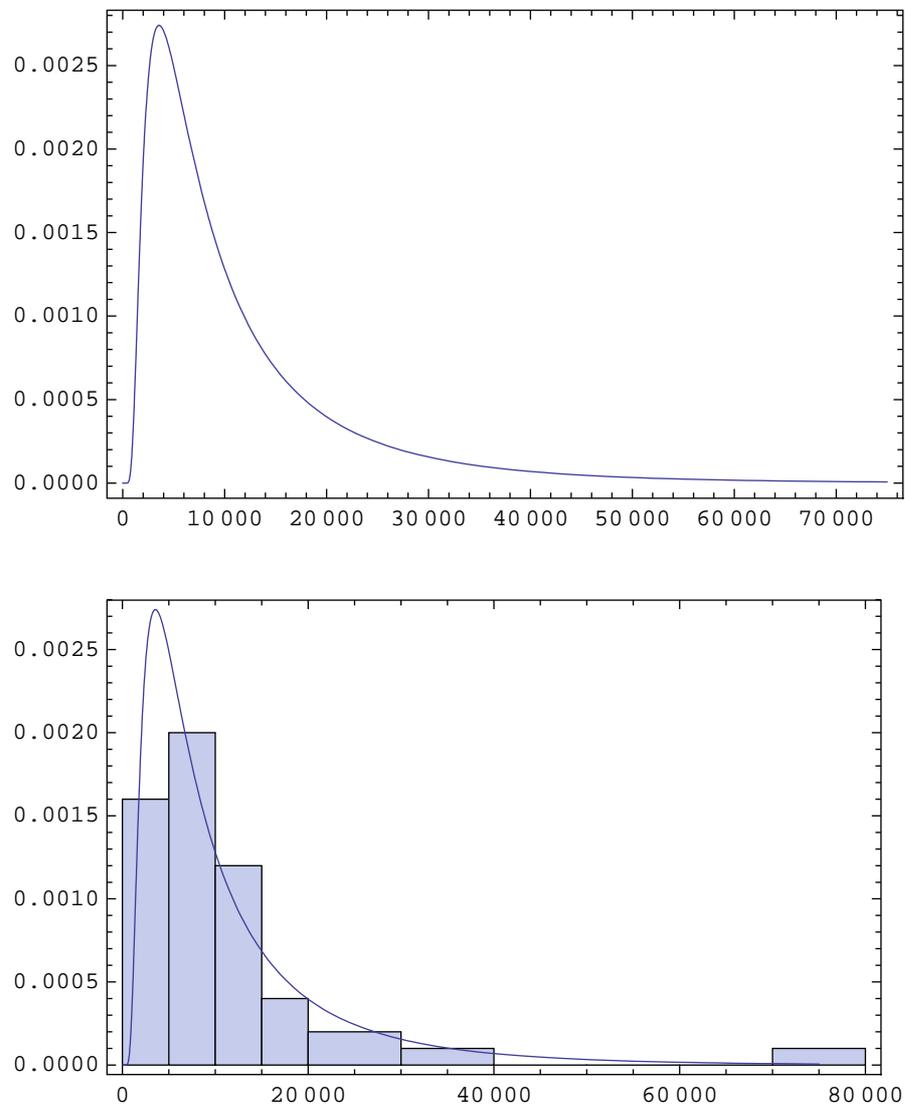


Figura 3.9: Función de densidad ajustada de la distribución inversa Gaussiana a los datos normalizados de daños USA 1925-95.

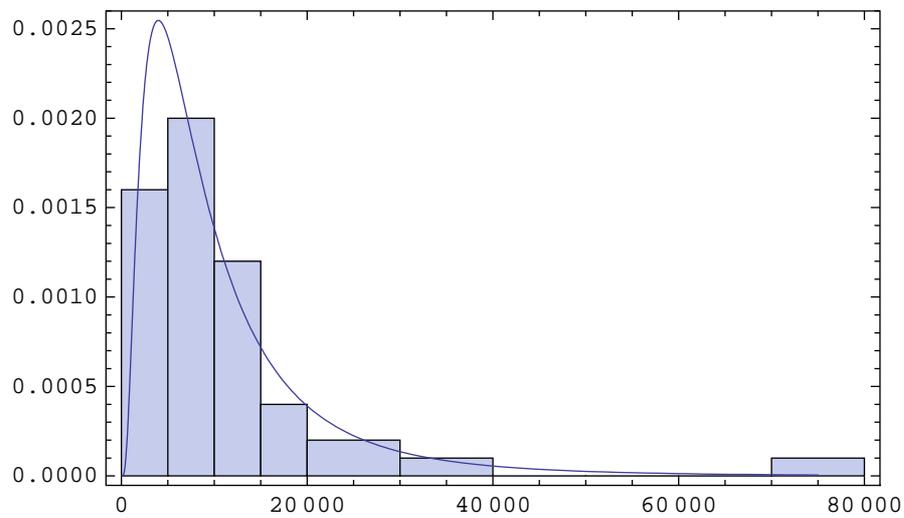
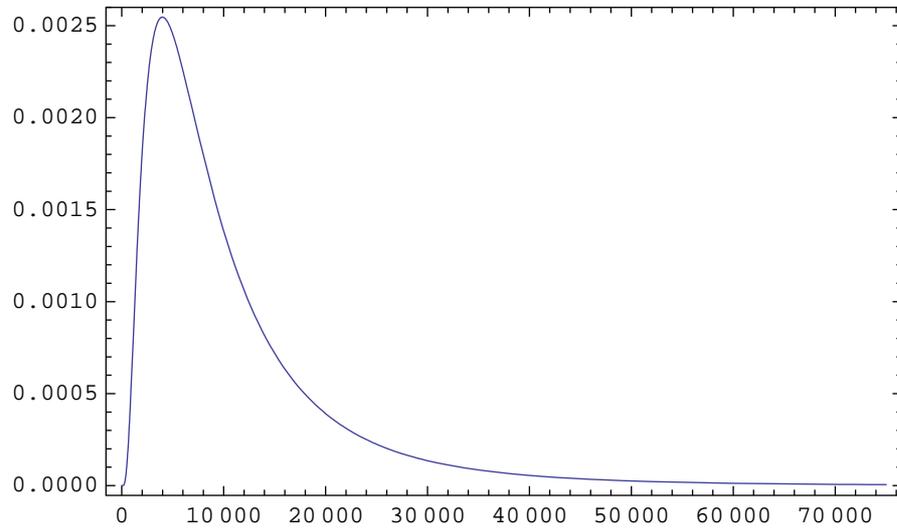


Figura 3.10: Función de densidad ajustada de la distribución lognormal a los datos normalizados de daños USA 1925-95.

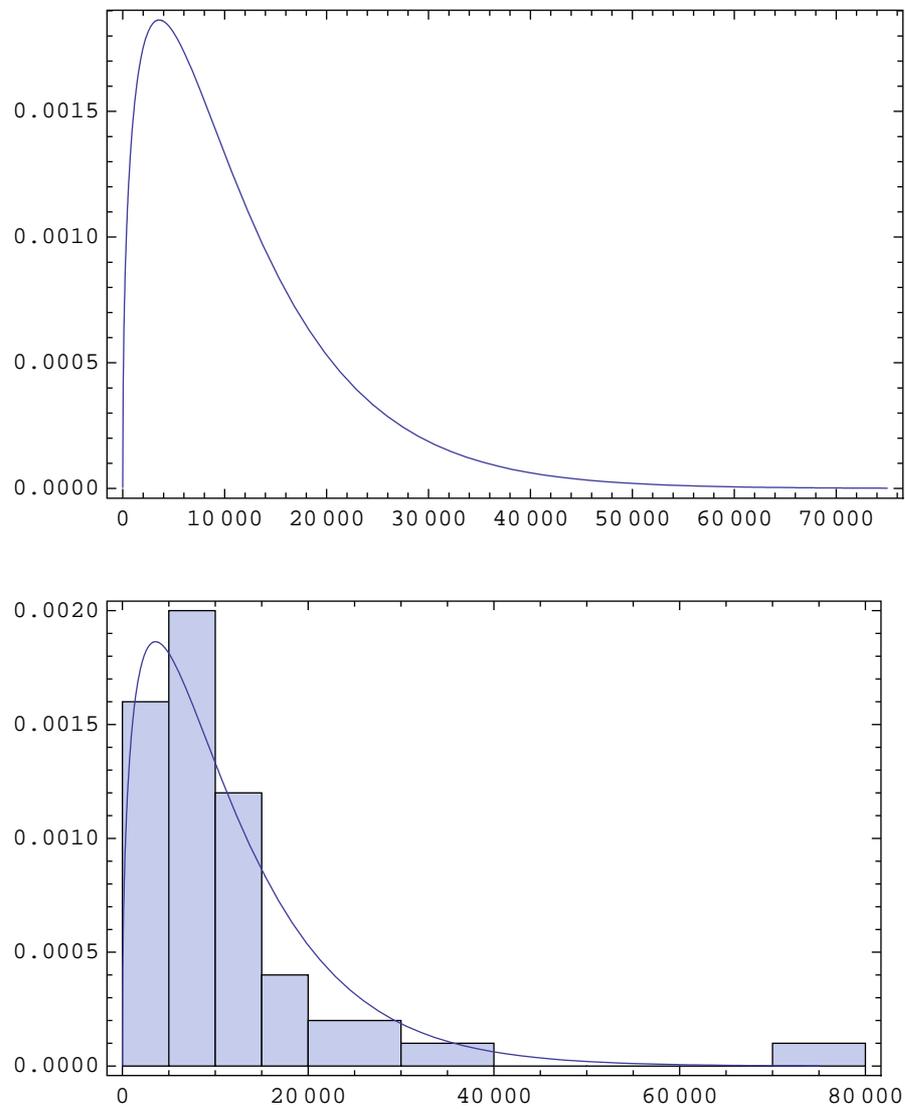


Figura 3.11: Función de densidad ajustada de la distribución gamma a los datos normalizados de daños USA 1925-95.

## Capítulo 4

# Instrumentos de Estadística Bayesiana y Primas Bayes

### 4.1. Introducción

La estadística Bayesiana constituye una de las herramientas básicas en el estudio de la teoría de la credibilidad. La idea de conjugar la información del individuo con la información del colectivo para obtener una fórmula de credibilidad, puede realizarse de una manera natural por medio de la metodología Bayesiana. En estadística Bayesiana la información del individuo (información de siniestralidad) se combina con la información del colectivo (información a priori) por medio del teorema de Bayes.

Por otro lado, muchos de los métodos clásicos de la estadística dejan de funcionar cuando no se dispone de información muestral o cuando el tamaño muestral no es suficientemente grande como para aplicar resultados asintóticos que llevan al uso de la distribución normal.

La metodología Bayesiana se está aplicando con éxito en la práctica actuarial moderna y en la teoría de la credibilidad. El libro de Klugman (1992) hace uso de la estadística Bayesiana para la resolución de diversas cuestiones actuariales. Los libros de Berger (1985), Bernardo y Smith (1994) y Press (2003) contienen desarrollos sistemáticos de la técnica Bayesiana en la actualidad. En los últimos años se ha producido un importante avance en las técnicas Bayesianas de computación. En este sentido, el programa de computación Bayesiana WINBUGS se ha convertido en una herramienta de utilidad en el cálculo Bayesiano. Scollnik (2001) ha utilizado técnicas Bayesianas de computación intensiva con el programa WINBUGS para la resolución de diversos problemas actuariales. Una introducción a este programa junto con diversas aplicaciones actuariales puede encontrarse en Sarabia et al. (2006).

Los contenidos de este capítulo son los siguientes. En las secciones 4.2 y 4.3 introduciremos las herramientas necesarias para la implementación de las técnicas de inferencia Bayesianas. En la sección 4.4 estudiaremos dos aplicaciones en el ámbito actuarial. El análisis Bayesiano de la distribución de Pareto, de uso imprescindible como distribución de pérdidas será visto en la sección 4.5 junto con una aplicación. En la sección 4.6 presentamos la estimación Bayesiana de una importante clase de distribuciones de pérdida introducida por Haberman y Renshaw (1996), cuya estimación se puede escribir en términos de una fórmula de credibilidad. En la sección 4.7 se utiliza la metodología Bayesiana para estimar la prima de riesgo a través, primero de la prima colectiva y, posteriormente mediante la prima Bayes. Finalmente, las secciones 4.9 y 4.10 se dedican al estudio de las familias de distribuciones exponencial y exponencial de dispersión, respectivamente.

## 4.2. Definiciones básicas y teorema de Bayes

Partimos de un riesgo  $X$  dependiente de un parámetro  $\theta$ . En estadística clásica, el parámetro es una constante. Sin embargo, en estadística bayesiana  $\theta$  no es constante, sino que varía según una distribución de probabilidad. La distribución a priori o inicial es la distribución de referencia que se elige para el parámetro.

**Definición 4.1** *La distribución a priori es una distribución de probabilidad definida sobre el espacio paramétrico que recoge las creencias a priori sobre el modelo. Se representa por  $\pi(\theta)$ .*

Esta función de densidad a priori puede ser tanto de naturaleza discreta como continua. Puede depender a su vez de parámetros que reciben el nombre de hiperparámetros. En estadística bayesiana se trabaja a menudo con las distribuciones a priori impropias, es decir, que la integral no es uno. Por ejemplo, si queremos estimar la media de una normal  $X|\mu \sim \mathcal{N}(\mu, \sigma^2)$  podemos elegir,

$$\pi(\mu) = 1 \quad \text{si} \quad -\infty < \mu < \infty.$$

que expresa que todos los valores de  $\mu$  se presentan con la misma probabilidad. Las distribuciones impropias suelen dar lugar a distribuciones a posteriori genuinas, además de aparecer como caso límite de distribuciones a priori genuinas. Este tipo de distribuciones a priori tienen la ventaja adicional de no requerir preasignación de hiperparámetros.

**Definición 4.2** *El modelo de distribución de la población, es la distribución de probabilidad que se elige para los datos, y depende del valor del parámetro  $\theta$ . Suele venir descrita en términos de la función de densidad y se representa por*

$$f_{X|\Theta}(x|\theta).$$

Si  $X = (X_1, \dots, X_n)$  representa un vector de observaciones, la función de verosimilitud basada en  $X$  viene dada por,

$$f_{X|\Theta}(x|\theta) = f_{X|\Theta}(x_1|\theta) \cdots f_{X|\Theta}(x_n|\theta) = \prod_{i=1}^n f_{X|\Theta}(x_i|\theta).$$

Diferentes tipos de muestras (truncadas, censuradas etc.) darán lugar a diferentes tipos de verosimilitudes.

**Definición 4.3** La función de densidad conjunta de  $(X, \Theta)$  viene dada por,

$$f_{X,\Theta}(x, \theta) = f_{X|\Theta}(x|\theta)\pi(\theta). \quad (4.1)$$

La siguiente definición establece un concepto importante.

**Definición 4.4** La función de densidad marginal de  $X$  viene definida por medio de la expresión,

$$f_X(x) = \int f_{X|\Theta}(x|\theta)\pi(\theta)d\theta. \quad (4.2)$$

La siguiente definición introduce la distribución a posteriori.

**Definición 4.5** La distribución a posteriori es la distribución condicionada de  $\Theta$  después de haber observado los datos. Se denota por  $\pi_{\Theta|X}(\theta|x)$ .

Las definiciones anteriores son los puntos claves de la estadística Bayesiana. El teorema de Bayes, permite obtener la distribución a posteriori por medio de la función de verosimilitud y la densidad a priori.

**Teorema 4.1** Si  $\theta$  es una variable aleatoria continua, la función de densidad a posteriori se puede obtener como,

$$\pi_{\Theta|X}(\theta|x) = \frac{f_{X|\Theta}(x|\theta)\pi(\theta)}{\int f_{X|\Theta}(x|\theta)\pi(\theta)d\theta}. \quad (4.3)$$

Otro concepto importante en estadística Bayesiana lo constituye la distribución predictiva.

**Definición 4.6** La distribución predictiva es la distribución de una observación futura  $X_{n+1}$  dada la historia pasada  $X$ .

La función de densidad de la distribución predictiva viene dada por,

$$f_{X_{n+1}|X}(x_{n+1}|x) = \int f_{X_{n+1}|\Theta}(x_{n+1}|\theta)\pi_{\Theta|X}(\theta|x)d\theta. \quad (4.4)$$

El teorema de Bayes describe, fundamentalmente, el proceso de aprendizaje a partir de la experiencia, y prueba cómo el conocimiento sobre el estado de la naturaleza representado por  $\theta$  es modificado a medida que se recibe un nuevo dato. Este aspecto es clave en teoría de la credibilidad. Con una sola observación tenemos que:

$$\pi(\theta|x_1) \propto \pi(\theta)f(x_1|\theta)$$

Si observamos un nuevo dato:

$$\begin{aligned} \pi(\theta|x_1, x_2) &\propto \pi(\theta)f(x_1, x_2|\theta) \\ &\propto \pi(\theta)f(x_1|\theta)f(x_2|\theta) \\ &\propto \pi(\theta|x_1)f(x_2|\theta), \end{aligned}$$

y en general:

$$\pi(\theta|x_1, \dots, x_m) \propto \pi(\theta|x_1, \dots, x_{m-1})f(x_m|\theta).$$

Por otro lado, la distribución a posteriori se puede considerar como un punto medio entre los datos y la información a priori. Se puede esperar que los momentos a posteriori ponderen las opiniones y los datos. Se verifica que:

$$\begin{aligned} E[\Theta] &= E[E(\Theta|X)] \\ \text{Var}[\Theta] &= E[\text{Var}(\Theta|X)] + \text{Var}[E(\Theta|X)] \end{aligned}$$

y por tanto:

$$\text{Var}[\Theta] > E[\text{Var}(\Theta|X)],$$

lo que significa que la distribución a posteriori proporciona menor variabilidad (en media) que la distribución a priori.

### 4.3. Inferencia y predicción Bayesianas

La inferencia y predicción Bayesianas requieren del uso del concepto de función de pérdida, que también fue utilizado en el capítulo 2.

**Definición 4.7** *Una función de pérdida  $L(\theta, a)$  es una función que indica la pérdida incurrida si se elige a  $\theta$  como el verdadero estado de la naturaleza (valor del parámetro), ante una determinada acción  $a$  (un estimador).*

Valores grandes de  $L(\theta, a)$  indican que  $a$  es más incorrecto, mientras que valores pequeños indican que  $a$  es más correcto. En vez de una función de pérdida podemos hablar de una función de utilidad. En una función de utilidad hablamos de ganancias y no de pérdidas, por lo que la utilidad es simplemente  $-L(\theta, a)$ . (Berger, 1985). Algunas propiedades de las funciones de pérdida son las siguientes:

1.

$$L(\theta, a) \geq 0,$$

2.

$$L(\theta, \theta) = 0,$$

3.

$$L(\theta, a) \leq L(\theta, a') \text{ si } |\theta - a| \leq |\theta - a'|.$$

Algunos ejemplos de funciones de pérdida son:

- Función de pérdida cuadrática:

$$L(\theta, a) = (\theta - a)^2.$$

- Error absoluto:

$$L(\theta, a) = |\theta - a|.$$

- La función,

$$L(\theta, a) = \begin{cases} k_2(\theta - a), & \theta > a \\ k_1(a - \theta), & \theta \leq a, \end{cases}$$

asigna un peso de  $k_2$  a los valores de la izquierda del óptimo y  $k_1$  a los valores a la derecha.

- La función,

$$L(\theta, a) = \begin{cases} (a - \theta)^2, & a < \theta \\ 10(a - \theta)^2, & a \geq \theta. \end{cases}$$

penaliza la sobreestimación, más que la subestimación.

- La función,

$$L(\theta, a) = \frac{(a - \theta)^2}{|\theta| + 1},$$

penaliza errores en la estimación más cuando  $\theta$  está cerca de cero, que cuando  $|\theta|$  es grande.

El estimador de Bayes permite realizar inferencias sobre el parámetro  $\theta$  dada una función de pérdida  $L$ . Las inferencias que se realizan en este tipo de estadística se basan en la distribución a posteriori del parámetro, eligiendo una adecuada función de pérdida.

**Definición 4.8 (Estimador de Bayes)** *El estimador de Bayes (también llamado regla de Bayes), es la cantidad que minimiza la expresión en  $a \in \mathcal{A}$ ,*

$$E[L(\theta, a)] = \int L(\theta, a)\pi_{\Theta|X}(\theta|x)d\theta. \quad (4.5)$$

El estimador de Bayes se suele representar por  $\delta^\pi(x)$  o simplemente por  $\hat{\theta}_{Bayes}$ . Cada función de pérdida junto con la correspondiente distribución a priori, da lugar un tipo diferente de estimador.

Veamos un ejemplo de un estimador Bayes con una función de pérdida cuadrática.

**Teorema 4.2** *En el caso de una función de pérdida cuadrática, el estimador de Bayes viene dado por la media de la distribución a posteriori:*

$$\delta^\pi(x) = E[\theta|X] = \frac{\int \theta f(x|\theta)\pi(\theta)d\theta}{\int f(x|\theta)\pi(\theta)d\theta}$$

**Demostración:**

$$E[L(\theta, a)] = \int_{\Theta} (\theta - a)^2 \pi(\theta|x) d\theta$$

Derivando respecto de  $a$  (aplicando el teorema de derivación bajo el signo integral):

$$\frac{\partial}{\partial a} E[L(\theta, a)] = \int_{\Theta} -2(\theta - a)\pi(\theta|x) d\theta = 0$$

de donde:

$$a = \int_{\Theta} \theta \pi(\theta|x) d\theta = E[\theta|X].$$

Ahora:

$$\frac{\partial^2}{\partial a^2} E[L(\theta, a)] = 2 \int_{\Theta} \pi(\theta|x) d\theta > 0.$$

■

El siguiente resultado permite la estimación con otra función de pérdida.

**Teorema 4.3** *La función de pérdida porcentaje de error relativo,*

$$L(\theta, a) = \left( \frac{a - \theta}{\theta} \right)^2$$

*da lugar al estimador de Bayes:*

$$\delta^\pi(x) = \frac{E[\theta^{-1}|X]}{E[\theta^{-2}|X]}.$$

El cálculo de intervalos de confianza se reduce al cálculo de percentiles de la distribución a posteriori.

**Definición 4.9** *Se dice que el intervalo  $(a, b)$  es un intervalo de credibilidad para el parámetro  $\theta$  a nivel  $1 - \alpha$  si,*

$$\Pr(a \leq \Theta \leq b|X) = 1 - \alpha, \quad (4.6)$$

*donde  $0 \leq \alpha \leq 1$ .*

Las distribuciones conjugadas son un tipo de distribución a priori que facilitan los cálculos de estimación.

**Definición 4.10** Una clase  $\mathcal{C}$  de distribuciones a priori para un parámetro  $\theta$  se dice que es conjugada para una determinada verosimilitud, si tanto la distribución a priori como la distribución posteriori pertenecen a la misma clase, es decir si

$$\pi(\theta) \in \mathcal{C}$$

entonces,

$$\pi_{\Theta|X}(\theta|x) \in \mathcal{C}.$$

El siguiente resultado es especialmente importante en credibilidad, dado que permite obtener la predicción de una nueva observación a partir de los datos disponibles, es decir, fórmulas Bayesianas de credibilidad.

**Teorema 4.4** Sean  $X_1, X_2, \dots, X_n$  variables aleatorias independientes e igualmente distribuidas de modo que  $E[X_i|\theta] = \theta$ ,  $i = 1, 2, \dots, n$ . Entonces:

$$E[X_{n+1}|X_1 = x_1, \dots, X_n = x_n] = E[\Theta|X_1 = x_1, \dots, X_n = x_n]. \quad (4.7)$$

**Demostración:** Supongamos que se trata de variables aleatorias de tipo continuos. Calculando directamente la esperanza tenemos que,

$$\begin{aligned} E[X_{n+1}|X_1 = x_1, \dots, X_n = x_n] &= \int_{\mathcal{X}} x f(x_{n+1}|x_1, x_2, \dots, x_n) dx \\ &= \int_{\mathcal{X}} x \left( \int_{\Theta} f(x_{n+1}|\theta) \pi(\theta|x_1, x_2, \dots, x_n) d\theta \right) dx \\ &= \int_{\Theta} \left( \int_{\mathcal{X}} x f(x_{n+1}|\theta) dx \right) \pi(\theta|x_1, \dots, x_n) d\theta \\ &= E[\Theta|X_1 = x_1, \dots, X_n = x_n]. \end{aligned}$$

■

Existen diferentes técnicas de estimación Bayesianas, dependiendo de los objetivos de la estimación y de la información disponible. Podemos considerar cuatro tipos de estimación Bayesiana (Lehmann y Casella, 1998):

- *Estimación Bayesiana clásica.* La distribución a priori se obtiene mediante información extramuestral (experiencia).
- *Estimación Bayes empírica.* En esta situación los parámetros de la distribución a priori son estimados a partir de los datos.
- *Estimación Bayesiana jerárquica.* Los parámetros de la distribución a priori son modelizados por medio de otra distribución a priori, a veces llamada hiper-prior.

- *Estimación Bayesiana robusta.* La precisión de un estimador se evalúa para cada miembro de la distribución a priori, con el fin de encontrar un estimador que mejore a los de toda la clase.

Estos cuatro métodos Bayesianos de estimación serán utilizados a lo largo de la presente investigación.

## 4.4. Aplicaciones Bayesianas en seguros

En esta sección veremos dos aplicaciones importantes de la estadística Bayesiana en seguros: estimación de una cuota de mercado y estimación del fraude.

### 4.4.1. Estimación de una cuota de mercado

Un analista está interesado en estimar la cuota de mercado  $\Theta$  de un nuevo producto. A partir de estudios previos de naturaleza similar y haciendo uso de otras fuentes de información sobre el sector, el analista es capaz de construir una distribución de probabilidad a priori  $\pi(\theta)$  que viene dada en la tabla 4.1. Con objeto de completar el estudio, el analista realiza una encuesta sobre el nivel de aceptación del producto. Realiza un total de 30 encuestas y obtiene que 5 individuos que están dispuestos a adquirir el nuevo producto. ¿Cómo reestimar la distribución de las cuotas de mercado?

Tabla 4.1: Cuota de mercado y probabilidades a priori.

$\Theta$ : cuota de mercado	Probabilidades a priori $\pi(\theta)$
0.05	0.16
0.10	0.24
0.15	0.18
0.20	0.15
0.25	0.12
0.30	0.10
0.35	0.05

A partir de los datos de la encuesta, tenemos que la función de verosimilitud viene dada por,

$$f_{x|\Theta}(x|\theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x},$$

donde  $n = 30$  encuestas y  $x = 5$ . Si consideramos la versión discreta de (4.3)

tenemos que,

$$\pi_{X|\Theta}(x|\theta) = \frac{f_{x|\Theta}(x|\theta)\pi(\theta)}{\sum_{\theta} f_{x|\Theta}(x|\theta)\pi(\theta)}.$$

Los valores de la función de verosimilitud, junto con las probabilidades a priori y posteriori están incluidas en la tabla 4.2. Asimismo, la figura 4.1 representa mediante un diagrama de barras las probabilidades a priori. La cuota de mercado más probable es 0.15

Tabla 4.2: Probabilidades a priori, a posteriori y verosimilitudes de las cuotas de mercado.

$\Theta$	$\pi(\theta)$	$f_{X \Theta}(x \theta)$	$\pi_{\Theta X}(\theta x)$
0.05	0.16	0.191	0.205
0.10	0.24	0.119	0.191
0.15	0.18	0.189	0.227
0.20	0.15	0.186	0.187
0.25	0.12	0.145	0.117
0.30	0.10	0.102	0.069
0.35	0.05	0.012	0.004

#### 4.4.2. Estimación del fraude

Un actuario desea estimar la proporción  $\theta$  de fraude en un determinado tipo de pólizas. Para ello se realizan inspecciones aleatorias en las reclamaciones realizadas. Se conoce que el fraude en el sector es del 4 por ciento con una desviación típica del 2 por ciento.

Se inspeccionan un total de  $n = 240$  pólizas, y se concluye que 22 de ellas son fraudulentas. El actuario está interesado en combinar ambas fuentes de información (la del sector y la de su compañía) para obtener un estimador final del fraude.

Si  $X$  representa el número pólizas fraudulentas sobre un total de  $n$  inspecciones, es claro que la distribución de  $X$  dada la verdadera proporción de fraude  $\theta$ , sigue una distribución binomial,

$$f_{X|\Theta}(x|\theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, \quad x = 0, 1, \dots, n. \quad (4.8)$$

Como distribución a priori para  $\theta$  se elige una distribución beta de primera especie clásica con función de densidad,

$$\pi_{\Theta}(\theta; a, b) = \frac{\theta^{a-1}(1 - \theta)^{b-1}}{B(a, b)}, \quad 0 < \theta < 1, \quad (4.9)$$

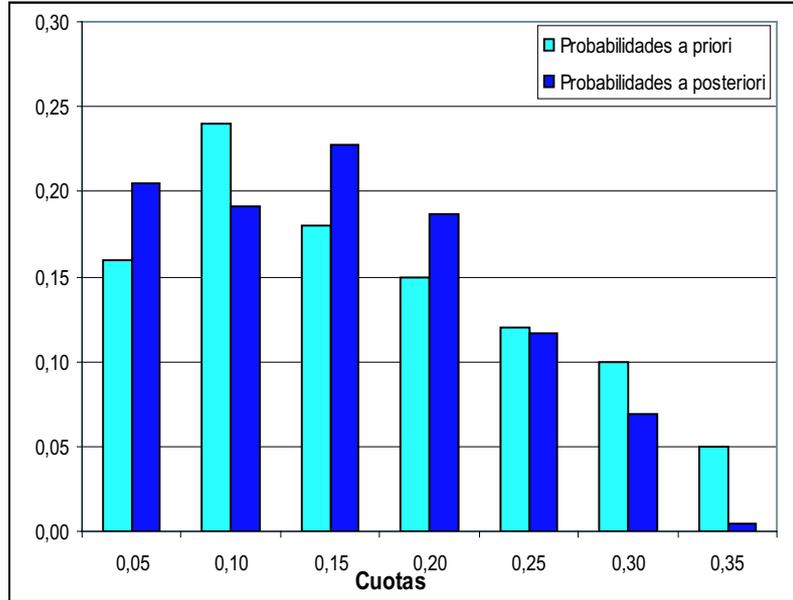


Figura 4.1: Probabilidades a priori y a posteriori de la cuota de mercado.

donde  $a, b > 0$  son hiperparámetros que se estiman a partir de los datos. La información sobre el sector corresponde a información a priori sobre  $\Theta$ , por lo que podemos conocer los valores de los parámetros. Si  $E[\Theta] = m$  y  $\text{Var}[\Theta] = s^2$ , se trata de resolver en  $a$  y  $b$  el sistema anterior. La solución viene dada por,

$$a = \frac{m^2 - m(m^2 + s^2)}{s^2}, \quad (4.10)$$

$$b = \frac{m - (m^2 + s^2)}{s^2} - a. \quad (4.11)$$

A continuación se trata de obtener la distribución a posteriori de los datos dada la distribución a priori beta definida en (4.9). Utilizando el resultado (4.3) tenemos que,

$$\pi_{\Theta|X}(\theta|x) = \frac{f_{X|\Theta}(x|\theta)\pi(\theta; a, b)}{\int_0^1 f_{X|\Theta}(x|\theta)\pi(\theta; a, b)d\theta}$$

$$\begin{aligned}
&= \frac{\binom{n}{x} \theta^x (1-\theta)^{n-x} \frac{\theta^{a-1} (1-\theta)^{b-1}}{B(a,b)}}{\int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} \frac{\theta^{a-1} (1-\theta)^{b-1}}{B(a,b)} d\theta} \\
&= \frac{\theta^{x+a-1} (1-\theta)^{n-x+b-1}}{B(x+a, n-x+b)}.
\end{aligned}$$

La última expresión coincide nuevamente con una distribución tipo beta, por tanto,

$$\Theta|x \sim \mathcal{Be}(x+a, n-x+b). \quad (4.12)$$

El estimador de Bayes bajo función de pérdida cuadrática:

$$\begin{aligned}
E[\Theta|x] &= \frac{x+a}{a+n+b} \\
&= \frac{n}{n+a+b} \cdot \frac{x}{n} + \frac{a+b}{n+a+b} \cdot \frac{a}{a+b} \\
&= \frac{n}{n+a+b} \cdot \bar{x} + \frac{a+b}{n+a+b} \cdot E[\Theta].
\end{aligned} \quad (4.13)$$

Notar que (4.13) es una fórmula de credibilidad, donde se combinan la experiencia del sector (colectivo) con la información de la compañía (el individuo), y donde el factor de credibilidad viene dado por:

$$Z(n) = \frac{n}{n+a+b}.$$

Los intervalos de credibilidad del parámetro  $\theta$  se obtienen a partir de la distribución a posteriori (4.12). Se trata de obtener dos límites  $(l, u)$  tales que,

$$\Pr(l \leq \theta | X \leq u) = 1 - q,$$

donde  $0 < q < 1$  es el error. La tabla 4.3 recoge las estimaciones con la información disponible. Se puede estimar un fraude del 7.7 por ciento con una desviación típica del 1.45 por ciento. La figura 4.2 representa las distribuciones a priori y a posteriori del problema.

Tabla 4.3: Estimadores Bayesianos y parámetros con los datos de fraude.

Parámetros y estimadores	Estimaciones numéricas
$a$ inicial	3.8000
$b$ inicial	91.200
$a$ final	25.800
$b$ final	309.800
$E[\Theta]$	0.04
$DT[\Theta]$	0.02
$\hat{\theta}(\text{EMV})$	0.0917
$E[\Theta \underline{X}]$	0.0770
$DT[\Theta \underline{X}]$	0.0145
Factor de credibilidad	0.4174
Intervalo de credibilidad 90 %	0.0546–0.1023
Intervalo de credibilidad 95 %	0.0510–0.1078
Intervalo de credibilidad 99 %	0.0444–0.1199

## 4.5. Análisis Bayesiano de la distribución de pérdidas Pareto

En este apartado realizaremos un análisis Bayesiano de la distribución clásica de Pareto, según la definición vista en la sección 3.2.1. Suponemos entonces que los datos de pérdidas se encuentran por encima de un límite inferior  $L > 0$ , y pueden ser modelizadas según una distribución clásica de Pareto, que representaremos por  $X \sim \mathcal{P}(\alpha, L)$ . La función de densidad de las pérdidas es,

$$f(x|\alpha) = \frac{\alpha}{L} \left(\frac{x}{L}\right)^{-(\alpha+1)}, \quad x \geq L > 0, \quad (4.14)$$

donde suponemos que el parámetro  $L$  es una constante conocida predeterminada por el actuario. Supongamos ahora que se dispone de una muestra de tamaño  $n$ ,  $X = (X_1, \dots, X_n)$  de la severidad del riesgo. La función de verosimilitud de (4.14) viene dada por,

$$\begin{aligned} L(X|\alpha) &= \prod_{i=1}^n \frac{\alpha}{L} \left(\frac{X_i}{L}\right)^{-(\alpha+1)} \\ &\propto \alpha^n \exp\left(-\alpha \sum_{i=1}^n \log(X_i/L)\right) \end{aligned} \quad (4.15)$$

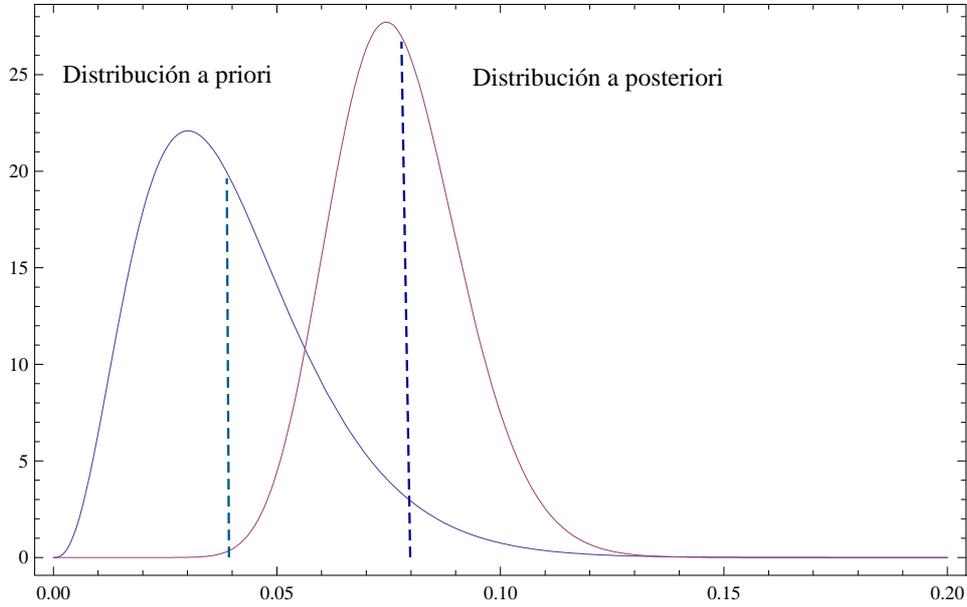


Figura 4.2: Distribuciones a priori y a posteriori de los datos del fraude. Las líneas verticales discontinuas representan la media a priori (izquierda) y la media a posteriori (derecha).

A continuación elegimos para el parámetro  $\alpha$  una distribución gamma con función de densidad,

$$\pi(\alpha; \alpha_0, \sigma) = \frac{\alpha^{\alpha_0-1} \exp(-\alpha/\sigma)}{\sigma^{\alpha_0} \Gamma(\alpha_0)}, \quad \alpha > 0 \quad (4.16)$$

para de este modo disponer de una distribución conjugada. La distribución a priori la representaremos por  $\alpha \sim \mathcal{G}(\alpha_0, \sigma)$ .

En esta etapa un aspecto importante es la elección de los hiperparámetros de la distribución a priori. En este caso, suponemos que el actuario es capaz de proporcionar dos percentiles de la distribución a priori. Si  $(p_i, \alpha_{p_i})$ ,  $p_i \in (0, 1)$ ,  $i = 1, 2$  representan los dos percentiles, los valores de  $\alpha_0$  y  $\sigma$  se obtienen como solución del sistema,

$$\Pr(\mathcal{G}(\alpha_0, \sigma) \leq \alpha_{p_i}) = p_i, \quad i = 1, 2.$$

La distribución a posteriori viene dada por,

$$\pi(\alpha|X) \propto L(X|\alpha) \cdot \pi(\alpha; \alpha_0, \sigma)$$

$$\begin{aligned} &\propto \alpha^n \exp\left(-\alpha \sum_{i=1}^n \log(X_i/L)\right) \cdot \alpha^{\alpha_0-1} \exp(-\alpha/\sigma) \\ &= \alpha^{n+\alpha_0-1} \exp\left[-\alpha \left(\sum_{i=1}^n \log(X_i/L) + 1/\sigma\right)\right], \end{aligned}$$

por tanto, tenemos que,

$$\alpha|X \sim \mathcal{G}\left(\alpha_0 + n, \frac{1}{A(n, L) + 1/\sigma}\right), \quad (4.17)$$

donde  $A(n, L)$  viene dado por,

$$A(n, L) = \sum_{i=1}^n \log(X_i/L). \quad (4.18)$$

A continuación usando (4.17), y utilizando una función de pérdida cuadrática, de acuerdo con el teorema 4.2, el estimador Bayes del parámetro  $\alpha$  viene dado por,

$$\begin{aligned} E[\alpha|X] &= \frac{\alpha_0 + n}{A(n, L) + 1/\sigma} \\ &= \frac{A(n, L)}{A(n, L) + 1/\sigma} \cdot \frac{n}{A(n, L)} + \frac{1/\sigma}{A(n, L) + 1/\sigma} \cdot \alpha_0 \cdot \sigma \\ &= Z \cdot \hat{\alpha}_{EMV} + (1 - Z) \cdot E[\alpha] \end{aligned} \quad (4.19)$$

donde  $Z$  es el factor de credibilidad dado por,

$$Z = \frac{A(n, L)}{A(n, L) + 1/\sigma},$$

y  $\alpha_{EMV} = n/A(n, L)$  el estimador de máxima verosimilitud de  $\alpha$ . En este caso (4.19) no es estrictamente una fórmula de credibilidad, puesto que  $Z$  depende del valor de las severidades  $X_i$ .

#### 4.5.1. Intervalos de credibilidad y distribución predictiva

Para la obtención de los intervalos de credibilidad basta tener en cuenta la relación de la distribución gamma con la distribución chi-cuadrado de Pearson. Se verifica entonces que,

$$2\alpha(A(n, L) + 1/\sigma)|X, \sigma \sim \chi^2(2(n + \alpha_0)). \quad (4.20)$$

donde  $\chi^2(n)$  representa una distribución chi-cuadrado con  $n$  grados de libertad. Por tanto, el intervalo de credibilidad  $(a, b)$  debe de verificar que,

$$\Pr(a \leq 2\alpha(A(n, L) + 1/\sigma)|X, \sigma \leq b) = 1 - q, \quad (4.21)$$

donde  $0 \leq q \leq 1$ . A partir de (4.21), se obtiene que los límites del intervalo de credibilidad vienen dados por,

$$a = \frac{\chi^2(2(n + \alpha_0); q/2)}{2(A(n, L) + 1/\sigma)}, \quad (4.22)$$

$$b = \frac{\chi^2(2(n + \alpha_0); 1 - q/2)}{2(A(n, L) + 1/\sigma)}, \quad (4.23)$$

donde ahora donde  $\chi^2(n; q)$  representa el cuantil de orden  $q$  de una distribución chi-cuadrado con  $n$  grados de libertad. Usando la fórmula general (4.4), la distribución predictiva viene dada por,

$$\begin{aligned} f_{y|X}(y|x) &= \int_0^\infty f(y|\alpha)\pi(\alpha|X)d\alpha \\ &= \frac{\tilde{\alpha}_0 \tilde{\sigma}^{-\tilde{\alpha}_0}}{y(1/\tilde{\sigma} + \log(y/L))^{\tilde{\alpha}_0+1}}. \end{aligned}$$

#### 4.5.2. Caso de estudio: datos de pérdidas por vientos

En esta sección utilizaremos los resultados previos para el análisis de los datos de pérdidas de vientos que aparecen en Hogg y Klugman (1984). Los datos se recogen en la tabla 4.4 y el histograma aparece representado en la figura 4.3. Los datos presentan las características ya comentadas de asimetría, cero-modalidad y colas largas. Puesto que el mínimo de los datos es 2, elegimos  $L = 2$ , que coincide con el estimador de máxima verosimilitud. La función de verosimilitud de los datos es,

$$L(X; \alpha) = \alpha^n L^{n\alpha} \left( \prod_{i=1}^n X_i \right)^{-(\alpha+1)},$$

que aparece representada en la Figura 4.4. El estimador de máxima verosimilitud es entonces,

$$\hat{\alpha}_{EMV} = 0,9763$$

con error estándar:

$$\text{ste}(\hat{\alpha}) = \left( -\frac{\partial^2 L(X|\alpha)}{\partial \alpha^2} \right)^{-1/2} = 0,1544$$

al sustituir en el valor del máximo. El valor del logaritmo de la verosimilitud es  $-109.658$ . La función de densidad ajustada junto con el histograma de los datos, aparecen representadas en la figura 4.5.

Para la estimación Bayesiana de  $\alpha$  necesitamos asignar los parámetros  $\alpha_0$  y  $\sigma$  en (4.16). Se presentan dos escenarios:

- *Escenario 1.* Incertidumbre baja. Asignamos mediante dos percentiles:

$$\Pr(\alpha \leq 1,5) = 0,95, \quad \Pr(\alpha \leq 0,5) = 0,05.$$

- *Escenario 2.* Incertidumbre alta. Obtenemos los valores de los parámetros a partir de las especificaciones:

$$\Pr(\alpha \leq 2,5) = 0,95, \quad \Pr(\alpha \leq 0,3) = 0,05.$$

Los valores obtenidos bajo los dos escenarios, aparecen recogidos en la tabla 4.5. Se han incluido además los estimadores Bayesianos (fórmula (4.19)), el factor de credibilidad y los intervalos de credibilidad (fórmulas (4.22)-(4.23)) al 90, 95 y 99 por ciento. Se trata de resultados bastante robustos, pues a pesar de partir de distribuciones a priori muy diferentes, se llegan a estimaciones finales muy similares. La figura 4.6 representa las funciones de densidad a priori y a posteriori bajos los dos escenarios.

Tabla 4.4: Datos de pérdidas por vientos (Hogg y Klugman, 1984).

2	2	2	2	2	2	2	2	2	2
2	2	3	3	3	3	4	4	4	5
5	5	5	6	6	6	6	8	8	9
15	17	22	23	24	24	25	27	32	43

### 4.5.3. Otras distribuciones a priori

En la estimación Bayesiana de datos de Pareto existen otras posibilidades para la elección de la distribución a priori. Esta elección dependerá de la naturaleza de los datos disponibles (por ejemplo censurados) y de la información a priori disponible. En este sentido Arnold y Press (1989) han considerado diversas distribuciones a priori para  $(\alpha, \sigma)$  bajo diversos tipos de censura de los datos. Distribuciones a priori basadas en especificación condicional han sido consideradas por Arnold et al. (1998) y Sarabia et al. (2005), de acuerdo con la metodología de las distribuciones especificadas condicionalmente (Arnold et al, 1999).

## 4.6. Estimación Bayesiana de una clase de distribuciones de pérdida

Las distribuciones paramétricas utilizadas como modelos de pérdidas (es decir, distribuciones para la magnitud de la pérdida) en los seguros no-vida, tienen un

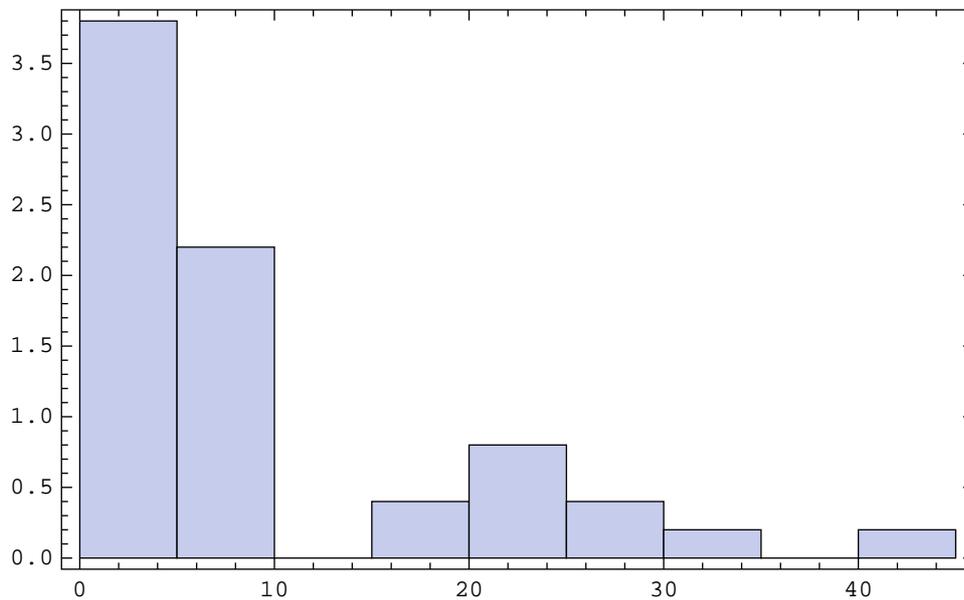


Figura 4.3: Histograma de los datos de pérdidas por vientos (Hogg y Klugman, 1984).

Tabla 4.5: Estimadores Bayesianos y parámetros con los datos de pérdidas por vientos bajo dos escenarios.

Parámetros y estimadores	Escenario 1	Escenario 2
$\alpha_0$	9.3869	2.8079
$\sigma$	0.1005	0.4162
$\hat{\alpha}_0$	49.3869	42.8079
$\hat{\sigma}$	0.0196	0.0231
$E[\alpha]$	0.9432	1.1687
$E[\alpha X]$	0.9698	0.9869
$\hat{\alpha}(EMV)$	0.9763	0.9763
Factor de credibilidad	0.8046	0.9446
Intervalo de credibilidad 90 %	0.7545–1.2074	0.7526–1.4186
Intervalo de credibilidad 95 %	0.7184–1.2584	0.7137–1.3038
Intervalo de credibilidad 99 %	0.6512–1.3620	0.6417–1.4186

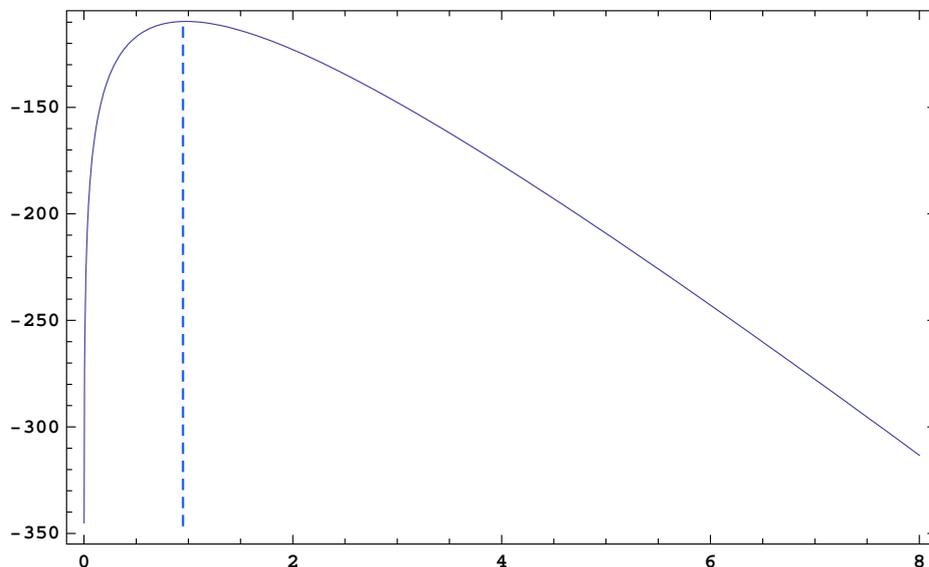


Figura 4.4: Función de verosimilitud de Pareto para los datos de pérdidas por vientos (Hogg y Klugman, 1984).

soporte positivo, colas pesadas y presentan asimetría positiva, tal como hemos evidenciado con los datos reales de pérdidas (secciones 3.3, 4.5.2 y Hogg y Klugman (1984)).

En este apartado vamos a realizar un análisis Bayesiano empírico sobre una clase amplia de distribuciones de pérdida, considerada por Haberman y Renshaw (1996). Este tipo de análisis no ha sido hecho hasta el momento. Además, obtendremos un estimador Bayesiano del parámetro que tiene la propiedad de poder representarse como una fórmula de credibilidad.

Nos fijamos entonces en la clase de distribuciones con función de densidad,

$$f(x; \alpha, \theta) = \alpha \lambda(x; \theta) \exp \{-\alpha \Lambda(x; \theta)\}, \quad x > d \geq 0, \quad (4.24)$$

y función de supervivencia,

$$S(x; \alpha, \theta) = \exp \{-\alpha \Lambda(x; \theta)\}, \quad x > d \geq 0, \quad (4.25)$$

donde  $\lambda(x; \theta)$  es proporcional a la tasa de fallo y

$$\Lambda(x; \theta) = \int_d^x \lambda(u; \theta) du, \quad (4.26)$$

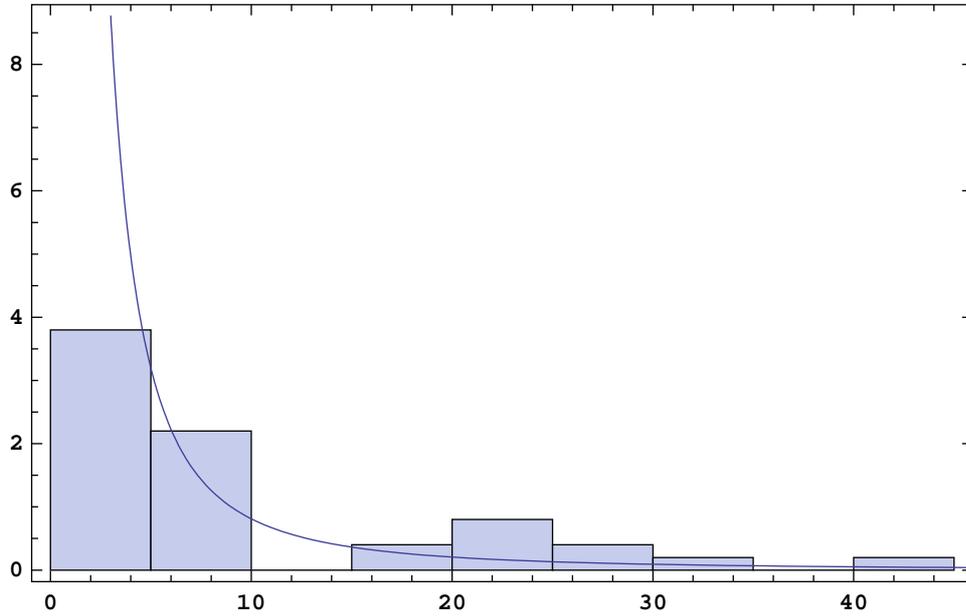


Figura 4.5: Histograma y función de densidad de Pareto ajustada a los datos de pérdidas por vientos (Hogg y Klugman, 1984).

representa la tasa de fallo integrada, donde  $\theta$  es un parámetro (que puede ser de dimensión mayor o igual que uno) y  $\alpha > 0$ . Casos especiales de (4.24)-(4.25) incluyen a las distribuciones de Burr, Pareto generalizada y Weibull (ver Tabla 4.6). Disponemos de una muestra de datos de pérdidas  $X = (X_1, \dots, X_n)$  para la estimación de (4.24). Procedemos en dos etapas:

- Etapa 1: estimación de  $\theta$  mediante métodos clásicos.
- Etapa 2: estimación de  $\alpha$  mediante técnicas Bayesianas.

