ÍNDICE

ESTUDIOS DE ECONOMÍA APLICADA N° 19, 2001. PÁGS. 5-20

Fijación de primas de seguros bajo técnicas de robustez bayesiana

[*]GÓMEZ DÉNIZ, E. y [**]PÉREZ SÁNCHEZ, J. M.
Departamento de Métodos Cuantitativos en Economía y Gestión. Universidad de Las Palmas de G.C. Fac. CC. Económicas. Módulo D. Campus de Tafira.
Telf.: +34 28451803-Fax: +34 28451829 • [*] e-mail: emilio@empresariales.ulpgc.es - [**] e-mail: josema@empresariales.ulpgc.es

RESUMEN

La estadística actuarial ha abordado el problema de la tarificación en los seguros de no vida desde un punto de vista clásico y bayesiano clásico. En los primeros el parámetro de riesgo se considera conocido, mientras que en los segundos se considera aleatorio. En este trabajo se estudia la prima obtenida siguiendo ambas metodologías en el modelo colectivo de la teoría del riesgo. La utilización de la metodología bayesiana supone una confianza absoluta en la distribución a priori del parámetro de riesgo, y esto ha sido ampliamente criticado por los estadísticos no bayesianos. Para salvar esta situación, y utilizando la metodología de robustez bayesiana, mediremos la sensibilidad de la prima obtenida en un contexto bayesiano con respecto a perturbaciones en la distribución a priori del parámetro de riesgo, utilizando la clase de ε-contaminación.

Palabras Clave: Teoría de la Credibilidad, Principio de Cálculo de Prima, Robustez Bayesiana, Clase de ε-contaminación.

ABSTRACT

Actuarial statistics have approached tariffication problem in no-life insurances from a Bayesian and classical point of view. From the classical point of view, parameter is known, while Bayesian statistics considere it random. In this work, we studied risk premium under both methodologies in the collective model of the Risk Theory. Bayesian methodology supposes an absolute confidence in the prior distribution of the risk parameter, and it has been widely criticized by the classical statisticians. To save this situation, and using the Bayesian sensitivity methodology, we will measure the sensitivity of Bayesian Premium with respect to disturbances in the prior distribution from the risk parameter using ε-contaminated class.

Key words: Credibility Theory, Premium Calculation Principle, Bayesian Robustness, ε-Contamination Class.

Clasificación AMS: 62F15, 62P05.

Artículo recibido el 15 de julio de 2000. Aceptado el 14 de diciembre de 2000.

1. Introducción

Los actuarios son las personas que tratan con toda clase de problemas matemáticos y estadísticos en seguros. Las compañías de seguros aceptan riesgos de sus clientes, los asegurados, frente a un cierto precio denominado prima. La teoría de la credibilidad se basa en agrupar las pólizas referentes a un mismo riesgo con una serie de características comunes en un colectivo, al cual le corresponde como tal una determinada *prima colectiva*. Pero cada póliza, a su vez, tiene un conjunto de características específicas que la diferencia de las demás pólizas. Estas características, en la mayoría de los casos, son inobservables o difíciles de cuantificar, pero obviamente han de tenerse en cuenta para calcular las *primas de riesgo individuales*. La teoría de la credibilidad estima dichas primas basándose en la información pasada de la experiencia de siniestralidad y las fórmulas obtenidas son, en muchas ocasiones, una suma ponderada de la prima colectiva del riesgo y la media empírica de las indemnizaciones pagadas. El factor de ponderación utilizado se conoce con el nombre de *factor de credibilidad*.

Hasta hace poco, se intentaba determinar la prima para el colectivo sin preocuparse excesivamente por la heterogeneidad de la cartera. Sin embargo, la tendencia actual y futura parece considerar también las características particulares de cada riesgo. Los métodos bayesianos juegan aquí un papel muy importante, pues permiten incorporar la información resultante de la historia particular de cada riesgo.

En la práctica, el cálculo de la prima requiere que la distribución del riesgo X bajo consideración sea conocida (o al menos ciertos momentos, cuantiles, etc.). Nosotros consideraremos, como es usual en teoría de la credibilidad, el caso en el que la distribución de X está especificada salvo un parámetro desconocido.

La terminología que utilizaremos será la siguiente. Supondremos variables aleatorias q,X,X_1,X_2 tales que q toma valores en algún espacio paramétrico Θ (generalmente un subconjunto de la recta real), y dado $\Theta=q,X,X_1,X_2$ son independientes e idénticamente distribuidas con función de densidad de probabilidad $f(x|è)$. La *prima neta de riesgo* que la compañía aseguradora cobra al riesgo X viene dada por $\mathsf{P}(\theta)=\mathsf{E}[X|\theta]$, bajo el supuesto de que el parámetro q sea conocido. Uno de los principales tópicos en teoría de la credibilidad es precisamente la incertidumbre de dicho parámetro (Eichenauer, J. et al. (1988); Gómez, E. et al. (1999); Klugman, S. et al. (1998); entre otros). Asumiremos, pues, una distribución a priori $p_0(q)$ (*función estructura* en términos actuariales) para dicho parámetro. Ahora, suponiendo para el riesgo alguna experiencia de siniestralidad m, la prima Bayesiana vendrá dada por $\mathsf{P}^*_{\pi_0}(\mathsf{m})=\mathsf{E}\{\mathsf{E}[X|\theta]\}$.