En la primera etapa estimamos  $\theta$  mediante máxima verosimilitud. De modo alternativo, podemos obtener  $\theta$  mediante un método de estimación basado en percentiles (ver Castillo et al., 2005). Si  $x_{q_i}$ ,  $0 < q_i < 1$ ,  $i = 1, 2$  representan dos percentiles de (4.25) tenemos que,

$$\log(1 - q_i) = -\alpha \Lambda(x_{q_i}; \theta), \quad i = 1, 2.$$

Eliminando  $\alpha$  tenemos que,

$$\frac{\Lambda(x_{q_1}; \theta)}{\Lambda(x_{q_2}; \theta)} = \frac{\log(1 - q_1)}{\log(1 - q_2)}, \quad (4.27)$$

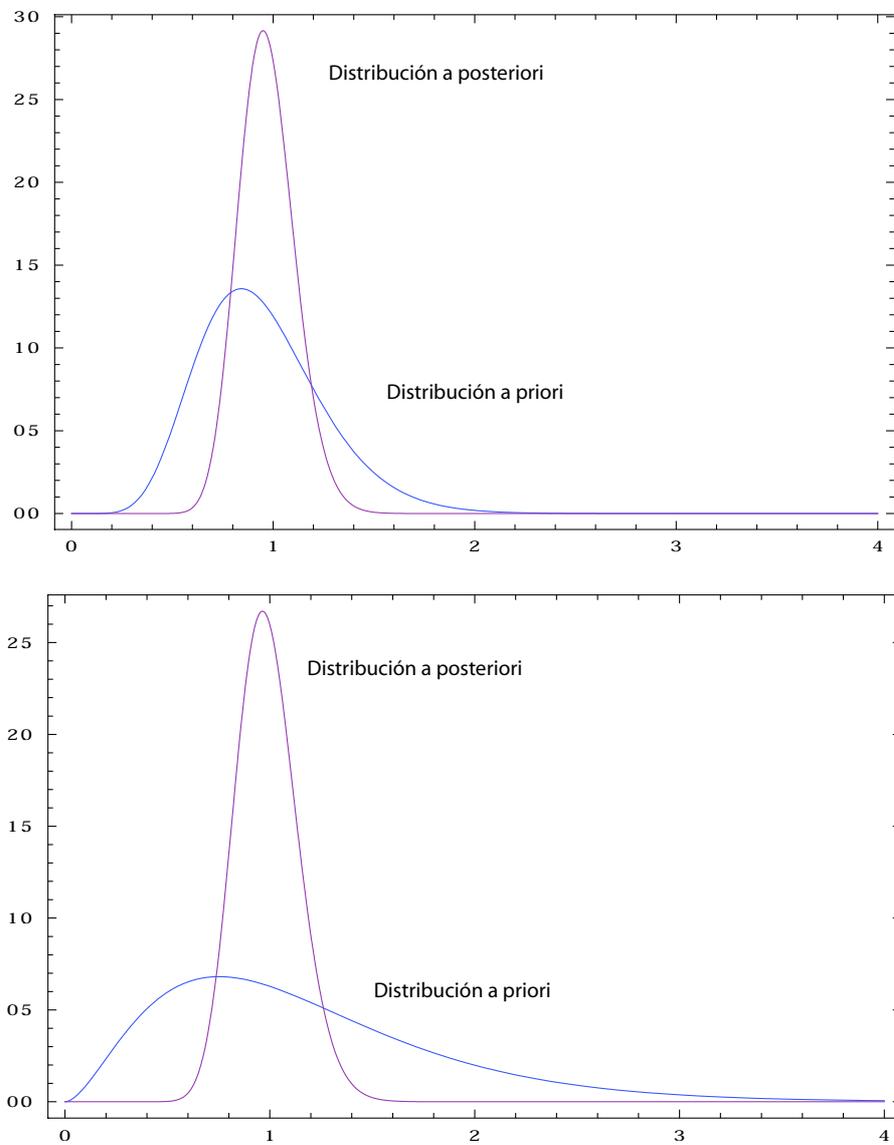


Figura 4.6: Funciones de densidad a priori y a posteriori para los datos de pérdidas por vientos Escenario 1 (gráfico superior) y Escenario 2 (gráfico inferior).

que sólo depende de  $\theta$ . Si suponemos que  $\theta$  es un escalar se puede obtener un estimación mediante un procedimiento iterativo. A partir de la estimación de  $\theta$  y en una segunda etapa, realizamos un análisis Bayesiano clásico eligiendo como distribución a priori para  $\alpha$  una gamma con parámetros  $(a, b)$  y media  $E[\alpha] = a \cdot b$ . La función de verosimilitud viene dada por,

$$L(X|\alpha, \theta) = \alpha^n \prod_{i=1}^n \lambda(X_i; \theta) \exp \left\{ -\alpha \sum_{i=1}^n \Lambda(X_i; \theta) \right\}. \quad (4.28)$$

La función de densidad a posteriori es,

$$\pi_{X|\alpha, \theta}(X|\alpha, \theta) \propto L(X|\alpha, \theta) \cdot \pi(\alpha) \quad (4.29)$$

$$\propto \alpha^{a+n-1} \exp \left\{ -\alpha \left( \sum_{i=1}^n \Lambda(X_i; \theta) + 1/b \right) \right\}, \quad (4.30)$$

que es una distribución gamma. Los parámetros de esta distribución se actualizan según la tabla 4.7. Utilizando una función de pérdida cuadrática, el estimador Bayesiano de  $\alpha$  viene dado por,

$$\begin{aligned} E[\alpha|X, \theta] &= \frac{a+n}{\sum_{i=1}^n \Lambda(X_i; \theta) + 1/b} \\ &= \frac{\sum_{i=1}^n \Lambda(X_i; \theta)}{\sum_{i=1}^n \Lambda(X_i; \theta) + 1/b} \cdot \frac{n}{\sum_{i=1}^n \Lambda(X_i; \theta)} + \frac{1/b}{\sum_{i=1}^n \Lambda(X_i; \theta) + 1/b} \cdot a \cdot b \\ &= Z \cdot \tilde{\alpha}_{EMV} + (1-Z) \cdot E[\alpha], \end{aligned} \quad (4.31)$$

que corresponde a una fórmula de credibilidad (no estricta). En (4.31), el factor de credibilidad es

$$Z = \frac{\sum_{i=1}^n \Lambda(X_i; \theta)}{\sum_{i=1}^n \Lambda(X_i; \theta) + 1/b}. \quad (4.32)$$

## 4.7. Estimación de la prima de riesgo

La prima calculada, de acuerdo a los principios de cálculo de primas estudiados en el capítulo anterior, dependerá en la práctica de algún parámetro del que dependa la función de distribución que sigue la variable aleatoria  $X$  asociada al riesgo.

Tabla 4.6: Distribuciones de pérdida del tipo (4.24).

	Burr, $\alpha, \lambda, \gamma > 0$	Pareto gen., $\alpha, \lambda > 0$	Weibull, $\alpha, \gamma > 0$
$f(x)$	$\alpha \gamma x^{\gamma-1} \frac{(\lambda + d^\gamma)^\alpha}{(\lambda + x^\gamma)^{\alpha+1}}$	$\alpha \frac{(\lambda + d)^\alpha}{(\lambda + x)^{\alpha+1}}$	$\alpha \gamma x^{\gamma-1} \exp\{-\alpha(x^\gamma - d^\gamma)\}$
$S(x)$	$\left(\frac{\lambda + d^\gamma}{\lambda + x^\gamma}\right)^\alpha$	$\left(\frac{\lambda + d}{\lambda + x}\right)^\alpha$	$\exp\{-\alpha(x^\gamma - d^\gamma)\}$
$\lambda(x)$	$\frac{\gamma x^{\gamma-1}}{\lambda + x^\gamma}$	$\frac{1}{\lambda + x}$	$\gamma x^{\gamma-1}$
$\Lambda(x)$	$\log\left(\frac{\lambda + d^\gamma}{\lambda + x^\gamma}\right)$	$\log\left(\frac{\lambda + d}{\lambda + x}\right)$	$x^\gamma - d^\gamma$

Tabla 4.7: Actualización de parámetros de la familia de distribuciones de pérdida (4.24).

Parámetros	Valores a priori	Valores actualizados
$a$	$a$	$\tilde{a} = a + n$
$b$	$b$	$\tilde{b} = \frac{1}{\sum_{i=1}^n \Lambda(X_i; \theta) + 1/b}$

Por ejemplo, si el riesgo sigue una distribución de Poisson de parámetro  $\theta > 0$ , la prima calculada conforme al principio de prima neta vendrá dada por

$$H(X) = E[X] = \theta.$$

En la literatura actuarial es habitual considerar que todos o algunos de los parámetros de los que depende la función de distribución de probabilidad son desconocidos.

De ahí que, dado que el parámetro  $\theta$  se desconoce, a la prima se le denomine prima pura de riesgo o simplemente prima de riesgo. De aquí en adelante suponemos que la prima dependerá de al menos un parámetro desconocido y escribiremos  $P_R$  para denotar a la prima de riesgo que, insistimos dependerá de al menos un parámetro.

En el ejemplo anterior se verifica que:

$$P_R \equiv H[X|\theta] = \theta.$$

El problema que surge ahora es estimar la prima suponiendo que el parámetro del que depende es desconocido y aleatorio. De ello nos ocupamos en los siguientes apartados.

### 4.7.1. Prima colectiva o a priori

Si  $f(x)$  es la función de densidad asociada al riesgo  $X$  dependiente de un parámetro  $\theta$ , escribiremos  $f(x; \theta)$  o  $f(x|\theta)$  según que el parámetro sea fijo o aleatorio. Podemos ahora suponer que dicho parámetro se distribuye entre toda la cartera de seguros de acuerdo a cierta función de densidad  $\pi(\theta)$ . Desde un punto de vista Bayesiano ésta no es más que la distribución a priori, denominada en el escenario actuarial función estructura.

Para clarificar el tema pensemos en una cartera de seguro de automóviles. Cada asegurado tiene una propensión diferente a experimentar reclamación porque cada uno de ellos poseen diferentes factores de riesgo. Como ejemplo de factores de riesgo podemos citar la edad, los años de vigencia de la licencia de conducir, la potencia del vehículo, la zona en la que circula, etc. El actuario puede desconocer la propensión de un asegurado a experimentar reclamación, pero puede tener cierta idea acerca de cómo se distribuye dicha propensión en la cartera de seguros. Es decir, puede asignar una distribución de probabilidad al parámetro que recoge esta propensión. En el ejemplo considerado anteriormente, la propensión viene dada por el parámetro  $\theta$ .

En principio, la mejor estimación que puede obtenerse de la prima de riesgo es la prima colectiva que se define de la siguiente forma.

**Definición 4.11** *Dados un riesgo  $X$  con función de densidad  $f(x|\theta)$ , siendo  $\theta$  un parámetro desconocido con función de densidad a priori  $\pi(\theta)$  y una función de pérdida  $L : \mathbb{R}^2 \rightarrow \mathbb{R}$ , la prima colectiva es el valor  $P_C$  que minimiza la pérdida esperada*

$$\int_{\Theta} L(P_R, P_C) \pi(\theta) d\theta, \quad (4.33)$$

siendo  $P_R$  la prima de riesgo definida en (2.16).

Obsérvese que en (4.33)  $P_R$  depende del parámetro desconocido  $\theta$ .

La prima colectiva tal y como aparece definida más arriba representa la mejor decisión que estima la prima de riesgo (obviamente desconocida). Observemos que para calcularla se necesitará que el actuario defina una distribución de probabilidad, la distribución a priori, para el valor del parámetro desconocido  $\theta$ . Para ello será fundamental la experiencia de lo acontecido en los períodos precedentes o en otros contratos similares.

**Ejemplo 4.1** Sea un riesgo  $X$  con función de densidad de probabilidad  $f(x|\theta)$ , dependiente del parámetro  $\theta$  desconocido y aleatorio y con distribución a priori (función estructura)  $\pi(\theta)$ . Vamos a obtener la prima colectiva para los principios de prima neta, exponencial, Esscher y de varianza.

Usando los diferentes principios tenemos que:

1. Para el principio de prima se verifica  $L(P_R, P_C) = (P_R - P_C)^2$ , entonces  $P_R$  es como en (2.17) y se obtiene,

$$P_C = \int_{\Theta} \left[ \int_{\mathcal{X}} xf(x|\theta)dx \right] \pi(\theta)d\theta = E_{\pi} [E_f [X|\theta]]. \quad (4.34)$$

2. Para el principio de utilidad exponencial tenemos:

$$L(P_R, P_C) = \frac{1}{\alpha}(e^{\alpha P_R} - e^{\alpha P_C})^2,$$

con  $\alpha > 0$ , entonces  $P_R$  es como en (2.18), de donde,

$$P_C = \frac{1}{\alpha} \log \int_{\Theta} e^{\alpha P_R} \pi(\theta) d\theta = \frac{1}{\alpha} \log E_{\pi} [e^{\alpha P_R}]. \quad (4.35)$$

3. Para el principio Esscher tenemos que

$$L(P_R, P_C) = e^{\alpha P_R} (P_R - P_C)^2, \quad \alpha > 0,$$

entonces  $P_R$  viene dada en (2.19), y se obtiene,

$$P_C = \frac{\int_{\Theta} P_R e^{\alpha P_R} \pi(\theta) d\theta}{\int_{\Theta} e^{\alpha P_R} \pi(\theta) d\theta} = \frac{E_{\pi} [P_R e^{\alpha P_R}]}{E_{\pi} [e^{\alpha P_R}]}. \quad (4.36)$$

4. Finalmente, para el principio de varianza tenemos que

$$L(P_R, P_C) = P_R (P_R - P_C)^2,$$

entonces  $P_R$  viene dada como en (2.20), y se deduce:

$$P_C = \frac{\int_{\Theta} P_R^2 \pi(\theta) d\theta}{\int_{\Theta} P \pi(\theta) d\theta} = \frac{E_{\pi} [P_R^2]}{E_{\pi} [P_R]}. \quad (4.37)$$

**Ejemplo 4.2** Supongamos que el número de reclamaciones de una cartera de seguros sigue una distribución binomial negativa con parámetros  $r > 0$  y  $\theta > 0$  de la forma:

$$\Pr(X = x|\theta) = \binom{r+x-1}{x} \left( \frac{r}{r+\theta} \right)^r \left( \frac{\theta}{r+\theta} \right)^x, \quad x = 0, 1, \dots; \quad r, \theta > 0,$$

y  $\theta$  sigue una distribución beta de segunda especie con parámetros  $r, b, \nu > 0$ , de modo que

$$\pi(\theta) = \frac{r^b}{B(\nu, b)} \frac{\theta^{\nu-1}}{(r + \theta)^{b+\nu}}.$$

Vamos a calcular, bajo el principio de prima neta la prima de riesgo y colectiva.

La prima de riesgo es

$$P_R = E[X|\theta] = \theta,$$

mientras que la prima colectiva viene dada por

$$P_C = E_\pi[P_R] = E_\pi(\theta) = \frac{r\nu}{b-1}, \quad b > 1. \quad (4.38)$$

**Ejemplo 4.3** Veamos cómo se calcula la prima de riesgo y colectiva bajo el principio de varianza suponiendo que el riesgo  $X$  sigue la distribución gamma con parámetros  $\theta > 0$ , y  $\nu > 0$  conocido,

$$f(x|\theta) = \frac{\theta^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\theta x}, \quad \nu > 0, \theta > 0$$

y  $\theta$  sigue en el colectivo la distribución gamma con parámetros  $a, b$ :

$$\pi(\theta) = \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta}, \quad a > 0, b > 0.$$

La prima de riesgo es

$$P_R = \frac{\int_0^\infty x^2 f(x|\theta) dx}{\int_0^\infty x f(x|\theta) dx} = \frac{\nu + 1}{\theta},$$

y la prima colectiva

$$P_C = \frac{\int_0^\infty P_R^2 \pi(\theta) d\theta}{\int_0^\infty P_R \pi(\theta) d\theta} = \frac{b(\nu + 1)}{a - 2}, \quad a > 2. \quad (4.39)$$

#### 4.7.2. Prima Bayes o a posteriori

Los modelos de decisión más habituales son aquellos en los que el decisor dispone de experiencia previa: son los llamados problemas de decisión con experimentación. En el problema de decisión con experimentación, la decisión elegida depende del vector observado,  $\underline{X} = (X_1, \dots, X_n)$ , donde la observación consiste

en una muestra de tamaño  $n$ ,  $X_1, \dots, X_n$ , procedente de una población  $X$  cuya función de distribución depende del estado  $\Theta$  que se presente.

En este caso, el análisis Bayesiano nos permitirá combinar la información inicial o a priori que se tiene sobre el parámetro  $\theta \in \Theta$  con la información muestral,  $\underline{X}$ , para obtener la distribución a posteriori del parámetro.

Así, si  $\pi(\theta)$  es la densidad a priori, que refleja las creencias sobre  $\theta$  antes de obtener la información muestral  $\underline{X} = (X_1, \dots, X_n)$ , el teorema de Bayes, como se mostró anteriormente, permite combinar estas dos fuentes de información para, en nuestro caso, obtener otra estimación de la prima de riesgo, distinta de la prima colectiva. El resultado es la prima Bayes o a posteriori que se define a continuación.

**Definición 4.12** *Dados un riesgo  $X$  con función de densidad de probabilidad dada por  $f(x|\theta)$ , siendo  $\theta$  un parámetro desconocido con distribución a priori  $\pi(\theta)$ , una función de pérdida  $L : \mathbb{R}^2 \rightarrow \mathbb{R}$  y un vector de datos observados  $\underline{X} = (X_1, \dots, X_n)$ , la prima Bayes es el valor  $P_B$  que minimiza:*

$$\int_{\Theta} L(P_R, P_B) \pi(\theta|\underline{X}) d\theta, \quad (4.40)$$

siendo  $\pi(\theta|\underline{X})$  la distribución a posteriori de  $\theta$  dada la muestra y  $P_R$  la prima de riesgo definida anteriormente.

Nuevamente  $P_R$  en (4.40) depende del parámetro  $\theta$ .

**Ejemplo 4.4** Sea un riesgo  $X$  con función de densidad de probabilidad  $f(x|\theta)$ , dependiente del parámetro  $\theta$  desconocido y aleatorio con distribución a priori  $\pi(\theta)$ . Vamos a obtener la prima Bayes para los cuatro principios de prima neta, exponencial, Esscher y de varianza.

Las expresiones de las primas Bayes se obtienen intercambiando en (4.34), (4.35), (4.36) y (4.37)  $\pi(\theta)$  por  $\pi(\theta|\underline{X})$ , obteniéndose:

$$P_B = \int_{\Theta} \left[ \int_{\mathcal{X}} x f(x|\theta) dx \right] \pi(\theta|\underline{X}) d\theta = E_{\pi(\theta|\underline{X})} [E_f(X|\theta)], \quad (4.41)$$

para el principio de prima neta. Entonces:

$$P_B = \frac{1}{\alpha} \log \int_{\Theta} e^{\alpha P_R} \pi(\theta|\underline{X}) d\theta = \frac{1}{\alpha} \log E_{\pi(\theta|\underline{X})} [e^{\alpha P_R}], \quad (4.42)$$

para el principio exponencial.

$$P_B = \frac{\int_{\Theta} P_R e^{\alpha P_R} \pi(\theta|\underline{X}) d\theta}{\int_{\Theta} e^{\alpha P_R} \pi(\theta|\underline{X}) d\theta} = \frac{E_{\pi(\theta|\underline{X})} [P_R e^{\alpha P_R}]}{E_{\pi(\theta|\underline{X})} [e^{\alpha P_R}]}, \quad (4.43)$$

para el principio Esscher. Finalmente:

$$P_B = \frac{\int_{\Theta} P_R^2 \pi(\theta|\underline{X}) d\theta}{\int_{\Theta} P \pi(\theta|\underline{X}) d\theta} = \frac{E_{\pi(\theta|\underline{X})} [P_R^2]}{E_{\pi} [P_R]}, \quad (4.44)$$

para el principio de varianza.

**Ejemplo 4.5** Supongamos que disponemos de una muestra aleatoria simple  $\underline{X} = (X_1, \dots, X_n)$  procedente de la población del ejemplo 4.2. Vamos a calcular la prima neta Bayes.

La distribución a posteriori de  $\theta$  dada la muestra  $\underline{X} = (X_1, \dots, X_n)$  es:

$$\begin{aligned} \pi(\theta|\underline{X}) &\propto \left(\frac{r}{r+\theta}\right)^{nr} \left(\frac{\theta}{r+\theta}\right)^{n\bar{X}} \frac{\theta^{\nu-1}}{(r+\theta)^{b+\nu}} \\ &\propto \frac{\theta^{\nu+n\bar{X}-1}}{(r+\theta)^{b+\nu+n\bar{X}+nr}}, \end{aligned}$$

que es de nuevo una distribución beta de segunda especie con parámetros  $r, b + nr, \nu + n\bar{X}$ . Luego, la prima neta Bayes se obtiene reemplazando en (4.38) los parámetros  $r, b$  y  $\nu$  por los parámetros actualizados, resultando:

$$P_B = \frac{r(\nu + n\bar{X})}{b + nr - 1}. \quad (4.45)$$

**Ejemplo 4.6** Supongamos que disponemos de una muestra aleatoria simple  $\underline{X} = (X_1, \dots, X_n)$  procedente de la población del ejemplo 4.3. Vamos a calcular la prima Bayes de varianza.

Es fácil deducir que la distribución a posteriori es en este caso de nuevo una distribución  $\mathcal{G}(a + n\nu, b + n\bar{X})$ , luego la prima Bayes de varianza se obtiene actualizando en (4.39) los parámetros correspondientes, obteniendo:

$$P_B = \frac{(\nu + 1)(b + n\bar{X})}{a + n\nu - 2}.$$

Resulta interesante destacar que para calcular las primas colectiva y Bayes podemos utilizar un principio de cálculo de prima diferente que el utilizado para calcular la prima de riesgo,

**Ejemplo 4.7** Supongamos que el número de reclamaciones de una cartera de seguros sigue una distribución de Poisson con parámetros  $\theta > 0$  y  $\theta$  sigue una distribución  $\mathcal{G}(a, b)$ ,  $a > 0$ ,  $b > 0$ . Vamos a calcular la prima Bayes de acuerdo al principio Esscher cuando la prima de riesgo se ha calculado bajo el principio de prima neta.

Ya sabemos que bajo el principio de prima neta la prima de riesgo es  $P_R = \theta$ . La prima colectiva es:

$$\begin{aligned} P_C &= \frac{\int_0^\infty \theta e^{\alpha\theta} \pi(\theta) d\theta}{\int_0^\infty e^{\alpha\theta} \pi(\theta) d\theta} = \frac{\int_0^\infty \theta^a e^{-(\alpha+b)\theta} d\theta}{\int_0^\infty \theta^{a-1} e^{-(\alpha+b)\theta} d\theta} \\ &= \frac{a}{\alpha + b}. \end{aligned} \quad (4.46)$$

Reemplazando en (4.46)  $a$  y  $b$  por  $a + n\bar{X}$  y  $b + n$ , respectivamente, se obtiene la prima Bayes, dada por:

$$P_B = \frac{a + n\bar{X}}{\alpha + b + n}. \quad (4.47)$$

Nótese que (4.47) es una fórmula de credibilidad, puesto que

$$P_B = Z(n)\bar{X} + [1 - Z(n)]P_C,$$

con:

$$Z(n) = \frac{n}{\alpha + b + n}.$$

## 4.8. Sistemas bonus-malus

Con objeto de reducir el número de reclamaciones en el sector de seguro de automóviles, las compañías Europeas introdujeron el sistema de tarificación *bonus-malus*. Este sistema penaliza a los conductores que no experimentan reclamación, los buenos, penalizando a los malos. Por tanto, según este mecanismo la prima se calcula en función del número de reclamaciones que experimenta el asegurado, lo que favorece la utilización de la metodología desarrollada en teoría de la credibilidad.

Es conocido que en los últimos años se ha fomentado entre las entidades que operan en el sector del automóvil, el uso del fichero histórico de siniestralidad de conductores (SINCO), que incluye el historial de siniestralidad de los últimos cinco años de cada póliza. El fichero SINCO tiene la ventaja de proporcionar a las entidades aseguradoras adheridas al mismo, los historiales de cada conductor antes de aceptarlo. Este hecho favorece, por un lado, ante un posible cambio de entidad aseguradora, la fijación del precio que corresponde al riesgo contratado, y desde un punto de vista metodológico, la posibilidad de contar siempre con una muestra personalizada de un asegurado en concreto, lo que facilita y favorece el uso de determinadas técnicas estadísticas, que sin la misma no podría llevarse a cabo.

Para construir un sistema bonus-malus se parte de un nivel  $X$  neutro, de modo que para niveles inferiores a  $X$  el asegurado entra en la escala *bonus*, mientras que para niveles superiores a  $X$ , el asegurado se incorpora a la escala *malus*.

Este tipo de sistemas está generalmente basado en el número de reclamaciones, y no en la cuantía. Expresando la prima como una función del número medio de reclamaciones  $\bar{X}$ ,  $n\bar{X} = \sum_{i=1}^n X_i$ , y del período de tiempo  $n$ , representaremos la prima *bonus-malus* mediante  $P_{BM} \equiv P_{BM}(\bar{X}, n)$ .

Lo relevante de este sistema de tarificación es que un asegurado que en el período actual no presente reclamación se verá bonificado en el siguiente período mediante un descuento en la prima a pagar. Por el contrario, si experimenta reclamación se verá penalizado con un incremento de la prima. Luego, tendrán que verificarse las siguientes reglas de transición:

$$\frac{\partial P_{BM}(\bar{X}, n)}{\partial \bar{X}} > 0, \quad \frac{\partial P_{BM}(\bar{X}, n)}{\partial n} < 0.$$

Existen diversas metodologías para el cálculo de primas bonus-malus, entre las que destacamos:

1. *Método Markoviano*. En este caso el sistema bonus-malus se contempla como un proceso de Markov en el que el asegurado se mueve de un estado a otro en el tiempo. Ver Centeno y Silva (2002) y Lemaire (1995).
2. *Programas de optimización multiobjetivo*. Desarrollados por Heras et al. (2002, 2004).
3. *Métodos Bayesianos*. Estos métodos serán estudiados en esta investigación. Algunas referencias son: Gómez-Déniz et al. (2002b, 2006b, 2008c, 2008d, 2008e), Gómez-Déniz y Vázquez (2005c), Lemaire (1979, 1985, 1995), Meng et al. (1999) y Sarabia et al. (2004).

#### 4.8.1. Cálculo de primas bonus-malus. Método Bayesiano

Una forma de obtener primas que cumplan las *reglas de transición*, utilizando la metodología Bayesiana, consiste en dividir la prima Bayes entre la prima colectiva para los principios de cálculo de prima estudiados. Denotando ahora mediante  $P_B$  a la prima Bayes, la prima bonus-malus puede obtenerse como:

$$P_{BM} = \frac{P_B}{P_C}, \quad (4.48)$$

donde  $P_C$  representa la prima colectiva.

**Ejemplo 4.8** Vamos a obtener las expresiones de la prima bonus-malus para los principios de cálculo de prima neta, exponencial, Esscher y varianza.

Utilizando las expresiones (4.48), (4.34) y (4.41), se deduce que la prima bonus-malus bajo el principio de prima neta es:

$$P_{BM} = \frac{E_{\pi(\theta|\underline{X})} [E_{f(x|\theta)} [X|\theta]]}{E_{\pi(\theta)} [E_{f(x|\theta)} [X|\theta]]}. \quad (4.49)$$

Utilizando (4.48), (4.35) y (4.42), se deduce que la prima bonus-malus bajo el principio de prima exponencial es:

$$P_{BM} = \frac{\log \{E_{\pi(\theta|\underline{X})} [e^{\alpha P_R}]\}}{\log \{E_{\pi(\theta)} [e^{\alpha P_R}]\}}, \quad (4.50)$$

donde

$$P_R = \frac{1}{\alpha} \log \int_{\Theta} e^{\alpha x} f(x|\theta) dx.$$

Utilizando (4.48), (4.36) y (4.43), se deduce que la prima bonus-malus bajo el principio de prima Esscher es:

$$P_{BM} = \frac{E_{\pi(\theta|\underline{X})} [P_R e^{\alpha P_R}]}{E_{\pi(\theta|\underline{X})} [e^{\alpha P_R}]} \frac{E_{\pi(\theta)} [e^{\alpha P_R}]}{E_{\pi(\theta)} [P_R e^{\alpha P_R}]},$$

donde

$$P_R = \frac{E_{f(x|\theta)} [X e^{\alpha X} | \theta]}{E_{f(x|\theta)} [e^{\alpha X} | \theta]}.$$

Finalmente, utilizando (4.48), (4.37) y (4.44), se deduce que la prima bonus-malus bajo el principio de prima de varianza es:

$$P_{BM} = \frac{E_{\pi(\theta|\underline{X})} [P_R^2]}{E_{\pi(\theta|\underline{X})} [P_R]} \frac{E_{\pi(\theta)} [P_R]}{E_{\pi(\theta)} [P_R^2]},$$

donde

$$P_R = \frac{E_{f(x|\theta)} [X^2 | \theta]}{E_{f(x|\theta)} [X | \theta]}.$$

**Ejemplo 4.9** La tabla 4.8 muestra el número de asegurados para las reclamaciones  $x = 0, 1, \dots$  en una cartera de seguros de automóviles en Alemania en 1960. Esta cartera de seguros aparece en Willmot (1987) y en Gómez-Déniz et al. (2008c). Suponiendo los siguientes modelos:

1. Función de verosimilitud Poisson de parámetro  $\theta > 0$  y una distribución a priori gamma de parámetros  $a > 0$ ,  $b > 0$ .
2. Función de verosimilitud binomial negativa con parámetros  $r > 0$ ,  $0 < \theta < 1$  y una distribución a priori para el parámetro  $\theta$  beta con parámetros  $a > 0$ ,  $b > 0$ .

Para ambos modelos, se trata de obtener:

1. La prima bonus-malus bajo el principio de prima neta.

2. Comprobar que dicha prima puede escribirse como una fórmula de credibilidad.
3. La distribución incondicional del número de reclamaciones.
4. Estimar mediante el método de máxima verosimilitud los parámetros de la distribución a priori.
5. Obtener la tabla de frecuencias ajustadas.
6. Estudiar la bondad del ajuste mediante el test  $\chi^2$  con un tamaño del 5 por ciento. ¿Qué modelo ajusta mejor los datos?
7. Obtener una matriz de orden  $3 \times 4$  en la que el elemento  $a_{ij}$ ,  $i = 1, 2, 3$ ;  $j = 1, 2, \dots, 4$  sea el porcentaje de prima a aplicar en el período  $i$  con  $j$  reclamaciones, bajo un sistema de tarificación bonus-malus y el principio de prima neta.

Tabla 4.8: Número de siniestros observados.

Número de siniestros	Frecuencias absolutas
	Observadas
0	20592
1	2651
2	297
3	41
4	7
5	0
6	1
Total	23589

Veamos la resolución de las diferentes cuestiones planteadas:

1. Utilizando (4.49) se deduce, para el caso Poisson-gamma:

$$P_{BM} = \frac{a + n\bar{X}}{b + n} \frac{b}{a}, \quad \text{con } \sum_{i=1}^n x_i = n\bar{X}. \quad (4.51)$$

Bajo el modelo binomial negativa-beta, las primas netas colectiva y Bayes son, respectivamente:

$$P_C = \frac{rb}{a-1},$$

$$P_B = \frac{r(b + n\bar{X})}{a + nr - 1},$$

de donde se deduce que la prima neta bonus-malus es:

$$P_{BM} = \frac{b + n\bar{X}}{a + nr - 1} \frac{a - 1}{b}. \quad (4.52)$$

2. Es inmediato probar que (4.51) puede reescribirse como:

$$P_{BM} = \frac{n}{b+1} \left( \frac{b}{a} \bar{X} \right) + \frac{b}{b+1} \left( \frac{b}{a} P_C \right) = Z(n)g(\bar{X}) + [1 - Z(n)]g(P_C),$$

siendo  $Z(n) = n/(b+n)$  el factor de credibilidad,  $P_C$  la prima neta colectiva y  $g(x) = \frac{b}{a}x$ . La expresión (4.52) puede reescribirse como

$$P_{BM} = Z(n)g(\bar{X}) + [1 - Z(n)]g(P_C),$$

donde:

$$\begin{aligned} Z(n) &= \frac{rn}{a + nr - 1}, \\ g(x) &= \frac{a - 1}{rb}x. \end{aligned}$$

3. Para el modelo Poisson-gamma, comprobaremos que la distribución incondicional del número de reclamaciones es una binomial negativa con parámetros  $a$  y  $b/(b+1)$ . En efecto:

$$\begin{aligned} \Pr(X = x) &= \int_0^\infty f(x|\theta)\pi(\theta)d\theta \\ &= \int_0^\infty \frac{e^{-\theta}\theta^x}{x!} \cdot \frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}d\theta \\ &= \frac{b^a}{x!\Gamma(a)} \int_0^\infty \theta^{a+x-1}e^{-(b+1)\theta}d\theta \\ &= \frac{b^a}{x!\Gamma(a)} \frac{\Gamma(a+x)}{(b+1)^{a+x}} \\ &= \binom{a+x-1}{x} \left( \frac{b}{b+1} \right)^a \left( \frac{1}{b+1} \right)^x. \quad (4.53) \end{aligned}$$

Procediendo de forma análoga para el modelo binomial negativa-beta tenemos:

$$\begin{aligned} \Pr(X = x) &= \binom{r+x-1}{x} \frac{1}{B(a,b)} \int_0^1 \theta^{a+r-1}(1-\theta)^{b+x-1}d\theta \\ &= \binom{r+x-1}{x} \frac{B(a+r, b+x)}{B(a,b)}. \quad (4.54) \end{aligned}$$

4. A partir de una muestra aleatoria simple de tamaño  $n$  procedente de (4.53), el logaritmo de la función de verosimilitud es,

$$\ell(a, b) = \sum_{i=1}^n \log \binom{a + X_i - 1}{X_i} + na \log \left( \frac{b}{b+1} \right) - n\bar{X} \log(b+1). \quad (4.55)$$

Ahora teniendo en cuenta que

$$\binom{a + X_i - 1}{X_i} = \frac{1}{X_i!} \prod_{j=0}^{X_i-1} (a + X_i - 1 - j),$$

tenemos que (4.55) puede reescribirse como

$$\begin{aligned} \ell(a, b) &= \sum_{i=1}^n \left[ \sum_{j=1}^{X_i-1} \log(a + X_i - 1 - j) - \log X_i! \right] + na \log \left( \frac{b}{b+1} \right) \\ &\quad - n\bar{X} \log(b+1). \end{aligned} \quad (4.56)$$

Derivando parcialmente (4.56) con respecto a  $b$  e igualando a cero resulta:

$$\hat{b} = \frac{a}{\bar{X}}, \quad (4.57)$$

mientras que derivando (4.56) respecto a  $a$  e igualando a cero se obtiene

$$\sum_{i=1}^n \sum_{j=1}^{i-1} \frac{1}{a + X_i - 1 - j} + n \log \left( \frac{b}{b+1} \right) = 0. \quad (4.58)$$

Finalmente, sustituyendo (4.57) en (4.58) resulta una ecuación en  $a$  que puede resolverse numéricamente, obteniéndose los estimadores  $\hat{a} = 1,1179$  y  $\hat{b} = 7,7513$ .

Para el modelo binomial negativa-beta, a partir de una muestra aleatoria simple de tamaño  $n$  procedente de (4.54), el logaritmo de la función de verosimilitud es:

$$\ell(r, a, b) = \sum_{i=1}^n \log \left[ \binom{r + X_i - 1}{X_i} \frac{B(a+r, b+X_i)}{B(a, b)} \right]. \quad (4.59)$$

Calculando las derivadas parciales de (4.59) con respecto a  $r$ ,  $a$  y  $b$  e igualando a cero se obtiene un sistema de ecuaciones que puede resolverse fácilmente utilizando, por ejemplo, el programa Mathematica. La solución que se obtiene para el caso que nos ocupa es:  $r = 2,6895$ ;  $a = 51,1597$ ;  $b = 2,6895$ .

5. Las frecuencias observadas y esperadas del modelo aparecen en la tabla 4.9.

Tabla 4.9: Número de siniestros observados y ajustados.

Número de siniestros	Frecuencias absolutas		
	Observadas	Poisson-gamma	Binomial negativa-beta
0	20592	20596.80	20596.80
1	2651	2631.03	2635.24
2	297	318.37	311.73
3	41	37.81	39.03
4	7	4.44	5.30
5	0	0.52	0.78
6	1	0.06	0.12
Total	23589	23589	23589

6. Para el modelo Poisson-gamma, el logaritmo de la función de verosimilitud es  $-10223,4$ , mientras que el valor de  $\chi^2$  con 2 grados de libertad es 3,62. El  $p$ -valor correspondiente resulta 0,1636, por lo que aceptamos que la distribución binomial negativa ajusta los datos.

Para el modelo binomial negativa-beta, el logaritmo de la función de verosimilitud es  $-10222,2$  y el valor de  $\chi^2$  con 1 grado de libertad es 1,41, que le corresponde un  $p$ -valor de 0,2350. Se acepta también que la distribución binomial negativa beta ajusta los datos. Este último modelo da lugar a un mejor ajuste.

7. Utilizando (4.51) y (4.52) se obtiene la matriz pedida, que aparece representada en las tablas 4.10 y 4.11, para los modelos Poisson-gamma y binomial negativa-beta, respectivamente.

Como se aprecia, las primas son sensiblemente mayores en el grupo de los buenos conductores ( $n\bar{X} = 0$ ) para el modelo binomial negativa-beta y menores para el resto de los grupos.

#### 4.8.2. Penalización de las sobrecargas

El principal objetivo de los sistemas de tarificación bonus-malus es que los asegurados paguen una prima justa, esto es, la prima que corresponda a su propia experiencia de reclamación. Sin embargo, muchos de estos sistemas penalizan injustamente a determinados asegurados, haciéndoles pagar más de lo que realmente les corresponde. Por otro lado, en muchas ocasiones las bonificaciones establecidas son pequeñas, lo que puede acarrear serios problemas de competitividad y, en consecuencia, de equilibrio financiero a la compañía aseguradora (Baione et al., 2002,

Tabla 4.10: Primas bonus-malus. Modelo Poisson-gamma.

$n$	$n\bar{X}$				
	0	1	2	3	4
0	100				
1	88.57	167.80	247.03	326.29	405.50
2	79.49	150.59	221.70	292.81	363.91
3	72.03	136.59	201.08	265.57	330.07

Tabla 4.11: Primas bonus-malus. Modelo binomial negativa-beta.

$n$	$n\bar{X}$				
	0	1	2	3	4
0	100				
1	94.91	130.20	165.48	200.77	236.06
2	90.31	123.89	157.47	191.05	224.63
3	86.14	118.17	150.20	181.22	214.25

pp. 159-160). Además, es costumbre en el mercado de seguro de automóviles, que el asegurado cambie de compañía aseguradora buscando precios más competitivos.

Esto representa un problema para la compañía aseguradora, pues los buenos asegurados pueden optar por abandonar dicha compañía para asegurarse en otra con precios más bajos. Además, para la compañía, la pérdida de los ingresos que le reporta estos clientes supone no poder compensar el balance de la empresa, pues aunque los asegurados catalogados como malos asegurados paguen más, es generalmente una población menos numerosa la que figura en estas clases. Como muestra, obsérvese la cartera de seguros de automóviles que aparece en la tabla 4.12 (extraída de Gómez-Déniz et al., 2005a) y que corresponde a datos reales de una aseguradora española. En esta cartera, los asegurados se clasificaron de acuerdo a dos categorías: edad del conductor ( $q$ ) y la potencia del vehículo asegurado ( $p$ ), como se aprecia en la tabla 4.13. Así,  $q = 1$  corresponde a conductores jóvenes con edad inferior a 35 años;  $q = 2$ , conductores con edades comprendidas entre 35 y 49 años;  $q = 3$ , conductores mayores de 50 años. Por otro lado,  $p = 1$  corresponde a vehículos con una potencia entre 54 y 75 cv;  $p = 2$ , vehículos con una potencia entre 76 y 118; finalmente  $p = 3$ , corresponde a vehículos con una potencia mayor de 119 cv.

Por otro lado, la experiencia demuestra que los buenos asegurados prefieren pagar inicialmente una prima ligeramente mayor, de modo que si incurren en una

Tabla 4.12: Distribución de los asegurados en las distintas subcarteras. Gómez-Déniz et al., 2005a.

$n$	$k$					$k$			
	0	1	2	$\geq 3$		0	1	2	$\geq 3$
0	3945					9023			
1	3316	548	61	20		7797	1063	140	23
2	2756	965	182	42		6668	1952	344	59
3	2290	1226	339	90		5702	2562	625	135
4	1903	1371	503	168		4876	2954	937	257
	11947					25719			
	9470	1916	445	116		21031	3775	720	193
	7273	3391	984	299		16789	6796	1680	455
	5585	4079	1638	645		13402	8444	2903	969
	4289	4265	2234	1159		10699	9151	4108	1761
	8447					19609			
	6570	1423	321	133		15702	3112	603	92
	4908	2495	759	286		12234	5478	1456	441
	3666	2933	1266	582		9532	6647	2496	934
	2738	2990	1701	1017		7426	7032	3469	1681
	1486					5762			
	1125	274	69	18		4554	902	224	82
	822	458	154	52		3452	1640	492	178
	601	523	247	115		2617	1968	817	359
	439	520	322	205		1984	2041	1109	628
	11758					27287			
	10437	1159	143	19		22788	3766	591	142
	9185	2193	331	48		18714	6778	1457	337
	8354	2693	604	107		15369	8594	2591	733
	7114	3517	923	204		12622	9543	3763	1360
	18688					5812			
	15158	2848	510	172		4680	900	187	45
	11988	5065	1262	373		3672	1597	426	117
	9482	6230	2198	778		2881	1952	723	256
	7499	6686	3101	1402		2261	2080	1005	466

reclamación, en el siguiente período (año) su prima no se vea incrementada en una cantidad grande.

Se han elaborado diversos modelos tratando de resolver estos problemas. El primero de estos modelos fue propuesto por Ferreira (1977), que puede verse tam-

Tabla 4.13: Número de asegurados y número medio de reclamaciones en las distintas subcarteras.

	$q = 1$	$q = 2$	$q = 3$
$p = 1$	3945	9023	11758
	0.1850	0.1564	0.1277
$p = 2$	11947	25719	27287
	0.2639	0.2252	0.1969
$p = 3$	8447	19609	18688
	0.2917	0.2495	0.2345
$p = 4$	1486	5762	5812
	0.3135	0.2769	0.2424

bién en Lemaire (1979, 1985 y 1995), en el que se propone un modelo que aumente ligeramente las primas que pagan los buenos asegurados como contrapartida a la disminución de las primas pagadas por los asegurados malos y que garantice el equilibrio financiero. En Gómez-Déniz y León (2005b) y Gómez-Déniz et al. (2005a) se propone un mecanismo similar con la función de pérdida Esscher que realiza el proceso contrario. Ver también Sarabia et al. (2004), haciendo uso de los modelos de especificación condicional.

Las ideas sobre penalización fueron inicialmente propuestas por Lemaire (1979), donde el autor expone un procedimiento para reducir el porcentaje de incremento de la prima que permite mantener el ajuste presupuestario de la compañía aseguradora. Para ello, en este texto se utilizará el principio de prima neta, y minimizaremos la diferencia entre la prima neta Bayes  $P(k, n)$ , siendo  $k$  el número de reclamaciones y  $n$  el tiempo, y el valor del parámetro  $\theta$ , sujeto a la restricción presupuestaria. Así, denominando  $m$  al número máximo de clases en la cartera,  $N_k$  al número de asegurados en dichas clases y  $N = \sum_{k=0}^m N_k$ , formalmente el problema a resolver tendrá la siguiente formulación:

$$\left. \begin{aligned} \text{mín } & \frac{1}{N} \sum_{k=0}^m N_k \int_{\Theta} (P_R - P(k, n))^2 \pi(\theta|k, n) d\theta, \\ \text{s.a } & \frac{1}{N} \sum_{k=0}^m N_k P(k, n) = E(\theta). \end{aligned} \right\} \quad (4.60)$$

Mediante la formulación de este problema el asegurado minimiza la pérdida esperada de la operación de aseguramiento sujeta a la restricción que permite el equilibrio financiero.

La solución a este problema de optimización restringida se presenta en la sigu-

iente proposición.

**Proposición 4.1** *La solución del problema de optimización (4.60) viene dada por:*

$$\begin{aligned} P(k, n) &= \int P_R \pi(\theta|k, n) d\theta + \int P_R \pi(\theta) d\theta \\ &- \frac{1}{N} \sum_{k=0}^m \int P_R \pi(\theta|k, n) d\theta. \end{aligned} \quad (4.61)$$

**Demostración:** La función lagrangiana es:

$$\begin{aligned} \psi &= \frac{1}{N} \sum_{k=0}^m P(k, n) \int (P_R - P(k, n))^2 \pi(\theta|k, n) d\theta \\ &- \beta \left[ \frac{1}{N} \sum_{k=0}^m P(k, n) - \bar{\theta} \right] \end{aligned}$$

en la que  $\beta$  es el correspondiente multiplicador de Lagrange. Ahora,

$$\frac{\partial \psi}{\partial \beta} = 0 \implies \frac{1}{N} \sum_{k=0}^m N_k P(k, n) = \bar{\theta}, \quad (4.62)$$

$$\frac{\partial \psi}{\partial P(k, n)} = 0 \implies P(k, n) = \frac{1}{2} \beta + \int P_R \pi(\theta|k, n), \quad k = 0, 1, \dots, m. \quad (4.63)$$

Tras algunas manipulaciones algebraicas se obtiene:

$$\frac{1}{N} \sum_{k=0}^m N_k P(k, n) = \frac{1}{N} \sum_{k=0}^m N_k \int P_R \pi(\theta) d\theta + \frac{1}{2} \beta.$$

Finalmente, despejando de esta expresión  $\frac{1}{2} \beta$ , sustituyendo en (4.63) y teniendo en cuenta (4.62) se obtiene el resultado. ■

Ahora podemos construir una prima bonus-malus como el cociente:

$$P_{BM} = P_{BM}(k, n) = \frac{P(k, n)}{P(0, 0)}. \quad (4.64)$$

**Ejemplo 4.10** Considérese la cartera de seguros de automóviles que aparece en la tabla 4.12. Calcular las primas netas bonus-malus para la subcartera  $p = 1$ ,  $q = 1$  obtenidas cuando la verosimilitud es Poisson de parámetro  $\theta$  y la distribución a priori es gamma con parámetros  $a = 1$ ,  $b = 5$  para el modelo estándar Poisson-Gamma y el modelo modificado de las sobrecargas. Obsérvese que con estos parámetros se está considerando que la media del número medio de reclamaciones es  $a/b = 0,25$ , lo que parece acorde vista la media de reclamaciones en las distintas subcarteras que aparecen en la tabla 4.13. Vamos a comprobar que se resuelve el problema de las sobrecargas.

Bajo el modelo estándar Poisson-Gamma las primas bonus-malus se calculan de acuerdo a la expresión

$$P_{BM} = \frac{a+k}{b+n} \frac{b}{a},$$

siendo  $k = n\bar{k}$ . Bajo el modelo modificado, utilizando (4.61), resulta, después de algunos cálculos:

$$P(k, n) = \frac{a+k}{b+n} + \frac{a}{b} - \frac{1}{N} \sum_{k=0}^m N_k \frac{a+k}{b+n}. \quad (4.65)$$

Finalmente, las primas netas bonus-malus en el modelo modificado se obtienen llevando (4.65) a (4.64).

Las primas correspondientes (multiplicadas por 100) aparecen en la tabla 4.14, y los resultados del modelo modificado aparecen en negrita. Se observa que las primas, bajo el modelo modificado, aumentan ligeramente para los asegurados en la clase  $k = 0$  y disminuyen para el resto de las clases. En la figura 4.7 aparece el porcentaje de incremento de la prima cuando el asegurado se mueve de la clase  $k = 0$  a la clase  $k = 1$ . Este porcentaje es mucho menor en el modelo modificado que bajo el modelo estándar.

Tabla 4.14: Prima bonus-malus basada en (4.61) en la subcartera  $p = 1$ ,  $q = 1$ .

$n$	$k$			
	0	1	2	$\geq 3$
0	100			
1	83.33 <b>90.32</b>	166.66 <b>135.82</b>	250.00 <b>176.77</b>	333.33 <b>217.42</b>
2	71.42 <b>84.09</b>	142.85 <b>122.41</b>	214.28 <b>157.86</b>	285.71 <b>193.13</b>
3	62.50 <b>79.66</b>	125.00 <b>112.87</b>	187.50 <b>144.07</b>	250.00 <b>175.11</b>
4	55.55 <b>76.31</b>	111.11 <b>105.67</b>	166.67 <b>133.48</b>	222.22 <b>161.16</b>

## 4.9. Análisis Bayesiano de la familia exponencial natural

La familia exponencial constituye una de las principales familias de distribuciones de uso en credibilidad. En las décadas de los 50 y 60 diversos actuarios

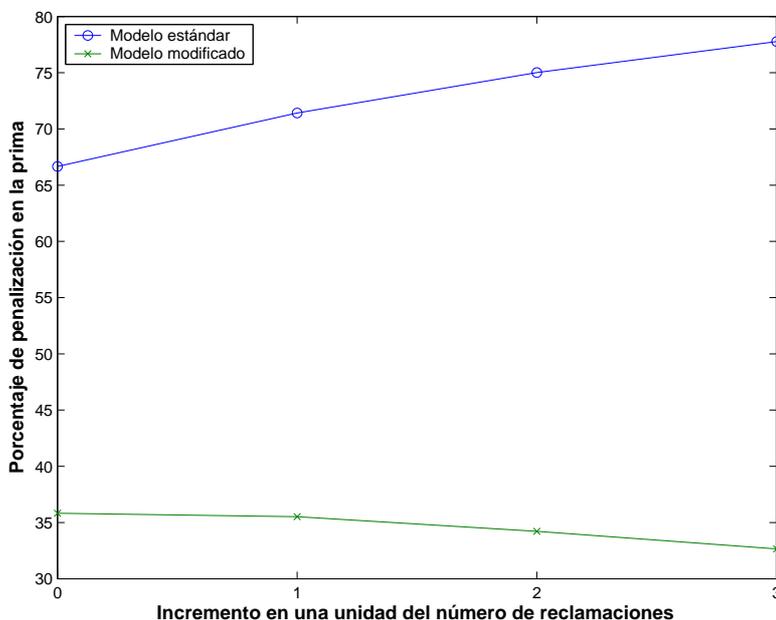


Figura 4.7: Porcentaje de penalización en la prima para los buenos asegurados.

probaron que la fórmula de credibilidad era el estimador Bayes para determinadas combinaciones de verosimilitudes y distribuciones a priori. Este es el caso de la verosimilitud Poisson con distribución a priori gamma o la verosimilitud binomial con distribución a priori tipo beta. Jewell (1974) en un resultado clásico, demostró que estas situaciones no eran más que casos particulares de una verosimilitud más general conocida como familia exponencial. Como veremos en detalle más adelante, para los miembros de esta familia la fórmula de credibilidad es el estimador de Bayes (prima Bayes). Tenemos la siguiente definición.

**Definición 4.13** *Se dice que una variable aleatoria  $X$  (riesgo) pertenece a la familia exponencial natural, si su función de densidad se puede escribir como,*

$$f(x; \theta) = \frac{a(x) \exp(-\theta x)}{c(\theta)}, \quad \theta \in \Theta, \quad (4.66)$$

donde  $\Theta \subset \mathbb{R}$  y  $a(\cdot)$  es una función positiva.

La constante  $c(\theta)$  es la constante de normalización, y viene por tanto determinada por la función  $a(x)$ . Muchas variables aleatorias tanto discretas como continuas

pertenecen a la familia (4.66) y a su extensión

$$f(x; \theta) = \frac{a(x) \exp(-\theta t(x))}{c(\theta)}, \quad \theta \in \Theta, \quad (4.67)$$

donde además es posible que el parámetro  $\theta$  sea vectorial, para tener la llamada familia exponencial  $k$ -paramétrica. Por ejemplo, la distribución exponencial,

$$f(x; \theta) = \theta \exp(-\theta x), \quad x > 0$$

es del tipo (4.66) con  $a(x) = 1$ . Igualmente la distribución de Poisson,

$$f(x; \lambda) = \frac{\exp(-\lambda) \lambda^x}{x!}, \quad x = 0, 1, \dots$$

pertenece a la familia exponencial con  $a(x) = 1/x!$  y  $\theta = \log \lambda$ . Otras distribuciones como la normal con varianza conocida o la distribución de Weibull con parámetro de forma conocido, pertenecen a la familia exponencial (4.67). Para la distribución (4.66) se verifica que,

$$E[X^k] = \frac{c^{(k)}(\theta)}{c(\theta)}, \quad k = 1, 2, \dots \quad (4.68)$$

donde  $c^{(k)}(\theta)$  es la derivada  $k$ -ésima de  $c(\theta)$  respecto  $\theta$ . Para una muestra de tamaño  $n$ , la función de verosimilitud de (4.66) viene dada por,

$$\begin{aligned} L(X|\theta) &= \prod_{i=1}^n \frac{a(x_i) \exp(-\theta x_i)}{c(\theta)} \\ &\propto \frac{\exp(-\theta \sum_{i=1}^n x_i)}{[c(\theta)]^n}. \end{aligned} \quad (4.69)$$

Si consideramos como distribución a priori para  $\theta$  la familia,

$$\pi(\theta) = \frac{[c(\theta)]^{-n_0} e^{-\theta x_0}}{d(n_0, x_0)}, \quad (4.70)$$

se trata de una distribución conjugada para (4.66), donde  $n_0$  y  $x_0$  son parámetros y  $d(n_0, x_0)$  la constante de normalización. Combinando (4.66) con (4.70) obtenemos la distribución a posteriori,

$$\begin{aligned} \pi(\theta|\underline{X}) &\propto L(X|\theta) \cdot \pi(\theta) \\ &\propto \frac{\exp(-\theta \sum_{i=1}^n x_i)}{[c(\theta)]^n} \cdot \frac{[c(\theta)]^{-n_0} e^{-\theta x_0}}{d(n_0, x_0)} \\ &= [c(\theta)]^{-(n_0+n)} \exp \left[ -\theta \left( x_0 + \sum_{i=1}^n x_i \right) \right], \end{aligned}$$

que pertenece a la familia (4.70), y donde los parámetros se actualizan según las reglas:

$$\begin{aligned} n_0 &\rightarrow n_0 + n, \\ x_0 &\rightarrow x_0 + \sum_{i=1}^n x_i. \end{aligned}$$

## 4.10. Análisis Bayesiano de la familia exponencial de dispersión

La motivación de porqué utilizar la siguiente familia de distribuciones procede directamente de la investigación actuarial en credibilidad. La fórmula clásica de credibilidad (Bayesiana) establece que,

$$E[X_{n+1}|X] = E[\mu(\theta)|X] = \frac{n}{n_0 + n} \cdot \bar{X} + \frac{n_0}{n_0 + n} \cdot m, \quad (4.71)$$

donde  $X$  representa los datos disponibles,  $\bar{X}$  la media muestral,  $\theta$  el parámetro de riesgo,  $\mu(\theta) = E[X|\theta]$ ,  $m = \int \mu(\theta)d\pi(\theta)$  y  $\pi(\theta)$  la distribución a priori. Esta fórmula se cumple para la familia exponencial vista en la sección anterior. Sin embargo, la familia exponencial no contiene algunas de las distribuciones habitualmente usadas en riesgos. Entonces, ¿es posible extender el resultado (4.71) a otras clases de distribuciones? La respuesta es que sí es posible, y Kass et al. (1997), Nelder y Verral (1997) y Landsman y Makov (1998) probaron que este fenómeno se cumplía para una clase más general de distribuciones denominada *familia exponencial de dispersión*. Esta familia ya era conocida en el mundo estadístico como parte de los llamados *modelos lineales generalizados*, introducidos por los estadísticos ingleses Nelder y Wedderburn (1972) y también estudiados por Tweedie (1984) y Jorgensen (1986). Tenemos la siguiente definición

**Definición 4.14** *Se dice que una variable aleatoria  $X$  pertenece a la familia exponencial de dispersión si su función de densidad se puede escribir como,*

$$f(x; \theta, \lambda) = \frac{p(\lambda, x) \exp(-\lambda \theta x)}{[q(\theta)]^\lambda}, \quad (4.72)$$

donde  $\theta \in \Theta \subset \mathbb{R}$  y  $\lambda \in \Lambda \subset \mathbb{R}^+$  y  $p(\cdot, \cdot)$  es una función positiva.

La media de (4.72) viene dada por,

$$E[X] = \mu(\theta) = -\frac{q'(\theta)}{q(\theta)} = -\frac{d}{d\theta} \log q(\theta), \quad (4.73)$$

al igual que la familia exponencial estándar. La varianza viene dada por,

$$\text{Var}[X] = \frac{v(\theta)}{\lambda} = -\frac{\mu'(\theta)}{\lambda} = \frac{1}{\lambda} \frac{d^2}{d\theta^2} \log q(\theta). \quad (4.74)$$

Por otro lado si  $\lambda$  es un parámetro conocido, (4.72) es una familia de tipo exponencial clásica.

La distribución inversa Gaussiana con función de densidad:

$$f(x; \theta, \mu) = \left( \frac{\theta}{2\pi x^3} \right)^{1/2} \exp \left\{ -\frac{\theta(x - \mu)^2}{2x\mu^2} \right\},$$

pertenece a la familia exponencial de dispersión donde:

$$\begin{aligned} p(\lambda, x) &= \left( \frac{\lambda}{2\pi x^3} \right)^{1/2} \exp \left( \frac{\lambda}{2x} \right) \\ q(\theta) &= \exp[-(2\theta)^{1/2}], \end{aligned}$$

y por tanto es de la forma de una exponencial de dispersión.

Para el análisis Bayesiano, consideramos la distribución a priori,

$$\pi(\theta) = \frac{[q(\theta)]^{-n_0} \exp(-\theta\mu n_0)}{c(\mu, n_0)}, \quad \theta_0 < \theta < \theta_1, \quad (4.75)$$

donde  $\pi(\theta_0) = \pi(\theta_1)$ . Suponiendo que  $\lambda$  es un parámetro conocido, la función de verosimilitud viene dada por,

$$\begin{aligned} L(X|\theta) &= \prod_{i=1}^n \frac{p(\lambda, x_i) \exp(-\lambda\theta x_i)}{[q(\theta)]^\lambda} \\ &\propto \frac{\exp(-\lambda\theta \sum_{i=1}^n x_i)}{[q(\theta)]^{\lambda n}}. \end{aligned} \quad (4.76)$$

Combinando (4.75) con (4.76) obtenemos la distribución a posteriori,

$$\begin{aligned} \pi(\theta|\underline{X}) &\propto L(X|\theta) \cdot \pi(\theta) \\ &\propto \frac{\exp(-\lambda\theta \sum_{i=1}^n x_i)}{[q(\theta)]^{\lambda n}} \cdot \frac{[q(\theta)]^{-n_0} \exp(-\theta\mu n_0)}{c(\mu, n_0)} \\ &= [q(\theta)]^{-n_0 - \lambda n} \exp \left[ -\theta \left( \lambda \sum_{i=1}^n x_i + \mu n_0 \right) \right], \end{aligned}$$

que pertenece a la misma clase de distribuciones (4.75), y por tanto se trata de una clase conjugada. Los parámetros de la distribución se actualizan de acuerdo con la Tabla 4.15.

Se verifica el siguiente resultado:

**Teorema 4.5** *En la familia exponencial de dispersión y usando la distribución a priori (4.75), la prima Bayes viene dada por,*

$$E[X_{n+1}|X] = \frac{\lambda \sum_{i=1}^n X_i + \mu n_0}{\lambda n + n_0} = \frac{\lambda n}{\lambda n + n_0} \cdot \bar{X} + \frac{n_0}{\lambda n + n_0} \cdot \mu, \quad (4.77)$$

Tabla 4.15: Actualización de parámetros en la familia exponencial de dispersión.

Parámetros	Valores a priori	Valores actualizados
$n_0$	$n_0$	$\tilde{n}_0 = n_0 + \lambda n$
$\mu$	$\mu$	$\tilde{\mu} = \frac{\lambda \sum_{i=1}^n X_i + \mu n_0}{\lambda n + n_0}$

que es una fórmula de credibilidad exacta. Además, la prima de Bühlmann coincide con la prima Bayes.

**Demostración:** Se verifica que,

$$E[X_{n+1}|X] = \int_{\theta_0}^{\theta_1} \mu(\theta) \frac{[q(\theta)]^{-\tilde{n}_0} \exp(-\theta \tilde{\mu} \tilde{n}_0)}{c(\tilde{\mu}, \tilde{n}_0)} d\theta,$$

donde  $\mu(\theta)$  viene dada en (4.73). Mediante integración por partes, se obtiene fácilmente (5.22). La prima de Bühlmann coincide con la prima Bayes, puesto que ésta última es lineal en las observaciones. ■

## Capítulo 5

# Modelos de Credibilidad

### 5.1. Introducción

Este capítulo está dedicado a los diferentes modelos de credibilidad. Como ya se comentó detalladamente en el capítulo 1, la teoría de la credibilidad constituye uno de los campos más atractivos y más desarrollados de la estadística actuarial, y dada su utilidad práctica, es posible encontrarse con modelos de credibilidad en otros terrenos científicos, como así ocurre en riesgos operacionales.

Bajo esta teoría, y en el marco actuarial, se propone que la prima que se debe pagar por parte del asegurado incluya tanto la experiencia individual del propio asegurado como la del colectivo al que pertenece, la cartera de seguros. Así, el estimador de la prima se calcularía como una combinación lineal convexa de la prima del colectivo y de la experiencia del asegurado, escrito en términos de una expresión que recibe el nombre de fórmula de credibilidad. Cuando ocurre esto, se dice entonces que la fórmula de credibilidad es exacta. Sin embargo, hoy en día se habla de credibilidad en términos mucho más generales, incluyéndose en esta teoría aquéllos estimadores de la primas que para su cálculo, haga uso de la experiencia individual y de cualquier información colateral disponible.

La incorporación de estas dos fuentes de información, la muestral y la colateral, parece predisponer el asunto a ser estudiado bajo la óptica Bayesiana. De esta manera, y a partir de la utilización de esta metodología, es cuando la teoría de la credibilidad ha tenido un despegue que hace de ella una de las disciplinas más estudiadas de la ciencia actuarial.

El capítulo está estructurado de acuerdo a los siguientes contenidos. La sección 5.2 está dedicada a los modelos clásicos de credibilidad desarrollados desde comienzos del siglo XX hasta finales de los años 60. En la sección 5.3 se estudian los modelos de Bühlmann (1967) y de Bühlmann-Straub (1972), que constituyen el punto de partida de la moderna teoría de la credibilidad. La sección 5.4 incluye

los modelos Bayesianos desarrollados en teoría de la credibilidad. Entre ellos estudiamos, el modelo Bayesiano estándar, el modelo jerárquico y el modelo de Jewell (1974), que constituye una extensión natural del modelo de Bühlmann (1967) cuando se considera la familia exponencial de distribuciones. Un caso más general de éste lo constituye el modelo de Landsman y Makov (1998) que se analiza a continuación. Posteriormente se muestra el caso más general, que incluye a todos los modelos anteriores, desarrollado en Gómez-Déniz (2008a) y que se obtiene utilizando una función de pérdida que incluye como caso particular a la pérdida cuadrática ponderada. El capítulo finaliza exponiendo algunos modelos de credibilidad desarrollados también bajo una metodología Bayesiana no estándar. Este análisis desarrolla modelos de credibilidad cuando se admite que el investigador no conoce exactamente la distribución a priori del parámetro de riesgo.

## 5.2. Credibilidad total y parcial

Desarrollamos en esta sección dos modelos de credibilidad que fueron el origen de la actual teoría. Se trata de los modelos de credibilidad total y parcial.

### 5.2.1. Credibilidad total

Parece lógico que los asegurados con una experiencia de reclamación que les sea favorable quieran que la prima que tengan que pagar esté basada únicamente en su propia experiencia de siniestralidad, es decir que la aseguradora le asigna a ésta un 100 % de credibilidad. Sin embargo, desde el punto de vista de la aseguradora, esto sólo será posible si la experiencia de reclamación es estable. Una manera de resolver este problema es suponer que  $\bar{X}$  es estable si existe una probabilidad alta de que la diferencia entre  $\bar{X}$  y  $\varepsilon$  sea pequeña. En términos probabilísticos esto requeriría admitir credibilidad total si se verifica:

$$\Pr(|\bar{X} - \varepsilon| \leq c\varepsilon) = \Pr((1 - c)\varepsilon \leq \bar{X} \leq (1 + c)\varepsilon) \geq p, \quad (5.1)$$

siendo  $0 < p < 1$  y  $c > 0$ . En la práctica lo razonable es elegir un valor de  $p$  cercano a 1 y un valor de  $c$  cercano a 0. Normalmente suelen considerarse 0,9 y 0,05 para  $p$  y  $c$ , respectivamente.

Reescribimos (5.1) en la forma:

$$\Pr\left(\left|\frac{\bar{X} - \varepsilon}{\sigma/\sqrt{n}}\right| \leq \frac{c\varepsilon\sqrt{n}}{\sigma}\right) \geq p,$$

y definimos ahora  $x_p$  como

$$x_p = \inf_x \left\{ \Pr\left(\left|\frac{\bar{X} - \varepsilon}{\sigma/\sqrt{n}}\right| \leq x\right) \geq p \right\}.$$

Suponiendo que  $\bar{X}$  sigue una distribución de tipo continua, esta última expresión es equivalente a:

$$\Pr\left(\left|\frac{\bar{X} - \varepsilon}{\sigma/\sqrt{n}}\right| \leq x_p\right) = p. \quad (5.2)$$

Por tanto, la condición que ha de verificarse para suponer credibilidad total es

$$\frac{c\varepsilon\sqrt{n}}{\sigma} \geq x_p,$$

o de forma equivalente:

$$\frac{\sigma}{\varepsilon} \leq \frac{c}{x_p} \sqrt{n} = \sqrt{\frac{n}{\lambda_0}}, \quad (5.3)$$

siendo  $\lambda_0 = (x_p/c)^2$ . La expresión (5.3) puede interpretarse en el sentido siguiente: se supone credibilidad total si el coeficiente de variación es menor o igual que  $\sqrt{n/\lambda_0}$ . También, a partir de (5.3) se observa que se supone credibilidad total si:

$$\text{Var}[\bar{X}] = \frac{\sigma^2}{n} \leq \frac{\varepsilon^2}{\lambda_0},$$

y, por otro lado, el valor que ha de tomar  $n$  para suponer credibilidad total ha de cumplir:

$$n \geq \lambda_0 \left(\frac{\sigma}{\varepsilon}\right)^2$$

En la práctica, si la experiencia del asegurado es suficientemente grande, de acuerdo al teorema central del límite, entonces la variable aleatoria  $(\bar{X} - x)/(\sigma\sqrt{n})$  sigue aproximadamente una distribución normal con media cero y desviación típica 1. Así (5.2) puede entonces escribirse como  $p = 2\Phi(x_p) - 1$ , donde  $\Phi(x)$  es la función de distribución de la normal tipificada, luego  $x_p$  es el  $(1+p)/2$  percentil de la distribución normal tipificada.

En la tabla 5.1 se incluye el valor de  $\lambda_0$  correspondiente a distintas combinaciones de  $c$  y  $p$ .

**Ejemplo 5.1** Supongamos que se cuenta con la experiencia  $X_j$ ,  $j = 1, \dots, n$  de un contrato de seguro perteneciente a una cartera de seguros y que  $X_1, \dots, X_n$  son variables aleatorias independientes e idénticamente distribuidas tipo Poisson de parámetro  $\theta = 200$ . Vamos a obtener el valor más pequeño de  $n$  para suponer credibilidad total a la experiencia observada, suponiendo  $c = 0,04$ ,  $p = 0,95$  y que la aseguradora tarifica atendiendo sólo al número de reclamaciones.

En este caso  $\varepsilon = \sigma^2 = 200$ . Dados los valores de  $c$  y  $p$  resulta  $x_p = 1,96$ , luego:

$$\lambda_0 = \left(\frac{x_p}{c}\right)^2 = 2401,$$

Tabla 5.1: Valor de  $\lambda_0$  correspondientes a distintas combinaciones de  $c$  y  $p$ . Fuente, Hossack et al. (1999).

$p/c$	0.3	0.2	0.1	0.05	0.01
0.9	30	68	271	1082	27060
0.95	43	96	384	1537	38416
0.99	74	166	663	2654	66358
0.999	120	271	1083	4331	108274

y

$$n \geq 2401 \frac{200}{200^2} = 12,005.$$

**Ejemplo 5.2** La tabla 5.2, extraída de Hossack et al. (1999, 2001), recoge la cantidad reclamada y el número de reclamaciones en una cartera de seguros. Vamos a obtener el valor más pequeño de asegurados en la cartera para dar credibilidad total a la experiencia observada, suponiendo que  $c = 0,05$  y  $p = 0,95$ .

Tabla 5.2: Distribución de la cantidad reclamada. Fuente, Hossack et al. (1999)

Cantidad reclamada	Número de reclamaciones
0 – 2000	488
2000 – 4000	115
4000 – 6000	92
6000 – 8000	54
8000 – 10000	33
10000 – 12000	19
12000 – 14000	15
14000 – 16000	15
16000 – 18000	7
18000 – 20000	4
Total	842

La media y la varianza de la distribución dada en la tabla 5.2 es:

$$\begin{aligned} \varepsilon &= \frac{1}{842}(1000 \cdot 488 + \dots + 19000 \cdot 20000) = 3351,91, \\ \sigma^2 &= \frac{1}{842}(1000^2 \cdot 488 + \dots + 19000^2 \cdot 20000) - \varepsilon^2 = 13729176,66. \end{aligned}$$

En consecuencia, el coeficiente de variación es  $\sigma/\varepsilon = 1,117$ . De la tabla 5.1 se desprende que  $\lambda_0 = 1537$ , por lo tanto  $n = 1537 \cdot 1,117^2 = 1918$ . Teniendo en cuenta que la media del número de reclamaciones es aproximadamente 0.0012, concluimos que se requerirán de  $1918/0,0012 = 1,598,334$  asegurados para admitir que la experiencia obtenida es creíble al cien por cien.

### 5.2.2. Credibilidad parcial

Para muchos asegurados la experiencia de siniestralidad es insuficiente para suponer credibilidad total, es decir que el factor  $Z(n)$  sea igual a 1. Ahora se supone que la prima a cargar sea una combinación lineal entre la experiencia del asegurado y la experiencia del colectivo o prima manual, de modo que:

$$P = Z(n)\bar{X} + [1 - Z(n)]M,$$

y habrá, por tanto, que determinar el valor de  $Z(n)$  para obtener la prima. Dado que:

$$\text{Var}[P] = \text{Var} [Z(n)\bar{X} + [1 - Z(n)]M] = Z(n)^2 \text{Var}(\bar{X}) = Z(n)^2 \frac{\sigma^2}{n},$$

igualando este último término a  $\varepsilon^2/\lambda_0$  resulta,

$$Z(n) = (\varepsilon/\sigma)\sqrt{n/\lambda_0},$$

de modo que se elige  $Z(n)$  de acuerdo a la expresión:

$$Z(n) = \min \left\{ \frac{\varepsilon}{\sigma} \sqrt{\frac{n}{\lambda_0}}, 1 \right\}. \quad (5.4)$$

**Ejemplo 5.3** Con los datos del ejemplo 5.1 vamos a calcular el valor del factor de credibilidad para una experiencia de reclamaciones correspondiente a 10 años.

Haciendo uso de (5.4) resulta:

$$Z(10) = \min \left\{ \sqrt{\frac{10 \times 200}{2401}}, 1 \right\} = 0,9126.$$

**Ejemplo 5.4** Con los datos del ejemplo 5.1 vamos a calcular el factor de credibilidad cuando se han observado 850 reclamaciones para ese grupo de asegurados.

En este caso resulta conveniente estimar  $n\theta$  por  $N_1 + \dots + N_n$ , de manera que:

$$Z(n) = \min \left\{ \sqrt{\frac{850}{2401}}, 1 \right\} = 0,595.$$

### 5.3. Modelos de distribución libre

Dentro de los modelos de credibilidad clásicos más importantes figuran el modelo de Bühlmann de distribución libre y el modelo de Bühlmann-Straub. Ambos modelos constituyen el punto de partida de la moderna teoría de la credibilidad.

El objetivo de ambos modelos es estimar la prima correspondiente a un asegurado o grupo de asegurados que conforman una póliza en una cartera de seguros, restringiéndose a las primas lineales y utilizando el método de los mínimos cuadrados. La diferencia fundamental entre ambos modelos radica en que el segundo admite observaciones ponderadas.

Lo relevante de ambos modelos es la no necesidad de establecer hipótesis alguna, ni sobre la distribución que gobierna los riesgos individuales, ni sobre la distribución a priori de los parámetros de riesgo, de ahí el nombre de modelos de distribución libre. Para su descripción consideremos la cartera de seguros que aparece en la tabla 5.3. Los elementos de la tabla son los siguientes:

1.  $X_{ij}$  es una variable aleatoria que representa el riesgo (cantidad reclamada, número de reclamaciones, etc.) del asegurado  $i$  en el año  $j$ .
2. La tabla completa describe la evolución de  $k$  asegurados o contratos de seguros (pólizas) con características comunes durante  $n$  años.
3. Las columnas, por tanto representan variables aleatorias pertenecientes al mismo riesgo.
4. Las filas representan variables aleatorias pertenecientes al mismo año.
5. La función de distribución de la variable aleatoria  $X_{ij}$  depende de un parámetro desconocido  $\theta_{ij}$ . Es usual denominarlo parámetro de riesgo.

Tabla 5.3: Datos de la cartera de seguros.

Años	Pólizas o grupos					
	1	2	...	$j$	...	$k$
1	$X_{11}$	$X_{21}$	...	$X_{j1}$	...	$X_{k1}$
2	$X_{12}$	$X_{22}$	...	$X_{j2}$	...	$X_{k2}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$n$	$X_{1n}$	$X_{2n}$	...	$X_{jn}$	...	$X_{kn}$

y establecemos las siguientes hipótesis:

1. Homogeneidad en el tiempo, esto es  $\theta_{ij} = \theta_j$ . Esto se traduce en que todas las variables aleatorias pertenecientes a la misma columna tienen la misma distribución de probabilidad.
2. Los parámetros  $\theta_j$  son variables aleatorias independientes y con la misma distribución  $\pi(\theta)$ .
3. Independencia de riesgos. Esto es, las columnas son independientes entre sí.
4. Dado el valor de  $\theta$  particular  $\theta_j$  las variables aleatorias dentro de la columna  $j$  son independientes.

En la tabla,  $\theta_j$  representa el parámetro de riesgo para la póliza  $j$ -ésima. Se trata de una variable estructural que describe las características de riesgo del contrato  $j$ -ésimo. En la ciencia actuarial es costumbre considerar a dicho parámetro desconocido y aleatorio.

La variable  $X_{ij}$  representa la experiencia de reclamaciones para la póliza  $i$ -ésima en el período  $j$ -ésimo. Se trata de una variable aleatoria con realizaciones observables.

Denotaremos mediante

$$\mu(\theta_j) = E[X_{ij}|\theta_j],$$

a la prima de riesgo para la póliza  $j$ ,

$$m = E[P(\theta_j)],$$

al valor esperado de todas las primas de riesgo, es decir, la prima colectiva. Finalmente,

$$a = \text{Var}[\mu(\theta_j)],$$

es la varianza de las primas de riesgo, que es un indicador de la heterogeneidad de la cartera.

### 5.3.1. Modelo de Bühlmann

El objetivo del modelo de Bühlmann consiste en calcular la mejor prima lineal

$$H[\mu(\theta_j)|X_{j1}, X_{j2}, \dots, X_{jn}],$$

dependiente de los datos observados, mediante el método de los mínimos cuadrados. Para ello establecemos la siguiente notación previa, en la que prescindiremos del subíndice  $j$ :

- $\mu(\theta) = E[X|\theta]$ : Prima de riesgo individual
- $m = E_{\text{Total}}[X] = E[\mu(\theta)]$ : Prima de riesgo colectiva. Valor esperado de todas las primas de riesgo individuales.

- $a = \text{Var} [E(X|\theta)] = \text{Var} [\mu(\theta)]$ : Varianza de las primas de riesgo individuales, indicador de la heterogeneidad de la cartera.
- $s^2 = E[\text{Var}(X|\theta)]$ : Medida global de la dispersión de la siniestralidad individual.

Se supone también que  $X_1|\theta, X_2|\theta, \dots, X_n|\theta$  son variables aleatorias idénticamente distribuidas con media y varianza comunes  $\mu(\theta)$  y  $\sigma^2(\theta)$ , respectivamente. Antes de comenzar con el modelo de Bühlmann necesitamos el siguiente resultado.

**Proposición 5.1** Si  $X$  e  $Y$  son variables aleatorias con distribución conjunta dependiente de la variable aleatoria  $\Theta$  se verifica:

$$\begin{aligned} E(X) &= E_{\Theta} [E_X(X|\Theta)], \\ \text{Cov}(X, Y) &= E[\text{Cov}(X, Y|\Theta)] + \text{Cov}[E(X|\Theta), E(Y|\Theta)]. \end{aligned} \quad (5.5)$$

**Demostración:** La primera relación es bien conocida. En cuanto a la segunda tenemos:

$$\begin{aligned} \text{Cov}(X, Y) &= E\{[X - E(X)][Y - E(Y)]\} \\ &= E\{[X - E(X|\Theta) + E(X|\Theta) - E(X)] \\ &\quad \times [Y - E(Y|\Theta) + E(Y|\Theta) - E(Y)]\} \\ &= E_{\Theta} E_{X, Y|\Theta} \{[X - E(X|\Theta)][Y - E(Y|\Theta)]\} \\ &\quad + E_{\Theta} E_{X, Y|\Theta} \{[X - E(X|\Theta)][E(Y|\Theta) - E(Y)]\} \\ &\quad + E_{\Theta} E_{X, Y|\Theta} \{[E(X|\Theta) - E(X)][Y - E(Y|\Theta)]\} \\ &\quad + E_{\Theta} E_{X, Y|\Theta} \{[E(X|\Theta) - E(X)][E(Y|\Theta) - E(Y)]\} \\ &= E[\text{Cov}(X, Y|\Theta)] + 0 + 0 + \text{Cov}[E(X|\Theta), E(Y|\Theta)]. \end{aligned}$$

■

Obsérvese que de (5.5) se deduce:

$$\text{Var}[X] = E[\text{Var}(X|\theta)] + \text{Var}[E(X|\theta)], \quad (5.6)$$

haciendo  $X = Y$ .

Vamos a probar el resultado clásico de Bühlmann(1967) .

**Teorema 5.1** La mejor aproximación lineal a  $H[\mu(\theta)|X_1, \dots, X_n]$  es:

$$a + b\bar{X} = a + b\frac{1}{n} \sum_{i=1}^n X_i,$$

donde:

$$\begin{aligned} a &= (1 - b)m, \\ b &= \frac{n}{n + k}, \\ k &= \frac{E[\sigma^2(\theta)]}{\text{Var}[\mu(\theta)]}. \end{aligned}$$

**Demstración:** Queremos encontrar la mejor estimación de la prima neta de riesgo que dependa linealmente de los datos observados, esto es:

$$H[\mu(\theta)|X_1, \dots, X_n] = c_0 + \sum_{s=1}^n c_s X_s.$$

Para ello hagamos mínima la esperanza del cuadrado de la desviación de la prima de riesgo individual respecto a  $H[\mu(\theta)|X_1, \dots, X_n]$ . Esto es:

$$\min_{c_i} E \left[ \left( \mu(\theta) - c_0 - \sum_{s=1}^n c_s X_s \right)^2 \right].$$

Calculando las correspondientes derivadas parciales e igualando a cero se obtiene el sistema de ecuaciones:

$$\left. \begin{aligned} E \left[ \mu(\theta) - c_0 - \sum_{s=1}^n c_s X_s \right] &= 0, \\ E \left[ X_r \left( \mu(\theta) - c_0 - \sum_{s=1}^n c_s X_s \right) \right] &= 0, \quad r = 1, 2, \dots, n. \end{aligned} \right\}$$

Multiplicando la primera ecuación por  $E[X_r]$  y restándosela a la segunda obtenemos:

$$\text{Cov}[\mu(\theta), X_r] = \sum_{s=1}^n c_s \text{Cov}(X_r, X_s), \quad r = 1, 2, \dots, n. \quad (5.7)$$

Teniendo en cuenta ahora que

$$\begin{aligned} \text{Cov}(X_r, X_s) &= E[\text{Cov}(X_r, X_s|\theta)] + \text{Cov}[E[X_r|\theta], E[X_s|\theta]] = s^2 + a, \quad r \neq s \\ \text{Cov}(X_r, X_s) &= \text{Cov}[\mu(\theta), \mu(\theta)] = \text{Var}[\mu(\theta)] = a, \quad r = s, \end{aligned}$$

el sistema (5.7) puede reescribirse como:

$$\left. \begin{aligned} s^2 c_r + \sum_{s=1}^n c_s a &= a, \\ c_0 &= m - m \sum_{s=1}^n c_s, \end{aligned} \right\}$$

donde se ha tenido en cuenta que  $E[\mu(\theta) - c_0 - \sum_{s=1}^n c_s X_s] = 0$ , y por tanto:

$$c_0 = E[\mu(\theta)] - \sum_{s=1}^n c_s E[X_s] = m - m \sum_{s=1}^n c_s.$$

Debido a la simetría del sistema resulta que  $c_1 = \dots = c_n$ , luego:

$$\left. \begin{aligned} s^2 c + anc &= a, \\ c_0 + mnc &= m. \end{aligned} \right\}$$

Finalmente se deduce:

$$\left. \begin{aligned} c &= \frac{a}{s^2 + an}, \\ c_0 &= m(1 - nc) = m \left( 1 - \frac{na}{s^2 + an} \right) = m \frac{s^2}{s^2 + an}. \end{aligned} \right\}$$

Por tanto,

$$\begin{aligned} H[\mu(\theta)|X_1, \dots, X_n] &= c_0 + \sum_{s=1}^n c_s X_s = m \frac{s^2}{s^2 + an} + cn\bar{X} \\ &= m \frac{s^2}{s^2 + an} + \frac{an}{s^2 + an} \bar{X} = [1 - Z(n)]m + Z(n)\bar{X}, \end{aligned}$$

con:

$$Z(n) = \frac{an}{an + s^2} = \frac{n \text{Var}[\mu(\theta)]}{n \text{Var}[\mu(\theta)] + E[\sigma^2(\theta)]}. \quad (5.8)$$

Obsérvese que el resultado no depende de la distribución de probabilidad de  $X$  ni de la distribución de probabilidad de  $\theta$ , de ahí el término de *distribución libre*. Las cantidades

$$\begin{aligned} a &= \text{Var}[\mu(\theta)], \\ s^2 &= E[\sigma^2(\theta)], \\ m &= \bar{X}, \end{aligned}$$

suelen llamarse parámetros estructurales del modelo y pueden estimarse a partir de:

$$\hat{m} = \frac{1}{k} \sum_{j=1}^k \bar{X}_j = \frac{1}{k} \sum_{j=1}^k \sum_{s=1}^n \frac{X_{js}}{n}, \quad (5.9)$$

$$\hat{s}^2 = \frac{1}{k} \sum_{j=1}^k \hat{s}_j^2, \quad \hat{s}_j^2 = \frac{1}{n-1} \sum_{s=1}^n (X_{js} - \bar{X}_j)^2, \quad (5.10)$$

$$\hat{a} = \frac{1}{k-1} \sum_{j=1}^k (\bar{X}_j - \bar{X})^2 - \frac{1}{n} \hat{s}^2. \quad (5.11)$$

Finalmente, puede probarse que estos estimadores son insesgados y consistentes, es decir:

$$\begin{aligned} E(\hat{m}) &= m, \\ E(\hat{s})^2 &= s^2, \\ E(\hat{\alpha}) &= \alpha, \\ (\hat{m}, \hat{s}, \hat{\alpha}) &\longrightarrow (m, s^2, \alpha), \text{ cuando } n \rightarrow \infty. \end{aligned}$$

Resulta interesante destacar las siguientes propiedades del factor de credibilidad. Se verifica que:

- $Z(n)$  es una función creciente en  $n$ , de modo que  $Z(n) \rightarrow 1$  cuando  $n \rightarrow \infty$ , mientras que  $Z(n) \rightarrow 0$  cuando  $n \rightarrow 0$ . Por tanto  $n = 0$  supone que no se dispone de experiencia para el asegurado (se trata de un contrato nuevo), y la prima a cobrar en este caso es simplemente la prima colectiva. En la medida en que aumenta  $n$ , y por tanto se dispone de más datos, la información individual tiene mayor peso.
- $Z(n)$  es también una función creciente de la varianza de las medias teóricas,  $a = \text{Var}[E(X|\theta)]$ , con límite 1 cuando aquélla tiende a infinito y cero cuando tiende a cero. Esto es lógico, pues si la cartera no es heterogénea,  $a = 0$ , entonces la prima colectiva es el mejor estimador de la prima individual, mientras que una mayor heterogeneidad de la cartera debe suponer dar mayor peso a la información individual.
- $Z(n)$  es una función decreciente respecto al valor esperado de la varianza teórica,  $s^2 = E[\text{Var}(X|\theta)]$ , de modo que cuanto mayor sea la varianza del individuo menor peso se da a su experiencia individual y mayor a la del colectivo.

**Ejemplo 5.5** El número de reclamaciones,  $X_i$ , durante el año  $i$ -ésimo ( $i = 1, 2$ ) de vigencia de una póliza de seguros sigue una distribución geométrica de parámetro  $\theta$ ,  $0 < \theta < 1$ . La distribución a priori de este parámetro es la distribución beta,  $Be(2, 5)$ . Si el asegurado presenta dos reclamaciones durante el primer año de vigencia de la póliza, vamos a utilizar el estimador de Bühlmann para estimar el número de reclamaciones para el segundo año.

El número de reclamaciones para el segundo año de la póliza vendrá dado por:

$$Z(n)\bar{X} + [1 - Z(n)]E[E(X|\theta)],$$

donde  $\bar{X} = 2$ ,  $n = 1$  y  $Z(n)$  viene dado por (5.8). Se precisa de los siguientes cálculos:

$$E[X|\theta] = \frac{\theta}{1 - \theta},$$

$$E[E(X|\theta)] = E\left(\frac{\theta}{1-\theta}\right) = \frac{1}{B(2,5)} \int_0^1 \frac{\theta}{1-\theta} \theta(1-\theta)^4 d\theta = \frac{B(3,4)}{B(2,5)} = 0,5,$$

$$\text{Var}[X|\theta] = \frac{\theta}{(1-\theta)^2},$$

donde  $B(x, y)$  representa la función beta. Por otro lado,

$$\text{Var}[E(X|\theta)] = E\left[\frac{\theta}{1-\theta}\right]^2 - E^2\left[\frac{\theta}{1-\theta}\right],$$

donde

$$E\left[\frac{\theta}{1-\theta}\right]^2 = \frac{1}{B(2,5)} \int_0^1 \frac{\theta^2}{(1-\theta)^2} \theta(1-\theta)^4 d\theta = \frac{B(4,3)}{B(2,5)} = 0,5,$$

y procediendo de forma similar tenemos,

$$E^2\left[\frac{\theta}{1-\theta}\right] = \left[\frac{B(3,4)}{B(2,5)}\right]^2 = 0,25.$$

Además,

$$E[\text{Var}(X|\theta)] = E\left[\frac{\theta}{(1-\theta)^2}\right] = \frac{B(3,3)}{B(2,5)} = 1.$$

Por tanto, el factor de credibilidad es:

$$Z(2) = \frac{1 \cdot 0,4375}{1 \cdot 0,4375 + 1} = 0,3043.$$

Finalmente el estimador de Bühlmann del número de reclamaciones para el segundo año viene dado por:

$$0,3043 \cdot 2 + (1 - 0,3043) \cdot 0,5 = 0,9565.$$

**Ejemplo 5.6** Una compañía de seguros Canadiense tiene dividido en cuatro territorios las zonas en las que operan los vehículos industriales que tiene asegurados. Se dispone de información sobre el porcentaje de accidentes en cada uno de los territorios, que se muestran en la tabla 5.4. Vamos a estimar, utilizando el modelo de Bühlmann, la prima neta basada en el porcentaje de accidentes en cada uno de los territorios para el año 1997.

El estimador del factor de credibilidad es:

$$\hat{Z}(n) = \frac{\hat{a}n}{\hat{s}^2 + \hat{a}n}.$$

Tabla 5.4: Porcentaje de accidentes por territorio (Herzog, 1996).

Territorios	Años		
	1994	1995	1996
1	6.23 %	7.14 %	9.84 %
2	4.55 %	4.98 %	6.69 %
3	6.13 %	5.38 %	6.19 %
4	7.91 %	6.67 %	6.59 %

Utilizando (5.9), (5.10) y (5.11) obtenemos:

$$\hat{m}_1 = 7,737, \quad \hat{m}_2 = 5,406, \quad \hat{m}_3 = 5,900, \quad \hat{m}_4 = 7,057, \quad \hat{m} = \bar{x} = 6,525,$$

$$\hat{s}_1^2 = 3,525, \quad \hat{s}_2^2 = 1,281, \quad \hat{s}_3^2 = 0,2037, \quad \hat{s}_4^2 = 0,5477, \quad \hat{s}^2 = 1,38935,$$

$$\hat{a} = 0,6684, \quad \hat{Z}(t) = 0,59071.$$

El estimador de la prima neta para los cuatro territorios es:

$$\text{Territorio 1: } \hat{Z}(n)\bar{x}_1 + [1 - \hat{Z}(n)]\hat{m} = 0,59071 \times 0,4093 \times 6,525 = 7,240,$$

$$\text{Territorio 2: } 5,863,$$

$$\text{Territorio 3: } 6,155,$$

$$\text{Territorio 4: } 6,839.$$

Obsérvese que la experiencia individual contribuye a la prima en un 59.07 por ciento y el colectivo de asegurados en un 40.93 por ciento.

### 5.3.2. Modelo de Bühlmann-Straub

El modelo de Bühlmann-Straub es una generalización del modelo de Bühlmann estudiado en la sección anterior. Este modelo introduce observaciones ponderadas con factor de ponderación  $m_{ij}$ , como se aprecia en la tabla 5.5.

En consecuencia, se tiene en cuenta que no todas las pólizas ni todos los períodos de tiempo de observación tienen la misma importancia. Para ello se introduce el parámetro conocido  $m_{ij}$ ,  $i = 1, \dots, k; j = 1, \dots, n$  y se considera que las observaciones no son homogéneas en el tiempo, modelizando a través de los pesos  $m_{ij}$ ,  $i = 1, \dots, k; j = 1, \dots, n$ . Establecemos la siguiente notación:

$$\nu = \text{Var}[\theta_i] = m_{ij} \text{Var}[X_{ij}|\theta_i].$$

En la práctica este factor de ponderación suele representar el número de reclamaciones, mientras que  $X_{ij}$ ,  $i = 1, \dots, k; j = 1, \dots, n$  suele representar la cantidad

Tabla 5.5: Datos de la cartera de seguros.

Años	Pólizas o grupos de asegurados					
	1	2	...	$j$	...	$k$
1	$X_{11}$	$X_{21}$	...	$X_{j1}$	...	$X_{k1}$
	$m_{11}$	$m_{21}$	...	$m_{j1}$	...	$m_{k1}$
2	$X_{12}$	$X_{22}$	...	$X_{j2}$	...	$X_{k2}$
	$m_{12}$	$m_{22}$	...	$m_{j2}$	...	$m_{k2}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$n$	$X_{1n}$	$X_{2n}$	...	$X_{jn}$	...	$X_{kn}$
	$m_{1n}$	$m_{2n}$	...	$m_{jn}$	...	$m_{kn}$

reclamada. Siguiendo un argumento similar al elaborado en el modelo de Bühlmann de la sección anterior se concluye (se ha prescindido de nuevo del subíndice  $j$ ) que

$$H[\mu(\theta_i)|X_1, \dots, X_n] = Z_i(n)\bar{X}_i + [1 - Z_i(n)]\bar{X},$$

donde

$$Z(n) = \frac{ma}{ma + s^2},$$

$$m = \sum_{i=1}^k m_i,$$

$$\bar{X} = \frac{1}{m} \sum_{j=1}^n m_j X_j.$$

Como se observa, bajo este modelo el factor de credibilidad es diferente para cada póliza.

Los parámetros estructurales del modelo se estiman a partir de:

$$\hat{\mu} = \bar{X} = \frac{1}{m} \sum_{i=1}^k m_i \bar{X}_i = \frac{1}{m} \sum_{i=1}^k \sum_{j=1}^n m_{ij} X_{ij}, \quad (5.12)$$

$$\bar{X}_i = \frac{1}{m_i} \sum_{j=1}^n m_{ij} X_{ij}, \quad (5.13)$$

$$\hat{\nu} = \frac{1}{\sum_{i=1}^k (n_i - 1)} \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2, \quad (5.14)$$

$$\hat{a} = \frac{1}{m - \frac{1}{m} \sum_{i=1}^k m_i^2} \left[ \sum_{i=1}^k m_i (\bar{X}_i - \bar{X})^2 - \nu(k-1) \right]. \quad (5.15)$$

Puede probarse que  $\hat{\mu}$ ,  $\hat{\nu}$  y  $\hat{a}$  son estimadores insesgados de  $\mu$ ,  $\nu$  y  $a$ , respectivamente. El estimador de  $K$  es  $\hat{K} = \hat{\nu}/\hat{a}$ , mientras que el estimador del factor de credibilidad para el grupo o póliza  $i$ -ésimo es:

$$\hat{Z}(n) = \frac{n}{n + \hat{K}}. \quad (5.16)$$

Por lo tanto, se tiene que el estimador de Bühlmann-Straub para el número medio de reclamaciones es

$$\hat{X}_i = \hat{Z}(n)\bar{X}_i + [1 - \hat{Z}(n)]\hat{\mu}, \quad (5.17)$$

y el estimador para la cantidad total reclamada

$$\hat{C}_i = m_{i,n_i+1}\hat{X}_i. \quad (5.18)$$

Existe una desventaja de este modelo y es que el estimador  $\hat{a}$  puede en ocasiones ser negativo, y el modelo no sería viable.

**Ejemplo 5.7** La tabla 5.6 recoge las cantidades totales reclamadas, así como el número de vehículos asegurados (entre paréntesis) en cada uno de los territorios en los que operan vehículos industriales con una póliza suscrita en una compañía de seguros. Mediante el modelo de Bühlmann-Straub, vamos a estimar la cantidad total reclamada para el año 2009 en cada uno de los territorios.

Tabla 5.6: Número de vehículos expuestos al riesgo (Herzog, 1996).

Territorios	Años			
	2006	2007	2008	2009
1	8000 (40)	11000 (50)	15000 (75)	(75)
2	20000 (100)	24000 (120)	18000 (120)	(95)
3	10000 (50)	15000 (60)	13500 (60)	(60)

Es claro que  $n_1 = n_2 = n_3 = 3$ . Además, tenemos para cada uno de los territorios lo siguiente:

Territorio 1:  $m_{11} = 40$ ,  $m_{12} = 50$ ,  $m_{13} = 75$ . Luego,  $m_1 = \sum_{j=1}^3 m_{1j} = 40 + 50 + 75 = 165$ . Las cantidades medias reclamadas son  $x_{11} = 8000/40 = 200$ ;  $x_{12} = 11000/50 = 220$  y  $x_{13} = 15000/75 = 200$ . Por tanto,

$$\bar{x}_1 = \frac{1}{m_1} \sum_{j=1}^3 m_{1j} x_{1j} = \frac{1}{165} (8000 + 11000 + 15000) = 206,060.$$

Territorio 2:  $m_{21} = 100$ ,  $m_{22} = 120$ ,  $m_{23} = 120$ . Luego,  $m_2 = 340$ . Las cantidades medias reclamadas son ahora  $x_{21} = 200$ ;  $x_{22} = 200$  y  $x_{23} = 150$ . Por tanto,

$$\bar{x}_2 = \frac{1}{m_2} \sum_{j=1}^3 m_{2j} x_{2j} = \frac{1}{340} (20000 + 24000 + 18000) = 182,353.$$

Territorio 3:  $m_{31} = 50$ ,  $m_{32} = 60$ ,  $m_{33} = 60$ . De donde,  $m_3 = 170$ . Las cantidades medias reclamadas son  $x_{31} = 200$ ;  $x_{32} = 250$  y  $x_{33} = 225$ . Por tanto,  $\bar{x}_3 = 226,470$ .

De todo ello, y utilizando (5.12), (5.13), (5.14) y (5.15), se deduce:

$$\begin{aligned} m &= m_1 + m_2 + m_3 = 675, \\ \hat{\mu} &= \frac{1}{m} \sum_{i=1}^3 m_i \bar{X}_i = 199,259, \\ \hat{\nu} &= \frac{1}{\sum_{i=1}^3 (tn_i - 1)} \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} (X_{ij} - \bar{X}_i)^2 = 46073,23, \\ \hat{a} &= 328,331. \end{aligned}$$

Finalmente:  $\hat{k} = \hat{\nu} \hat{a} = 140,325$ . El factor de credibilidad se obtiene a partir de (5.16), el número medio de reclamaciones por medio de (5.17), mientras que la cantidad total reclamada predicha para el año 2009 se obtiene usando (5.18). Las cantidades estimadas aparecen en la tabla 5.7.

Tabla 5.7: Número estimado de vehículos expuestos al riesgo.

Territorio	$\hat{Z}(n)$	$\hat{X}$	$\hat{C}$
1	0.5404	202.935	15220.18
2	0.7078	187.302	17793.76
3	0.5478	214.160	12850.08

## 5.4. Modelos Bayesianos

El uso de distribuciones a priori, con un marcado carácter subjetivo, resulta útil en el mercado de seguros, sobre todo si se tiene en cuenta que cuando se quiere tarificar un riesgo nuevo no se dispone de información.

La visión Bayesiana se incorporó rápidamente a la disciplina actuarial no tardando en demostrarse (Bailey (1945) y Mayerson (1964)) que algunas primas obtenidas mediante la metodología Bayesiana podían escribirse como fórmulas de credibilidad. Así fue como surgió dentro de la ciencia actuarial la *teoría de la credibilidad*, que hoy en día abarca un terreno mucho más amplio del inicialmente establecido. El problema de la teoría de la credibilidad consiste en determinar las ponderaciones que afectan a la experiencia de siniestralidad de una póliza respecto a la experiencia de un colectivo al que pertenece dicha póliza.

Consideremos la cartera de seguros que aparece en la tabla 5.3, que consta de  $k$  pólizas o asegurados y  $n$  períodos de observación de las mismas.

En este contexto, la cuestión básica de la teoría de la credibilidad es determinar una prima establecida como una combinación lineal convexa entre la experiencia particular de un asegurado y la experiencia del colectivo, esto es de toda la cartera. Una expresión válida sería:

$$P_j = Z(n)\tilde{P}_j + [1 - Z(n)]P_0,$$

donde:

- $P_j$  : Prima a aplicar a los asegurados al riesgo  $j$ .
- $P_0$  : Prima a aplicar a un colectivo al que pertenece el asegurado  $j$ .
- $\tilde{P}_j$  : Prima obtenida en base a la experiencia del asegurado  $j$ .
- $Z(n)$  : Factor de credibilidad que debe verificar  $\lim_{n \rightarrow \infty} Z(n) = 1$ , siendo  $n$  el número de expuestos al riesgo  $j$  o el período de observación de la póliza  $j$ . Por tanto, si  $Z(n) = 1$  la experiencia del asegurado es creíble al 100%, mientras que si  $Z(n) = 0$ ,  $P_j = P_0$  y la prima del asegurado  $j$  coincide con la del colectivo al que pertenece dicha póliza.

La fórmula de credibilidad puede además interpretarse de la siguiente manera: se puede considerar a  $P_0$  como la información a priori;  $\tilde{P}_j$  la nueva información obtenida mediante la observación de la siniestralidad del riesgo  $j$  y  $P_j$  el resultado de combinar la información a priori con la información adquirida. Por tanto,

$$\text{Prima (a posteriori)} = [1 - Z(n)]\text{Prima a priori} + Z(n)\text{Experiencia observada.}$$

Visto de esta manera, la teoría de la credibilidad sigue un esquema Bayesiano, donde se da entrada a la información a priori con la información muestral, para obtener finalmente un estimador revisado de la prima.

Una de las principales aplicaciones de la teoría de la credibilidad se presenta en el seguro de automóviles, en el que la prima inicial se va modelando sucesivamente a medida que se incorpora la información de la siniestralidad. Son los denominados *sistemas de tarificación bonus-malus*. Se parte de un nivel  $\mathbf{x}$  neutro, de modo que para niveles superiores a  $\mathbf{x}$  el asegurado entra en la escala *malus* y para niveles inferiores a  $\mathbf{x}$  el asegurado entra en la escala *bonus*.

**Ejemplo 5.8** Veamos que la prima Bayes obtenida en el ejemplo 4.6 se puede escribir como una fórmula de credibilidad.

La prima Bayes obtenida fue:

$$P_B = (\nu + 1) \frac{a + n\bar{X}}{b + n\nu - 2},$$

que puede reescribirse como,

$$P_B = \frac{(\nu + 1)n}{b + n\nu - 2} \bar{X} + \frac{a(\nu + 1)}{b + n\nu - 2}.$$

Ahora, multiplicando y dividiendo el primer sumando de esta expresión por  $\nu$  y el segundo sumando por  $b - 2$  resulta,

$$P_B = \frac{\nu n}{b + n\nu - 2} \frac{\nu + 1}{\nu} \bar{X} + \frac{b - 2}{b + n\nu - 2} \frac{a(\nu + 1)}{b - 2},$$

que puede escribirse como

$$P_B = Z(n) \left( \frac{\nu + 1}{\nu} \bar{X} \right) + [1 - Z(n)] P_C,$$

con:

$$Z(n) = \frac{\nu n}{b + n\nu - 2},$$

siendo  $P_C$  la prima colectiva.

**Ejemplo 5.9** Sean  $X_i$  variables aleatorias independientes con distribución común de Poisson de media  $\theta$ , con función estructura gamma. Veamos que la prima calculada bajo el principio de prima neta se puede escribir como una fórmula de credibilidad.

Si la distribución a priori es  $\mathcal{G}(a, b)$ , sabemos que la distribución a posteriori es  $\mathcal{G}(a + n\bar{X}, b + n)$ . Utilizando (4.41) resulta:

$$P_B = \frac{a + n\bar{X}}{b + n},$$

que puede reescribirse como

$$P_B = \frac{n}{b+n} \bar{X} + \frac{b}{b+n} \frac{a}{b},$$

donde se ha multiplicado y dividido el segundo sumando por  $b$ . Entonces resulta:

$$P_B = Z(n) \bar{X} + [1 - Z(n)] P_C,$$

con:

$$Z(n) = \frac{n}{b+n},$$

siendo  $P_C = a/b$  la prima colectiva.

**Ejemplo 5.10** Vamos a comprobar que la prima Bayes obtenida en el ejemplo 4.5 se puede escribir como una fórmula de credibilidad.

Es inmediato probar que (4.45) puede reescribirse como

$$Z(n) \bar{X} + [1 - Z(n)] P_C,$$

con:

$$Z(n) = \frac{rn}{b + nr - 1},$$

siendo  $P_C$  la prima colectiva.

### 5.4.1. Modelo de Jewell

Comencemos con el siguiente resultado relativo a la familia de distribuciones exponencial.

**Teorema 5.2** *Dados un riesgo  $X$  con función de densidad  $f(x|\theta)$ , y la distribución a priori del parámetro conjugada para esa verosimilitud, entonces el estimador de Bühlmann de la prima neta y el estimador Bayesiano (la prima neta Bayes) coinciden cuando ambas distribuciones pertenecen a la familia exponencial.*

**Demostración:** La demostración se realizará considerando la familia exponencial continua. El caso discreto es similar. Así, dada la familia exponencial con función de densidad,

$$f(x|\theta) = \frac{a(x)e^{-\theta x}}{c(\theta)}, \quad \theta \in \Theta, \quad (5.19)$$

en la que  $c(\theta)$  es la constante de normalización. La distribución a priori conjugada natural para esta verosimilitud es:

$$\pi(\theta) = \frac{[c(\theta)]^{-n_0} e^{-\theta x_0}}{d(n_0, x_0)}, \quad (5.20)$$

donde  $d(n_0, x_0)$  es de nuevo una constante de normalización y  $n_0$  y  $x_0$  dos parámetros de la que depende. La distribución a posteriori es de nuevo del tipo (5.20), pero con los parámetros actualizados:

$$\begin{aligned} n_0 &\rightarrow n_0 + n, \\ x_0 &\rightarrow x_0 + \sum_{i=1}^n X_i. \end{aligned}$$

La prima neta de riesgo y la varianza de  $X$  vienen dadas por:

$$\begin{aligned} P(\theta) = \mu(\theta) &= -\frac{c'(\theta)}{c(\theta)}, \\ \text{Var}[X|\theta] &= \frac{c''(\theta)c(\theta) - c'(\theta)^2}{c(\theta)^2} = -\frac{d}{d\theta} [P(\theta)]. \end{aligned}$$

Derivando (5.20) con respecto a  $\theta$  se obtiene:

$$\begin{aligned} \pi'(\theta) &= \frac{1}{d(n_0, x_0)} \left\{ -n_0 [c(\theta)]^{-n_0-1} c'(\theta) e^{-\theta x_0} - x_0 e^{-\theta x_0} [c(\theta)]^{-n_0} \right\} \\ &= \pi(\theta) \left[ \frac{-n_0 c'(\theta)}{c(\theta)} - x_0 \right] = \pi(\theta) [n_0 \mu(\theta) - x_0]. \end{aligned} \quad (5.21)$$

Integrando ahora (5.21) sobre  $\Theta$  tenemos:

$$\pi(\theta)|_{\Theta} = n_0 \int_{\Theta} \mu(\theta) \pi(\theta) d\theta - x_0,$$

y suponiendo que  $\pi(\theta)$  se anula en los extremos de  $\Theta$  resulta:

$$\int_{\Theta} \mu(\theta) \pi(\theta) d\theta = \frac{x_0}{n_0} = m.$$

Entonces:

$$\int_{\Theta} \mu(\theta) \pi(\theta | X_1, \dots, X_n) d\theta = \frac{x_0 + \sum_{i=1}^n X_i}{n_0 + n} = [1 - Z(n)] m + Z(n) \bar{X},$$

con factor de credibilidad:

$$Z(n) = \frac{n}{n_0 + n}.$$

Derivando (5.21) con respecto a  $\theta$  queda:

$$\begin{aligned} \pi''(\theta) &= \pi'(\theta) [n_0 \mu(\theta) - x_0] + \pi(\theta) n_0 \frac{-c''(\theta)c(\theta) + c'(\theta)^2}{c(\theta)^2} \\ &= \pi(\theta) [n_0 \mu(\theta) - x_0]^2 - \pi(\theta) n_0 \text{Var}(X|\theta). \end{aligned}$$

Finalmente, integrando (5.22) con respecto a  $\Theta$  resulta:

$$\begin{aligned} \int_{\Theta} \pi''(\theta) d\theta &= \int_{\Theta} [n_0 (\mu(\theta) - m)]^2 \pi(\theta) d\theta - n_0 E[\text{Var}(X|\theta)] \\ &= n_0 \{n_0 \text{Var}[\mu(\theta)] - E[\text{Var}(X|\theta)]\}. \end{aligned}$$

Teniendo en cuenta que si  $\pi(\theta)$  se anula en los extremos de  $\Theta$ , también lo hará su derivada, concluimos que

$$n_0 = \frac{E[\text{Var}(X|\theta)]}{\text{Var}[E(X|\theta)]} = \frac{s^2}{a},$$

y, por tanto, el factor de credibilidad  $Z(n)$  es igual al de Bühlmann. ■

En definitiva, el estimador de credibilidad de Bühlmann coincide con el estimador Bayesiano de la prima en un gran número de situaciones. Esto ocurre, por ejemplo, si la distribución a priori es conjugada y la verosimilitud es un miembro de la familia exponencial, entonces el estimador de credibilidad de Bühlmann de la prima neta coincide con el estimador de credibilidad Bayesiano. Así ocurre con los pares de distribuciones beta-binomial, normal-normal, gamma-exponencial y Poisson-gamma, como se puso de manifiesto en el ejemplo 5.12.

**Ejemplo 5.11** Vamos a obtener las distribuciones a posteriori para los pares de verosimilitudes- distribución a priori Poisson-gamma, binomial negativa-beta, binomial-beta, gamma-gamma y normal-normal. Así mismo dar las expresiones de las primas netas de riesgo, colectiva y Bayes para todos los casos y comprobar que pueden escribirse como una fórmula de credibilidad.

La tabla 5.8 recoge los principales pares de verosimilitudes y distribuciones a priori pertenecientes a la familia exponencial de distribuciones. A continuación se recogen las primas netas de riesgo, colectivas y Bayes obtenidas para dichos pares de verosimilitudes-distribuciones a priori. Como se observa, las primas Bayes pueden reescribirse como una fórmula de credibilidad, con factor de credibilidad igual al de Bühlmann (1967) en todos los casos considerados.

1. Poisson-gamma:

$$P_R = \theta, \quad \theta > 0,$$

$$P_C = \frac{a}{b},$$

$$P_B = \frac{a + n\bar{X}}{b + n} = \frac{n}{b + n} \bar{X} + \frac{b}{b + n} \frac{a}{b}$$

Tabla 5.8: Distribuciones conjugadas en la familia exponencial.