Muchos trabajos se han publicado en teoría de la credibilidad acerca de la aproximación Bayesiana (Herzog, T. (1994); Gómez, E. et al. (1999); Klugman, S. et al. (1998); entre otros). Sin embargo muy pocos han tratado el asunto de la robustez bayesiana. Eichenauer,

J. et al. (1988); Gómez, E. et al. (1998); Gómez, E. et al. (1999), Heilmann, W. y Schröter, K. (1987) son algunos ejemplos de ellos.

El problema de interés en este artículo será el análisis bayesiano basado en información a priori parcial en teoría de la credibilidad. El análisis de robustez bayesiano ha recibido una atención considerable en las últimas dos décadas y numerosos autores han propuesto soluciones para este problema (Berger, J. (1994); Moreno, E. y Cano, J. (1991); Sivaganesan, S. y Berger, J. (1989); Sivaganesan, S. (1991); entre otros).

Bajo dicho análisis, el actuario será incapaz de especificar una única distribución a priori para el parámetro de riesgo. Aunque existen numerosos tipos de clases de distribuciones a priori que pueden utilizarse, dedicaremos nuestra atención a la clase de ε-contaminación $\Gamma_\varepsilon = \{\pi(\theta) = (1-\varepsilon)\pi_0(\theta) + \varepsilon\, q(\theta) : q \in \mathrm{Q}\}$, donde ε mide la incertidumbre que se tiene en la distribución a priori inicial y Q es una clase plausible de contaminaciones. Esta clase ha sido utilizada en situaciones diversas (Berger, J. (1994); Moreno, E. y Cano, J. (1991); Sivaganesan, S. y Berger, J. (1989); entre otros) para estudiar la sensibilidad de alguna cantidad a posteriori de interés, medida generalmente por el rango de variación de la misma.

El artículo está organizado como sigue. En la sección 2 se realiza una introducción a la teoría de la credibilidad y se procede al cálculo de la prima neta en sus distintas versiones. En la sección 3 exponemos el modelo gamma-gamma de la teoría del riesgo. En la sección 4 procedemos a flexibilizar las entradas que exige un análisis bayesiano intercambiando la distribución a priori inicial por toda una clase de distribuciones posibles. En la sección 5 se ilustran todas las ideas anteriores con ejemplos numéricos. En la sección 6 presentamos las propiedades exigibles a un principio de cálculo de prima y proponemos un nuevo sistema de tarificación basado en la robustez bayesiana. Finalmente la sección 7 concluye con algunos comentarios y posibles líneas abiertas para futuras investigaciones.

2. La teoría de la credibilidad y el principio de prima neta

El problema de la credibilidad se basa en estimar las ponderaciones que afectan a la experiencia de siniestralidad de una póliza respecto a la experiencia de un colectivo al que pertenece el suscriptor de dicha póliza. La cuestión básica es determinar hasta qué punto es creíble la experiencia observada de un asegurado individual en relación a la experiencia de un colectivo al que el asegurado pertenece.

En lo que sigue asumiremos que la siniestralidad de un riesgo o asegurado es una variable aleatoria con función de densidad de probabilidad $f(x|q)$, y que el valor de q es fijo para un riesgo dado, aunque desconocido. Si deseamos distinguir en qué año o período ocurre la siniestralidad X escribimos X_i para la siniestralidad en el año o período i=1,2,...,t.

Luego supondremos variables aleatorias $q, X, X_1, X_2, \ldots$ tales que X_i son independientes dado q e idénticamente distribuidas. Denotaremos, como es usual en teoría de la credibilidad, mediante $p_0(q)$ la función de densidad de q a la que se le llama función estructura.

En términos bayesianos, esta función de densidad representa una opinión a priori subjetiva acerca del parámetro desconocido q, que puede representar, por ejemplo, la propensión de un conductor a reclamar un siniestro y $p_0(q)$ puede describir de qué modo esa propensión se distribuye a través de la población de conductores asegurados. Luego $p_0(q)$ representa nuestra opinión a priori acerca de un conductor seleccionado aleatoriamente de la cartera.

$f(x)$ describe la distribución de la variable experiencia de siniestralidad para un contrato elegido aleatoriamente de la cartera, y es

$$f(x) = \int_Q f(x|q) p_0(q) dq, \tag{1}$$

que se trata de la densidad de X incondicional de q.

En teoría de la credibilidad se usan los términos individual y colectivo como sinónimos de contrato y cartera, y se distingue entre prima de riesgo, prima colectiva y prima bayesiana.