Función de verosimilitud Distribución a priori	Distribución a posteriori
$X   \theta \sim \mathcal{P}o(\theta)$ $\theta \sim \mathcal{G}(a, b)$	$\mathcal{G}(a + n, b + n\bar{X})$
$X   \theta \sim \mathcal{NB}(r, \theta)$ $\theta \sim \mathcal{B}(a, b)$	$\mathcal{B}(a + nr, b + n\bar{X})$
$X   \theta \sim \mathcal{Bi}(m, \theta)$ $\theta \sim \mathcal{B}(a, b)$	$\mathcal{B}(a + n\bar{X}, b + mn - n\bar{X})$
$X   \theta \sim \mathcal{G}(\nu, \theta)$ $\theta \sim \mathcal{G}(a, b)$	$\mathcal{G}(a + n\nu, b + n\bar{X})$
$X   \theta \sim \mathcal{N}(\theta, \sigma^2)$ $\theta \sim \mathcal{N}(a, \tau^2)$	$\mathcal{N}\left(\frac{a\sigma^2 + n\bar{X}\tau^2}{\sigma^2 + n\tau^2}, \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2}\right)$

$\mathcal{P}o$ : Poisson,  $\mathcal{G}$ : gamma,  $\mathcal{NB}$ : Binomial negativa,  
 $\mathcal{Bi}$ : Binomial,  $\mathcal{B}$ : Beta,  $\mathcal{N}$ : Normal

$$= Z(n)\bar{X} + [1 - Z(n)]P_C,$$

$$Z(n) = \frac{n}{b + n}.$$

2. Binomial negativa-beta:

$$P_R = \frac{r\theta}{1 - \theta}, \quad r > 0, \quad 0 < \theta < 1,$$

$$P_C = \frac{ra}{b - 1}, \quad a > 0, \quad b > 1,$$

$$\begin{aligned} P_B &= \frac{r(a + n\bar{X})}{b + nr - 1} = \frac{rn}{b + nr - 1}\bar{X} + \frac{b - 1}{b + nr - 1}\frac{ra}{b - 1} \\ &= Z(n)\bar{X} + [1 - Z(n)]P_C, \quad b + nr > 1, \end{aligned}$$

$$Z(n) = \frac{rn}{b + nr - 1}.$$

3. Binomial-beta:

$$P_R = m\theta, \quad m > 0, \quad 0 < \theta < 1,$$

$$P_C = m \frac{a}{a+b}, \quad a > 0, \quad b > 0,$$

$$\begin{aligned} P_B &= \frac{m(a+n\bar{X})}{a+b+mn} = \frac{mn}{a+b+mn} \bar{X} + \frac{a+b}{a+b+mn} \frac{ma}{a+b} \\ &= Z(n)\bar{X} + [1 - Z(n)] P_C, \end{aligned}$$

$$Z(n) = \frac{mn}{a+b+mn}.$$

4. Gamma-gamma:

$$P_R = \frac{\nu}{\theta}, \quad \nu > 0, \quad \theta > 0,$$

$$P_C = \frac{\nu b}{a-1},$$

$$\begin{aligned} P_B &= \frac{\nu(b+n\bar{X})}{a+n\nu-1} = \frac{\nu n}{a+n\nu-1} \bar{X} + \frac{a-1}{a+n\nu-1} \frac{\nu b}{a-1} \\ &= Z(n)\bar{X} + [1 - Z(n)] P_C, \end{aligned}$$

$$Z(n) = \frac{\nu n}{a+n\nu-1}.$$

5. Normal-normal:

$$P_R = \theta, \quad -\infty < \theta < \infty,$$

$$P_C = a, \quad a > 0,$$

$$P_B = \frac{a\sigma^2 + n\bar{X}\tau^2}{\sigma^2 + n\tau^2} = \frac{n\tau^2}{\sigma^2 + n\tau^2} \bar{X} + \frac{\sigma^2}{\sigma^2 + n\tau^2} a$$

$$= Z(n)\bar{X} + [1 - Z(n)]P_C,$$

$$Z(n) = \frac{n\tau^2}{\sigma^2 + n\tau^2}$$

**Ejemplo 5.12** Vamos a comprobar que el factor de credibilidad obtenido en el ejemplo 5.9 mediante la metodología Bayesiana coincide con el de Bühlmann (1967).

El factor de credibilidad obtenido fue,

$$Z(n)_{\text{Bayes}} = \frac{n}{b+n}.$$

Por otro lado, el factor de credibilidad de Bühlmann viene dado por:

$$Z(n)_{\text{Bühlmann}} = \frac{n\text{Var}[E(X|\theta)]}{n\text{Var}[E(X|\theta)] + E[\text{Var}(X|\theta)]}.$$

Ahora,

$$E[X|\theta] = \text{Var}[X|\theta] = \theta, \quad E[\text{Var}(X|\theta)] = E[\theta] = \frac{a}{b},$$

$$\text{Var}[E(X|\theta)] = \text{Var}[\theta] = E[\theta^2] - E^2[\theta] = \frac{a(a+1)}{b^2} - \left(\frac{a}{b}\right)^2 = \frac{a}{b^2}.$$

Por tanto,

$$Z(n)_{\text{Bühlmann}} = \frac{n\frac{a}{b^2}}{n\frac{a}{b^2} + \frac{a}{b}} = \frac{n}{b+n} = Z(n)_{\text{Bayes}}.$$

### 5.4.2. Modelo de Landsman y Makov

Landsman y Makov (1998) extienden el resultado de Jewell a la familia exponencial de dispersión estudiada en el capítulo anterior. De modo que se verifica el siguiente resultado.

**Teorema 5.3** *En la familia exponencial de dispersión y usando la distribución a priori (4.75), la prima Bayes viene dada por,*

$$P_B = \frac{\lambda \sum_{i=1}^n X_i + \mu n_0}{\lambda n + n_0} = \frac{\lambda n}{\lambda n + n_0} \cdot \bar{X} + \frac{n_0}{\lambda n + n_0} \cdot \mu, \quad (5.22)$$

que es una fórmula de credibilidad exacta, con factor de credibilidad dado por

$$Z(t) = \frac{\lambda n}{\lambda n + n_0}.$$

Además, la prima de Bühlmann coincide con la prima Bayes.

**Demostración:** Se verifica que,

$$E[X_{n+1}|X] = \int_{\theta_0}^{\theta_1} \mu(\theta) \frac{[q(\theta)]^{-\tilde{n}_0} \exp(-\theta \tilde{\mu} \tilde{n}_0)}{c(\tilde{\mu}, \tilde{n}_0)} d\theta,$$

donde  $\mu(\theta)$  viene dada en (4.73). Mediante integración por partes, se obtiene fácilmente (5.22). La prima de Bühlmann coincide con la prima Bayes, puesto que ésta última es lineal en las observaciones. ■

La expresión (5.22) depende del parámetro  $\lambda$ , y resulta evidente que si asignamos una distribución a priori  $\pi(\lambda)$  a dicho parámetro, la prima neta Bayes resultante es:

$$P_B = \bar{X} \int \frac{\lambda n}{\lambda n + n_0} \pi(\lambda) d\lambda + \mu \int \frac{n_0}{\lambda n + n_0} \pi(\lambda) d\lambda, \quad (5.23)$$

para lo que se ha asumido independencia entre los parámetros  $\theta$  y  $\lambda$ . Si la distribución a priori del parámetro  $\lambda$  se desconoce podemos incorporar una estructura jerárquica, que veremos más adelante, o bien elegir una clase amplia de distribuciones de probabilidad en el escenario del análisis Bayesiano robusto. Este tema también se abordará más adelante. Otra posibilidad, y así lo llevan a cabo Landsman y Makov (1999) es elegir una distribución a priori de máxima entropía.

La entropía de una variable aleatoria  $\lambda$  con distribución a priori  $\pi(\lambda)$  se define como

$$E_\pi(\lambda) = - \int_0^\infty \pi(\lambda) \log \pi(\lambda) d\lambda. \quad (5.24)$$

La distribución de máxima entropía es aquella que maximiza (5.24) sujeta a que el parámetro satisfaga cierta restricción. Ahora tenemos el siguiente resultado que presentamos sin prueba.

**Teorema 5.4** *Si tomamos en (5.23) la distribución de máxima entropía que satisface la restricción*

$$\int_0^\infty \lambda \pi(\lambda) d\lambda = \lambda_0,$$

entonces la prima neta Bayes viene dada por

$$P_B = Z(n) \bar{X} + [1 - Z(n)] \mu,$$

donde

$$Z(n) = 1 - \frac{n_0}{n\lambda_0} \exp\{n_0/(n\lambda_0)\} \Gamma\left(0, \frac{n_0}{n\lambda_0}\right),$$

siendo:

$$\Gamma(\alpha, x) = \int_0^x s^{\alpha-1} \exp(-s) ds,$$

la función gamma incompleta.

### 5.4.3. Modelos Bayesianos jerárquicos

Con objeto de motivar esta sección, continuamos con el ejemplo Poisson-gamma, en el que se supone que la distribución del número de siniestros para una póliza de seguros sigue una distribución de Poisson con parámetro  $\theta > 0$ . El camino que hemos seguido hasta ahora es asumir que el actuario no tiene ningún problema en especificar una distribución a priori para este parámetro, eligiendo, como hasta ahora, la distribución a priori gamma con parámetros  $a > 0$  y  $b > 0$ . El estimador Bayesiano del parámetro  $\theta > 0$  y la prima Bayes, como ya sabemos, vienen dados por

$$P_B = \frac{a + n\bar{X}}{b + n}, \quad (5.25)$$

donde  $\bar{X}$  es la media muestral y  $n$  el número de períodos de tiempo observados.

El problema que tiene esta aproximación es que en ocasiones el actuario no tiene por qué conocer de manera exacta dicha distribución. Existen diversas soluciones alternativas:

1. Elegir una clase amplia de distribuciones a priori, que contenga a dicha distribución a priori inicial en la que se confía, y sobre dicha clases estudiar la variación de  $P_B$ . Esto se conoce en la literatura como análisis de robustez Bayesiano. Este tema será tratado en una sección posterior.
2. Llevar a cabo un análisis empírico Bayes. Es decir, estimar los parámetros de la distribución a priori a partir de los datos utilizando para ello la distribución incondicional o marginal de los datos, esto es:

$$f(x) = \int_0^{\infty} f(x|\theta)\pi(\theta)d\theta,$$

3. Suponer que no se conocen exactamente los parámetros de los que depende la distribución a priori y admitir que uno o todos los parámetros de la misma (hiper-parámetros) siguen una determinada distribución de probabilidad (llamada distribución hiperpriori). En este caso, el modelo de verosimilitud-distribución a priori-distribución hiper-priori, se denomina modelo jerárquico.

La estructura de un modelo jerárquico suele expresarse en la forma de estados, de la siguiente forma, elaborada para el desconocimiento de un sólo parámetro:

- Primer nivel: Verosimilitud  $f(x|\theta)$ ,
- Segundo nivel: Distribución a priori  $\pi(\theta|\lambda)$ , hiper-parámetro  $\lambda$ ,
- Tercer nivel: Distribución hiperpriori  $\pi(\lambda)$ .

La estructura Bayesiana de verosimilitud-distribución a priori, analizada hasta el momento puede contemplarse como una estructura jerárquica con sólo dos estados.

Veamos ahora un modelo jerárquico en el que la prima neta Bayes puede reescribirse como una fórmula de credibilidad. El ejemplo es un caso más general del modelo planteado en Klugman (1992). Para ello supongamos que disponemos de una cartera de seguros con  $k > 0$  asegurados en un año. Supongamos también que el número de reclamaciones del asegurado  $i$ -ésimo sigue una distribución de Poisson de parámetro  $\theta_i > 0$ , con distribución a priori  $\theta_i \sim \mathcal{G}(\lambda, b)$ ,  $i = 1, \dots, k$ . Supongamos ahora que el hiper-parámetro  $a > 0$  es desconocido, aleatorio y con distribución hiper-priori impropia  $\pi_1(\lambda) \propto 1/\lambda$ . Nos encontramos con una estructura jerárquica con los siguiente niveles:

- Primer nivel:  $X_i \sim \mathcal{P}o(\theta_i)$ ,  $\theta_i > 0$ ,  $i = 1, \dots, k$
- Segundo nivel:  $\theta_i \sim \mathcal{G}(\lambda, b)$ ,  $\lambda > 0$ ,  $b$  conocido,  $i = 1, \dots, k$
- Tercer nivel:  $\pi(\lambda) \propto 1/\lambda$ .

Las distribuciones marginales que se precisan para el análisis son los siguientes:

$$\begin{aligned}\pi(\theta|X_1, \dots, X_n) &= \frac{f(X_1, \dots, X_n|\theta) \int \pi(\theta|\lambda)\pi(\lambda)d\lambda}{\int \int f(X_1, \dots, X_n|\theta)\pi(\theta|\lambda)\pi(\lambda)d\theta d\lambda}, \\ \pi(\lambda|X_1, \dots, X_n) &= \frac{\pi(\lambda) \int f(X_1, \dots, X_n|\theta)\pi(\theta|\lambda)d\theta}{\int \int f(X_1, \dots, X_n|\theta)\pi(\theta|\lambda)\pi(\lambda)d\theta d\lambda},\end{aligned}\quad (5.26)$$

donde  $\theta = (\theta_1, \dots, \theta_k)$ . De aquí en adelante, cuando no se especifique, tanto los productos como las sumatorias se extienden desde 1 hasta  $k$ . Denotando mediante  $\underline{X} = (X_1, \dots, X_n)$  la información muestral, tenemos:

$$\begin{aligned}\int \pi(\theta|\lambda)\pi(\lambda)d\lambda &= \frac{(\prod \theta_i^{b-1})}{\Gamma(b)} \int_0^\infty \lambda^{kb-1} e^{-\lambda \sum \theta_i} d\lambda = \frac{(\prod \theta_i^{b-1})}{(\sum \theta_i)^{kb}}, \\ \int f(\underline{X}|\theta)\pi(\theta|\lambda)d\theta &= \int \dots \int \left( \prod_{i=1}^k \frac{e^{-\theta_i} \theta_i^{X_i}}{X_i!} \right) \prod_{i=1}^k \frac{\lambda^b}{\Gamma(b)} \theta_i^{b-1} e^{-\lambda \theta_i} d\theta_1 \dots d\theta_k \\ &= \frac{\lambda^{kb}}{\Gamma(b) \prod X_i!} \frac{\prod \Gamma(X_i + b)}{(\lambda + 1)^{kb + \sum X_i}},\end{aligned}\quad (5.27)$$

Ahora:

$$\int f(\underline{X}|\theta)\pi(\theta|\lambda)\pi(\lambda)d\lambda = \int_0^\infty \frac{\lambda^{kb}}{\Gamma(b) \prod X_i!} \frac{\prod \Gamma(X_i + b)}{(\lambda + 1)^{kb + \sum X_i}} d\lambda.$$

Haciendo ahora el cambio de variable  $\mu = \lambda/(1 + \lambda)$  se llega, después de algunos cálculos:

$$\int f(\underline{X}|\theta)\pi(\theta|\lambda)\pi(\lambda)d\lambda = \frac{\prod \Gamma(X_i + b)}{\Gamma(b) \prod X_i!} B(kb, \sum X_i). \quad (5.28)$$

Por tanto, tenemos que:

$$\begin{aligned} \pi(\theta|\underline{X}) &\propto \frac{\Gamma(b) \prod X_i!}{\prod \Gamma(X_i + b)} \frac{1}{B(kb, \sum X_i)} \prod \frac{e^{-\theta_i} \theta_i^{X_i}}{X_i!} \frac{(\prod \theta_i^{b-1})}{(\sum \theta_i)^{kb}} \\ &= \frac{(\prod \theta_i^{X_i + b - 1}) e^{-\sum \theta_i}}{(\sum \theta_i)^{kb} \prod \Gamma(X_i + b)} \frac{\Gamma(b)}{B(kb, \sum X_i)}. \end{aligned} \quad (5.29)$$

Puede probarse que (5.29) constituye una función de densidad de probabilidad genuina si  $\Gamma(kb)$  se suprime del denominador (es decir, conocemos la constante de normalización de la integral).

Ahora, utilizando (5.26), (5.27), (5.28), y después de algunos cálculos, se obtiene la distribución a posteriori del parámetro  $\lambda$ , que viene dada por:

$$\pi(\lambda|\underline{X}) \propto \frac{1}{B(kb, k\bar{X})} \frac{\lambda^{kb-1}}{(\lambda + 1)^{k\bar{X} + kb}}. \quad (5.30)$$

Se verifica que  $\pi(\lambda|\underline{X})$  es una distribución de Pareto generalizada con parámetros  $kb$ ,  $k\bar{X}$ , también denominada inversa Beta. Por tanto, en (5.30) el símbolo  $\propto$  se puede reemplazar por el de igualdad.

Ahora tenemos que para el asegurado  $r$ -ésimo la prima neta Bayes vendrá dada por

$$\begin{aligned} P_B = E[\theta_r|\underline{X}] &= \frac{\Gamma(b)\Gamma(kb + k\bar{X})}{\prod \Gamma(X_i + b)\Gamma(k\bar{X})} \frac{\prod \Gamma(xX_r + b + 1)\Gamma(k\bar{X} + 1)}{\Gamma(b)\Gamma(kb + k\bar{X} + 1)} \\ &= \frac{\bar{X}(X_r + b)}{\bar{X} + b}, \end{aligned} \quad (5.31)$$

donde  $X_r$  es la media del asegurado  $r$ -ésimo, ya que se dispone sólo de una observación (1 año). Ahora tenemos que (5.31) puede reescribirse como

$$P_B = ZX_r + (1 - Z)\bar{X},$$

donde, el factor de credibilidad es:

$$Z = \frac{\bar{X}}{b + \bar{X}}.$$

Puesto que,

$$E[\bar{X}|\lambda] = \frac{b}{\lambda}$$

tenemos que el factor de credibilidad es en términos de  $\lambda$ ,  $Z = 1/(1 + \lambda)$ .

### El modelo jerárquico Poisson-gamma-gamma

En Gómez-Déniz et al. (2008c) se lleva a cabo el desarrollo del modelo jerárquico Poisson-gamma-gamma que permite construir una nueva distribución de probabilidad discreta. Una propiedad importante de esta distribución es que resulta sobredispersa (varianza mayor que la media), lo que la hace una buena candidata para ser utilizada en el ramo de seguro de automóviles, donde empíricamente está comprobado que se manifiesta esta propiedad (ver Meng et al., 1999). Esta distribución se puede utilizar para calcular primas bonus-malus, pareciendo responder también bastante bien el modelo para resolver problemas de sobrecargas.

En esta sección presentaremos dicho modelo y lo utilizaremos para ajustar nuevos datos de reclamaciones de seguro de automóviles y calcular a partir de dichos datos, las primas bonus-malus correspondientes.

El modelo jerárquico está construido partiendo de los siguientes niveles:

$$\text{Nivel 1: } X_i|\lambda \sim \mathcal{P}o(\lambda), \quad i = 1, \dots, n, \quad \text{independientes} \quad (5.32)$$

$$\text{Nivel 2: } \lambda|b \sim \pi_1(\lambda|a, b) = \mathcal{G}(a, b), \quad a, b > 0 \quad (5.33)$$

$$\text{Nivel 3: } b \sim \pi_2(b|\alpha, \beta) = \mathcal{G}(\alpha, \beta), \quad \alpha, \beta > 0 \quad (5.34)$$

donde  $\mathcal{P}o(\lambda)$  denota la distribución de Poisson con media  $\lambda$ , y  $\mathcal{G}(a, b)$  la distribución gamma. Utilizando (5.33) y (5.34) se obtiene la distribución incondicional para el parámetro  $\lambda$ :

$$\begin{aligned} \pi_1(\lambda|a, \alpha, \beta) &= \int_0^\infty \pi_1(\lambda|b)\pi_2(b)db \\ &= \int_0^\infty \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda} \frac{\beta^\alpha}{\Gamma(\alpha)} b^{\alpha-1} e^{-\beta b} db \\ &= \frac{1}{B(a, \alpha)\beta} \frac{(\lambda/\beta)^{a-1}}{(1 + \lambda/\beta)^{a+\alpha}} \end{aligned} \quad (5.35)$$

donde  $B(x, y)$  denota la función beta. Esta distribución se corresponde con la distribución Tipo IV de Pearson, también llamada distribución beta de segunda especie (Stuart y Ord (1987), capítulo 6 y Johnson et al. (1995), capítulo 27). Escribiremos  $\lambda \sim \mathcal{B}2(a, \alpha; \beta)$ .

La distribución predictiva viene expresada en términos de la función confluyente hipergeométrica, como aparece a continuación.

$$\begin{aligned} p(x_1, \dots, x_n) &= \Pr(X_1 = x_1, \dots, X_n = x_n) \\ &= \int_0^\infty f(x_1, \dots, x_n|\lambda)\pi(\lambda|a, \alpha, \beta)d\lambda \\ &= \frac{1}{\prod_{i=1}^n x_i!} \int_0^\infty e^{-t\lambda} \lambda^x \frac{1}{B(a, \alpha)\beta} \frac{(\lambda/\beta)^{a-1}}{(1 + \lambda/\beta)^{a+\alpha}} d\lambda \end{aligned}$$

$$= \frac{\beta^x \Gamma(a+x)}{\prod_{i=1}^t x_i! \mathcal{B}(a, \alpha)} \mathcal{U}(a+x, x-\alpha+1, \beta n),$$

donde  $\mathcal{U}(a, b, z)$  representa la función confluyente hipergeométrica definida (Goovaerts y De Pril, 1980) como:

$$\mathcal{U}(m, p, z) = \frac{1}{\Gamma(m)} \int_0^\infty e^{-zs} s^{m-1} (1+s)^{p-m-1} ds.$$

Si  $n = 1$  obtenemos la distribución predictiva:

$$\Pr(X = x) = \frac{\beta^x \Gamma(a+x)}{x! \mathcal{B}(a, \alpha)} \mathcal{U}(a+x, x-\alpha+1, \beta). \quad (5.36)$$

La distribución (5.36) es unimodal, puesto que se trata de una mezcla de una distribución de Poisson con una distribución unimodal (Holgate, 1970).

Los valores (5.36) se pueden obtener de manera recursiva teniendo en cuenta que:

$$\beta(a+\alpha)p_{a,\alpha,\beta}(x) - \alpha(x+1)p_{a,\alpha+1,\beta}(x+1) - \alpha\beta p_{a,\alpha+1,\beta}(x) = 0.$$

También, es posible obtener que los momentos factoriales de (5.36) vienen dados por:

$$\mu'_{[r]}(X) = \beta^r \cdot \frac{\Gamma(a+r)\Gamma(\alpha-r)}{\Gamma(a)\Gamma(\alpha)}, \quad \text{if } \alpha > r, \quad (5.37)$$

donde  $\mu'_{[r]}(X) = E[X!/(X-r)!]$ . Se obtiene:

$$\begin{aligned} E[X] &= \mu'_{[1]}(X), \\ E[X^2] &= \mu'_{[2]}(X) + \mu'_{[1]}(X), \\ E[X^3] &= \mu'_{[3]}(X) + 3\mu'_{[2]}(X) + \mu'_{[1]}(X). \end{aligned}$$

Ahora, los hiper-parámetros del modelo pueden estimarse por el método de los momentos resolviendo el sistema de ecuaciones:

$$\begin{aligned} f_1 &= \mu'_{[1]}(X) = \frac{a\beta}{\alpha-1}, \\ f_2 &= \mu'_{[2]}(X) = \frac{a(a+1)\beta^2}{(\alpha-1)(\alpha-2)}, \\ f_3 &= \mu'_{[3]}(X) = \frac{a(a+1)(a+2)\beta^3}{(\alpha-1)(\alpha-2)(\alpha-3)}. \end{aligned}$$

que proporciona los estimadores consistente:

$$\hat{a} = \frac{2f_1(f_2^2 - f_1f_3)}{-f_1f_2^2 + 2f_1^2f_3 - f_2f_3}, \quad (5.38)$$

$$\hat{\alpha} = \frac{f_1^2 f_2 - 4f_2^2 + 3f_1 f_3}{f_1^2 f_2 - 2f_2^2 + f_1 f_3}, \quad (5.39)$$

$$\hat{\beta} = \frac{f_1 f_2^2 - 2f_1^2 f_3 + f_2 f_3}{f_1^2 f_2 - 2f_2^2 + f_1 f_3}. \quad (5.40)$$

Finalmente, la distribución es sobredispersa (5.36), puesto que:

$$\frac{\text{Var}[X]}{E[X]} = 1 + \frac{\beta}{\alpha - 2} \left( 1 + \frac{a}{\alpha - 1} \right) > 1.$$

Para el modelo desarrollado pueden obtenerse expresiones cerradas para las primas neta bonus-malus, como mostramos en el siguiente resultado.

**Proposición 5.2** *Dado el modelo Poisson-gamma-gamma presentado, la prima neta bonus-malus viene dada por:*

$$P_{BM} = \frac{(a+x)(\alpha-1)}{a} \frac{\mathcal{U}(a+x+1, x-\alpha+2, \beta n)}{\mathcal{U}(a+x, x-\alpha+1, \beta n)}. \quad (5.41)$$

**Demostración:** Se obtiene, de manera inmediata, teniendo en cuenta que la media de  $\lambda$ , así como la media a posteriori del parámetro dados los datos  $\underline{X} = (X_1, \dots, X_n)$ , vienen dadas por:

$$E[\lambda] = \frac{a\beta}{\alpha-1}, \quad \text{if } \alpha > 1,$$

$$E[\lambda|\underline{X}] = \beta(a+x) \frac{\mathcal{U}(a+x+1, x-\alpha+2, \beta n)}{\mathcal{U}(a+x, x-\alpha+1, \beta n)},$$

respectivamente. ■

**Ejemplo 5.13** La tabla 5.9 muestra el número de asegurados para las reclamaciones  $x = 0, 1, \dots$  en dos carteras de seguros de automóviles en Suiza (1961) y Bélgica (1993). Estas carteras de seguros aparecen en Denuit (1997) y Willmot (1987). Suponiendo la estructura jerárquica (5.32)-(5.34) así como la estructura Bayesiana estándar Poisson-gamma se trata de:

1. Estimar mediante el método de los momentos los parámetros del modelo jerárquico y mediante el método de máxima verosimilitud los parámetros en el modelo Poisson-gamma.
2. Obtener las tablas de frecuencias ajustadas y estudiar la bondad de los ajustes mediante el test  $\chi^2$ .
3. Obtener para ambos modelos y ambas carteras de seguros una matriz de orden  $3 \times 4$  en la que el elemento  $a_{ij}$ ,  $i = 1, 2, 3$ ;  $j = 1, 2, 3, 4$ , sea el porcentaje de prima a aplicar en el período  $i$  con  $j$  reclamaciones, bajo un sistema de tarificación bonus-malus y el principio de prima neta.

4. Calcular para ambos modelos y ambas carteras de seguros el porcentaje de penalización de la prima al incrementarse, en la clase 0, el número de reclamaciones en 1 unidad.

Tabla 5.9: Reclamaciones en dos carteras de seguros de automóviles.

Número de reclamaciones	Suiza, 1961	Bélgica, 1993
	Observadas	Observadas
0	103704	57178
1	14075	5617
2	1766	446
3	255	50
4	45	8
5	6	0
6	2	—

Veamos la respuesta a cada una de estas cuestiones:

1. A partir de las ecuaciones (5.38)-(5.40) se obtienen los estimadores por momentos del modelo Poisson-gamma-gamma. Para la estimación por máxima verosimilitud de los parámetros del modelo Poisson-gamma se procede igual que en el ejemplo 4.9. Los valores obtenidos aparecen en la tabla 5.10.

Tabla 5.10: Resumen datos correspondientes a la tabla 5.11

	PGG		PG	
	Suiza, 1961	Bélgica, 1993	Suiza, 1961	Bélgica, 1993
$\bar{X}$	0.15	0.10	0.15	0.10
$s^2$	0.18	0.11		
$\hat{\alpha}$	6.52	4.69		
$\hat{a}$	1.55	3.05	1.03	1.27
$\hat{b}$			6.65	12.10
$\hat{\beta}$	0.55	0.12		
$\chi^2$	1.71	2.39	12.02	11.15
g.l.	2	1	2	1
$p$ -valor	0.42	0.12	0.00	0.00

Tabla 5.11: Ajuste de reclamaciones en el ramo de automóviles.

Número de reclamaciones	Suiza, 1961		Bélgica, 1993	
	Observadas	Estimadas	Observadas	Estimadas
0	103704	103715.0	57178	57183.9
1	14075	14041.4	5617	5594.3
2	1766	1799.1	446	467.6
3	255	249.5	50	43.8
4	45	39.2	8	5.2
5	6	7.1	0	0.55
6	2	1.5	—	—

- Los valores ajustados aparecen en la tabla 5.11. En la tabla 5.10 aparecen también los valores de la  $\chi^2$ , el número de grados de libertad y el  $p$ -valor. Es evidente que el modelo Poisson-gamma-gamma mejora las estimaciones obtenidas mediante el modelo Poisson-gamma.
- Sustituyendo los valores de los parámetros de la tabla 5.10 en la expresión (5.41) se obtienen las primas netas bonus-malus que se muestran en las tablas 5.12 y 5.13. Los valores en negrita corresponden al modelo Poisson-gamma y los otros al modelo Poisson-gamma-gamma. Puede observarse, a la vista de las mismas, que las primas son ligeramente mayores para los asegurados en la clase  $x = 0$  en el modelo Poisson-gamma-gamma y menores para los asegurados en la clase  $x = 1$ . Insistimos que las carteras de seguros de automóviles el mayor número de asegurados está en la clase  $x = 0$  y después en la clase  $x = 1$ . El resto de las clases tienen un porcentaje de asegurados muy bajo en relación con las dos clases anteriores.
- Los porcentajes de penalización pedidos aparecen en la tabla 5.14. Se observa que este porcentaje es mucho menor para el modelo Poisson-gamma-gamma. Es decir, este modelo resuelve el problema de las sobrecargas estudiado en el capítulo anterior. Por un lado, los asegurados en la clase  $x = 0$  pagan una prima ligeramente mayor a cambio de que si experimentan una reclamación el porcentaje de penalización en la prima no sea demasiado grande.

En Gómez-Déniz et al. (2006b) los autores realizan un estudio de sensibilidad del modelo anterior, mientras que Pérez et al. (2006) han desarrollado el modelo anterior para el principio de varianza.

Tabla 5.12: Primas netas bonus-malus. Modelo Poisson-gamma-gamma y Poisson-gamma (en negrita). Suiza, 1961.

$n$	$x$				
	0	1	2	3	4
0	100.00				
1	87.26	165.54	269.07	407.00	590.05
	<b>86.92</b>	<b>171.32</b>	<b>255.72</b>	<b>340.11</b>	<b>425.51</b>
2	78.14	143.81	224.87	324.40	445.17
	<b>76.87</b>	<b>226.15</b>	<b>375.43</b>	<b>524.71</b>	<b>674.00</b>
3	71.09	128.29	196.08	275.75	368.27
	<b>68.91</b>	<b>269.62</b>	<b>470.34</b>	<b>671.05</b>	<b>871.76</b>

Tabla 5.13: Primas netas bonus-malus. Modelo Poisson-gamma-gamma y Poisson-gamma (en negrita). Bélgica, 1993.

$n$	$x$				
	0	1	2	3	4
0	100.00				
1	92.99	159.22	269.51	460.66	785.27
	<b>92.36</b>	<b>165.09</b>	<b>237.82</b>	<b>310.55</b>	<b>383.28</b>
2	87.50	144.62	230.83	361.32	552.35
	<b>85.81</b>	<b>220.95</b>	<b>356.10</b>	<b>491.24</b>	<b>626.38</b>
3	82.95	133.76	206.00	307.76	446.38
	<b>80.13</b>	<b>269.42</b>	<b>458.71</b>	<b>648.00</b>	<b>837.28</b>

Tabla 5.14: Porcentaje de penalización al pasar de  $x_0$  a  $x_1$  en la clase 0.

$x_0 - x_1$	Suiza, 1961	Bélgica, 1993
0 - 1	65.54%	59.22%
	<b>71.32%</b>	<b>65.09%</b>
1 - 2	64.80%	55.52%
	<b>160.18%</b>	<b>139.22%</b>
2 - 3	64.17%	52.86%
	<b>250.74%</b>	<b>213.97%</b>

### El modelo jerárquico general

Abordamos ahora el modelo jerárquico general que resulta útil para jerarquizar carteras y subcarteras de seguros.

El modelo jerárquico constituye una generalización del modelo de Bühlmann-Straub (1972). Se supone que cada cartera puede estar dividida en un cierto número de subcarteras, cada una de ellas caracterizada por un parámetro de riesgo desconocido que describe como difiere una subcartera de las otras. El modelo más flexible es el modelo jerárquico de múltiples niveles. El datos del esquema para dos niveles (dos subcarteras) aparece en la tabla 5.15.

Tabla 5.15: Datos de modelo jerárquico de dos niveles.

$p = 1$			$p = 2$		
$\theta_1$			$\theta_2$		
$\theta_{11}$	$\theta_{12}$	...	$\theta_{11}$	$\theta_{12}$	...
$x_{111}$	$x_{121}$	...	$x_{211}$	$x_{221}$	...
$x_{112}$	$x_{122}$	...	$x_{212}$	$x_{222}$	...

En la práctica  $\theta_{ij}$  representa las características del asegurado  $j$  dentro de la subcartera  $i$  y  $\theta_i$  puede representar, por ejemplo, la zona geográfica  $i$ -ésima.

Para desarrollar el modelo podemos situarnos en el ramo de los seguros de automóviles en España. El ramo de automóviles es el más importante después del seguro de vida. El procedimiento de tarificación de un asegurado se realiza en función de la zona geográfica española donde se circula habitualmente. Así, determinadas regiones españolas tienen una siniestralidad mayor que otras, debido sobre todo a sus condiciones geográficas, climatológicas, número de vehículos que componen el parque automovilístico, etc.

En este sentido, en España la división se efectúa en tres zonas diferentes (A, B y C) según el riesgo atribuido a las mismas. En la tabla 5.17 aparecen algunas de las provincias pertenecientes a cada una de las tres zonas.

La zona A se considera de siniestralidad alta, la zona B de siniestralidad media, y la zona C de siniestralidad baja.

Por lo que se refiere a las características individuales de los conductores asegurados podemos mencionar el sexo, la edad, estado civil, antigüedad de posesión del carnet de conducir, número de veces que precisó examinarse para la obtención del mismo, etc.

El objetivo del modelo vuelve a ser encontrar los estimadores de credibilidad para las primas de riesgo individuales, es decir las primas a cobrar en cada uno de los contratos.

El modelo jerárquicos más simple ha sido planteado por Klugman (1992), y se

Tabla 5.16: Zonas geográficas en el ramo de automóviles.

A	B	C
Asturias	Alicante	Albacete
Guipúzcoa	Barcelona	Almería
Las Palmas	Madrid	Granada
Tenerife	Málaga	Soria
Vizcaya	Sevilla	Teruel
⋮	⋮	⋮

denomina *one-way*, que se corresponde con el modelo clásico de Bühlmann-Straub (1972). Recordemos que una cartera consiste en pólizas o contratos más o menos similares, aunque nunca completamente idénticas. Como ya hemos hecho anteriormente, esto lo modelizamos mediante la variable aleatoria  $\theta_{js}$ , que describe totalmente las características del riesgo  $j$ ,  $j = 1, 2, \dots, k$  en el período  $s$ ,  $s = 1, 2, \dots, n$ . Intuitivamente uno asume que las diferencias entre los contratos y períodos son causadas por parámetros diferentes  $\theta_{11}, \theta_{12}, \dots, \theta_{kn}$ .

Sin embargo en la mayoría de los modelos de teoría de la credibilidad se supone que los parámetros  $\theta_{js}$  son homogéneos en el tiempo, i.e. los parámetros de riesgo para un contrato fijo no cambian con el mismo; luego desaparecen los subíndices  $s$ , en los parámetros de riesgo, y escribimos  $\theta_1, \theta_2, \dots, \theta_k$ . Consideraremos de nuevo a estos parámetros como variables aleatorias.

Resulta obvio que este modelo es más realista que los tratados hasta este momento pero como contrapartida presenta el inconveniente de su mayor complejidad matemática.

El modelo jerárquico general consta de los siguientes niveles:

- Primer nivel:  $f(x|\theta, F)$ ,
- Segundo nivel:  $\pi_1(\theta|\mu, G)$ ,
- Tercer nivel:  $\pi_2(\mu, F, G)$ .

El primer nivel indica que los datos dependen de los parámetros de interés  $\theta$  y algún otro parámetro  $F$ . El segundo nivel indica cómo varían los parámetros adicionales desconocidos  $\mu$  y  $G$ . Finalmente se incorpora un tercer nivel donde se considera una distribución conjunta de los parámetros. El objetivo del modelo consiste en estimar la media y la varianza a posteriori dada la observación muestral  $\underline{X} = (X_1, \dots, X_n)$ , para lo que se precisa de la distribución a posteriori de  $\theta$ , que viene dada por (Klugman, 1992),

$$\pi(\theta|\underline{X}) = \int \int \int \pi_1(\theta|\underline{X}, \mu, F, G) \pi_2(\mu, F, G|\underline{X}) d\mu dF dG,$$

donde

$$\begin{aligned}\pi_1(\theta|\underline{X}, \mu, F, G) &= \frac{f(\underline{X}|\theta, F)\pi_1(\theta|\mu, G)}{\int f(\underline{X}|\theta, F)\pi_1(\theta|\mu, G)d\theta}, \\ \pi_2(\mu, F, G|\underline{X}) &= \frac{\pi_2(\mu, F, G) \int f(\underline{X}|\theta, F)\pi_1(\theta|\mu, G)d\theta}{\int \int \int f(\underline{X}|\theta, F)\pi_1(\theta|\mu, G)\pi_2(\mu, F, G)d\mu dF dG},\end{aligned}$$

siendo  $\pi_1(\theta|\mu, G)$  la distribución a priori de  $\theta$  (elegida por el actuario),  $\pi_2(\mu, F, G)$  la distribución conjunta de los parámetros implicados y  $\pi_1(\theta|\mathbf{x}, \mu, F, G)$  y  $\pi_2(\mu, F, G|\mathbf{x})$  las distribuciones a posteriori dados los datos.

La media y la varianza a posteriori de los elementos de  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$  son de especial importancia. Si  $\theta_i$  es uno de esos elementos, la media y varianza a posteriori de  $\theta_i$  pueden obtenerse a partir de,

$$\begin{aligned}E[\theta_i|\underline{X}] &= \int \theta_i \pi(\theta|\underline{X})d\theta, \\ E[\theta_i^2|\underline{X}] &= \int \theta_i^2 \pi(\theta|\underline{X})d\theta.\end{aligned}$$

Es obvio que el cálculo de estas integrales múltiples puede resultar complejo, de ahí que generalmente se impongan restricciones al modelo general para hacer el asunto más tratable, trabajando, en el escenario Bayesiano, con el modelo jerárquico normal, que presentamos a continuación.

### El modelo jerárquico normal

En este modelo suponemos las distribuciones en el primer y segundo nivel son distribuciones del tipo normal multivariantes. Los siguientes argumentos justifican la elección de esta distribución:

1. El análisis se hace frecuentemente con indemnizaciones medias y no con indemnizaciones totales (o con número medio de siniestros y no con siniestros). Además, eligiendo una distribución normal para la que la media sea grande en comparación con la desviación típica de la misma, se asegura que la distribución concentra prácticamente toda su masa de probabilidad en  $(0, \infty)$ .
2. Es computacionalmente fácil trabajar con la distribución normal.
3. Podría darse el caso en que el modelo incluya observaciones dependientes. El modelo normal multivariante es uno de los pocos en que es fácilmente tratable la dependencia.

El más sencillo de los modelos jerárquicos, como comentamos anteriormente, es el denominado por Klugman (1992) *one-way*, que se corresponde en su versión clásica con el modelo de Bühlmann-Straub (1972). El modelo se especifica por medio de tres niveles:

- Primer nivel:  $X_{ij} \sim \mathcal{N}(\theta_i, \sigma^2/m_{ij})$ ,  $i = 1, 2, \dots, k$ ,  $j = 1, 2, \dots, n$ ,
- Segundo nivel:  $\theta_i \sim \mathcal{N}(\mu, \tau^2)$ ,  $i = 1, 2, \dots, k$ ,
- Tercer nivel:  $\pi_2(\mu, \sigma^2, \tau^2)$ .

El primer nivel consiste de  $n$  observaciones para cada uno de los  $k$  grupos. Sea  $X_{ij}$  la  $j$ -ésima observación del grupo  $i$ -ésimo. Las cantidades  $m_{ij}$  son conocidas, y aquí representan el número de siniestros. Se supone que la varianza  $\sigma^2$  es constante. En el segundo nivel,  $\theta_1, \theta_2, \dots, \theta_k$  son independientes dados  $\mu$  y  $\tau^2$ .

La metodología Bayesiana que adoptaremos en esta investigación diferirá sustancialmente de la adoptada por Klugman (1992). El nuevo modelo considerado en esta investigación es más sencillo de aplicar, en el sentido de que sólo se precisarán calcular integrales numéricas sobre  $\tau^2$ , mientras que en el modelo de Klugman (1992) es necesario elaborar complejas integrales múltiples. En este sentido se pronuncia Berger (1985), quien además señala la ventaja de que no se precisa de la manipulación de complejas matrices.

El modelo propuesto en esta investigación consta de los siguientes niveles:

- Primer nivel:  $X_i \sim \mathcal{N}(\theta_i, \sigma^2/n)$ ,  $\sigma^2$  conocida,  $i = 1, 2, \dots, p$ ,
- Segundo nivel:  $\theta_i \sim \mathcal{N}(\mu, \tau^2)$ ,
- Tercer nivel:  $\pi(\mu, \tau^2)$ .

El primer nivel indica que las variables  $X_{ij} \sim \mathcal{N}(\theta_i, \sigma^2)$ ; ahora bien nosotros consideraremos las variables aleatorias  $\bar{X}_i = \sum_{j=1}^n X_{ij}/n$  a las que denominaremos por simplicidad  $X_i$  y que obviamente seguirán una distribución  $\mathcal{N}(\theta_i, \sigma^2/n)$ . Por tanto el valor observado  $\underline{X}$  será para nosotros la media observada para cada clase, esto es

$$\underline{X} = (X_1, X_2, \dots, X_p), \quad \text{con} \quad X_i = \sum_{j=1}^n X_{ij}/n.$$

Suponemos también que la distribución conjunta  $\pi(\mu, \tau^2)$  se puede factorizar como

$$\pi(\mu, \tau^2) = \pi_1(\mu) \cdot \pi_2(\tau^2),$$

es decir, los parámetros son independientes. Para el cálculo de la media a posteriori de  $\theta$  se requiere del siguiente resultado.

**Corolario 5.1 (Berger (1985))** Si  $\pi(\mu, \tau^2) \equiv 1$  y  $p \geq 4$  (de modo que  $\pi_2(\tau^2|\underline{X})$  que aparece abajo sea propia), tomando:

$$M = \frac{\sigma^2}{\sigma^2 + \tau^2},$$

se tiene:

$$E[\theta|\underline{X}] = \underline{X} - E(M)(\underline{X} - \underline{X} \cdot \mathbf{1}),$$

y

$$\text{Var}[\theta|\underline{X}] = \sigma^2 I - \sigma^2 E(M) \left( I - \frac{1}{p} \mathbf{1} \right) + \text{Var}[M](\underline{X} - \bar{X} \cdot \mathbf{1})(\underline{X} - \bar{X} \cdot \mathbf{1})^\top,$$

donde  $E$  es la esperanza respecto a la distribución:

$$\pi_{22}(\tau|\underline{X}) = k(\sigma^2 + \tau^2)^{-(p-1)/2} \exp \left\{ -\frac{s^2}{2(\sigma^2 + \tau^2)} \right\},$$

$\mathbf{1} = (1, \dots, 1)^\top$ ,  $(\mathbf{1}) \equiv$  matriz toda de unos,  $s^2 = \sum_{i=1}^p (X_i - \bar{X})^2$  y  $k$  es la constante de normalización que resulta al integrar sobre  $\tau^2$  y normalizar.

Observemos que según el resultado del corolario anterior la media a posteriori de  $\theta$ ,  $E[\theta|\underline{X}]$ , y por tanto la prima neta Bayes  $P_B$  puede escribirse como:

$$\begin{aligned} P_B &= X_i - \frac{\int_0^\infty \frac{\sigma^2(X_i - \bar{X})}{\sigma^2 + \tau^2} \frac{\exp\{-\frac{s^2}{2(\sigma^2 + \tau^2)}\}}{(\sigma^2 + \tau^2)^{(p-1)/2}} g(\tau^2) d\tau^2}{\int_0^\infty \frac{\exp\{-\frac{s^2}{2(\sigma^2 + \tau^2)}\}}{(\sigma^2 + \tau^2)^{(p-1)/2}} g(\tau^2) d\tau^2} \\ &= X_i - \sigma^2(X_i - \bar{X}) \frac{\int_0^\infty \frac{1}{\sigma^2 + \tau^2} \frac{\exp\{-\frac{s^2}{2(\sigma^2 + \tau^2)}\}}{(\sigma^2 + \tau^2)^{(p-1)/2}} g(\tau^2) d\tau^2}{\int_0^\infty \frac{\exp\{-\frac{s^2}{2(\sigma^2 + \tau^2)}\}}{(\sigma^2 + \tau^2)^{(p-1)/2}} g(\tau^2) d\tau^2} \\ &= Z \cdot X_i + (1 - Z) \cdot \bar{X}, \end{aligned}$$

con

$$Z = 1 - \sigma^2 \frac{\int_0^\infty \frac{1}{\sigma^2 + \tau^2} \frac{\exp\{-\frac{s^2}{2(\sigma^2 + \tau^2)}\}}{(\sigma^2 + \tau^2)^{(p-1)/2}} g(\tau^2) d\tau^2}{\int_0^\infty \frac{\exp\{-\frac{s^2}{2(\sigma^2 + \tau^2)}\}}{(\sigma^2 + \tau^2)^{(p-1)/2}} g(\tau^2) d\tau^2},$$

es decir, una fórmula de credibilidad

En lo que sigue se requiere de la marginal de  $\underline{X}$  dada  $\pi_0$ , que calculamos a continuación.

$$\begin{aligned} m(\underline{X}|\pi_0) &= \int_{\Theta} f(\underline{X}|\theta) \pi_0(\theta) d\theta \\ &= \int_{\Theta} \frac{1}{(\sigma\sqrt{2\pi})^n} \exp \left\{ -\frac{\sum (X_i - \theta)^2}{2\sigma^2} \right\} \frac{1}{\tau\sqrt{2\pi}} \exp \left\{ -\frac{(\theta - a)^2}{2\tau^2} \right\} d\theta. \end{aligned}$$

A continuación, utilizando la identidad:

$$\sum_{i=1}^n (X_i - \theta)^2 = \sum_{i=1}^n (X_i - \bar{x})^2 + n(\bar{X} - \theta)^2,$$

la densidad predictiva,  $m(\underline{X}|\pi_0)$ , puede escribirse como

$$\begin{aligned} m(\underline{X}|\pi_0) &= \frac{1}{(\sigma\sqrt{2\pi})^n} \frac{1}{\tau\sqrt{2\pi}} \exp\left\{-\frac{\sum(X_i - \bar{X})^2}{2\sigma^2}\right\} \\ &\times \int_{\Theta} \exp\left\{-\frac{n\tau^2(\bar{X} - \theta)^2 + \sigma^2(\theta - a)^2}{2\sigma^2\tau^2}\right\} d\theta \\ &= k \int_{\Theta} \exp\left\{-\frac{n\tau^2(\bar{X}^2 - 2\bar{X}\theta + \theta^2) + \sigma^2(\theta^2 - 2a\theta + a^2)}{2\sigma^2\tau^2}\right\} d\theta, \end{aligned}$$

donde

$$k = \frac{1}{(\sigma\sqrt{2\pi})^n} \frac{1}{\tau\sqrt{2\pi}} \exp\left\{-\frac{\sum(X_i - \bar{X})^2}{2\sigma^2}\right\},$$

luego

$$\begin{aligned} m(\underline{X}|\pi_0) &= k \int_{\Theta} \exp\left\{-\frac{(n\tau^2 + \sigma^2)\theta^2 - (2\tau^2 n\bar{X} + 2a\sigma^2)\theta + (n\tau^2\bar{X}^2 + a^2\sigma^2)}{2\sigma^2\tau^2}\right\} \\ &= k \int_{\Theta} \exp\left\{-\frac{\theta^2 - \frac{2n\tau^2\bar{X} + 2a\sigma^2}{n\tau^2 + \sigma^2}\theta + \frac{n\tau^2\bar{X}^2 + a^2\sigma^2}{n\tau^2 + \sigma^2}}{\frac{2\sigma^2\tau^2}{n\tau^2 + \sigma^2}}\right\} d\theta \\ &= k \int_{\Theta} \exp\left\{-\frac{\left(\theta - \frac{n\tau^2\bar{X} + a\sigma^2}{n\tau^2 + \sigma^2}\right)^2 + \frac{\tau^2\bar{X}^2 + a^2\sigma^2}{n\tau^2 + \sigma^2} - \left(\frac{\tau^2\bar{X} + a\sigma^2}{n\tau^2 + \sigma^2}\right)^2}{\frac{2\sigma^2\tau^2}{n\tau^2 + \sigma^2}}\right\} d\theta \\ &= k' \int_{\Theta} \exp\left\{-\frac{\left(\theta - \frac{n\tau^2\bar{X} + a\sigma^2}{n\tau^2 + \sigma^2}\right)^2}{\frac{2\sigma^2\tau^2}{n\tau^2 + \sigma^2}}\right\} d\theta, \end{aligned}$$

con

$$k' = k \exp\left\{-\frac{\frac{n\tau^2\bar{X}^2 + a^2\sigma^2}{n\tau^2 + \sigma^2} - \left(\frac{n\tau^2\bar{X} + a\sigma^2}{n\tau^2 + \sigma^2}\right)^2}{\frac{2\sigma^2\tau^2}{n\tau^2 + \sigma^2}}\right\}. \quad (5.42)$$

Así,

$$\begin{aligned} m(\underline{X}|\pi_0) &= k' \frac{\sigma\tau\sqrt{2\pi}}{\sqrt{n\tau^2 + \sigma^2}} \int_{\Theta} \frac{1}{\frac{\sigma\tau\sqrt{2\pi}}{\sqrt{n\tau^2 + \sigma^2}}} \exp\left\{-\frac{\left(\theta - \frac{n\tau^2\bar{X} + a\sigma^2}{n\tau^2 + \sigma^2}\right)^2}{\frac{2\sigma^2\tau^2}{n\tau^2 + \sigma^2}}\right\} d\theta \\ &= k' \frac{\sigma\tau\sqrt{2\pi}}{\sqrt{n\tau^2 + \sigma^2}}. \end{aligned}$$

Sustituyendo ahora  $k'$  por (5.42) resulta,

$$\begin{aligned}
m(\underline{X}|\pi_0) &= \frac{1}{(\sigma\sqrt{2\pi})^n} \frac{\sigma}{\sqrt{\tau^2 t + \sigma^2}} \exp \left\{ \frac{\left( \frac{n\tau^2 \bar{X}^2 + a^2 \sigma^2}{n\tau^2 + \sigma^2} - \left( \frac{n\tau^2 \bar{X} + a\sigma^2}{n\tau^2 + \sigma^2} \right)^2 \right)}{\frac{2\sigma^2 \tau^2}{n\tau^2 + \sigma^2}} \right\} \\
&\times \exp \left\{ -\frac{\sum (X_i - \bar{X})^2}{2\sigma^2} \right\} \\
&= \frac{\sigma}{(\sigma\sqrt{2\pi})^n \sqrt{n\tau^2 + \sigma^2}} \exp \left\{ \frac{\left( \frac{n\tau^2 \bar{X}^2 + a^2 \sigma^2}{n\tau^2 + \sigma^2} - \left( \frac{n\tau^2 \bar{X} + a\sigma^2}{n\tau^2 + \sigma^2} \right)^2 \right)}{\frac{2\sigma^2 \tau^2}{n\tau^2 + \sigma^2}} \right\} \\
&\times \exp \left\{ -\frac{\sum (X_i - \bar{X})^2}{2\sigma^2} \right\} \\
&= \frac{\sigma}{(\sigma\sqrt{2\pi})^n \sqrt{n\tau^2 + \sigma^2}} \exp \left\{ -\frac{\sum (X_i - \bar{X})^2}{2\sigma^2} \right\} \\
&\times \exp \left\{ \frac{n [a(a - 2\bar{X}) + \bar{X}^2]}{2(n\tau^2 + \sigma^2)} \right\}.
\end{aligned}$$

A continuación desarrollaremos dos ejemplos con datos de Klugman (1992). El objetivo será obtener las primas neta Bayes, y realizar un análisis comparativo con los resultados obtenidos por Klugman (1992) y los que se obtendrían con el modelo clásico de Bühlmann-Straub (1972). En nuestro trabajo, y por simplicidad supondremos una distribución no informativa para el parámetro  $\tau^2$ , esto es,  $g(\tau^2) \equiv 1$ .

**Ejemplo 5.14** Los datos de este primer ejemplo están tomados de Klugman (1992). Los datos son simulados, con el objetivo de que conociéndose los verdaderos resultados, puedan hacerse análisis comparativos. El ejemplo consiste de  $k = 10$  grupos con  $n = 5$  observaciones para cada uno de los grupos. Los datos  $m_{ij} = 1$  para las 50 observaciones, que se encuentran en la tabla 5.17.

Puesto que el ejemplo es simulado, los verdaderos valores  $\theta_1, \theta_2, \dots, \theta_{10}$  se conocen. La observación muestral consistirá de las medias observadas para cada clase, esto es:

$$\begin{aligned}
\underline{X} &= (109,636, 101,240, 90,934, 110,048, 100,350, 101,802, 106,630, 117,604, \\
&\quad 87,472, 105,126),
\end{aligned}$$

siendo  $\bar{X} = 103,1842$ ,  $s^2 = 737,70354$ ,  $p = 10$  y  $\sigma^2 = 100/5 = 20$ .

El factor de credibilidad es 0,810226 en nuestro modelo y 0,8218 en el modelo de Klugman (1992). En el modelo clásico el valor de  $Z(n)$  es 0,8129, y la media a posteriori, primas netas Bayes, para cada una de las clases están expuestos en

Tabla 5.17: Datos escenario 1. (Klugman, 1992)

Clases	Años					$\theta_i$	$\bar{X}_i$
	1	2	3	4	5		
1	124.930	110.670	106.930	104.05	101.60	111.97	109.636
2	97.160	89.28	102.88	111.10	105.78	98.87	101.240
3	103.570	86.82	92.49	87.99	83.80	90.98	90.934
4	119.530	125.92	98.05	117.57	94.17	108.44	111.048
5	95.680	110.43	83.59	110.54	101.51	99.06	100.350
6	102.040	93.60	106.12	98.70	108.55	95.44	101.802
7	112.710	101.64	106.50	111.71	100.59	105.09	106.630
8	119.160	111.54	127.24	115.02	115.06	90.51	117.604
9	81.710	90.45	91.51	84.74	88.95	95.99	87.472
10	102.510	123.81	113.83	94.97	90.51	105.60	105.126

la Tabla 5.18, donde también aparecen las primas Bayes estimadas a partir del modelo de Bühlmann-Straub (1972).

En las tablas adjuntas se ha utilizado la siguiente notación:

- I: Modelo propuesto
- II: Modelo de Klugman (1992)
- III: Modelo de Bühlmann-Straub
- IV: Primas Bayes verdaderas

A continuación calcularemos la varianza a posteriori que nos permitirá obtener los intervalos de credibilidad. Para el cálculo de la varianza utilizamos el resultado del corolario 5.1. En nuestro caso, y puesto que,  $E[M] = 0,189774$ ,  $E[M^2] = 0,0463006$ , resulta:

$$\begin{aligned} \text{Var}[\theta|\underline{X}] &= 20I - 3,8 \left( I - \frac{1}{10}(\underline{1}) \right) \\ &+ 0,01028(\underline{X} - 103,1842 \cdot \underline{1})(\underline{X} - 103,1842 \cdot \underline{1})^\top \\ &= 16,2I + 0,38(\underline{1}) + 0,01028(\underline{X} - 103,1842 \cdot \underline{1})(\underline{X} - 103,1842 \cdot \underline{1})^\top, \end{aligned}$$

por tanto

$$\text{Var}(\theta|\underline{X}) = 16,58 + 0,01028(\underline{X} - 103,1842)^2.$$

En la tabla 5.19 se recogen las varianzas a posteriori para cada una de las clases así como los intervalos de credibilidad al 90 por ciento para nuestro modelo

Tabla 5.18: Primas Bayes.

Clase	I	II	III	IV
1	108.412	108.486	108.42	111.970
2	101.608	101.586	98.870	101.600
3	93.258	93.117	93.220	90.980
4	108.745	109.646	109.570	108.440
5	100.887	100.855	100.88	99.060
6	102.064	102.048	102.060	95.440
7	105.976	106.016	105.980	105.090
8	114.867	115.034	114.900	111.990
9	90.453	90.271	90.410	95.990
10	104.757	104.780	105.600	104.760

Tabla 5.19: Varianzas a posteriori e intervalos de credibilidad.

Clase	Varianza, I	Intervalos, I	Varianza, II	Intervalo, II
1	17.008	102.851;116.420	16.075	101.890;115.081
2	16.618	94.534;107.945	15.646	95.079;108.093
3	18.122	83.931;97.936	17.304	86.274;99.960
4	17.064	103.252;116.843	16.304	103.004;116.288
5	16.662	93.635;107.064	15.694	94.338;107.371
6	16.600	95.099;108.504	15.624	95.546;108.550
7	16.702	99.907;113.352	15.737	99.490;112.540
8	18.717	110.487;124.720	17.961	108.062;122.005
9	19.117	80.279;94.664	18.402	83.215;97.328
10	16.618	98.420;111.831	15.645	98.273;111.286

y el de Klugman (1992). A la vista de la tabla 5.18 se desprende que las primas Bayes del modelos propuesto están más próximas a las obtenidas bajo el modelo de Bühlmann-Straub (1972) para todas las clases excepto para la clase 4 que está más próximo el de Klugman (1992). Esto era de esperar, pues el factor de credibilidad en nuestro modelo está más cercano al factor de credibilidad de Bühlmann-Straub (1972) que el de Klugman (1992). Estas diferencias se ilustran en los gráficos 5.1 y 5.2.

Si comparamos con el verdadero valor del parámetro, el modelo de Klugman (1992) se acerca más a este para las clases 1, 2, 3, 5, 6 y 10 y el nuestro para las restantes

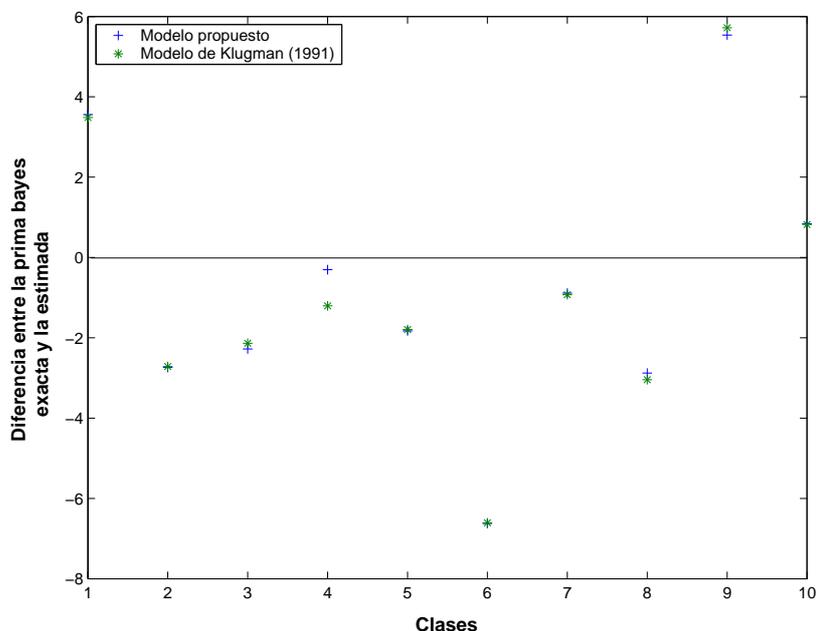


Figura 5.1: Diferencias entre la verdadera prima Bayes y la estimada en ambos modelos.

clases. A la vista de la tabla 5.19 se desprende que los intervalos de credibilidad contienen al estimador de Bühlmann-Straub en los dos modelos y para todas las clases. Los intervalos de credibilidad en nuestro modelo contienen a los verdaderos valores en todas las clases salvo para la clase 9, que es también la que presenta mayor varianza. Los intervalos de credibilidad de Klugman (1992) contienen a los verdaderos valores para todas las clases excepto para la clase 6.

**Ejemplo 5.15** Los datos de este ejemplo han sido nuevamente extraídos de Klugman (1992, pp. 125-129), y son datos reales procedentes de *National Council on Compensation Insurance*, y correspondientes a un estado Norteamericano. Se muestran en el apéndice de este texto y corresponden a 133 clases de empleos sobre un período de siete años de observación. Los datos  $Y_{ij}$  son número de siniestros que provocaron incapacidad parcial permanente.  $m_{ij}$  representa los costes o pagos graduados que fueron corregidos en términos inflacionarios para representar unidades monetarias constantes. Los datos utilizados para  $X_{ij}$  son  $Y_{ij}/m_{ij}$ , el número de siniestros relativo por cada clase y año.

Se utilizarán, igual que lleva a cabo Klugman (1992), los primeros seis años de los siete con el propósito de estimar tanto las primas netas Bayes como el

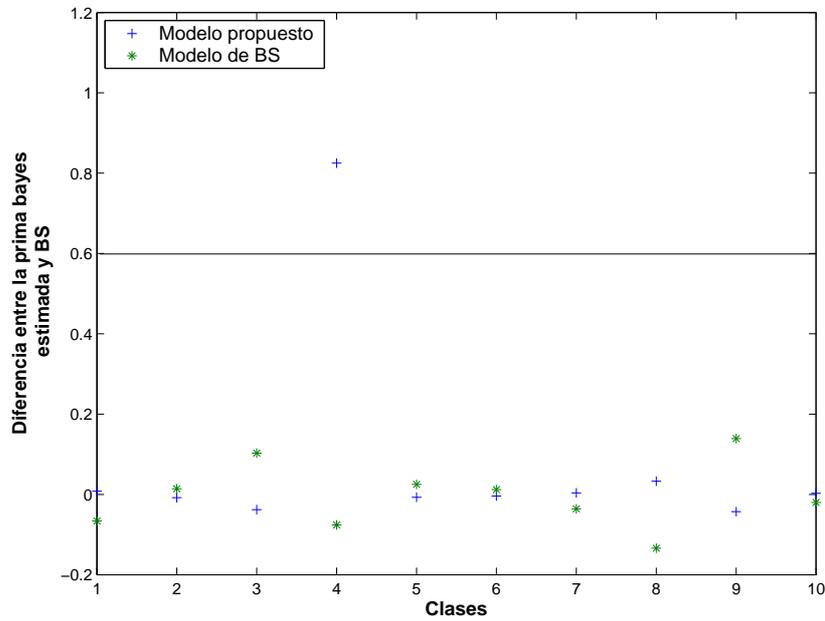


Figura 5.2: Diferencias entre la verdadera prima Bayes y el estimador de Bühlmann-Straub, BS.

número de siniestros para el séptimo año, y para las clases 4, 11, 112, 70, 20 y 89. Al igual que en el escenario 1 los datos que utilizaremos será la media para cada clase, aunque en este caso para los primeros seis años, esto es,  $X_i \equiv \bar{X}_i = (\sum_{j=1}^6 Y_{ij}/m_{ij})/n$ ,  $i = 1, 2, \dots, k = 133$ .

Hay que tener en cuenta que de las 133 clases sólo se han tomado las 130 para las que  $m_i = \sum m_{ij} > 0$ . Es decir, se excluyen las clases 7, 18 y 128. Además se excluyen también del estudio los 13 años para los que no se producen pagos. Estos son:

- Clase 4: años 1, 2, 3, 5 y 6
- Clase 54: años 1, 2, 3, 5 y 6
- Clase 61: años 1 y 2
- Clase 86: año 1

Por tanto, disponemos de un total de  $130 \times 6 - 13 = 767$  observaciones. Los tres niveles del modelo son:

- Primer nivel:  $X_i \sim \mathcal{N}(\theta_i, \sigma^2/n)$ ,  $\sigma^2 = 0.002$
- Segundo nivel:  $\theta_i \sim \mathcal{N}(\mu, \tau^2)$ ,
- Tercer nivel:  $\pi(\mu, \tau^2)$ ,  $\pi(\mu) \equiv 1$ ,  $\pi(\tau^2) \equiv 1$ .

Ahora  $p = 130$ ,  $\bar{X} = 0,05317$  y  $s^2 = 1,89251804$ ,  $Z = 0,977631$ .

Las primas netas Bayes para nuestro modelo, así como las que aparecen en Klugman (1992) y las calculadas bajo el modelo de Bühlmann-Straub (1972) aparecen recogidas en la tabla 5.20. El siguiente paso consiste en predecir el número de

Tabla 5.20: Primas Bayes.

Clase	I	II	III
4	0.0000	0.0404	0.0395
11	0.0440	0.0442	0.0434
112	0.0030	0.0019	0.0020
70	0.0012	0.0114	0.0205
20	0.0316	0.0315	0.0316
89	0.4158	0.3697	0.2989

siniestros para el séptimo año. Para el séptimo año, en la clase  $i$ , el número esperado de siniestros es  $m_{i7} \cdot E[\theta|\mathbf{x}]$ , es decir  $m_{i7}$  veces el estimador puntual a posteriori. Los valores obtenidos se recogen en la tabla 5.21. Para el cálculo de la varianza hay

Tabla 5.21: Número de siniestros (estimados).

Clase	$m_{i7}$	$y_{i7}$	I	II	III
4	0.0	0	0	0	0
11	229.83	8	10.16	10.16	7.37
112	18809.67	45	57.12	36.34	66.82
70	54.81	0	0.065	0.63	3.51
20	1315.37	22	41.85	41.45	17.04
89	79.63	40	33.11	29.42	4.23

que tener en cuenta  $E[M] = 0,0223688$ ,  $E[M^2] = 0,000508242$ , de donde resulta

$$\text{Var}[\underline{X}|\theta] = 3,26 \cdot 10^{-4} + 7,87 \cdot 10^{-6}(\underline{X} - 0,05317)^2.$$

A la vista de la tabla 5.20 observamos que la prima Bayes está más próxima a la estimada bajo el modelo de Bühlmann-Straub (1972) en nuestro modelo para

Tabla 5.22: Varianzas e intervalos de credibilidad al 90 por ciento.

Clase	$m_{i7}^2 \cdot \text{Var}[X_i \theta]$	Intervalo
4	0	0-0
11	17.22	3.33-16.98
112	115347.32	-501.56-615.80
70	0.97	-1.55-1.68
20	564.05	2.51-80.64
89	2.07	30.74-35.47

las clases 11 y 20, mientras que para el resto de las clases está más próximo el estimador de Klugman (1992). En la tabla 5.21 se observa que el número de siniestros predicho en nuestro modelo está más próximo al verdadero valor,  $y_{i7}$ , para las clases 70 y 89, mientras que el de Klugman (1992) lo está para las clases 112 y 20. Ambos modelos coinciden en las clases 4 y 11. Sin embargo el valor que prevee el modelo de Bühlmann-Straub (1972) está más cercano a nuestro modelo en la clase 112 y al de Klugman (1992) en las clases 70, 20 y 89.

Los intervalos de credibilidad, que pueden verse en la tabla 5.22 contienen al verdadero valor salvo para la clase 89 (en el modelo de Klugman (1992) ocurre igual). De nuevo el intervalo de credibilidad para la clase 89 no contiene al estimador de Bühlmann-Straub (1972).

#### 5.4.4. Un modelo basado en una función general de pérdida

El modelo propuesto en Gómez-Déniz (2008a) incluye a los modelos de Bühlmann (1967) y al modelo de Jewell (1974) como casos particulares, dando lugar a nuevas fórmulas de credibilidad dependientes de un parámetro adicional.

Hemos visto que bajo determinadas formas funcionales  $h(x)$  y mediante el procedimiento de minimizar la pérdida esperada a posteriori, la familia de funciones de pérdida:

$$L_1(a, x) = h(x)(x - a)^2, \quad (5.43)$$

bajo determinados pares de verosimilitud y distribuciones a priori da lugar a primas Bayes que pueden expresarse como fórmulas de credibilidad, es decir:

$$P_B^{L_1} = Z(n)g(\bar{\mathbf{x}}) + [1 - Z(n)]P_C^{L_1}, \quad (5.44)$$

donde se ha denotado mediante  $P_B^{L_1}$  y  $P_C^{L_1}$  las primas Bayes y colectivas que se obtienen al utilizar la función de pérdida (5.43), respectivamente.

Comprobaremos ahora que una generalización de la expresión (5.44) puede obtenerse utilizando la denominada función de pérdida ponderada y equilibrada, dada por:

$$L_2(a, x) = wh(x)(\delta_0(x) - a)^2 + (1 - w)h(x)(x - a)^2, \quad (5.45)$$

donde  $0 \leq w \leq 1$  es un factor peso determinado por el investigador,  $h(x)$  es, como en el caso de la función de pérdida (5.43), una función peso positiva y  $\delta_0(x)$  es una función de los datos.

La función de pérdida ponderada y equilibrada es una función de pérdida generalizada que fue introducida por Zellner (1994, pp.371–390) y que también aparece en Dey et al. (1999), Farsipour y Asgharzadeh (2004) y Jafari et al. (2006) cuando  $h(x) = 1$  en (5.45). Ahora, la función de pérdida (5.43) es un caso particular de (5.45), que se obtiene cuando  $w$  es igual a 0.

Además, se obtendrá una generalización del modelo de Bühlmann (1967), así como del modelo de Jewell (1974).

En esta sección se obtendrán nuevas y más ricas fórmulas de credibilidad utilizando la función de pérdida (5.45) y el principio de prima neta. Para ello se requiere del siguiente resultado.

**Teorema 5.5** *Las primas de riesgo y colectiva bajo la función de pérdida (5.45) vienen dadas por:*

$$\begin{aligned} P_R^{L_2}(\theta) \equiv P_R^{L_2} &= w \frac{E_{f(x|\theta)} [\delta_0(X)h(X)|\theta]}{E_{f(x|\theta)} [h(X)|\theta]} \\ &+ (1 - w) \frac{E_{f(x|\theta)} [Xh(X)|\theta]}{E_{f(x|\theta)} [h(X)|\theta]}, \end{aligned} \quad (5.46)$$

$$P_C^{L_2} = w\delta_0^* + (1 - w) \frac{E_\pi [P_R^{L_2}h(P_R^{L_2})]}{E_\pi [h(P_R^{L_2})]}, \quad (5.47)$$

respectivamente, donde  $\delta_0^*$  es un estimador objetivo para la prima de riesgo  $P_R^{L_2}$ .

**Demostración:** Minimizando  $E_{f(x|\theta)} [L_2(\theta, P_R^{L_2})]$  y  $E_{\pi(\theta)} [L_2(P_R^{L_2}, P_C^{L_2})]$  con respecto a  $P_R^{L_2}$  y  $P_C^{L_2}$  respectivamente, se obtiene el resultado. ■

La prima Bayes,  $P_B^{L_2}$ , se obtiene reemplazando en (5.47) la distribución a priori  $\pi(\theta)$  por la distribución a posteriori  $\pi(\theta|\underline{X})$ .

Obsérvese ahora que poniendo  $\gamma = E_\pi [P_R^{L_2}h(P_R^{L_2})] / E_\pi [h(P_R^{L_2})]$ ,

$$\begin{aligned} P_C^{L_2} &\in (\delta_0^*, \gamma), \quad \text{si } \delta_0^* < \gamma, \\ P_C^{L_2} &\in (\gamma, \delta_0^*), \quad \text{si } \delta_0^* > \gamma, \end{aligned}$$

y el mismo resultado ocurre cuando  $C$  se reemplaza por  $B$ . Por lo tanto, el actuario puede elegir el valor de  $\delta_0^*$  para obtener una prima de acuerdo a sus preferencias.

El siguiente resultado proporciona primas netas de credibilidad obtenidas cuando utilizamos la función de pérdida (5.45) y establecemos que  $h(x) = 1$ .

**Proposición 5.3** *Si la prima Bayes obtenida cuando se utiliza la función de pérdida (5.43) es una fórmula de credibilidad, entonces la prima Bayes que se obtiene al utilizar la función de pérdida (5.45) es también una fórmula de credibilidad, en la forma:*

$$P_B^{L2} = Z(n)l(P_C^{L1}) + [1 - Z(n)]l(\bar{X}), \quad (5.48)$$

con  $Z(n) \in [0, 1]$  y

$$l(x) = (1 - w)^2x + w(1 - w)E_{\pi(\theta|\underline{X})} [E_{f(x|\theta)} (\delta_0(X|\theta))] + w\delta_0^*.$$

**Demostración:** Utilizando (5.46) y (5.47) con  $h(x) = 1$  tenemos:

$$P_R^{L2} = wE[\delta_0(X)|\theta] + (1 - w)E_{f(x|\theta)} [X|\theta]$$

y

$$\begin{aligned} P_C^{L2} &= w\delta_0^* + (1 - w)E_{\pi(\theta)} [wE_{f(x|\theta)} [\delta_0(X)|\theta] + (1 - w)E_{f(x|\theta)} [X|\theta]] \\ &= w\delta_0^* + w(1 - w)E_{\pi(\theta)} \{E_{f(x|\theta)} [\delta_0(X|\theta)]\} \\ &\quad + (1 - w)^2E_{\pi(\theta)} [E_{f(x|\theta)} \delta_0(X)|\theta]. \end{aligned}$$

Por lo tanto,

$$P_B^{L2} = w\delta_0^* + w(1 - w)E_{\pi^*(\theta)} \left\{ E_{f(x|\theta)} [\delta_0(X)|\theta] + (1 - w)^2P_B^{L1} \right\}.$$

Ahora, si  $P_B^{L1}$  es una fórmula de credibilidad de la forma

$$P_B^{L1} = Z(n)P_C^{L1} + [1 - Z(n)]\bar{X},$$

entonces

$$\begin{aligned} P_B^{L2} &= w\delta_0^* + w(1 - w)E_{\pi(\theta|\underline{X})} \{E_{f(x|\theta)} [\delta_0(X)|\theta]\} \\ &\quad + (1 - w)^2 \left\{ Z(n)P_C^{L1} + [1 - Z(n)]\bar{X} \right\} \\ &= Z(n) \left[ (1 - w)^2P_C^{L1} + w(1 - w)E_{\pi(\theta|\underline{X})} \{E_{f(x|\theta)} [\delta_0(X)|\theta]\} + w\delta_0^* \right] \\ &\quad + [1 - Z(n)] \left[ (1 - w)^2\bar{X} + w\delta_0^* + w(1 - w)E_{\pi(\theta|\underline{X})} \{E_{f(x|\theta)} [\delta_0(X)|\theta]\} \right] \\ &= Z(n)l(P_C^{L1}) + [1 - Z(n)]l(\bar{X}). \end{aligned}$$

■

**Corolario 5.2** Si suponemos que la verosimilitud es un miembro de la familia exponencial de distribuciones y que la distribución a priori es la natural conjugada, entonces la prima Bayes obtenida cuando se utiliza la función de pérdida (5.45) es una fórmula de credibilidad exacta.

**Demostración:** Es inmediato, teniendo en cuenta que bajo la función de pérdida  $L(x, a) = (x - a)^2$  y las hipótesis de verosimilitud perteneciente a la familia exponencial de distribuciones y la natural conjugada como distribución a priori, la prima Bayes es una fórmula de credibilidad exacta. ■

Obsérvese que el factor de credibilidad en la proposición 5.3 es el mismo que el factor de credibilidad obtenido bajo la función de pérdida  $L(x, a) = (x - a)^2$ .

El siguiente resultado es consecuencia inmediata de aplicar el corolario 5.2 a los pares de verosimilitudes y distribuciones a priori que se mencionan.

**Corolario 5.3** La prima Bayes obtenida al utilizar la función de pérdida (5.45) puede reescribirse como una fórmula de credibilidad cuando se utilizan los pares de verosimilitudes y distribuciones a priori siguientes: Poisson-gamma, binomial negativa-beta, binomial-beta, normal-normal y gamma-gamma.

**Ejemplo 5.16** Considérense de nuevo los datos normalizados sobre daños de huracanes que aparecen en Pielke y Landsea (1998) y supongamos que los mismos siguen una distribución gamma de parámetros  $\theta > 0$  y  $\nu > 0$ , conocido. Suponiendo una distribución a priori  $\mathcal{G}(a, b)$ ,  $a > 0$ ,  $b > 0$  para el parámetro  $\theta$ , vamos a estimar por el método de los momentos los valores de los parámetros de la distribución a priori y utilizarlos para calcular las primas netas Bayes obtenidas bajo la función de pérdida ponderada y equilibrada. Elegimos:

1. Como valor de  $\delta_0(X) = \delta_0(X|\theta)$  la moda de la verosimilitud.
2. Como valor de  $w$ : 0; 0,15; 0,30; 0,45; 0,60; 0,75; 0,90.
3. Como valor de  $\nu$ : 10; 20; 30; 50; 100.

La verosimilitud y la distribución a priori son, respectivamente:

$$f(x|\theta) = \frac{\theta^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\theta x}, \quad x > 0, \theta > 0, \nu > 0, \quad (5.49)$$

$$\pi(\theta) = \frac{a^b}{\Gamma(b)} \theta^{b-1} e^{-a\theta}, \quad \theta > 0, a > 0, b > 0. \quad (5.50)$$

La distribución marginal de  $X$  se calcula como:

$$f(x) = \int_0^\infty \frac{\theta^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\theta x} \frac{a^b}{\Gamma(b)} \theta^{b-1} e^{-a\theta} = \frac{a^b \Gamma(\nu + b)}{\Gamma(\nu) \Gamma(b)} \frac{x^{\nu-1}}{(a+x)^{\nu+b}}.$$

Se trata de una distribución de Pearson tipo VII con parámetros  $a$ ,  $b$  y  $\nu$ . La media y varianza de esta distribución vienen dadas por:

$$E[X] = \frac{\nu a}{b-1}, \quad b > 1,$$

$$\text{Var}[X] = \frac{a^2 \nu (b-1) + 2\nu^2 a^2}{(b-1)^2 (b-2)}, \quad b > 2.$$

Los estimadores por momentos, obtenidos a partir de la media y la varianza de los datos, aparecen en la tabla 5.23. Las primas colectiva y Bayes, para el modelo

Tabla 5.23: Modelo ajustado mediante método de los momentos.

Caso	$\nu$	$a$	$b$
1	10	3231,54	3,7502
2	20	1551,13	3,6402
3	30	1020,48	3,6054
4	50	605,908	3,5783
5	100	300,605	3,5583
6	1000	29,8522	3,5406

gamma-gamma, son:

$$P_C = \frac{\nu a}{b-1}, \quad b > 1,$$

$$P_B = \frac{\nu(a + n\bar{X})}{b + n\nu - 1}, \quad b + n\nu > 1.$$

Para el cálculo de las primas pedidas se requiere de

$$E[\delta_0(X|\theta)] = \frac{(\nu-1)a}{b-1}, \quad b > 1.$$

Por lo tanto,

$$E_{\pi(\theta|\underline{X})}[\delta_0(X|\theta)] = \frac{(\nu-1)(a + n\bar{X})}{b + n\nu - 1}.$$

Finalmente, utilizando (5.48) se obtienen los valores de las primas que aparecen en la tabla 5.24.

### Una generalización del modelo de Bühlmann

Según lo estudiado en secciones precedentes, bajo el modelo de Bühlmann (1967) la prima Bayes puede escribirse como una fórmula de credibilidad con el

Tabla 5.24: Primas Bayes basadas en la pérdida cuadrática ponderada.

$w$	0,0	0,15	0,3	0,45	0,6	0,75	0,9
1	11749.9	11423.9	11150.7	10930.4	10762.9	10648.4	10586.7
2	11749.9	11586.9	11450.3	11340.1	11256.4	11199.1	11168.3
3	11749.9	11641.2	11550.2	11476.7	11420.9	11382.7	11362.2
4	11749.9	11684.7	11630.1	11586.0	11552.5	11529.6	11517.3
5	11749.9	11717.3	11690.0	11668.0	11651.2	11639.8	11633.6
6	11749.9	11746.7	11743.9	11741.7	11740.1	11738.9	11738.3

factor de credibilidad verificando la fórmula:

$$Z(n) = \frac{n \operatorname{Var} E[X|\theta]}{n \operatorname{Var} E(X|\theta) + E \operatorname{Var}[X|\theta]}. \quad (5.51)$$

Hemos probado además que este factor coincide con el obtenido por Jewell (1974). En esta sección generalizaremos el resultado de Bühlmann (1967) utilizando la función de pérdida (5.45).

La idea central es reemplazar la prima teórica (desconocida) por una expresión lineal de la forma  $c_0 + \sum_{s=1}^n c_s X_s$  dependiente de la experiencia de reclamaciones  $X_i$ ,  $i = 1, \dots, n$ , usando la función de pérdida (5.45). Para ello supondremos que las variables  $X_1|\theta, X_2|\theta, \dots, X_n|\theta$  son independientes e idénticamente distribuidas. Se utilizará la siguiente notación:

$$\begin{aligned} \mu(\theta) &= w E[\delta_0(X)|\theta] + (1-w) E(X_s|\theta), \\ m &= w \delta_0^* + (1-w) \{w E[E(\delta_0(X)|\theta)] + (1-w) E[E(X_s|\theta)]\}, \\ a &= \operatorname{Var}[E(X_s|\theta)], \\ s^2 &= E[\operatorname{Var}(X_s|\theta)], \end{aligned} \quad (5.52)$$

que ha sido tradicionalmente usada en la literatura actuarial (ver Bühlmann, 1967 y 1969, Herzog, 1996 y Goovaerts et al., 1990).

Los coeficientes  $c_i$ ,  $i = 0, \dots, n$  deberán calcularse, atendiendo al criterio de los mínimos cuadrados (reemplazando la pérdida cuadrática por la pérdida general (5.45) como:

$$\min_{c_i} E \left[ w \left( \delta_0(X) - c_0 - \sum_{s=1}^n c_s X_s \right)^2 + (1-w) \left( \mu(\theta) - c_0 - \sum_{s=1}^n c_s X_s \right)^2 \right].$$

Calculando las correspondientes derivadas parciales e igualando a cero, obten-

emos el sistema de ecuaciones:

$$c_0 = wE[\delta_0(X)] + (1-w)E[\mu(\theta)] - \sum_{s=1}^n c_s E(X_s),$$

$$c_0 E(X_r) = wE(X_r \delta_0(X)) + (1-w)E[X_r \mu(\theta)] - \sum_{s=1}^n c_s E(X_r X_s), \quad r = 1, \dots, n,$$

que equivale al sistema:

$$\sum_{s=1}^n c_s \text{Cov}(X_r, X_s) = (1-w)\text{Cov}[X_r, \mu(\theta)] + w\text{Cov}[X_r, \delta_0(X)],$$

$$r = 1, 2, \dots, n. \quad (5.53)$$

Teniendo en cuenta ahora:

$$\begin{aligned} \text{Cov}(X_r, X_s) &= E[\text{Cov}(X_r, X_s|\theta)] + \text{Cov}[E[X_r|\theta], E[X_s|\theta]] \\ &= E(0) + \text{Var}\left[\frac{\mu(\theta) - wE(\delta_0(X))}{1-w}\right] = \text{Var}[\mu^*(\theta)] = a^*, \end{aligned}$$

donde:

$$\mu^*(\theta) = \frac{\mu(\theta) - wE[\delta_0(X)]}{1-w},$$

$$\text{Var}[X_r] = E[\text{Var}(X_r|\theta)] + \text{Var}[E(X_r|\theta)] = s^2 + a,$$

y

$$\begin{aligned} \text{Cov}[X_r, \mu(\theta)] &= E[\text{Cov}(X_r, \mu(\theta)|\theta)] + \text{Cov}[E[X_r|\theta], E[\mu(\theta)|\theta]] \\ &= E(0) + \text{Cov}[E(X_r|\theta), E(\mu(\theta)|\theta)] \\ &= \frac{\text{Var}[\mu(\theta)] - w\text{Cov}[\delta_0(X), \mu(\theta)]}{1-w}, \end{aligned}$$

el sistema (5.53) se reduce a:

$$cs^2 + a^*cn = \text{Var}[\mu(\theta)] - w\text{Cov}[\delta_0(X), \mu(\theta)] + w\text{Cov}[X_r, \delta_0(X)],$$

de donde obtenemos:

$$c = \frac{\text{Var}[\mu(\theta)] + w\{w\text{Cov}[X_r, \delta_0^*] - \text{Cov}[\delta_0(X), \mu(\theta)]\}}{s^2 + a^*n}.$$

y por lo tanto:

$$\begin{aligned} H[\mu(\theta)|X_1, X_2, \dots, X_n] &= c_0 + cn\bar{X} = wE(\delta_0(X)) + (1-w)E[\mu(\theta)] \\ &\quad - cnE(X_s) + cn\bar{X} \\ &= wE[\delta_0(X)] + (1-w)E[wE(\delta_0(X)) \\ &\quad + (1-w)E(X|\theta)] - cnE(X) + cn\bar{X}. \end{aligned} \quad (5.54)$$

Si suponemos ahora que  $\delta_0^* = E[\delta_0(X)]$ , utilizando (5.52), la expresión (5.54), después de poner  $c^* = c/(1-w)^2$ , puede reescribirse como:

$$\begin{aligned} H[\mu(\theta)|X_1, X_2, \dots, X_n] &= m - \left\{ \frac{m - w\delta_0^*}{(1-w)^2} - \frac{wcnE[E(\delta_0(X|\theta))]}{1-w} \right\} + cn\bar{X} \\ &= m - c^*(m - w\delta_0^*)n + w(1-w)E[E(\delta_0(X|\theta))]nc^* \\ &\quad + c^*n(1-w)^2\bar{X} \\ &= (1 - c^*n)m + \{w\delta_0^* + w(1-w)E[E(\delta_0(X|\theta))]\} \\ &\quad + (1-w)^2\bar{X}\}c^*n. \end{aligned}$$

Para finalizar, denotando mediante  $c^*n$  por  $Z(n)$  se tiene:

$$H[\mu(\theta)|X_1, X_2, \dots, X_n] = Z(n)l(\bar{X} + [1 - Z(n)]l(P_C^{L^2})).$$

Obsérvese que si  $w = 0$  entonces  $\mu^*(\theta) = \mu(\theta)$ ,  $a^* = a$  y por lo tanto:

$$Z(n) = a/(s^2 + na),$$

que corresponde con el factor de credibilidad que se obtiene usando la pérdida (5.43) y por lo tanto con el factor de credibilidad de Bühlmann (1967) y el de Jewell (1974).

El siguiente resultado es consecuencia inmediata de aplicar el corolario 5.2 a los pares de verosimilitudes y distribuciones a priori que se mencionan.

**Corolario 5.4** *La prima Bayes obtenida al utilizar la función de pérdida (5.45) puede reescribirse como una fórmula de credibilidad cuando se utilizan los pares de verosimilitudes y distribuciones a priori siguientes: Poisson-gamma, binomial negativa-beta, binomial-beta, normal-normal and gamma-gamma.*

### 5.4.5. Modelos de credibilidad basados en robustez

Como hemos visto en los apartados anteriores, la teoría moderna de la credibilidad puede interpretarse como un proceso Bayesiano en el que las medidas de riesgo se actualizan en la medida en que se obtiene la experiencia del riesgo. Los primeros investigadores actuariales, como Margolin (1975) se oponían abiertamente a esta metodología. Sin embargo, otros ya señalaban (Kahn, 1975) que el procedimiento de tarificación a posteriori (metodología Bayes) constituye uno de los mejores escenarios en los que la aproximación Bayesiana, esto es el uso de probabilidades a priori subjetivas, no resulta apropiada.

Sin embargo, la mayoría de los investigadores, entre los que cabe citar a Bühlmann (1967), Freifelder (1974), Klugman (1992) y Makov (2001), defienden abiertamente la aproximación Bayesiana. A pesar de esta vieja controversia entre estadística clásica y Bayesiana, sería demasiado atrevido asegurar que existen dos escuelas en

teoría de la credibilidad, la clásica y la Bayesiana, sino que existe una separación que se debe en parte a las distintas situaciones prácticas en que nos encontremos. En algunas ocasiones los juicios subjetivos no pueden evitarse; por ejemplo cuando una compañía introduce una nueva clase de cobertura. En este caso, el actuario deberá realizar una valoración inicial del riesgo en base al conocimiento de otros riesgos similares. La prima obtenida inicialmente se irá ajustando en la medida de que se disponga de experiencia de siniestralidad.

La crítica que surgió hace ya bastantes décadas en relación con la metodología Bayesiana, y no sólo en el ámbito actuarial, provocó en los años 80 multitud de trabajos en el escenario de lo que ha venido denominándose análisis de robustez o sensibilidad Bayesiano.

Consiste éste en admitir que el conocimiento que el investigador tiene sobre la distribución a priori no es completo y admitir cualquier distribución a priori parecida (próxima) a una inicial en la que el investigador tiene cierta seguridad.

La idea de que el investigador no tiene conocimiento exacto sobre la distribución a priori puede venir marcada por el hecho de que la decisión sobre la elección de la distribución a priori deba ser tomada por dos decisores que no se pongan de acuerdo. Por ejemplo, sería difícil distinguir entre la distribución normal y la de Cauchy, ya que son muy parecidas. En el gráfico 5.3 aparecen representadas algunas densidades tipo normal y de Cauchy con parámetros de localización y escala. Excepto por las colas, se trata de distribuciones muy similares.

En otra posible situación, el investigador confía en que la distribución a priori es unimodal pero no está seguro si utilizar (para el caso continuo) una distribución gamma o una distribución inversa gaussiana.

La solución que la estadística Bayesiana robusta ha encontrado para estas situaciones de incertidumbre ha sido considerar clases de distribuciones de probabilidad que contengan, por un lado la distribución a priori inicial en la que se confía inicialmente y, por otro lado, alguna característica en la que confíe el investigador; o bien, esta última característica solamente. Sobre dicha clase se calculará el rango de variación de la magnitud a posteriori de interés, de modo que un rango pequeño señalará que el investigador puede garantizar de modo objetivo su especificación a priori inicial. Un rango grande dará señales de carencia de robustez y, por tanto, de prudencia en las decisiones tomadas a partir de la magnitud a posteriori calculada. El lector interesado en la estadística Bayesiana robusta aplicada en las ciencias actuariales puede consultar a Calderín et al. (2007), Gómez-Déniz et al. (1999a, 1999b, 2000, 2002a, 2002b, 2005a, 2005b, 2006c y 2008d), Makov (1995) y Ríos et al. (1999)

En líneas generales, y en el ámbito actuarial, el planteamiento es el siguiente, dada la clase

$$\Gamma = \{\text{Distribuciones a priori del parámetro } \theta \in \Theta\},$$

a la que pertenece la distribución a priori asignada al parámetro,  $\pi_0(\theta) \in \Gamma$ , y dada la observación muestral  $\underline{X} = (X_1, \dots, X_n)$ , se trata de calcular los extremos de la

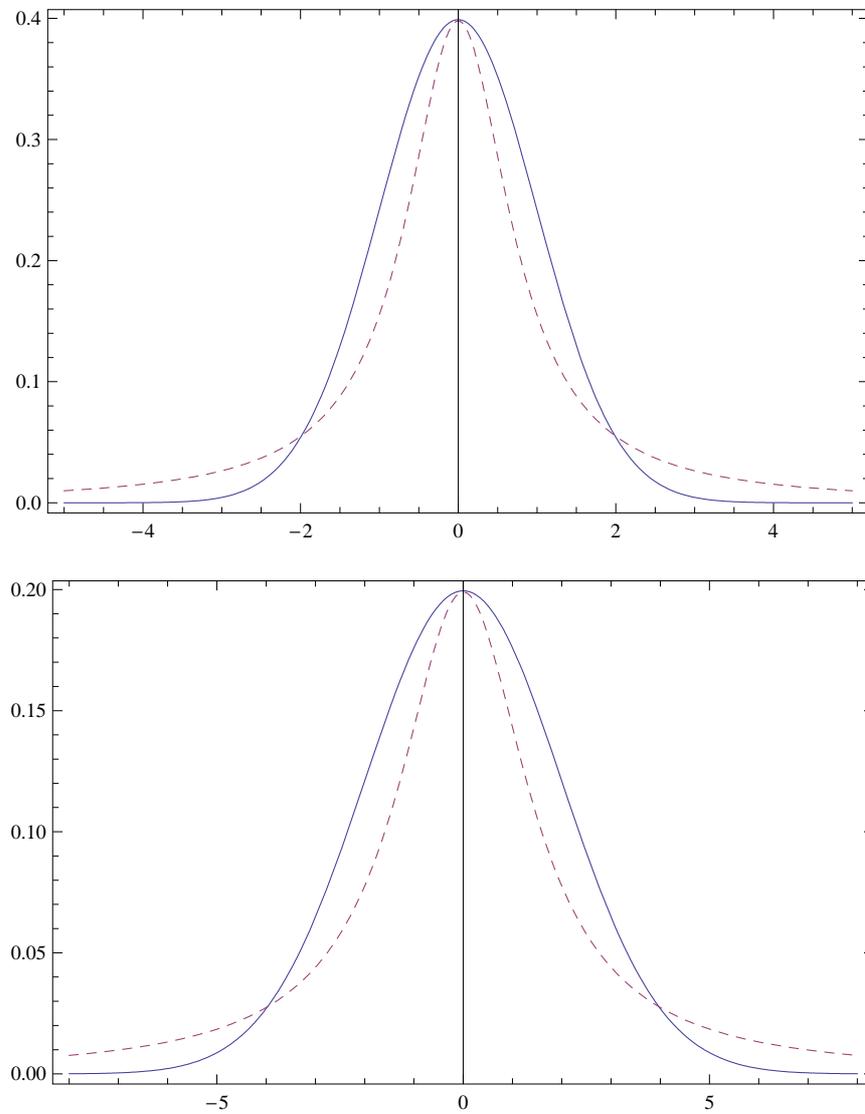


Figura 5.3: Función de densidad normal (trazo continuo) y función de densidad de Cauchy (trazo discontinuo) para los valores: gráfico superior  $\mu = 0$  y  $\sigma = 1$  para la normal y  $\sigma = 0,8$  para la Cauchy y gráfico inferior  $\mu = 0$  y  $\sigma = 2$  para la normal y  $\sigma = 1,6$  para la Cauchy.

medida de riesgo Bayes sobre la clase  $\Gamma$ . Si la magnitud a posteriori (Bayes) de interés es  $\varrho_B$ , el investigador deberá examinar el comportamiento de esa magnitud cuando se incorpora incertidumbre sobre la distribución a priori. Especificada una clase de distribuciones a priori plausibles, denotada por  $\Gamma$ , el análisis se dirige a calcular las oscilaciones de  $\varrho_B$  en esa clase, cuantificando esas oscilaciones por el intervalo:

$$\left( \inf_{\pi \in \Gamma} \varrho_B, \sup_{\pi \in \Gamma} \varrho_B \right).$$

La lectura de este intervalo da una idea al usuario de cuánto pueden variar sus conclusiones si su incertidumbre sobre la distribución a priori está bien modelizada por  $\Gamma$ .

Dos modelos desarrollados basados en esta idea, y que conducen a nuevas fórmulas de credibilidad, serán analizados en las dos secciones siguientes. Se trata de las aproximaciones gamma-minimax (Eichenauer et al., 1988) y posterior regret gamma-minimax.

### Modelo gamma-minimax

La metodología Bayes basa su decisión final en el objetivo de minimizar la pérdida esperada con respecto a una distribución a priori. Parece razonable admitir que el investigador no tiene por qué tener un conocimiento exacto sobre la distribución a priori. Se han propuesto diversas respuestas a este planteamiento, entre las que figura el principio gamma-minimax. Este principio restringe el conjunto de acciones Bayes posibles a aquellas que sean compatibles con la información de que se dispone.

El primer, y hasta el momento, el único trabajo referido a fórmulas de credibilidad utilizando la aproximación gamma-minimax se debe a Eichenauer et al. (1988). La idea, que se expone aquí de manera escueta, responde a lo siguiente.

Sea un riesgo  $X$  con función de densidad  $f(x|\theta)$  dependiente de un parámetro de riesgo  $\theta$ . Sabemos, y nos restringiremos al principio de prima neta, que si el actuario elige la prima Bayes  $P_B$  ha de ser consciente de que el riesgo asumido es:

$$R(\theta, P_B) = \int L(x, P_B) f(x|\theta) dx,$$

en la que la integral se cambiará por la sumatoria para el caso en el que  $X$  sea discreta. Introduzcamos ahora las siguientes definiciones que será útiles en adelante:

**Definición 5.1** *La cantidad*

$$r(\pi, P(\theta)) = \int R(\theta, P(\theta)) \pi(\theta|\underline{X}) d\theta$$

*se denomina riesgo Bayes en un problema de decisión Bayesiano con experimentación.*

**Definición 5.2** Dada la distribución a priori  $\pi(\theta)$ , una acción  $\delta^* \in \Delta$ , donde  $\Delta$  es el conjunto de acciones posibles, es Bayes si

$$r(\pi, \delta^*) = \inf_{\delta \in \Delta} r(\pi, \delta).$$

Supongamos ahora que el actuario no tiene conocimiento exacto sobre la distribución a priori y sólo es capaz de establecer que pertenece a cierta clase de distribuciones de probabilidad  $\Gamma$ . Por ejemplo, el actuario podría tener conocimiento de que la distribución a priori es unimodal y considerar la clase de distribuciones de probabilidad unimodales.

**Definición 5.3** Dada la distribución a priori  $\pi(\theta)$ , una acción  $\delta^* \in \Delta$ , donde  $\Delta$  es el conjunto de acciones posibles, es gamma-minimax si

$$\sup_{\pi \in \Gamma} r(\pi, \delta^*) = \inf_{\delta \in \Delta} \sup_{\pi \in \Gamma} r(\pi, \delta).$$

Presentamos ahora el siguiente resultado debido a Eichenauer et al. (1988), planteado bajo el modelo de verosimilitud gamma y distribución a priori gamma y el principio de prima neta.

**Teorema 5.6** Sea un riesgo  $X$  con distribución gamma con parámetros  $\theta > 0$ ,  $\nu > 0$  y  $\pi_1(\theta)$  la distribución (a priori del parámetro  $\theta$ ) gamma con parámetros  $a > 0$ ,  $b > 0$ , siendo

$$a = \frac{cd}{\nu(d - c^2)}, \quad (5.55)$$

$$b = \frac{2d - c^2}{d - c^2}. \quad (5.56)$$

Entonces, dadas

$$P_B = \frac{cd + n\nu(d - c^2)\bar{X}}{d + n\nu(d - c^2)}, \quad (5.57)$$

donde  $n$  es el número de períodos de tiempo de observación y  $\bar{X}$  es la media muestral, y la clase de distribuciones a priori

$$\Gamma = \{ \pi(\theta) : E_\pi(P(\theta)) \geq c, E_\pi(P(\theta)^2) \leq d, c > 0, d > 0, d > c^2 \},$$

se tiene:

1.  $P_B$  es una prima neta Bayes con respecto a  $\pi_1(\theta)$ .
2.  $P_B$  es una prima neta gamma-minimax.
3.  $P_B$  es una fórmula de credibilidad.

**Demostración:** Procederemos de la siguiente manera. La verosimilitud y la distribución a priori son como en (5.49) y (5.50), respectivamente.

Como ya sabemos, la distribución a posteriori del parámetro  $\theta$  dada una muestra de tamaño  $n$  con media  $\bar{X}$  es de nuevo una distribución gamma con los parámetros actualizados,  $a + n\bar{X}$  y  $b + n\nu$ .

1. Se tiene que la prima neta de riesgo viene dada por  $P_R = E[X|\theta] = \nu/\theta$ , mientras que la prima neta colectiva es  $P_C = E[P(\theta)] = \nu a/(b-1)$ ,  $b > 1$ . La prima neta Bayes es  $P_B = \nu(a + n\bar{X})/(b + n\nu - 1)$ ,  $b > n\nu - 1$ . Reemplazando ahora  $a$  y  $b$  por las expresiones (5.55) y (5.56) se obtiene (5.57).
2. Ahora tenemos:

$$\begin{aligned}
 R(P(\theta), P_B) &= E_{f(x|\theta)} \left\{ \left[ \frac{\nu}{\theta} - \frac{\nu(a + n\bar{X})}{b + n\nu - 1} \right]^2 \right\} \\
 &= \frac{\nu^2}{\theta^2} + \frac{\nu^2}{(b + n\nu - 1)^2} E_{f(x|\theta)} (a^2 + 2at\bar{X} + n^2\bar{X}^2) \\
 &\quad - \frac{2\nu^2}{\theta(b + n\nu - 1)} E_{f(x|\theta)} (a + n\bar{X}) \\
 &= \frac{\nu^2}{\theta^2} + \frac{\nu^2}{(b + n\nu - 1)^2} \left[ a^2 + 2an\frac{\nu}{\theta} + \frac{n(\nu + n\nu^2)}{\theta^2} \right] \\
 &\quad - \frac{2\nu^2}{\theta(b + n\nu - 1)} \left( a + \frac{n\nu}{\theta} \right) \\
 &= \frac{\nu^2}{(b + n\nu - 1)^2} \left\{ \frac{1}{\theta^2} [n\nu + (b-1)^2] - \frac{1}{\theta} [2a(b-1) + a^2] \right\}.
 \end{aligned}$$

Mientras que:

$$\begin{aligned}
 r(\pi, P_B) &= E_\pi [R(P(\theta), P_B)] \\
 &= \frac{1}{(b + n\nu - 1)^2} \{ [n\nu + (b-1)^2] E_{\pi(\theta)} [P(\theta)^2] \\
 &\quad - 2\nu a(b-1) E_{\pi(\theta)} [P(\theta)] + \nu^2 a^2 \}.
 \end{aligned}$$

Por otro lado, tenemos:

$$\begin{aligned}
 E_{\pi_1} [P(\theta)] &= \frac{\nu a}{b-1}, \quad b > 1, \\
 E_{\pi_1} [P(\theta)^2] &= \frac{\nu^2 a^2}{(b-1)(b-2)}, \quad b > 2.
 \end{aligned}$$

Se trata de probar que

$$\sup_{\pi \in \Gamma} r(\pi, P_B) < \sup_{\pi \in \Gamma} r(\pi, \delta), \quad \forall \delta \in \Delta,$$

pero  $r(\pi_1, P_B) \geq r(\pi, P_B)$ ,  $\pi \in \Gamma$ , mientras que  $r(\pi, P_B) \leq r(\pi, \delta)$ ,  $\delta \in \Delta$ , por ser  $P_B$  acción Bayes. Luego,

$$r(\pi, P_B) \leq r(\pi_1, P_B) \leq r(\pi_1, \delta) \leq \sup_{\pi \in \Gamma} r(\pi, \delta).$$

Por lo tanto,  $\sup_{\pi \in \Gamma} r(\pi, P_B) \leq \sup_{\pi \in \Gamma} r(\pi, \delta)$ .

3. La prima neta colectiva bajo  $\pi_1(\theta)$  es  $P_C = c$  y la prima neta Bayes (5.57) puede ahora reescribirse como:

$$P_B = Z(n)\bar{X} + [1 - Z(n)]P_C,$$

donde el factor de credibilidad es

$$Z(n) = \frac{n\nu(d - c^2)}{d + n\nu(d - c^2)}.$$

Resulta conveniente resaltar que si el actuario conoce perfectamente la distribución a priori, la acción que debe elegir es la prima Bayes. Si no la conoce totalmente pero conoce alguna particularidad sobre la misma elegirá la prima gamma-minimax, como se detalla en la ilustración 5.4. Esto no es descabellado

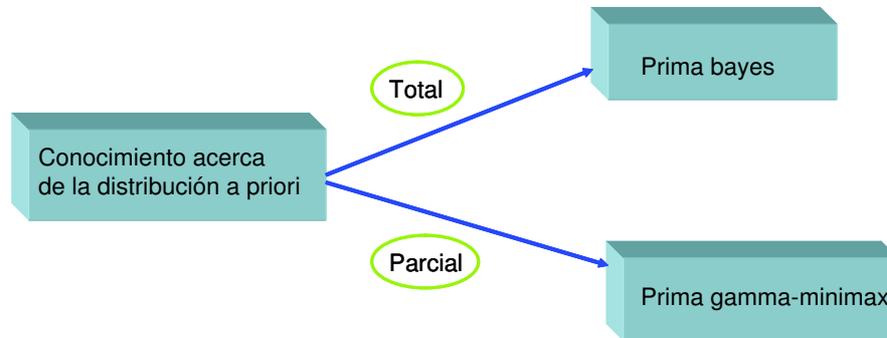


Figura 5.4: Ilustración de la acción a elegir ante certeza o imprecisión en la distribución a priori.

puesto que podemos desconocer para un determinado riesgo, un asegurado o un grupo de ellos, la media del número de reclamaciones por ejemplo, pero conocer en cierta medida como se distribuye la misma. Por ejemplo, si tiene determinada moda. Así, en seguro de automóviles el valor más frecuente del número medio de

reclamaciones es cero, y por tanto la moda es cero. Luego, si el número de reclamaciones se modela mediante la distribución de Poisson, parece lógico elegir como distribución a priori del parámetro una clase que contenga a las distribuciones unimodales.

En el resultado que aparece en el teorema 5.6 la clase  $\Gamma$  recoge información acerca de características del parámetro de riesgo en el colectivo de asegurados, la cartera de seguros.

### Modelo posterior regret-gamma minimax

Una aproximación alternativa al método Bayes y a la metodología gamma-minimax, y también conectada con el análisis de robustez, la constituye la aproximación posterior regret gamma-minimax. Admitimos de nuevo que el actuario no conoce exactamente la distribución a priori, pero es capaz de dar cierta clase  $\Gamma$  de distribuciones a priori posibles.

Si  $\rho(\pi^x, P)$  es la pérdida esperada a posteriori de una acción  $P$ , perteneciente al conjunto de acciones posibles  $\Delta$ , obtenida cuando se utiliza como distribución a posteriori  $\pi(\theta|\underline{X})$ , la acción posterior regret de  $P$ , que será notada como  $RP_B$ , se define (ver Ríos et al. (1995) y Zen y DasGupta (1993)) como:

$$r(\pi(\theta|\underline{X}), P) = \rho(\pi(\theta|\underline{X}), P) - \rho(\pi(\theta|\underline{X}), RP_B),$$

que mide la pérdida de optimalidad por elegir  $P$  en lugar de la acción óptima  $RP_B$ .

Ahora se tiene la siguiente definición.

**Definición 5.4**  $RP_B \in \Delta$  es una acción posterior regret gamma-minimax si

$$\inf_{P \in \Delta} \sup_{\pi \in \Gamma} r(\pi(\theta|\underline{X}), P) = \sup_{\pi \in \Gamma} r(\pi(\theta|\underline{X}), RP_B).$$

Es fácil deducir (ver Ríos et al. (1995) y Zen y DasGupta (1993)) que bajo pérdida cuadrática o principio de prima neta, la acción posterior regret gamma-minimax es el punto medio del intervalo  $[\inf_{\pi \in \Gamma} P_B, \sup_{\pi \in \Gamma} P_B]$ , esto es:

$$RP_B = \frac{1}{2} \left( \inf_{\pi \in \Gamma} P_B + \sup_{\pi \in \Gamma} P_B \right). \quad (5.58)$$

Esta acción es óptima en el sentido que minimiza el máximo de la función de riesgo sobre la clase de distribuciones  $\Gamma$ . Por lo tanto, el investigador no tendrá problemas asegurándose el menor valor del máximo riesgo que pueda obtener.

Ahora la situación para el caso en que no se conozca exactamente la distribución a priori permite decidir entre la acción gamma-minimax y la acción posterior regret gamma-minimax, como se muestra en la figura 5.5.

Supongamos para fijar ideas, que la verosimilitud es un miembro de la familia exponencial de distribuciones (5.19) y que la distribución a priori sobre el

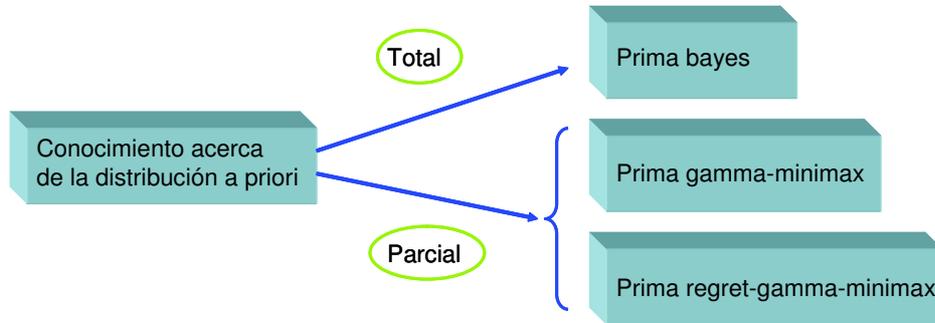


Figura 5.5: Ilustración de la acción a elegir ante certeza o imprecisión en la distribución a priori.

parámetro  $\theta$  es la conjugada natural (5.20), pero que el investigador es incapaz de especificarla completamente. De ahí que use la siguiente clase de distribuciones a priori:

$$\Gamma_1 = \left\{ \pi(\theta) : x_0^{(1)} \leq x_0 \leq x_0^{(2)}, n_0 \text{ conocido} \right\}, \quad (5.59)$$

$$\Gamma_2 = \left\{ \pi(\theta) : \gamma_1 \leq m \leq \gamma_2, n_0 \text{ conocido}, m = \frac{x_0}{n_0} \right\}, \quad (5.60)$$

$$\Gamma_3 = \left\{ \pi(\theta) : x_0^{(1)} \leq x_0 \leq x_0^{(2)}, n_0^{(1)} \leq n_0 \leq n_0^{(2)} \right\}.$$

El siguiente resultado proporciona las primas neta *posterior regret gamma-minimax* cuando se utiliza las clases  $\Gamma_j$ ,  $j = 1, 2, 3$ .

**Teorema 5.7** *Dada la verosimilitud perteneciente a la familia exponencial de distribuciones (5.19) y la distribución a priori conjugada (5.20), entonces la prima posterior regret gamma-minimax bajo el principio de prima neta cuando se utiliza  $\Gamma_j$ ,  $j = 1, 2, 3$ , viene dada por:*

$$RP_{B_j} = \frac{X_j + n\bar{X}}{N_j + n}, \quad j = 1, 2, 3, \quad (5.61)$$

donde,

$$X_1 = \frac{1}{2} (x_0^{(1)} + x_0^{(2)}), \quad N_1 = n_0, \quad (5.62)$$

$$X_2 = \frac{1}{2} (\gamma_1 + \gamma_2) n_0, \quad N_2 = n_0, \quad (5.63)$$

$$X_3 = \frac{x_0^{(1)}n_0^{(1)} + x_0^{(2)}n_0^{(2)} + (x_0^{(1)} + x_0^{(2)})n}{n_0^{(1)} + n_0^{(2)} + 2n}, \quad N_3 = \frac{2n_0^{(1)}n_0^{(2)} + (n_0^{(1)} + n_0^{(2)})n}{n_0^{(1)} + n_0^{(2)} + 2n}. \quad (5.64)$$

**Demostración:** La prima neta Bayes, cuando la verosimilitud es un miembro de la familia exponencial y la distribución a priori es la natural conjugada viene dada por:

$$P_B = \frac{x_0 + n\bar{X}}{n_0 + n}.$$

Ahora resulta claro que

$$\begin{aligned} \inf_{\pi \in \Gamma_1} P_B &= \frac{x_0^{(1)} + n\bar{X}}{n_0 + n}, \\ \sup_{\pi \in \Gamma_1} P_B &= \frac{x_0^{(2)} + n\bar{X}}{n_0 + n}, \end{aligned}$$

de donde, utilizando (5.58) se obtiene:

$$RP_B = \frac{X_1 + n\bar{X}}{n_0 + n}.$$

Dada la restricción impuesta en la clase de distribuciones  $\Gamma_2$  es inmediato obtener que  $n_0\gamma_1 \leq x_0 \leq n_0\gamma_2$ . Luego:

$$\begin{aligned} \inf_{\pi \in \Gamma_2} P_B &= \frac{n_0\gamma_1 + n\bar{X}}{n_0 + n}, \\ \sup_{\pi \in \Gamma_2} P_B &= \frac{n_0\gamma_2 + n\bar{X}}{n_0 + n}, \end{aligned}$$

de donde, utilizando (5.58) se obtiene:

$$RP_B = \frac{X_2 + n\bar{X}}{n_0 + n}.$$

Finalmente, para la clase  $\Gamma_3$  resulta:

$$\begin{aligned} \inf_{\pi \in \Gamma_3} P_B &= \frac{x_0^{(1)} + n\bar{X}}{n_0^{(2)} + n}, \\ \sup_{\pi \in \Gamma_3} P_B &= \frac{x_0^{(2)} + n\bar{X}}{n_0^{(1)} + n}, \end{aligned}$$

de donde, utilizando (5.58) se obtiene:

$$\begin{aligned}
 RP_B &= \frac{1}{2} \left( \frac{x_0^{(1)} + n\bar{X}}{n_0^{(2)} + n} + \frac{x_0^{(2)} + n\bar{X}}{n_0^{(1)} + n} \right) \\
 &= \frac{1}{2} \frac{x_0^{(1)} n_0^{(1)} + n_0^{(2)} x_0^{(2)} + [x_0^{(1)} + x_0^{(2)}] n + n\bar{X} [n_0^{(1)} + n_0^{(2)}]}{n_0^{(1)} n_0^{(2)} + [n_0^{(1)} + n_0^{(2)}] n + n^2} \\
 &= \frac{X_3 + n\bar{X}}{2 \left[ \frac{n_0^{(1)} n_0^{(2)} + [n_0^{(1)} + n_0^{(2)}] n + n^2}{n_0^{(1)} + n_0^{(2)} + 2n} + n - n \right]} \\
 &= \frac{X_3 + n\bar{X}}{\frac{2n_0^{(1)} n_0^{(2)} + 2[n_0^{(1)} + n_0^{(2)}] n + 2n^2 - [n_0^{(1)} + n_0^{(2)}] n - 2n^2}{n_0^{(1)} + n_0^{(2)} + 2n} + n} \\
 &= \frac{X_3 + n\bar{X}}{N_3 + n}.
 \end{aligned}$$

Puesto que los intervalos cerrados en la recta real son conjuntos conexos, utilizando la proposición 3.2 en Ríos et al. (1995) se deduce que  $RP(\pi; \Gamma_j)$ ,  $j = 1, 2, 3$ , son acciones Bayes.

El siguiente resultado prueba que las primas Bayes obtenidas en el teorema anterior pueden expresarse como fórmulas de credibilidad.

**Corolario 5.5** *Si se supone que la verosimilitud es un miembro de la familia exponencial de distribuciones (5.19) y que la distribución a priori es la natural conjugada (5.20) con parámetros  $(X_j, T_j)$ ,  $j = 1, 2, 3$  dados en (5.62), (5.63) y (5.64) respectivamente, entonces las primas netas regret gamma-minimax  $RP_{B_j}$ ,  $j = 1, 2, 3$  en (5.61) pueden reescribirse como fórmulas de credibilidad, siendo el factor de credibilidad como en (5.8).*

**Demostración:** Es inmediato que en todos los casos contemplados se tiene que

$$RP_{B_j} = Z(n)\bar{X} + [1 - Z(n)]m, \quad j = 1, 2, 3,$$

donde  $Z(n) = n/(N_j + n)$  y  $m = X_j/N_j$ ,  $j = 1, 2, 3$ . ■

**Ejemplo 5.17** Considérese la cartera de seguro de automóviles en Bélgica (1993) expuesta en el ejemplo 5.13. Vamos a calcular las primas netas Bayes cuando se supone que la distribución del número de reclamaciones es Poisson con parámetro  $\theta$  y la distribución a priori es gamma con parámetros  $a > 0$ ,  $b > 0$ . Vamos a calcular también las primas netas posterior regret gamma-minimax para las clases (5.59) y (5.60), tomando  $x_0^{(1)} = 1$ ,  $x_0^{(2)} = 2$ ,  $\gamma_1 = 0,1$ ,  $\gamma_2 = 0,2$  y  $n_0 = \hat{b}$ , siendo  $\hat{b}$  el estimador máximo verosímil del parámetro  $b$ . Considerar los períodos de tiempo de observación  $n = 1$ ,  $n = 5$  y las medias muestrales  $\bar{X} = \{0, 2, 4, 6, 8, 10\}$ .

Los estimadores máximo verosímiles de los parámetros  $a$  y  $b$  fueron calculados en el ejemplo 5.13,  $\hat{a} = 1,27$  y  $\hat{b} = 12,10$ . Las primas netas Bayes se obtienen sustituyendo estos valores en (5.25). Las primas posterior regret se calculan atendiendo al resultado el teorema anterior. En este caso se obtienen:

$$RP_B = \frac{X_1 + n\bar{X}}{b + n}, \quad X_1 = \frac{1}{2}(a_1 + a_2),$$

$$RP_B = \frac{X_2 + n\bar{X}}{b + n}, \quad X_2 = \frac{1}{2}(b(\gamma_1 + \gamma_2)),$$

para las clases  $\Gamma_1$  y  $\Gamma_2$ , respectivamente. Las primas resultantes aparecen en la tabla 5.25.

Tabla 5.25: Factor de credibilidad, prima neta Bayes y prima neta Bayes regret gamma-minimax para la clase  $\Gamma_1$ .

$\bar{X}, n = 1$	$P_B$	$RP_{B_1}$	$RP_{B_2}$
0	0.096	0.114	0.138
2	0.249	0.267	0.291
4	0.402	0.419	0.443
6	0.555	0.572	0.596
8	0.707	0.725	0.749
10	0.860	0.877	0.901
$\bar{X}, n = 5$	$P_B$	$RP_{B_1}$	$RP_{B_2}$
0	0.074	0.087	0.106
2	0.659	0.672	0.696
4	1.243	1.257	1.275
6	1.828	1.842	1.860
8	2.413	2.426	2.440
10	2.998	3.011	3.030

Puede verse que las primas posterior regret gamma-minimax son más conservadoras que las primas Bayes respectivas en todos los casos considerados. La diferencia es mayor para la clase  $\Gamma_2$ . El factor de credibilidad que resulta es  $Z(1) = 0,076$  y  $Z(5) = 0,293$ .

## Capítulo 6

# Aplicaciones en Riesgos Operacionales

### 6.1. Introducción

La cuantificación de pérdidas debidas a riesgos operacionales está recibiendo una considerable atención en la investigación financiera actual. Debido a las nuevas indicaciones regulatorias conocidas como Basilea II (BIS, 2005) en el sector bancario, se está realizando un considerable esfuerzo en la implementación de procedimientos tanto cualitativos como cuantitativos para modelizar el riesgo operacional.

El riesgo operacional de una compañía financiera o de seguros tiene su origen en sucesos que no pueden ser clasificados como riesgos de mercado o de crédito. El riesgo operacional se define como: *el riesgo de pérdidas resultantes de una falta de adecuación o un fallo de los procesos, el personal o los sistemas internos o bien de acontecimientos externos.*

El Comité de Basilea (2001) ha propuesto tres metodologías para la medición del riesgo operacional en términos del capital económico. Dichas metodología son:

- Método del indicador básico,
- Método estándar,
- Procedimientos de medición avanzada.

Dentro del último apartado tenemos el modelo de medición interna, los cuadros de mando y el modelo LDA (*loss distribution approach*) de pérdidas agregadas, que estudiaremos detalladamente. Las características generales de las metodologías anteriores han sido ampliamente descritas en la literatura y no serán comentadas

aquí. Textos sobre modelización de riesgos operacionales son (entre otros): Cruz (2002 y 2004), Panjer (2006), Chernobai et al. (2007) y los artículos de revisión de Chavez-Demoulin et al. (2006) y Moosa (2007). En castellano destacamos los textos editados por Fernández-Laviada (2007) y por Sarabia y Guillén (2007), donde aparecen diversas aplicaciones.

En este capítulo propondremos metodologías avanzadas para la medición del riesgo operacional en tres vertientes diferentes: frecuencia, severidad y modelos agregados. Estos modelos harán uso de las ideas de la teoría de la credibilidad.

Los contenidos del capítulo son los siguientes. En la sección 6.2 realizaremos algunas reflexiones metodológicas sobre la metodología en riesgos operacionales. La sección 6.3 se dedica a la modelización de la frecuencia en riesgos operacionales. Diversos modelos para la cuantificación de la severidad en riesgos operacionales aparecen descritos en la sección 6.4. La sección 6.5 se dedica a la modelización de la distribución de pérdidas agregadas (LDA). Finalmente, en la sección 6.6.2 veremos varias aplicaciones de los modelos aquí descritos.

## 6.2. Metodología Bayesiana en riesgos operacionales

Uno de los principales problemas en la aplicabilidad de riesgos operacionales es la falta de datos, tanto de frecuencia de un riesgo como de la severidad. Este hecho afecta obviamente a la estimación de las distribuciones marginales de tales pérdidas. La principal razón de esta falta de datos es que las instituciones financieras comenzaron a recoger datos relativos a riesgos operacionales hace pocos años, debido a la relativamente reciente definición del concepto de riesgo operacional. Teniendo en cuenta este inconveniente, la metodología Bayesiana junto con las técnicas de simulación, son las herramientas naturales para el tratamiento de este tipo de información. Es bien conocido que en análisis de datos, se recomienda el uso de la metodología Bayesiana cuando se dispone de poca información, además de cuando se quiere combinar información de expertos con la información proporcionada por los datos. La metodología Bayesiana ha sido descrita en el capítulo 4. En este nuevo capítulo, aplicaremos las herramientas Bayesianas a la modelización de la frecuencia y severidad datos de pérdidas en un contexto de riesgos operacionales.

## 6.3. Riesgos operacionales: modelización de la frecuencia

La elección de un modelo para la distribución de la frecuencia depende de varios factores: por un lado del tipo de datos disponibles y por otro de la fuente o

fuentes de la que proceden (datos de pérdidas internos y/o externos y/o auto-evaluaciones). Puede resultar complicado (al menos desde el punto de vista psicológico) diseñar preguntas en una auto-evaluación del riesgo que sean compatible con una distribución, por ejemplo, binomial negativa (Alexander, 2007). Si nos planteamos la pregunta ¿cuál es el número esperado de sucesos de pérdida para el próximo año?, la respuesta es clara y compatible con un modelo de Poisson de media  $\lambda$ . Sin embargo, puede resultar un poco más difícil aplicar la distribución binomial para datos externos de algún consorcio. La información disponible de un consorcio no suele registrar el valor  $N$  (número total de eventos) para cada banco, por lo que se debería utilizar alguna aproximación razonable, como es el caso de la distribución de Poisson.

En esta sección desarrollaremos dos modelos para la predicción de la frecuencia del riesgo operacional. Comenzaremos con el modelo Poisson-gamma, que puede ser el de mayor utilidad. A continuación, presentamos el binomial-beta, que puede ser utilizado en caso de disponer de información más detallada de cada entidad bancaria.

### 6.3.1. Modelo Poisson-gamma

En este modelo la frecuencia temporal (diaria, mensual o anual)  $X$  de pérdidas debidas a riesgos operacionales se puede modelizar según una distribución de Poisson  $X_i \sim \mathcal{P}o(\lambda)$  con función de cuantía

$$f_{X_i|\lambda}(x_i|\lambda) = \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}, \quad x_i = 0, 1, 2, \dots \quad (6.1)$$

Notar que la distribución de probabilidad (6.1) es condicionada para cada valor del parámetro  $\lambda$ . En consecuencia, la frecuencia temporal media no es constante de un período a otro puesto que  $\lambda$  es un parámetro aleatorio. Establecemos que  $\lambda$  se distribuye según una variable aleatoria tipo gamma de modo que:

$$\lambda \sim \mathcal{G}(\alpha, \sigma),$$

con función de densidad:

$$\pi(\lambda; \alpha, \sigma) = \frac{\lambda^{\alpha-1} \exp(-\lambda/\sigma)}{\sigma^\alpha \Gamma(\alpha)}, \quad \lambda > 0. \quad (6.2)$$

Supongamos ahora que la entidad bancaria dispone de datos procedentes de algún consorcio (o del propio sector o de alguna auto-evaluación) de modo que sea capaz de asignar la media y varianza del número de sucesos:

$$\begin{aligned} E[\lambda] &= m, \\ \text{Var}[\lambda] &= s^2, \end{aligned}$$

que conducen a los siguientes valores iniciales de los parámetros:

$$\alpha = m^2/s^2, \quad (6.3)$$

$$\sigma = s^2/m. \quad (6.4)$$

A continuación dado  $\lambda$ , se dispone de una serie temporal de frecuencias  $X = (X_1, X_2, \dots, X_n)$ , cuya función de verosimilitud es,

$$L(X|\lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{X_i}}{X_i!}. \quad (6.5)$$

Por medio del teorema de Bayes, combinando (6.2) con (6.6) tenemos que,

$$\begin{aligned} \pi(\lambda|X) &\propto L(X|\lambda) \cdot \pi(\lambda; \alpha, \sigma) \\ &= \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{X_i}}{X_i!} \cdot \frac{\lambda^{\alpha-1} \exp(-\lambda/\sigma)}{\sigma^\alpha \Gamma(\alpha)} \\ &\propto \lambda^{\tilde{\alpha}-1} \exp(-\lambda/\tilde{\sigma}) \end{aligned}$$

donde los parámetros  $\tilde{\alpha}$  y  $\tilde{\sigma}$  se actualizan de acuerdo con la tabla 6.1. De modo alternativo, puesto que la gamma es una distribución conjugada tenemos que,

$$\lambda_n|X \sim \mathcal{G} \left( \tilde{\alpha}_n = \sum_{i=1}^n X_i, \tilde{\sigma}_n = \frac{1}{n + 1/\sigma} \right), \quad (6.6)$$

donde el subíndice  $n$  significa que la estimación está basada en  $n$  períodos de tiempo.

Tabla 6.1: Modelo de frecuencia para riesgos operacionales Poisson-gamma: actualización de parámetros.

Parámetros	Valores a priori	Valores actualizados
$\alpha$	$\alpha$	$\tilde{\alpha} = \alpha + \sum_{i=1}^n X_i$
$\sigma$	$\sigma$	$\tilde{\sigma} = \frac{1}{n + 1/\sigma}$

La cantidad de interés es el valor de una frecuencia futura  $X_{n+1}$  dada la serie temporal  $X$ . Tenemos que,

$$\begin{aligned} E[X_{n+1}|X] &= E[\lambda|X] = \frac{n\bar{X} + \alpha}{n + 1/\sigma} \\ &= \frac{n}{n + 1/\sigma} \cdot \bar{X} + \frac{1/\sigma}{n + 1/\sigma} \cdot \alpha \cdot \sigma \\ &= Z(n)\bar{X} + [1 - Z(n)]E[\lambda] \end{aligned} \quad (6.7)$$

Tabla 6.2: Registro mensual de sucesos de pérdida: frecuencia baja.

Mes	1	2	3	4	5	6	7	8	9	10	11	12
No. sucesos	2	1	0	0	2	0	0	2	0	3	1	0
Mes	13	14	15	16	17	18	19	20	21	22	23	24
No. sucesos	1	2	0	0	0	2	3	2	3	0	1	0

que es una fórmula de credibilidad, donde el factor de credibilidad viene dado por,

$$Z(n) = \frac{n}{n + 1/\sigma}. \quad (6.8)$$

Cuando el horizonte temporal  $n$  aumenta, el factor de credibilidad  $w$  se acerca a uno, y la predicción de una frecuencia se basa únicamente en la serie histórica. Por otro lado, si el valor de  $n$  es pequeño, el factor de credibilidad disminuye y la predicción se basará principalmente en la opinión del experto  $E[\lambda]$ .

**Ejemplo 6.1** (Estimación de la frecuencia de Poisson usando datos históricos) En esta situación vamos a aplicar el modelo Poisson-gamma antes visto en tres situaciones: riesgos de frecuencia baja, media y alta.

- *Riesgos de baja frecuencia.* La tabla 6.2 presenta datos históricos de frecuencias de pérdidas mensuales (media 1.04 por mes). La figura 6.1 presenta la estimación secuencial de las estimaciones de máxima verosimilitud, de credibilidad (fórmula 6.8) para el siguiente período, así como la evolución del factor de credibilidad (fórmula (6.8)).
- *Riesgos de baja media.* La tabla 6.3 recoge datos históricos de frecuencias de pérdidas mensuales, con un total de 90 sucesos y una media muestral de 3.75. La figura 6.2 presenta nuevamente las estimaciones secuenciales de máxima verosimilitud y de credibilidad (fórmula 6.8) para el siguiente período.
- *Riesgos de alta frecuencia.* La tabla 6.4 presenta datos históricos de 233 frecuencias de pérdidas, con una media mensual de 9.71. La figura 6.3 presenta la evolución secuencial de los estimadores de máxima verosimilitud y de credibilidad.

### 6.3.2. Modelo binomial-beta

En esta ocasión disponemos de datos de frecuencias de riesgos operacionales relativos a distribuciones binomiales, de modo que,

$$f_{X_i|\Theta}(x_i|\theta) = \binom{N_i}{x_i} \theta^{x_i} (1 - \theta)^{N_i - x_i}, \quad x_i = 0, 1, \dots, N_i \quad (6.9)$$

Tabla 6.3: Registro mensual de sucesos de pérdida: frecuencia media.

Mes	1	2	3	4	5	6	7	8	9	10	11	12
No. sucesos	5	7	5	9	1	3	5	3	2	3	1	0
Mes	13	14	15	16	17	18	19	20	21	22	23	24
N. sucesos	2	2	6	3	3	5	5	6	3	3	4	4

Tabla 6.4: Registro mensual de sucesos de pérdida: frecuencia alta.

Mes	1	2	3	4	5	6	7	8	9	10	11	12
No. sucesos	13	5	12	12	8	8	9	7	10	10	9	7
Mes	13	14	15	16	17	18	19	20	21	22	23	24
N. sucesos	6	7	12	10	11	12	9	10	10	7	10	19

donde  $i = 1, 2, \dots, n$  y  $\theta \in (0, 1)$ . Puesto que  $\theta$  representa una proporción, elegimos como distribución a priori una beta de primera especie, con función de densidad,

$$\pi(\theta; a, b) = \frac{1}{B(a, b)} \theta^{a-1} (1-\theta)^{b-1}, \quad 0 < \theta < 1, \quad (6.10)$$

donde  $B(a, b)$  representa la función beta clásica, y donde el experto tiene que asignar dos parámetros  $a$  y  $b$  a partir de la experiencia del consorcio en riesgos operacionales. Si establecemos que  $E[\theta] = m$  y  $\text{Var}[\theta] = s^2$  entonces:

$$a = \frac{m^2 - m(s^2 + m^2)}{s^2},$$

$$b = \frac{m - (s^2 + m^2)}{s^2} - a.$$

La función de verosimilitud basada en una muestra  $X_1, \dots, X_n$  viene dada por,

$$L(X|\theta) = \prod_{i=1}^n \binom{N_i}{X_i} \theta^{X_i} (1-\theta)^{N_i - X_i}$$

$$\propto \theta^{\sum_{i=1}^n X_i} (1-\theta)^{\sum_{i=1}^n N_i - \sum_{i=1}^n X_i}$$

Haciendo uso del teorema de Bayes visto en el capítulo 4, la función de densidad a posteriori es,

$$\pi(\theta|X) \propto L(X|\theta) \cdot \pi(\theta; a, b)$$

$$= \theta^{\tilde{a}-1} (1-\theta)^{\tilde{b}-1}$$

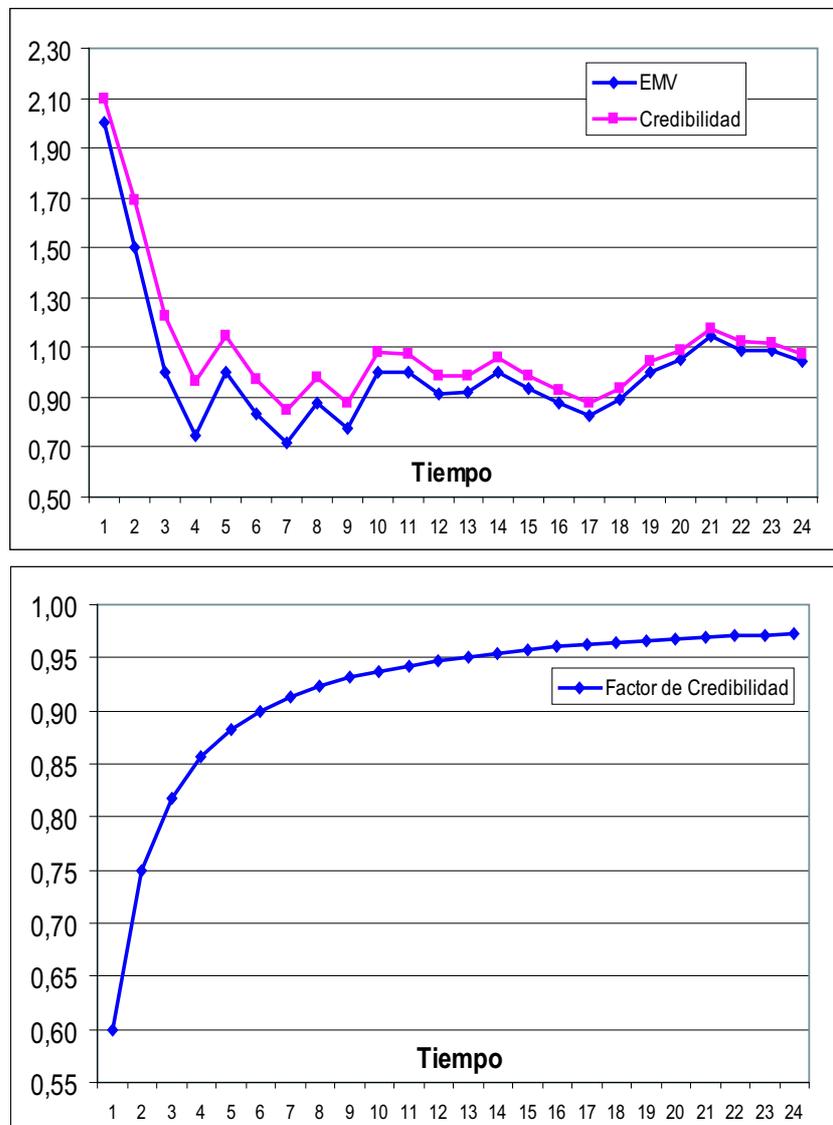


Figura 6.1: Modelo Poisson-gamma: frecuencia baja. Estimadores secuenciales en cada mes de máxima verosimilitud (EMV) y de credibilidad (gráfico superior). Factor secuencial de credibilidad (gráfico inferior).

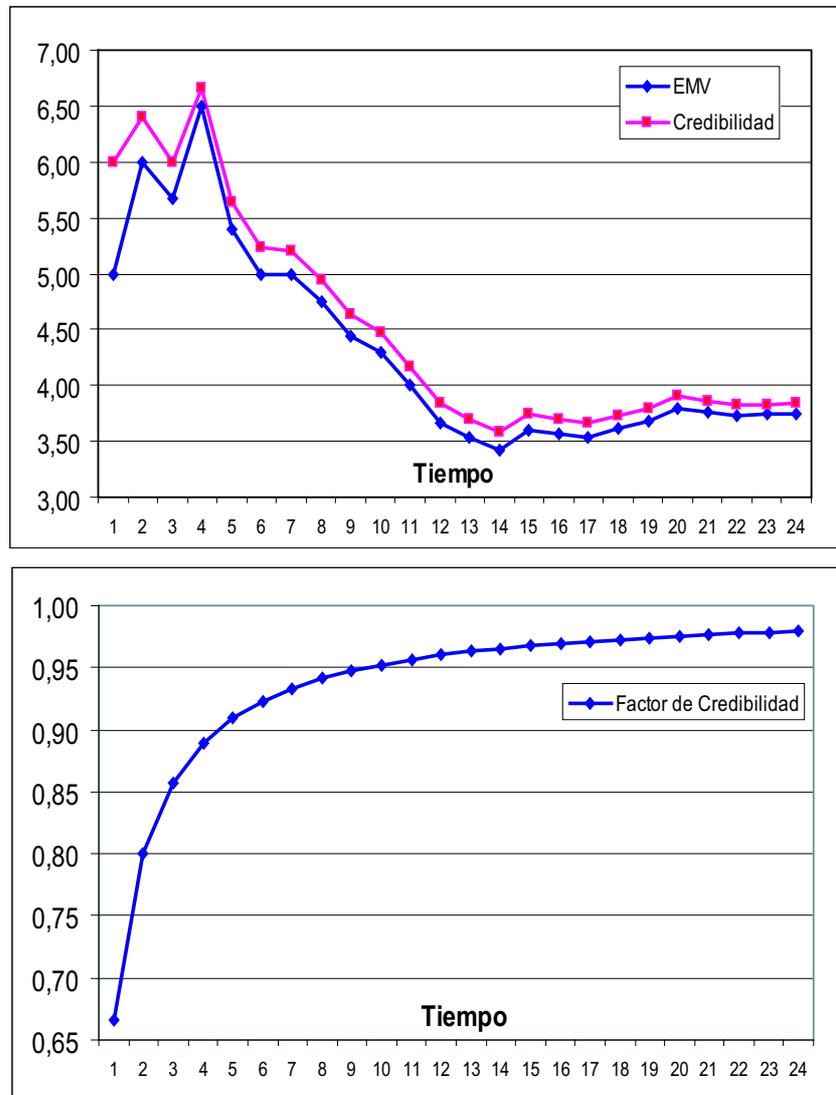


Figura 6.2: Modelo Poisson-gamma: frecuencia media. Estimadores secuenciales en cada mes de máxima verosimilitud (EMV) y de credibilidad (gráfico superior). Factor secuencial de credibilidad (gráfico inferior).

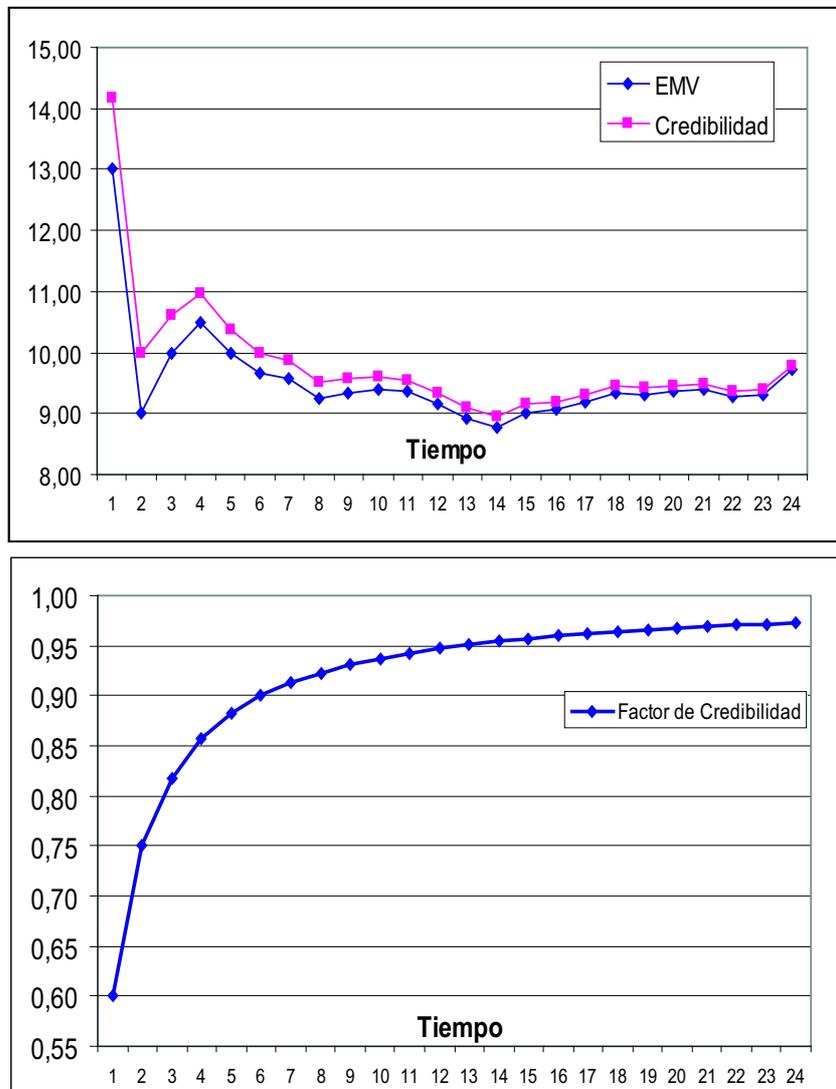


Figura 6.3: Modelo Poisson-gamma: frecuencia alta. Estimadores secuenciales en cada mes de máxima verosimilitud (EMV) y de credibilidad (gráfico superior). Factor secuencial de credibilidad (gráfico inferior).

donde las cantidades  $\tilde{a}$  y  $\tilde{b}$  se encuentran definidas en la tabla 6.5 y los parámetros  $a$  y  $b$  se actualizan de acuerdo con la citada tabla. La predicción de una observación futura viene dada por,

$$\begin{aligned} E[X_{n+1}|X] &= \int_0^1 \mu_{n+1}(\theta)\pi(\theta|X)d\theta \\ &= \int_0^1 N_{n+1}\theta \frac{\Gamma(\tilde{a} + \tilde{b})}{\Gamma(\tilde{a})\Gamma(\tilde{b})} \theta^{\tilde{a}-1}(1-\theta)^{\tilde{b}-1}d\theta \\ &= N_{n+1} \left( \frac{\sum_{i=1}^n N_i}{\sum_{i=1}^n N_i + a + b} \cdot \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n N_i} + \frac{a + b}{\sum_{i=1}^n N_i + a + b} \cdot \frac{a}{a + b} \right). \end{aligned}$$

Si

$$Z(n) = \frac{\sum_{i=1}^n N_i}{\sum_{i=1}^n N_i + a + b},$$

es el factor de credibilidad,

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n N_i},$$

es la media ponderada y  $\mu_\theta = \frac{a}{a+b}$ , tenemos que

$$E[X_{n+1}|X] = N_{n+1}[Z(n) \cdot \bar{X} + [1 - Z(n)] \cdot \mu_\theta],$$

que escribimos como,

$$E \left[ \frac{X_{n+1}}{N_{n+1}} | X \right] = Z(n) \cdot \bar{X} + [1 - Z(n)] \cdot \mu_\theta,$$

y es la prima de credibilidad.

Tabla 6.5: Modelo de frecuencia para riesgos operacionales binomial-beta: actualización de parámetros.

Parámetros	Valores a priori	Valores actualizados
$a$	$a$	$\tilde{a} = \sum_{i=1}^n X_i + a$
$b$	$b$	$\tilde{b} = \sum_{i=1}^n N_i - \sum_{i=1}^n X_i + b$

## 6.4. Riesgos Operacionales: modelización de la severidad

Para la modelización de la severidad presentamos cuatro modelos basados en las distribuciones lognormal, Pareto (clásica y generalizada) y exponencial, que

son cuatro de las distribuciones usuales para modelizar la severidad (capítulo 3). En los cuatro modelos obtendremos fórmulas de predicción basadas en la teoría de la credibilidad.

### 6.4.1. Modelo lognormal-normal

Supongamos que  $X$ , la severidad de las pérdidas debidas a riesgos operacionales, se pueden modelizar según una distribución lognormal  $X \sim \mathcal{LN}(\mu, \sigma^2)$ , definida en la sección 3.2.4. Específicamente, supongamos ahora que, dados  $\mu$  y  $\sigma$ , los datos de pérdidas operacionales  $X = (X_1, \dots, X_n)$  son variables aleatorias independientes  $\mathcal{LN}(\mu, \sigma^2)$  con función de densidad,

$$f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\}, \quad x > 0. \quad (6.11)$$

Definimos ahora el vector de las pérdidas transformadas,

$$Y = (Y_1, \dots, Y_n) = (\log X_1, \dots, \log X_n),$$

que por definición  $Y_i = \log X_i \sim \mathcal{N}(\mu, \sigma^2)$ . Mediante experiencias previas, el experto conoce el valor del parámetro  $\sigma$  y supone que la distribución a priori de  $\mu$  es una normal  $\mathcal{N}(\mu_0, \sigma_0^2)$  con función de densidad:

$$\pi(\mu; \mu_0, \sigma_0) = \frac{1}{\sigma_0\sqrt{2\pi}} \exp\left\{-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right\}, \quad \mu \in \mathbb{R}. \quad (6.12)$$

La función de verosimilitud de la muestra transformada  $Y$  es,

$$L(Y|\mu, \sigma) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(Y_i - \mu)^2}{2\sigma^2}\right\}. \quad (6.13)$$

Ahora, multiplicando la verosimilitud de los datos (6.13) por la función de densidad a priori (6.12) se obtiene que,

$$\begin{aligned} \pi(\mu|Y) &\propto \frac{1}{\sigma_0\sqrt{2\pi}} \exp\left\{-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right\} \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(Y_i - \mu)^2}{2\sigma^2}\right\} \\ &\propto \exp\left\{-\frac{(\mu - \tilde{\mu}_0)^2}{2\tilde{\sigma}_0^2}\right\}, \end{aligned}$$

que coincide con la distribución de una normal  $\mathcal{N}(\tilde{\mu}_0, \tilde{\sigma}_0^2)$ , y donde los parámetros  $\tilde{\mu}_0$  y  $\tilde{\sigma}_0$  vienen dados en la tabla 6.6.

A continuación nos interesa la predicción de una observación  $X_{n+1}$  (o de su transformada  $Y_{n+1} = -\log X_{n+1}$  en términos de la experiencia pasada  $X$ ). Tenemos

por tanto que,

$$\begin{aligned} E[Y_{n+1}|X] = E[\mu|X] &= \frac{\mu_0 + w \sum_{i=1}^n Y_i}{1 + n \cdot w} \\ &= Z(n) \cdot \bar{Y} + [1 - Z(n)] \cdot \mu_0, \end{aligned} \quad (6.14)$$

que es una fórmula de credibilidad, donde  $\bar{Y}$  es el estimador de  $\mu$  basado en las pérdidas operacionales históricas transformadas, mientras que  $\mu_0$  es la estimación basada en la opinión de un experto. El factor de credibilidad  $Z(n)$  viene dado por,

$$Z(n) = \frac{n}{n + \sigma^2/\sigma_0^2}, \quad (6.15)$$

cuyo valor se encuentra en el intervalo  $[0, 1]$ . Notar que si el tamaño muestral aumenta,  $n \rightarrow \infty$ , entonces  $w \rightarrow 1$ , y toda la predicción se basa en la experiencia acumulada de severidad. Sin embargo, si el tamaño de muestra es pequeño, el factor de credibilidad se convierte en próximo a cero y la predicción se basará en la opinión del experto  $\mu_0$ .

Tabla 6.6: Modelo de severidad para riesgos operacionales lognormal-normal: actualización de parámetros, donde  $w = \frac{\sigma_0^2}{\sigma^2}$ .

Parámetros	Valores a priori	Valores actualizados
$\mu_0$	$\mu_0$	$\tilde{\mu}_0 = \frac{\mu_0 + w \sum_{i=1}^n Y_i}{1 + n \cdot w}$
$\sigma_0^2$	$\sigma_0^2$	$\tilde{\sigma}_0^2 = \frac{\sigma_0^2}{1 + n \cdot w}$

### 6.4.2. Modelo Pareto generalizada-gamma

Supongamos ahora que la severidad de las pérdidas operacionales pueden ser modelizadas de acuerdo a una distribución generalizada de Pareto  $X \sim \mathcal{P}(\alpha, \sigma)$ , según la definición dada en la sección 3.2.1. La función de densidad de las pérdidas por riesgos operacionales es,

$$f(x|\alpha, \sigma) = \frac{\alpha}{\sigma} \left(1 + \frac{x}{\sigma}\right)^{-(\alpha+1)} 1(x > 0). \quad (6.16)$$

A pesar de la aparente sencillez de (6.16) existe una diferencia con la distribución de Pareto clásica. Mientras que la Pareto clásica dispone de un estadístico suficiente para la estimación de  $(\alpha, \sigma)$  dado por  $(\sum_{i=1}^n \log X_i, X_{1:n})$ , en este caso no es

posible una reducción de este tipo. Por tanto, independientemente de la distribución a priori que elijamos, no existe una expresión cerrada para la distribución a posteriori. Esto hace que los procedimientos Bayes empíricos estén recomendados.

Supongamos que se dispone de una muestra de tamaño  $n$ ,  $X = (X_1, \dots, X_n)$  de la severidad un riesgo. La función de verosimilitud de (6.16) viene dada por,

$$L(X|\alpha, \sigma) = \left(\frac{\alpha}{\sigma}\right)^n \prod_{i=1}^n \left(1 + \frac{X_i}{\sigma}\right)^{-(\alpha+1)} \quad (6.17)$$

Para la estimación establecemos dos supuestos. En el primer supuesto consideramos el parámetro  $\sigma$  conocido. En este caso, la distribución pertenece a la familia exponencial. Elegimos a priori una distribución gamma,

$$\alpha \sim \mathcal{G}(c, d),$$

donde los parámetros se asignan a partir de datos del consorcio, de acuerdo a la técnica descrita en la sección 6.3.1. Combinando la gamma con la verosimilitud, se obtiene que

$$\alpha|X \sim \mathcal{G}\left(n + c, \frac{1}{v_\sigma + 1/d}\right) \quad (6.18)$$

donde:

$$v_\sigma = \sum_{i=1}^n \log(1 + X_i/\sigma).$$

El estimador Bayes del parámetro  $\alpha$  viene dado por,

$$\begin{aligned} E[\alpha|X] &= \frac{n + c}{v_\sigma + 1/d} \\ &= \frac{v_\sigma}{v_\sigma + 1/d} \cdot \frac{n}{v_\sigma} + \frac{1/d}{v_\sigma + 1/d} \cdot cd \\ &= Z \cdot \hat{\alpha}_{EMV} + (1 - Z) \cdot E[\alpha] \end{aligned} \quad (6.19)$$

$$(6.20)$$

donde  $Z$  es el factor dado por,

$$Z = \frac{v_\sigma}{v_\sigma + 1/d},$$

y  $\alpha_{EMV} = n/v_\sigma$  el estimador de máxima verosimilitud de  $\alpha$ . En esta situación (6.20) no es estrictamente una fórmula de credibilidad, puesto que depende del valor de las severidades  $X_i$ .

En un segundo supuesto consideramos los dos parámetros  $\alpha$  y  $\sigma$  desconocidos. Si disponemos de datos a priori sobre la media  $m$  y la varianza muestral  $s^2$ , podemos estimar  $\sigma$  mediante la ecuación:

$$\hat{\sigma}_{EB} = \frac{m(s^2 + m^2)}{s^2 - m^2} \quad (6.21)$$

de donde el estimador Bayes empírico viene dado por,

$$E[\alpha|X] = \frac{n+c}{v_{EB}+1/d} = \frac{v_{EB}}{v_{EB}+1/d} \cdot \hat{\alpha}_{EMV} + \frac{1/d}{v_{EB}+1/d} \cdot E[\alpha], \quad (6.22)$$

donde:

$$v_{EB} = \sum_{i=1}^n \log(1 + X_i/\hat{\sigma}_{EB}).$$

Nuevamente (6.22) es una fórmula de credibilidad no estricta. La expresión (6.21) presupone la existencia del momento de orden dos. Si esta hipótesis no es creíble (véase el estudio de Moscadelli (2004) donde aparecen algunas evidencias empíricas sobre las colas de este tipo de distribuciones), podemos basar la estimación de  $\sigma$  en los percentiles de la distribución de las severidades, que obviamente siempre existen. Supongamos dos percentiles  $x_{p_i}$ ,  $i = 1, 2$  correspondientes a probabilidades  $p_i$ ,  $i = 1, 2$ . Se verifica entonces que:

$$\log(1 - p_i) = -\alpha \log(1 + x_{p_i}/\sigma), \quad i = 1, 2.$$

Eliminando  $\alpha$ , el estimador  $\hat{\sigma}_{PER}$  basado en percentiles es la solución en  $\sigma$  de la ecuación,

$$\frac{\log(1 - p_1)}{\log(1 - p_2)} = \frac{\log(1 + x_{p_1}/\sigma)}{\log(1 + x_{p_2}/\sigma)}.$$

Nuevamente el estimador de  $\alpha$  se puede escribir como una fórmula de credibilidad.

### 6.4.3. Modelo exponencial-gamma invertida

En este modelo suponemos que los datos de severidad debidos a riesgos operacionales se adecúan a una distribución exponencial de media  $\theta > 0$ , introducida en la sección 3.2.2. Si  $X = (X_1, \dots, X_n)$  es la muestra de pérdidas, la función de verosimilitud viene dada por,

$$L(X|\theta) = \prod_{i=1}^n \frac{\exp(-X_i/\theta)}{\theta}. \quad (6.23)$$

Para modelizar la media  $\theta$ , elegimos como distribución a priori una distribución gamma invertida,

$$\theta \sim \mathcal{GI}(a, b),$$

con función de densidad

$$\pi(\theta; a, b) = \frac{b^a \theta^{-a-1} \exp(-b/\theta)}{\Gamma(a)}, \quad \theta > 0 \quad (6.24)$$

con  $a, b > 0$ , introducida en la sección 3.2.3. Para la asignación de los hiperparámetros  $a$  y  $b$ , supongamos que disponemos de datos históricos del sector relativos a

este tipo de pérdidas operacionales. Si disponemos de la media  $m$  y la desviación típica  $s$  sectorial:

$$\begin{aligned} E[\theta] &= m, \\ \text{Var}[\theta] &= s^2, \end{aligned}$$

de donde:

$$a = \frac{2s^2 + m^2}{s^2}, \quad (6.25)$$

$$b = \frac{m(s^2 + m^2)}{s^2}. \quad (6.26)$$

Combinando (6.23) con (6.24) obtenemos la distribución a posteriori, que viene dada por,

$$\begin{aligned} \pi(\theta|X) &\propto L(X|\theta)\pi(\theta) \\ &\propto \theta^{-n} \exp\left(-\sum_{i=1}^n X_i/\theta\right) \theta^{-a-1} \exp(-b/\theta) \\ &= \theta^{-(n+a)-1} \exp\left(-\left(\sum_{i=1}^n X_i + b\right)/\theta\right), \end{aligned} \quad (6.27)$$

que se trata nuevamente de una distribución gamma invertida con parámetros:

$$\theta|X \sim \mathcal{GI}\left(\tilde{a} = a + n, \tilde{b} = b + \sum_{i=1}^n X_i\right) \quad (6.28)$$

Ahora, la predicción de la observación  $X_{n+1}$  según los datos  $X$  basada en credibilidad es:

$$\begin{aligned} E[\theta|X] &= \frac{b + \sum_{i=1}^n X_i}{a + n - 1} \\ &= \frac{n}{a + n - 1} \cdot \frac{\sum_{i=1}^n X_i}{n} + \frac{a - 1}{a + n - 1} \cdot \frac{b}{a - 1} \\ &= Z(n) \cdot \bar{X} + [1 - Z(n)]E[\theta], \end{aligned} \quad (6.29)$$

donde el factor de credibilidad es,

$$Z(n) = \frac{n}{a + n - 1}.$$

Una importante propiedad adicional de este modelo es que la distribución predictiva de una observación futura es del tipo Pareto generalizada, que posee una expresión cerrada por la función de distribución y por tanto facilita los cálculos.

#### 6.4.4. Modelo Pareto clásica-gamma

Este modelo resulta de especial utilidad cuando disponemos de datos de severidad de riesgos operacionales que exceden un umbral, como es el caso de la distribución clásica de Pareto (capítulo 3).

Supongamos entonces que la distribución de probabilidad de las pérdidas durante los últimos diez años (se supone independencia) es Pareto con función de densidad,

$$f(x|a, b) = \frac{a}{b} \left(\frac{x}{b}\right)^{-(a+1)}, \quad x \geq b > 0,$$

donde  $a > 0$ ,  $a = \alpha\beta$ , siendo  $\alpha > 0$  conocido y  $\beta > 0$ , para cada celda, según la tabla 6.7. Nótese que en este modelo se ha introducido un nuevo parámetro de distorsión. Suponemos que  $\beta$  tiene como distribución a priori una distribución gamma con parámetros  $\sigma > 0$ ,  $\lambda > 0$ , de modo que:

$$\pi(\beta|\sigma, \lambda) = \frac{\lambda^\sigma}{\Gamma(\sigma)} \beta^{\sigma-1} e^{-\lambda\beta}, \quad \beta > 0, \quad \sigma > 0, \quad \lambda > 0. \quad (6.30)$$

Vamos a obtener:

1. El estimador del parámetro  $\beta$  bajo función de pérdida cuadrática. Suponemos, al igual que Bühlmann et al. (2008), que  $\alpha = 1$ ,  $E[\beta] = 5$  y los valores  $\lambda$  son: 0.5; 1; 1.5; 2; 2.5; 3; 3.5; 4; 4.5 y 5.
2. Los estimadores que se obtienen cuando la distribución a priori es exponencial, i.e.  $\sigma = 1$  en (6.30).
3. La distribución marginal de los datos. Representarla gráficamente para  $\lambda = 0,5$ ;  $\lambda = 1$  y para el caso en que se adopte como distribución a priori la distribución exponencial.

Los estimadores y la distribución marginal se obtienen de la siguiente manera:

1. Por medio de la función de pérdida cuadrática el estimador de  $\beta$  es la media de la distribución a posteriori de este parámetro. La función de verosimilitud de los datos viene dada por:

$$\begin{aligned} L(X_1, \dots, X_n|\beta) &= \prod_{i=1}^n \left(\frac{\alpha\beta}{\lambda}\right) \left(\frac{X_i}{b}\right)^{-(\alpha\beta+1)} \\ &= \left(\frac{\alpha\beta}{b}\right)^n \prod_{i=1}^n \exp \left[ -(\alpha\beta + 1) \log \left(\frac{X_i}{b}\right) \right] \\ &= \left(\frac{\alpha\beta}{b}\right)^n \exp \left[ -(\alpha\beta + 1) \sum_{i=1}^n \log \left(\frac{X_i}{b}\right) \right]. \end{aligned}$$

Tabla 6.7: Pérdida en millones de dólares, excediendo 1 millón de dólares, en el Banco considerado (Bühlmann et al., 2008)

1	2	3	4	5	6	7	8	9	10
1.557	9.039	1.166	1.548	1.578	1.201	1.006	1.741	1.364	1.074
1.079	2.138	1.037	1.040	1.282	2.815	1.169	1.165	2.036	1.103
1.047	1.008	1.136	1.045	1.092	3.037	1.215	1.010	1.014	1.664
1.199	1.761	2.104	1.774	1.658	1.001	1.116	1.096	1.217	1.049
1.395	1.654	1.774	1.045	2.025	1.114	1.010	1.060	1.202	1.104
1.060	1.073	1.161	1.856	1.129	1.422	1.560	1.352	1.095	2.924
3.343	2.435	1.080	1.636	1.946	2.397	1.059	1.044	1.348	1.265
2.297	4.357	1.154	1.403	1.831	1.241	1.059	1.678	1.191	1.333
1.297	1.576	1.257	2.522	1.478	1.522	1.050	1.882	1.161	1.424
1.180	1.113	1.231	1.113	1.208	1.243	1.231	1.401	1.017	1.435

Luego, la distribución a posteriori del parámetro  $\beta$  dada los datos  $X_1, \dots, X_n$  resulta

$$\pi(\beta|X_1, \dots, X_n) \propto \beta^{\sigma+n-1} \exp \left[ -\beta \left( \alpha \sum_{i=1}^n \log \left( \frac{X_i}{b} \right) + \lambda \right) \right],$$

que es una distribución gamma con los parámetros actualizados:

$$\begin{aligned} \sigma &\longrightarrow \sigma + n, \\ \lambda &\longrightarrow \lambda + \alpha \sum_{i=1}^n \log \left( \frac{X_i}{b} \right). \end{aligned}$$

Por tanto, la media a posteriori del parámetro  $\beta$  es

$$\begin{aligned} E[\beta|X_1, \dots, X_n] &= \frac{\sigma + n}{\alpha \sum_{i=1}^n \log \left( \frac{X_i}{b} \right) + \lambda} \\ &= Z_\alpha \cdot \hat{\beta}_{EMV} + (1 - Z_\alpha) \cdot E[\beta], \end{aligned}$$

donde,

$$Z_\alpha = \frac{\alpha \sum_{i=1}^n \log(X_i/b)}{\alpha \sum_{i=1}^n \log(X_i/b) + \lambda},$$

que es nuevamente una fórmula de credibilidad. Ahora, puesto que

$$E[\beta] = \frac{\sigma}{\lambda} = 5,$$

se deduce que  $\sigma = 5\lambda$ . Finalmente, los valores del estimador de  $E[\beta|X]$ , para los distintos valores de  $\lambda$ , aparecen en la tabla 6.9, y en el gráfico 6.4

donde se ha tenido en cuenta, obviamente, que  $b = 1$ . Puede observarse que la secuencia de valores adopta la misma forma en todos los casos y que, salvo para la celda 7, el estimador calculado bajo la distribución a priori exponencial es menor que bajo los otros dos modelos. Destacar que el estimador de credibilidad es, en casi todas las celdas, menor que el estimador obtenido bajo la distribución a priori gamma.

Tabla 6.8: Estimador Bayesiano del parámetro de la distribución de Pareto. Distribución a priori del parámetro gamma.

$\lambda$	1	2	3	4	5	6	7	8	9	10
0.5	3.047	1.659	4.251	3.035	2.793	2.491	6.975	3.908	4.724	3.438
1.0	3.259	1.867	4.360	3.248	3.015	2.718	6.544	4.056	4.767	3.627
1.5	3.430	2.051	4.441	3.419	3.196	2.908	6.267	4.168	4.799	3.775
2.0	3.570	2.214	4.504	3.559	3.347	3.068	6.075	4.257	4.823	3.894
2.5	3.687	2.360	4.554	3.677	3.474	3.206	5.933	4.328	4.842	3.992
3.0	3.787	2.492	4.595	3.777	3.584	3.325	5.824	4.387	4.858	4.074
3.5	3.872	2.611	4.629	3.863	3.678	3.430	5.738	4.436	4.870	4.144
4.0	3.946	2.719	4.658	3.938	3.761	3.522	5.668	4.478	4.881	4.204
4.5	4.011	2.818	4.683	4.003	3.834	3.604	5.611	4.515	4.890	4.256
5.0	4.069	2.908	4.704	4.061	3.899	3.677	5.562	4.546	4.897	4.302

Tabla 6.9: Estimador Bayesiano del parámetro de la distribución de Pareto. Distribución a priori del parámetro exponencial.

$\lambda$	1	2	3	4	5	6	7	8	9	10
1	2.893	1.521	4.166	2.881	2.634	2.331	7.372	3.795	4.688	3.298

2. La distribución marginal de los datos se calcula de la siguiente forma ( $\alpha = 1$ ):

$$\begin{aligned}
 f(x|b, \sigma, \lambda) &= \int_b^{\infty} f(x|\beta, b)\pi(\beta|\sigma, \lambda)d\beta \\
 &= \frac{\lambda^{\sigma}}{\Gamma(\sigma)} \int_b^{\infty} \beta^{\sigma} e^{-\lambda\beta - (\beta+1)\log(x/b)} d\beta \\
 &= \frac{\sigma\lambda^{\sigma}b}{x[\lambda + \log(x/b)]^{\sigma+1}}.
 \end{aligned}$$

La gráfica de la función de densidad aparece representada en la figura 6.5. Se observa que las dos correspondientes a la distribución a priori gamma poseen colas más pesadas que la correspondiente a la distribución a priori

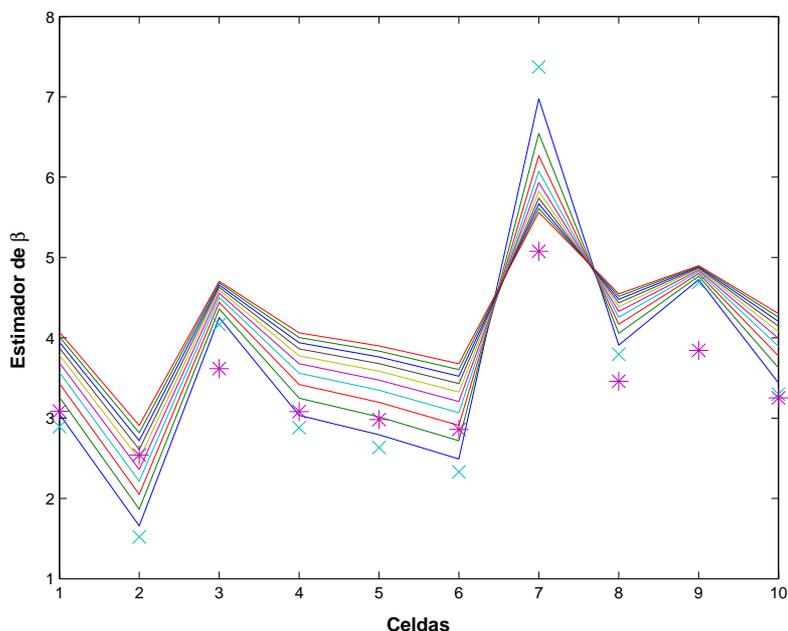


Figura 6.4: Valor del estimador en las distintas celdas para el caso de distribución a priori gamma y  $\lambda = 0, 0,5, \dots, 5$ , distribución a priori exponencial ( $\times$ ) y estimador de credibilidad (\*).

exponencial (caso  $\lambda = 0,2$ ,  $\sigma = 1$ ). De ahí que la primera resulte más adecuada para modelizar el riesgo operacional.

## 6.5. Modelo de distribución de pérdidas agregadas LDA

Según lo comentado en la sección 6.1, atendiendo a los requerimientos de Basilea II, la mayoría de los bancos adoptan una aproximación a la distribución agregada de pérdidas totales LDA. Para ello, al igual que bajo el modelo colectivo compuesto en estadística actuarial, se requiere la siguiente información:

1. El número de pérdidas operacionales,  $N$ . Se trata de sucesos que dan lugar a la materialización de un riesgo de carácter operacional.
2. Las pérdidas operacionales o severidades  $X_k$ . Constituyen éstas, impactos

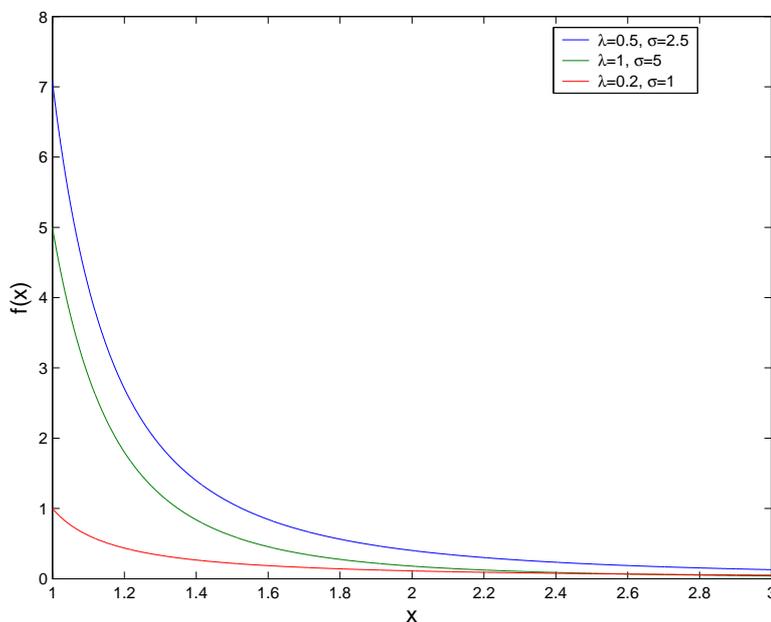


Figura 6.5: Distribución marginal en el modelo Pareto clásica-gamma de riesgo operacional.

negativos en resultados o en el patrimonio de la entidad, registrados contablemente, como consecuencia de cualquier evento de riesgo operacional.

El objetivo es la estimación del capital riesgo del banco correspondiente cumpliendo, simultáneamente, con los requerimientos de Basilea II. Este comité:

1. Aconseja calcular el capital riesgo (conocido también como capital regulatorio) como el percentil de la distribución de pérdidas totales para un nivel de probabilidad del 99.99 % y un horizonte temporal de un año. Este percentil suele conocerse como VaR (Valor en Riesgo operacional) o CaR (Capital en Riesgo operacional).
2. Establece ocho líneas de negocio:
  - a) Administración de activos.
  - b) Banca comercial.
  - c) Banca minorista.
  - d) Intermediación minorista.

- e) Finanzas corporativas.
- f) Negociación y ventas.
- g) Pagos y liquidación.
- h) Servicios de agencia.

y establece siete categorías de sucesos de riesgo:

- a) Fraude interno.
- b) Fraude externo.
- c) Relaciones laborales y fallos de seguridad en el puesto de trabajo.
- d) Clientes, productos y prácticas empresariales.
- e) Daños a activos materiales.
- f) Incidencias en el negocio y fallos en los sistemas.
- g) Ejecución, entrega y gestión de procesos.

Cada una de las pérdidas por riesgo operacional debe asignarse a una de las celdas de una matriz de riesgos y el capital en riesgo debe calcularse en este entorno.

La evidencia empírica pone de manifiesto que la magnitud del capital riesgo podría estar determinada sólo por pérdidas extremas y poco frecuentes. De ahí que para las estimaciones oportunas se tenga en cuenta solamente valores de pérdidas que superen un umbral determinado, generalmente elevado. Es por esto que la teoría de valores extremos juega un papel destacado en esta disciplina.

Teniendo en cuenta que la distribución de pérdidas totales, y los supuestos establecidos sobre las variables aleatorias que la conforman, responden exactamente al modelo colectivo compuesto, ampliamente utilizado en estadística actuarial, no puede sorprendernos que los métodos utilizados en esta disciplina se apliquen también para la estimación de riesgos operacionales. De ahí que se trabaje actualmente con metodologías de la estadística actuarial.

Siguiendo por tanto, las recomendaciones del comité de Basilea, el proceso de pérdidas agregadas puede ser modelizado mediante un proceso estocástico dado por una suma aleatoria, donde el número de sumandos viene gobernado por un proceso  $N(t)$  (frecuencia operacional del proceso) y cada uno de los sumandos es una variable aleatoria no negativa que representa la severidad debida al riesgo operacional. Más estrictamente tenemos la siguiente definición.

**Definición 6.1 (Proceso LDA)** *El proceso agregado de pérdidas operacionales viene definido por medio de:*

(1) *El proceso de severidad: Las severidades  $\{X_k\}_{k=1,2,\dots}$  son variables aleatorias no negativas independientes e igualmente distribuidas que representan la magnitud de cada pérdida operacional.*

(2) *El proceso de frecuencia: El número  $N(t)$  de sucesos de pérdidas en el intervalo*

de tiempo  $N(t)$  con  $t \geq 0$  es aleatorio. El proceso de conteo  $\{N(t)\}_{t \geq 0}$  viene generado por una sucesión de puntos  $\{T_1, T_2, \dots\}$  aleatorios no negativos tales que,

$$0 \leq T_1 \leq T_2 \leq \dots \text{ c.s.}$$

y

$$N(t) = \sup\{n \geq 1 : T_n \leq t\}, \quad t \geq 0$$

(3) Los procesos de severidad y de frecuencia son independientes.

(4) El proceso de pérdidas agregadas viene definido por,

$$S_N(t) = \sum_{k=1}^{N(t)} X_k, \quad t \geq 0. \quad (6.31)$$

Veamos algunos ejemplos típicos de procesos de frecuencia usuales en LDA.

**Ejemplo 6.2** El proceso LDA-Poisson estándar está basado en el proceso de Poisson homogéneo con parámetro de intensidad  $\lambda$  y función de cuantía,

$$\Pr(N(t) = n) = \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots$$

El proceso geométrico-LDA se basa en la distribución geométrica como medida de conteo,

$$\Pr(N(t) = n) = \frac{\gamma}{\gamma + t} \left( \frac{t}{\gamma + t} \right)^n, \quad n = 0, 1, 2, \dots$$

con  $\gamma > 0$ . Si disponemos de frecuencias debidas a riesgos operacionales con clara sobredispersión, el proceso binomial negativa-LDA es el adecuado:

$$\Pr(N(t) = n) = \binom{n+r-1}{n} \left( \frac{\gamma}{\gamma+t} \right)^r \left( \frac{t}{\gamma+t} \right)^n, \quad n = 0, 1, 2, \dots$$

con  $r > 0$ , que incluye como caso particular a la distribución geométrica.

Como distribuciones de probabilidad asociadas a  $N(t)$  se suele trabajar con las distribuciones Poisson y Binomial Negativa como acabamos de ver. En relación a las severidades, se puede trabajar con una gran número de distribuciones (capítulo 3). Por ejemplo, las distribuciones de exponencial, gamma, Weibull, lognormal y Pareto. Esta última suele ser la más utilizada porque permite modelizar de una manera natural y sencilla el hecho destacado anteriormente de que se trabaja con datos de pérdidas por encima de un umbral establecido. En caso que se adopte una distribución exponencial para las pérdidas individuales, esta distribución permite obtener una expresión cerrada para la distribución del proceso  $S$ , como veremos a continuación (Sarabia et al. (2006) y Rolski et al., (1999)).

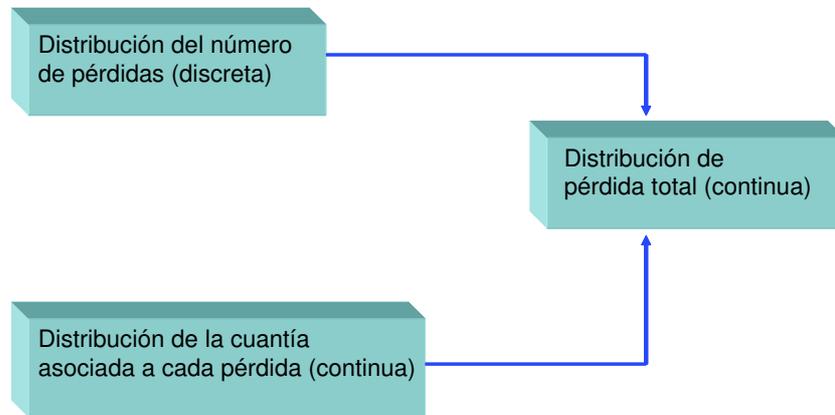


Figura 6.6: Esquema del proceso LDA: distribución del número de sucesos que dan lugar a pérdidas y distribución de las pérdidas para obtener la distribución de pérdidas agregadas.

Si bien los datos del propio banco que se analiza se suelen recoger durante varios años, en la práctica el horizonte temporal con el que se trabaja es de un año. Se trata entonces conjuntamente la información interna de la que se dispone del banco y la información colateral, vía opinión de expertos para, haciendo uso del teorema de Bayes, hacer las predicciones correspondientes. Esto obviamente responde a una metodología bayesiana pura en la que destaca el trabajo de Shvchenko y Wüthrich (2006).

Dado que el modelo básico de riesgo operacional es similar al modelo colectivo compuesto de estadística actuarial, Bühlmann et al. (2008) han aplicado el modelo de Bühlmann-Straub (1972) en este escenario. Para ello han tratado separadamente el número de sucesos con pérdidas y las pérdidas correspondientes en un banco dado. Así, para el caso de pérdidas han trabajado con un banco, denominado banco 1 (figura 6.7).

Estos autores han supuesto una distribución de Pareto para las pérdidas en cada una de las 10 celdas de riesgo, que aparecen en la tabla 6.7, y en la que las pérdidas observadas superan el millón de dólares. Usando esta información junto a la información de la estructura bancaria en su conjunto (información colateral a priori) han obtenido los estimadores de credibilidad de uno de los parámetros de la distribución de Pareto. En este caso el banco 1 tiene un parámetro de riesgo dado por  $\beta_1$ , mientras que el resto de bancos tiene el parámetro de riesgo  $\beta_j$ ,  $j = 2, \dots, M$ , para una estructura de  $M$  bancos. La información colateral la incorporan suponiendo que el valor esperado del parámetro en el colectivo de la estructura bancaria es conocido.

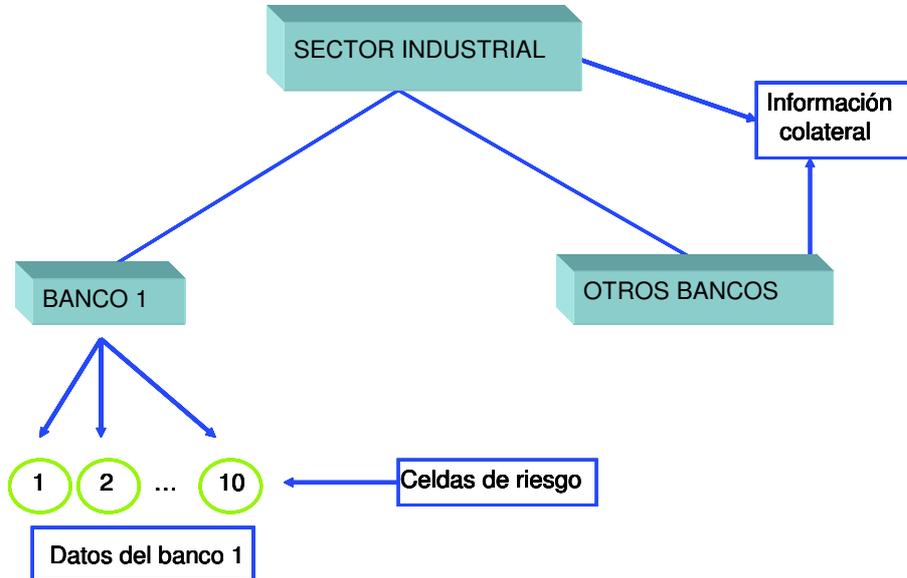


Figura 6.7: Niveles de información en riesgos operacionales.

### 6.5.1. Propiedades generales del proceso LDA

Algunas propiedades del proceso LDA son conocidas, y pueden obtenerse a partir de las propiedades del modelo clásico de riesgo colectivo, de amplia utilización en la ciencia actuarial (ver por ejemplo: Kass et al. (2001), Klugman et al. (2004), Panjer (2006), Sarabia et al. (2006)). Destacamos las siguientes propiedades del proceso LDA:

- *Función de distribución.* La función de distribución de (6.31) viene dada por,

$$F_{S_N(t)}(s) = \sum_{n=1}^{\infty} \Pr(N(t) = n) F_X^{*n}(s), \quad s > 0 \quad (6.32)$$

y

$$F_{S_N(t)}(0) = \Pr(N(t) = 0), \quad (6.33)$$

donde  $F_X^{*n}$  representa la función de distribución de la  $n$ -ésima convolución de la severidad  $X_k$ .

- *Media y varianza.* La media y la varianza vienen dadas por,

$$E[S_N(t)] = E[N(t)] \cdot E[X], \quad (6.34)$$

$$\text{Var}[S_N(t)] = E[N(t)] \cdot \text{Var}[X] + \text{Var}[N(t)] \cdot (E[X])^2. \quad (6.35)$$

- *Función característica.* La función característica del proceso viene dada por,

$$\varphi_{S_N}(t) = P_{N(t)}[\varphi_X(t)], \quad (6.36)$$

donde  $P_{N(t)}$  es la función generatriz del proceso de conteo y  $\varphi_X(t)$  la función característica de la severidad.

- *Normalidad Asintótica.* La siguiente propiedad se aplica al caso del proceso de Poisson, pero puede extenderse a otros procesos. Supongamos  $S_N = S_\lambda$  con  $\lambda \rightarrow \infty$  y  $t > 0$  fijo. Entonces si  $\lambda \rightarrow \infty$  se verifica,

$$\Pr\left(\frac{S_\lambda - E[S_\lambda]}{\sqrt{\text{Var}[S_\lambda]}} \leq x\right) \rightarrow \Phi(x),$$

siempre que  $E[X_k] < \infty$ , donde  $E[S_\lambda]$  y  $\sqrt{\text{Var}[S_\lambda]}$  vienen definidos en (6.34) y (6.35) y donde  $\Phi(\cdot)$  denota la función de distribución de una distribución normal estándar.

- *Comportamiento de la cola.* La cola superior de tiene una expresión sencilla en algunos casos especiales, cuando  $X_k$  pertenecen a la clase de distribuciones sub-exponenciales, que incluye a la distribución lognormal, Pareto, Weibull con colas pesadas etc. En este caso se verifica que,

$$\Pr(S_N > s) \sim E[N(t)] \cdot \Pr(X > s), \quad s \rightarrow \infty.$$

**Ejemplo 6.3 (Proceso LDA Pareto-Poisson)** Supongamos que las severidades siguen una distribución clásica de Pareto de parámetros  $\alpha$  y  $\sigma$ , y que el proceso de conteo es de tipo Poisson de modo que  $E[N(t)] = \text{Var}[N(t)] = \lambda t$ ,  $\lambda, t > 0$ . De acuerdo con las ecuaciones (6.34) y (6.35) la media y varianza del proceso LDA Pareto-Poisson vienen dadas por,

$$\begin{aligned} E[S_N(t)] &= \lambda t \cdot \frac{\alpha\sigma}{\alpha - 1}, \quad \alpha > 1, \\ \text{Var}[S_N(t)] &= \lambda t \cdot \frac{\alpha\sigma^2}{\alpha - 2}, \quad \alpha > 2. \end{aligned}$$

La función característica del proceso no tiene una expresión sencilla, y por tanto los cálculos de este proceso pueden ser algo laboriosos.

La obtención o cálculo de la distribución del LDA (6.32) no resulta sencilla. Para trabajar con ella se recurre a una serie de métodos exactos o aproximados:

1. El algoritmo de recursión de Panjer (Panjer, 1981 y Sarabia et al., 2006, pp. 428-435).

2. La transformada rápida de Fourier (Bertram, 1981 y Heckman y Meyers, 1983).
3. Método de DePril (Kuon et al., 1987).
4. Método de Kornya (Kuon et al., 1987).
5. La aproximación basada en métodos de Monte Carlo (Heckman y Meyers, 1983).
6. Aproximación a la pérdida total (Gerber, 1979, capítulo 4 y Sarabia et al., 2006, pp. 438-449).
7. Cálculo exacto mediante el uso de la convolución (Sarabia et al., 2006, pp. 419-424)

Pasemos a describir algunos procesos importantes cuya función de distribución o de densidad dispone de una expresión explícita.

### 6.5.2. Procesos LDA con distribución explícita

El siguiente resultado (Sarabia et al, (2006)) permite obtener la distribución exacta del proceso LDA, en el caso de severidades de tipo exponencial.

**Teorema 6.1** *Si el proceso de frecuencia tiene por función de probabilidad  $p_n$  y la distribución de la severidad es exponencial de parámetro  $\theta > 0$ , entonces la función de densidad, así como la función de distribución del proceso LDA vienen dadas por:*

$$f_S(x; \theta) = \frac{e^{-\theta x}}{x} \sum_{n=1}^{\infty} p_n \frac{(\theta x)^n}{(n-1)!}, \quad x > 0, \quad (6.37)$$

$$F_S(x; \theta) = 1 - e^{-\theta x} \sum_{j=0}^{\infty} \tilde{P}_j \frac{(\theta x)^j}{j!}, \quad x > 0,$$

donde  $\tilde{P}_j = \sum_{n=j+1}^{\infty} p_n$ ,  $j = 0, 1, \dots$

El siguiente resultado corresponde al caso en que el proceso de frecuencia es un proceso de Poisson homogéneo. Tenemos el siguiente teorema.

**Teorema 6.2 (Proceso LDA Poisson-Exponencial)** *La función de densidad exacta del modelo LDA con proceso de conteo Poisson y severidad exponencial viene dada por,*

$$f_S(s; \lambda, \theta) = \exp\{-(\lambda t + s/\theta)\} \sqrt{\frac{\lambda t}{\theta s}} I_1\left(2\sqrt{\lambda t s/\theta}\right), \quad s > 0 \quad (6.38)$$

y

$$f_S(0; \lambda, \theta) = \exp(-\lambda t), \quad (6.39)$$

donde

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{2k+\nu}}{k! \Gamma(\nu + k + 1)}, \quad x \in \mathbf{R}, \quad \nu \in \mathbf{R}$$

es la función modificada de Bessel de clase  $\nu$ .

### 6.5.3. Otros Procesos LDA

La importancia de la modelización de las pérdida agregadas por medio de los modelos LDA es evidente. Como acabamos de ver la aplicabilidad de estos modelos depende de la disponibilidad de una fórmula adecuada del tipo (6.32) o de la existencia de algún tipo de recurrencia que permita realizar los cálculos computacionales. En este sentido la investigación en riesgos ha realizado un importante esfuerzo, de modo que los resultados disponibles se pueden aplicar en este ámbito. En este apartado incluimos otros modelos LDA de utilidad que pueden ser implementados con relativa facilidad. La ventaja de los procesos que incluimos a continuación es que son válidos para cualquier tipo de distribución para la severidad (por ejemplo, no necesariamente con colas pesadas o no necesariamente de tipo absolutamente continuo). Incluiremos a continuación los resultados suponiendo la severidad absolutamente continua.

#### Proceso LDA de Panjer.

Supongamos que la distribución de conteo  $\{p_n\}_{n=0,1,\dots}$  pertenece a la clase de distribuciones de Panjer (clase  $(a, b, 0)$ ), donde la función de masa de probabilidad  $p_n = \Pr(N = n)$  verifica:

$$\frac{p_n}{p_{n-1}} = a + \frac{b}{n}, \quad n = 1, 2, \dots, \quad (6.40)$$

En este caso, la función de densidad de  $S_N$  satisface la siguiente fórmula recursiva:

$$f_S(x; a, b) = \begin{cases} P_N(f_S(0)), & x = 0, \\ p_1 f(x) + \int_0^x \left(a + \frac{by}{x}\right) f(y) f_S(x-y) dy, & x > 0, \end{cases} \quad (6.41)$$

siendo  $f(x)$  la función de densidad de la severidad  $X$  y el proceso de conteo  $N$  pertenece a la clase  $(a, b, 0)$ .

**Proceso LDA Poisson generalizada.**

En este proceso (Ambagaspitiya y Balakrishnan, 1994) la variable aleatoria de conteo sigue una distribución de Poisson generalizada con función de cuantía:

$$\Pr(N = n) = \lambda(\lambda + n\theta)^{n-1} \frac{\exp(-\lambda - n\theta)}{n!}, \quad n = 0, 1, \dots \quad (6.42)$$

y 0 en otro caso, donde  $\lambda > 0$ ,  $\max\{-1, -\lambda/m\} \leq \theta < 1$  y  $m \geq 4$  es el mayor entero tal que  $\lambda + n\theta > 0$  cuando  $\theta$  es negativo. En este proceso, no es nuevamente necesaria ninguna hipótesis sobre la distribución de las severidades, excepto que sea absolutamente continua, con función de densidad  $f(x)$  para  $x > 0$ . En este caso, la función de densidad  $g_S(x; \lambda, \theta)$  del modelo LDA verifica la ecuación integral

$$g_S(x; \lambda, \theta) = p_1(\lambda, \theta)f(x) + \frac{\lambda}{\lambda + \theta} \int_0^x \left( \theta + \frac{\lambda y}{x} \right) g_S(x - y; \lambda + \theta, \theta) f(y) dy, \quad (6.43)$$

donde  $p_1(\lambda, \theta) = \Pr(N = 1)$ .

**Proceso LDA binomial negativa-inversa Gaussiana.**

Gómez-Déniz et al. (2008e) han propuesto recientemente una distribución discreta altamente flexible, denominada distribución binomial negativa-inversa Gaussiana. Esta distribución se obtiene a partir de una mezcla de la distribución binomial negativa con una distribución inversa Gaussiana. La función de masa de probabilidad es

$$\Pr(N = n) = \binom{r+n-1}{n} \sum_{j=0}^n (-1)^j \binom{n}{j} \exp \left\{ \frac{\psi}{\mu} \left[ 1 - \sqrt{1 + \frac{2(r+j)\mu^2}{\psi}} \right] \right\}, \quad (6.44)$$

con  $n = 0, 1, 2, \dots$  y  $r, \mu, \psi > 0$ . Este modelo de conteo permite ajustar correctamente situaciones de sobredispersión, mejorando los ajustes de la distribución binomial negativa clásica. Si la distribución de las severidades es absolutamente continua con función de densidad  $f(x)$  para  $x > 0$ , entonces la función de densidad  $g_S(x; r)$  del correspondiente modelo LDA satisface la ecuación integral

$$g_S(x; r) = p_r(0) + \int_0^x \frac{ry + x - y}{x} g_S(x - y; r) f(y) dy - \int_0^x \frac{ry}{x} g_S(x - y; r + 1) f(y) dy. \quad (6.45)$$

## 6.6. Aplicaciones

Veamos dos aplicaciones de los modelos previamente estudiados.

### 6.6.1. Tipología en procesos LDA

En este apartado obtendremos diversas funciones de densidad exactas del proceso LDA Poisson-exponencial según tres escenarios, siguiendo las recomendaciones de Basilea:

- *Frecuencia baja*:  $\lambda = 1$ ,
- *Frecuencia media*:  $\lambda = 5$ ,
- *Frecuencia alta*:  $\lambda = 10$ .

Combinaremos las frecuencias anteriores con diversas severidades, cuya media individual  $\theta$  puede tomar seis valores: 100, 150, 200, 250 y 300 u.m. Las gráficas de las diferentes funciones de densidad aparecen representadas en las figuras 6.9 (frecuencia baja), 6.10 (frecuencia media) y 6.11 (frecuencia alta). Es importante señalar de nuevo que las gráficas de las funciones de densidad son exactas (teorema 6.2), y no han sido obtenidas mediante simulación. A partir de la información gráfica y de los resultados de la sección 6.5 podemos obtener las siguientes conclusiones:

- La media del proceso LDA aumenta tanto con la intensidad media de la frecuencia como con la media de la severidad (ecuación 6.34).
- La variabilidad del proceso LDA aumenta nuevamente tanto con la intensidad media, como con la los momentos de orden uno y dos de la severidad. Este aumento se hace especialmente visible para valores medios y grandes de la severidad (figuras 6.9 a 6.11).
- El coeficiente de asimetría del proceso disminuye al aumentar la magnitud de la frecuencia (figura 6.11). Para valores moderados y grandes de la severidad, la aproximación normal asintótica parece razonable.
- Las probabilidades asociadas al proceso (tanto centrales como las de las colas) se pueden obtener de forma exacta.

### 6.6.2. Aplicación del modelo exponencial-gamma invertida

En este apartado usaremos el modelo exponencial-gamma invertida para predecir la severidad debida a riesgos operacionales (sección 6.4.3). Para ello utilizaremos la información proporcionada por Guillén et al. (2007) correspondiente a una entidad financiera cuyo nombre se mantiene en el anonimato. Puesto que el modelo dispone de un estadístico suficiente (y por tanto incluye toda la información), para la aplicación de nuestros resultados sólo se requieren los dos primeros momentos. La asignación de los hiperparámetros y los estimadores de credibilidad son los que aparecen en la citada sección. La tabla 6.10 recoge los estadísticos descriptivos de siete categorías de sucesos debidos a riesgos operacionales: fraude

interno, fraude externo, prácticas de empleo y seguridad laboral, interrupción negocio, clientes, productos y fallos de sistemas, daños en activos físicos y ejecución, entrega y gestión de procesos.

Los resultados de las estimaciones aparecen en las tablas 6.11 (riesgos 1, 2 y 3) y 6.12 (riesgos 4, 5, 6 y 7). La figura 6.8 muestra las funciones de densidad a priori y a posteriori del modelo exponencial-gamma invertida en dos riesgos operacionales de diferente naturaleza. En el riesgo operacional 1 (fraude interno) la distribución a posteriori está muy concentrada debido a que existe una importante evidencia empírica, lo que hace disminuir la variabilidad, mientras que el riesgo 4 (interrupción negocio) muestra mayor variabilidad en la distribución a posteriori, puesto que se dispone de una menor evidencia muestral. Además de las predicciones del suceso siguiente mediante credibilidad se han incluido los valores  $\text{VaR}[X; q]$  y  $\text{TVaR}[X; q]$  de la distribución a posteriori, para valores  $q=0.95, 0.99, 0.997, 0.999$  y  $0.999$ , además del intervalo de credibilidad al 95 por ciento.

Tabla 6.10: Estadísticos descriptivos de siete categorías de riesgos debidas a riesgos operacionales (en millones de  $\mathcal{L}$ ).

<i>Tipo de riesgo</i>	<i>Número de pérdidas</i>	<i>Media muestral</i>	<i>D. típica muestral</i>
1. Fraude interno	1247	32.24	269.43
2. Fraude externo	538	15.60	69.68
3. Prácticas de empleo y seguridad laboral	721	7.84	20.04
4. Interrupción negocio	45	22.46	33.25
5. Clientes, productos y fallos de sistemas	6526	36.59	268.28
6. Daños en activos físicos	2395	74.91	1192.55
7. Ejecución, entrega y gestión de procesos	75	7.39	17.72

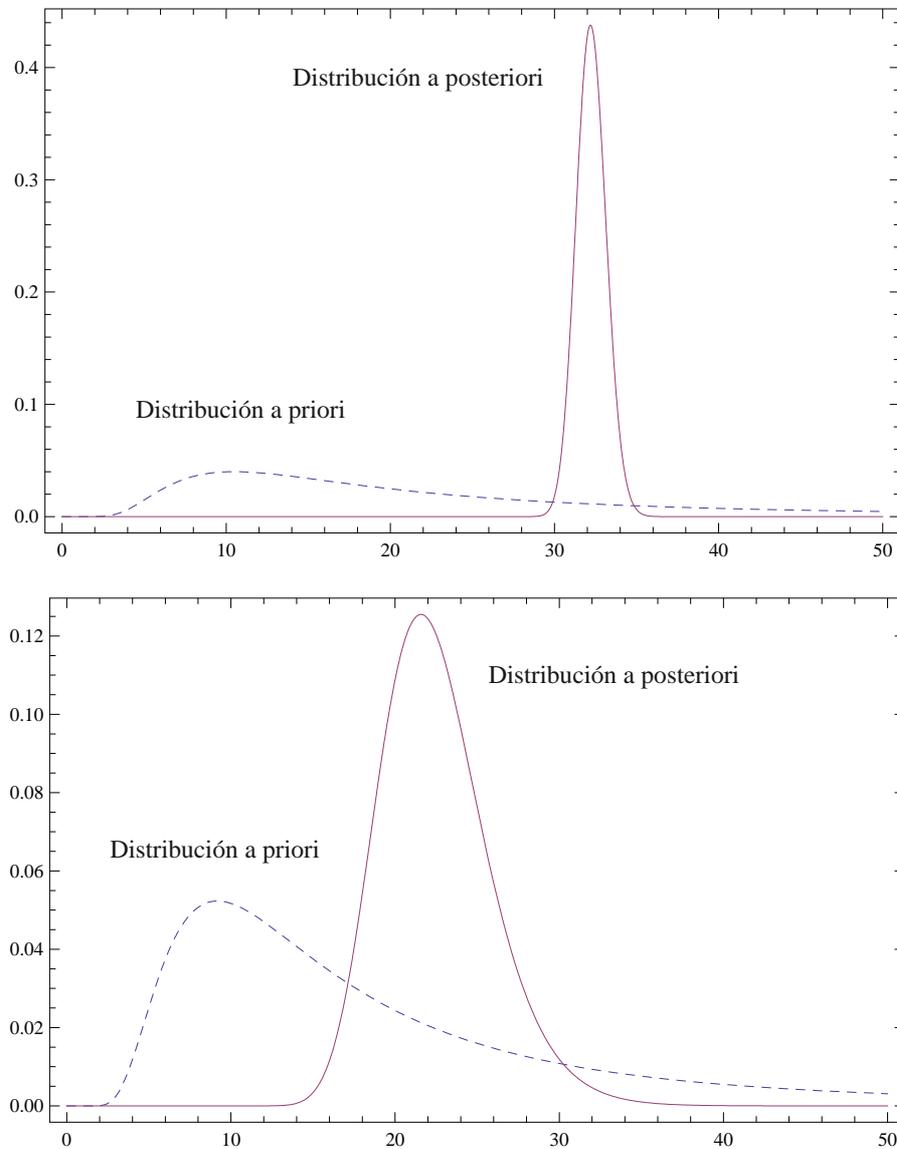


Figura 6.8: Funciones de densidad a priori (línea discontinua) y a posteriori (línea continua) del modelo exponencial-gamma invertida en riesgo operacional tipo 1 (gráfica superior) y riesgos operacionales tipo 4 (gráfico inferior).

Tabla 6.11: Estadísticos para los datos de riesgos operacionales para los tipos de riesgos 1, 2 y 3.

Cantidades	Riesgo 1	Riesgo 2	Riesgo 3
$a$ inicial	1.8129	1.8451	1.9378
$b$ inicial	29.4315	14.7437	8.1359
$a$ final	1248.8100	539.8450	722.9380
$b$ final	40232.7000	8407.5400	5660.7800
$E[\Theta]$	36.2061	17.4459	8.6761
$DT[\Theta]$	40.1575	18.9774	8.9594
$E[X_{n+1} X]$	32.2426	15.6029	7.8411
$DT[X_{n+1} X]$	0.9131	0.6728	0.2920
Factor de credibilidad	0.9993	0.9984	0.9987
$VaR[X; 0,95]$	33.7732	16.7413	8.3334
$VaR[X; 0,99]$	34.4436	17.2544	8.5528
$VaR[X; 0,997]$	34.8670	17.5823	8.6923
$VaR[X; 0,999]$	35.2161	17.8548	8.8079
$VaR[X; 0,9999]$	35.8694	18.3699	9.0256
$TVaR[X; 0,95]$	34.1848	17.0566	8.4682
$TVaR[X; 0,99]$	34.7854	17.5197	8.6656
$TVaR[X; 0,997]$	35.1762	17.8242	8.7949
$TVaR[X; 0,999]$	35.5032	18.0809	8.9035
$TVaR[X; 0,9999]$	36.1235	18.5726	9.1108
Int. credib. inf. 0.95	30.5020	14.3394	7.2894
Int. credib. sup. 0.95	34.0810	16.9760	8.4339

Tabla 6.12: Estadísticos para los datos de riesgos operacionales para los tipos de riesgos 4, 5, 6 y 7.

Cantidades	Riesgo 4	Riesgo 5	Riesgo 6	Riesgo 7
$a$ inicial	2.2107	1.8167	1.8036	1.9565
$b$ inicial	29.4373	33.5436	67.6850	7.8078
$a$ final	47.2107	6527.8200	2396.80	76.9565
$b$ final	1040.1400	238820.00	179477.00	562.0580
$E[\Theta]$	24.3152	41.0700	84.2324	8.1626
$DT[\Theta]$	22.0987	45.4446	93.9663	8.3460
$E[X_{n+1} X]$	22.5086	36.5906	74.9131	7.3997
$DT[X_{n+1} X]$	3.3476	0.4530	1.5308	0.8547
F. de credib.	0.9738	0.9999	0.999665	0.9874
$VaR[X; 0,95]$	28.4928	37.3419	77.4661	8.9049
$VaR[X; 0,99]$	31.8000	37.6608	78.5669	9.6858
$VaR[X; 0,997]$	34.1006	37.8599	79.2582	10.2143
$VaR[X; 0,999]$	36.1329	38.0227	79.8258	10.6719
$VaR[X; 0,9999]$	40.2951	38.3240	80.8822	11.5848
$TVaR[X; 0,95]$	30.5397	37.5376	78.1417	9.3870
$TVaR[X; 0,99]$	33.6952	37.8212	79.1243	10.1185
$TVaR[X; 0,997]$	35.9405	38.0037	79.7602	10.6259
$TVaR[X; 0,999]$	37.9452	38.1552	80.2903	11.0704
$TVaR[X; 0,9999]$	42.0907	38.4398	81.2907	11.9667
Int. credib. inf. 0.95	16.8858	35.7134	71.9722	5.9116
Int. credib. sup. 0.95	29.9627	37.4889	77.9725	9.2552

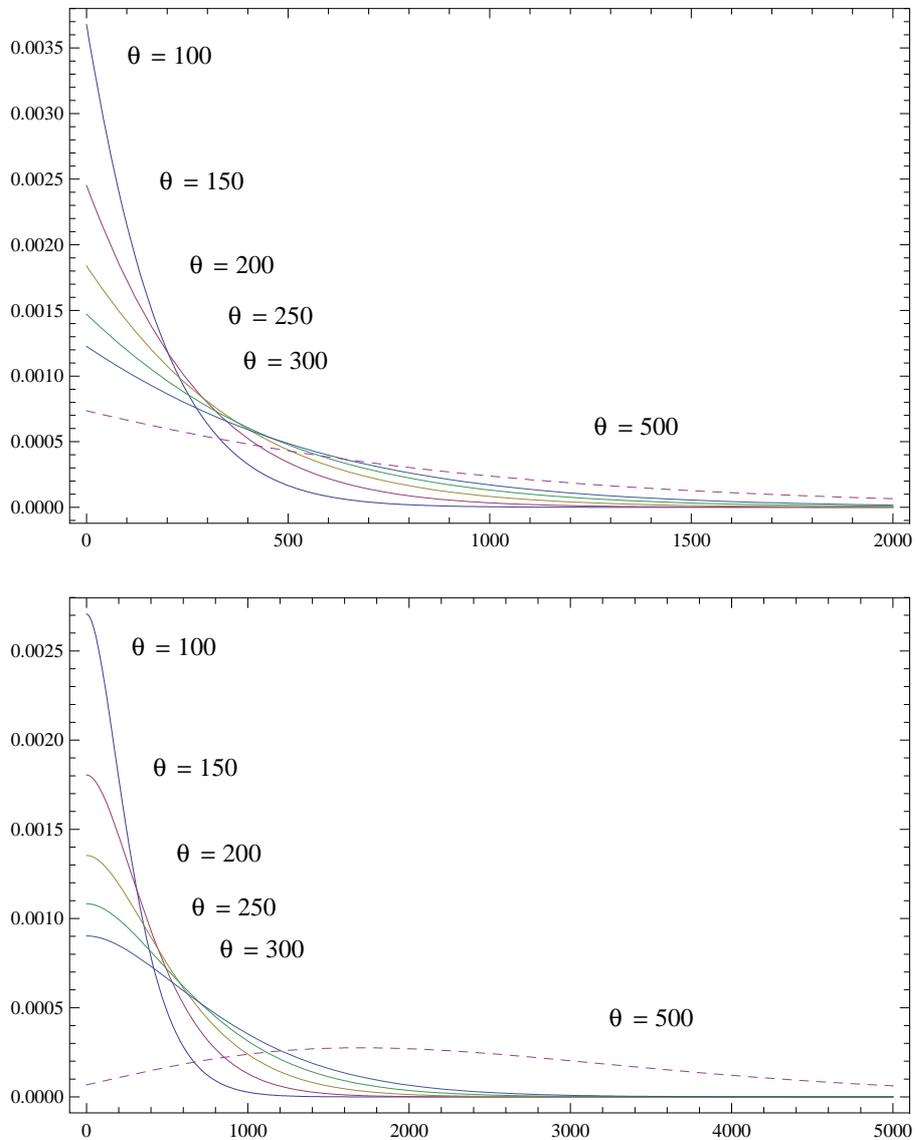


Figura 6.9: Funciones de densidad teóricas del modelo de riesgos operacionales LDA de pérdidas agregadas Poisson-exponencial de frecuencia baja  $\lambda = 1$  (gráfico superior) y  $\lambda = 2$  (gráfico inferior) para diferentes severidades ( $\theta$  es la pérdida media de un suceso). La gráfica con línea discontinua representa el modelo LDA con mayor severidad media.

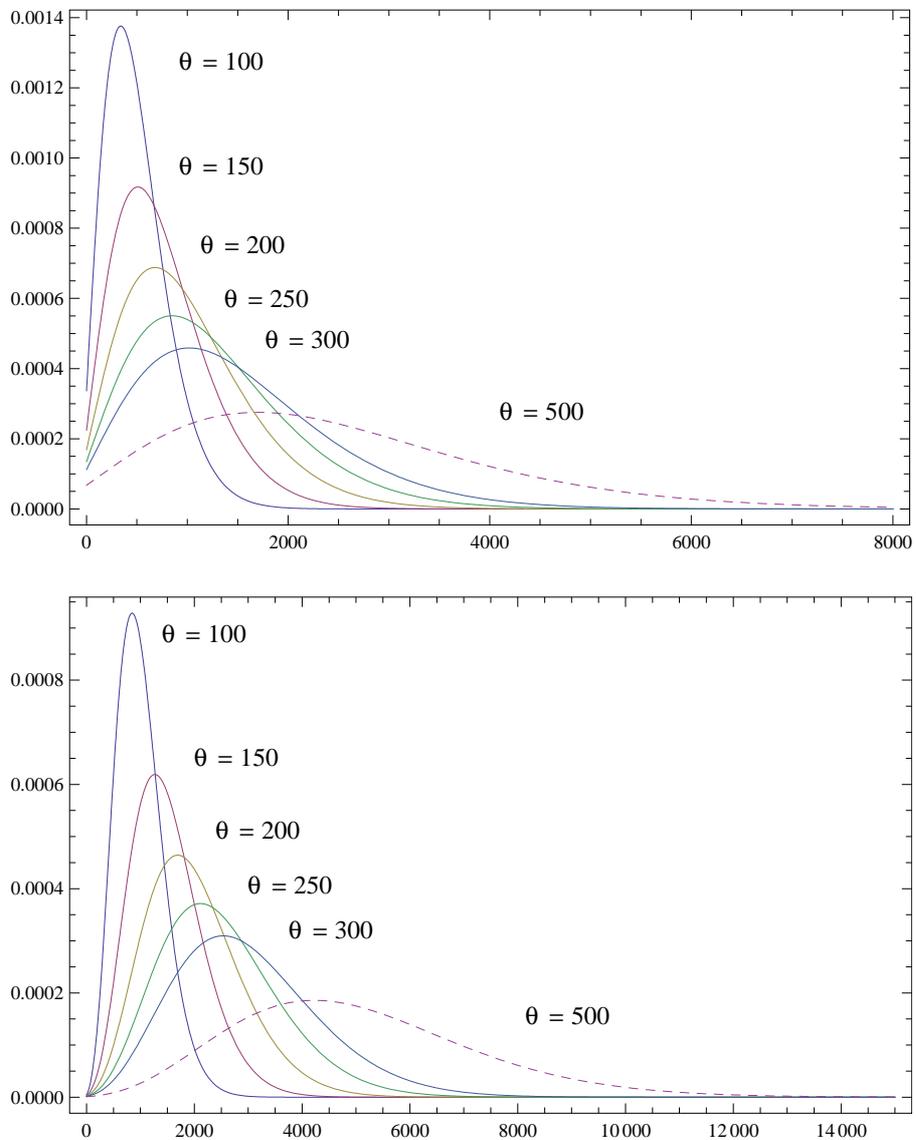


Figura 6.10: Funciones de densidad teóricas del modelo de riesgos operacionales LDA de pérdidas agregadas Poisson-exponencial de frecuencia media  $\lambda = 5$  (gráfico superior) y  $\lambda = 10$  (gráfico inferior) para diferentes severidades ( $\theta$  es la pérdida media de un suceso). La gráfica con línea discontinua representa el modelo LDA con mayor severidad media.

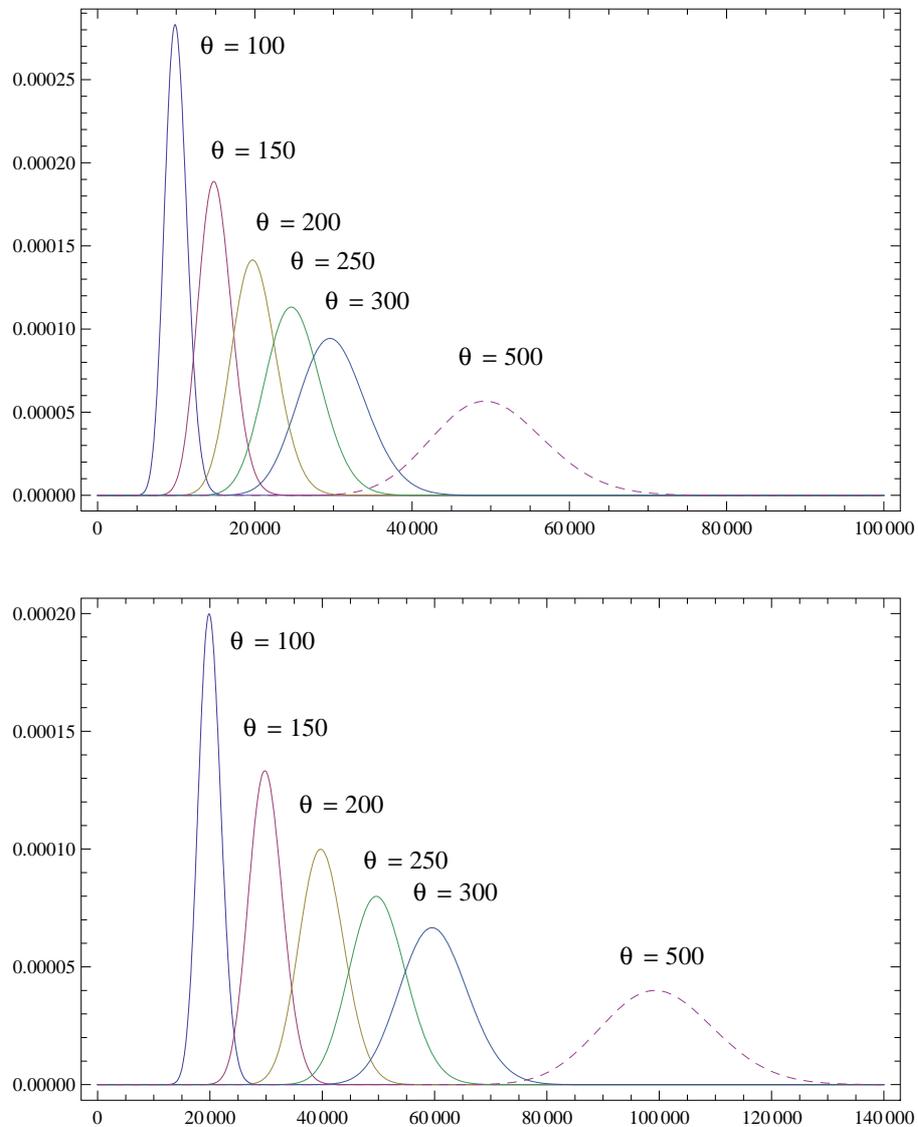


Figura 6.11: Funciones de densidad teóricas del modelo de riesgos operacionales LDA de pérdidas agregadas Poisson-exponencial de frecuencia alta  $\lambda = 100$  (gráfico superior) y  $\lambda = 200$  (gráfico inferior) para diferentes severidades ( $\theta$  es la pérdida media de un suceso). La gráfica con línea discontinua representa el modelo LDA con mayor severidad media.

## Capítulo 7

# Conclusiones

En este trabajo se ha realizado una investigación sobre el desarrollo, análisis y aplicación de modelos de credibilidad en el ámbito de los seguros y en riesgos operacionales. Se han presentado de una manera rigurosa, utilizando las herramientas adecuadas en cada caso, los diferentes modelos de credibilidad, tanto los modelos clásicos como los de desarrollo más reciente. Se han propuesto nuevas aportaciones al tema. La exposición y desarrollo de los diferentes modelos ha venido acompañada con su aplicación en problemas reales de naturaleza actuarial y de riesgos operacionales.

El capítulo 1 se ha dedicado al estudio del concepto de credibilidad, presentando una perspectiva histórica. Las conclusiones son las siguientes:

- La teoría de la credibilidad constituye en la actualidad, uno de los temas centrales de investigación y práctica en seguros.
- El desarrollo histórico de la teoría de la credibilidad ha venido asociado al desarrollo metodológico de obtención de primas de seguros.
- La utilización de la metodología Bayesiana en credibilidad, ha supuesto un importante avance metodológico de esta teoría, y ha estado presente desde sus orígenes.
- La investigación actual en teoría de la credibilidad hace uso intensivo de diversas metodologías de naturaleza matemática y estadística.

En el capítulo 2 se han estudiado principios de primas y medidas de riesgo. Las conclusiones de este capítulo son:

- Los diferentes principios de cálculo de primas, así como la metodología asociada a su obtención, constituye un elemento imprescindible en la tarificación mediante credibilidad.

- De acuerdo con la literatura más reciente, se han descrito un total de 16 propiedades que debería verificar un principio de cálculo de prima.
- Existen cuatro métodos para la obtención de primas: método ad-hoc, método basado en caracterizaciones, método económico y método de las funciones de pérdida.
- Las medidas de riesgo y primas basadas en la cola de la distribución (es decir valor en riesgo VaR y valor en riesgo en la cola TVaR) ocupan un lugar destacado en la investigación actual en seguros y serán igualmente estudiadas en este capítulo.

El capítulo 3 se ha dedicado a la modelización del riesgo en familias paramétricas son. Las conclusiones son:

- La modelización del riesgo por medio de familias paramétricas de distribuciones constituye la metodología clásica de establecer la población de referencia tanto en seguros como en teoría de la credibilidad.
- Se han estudiado de forma detallada las siguientes distribuciones como modelos de riesgo: distribuciones de Pareto, distribuciones exponencial, gamma y gamma invertida, distribuciones normal y lognormal, distribución de Weibull y distribución inversa Gaussiana.
- Para las distribuciones de riesgo anteriores, se han obtenido expresiones cerradas para el valor en riesgo VaR y el valor en riesgo en la cola TVaR.
- El análisis de las colas de una distribución es un aspecto clave en la decisión de cuál distribución elegir.
- Se ha introducido un modelo de regresión (modelización mediante covariables) basado en la distribución inversa Gaussiana.
- Los modelos anteriores han sido estimados y validados mediante técnicas adecuadas con datos reales de pérdidas por huracanes.

Los instrumentos de estadística Bayesiana y las primas Bayes han sido estudiadas en el capítulo 4. Las conclusiones son las siguientes:

- la metodología Bayesiana constituye el marco idóneo para el desarrollo de nuevos modelos de credibilidad. Dicha metodología se está aplicando con éxito tanto en la práctica actuarial moderna como en la teoría de la credibilidad.
- Se ha utilizado la metodología Bayesiana para la estimación de diversas familias paramétricas, entre ellas la distribución clásica de Pareto.

- Se ha presentado una importante clase de distribuciones de pérdida (sección 4.6) y se ha realizado su correspondiente estimación Bayesiana. Los estimadores obtenidos se puede escribir en términos de una fórmulas de credibilidad.
- La metodología Bayesiana ha sido utilizada para estimar la prima de riesgo a través, primero de la prima colectiva y, posteriormente mediante la prima Bayes.
- Las familias de distribuciones exponencial y exponencial de dispersión constituyen dos importantes herramientas para la obtención de primas de credibilidad.

En el capítulo 5 se han estudiado y propuestos diversos modelos de credibilidad. Las conclusiones son las siguientes:

- El modelo de Bühlmann (1967) de distribución libre constituye el punto de partida de la moderna teoría de la credibilidad. Dicho modelo no precisa de ninguna hipótesis sobre la distribución de los parámetros, y da lugar a la fórmula clásica de credibilidad, junto con la expresión del factor de credibilidad, donde se pondera tanto la experiencia individual como la información disponible del colectivo.
- El modelo de Bühlmann-Straub (1972) introduce observaciones ponderadas, mejorando el modelo precedente.
- Tomando como punto de partida el modelo de Bühlmann (1967), se han desarrollado numerosos modelos de credibilidad, basados en metodología Bayesiana. El análisis Bayesiano constituye una herramienta imprescindible cuando se quieren conjugar las dos fuentes de información de las que se disponen en el escenario actuarial, esto es, la información individual o de una póliza en particular y la información colateral, usualmente la de la cartera de seguros. Con esta metodología es posible obtener expresiones de credibilidad cuando se utilizan pares de verosimilitudes y distribuciones a priori ampliamente utilizadas en la práctica.
- En un trabajo pionero Jewell (1974) y posteriormente Landsman y Makov (1998) prueban que la expresión básica de credibilidad aparece siempre que se utilice la familia de distribución exponencial, en el primer caso, y la familia exponencial de dispersión en el segundo. También se ha demostrado que es posible obtener nuevas expresiones de credibilidad utilizando la función de pérdida cuadrática ponderada y equilibrada. Se obtiene entonces un modelo mucho más general que de nuevo incluye a los anteriores.

- Se han revisado modelos de credibilidad basados en funciones de pérdida generales, como el propuesto por Gómez-Déniz (2008a) construido a partir de la función de pérdida cuadrática ponderada y equilibrada. Este modelo es más general que muchos de los modelos anteriores, en el sentido que los incluye como caso particular.
- La metodología Bayesiana no está libre de críticas, debido en buena parte a la utilización de la distribución a priori, que puede ser elegida libremente por el actuario. Por esta razón se han propuesto otros modelos que tratan de salvar este posible inconveniente. El primero de ellos es el modelo jerárquico, que permite asignar una distribución a priori (hiper-priori) a alguno de los parámetros de la distribución a priori. De este modo la elección de ésta ya no puede ser vista como arbitraria, si la distribución hiper-priori que se elija es del tipo no informativa.
- En esta investigación se han desarrollado tres modelos jerárquicos diferentes. El primero de ellos así como el modelo jerárquico normal son totalmente novedosos, basados en algunos modelos de Klugman (1992). Los resultados obtenidos por los nuevos modelos mejoran los resultados obtenidos por Klugman.
- Otra vía alternativa para salvar la crítica a la metodología Bayesiana la constituyen los modelos de credibilidad basados en el análisis de robustez Bayesiana, los basados en la aproximación gamma-minimax y posterior regret gamma-minimax. En el último caso se ha presentado un nuevo modelo que constituye una generalización del modelo de Jewell (1974) y del de Landsman y Makov (1998), y que incluye a éstos como casos particulares.

Finalmente, el capítulo 6 se han presentado aplicaciones de los modelos de credibilidad en riesgos operacionales. Se ha realizado una reflexión sobre la metodología a utilidad, fundamentalmente de carácter Bayesiano, donde las fórmulas de predicción basadas en la teoría de la credibilidad juegan un papel importante. Destacamos las siguientes conclusiones:

- Para la modelización de la frecuencia de sucesos debidos a riesgos operacionales de debe combinar tanto la información de los datos como la experiencia del sector. A este respecto, los modelos presentados Poisson-gamma y binomial-beta recogen estas características.
- Los modelos para estudiar la severidad, deben igualmente combinar los dos tipos de información: muestral y del colectivo. Con este propósito, los modelos de credibilidad parecen adecuados. Los modelos lognormal-normal, Pareto generalizada-gamma, exponencial-gamma invertida y Pareto clásica-gamma resultan adecuados dependiendo de la naturaleza de la información disponible.

En todos estos casos existe una fórmula de credibilidad para realizar la predicción.

- Siguiendo las recomendaciones de Basilea II, los modelos de pérdidas agregadas (proceso LDA) son especialmente útiles. Para el ajuste y cómputo de la distribución agregada, se conocen muchas de sus propiedades, y es posible obtener las distribuciones exactas.
- Se han obtenido y descrito diversas distribuciones exactas de modelos LDA, que permiten obtener probabilidades y momentos exactos, sin necesidad de recurrir a técnicas e simulación.
- Los modelos anteriores han sido estimados con información real de pérdidas debidas a riesgos operacionales. Se han obtenido estimadores de credibilidad para la predicción del siguiente período.

# Apéndice A

## Datos de Klugman (1991)

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
1	1	32,322	1	5	1	310,389	15
1	2	33,779	4	5	2	292,464	10
1	3	43,548	3	5	3	262,560	18
1	4	46,686	5	5	4	273,257	25
1	5	34,713	1	5	5	372,730	27
1	6	32,857	3	5	6	263,443	20
1	7	36,600	4	5	7	355,826	28
2	1	45,995	3	6	1	291,784	27
2	2	37,888	1	6	2	267,439	44
2	3	34,581	0	6	3	288,555	40
2	4	28,298	0	6	4	308,364	31
2	5	45,265	2	6	5	387,095	40
2	6	39,945	0	6	6	351,445	34
2	7	39,322	4	6	7	342,518	37
3	1	289,047	5	7	1	0,000	0
3	2	392,176	8	7	2	0,000	0
3	3	368,982	8	7	3	0,000	0
3	4	323,770	8	7	4	0,000	0
3	5	385,222	16	7	5	0,000	0
3	6	346,390	8	7	6	0,000	0
3	7	324,132	9	7	7	0,000	0
4	1	0,000	0	8	1	41,403	3
4	2	0,000	0	8	2	34,066	11
4	3	0,000	0	8	3	32,729	4
4	4	0,037	0	8	4	32,235	0
4	5	0,000	0	8	5	32,777	2
4	6	0,000	0	8	6	28,620	4
4	7	0,000	0	8	7	24,263	1

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
9	1	6,452	0	15	1	215,898	11
9	2	6,927	1	15	2	224,229	11
9	3	5,851	0	15	3	224,306	10
9	4	3,033	1	15	4	212,232	6
9	5	1,787	0	15	5	245,198	11
9	6	1,074	1	15	6	203,698	16
9	7	0,685	0	15	7	210,496	15
10	1	537,311	11	16	1	57,265	4
10	2	569,041	12	16	2	62,662	3
10	3	597,146	15	16	3	66,984	6
10	4	570,295	15	16	4	63,211	3
10	5	1116,750	18	16	5	77,680	4
10	6	774,454	10	16	6	61,944	3
10	7	534,953	11	16	7	44,195	3
11	1	149,683	6	17	1	347,835	22
11	2	157,947	6	17	2	326,396	7
11	3	174,549	5	17	3	307,978	13
11	4	181,317	10	17	4	350,914	20
11	5	202,066	13	17	5	546,047	23
11	6	187,564	7	17	6	410,980	8
11	7	229,830	8	17	7	377,287	18
12	1	609,467	20	18	1	0,000	0
12	2	645,375	20	18	2	0,000	0
12	3	667,384	27	18	3	0,000	0
12	4	573,144	15	18	4	0,000	0
12	5	782,643	38	18	5	0,000	0
12	6	478,749	8	18	6	0,000	0
12	7	454,967	13	18	7	0,000	0
13	1	120,027	0	19	1	77,909	3
13	2	131,020	1	19	2	78,925	0
13	3	161,145	8	19	3	77,558	1
13	4	182,135	5	19	4	69,525	1
13	5	276,520	11	19	5	65,972	0
13	6	158,310	2	19	6	63,274	0
13	7	168,420	7	19	7	75,335	2
14	1	196,722	13	20	1	1787,463	50
14	2	209,923	12	20	2	2027,230	90
14	3	196,199	13	20	3	1853,812	54
14	4	188,820	14	20	4	1742,135	63
14	5	202,807	19	20	5	2119,499	56
14	6	180,979	8	20	6	1545,169	35
14	7	195,628	6	20	7	1315,368	22

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
21	1	258,885	10	27	1	131,306	1
21	2	314,825	10	27	2	130,015	2
21	3	299,999	10	27	3	138,982	1
21	4	314,081	11	27	4	131,359	5
21	5	258,487	17	27	5	187,921	6
21	6	235,082	8	27	6	148,775	3
21	7	230,142	8	27	7	134,179	3
22	1	86,381	4	28	1	103,473	3
22	2	95,036	4	28	2	102,280	0
22	3	134,285	4	28	3	111,964	1
22	4	113,491	8	28	4	115,853	2
22	5	141,717	0	28	5	164,759	3
22	6	128,956	0	28	6	138,420	3
22	7	105,489	1	28	7	123,096	0
23	1	766,953	7	29	1	129,504	5
23	2	912,516	8	29	2	150,853	13
23	3	866,413	9	29	3	180,207	8
23	4	898,687	15	29	4	226,038	21
23	5	1806,752	34	29	5	499,272	33
23	6	1018,684	14	29	6	344,642	13
23	7	984,843	11	29	7	436,146	21
24	1	62,153	1	30	1	18,708	0
24	2	84,116	1	30	2	16,761	2
24	3	75,524	1	30	3	9,341	0
24	4	109,237	1	30	4	11,269	0
24	5	119,034	2	30	5	2,485	0
24	6	116,794	1	30	6	2,056	0
24	7	136,571	3	30	7	1,556	0
25	1	3431,494	59	31	1	133,124	6
25	2	3882,069	68	31	2	156,835	13
25	3	3805,563	76	31	3	256,429	20
25	4	3919,527	77	31	4	253,184	9
25	5	4352,809	94	31	5	157,622	12
25	6	3949,550	73	31	6	75,804	2
25	7	3927,784	73	31	7	71,884	4
26	1	32,582	2	32	1	220,004	13
26	2	24,922	2	32	2	206,989	15
26	3	30,317	2	32	3	222,724	15
26	4	29,732	2	32	4	268,898	12
26	5	49,201	0	32	5	291,268	39
26	6	34,379	4	32	6	290,797	22
26	7	21,536	0	32	7	244,777	20

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
33	1	3,769	0	39	1	239,966	26
33	2	1,985	0	39	2	242,634	27
33	3	3,108	0	39	3	295,803	31
33	4	3,046	0	39	4	281,269	25
33	5	14,316	0	39	5	376,989	36
33	6	10,908	1	39	6	234,967	20
33	7	10,516	0	39	7	218,695	18
34	1	283,926	8	40	1	29,923	1
34	2	217,928	15	40	2	33,448	3
34	3	229,275	18	40	3	32,704	2
34	4	224,755	0	40	4	40,210	2
34	5	427,331	4	40	5	53,806	0
34	6	384,840	1	40	6	45,562	1
34	7	420,079	1	40	7	45,396	0
35	1	9,159	0	41	1	1100,645	54
35	2	4,492	0	41	2	1241,328	54
35	3	7,229	0	41	3	1185,711	48
35	4	9,169	0	41	4	1251,523	35
35	5	8,106	0	41	5	1420,391	44
35	6	12,594	0	41	6	1281,866	82
35	7	8,913	0	41	7	1170,673	30
36	1	981,952	22	42	1	43,101	1
36	2	527,044	17	42	2	34,826	0
36	3	533,930	21	42	3	47,346	2
36	4	451,839	19	42	4	70,784	2
36	5	352,529	20	42	5	26,081	0
36	6	230,163	16	42	6	77,421	0
36	7	243,936	19	42	7	7,120	2
37	1	543,295	12	43	1	1348,008	48
37	2	540,957	20	43	2	1408,717	59
37	3	515,061	22	43	3	1616,470	65
37	4	534,942	12	43	4	1485,584	73
37	5	620,379	13	43	5	1954,211	91
37	6	553,122	17	43	6	1528,900	62
37	7	510,214	9	43	7	1252,694	43
38	1	1029,563	67	44	1	1123,507	43
38	2	1097,433	58	44	2	1148,065	37
38	3	1272,550	68	44	3	1198,427	36
38	4	1357,884	92	44	4	1258,329	56
38	5	1622,744	68	44	5	1290,727	70
38	6	1160,147	77	44	6	1213,667	26
38	7	1056,315	57	44	7	1209,309	47

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
45	1	1558,521	50	51	1	14,229	0
45	2	1798,188	54	51	2	13,704	0
45	3	1804,204	48	51	3	15,445	1
45	4	1963,514	81	51	4	25,915	0
45	5	2758,736	122	51	5	38,275	0
45	6	2266,418	55	51	6	27,823	0
45	7	2135,638	77	51	7	18,472	0
46	1	1677,159	84	52	1	89,758	4
46	2	1818,022	132	52	2	118,457	1
46	3	1716,822	134	52	3	126,104	3
46	4	1828,931	94	52	4	106,947	5
46	5	2389,001	141	52	5	117,231	7
46	6	1838,996	93	52	6	101,401	2
46	7	1586,483	106	52	7	57,373	6
47	1	107,175	0	53	1	149,060	1
47	2	95,723	2	53	2	189,778	2
47	3	78,302	0	53	3	152,892	3
47	4	37,141	0	53	4	186,371	3
47	5	33,023	1	53	5	245,861	3
47	6	31,017	0	53	6	216,972	3
47	7	43,081	0	53	7	104,665	6
48	1	113,138	5	54	1	0,000	0
48	2	148,338	15	54	2	0,000	0
48	3	199,211	9	54	3	0,000	0
48	4	245,298	12	54	4	0,075	0
48	5	305,085	13	54	5	0,000	0
48	6	282,832	14	54	6	0,000	0
48	7	179,536	15	54	7	0,000	0
49	1	389,615	11	55	1	161,352	19
49	2	352,236	7	55	2	232,134	16
49	3	465,380	12	55	3	145,447	14
49	4	458,031	12	55	4	178,502	12
49	5	686,360	12	55	5	220,084	9
49	6	707,343	9	55	6	111,862	2
49	7	499,189	16	55	7	119,294	10
50	1	85,845	2	56	1	15,655	0
50	2	85,783	0	56	2	18,035	0
50	3	104,434	0	56	3	23,572	1
50	4	99,875	2	56	4	26,767	0
50	5	103,061	0	56	5	0,201	0
50	6	96,433	2	56	6	21,955	0
50	7	78,010	3	56	7	23,234	3

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
57	1	117,163	6	63	1	3,119	0
57	2	134,122	13	63	2	3,685	0
57	3	139,911	6	63	3	3,764	0
57	4	160,603	8	63	4	3,831	0
57	5	188,862	10	63	5	4,993	1
57	6	210,026	10	63	6	3,780	0
57	7	204,768	3	63	7	2,618	0
58	1	150,296	0	64	1	28,386	0
58	2	166,081	3	64	2	32,964	1
58	3	155,572	0	64	3	44,084	0
58	4	150,686	0	64	4	81,340	2
58	5	225,843	3	64	5	89,199	1
58	6	73,296	1	64	6	81,750	0
58	7	191,223	1	64	7	87,903	2
59	1	273,622	2	65	1	91,605	1
59	2	312,386	4	65	2	112,329	1
59	3	384,966	7	65	3	90,227	2
59	4	408,276	4	65	4	115,104	3
59	5	748,580	11	65	5	106,548	2
59	6	426,061	5	65	6	89,530	3
59	7	376,149	3	65	7	101,780	1
60	1	196,698	4	66	1	64,214	0
60	2	217,671	3	66	2	81,465	0
60	3	239,066	6	66	3	66,600	0
60	4	235,468	5	66	4	82,745	0
60	5	366,125	3	66	5	94,381	1
60	6	280,203	4	66	6	86,545	1
60	7	202,221	2	66	7	92,120	1
61	1	0,000	0	67	1	113,122	2
61	2	0,000	0	67	2	141,183	6
61	3	0,288	0	67	3	186,196	12
61	4	0,433	1	67	4	214,230	12
61	5	1,312	0	67	5	193,758	10
61	6	1,268	0	67	6	221,813	10
61	7	0,806	0	67	7	201,130	8
62	1	128,762	6	68	1	7,329	0
62	2	129,566	5	68	2	13,180	0
62	3	156,728	2	68	3	8,650	0
62	4	170,342	4	68	4	7,822	1
62	5	205,975	5	68	5	13,752	0
62	6	161,752	1	68	6	9,581	0
62	7	128,477	1	68	7	10,006	1

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
69	1	174,102	2	75	1	259,342	10
69	2	201,453	3	75	2	282,100	31
69	3	209,368	5	75	3	265,376	34
69	4	253,491	6	75	4	254,058	27
69	5	265,070	2	75	5	335,424	33
69	6	269,752	5	75	6	244,496	23
69	7	192,365	8	75	7	214,138	20
70	1	38,714	0	76	1	646,899	31
70	2	40,550	0	76	2	706,321	27
70	3	48,034	0	76	3	699,304	25
70	4	47,987	0	76	4	705,517	34
70	5	58,351	0	76	5	867,373	46
70	6	54,275	0	76	6	638,322	38
70	7	54,812	0	76	7	572,799	56
71	1	83,008	3	77	1	107,353	5
71	2	82,796	3	77	2	104,486	8
71	3	99,997	0	77	3	124,194	12
71	4	106,843	5	77	4	175,424	7
71	5	145,506	4	77	5	206,059	12
71	6	110,815	5	77	6	177,489	7
71	7	103,664	5	77	7	167,093	9
72	1	121,196	10	78	1	77,171	3
72	2	148,555	6	78	2	76,506	6
72	3	186,487	6	78	3	98,673	6
72	4	309,997	21	78	4	103,028	13
72	5	265,607	14	78	5	122,334	16
72	6	228,599	10	78	6	57,534	4
72	7	107,859	7	78	7	80,913	12
73	1	66,966	2	79	1	263,738	28
73	2	65,204	5	79	2	308,347	27
73	3	79,769	2	79	3	296,709	18
73	4	73,765	0	79	4	303,772	23
73	5	108,141	6	79	5	377,930	39
73	6	84,860	3	79	6	185,310	14
73	7	90,598	4	79	7	172,616	26
74	1	820,970	64	80	1	1013,867	39
74	2	1071,589	57	80	2	956,509	35
74	3	851,567	65	80	3	1041,898	39
74	4	772,685	41	80	4	1122,146	34
74	5	1126,973	65	80	5	1400,163	51
74	6	416,385	36	80	6	1034,801	41
74	7	308,005	16	80	7	891,256	56

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
81	1	111,090	9	87	1	28,988	3
81	2	106,213	10	87	2	21,153	8
81	3	102,541	7	87	3	19,404	4
81	4	113,139	12	87	4	13,989	1
81	5	168,177	13	87	5	5,164	1
81	6	149,237	10	87	6	10,733	0
81	7	89,148	13	87	7	3,488	0
82	1	2913,910	82	88	1	7,247	0
82	2	2941,158	106	88	2	8,407	0
82	3	3413,532	81	88	3	10,825	0
82	4	3402,478	104	88	4	6,140	0
82	5	4313,542	132	88	5	18,044	1
82	6	3024,050	96	88	6	6,229	1
82	7	2844,038	89	88	7	3,670	0
83	1	3175,937	224	89	1	106,135	30
83	2	3058,315	148	89	2	93,503	36
83	3	3226,647	160	89	3	96,455	26
83	4	3575,413	199	89	4	118,534	73
83	5	4517,677	228	89	5	112,757	54
83	6	3478,879	178	89	6	93,584	48
83	7	2593,355	178	89	7	79,629	40
84	1	79,582	4	90	1	76,405	6
84	2	58,306	7	90	2	80,802	1
84	3	100,149	11	90	3	82,884	9
84	4	78,744	10	90	4	101,123	11
84	5	231,495	24	90	5	164,261	6
84	6	276,840	11	90	6	95,666	7
84	7	459,901	34	90	7	89,050	4
85	1	84,886	9	91	1	1328,932	31
85	2	94,043	3	91	2	1334,186	33
85	3	98,463	14	91	3	1334,642	32
85	4	132,101	13	91	4	1439,621	55
85	5	211,462	17	91	5	1900,358	57
85	6	172,669	28	91	6	1575,137	33
85	7	194,281	19	91	7	1495,906	39
86	1	0,000	0	92	1	1283,025	54
86	2	21,319	0	92	2	1337,012	61
86	3	20,923	0	92	3	1355,122	68
86	4	18,530	0	92	4	1519,127	91
86	5	27,521	1	92	5	1963,594	53
86	6	10,468	0	92	6	1524,377	80
86	7	11,961	0	92	7	1419,202	76

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
93	1	306,705	3	99	1	1055,156	25
93	2	314,853	1	99	2	1065,292	30
93	3	357,296	0	99	3	1141,624	35
93	4	439,514	4	99	4	1193,582	48
93	5	602,841	4	99	5	1615,875	61
93	6	554,907	3	99	6	1337,935	49
93	7	571,591	4	99	7	1269,005	38
94	1	209,037	1	100	1	309,501	10
94	2	249,768	3	100	2	347,849	16
94	3	242,952	3	100	3	339,384	16
94	4	223,889	3	100	4	323,840	11
94	5	286,412	12	100	5	410,613	15
94	6	247,946	6	100	6	325,568	8
94	7	246,528	4	100	7	301,404	9
95	1	598,297	12	101	1	1768,157	12
95	2	633,966	15	101	2	1809,730	18
95	3	650,776	12	101	3	1096,204	20
95	4	761,772	12	101	4	2053,998	20
95	5	863,569	27	101	5	2942,043	25
95	6	784,155	15	101	6	2383,743	30
95	7	748,876	18	101	7	2327,415	27
96	1	511,401	41	102	1	317,952	6
96	2	609,635	27	102	2	318,712	18
96	3	518,000	11	102	3	305,967	16
96	4	561,834	9	102	4	342,672	7
96	5	617,042	10	102	5	463,343	19
96	6	448,895	6	102	6	340,908	13
96	7	501,484	8	102	7	329,476	10
97	1	591,968	26	103	1	183,162	12
97	2	609,840	23	103	2	204,070	11
97	3	641,186	32	103	3	215,898	15
97	4	673,982	41	103	4	254,076	22
97	5	801,838	40	103	5	309,034	16
97	6	592,050	44	103	6	270,101	21
97	7	516,338	24	103	7	244,806	24
98	1	3307,708	46	104	1	49,076	6
98	2	3156,211	42	104	2	46,452	3
98	3	3223,638	65	104	3	53,448	3
98	4	3269,966	60	104	4	45,917	0
98	5	4895,145	103	104	5	66,054	0
98	6	3694,214	75	104	6	52,828	3
98	7	3564,236	55	104	7	43,959	1

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
105	1	2126,661	63	111	1	63,785	6
105	2	2273,041	63	111	2	67,671	3
105	3	2325,099	77	111	3	77,918	2
105	4	2452,336	61	111	4	82,328	6
105	5	3209,569	93	111	5	105,884	4
105	6	2439,080	68	111	6	75,422	5
105	7	2236,104	62	111	7	70,851	10
106	1	39,867	1	112	1	12075,618	29
106	2	39,843	2	112	2	12618,147	24
106	3	42,824	1	112	3	13622,255	21
106	4	37,455	1	112	4	14833,854	26
106	5	54,205	3	112	5	21163,600	45
106	6	40,643	0	112	6	19070,066	31
106	7	37,574	2	112	7	18809,666	45
107	1	231,292	0	113	1	2582,738	18
107	2	250,281	1	113	2	3288,781	25
107	3	241,722	1	113	3	3107,033	25
107	4	245,193	0	113	4	3201,166	17
107	5	448,444	0	113	5	4361,026	46
107	6	341,276	0	113	6	3310,561	33
107	7	295,566	0	113	7	2020,386	12
108	1	679,884	14	114	1	3178,428	3
108	2	906,121	17	114	2	2294,945	11
108	3	950,964	7	114	3	2563,539	8
108	4	977,763	9	114	4	2828,968	6
108	5	1340,554	20	114	5	4422,089	9
108	6	1026,951	24	114	6	3809,180	14
108	7	969,722	31	114	7	4112,043	9
109	1	53,708	3	115	1	142,476	2
109	2	67,571	2	115	2	153,847	0
109	3	63,940	5	115	3	172,002	4
109	4	71,650	1	115	4	186,035	5
109	5	98,161	4	115	5	202,441	1
109	6	79,249	7	115	6	185,968	2
109	7	59,117	3	115	7	191,504	4
110	1	44,035	1	116	1	818,162	6
110	2	48,523	2	116	2	831,213	1
110	3	57,823	3	116	3	911,948	18
110	4	45,613	3	116	4	1031,709	13
110	5	72,047	1	116	5	1439,935	21
110	6	30,096	1	116	6	1184,156	8
110	7	24,968	2	116	7	1254,184	9

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
117	1	471,076	2	123	1	104,253	2
117	2	261,442	6	123	2	138,179	2
117	3	439,570	10	123	3	139,056	2
117	4	421,738	2	123	4	172,297	5
117	5	842,236	6	123	5	187,190	5
117	6	96,233	4	123	6	132,517	5
117	7	72,651	0	123	7	142,657	2
118	1	90,679	2	124	1	3,275	0
118	2	89,372	3	124	2	5,693	0
118	3	75,521	3	124	3	4,508	0
118	4	98,495	7	124	4	4,805	0
118	5	119,188	3	124	5	9,067	0
118	6	94,799	2	124	6	5,967	0
118	7	70,064	3	124	7	5,588	0
119	1	5100,537	23	125	1	13,182	3
119	2	5307,799	26	125	2	20,734	2
119	3	2885,003	31	125	3	15,334	3
119	4	6305,052	22	125	4	17,762	1
119	5	8304,123	41	125	5	17,228	0
119	6	7028,993	46	125	6	1,666	11
119	7	6541,694	37	125	7	1,740	1
120	1	1310,079	24	126	1	76,205	1
120	2	1596,148	38	126	2	70,900	0
120	3	1577,690	34	126	3	87,235	2
120	4	1671,204	38	126	4	100,693	0
120	5	2328,100	73	126	5	143,081	0
120	6	2066,956	63	126	6	120,950	0
120	7	1728,151	50	126	7	123,444	2
121	1	66,772	0	127	1	180,970	4
121	2	70,669	0	127	2	249,116	5
121	3	70,737	0	127	3	221,405	7
121	4	73,274	0	127	4	161,425	7
121	5	87,124	2	127	5	199,852	6
121	6	83,261	0	127	6	172,033	6
121	7	89,242	0	127	7	140,378	7
122	1	2113,111	17	128	1	0,000	0
122	2	2294,597	38	128	2	0,000	0
122	3	2585,560	38	128	3	0,000	0
122	4	2524,492	53	128	4	0,000	0
122	5	3616,607	78	128	5	0,000	0
122	6	2887,929	81	128	6	0,000	0
122	7	2870,043	67	128	7	0,000	0

Clase	Año	$P_{ij}$	$Y_{ij}$	Clase	Año	$P_{ij}$	$Y_{ij}$
129	1	795,167	19	132	1	2,317	0
129	2	1147,475	26	132	2	3,972	0
129	3	950,837	27	132	3	33,190	0
129	4	849,241	15	132	4	47,261	2
129	5	976,607	22	132	5	35,413	2
129	6	673,329	16	132	6	16,984	1
129	7	535,045	15	132	7	18,797	0
130	1	82,610	0	133	1	66,213	2
130	2	87,124	2	133	2	65,696	0
130	3	111,903	0	133	3	69,913	0
130	4	123,734	1	133	4	46,289	0
130	5	120,606	3	133	5	91,119	0
130	6	148,423	1	133	6	78,456	1
130	7	70,599	0	133	7	84,362	2
131	1	61,802	0				
131	2	116,093	2				
131	3	64,425	2				
131	4	56,817	0				
131	5	92,627	1				
131	6	119,006	6				
131	7	120,145	8				

# Bibliografía

- [1] Alexander, C. (2007). Valor en riesgo operacional. En: *La Gestión del Riesgo Operacional. De la Teoría a su Aplicación*, Fernández-Laviada, A. (Edit.), pp. 389–417. Ediciones 2010, Limusa, Madrid.
- [2] Ambagaspitiya, R.S., Balakrishnan, N. (1994). On the compound generalized Poisson distributions. *Astin Bulletin*, 24, 255–263.
- [3] Antal, P. (2001). *Quantitative Methods in Reinsurance*. Swiss Re, Switzerland.
- [4] Arnold, B.C. (1983). *Pareto Distributions*. International Co-operative Publishing House, Fairland, MD.
- [5] Arnold, B.C., Castillo, E., Sarabia, J.M. (1992). *Conditionally Specified Distributions*. Lecture Notes in Statistics, Vol. 73. Springer-Verlag, New York.
- [6] Arnold, B.C., Castillo, E., Sarabia, J.M. (1998). Bayesian analysis for classical distributions using conditionally specified priors. *Sankhya, Ser. B*, 60, 228–245.
- [7] Arnold, B.C., Castillo, E., Sarabia, J.M. (1999). *Conditional Specification of Statistical Models*. Springer-Verlag, New York.
- [8] Arnold, B.C., Press, S.J. (1989). Bayesian estimation and prediction for Pareto data. *Journal of the American Statistical Association*, 84, 1079–1084.
- [9] Artzner, P. (1999). Application of coherent risk measures to capital requirements in insurance. *North American Actuarial Journal*, 3, 11–29.
- [10] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). Coherent risk measures. *Mathematical Finance*, 9, 203–228.
- [11] Bailey, A. (1945). A generalized theory of credibility. *Proceedings of the Casualty Actuarial Society*, 32, 13–20.
- [12] Baione, F., Levantesi, S., Menzietti, M. (2002). The development of an optimal Bonus-Malus system in a competitive market. *Astin Bulletin*, 32, 159–169.

- [13] Balakrishnan, N., Nevzorov, V.B. (2003). *A Primer on Statistical Distributions*. John Wiley, New York.
- [14] Balbás, A., Gil, J.A., Heras, A. (1990). La desviación típica y la varianza como medidas del riesgo en un problema de reaseguro óptimo. *Previsión y Seguro*, **6**, 63–80.
- [15] Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Second Ed. Springer-Verlag, New York.
- [16] Bernardo, J.M., Smith, A.F.M. (1994). *Bayesian Theory*. John Wiley, New York.
- [17] Bertram, J. (1981). Numerische berechnung von gesamtschadenverteilungen. *Blatter der DGVM*, **15**, 175–194.
- [18] BIS (2005). *Basel II. International Convergence of Capital Measurement and Capital Standards: a revised framework Bank for International Settlements (BIS)*, [www.bis.org](http://www.bis.org)
- [19] Boucher, J.-P. Denuit, M. (2007). Credibility premiums for the zero-inflated Poisson model and new hunger for bonus interpretation. *Insurance: Mathematics and Economics* (en prensa).
- [20] Bühlmann, H. (1967). Experience rating and credibility. *Astin Bulletin*, **4**, 199–207.
- [21] Bühlmann, H. (1975). Minimax credibility. In Kahn, P.M. ed. *Credibility Theory and Applications*. Academic Press, New York, 1–22.
- [22] Bühlmann, H. (1995). *Mathematical Methods in Risk Theory*. Springer, Berlin.
- [23] Bühlmann, H., Straub, E. (1972). Credibility for loss ratios. (Translated by C.E. Brooks), *Actuarial Research Clearing House*, **2**.
- [24] Bühlmann, H., Shevchenko, P.V., Wüthrich, M.V. (2007). A "Toy" model for operational risk quantification using credibility theory. *Journal of Operational Risk*, **2**, 3–20.
- [25] Bury, K. (1999). *Statistical Distributions in Engineering*. Cambridge University Press, Cambridge.
- [26] Calderín, E, Gómez-Déniz, E., Cabrera, I. (2007). Bayesian local robustness under weighted squared-error loss function incorporating unimodality. *Statistics & Probability Letters*, **77**, 69–74.

- [27] Castillo, E., Hadi, A.S., Balakrishnan, N., Sarabia, J.M. (2005). *Extreme Value and Related Models with Applications in Engineering and Science*. John Wiley, New York.
- [28] Centeno, M.L., Silva, J.M. (2002). Optimal bonus scales under path-dependent bonus rules. *Scandinavian Actuarial Journal*, 2, 129–136.
- [29] Chavez-Demoulin, V., Embrechts, P., Neslehová, J. (2006). Quantitative models for operational risk: extremes, dependence and aggregation. *Journal of Banking & Finance*, 30, 2635–2658.
- [30] Chernobai, A.S., Rachev, S.T., Fabozzi, F.F. (2007). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*. John Wiley, New York.
- [31] Cruz, M. (2002). *Modelling, Measuring and Hedging Operational Risk*. John Wiley, Chichester.
- [32] Cruz, M. (Edit.) (2004). *Operational Risk Modelling and Analysis: Theory and Practice*. Risk Waters Group, London.
- [33] Denneberg, D. (1994). *Non-Additive Measure and Integral*. Kluwer, Dordrecht.
- [34] Denuit, M. (1997). A new distribution of Poisson-type for the number of claims. *Astin Bulletin*, 27, 229–242.
- [35] Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R., Laeven, R. (2006). Risk measurement with the equivalent utility principles. *Statistics and Decisions*, 24, 1-25.
- [36] Dey, D., G.; Ghosh, M., Strawderman, W. (1999). On estimation with balanced loss functions. *Statistics & Probability Letters*, 45, 97–101.
- [37] Dhaene, J., Goovaerts, M.J., Kaas, R. (2003). Economic capital allocation derived from risk measures. *North American Actuarial Journal*, 7, 44-59.
- [38] Eichenauer, J., Lehn, J., Rettig, S. (1988). A gamma-minimax result in credibility theory. *Insurance: Mathematics and Economics*, 7, 49–57.
- [39] Embrechts, P., Klüppelberg, C., Mikosch, T. (1997). *Modeling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin.
- [40] Farsipour, N.S., Asgharzadhe, A. (2004). Estimation of a normal mean relative to balanced loss functions. *Statistical Papers*, 45, 279–286.
- [41] Fernández-Laviada, A. (Edit.) (2007). *La Gestión del Riesgo Operacional. De la Teoría a su Aplicación*. Ediciones 2010, Limusa, Madrid.

- [42] Ferreira, J. (1977). Identifying equitable insurance premiums for risk classes: an alternative to the classical approach. *Twenty-third International Meeting of the Institute of Management Sciences. Athens.*
- [43] Freifelder, F. (1974). Statistical decision theory and procedures. In Kahn, P.M. ed. *Credibility theory and Applications*. Academic Press, New York, 71–78.
- [44] Fung, T., Seneta, E. (2007). Tailweight, quantiles and kurtosis: a study of competing models. *Operations Research Letters*, **35**, 448–454.
- [45] Furman, E., Landsman, Z. (2007). On some risk-adjusted tail-based premium calculation principles. *Journal of Actuarial Practice*, **13**, 175–191.
- [46] Furman, E., Zitikis, R. (2008). Weighted premium calculation principles. *Insurance: Mathematics and Economics*, **42**, 459–465.
- [47] Gajek, L., Miś, P., Slowińska (2007). Optimal streams of premiums in multi-period credibility models. *Applicationes Mathematicae*, **34**, 223–235.
- [48] Gerber, H.U. (1979). *An Introduction to Mathematical Risk Theory*. Huebner Foundation Monograph 8.
- [49] Gómez-Déniz, E. (2008a). A generalization of the credibility theory obtained by using the weighted balanced loss function. *Insurance: Mathematics and Economics*, **42**, 850–854.
- [50] Gómez-Déniz, E. (2008b). Deriving credibility premiums under different Bayesian methodology. In: *Advances in Mathematical and Statistical Modelling*. B.C. Arnold, N. Balakrishnan, J.M. Sarabia, R. Minguez Edit. Birkhäuser. Boston, 219–228.
- [51] Gómez-Déniz, E., Bermúdez, L., Morillo, I. (2005a). Computing bonus-malus premiums under partial prior information. *British Actuarial Journal*, **11**, 361–374.
- [52] Gómez-Déniz, E., Calderín, E., Cabrera, I. (2006c). A simple method to study sensitivity of BMP's. *Communications in Statistics: Theory and Methods*, **35**, 583–591.
- [53] Gómez-Déniz, E., Hernández, A., Vázquez-Polo, F. (1999a). The Variance Premium Principle: A Bayesian Robustness Analysis. *Actuarial Research Clearing House*, **1** 1–11.
- [54] Gómez-Déniz, E., Hernández, A., Vázquez-Polo, F. (1999b). The Esscher premium principle in risk theory: a Bayesian sensitivity study. *Insurance: Mathematics and Economics*, **25**, 387–395.

- [55] Gómez-Déniz, E., Hernández, A., Vázquez-Polo, F. (2000). Robust bayesian premium principles in actuarial science. *Journal of the Royal Statistical Society D*, **49**, 241–252.
- [56] Gómez-Déniz, E., Hernández, A., Vázquez-Polo, F. (2002a). Bounds for ratios of posterior expectations: applications in the collective risk model. *Scandinavian Actuarial Journal*, **1**, 37–44.
- [57] Gómez-Déniz, E., Pérez, J.M., Hernández, A., Vázquez-Polo, F. (2002b). Measuring sensitivity in a bonus-malus system. *Insurance: Mathematics and Economics*, **31**, 105–113.
- [58] Gómez-Déniz, E., Pérez, J.M., Vázquez, F. (2006a). On the use of posterior regret  $\Gamma$ -minimax actions to obtain credibility premiums. *Insurance: Mathematics and Economics*, **39**, 115–121.
- [59] Gómez-Déniz, E., León, M. (2005b). Un modelo de tarificación bonus-malus bajo el principio Esscher con tarifas más competitivas. *Estudios de Economía Aplicada*, **23**, 79–91.
- [60] Gómez-Déniz, E., Sarabia, J.M., Pérez, J.M., Vázquez, F. (2008c). Using a Bayesian hierarchical model for fitting automobile claim frequency data. *Communications in Statistics: Theory and Methods*, **37**, 1425–1435.
- [61] Gómez-Déniz, E., Sarabia, J.M., Vázquez, F. (2008d). Robust Bayesian bonus-malus premiums under the conditional specification model. *Statistical Papers*, en prensa.
- [62] Gómez-Déniz, E., Sarabia, J.M., Calderín-Ojeda, E. (2008e). Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications. *Insurance: Mathematics and Economics*, **42**, 39–49.
- [63] Gómez-Déniz, E., Vázquez-Polo, F. (2005c). Modelling uncertainty in insurance bonus-malus premium principles by using a Bayesian robustness approach. *Journal of Applied Statistics*, **32**, 771–784.
- [64] Gómez-Déniz, E., Vázquez, F., Pérez, J. (2006b). A note on computing bonus-malus insurance premiums using a hierarchical Bayesian framework. *Test*, **15**, 345–359.
- [65] Goovaerts, M.J., De Pril, N. (1980). Survival probabilities based on Pareto claim distributions. *Astin Bulletin*, **11**, 154–157.
- [66] Goovaerts, M.J., Kaas, R., Van Heerwaarden, A.E., Bauwelinckx, T. (1990). *Effective Actuarial Methods*. North-Holland, Amsterdam.

- [67] Guillén, M., Ayuso, M., Bolancé, C., Bermúdez, L., Morillo, I., Albarrán, I. (2005). *El Seguro de Automóviles: Estado Actual y Perspectiva de la Técnica Actuarial*. Editorial MAPFRE, S.A., Madrid.
- [68] Guillén, M., Gustafsson, J., Hansen, L., Nielsen, J.P. (2008). Multiline experience rating. *Astin Bulletin* (en prensa).
- [69] Guillén, M., Gustafsson, J., Nielsen, J.P., Pritchard, P. (2007). Using external data in operational risk. *The Geneva Papers on Risk and Insurance, Issues and Practice*, **32**, 178–189.
- [70] Haberman, S., Renshaw, A.E. (1996). Generalized linear models and actuarial science. *Journal of the Royal Statistical Society, D*, **45**, 407–436.
- [71] Hachermeister, C. (1975). Credibility for regression models with application to trend. In: *Credibility. Theory and Applications*. P.M. Kahn Ed. Academic Press, New York.
- [72] Heckman, P.E., Meyers, G.G. (1983). The calculation of the aggregate loss distribution for claim severity and claim count distributions. *Proceedings of the Casualty Actuarial Society*, **70**, 22–61.
- [73] Heilmann, W. (1989). Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics*, **8**, 77–95.
- [74] Heilpern, S. (2003). A rank-dependent generalization of zero utility principle. *Insurance: Mathematics and Economics*, **33**, 67–73.
- [75] Heras, A., Gil, J.A., García-Pineda, J.A., Vilar, J.L. (2004). An application of linear programming to bonus malus system design. *Astin Bulletin*, **34**, 435–456.
- [76] Heras, A., Vilar, J.L., Gil, J.A. (2002). Asymptotic fairness of bonus-malus systems and optimal scales of premiums. *The Geneva Papers on Risk and Insurance*, **27**, 61–82.
- [77] Herzog, T.N. (1996). *Introduction to Credibility Theory*. ACTEX, New York.
- [78] Hewitt, C. (1970). Credibility for severity. *Proceedings of the Casualty Actuarial Society*, **57**, 148–171.
- [79] Heyde, C.C., Kou, S.G. (2004). On the controversy over tailweight of distributions. *Operations Research Letters*, **32**, 399–408.
- [80] Hickmann, J. (1975). Introduction and historical overview of credibility. In: *Credibility. Theory and Applications*. P.M. Kahn Ed. Academic Press, New York.
- [81] Hogg, R.V., Klugman, S.A. (1984). *Loss Distributions*. John Wiley, New York.

- [82] Holgate, P., (1970). The modality of some compound Poisson distribution. *Biometrika*, 57, 666-667.
- [83] Hossack, I.B., Pollard, J.H., Zehnwirth, B. (1999). *Introductory Statistics with Applications in General Insurance (second edition)*. Cambridge University Press.
- [84] Hossack, I.B., Pollard, J.H., Zehnwirth, B. (2001). *Introducción a la Estadística con Aplicaciones a los Seguros Generales*. Editorial MAPFRE, Madrid.
- [85] Hürlimann, W. (1994). A note on experience rating, reinsurance and premium principles. *Insurance: Mathematics and Economics*, 14, 197-204.
- [86] Jafari, M.; Marchand, E., Parsian, A. (2006). On estimation with weighted balanced-type loss function. *Statistics & Probability Letters*, 76, 773-780.
- [87] Jewell, W.S. (1974). Credible means are exact Bayesian for exponential familie. *Astin Bulletin*, 8, 77-90.
- [88] Johnson, N.L., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Vol. 1, 2nd ed. John Wiley, New York.
- [89] Johnson, N.L., Kotz, S., Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Vol. 2, 2nd ed. John Wiley, New York.
- [90] Jorgensen, B. (1986). Some properties of exponential dispersion models. *Scandinavian Journal of Statistics*, 13, 187-198.
- [91] Kahn, P.M. (1975). *Credibility. Theory and Applications*. Academic Press, New York.
- [92] Kamps, U. (1998). On a class of premium principles including the Esscher premium. *Scandinavian Actuarial Journal*, 1, 75-80.
- [93] Kaas, R., Dannenburg, D., Goovaerts, M. (1997). Exact credibility for weighted observations. *Astin Bulletin*, 27, 287-295.
- [94] Kass, R., Goovaerts, M., Dhaene, J., Denuit, M. (2001). *Modern Actuarial Risk Theory*. Kluwer, Boston.
- [95] Kleiber, C., Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Science*. John Wiley, New York.
- [96] Klugman, S. (1992). *Bayesian Statistics in Actuarial Sciences with Emphasis on Credibility*. Kluwer, Boston.
- [97] Klugman, S.A., Panjer, H.H., Willmot, G.E. (2004). *Loss Models: From Data to Decisions*. Second Edition. John Wiley, New York.

- [98] Kremer, E. (1995). Finite formulae for the premium of the general reinsurance treaty based on ordered claims. *Insurance: Mathematics and Economics*, **4**, 233–238.
- [99] Kuon, S., Reich, A., Reimers, L. (1987). Panjer vs Kornya vs DePril: a comparison for a practical point of view. *Astin Bulletin*, **17**, 183–191.
- [100] Landsman, Z.M., Makov, U.E. (1998). Exponential dispersion models and credibility. *Scandinavian Actuarial Journal*, **1**, 89–96.
- [101] Landsman, Z., Sherris, M. (2001). Risk measures and insurance premiums principles. *Insurance: Mathematics and Economics*, **29**, 103–115.
- [102] Lehmann, E.L., Casella, G. (1998). *Theory of Point Estimation*. Second Edition. Springer Verlag, New York.
- [103] Lemaire, J. (1979). How to define a bonus-malus system with an exponential utility function. *Astin Bulletin*, **10**, 274–282.
- [104] Lemaire, J. (1985). *Automobile Insurance. Actuarial Models*. Kluwer-Nijhoff Publishing, Dordrecht.
- [105] Lemaire, J. (1995). *Bonus-Malus Systems in Automobile Insurance*. Kluwer Academic Publishers, Boston.
- [106] Linacre, N., Stewart-Oaten, A., Burgman, M., Ades, P. (2004). Incorporating collateral data in conservation biology. *Conservation Biology*, **18**, 768–774.
- [107] Mahler, H. (1991). An example of credibility and shifting risk parameters. *Proceedings of the Casualty Actuarial Society*, **76**, 225–308.
- [108] Makov, U.E. (1995). Loss robustness via Fisher-weighted squared error loss function. *Insurance: Mathematics and Economics*, **16**, 1–6.
- [109] Makov, U.E. (2001). Principal applications of Bayesian methods in Actuarial Science: a perspective. *North American Actuarial Journal*, **5**, 53–73.
- [110] Margolin, M.H. (1975). Are the classical and Bayesian approaches to credibility empirically valid?. In Kahn, P.M. ed. *Credibility theory and applications*. Academic Press, New York, 281–288.
- [111] Mayerson, A.L. (1964). A Bayesian view of credibility. *Proceedings of the Casualty Actuarial Society*, **51**, 85–104.
- [112] Meng, S., Wei, Y., Whitmore, G.A. (1999). Accounting for individual over-dispersion in a *bonus-malus* system. *Astin Bulletin*, **29**, 2, 327–337.

- [113] Mowbray, A. (1914). How extensive a payroll is necessary to give dependable pure premium? *Proceedings of the Casualty Actuarial Society*, **1**, 24–30.
- [114] Moosa, I.A. (2007). Operational risk: a survey. *Financial Markets, Institutions & Instruments*, **16**, 167–200.
- [115] Moscadelli, M. (2004). *The Modelling of Operational Risk: Experience with the Analysis of Data Collected by the Basel Committee*. Banca D'Italia, Termini di discussione, No. 517.
- [116] Nelder, J.A, Verrall, R.J. (1997). Credibility theory and generalized linear models. *Astin Bulletin*, **27**, 71–82.
- [117] Nelder, J.A, Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society A*, **135**, 370–384.
- [118] Panjer, H. (1981). Recursive evaluation of a family of compound distributions. *Astin Bulletin*, **12**, 22–26.
- [119] Panjer, H.H. (2006). *Operational Risks. Modeling Analytics*. John Wiley, New York.
- [120] Patil, G.P., Rao, C.R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, **34**, 179–189.
- [121] Pérez, J.M., Sarabia, J.M., Gómez-Déniz, E., Vázquez, F.J. (2006). Applying a Bayesian hierarchical model in Actuarial Science: inference and ratemaking. In: *Distribution Models Theory*. World Scientific.
- [122] Pielke, R.A., Landsea, C.W. (1998). Normalized Hurricane Damages in the United States: 1925-95. *Weather and Forecasting*, **13**, 621–631.
- [123] Press, S.J. (2003). *Subjective and Objective Bayesian Statistics: Principles, Models and Applications*. John Wiley, New York.
- [124] Ríos, D., Ruggeri, F., Vidakovic, B. (1995). Some results on posterior regret  $\Gamma$ -minimax estimation. *Statistics and Decisions*, **13**, 315–331.
- [125] Ríos, S., Martín, J., Ríos, D., Ruggeri, F. (1999). Bayesian forecasting for accident proneness evaluation. *Scandinavian Actuarial Journal*, **2**, 134–156.
- [126] Rolski, T., Schmidli, H., Schmidt, V., Teugel, J. (1999). *Stochastic Processes for Insurance and Finance*. John Wiley, New York.
- [127] Sarabia, J.M., Castillo, E., Gómez-Déniz, E., Vázquez-Polo, F. (2005). A class of conjugate priors for log-normal claims based on conditional specification. *Journal of Risk and Insurance*, **72**, 479–495.

- [128] Sarabia, J.M., Gómez-Déniz, E., Vázquez, F. (2004). On the use of conditional specification models in claim count distributions: an application to bonus-malus systems. *Astin Bulletin*, **34**, 85–98.
- [129] Sarabia, J.M., Gómez-Déniz, E., Vázquez, F. (2006). *Estadística Actuarial. Teoría y Aplicaciones*. Pearson-Prentice Hall, Madrid.
- [130] Sarabia, J.M., Guillén, M. (Edit.) (2007). *Investigaciones en Seguros y Gestión de Riesgos: RIESGO 2007*. Edita: TGD, Santander.
- [131] Scollnik, D. (2001). Actuarial modeling with MCMC and BUGS. *North American Actuarial Journal*, **5**, 96–125.
- [132] Shevchenko, P.V., Wüthrich, M.V. (2006). The structural modeling of operational risk via Bayesian inference: combining loss data with expert opinions. *Journal of Operational Risk*, **1**, 3–26.
- [133] Stuart, A., Ord, J.K. (1987). *Kendall's Advanced Theory of Statistics (vol. 1)*. Oxford University Press, New York.
- [134] Tsanakas, A., Desli, E. (2003). Risk measures and theories of choice. *British Actuarial Journal*, **9**, 959–991.
- [135] Tweedie, M.C.K. (1984). An index which distinguishes between some important exponential families. In: Ghosh, J.K., Roy, J. (Eds.), *Statistics: Applications and New Directions*. Proceedings of the Indian Statistical Golden Jubilee International Conference. Indian Statistical Institute, pp. 579–604.
- [136] Wang, S.S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, **17**, 43–54.
- [137] Wang, S.S. (1996). Premium calculation by transforming the layer premium density. *Astin Bulletin*, **26**, 71–92.
- [138] Wang, S.S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance*, **67**, 15–36.
- [139] Whitney, A. (1918). The theory of experience rating. *Proceedings of the Casualty Actuarial Society*, **4**, 274–292.
- [140] Willmot, G.E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial Journal*, 113–127.
- [141] Young, V. (2004). Premium principles. In: *Encyclopedia of Actuarial Science*. John Wiley, New York.

- [142] Zellner, A. (1994). Bayesian and non-Bayesian estimation using balanced loss function. In Gupta, S.S., Berger, J.O. (Eds.), *Statistical Decision Theory and Related Topics*, Springer New York, pp. 371–390.
- [143] Zen, M.M., DasGupta, A. (1993). Estimating a binomial parameter: is robust Bayes real Bayes?. *Statistics and Decisions*, **11**, 37–60.

# Índice alfabético

- aditividad, 19
- aproximación lineal, 112
- asegurado, 99
- aseguradora, 107
- aversión, 21
  
- Basilea II, 2, 171, 189, 190
- Bayes, 61–68, 71, 72, 75–79, 81, 83, 85–87, 89, 91, 97, 99, 100, 102–104, 123, 125
- bonus-malus, 1, 12, 88, 122, 135
  
- cartera
  - de seguros, 9, 15, 83, 84, 87, 107, 110, 111, 121
- coeficiente
  - de asimetría, 51, 199
  - de curtosis, 51
  - de variación, 107
- coeficiente de asimetría, 75
- cola de una distribución, 47, 195
- consistencia, 18
- constante
  - de Euler, 30
- credibilidad, 15, 105, 109, 110, 121–123, 125, 131, 139, 143, 151–155, 158, 159, 161, 162, 168
  - parcial, 109
  - total, 106–109
  
- dato de pérdidas, 35, 50, 75
- dato normalizado, 50
- diagram de barras, 69
  
- distribución
  - a posteriori, 64–67, 69–71, 73, 76, 80, 85–87, 101, 103
  - a priori, 62–64, 66–71, 73, 76, 80, 81, 83, 86, 90, 91, 98, 100–103, 110
  - beta, 69, 71, 91, 115, 125
  - beta de segunda especie, 85, 87, 133
  - binomial, 69, 125
  - binomial negativa, 84, 90–94
  - binomial negativa-inversa Gaussiana, 198
  - chi-cuadrado, 40, 74
  - conjugada, 73, 101
  - de Burr, 79
  - de Cauchy, 159
  - de Gumbel, 30
  - de máxima entropía, 129
  - de Pareto, 62, 72, 76, 78, 79
    - generalizada, 79
  - de Poisson, 82, 87, 90–92, 94, 95, 98–101, 125
  - de Poisson generalizada, 198
  - de Weibull, 79, 101
  - Erlang, 40
  - exponencial, 38, 40, 44, 49, 101, 125, 196
  - gamma, 38, 73, 74, 81, 85, 90–92, 94, 95, 98, 99, 125
  - gamma invertida, 40
  - inversa Gaussiana, 47, 103
  - lognormal, 33, 43

- máxima entropía, 129
- normal, 40, 43, 49, 61, 107, 125
  - tipificada, 107
- Pareto, 49
- Pearson tipo IV, 133
- poli-exponencial, 51
- predictiva, 63, 74, 75
- Rayleigh, 44
- sub-exponencial, 195
- Weibull, 44
  
- ecuación integral, 198
- error estándar, 51
- estadístico AIC de Akaike, 51
- estadístico chi-cuadrado, 135
- estimador
  - Bayes, 74, 76
  - consistente, 115, 134
  - de máxima verosimilitud, 75, 91
  - insesgado, 115, 119
  
- fórmula
  - de credibilidad, 61, 62, 71, 74, 78, 81, 91, 100, 104
- factor
  - de credibilidad, 6, 8, 10, 11, 115, 116, 118–121, 125, 128, 132, 145, 147, 154, 156, 158, 164, 168, 175, 180, 182, 185
- familia de distribuciones
  - exponencial, 99, 106, 123, 154, 168
  - exponencial de dispersión, 102, 128
- función
  - beta, 116, 133
  - confluente hipergeométrica, 133, 134
  - de densidad, 36–38, 40, 43, 47, 51
  - de distorsión de la cola, 22
  - de distribución, 40, 43, 44, 47
  - de distribución inversa, 29
  - de pérdida, 23, 24, 26, 27, 83, 86
  - de verosimilitud, 51, 52, 63, 68, 69, 72, 75, 81, 90, 93, 94, 174, 176, 181, 183, 184, 186
- estructura, 83
- exponencial integral, 30
- gamma incompleta, 129
- generatriz
  - de momentos, 21
- lagrangiana, 98
- modificada de Bessel, 197
- razón gamma incompleta, 40, 44
  
- hiper-priori, 67
- hiperparámetro, 62, 70, 73
- histograma, 50, 75
- homogeneidad positiva, 18
  
- inflación, 18
- intervalo
  - de credibilidad, 66, 71, 74–76
- iteratividad, 19
  
- método de estimación de
  - máxima verosimilitud, 51, 52
  - mínimos cuadrados, 110, 111
  - momentos, 134
- matriz, 135
- matriz de información, 51
- medida
  - coherente de riesgo, 28
  - de riesgo, 28
- medida de Arrow-Pratt, 21
- moda, 75
- modelo
  - de Bühlmann, 110–112, 115–118, 123, 125, 128, 152, 155, 156, 158
  - de Bühlmann-Straub, 110, 117
  - de distribución libre, 110
  - de Jewell, 123, 151, 152, 156, 158
  - gamma-minimax, 161, 162, 164, 165

- jerárquico, 130, 131, 133, 139–141
- regret gamma-minimax, 165, 166, 168
- regret gamma-minimax*, 166
- modelo de riesgo operacional
  - binomial-beta, 175
  - exponencial-gamma invertida, 184
  - lognormal-normal, 181
  - Pareto clásica-gamma, 186
  - Pareto generalizada-gamma, 182
  - Poisson-gamma, 173
- póliza, 15, 111, 121
- parámetro estructural, 114
- percentil, 43, 107
- prima, 12, 15, 16, 89, 94, 109
  - basada en estadísticos de orden, 33
  - Bayes, 62, 85–87, 89, 100, 103, 104, 122, 123
  - bonus-malus, 88–90, 92, 95, 98, 99, 122
  - colectiva, 6, 83, 85, 86, 89, 111, 115
  - de azar proporcional, 22
  - de desviación típica, 20
  - de Kamps, 32
  - de riesgo, 24, 81, 83, 85, 86, 111
  - de utilidad equivalente, 22
  - de utilidad exponencial, 20
  - de varianza, 20, 25, 83, 84, 86, 87, 89
  - de Wang, 22
  - del valor esperado, 20
  - Esscher, 21, 25, 27, 32, 83, 84, 86, 87, 89
  - exponencial, 25, 83, 86, 89
  - Holandesa, 22
  - manual, 109
  - neta, 20, 32, 83, 85, 86, 89, 90, 117, 122
  - ponderada, 30, 32
  - Suiza, 22
  - varianza en la cola, 32
- principio
  - de cálculo de prima, 16, 17
- probabilidad, 29, 31, 36, 37, 106, 111, 159
  - subjettiva, 158
- proceso
  - de frecuencia, 191
  - de Poisson homogéneo, 192
  - de severidad, 191
- proceso LDA, 189, 194, 196, 197
  - binomial negativa-inversa Gaussiana, 198
  - de Panjer, 197
  - Poisson generalizada, 198
  - Poisson-exponencial, 196, 199
  - Poisson-Pareto, 195
- rating, 17
- reaseguro, 16, 33
- recargo, 16, 18, 20, 25
- regresión
  - inversa-Gaussiana, 52
- riesgo, 15, 189–191
  - colectivo, 33
  - individual, 33
  - operacional, 171, 189–191, 193
- ruina, 18
- seguros, 16
- severidad, 17, 29, 36, 172, 180–182, 184, 186, 191
- siniestralidad, 5, 6, 9, 11, 12, 20, 61, 88, 106, 109, 121, 122, 139, 159
- sobrecarga, 94, 98
- sobredispersión, 133, 135, 192, 198
- tarificación, 89, 91, 94, 107
- teoría
  - de la credibilidad, 5–9, 11–13, 61, 62, 64, 67, 71, 74, 76, 78, 81, 91, 92, 99, 100, 102, 104, 105

- teorema
  - central del límite, 107
  - de Bayes, 63, 64, 86, 174
- valor en riesgo, 29
  - exponencial, 38
  - gamma, 40
  - inversa Gaussiana, 47
  - lognormal, 44
  - normal, 43
  - Pareto clásica, 37
  - Pareto II, 38
  - Weibull, 44
- valor en riesgo en la cola, 29
  - exponencial, 38
  - gamma, 40
  - lognormal, 44
  - normal, 43
  - Pareto clásica, 37
  - Pareto II, 38
  - Weibull, 44