La prima neta de riesgo viene dada por

$$\mathrm{P}(q) = \int x\, f(x|q) dx, \tag{2}$$

y la prima neta de riesgo colectiva se obtiene como

$$\mathrm{P}^*_{\pi_0} = \int x\, f(x) dx = \int x \left(\int_\Theta f(x|\grave{e}) \pi_0(\grave{e}) d\grave{e} \right) dx = \int_\Theta \mathrm{P}(\grave{e}) \pi_0(\grave{e}) d\grave{e} \tag{3}$$

Si ahora, en un período de tiempo t se observan las indemnizaciones $x_1, x_2, \ldots, x_t$, y asumiendo independencia de un período a otro, la distribución a posteriori viene dada, utilizando el teorema de Bayes, por

$$p_0(q|\mathrm{m}) = \frac{f(\mathrm{m}|q) p_0(q)}{\int f(\mathrm{m} q) p_0(q) dq}, \tag{4}$$

donde $f(\mathrm{m}|\grave{e}) \equiv f(x_1, x_2, \ldots, x_t|\grave{e})$ es la verosimilitud observada. Esta función estructura a posteriori (distribución a posteriori) nos permite obtener la prima neta bayesiana, que se calcula, de la misma forma que la prima colectiva, intercambiando en (3) la distribución a priori $p_0(\theta)$ por la distribución a posteriori $p_0(\theta|\mu)$. Ahora la prima neta bayesiana resulta

$$\Pi^*_{p_0}(\mu) = \int_\Theta \Pi(\theta)\, p_0(\theta|\mu)\, d\theta .$$

La prima de riesgo representa la tasa teórica que la compañía de seguros cobraría a un individuo dado al asegurarse. Para su cálculo, la compañía (el actuario) debe conocer la forma de la distribución de probabilidad del riesgo y los parámetros de esta distribución. Si se dispone de esta información la prima de riesgo se podrá calcular y, por lo tanto, no existirán motivos para hacer ajustes de credibilidad. Sin embargo, en teoría de la credibilidad se supone que esta información no está disponible. En este caso la prima que la compañía cobra es la colectiva. Para su cálculo se requiere que el actuario especifique una distribución a priori para el parámetro de riesgo. La información que para ello se necesita se puede obtener de los datos de una población de contratos similares. La prima bayesiana, como ya dijimos antes, es muy similar a la prima colectiva. Considera para su cálculo la información a priori acerca de los parámetros del proceso de reclamaciones y la información muestral o experiencia de siniestralidad. Utilizando ambas informaciones se calcula la distribución a posteriori para, siguiendo el mismo camino que en el cálculo de la prima colectiva, obtener la prima bayesiana.

Evidentemente la metodología seguida para el cálculo de la prima es solamente una posibilidad de actuar entre la amplia gama de principios de cálculo entre los que elegir. Algunos principios de cálculo de primas más comunes son el de *prima neta*, *exponencial*, *Esscher* y *varianza*, entre otros. Una excelente revisión y discusión de este tópico se encuentra en Heilmann, W. (1989).

Los actuarios gustan de utilizar primas de la forma

$$\zeta_t \Pi[g(\mu)] + (1-\zeta_t)\Pi^*_{\pi_0}, \qquad \textbf{(5)}$$

con $\zeta_t \in [0,1]$ y $\lim_{t\to\infty} \zeta_t = 1$,
donde g es un estimador máximo verosímil de θ. Aquí $\Pi[g(\mu)]$ podría denominarse prima de riesgo individual y $\Pi^*_{p_0}$ prima colectiva. La expresión (5) recibe el nombre de fórmula de credibilidad y ζ_t es el factor de credibilidad. Bühlmann, H. (1967) fue el primero en dar una formulación explícita para ζ_t basada en la aproximación por mínimos cuadrados, obteniendo

$$\zeta_t = \frac{t}{t+k}, \text{ donde } k = \frac{E[V(X|\theta)]}{V[E(X|\theta)]}. \qquad \textbf{(6)}$$

V *(.)* representa la varianza y Z_t y 1-Z_t se interpretan como la credibilidad parcial de los datos observados y de la información a priori.

3. El modelo. Cálculo de primas

Un modelo usado con frecuencia en los sistemas de tarificación en seguros asume que el riesgo tiene la distribución gamma, $G\left(q\,,J\right)$, para la cantidad de indemnización (Bühlmann, 1970), con una distribución también gamma, a lo largo de la población para el parámetro desconocido J . Bajo dicho modelo *f(x)* viene dada por,

$$f(x)=\int_Q f\left(x\middle|q\right)p_0(q)dq=\int_Q \frac{q^J}{G(J)}x^{J-1}e^{-qx}\frac{a^b}{G(b)}q^{b-1}e^{-aq}dq$$

$$=\frac{x^{J-1}a^b}{G(J)G(b)}\int_Q q^{J+b-1}e^{-(a+x)q}dq=\frac{x^{J-1}a^b}{G(J)G(b)}\frac{G(b+J)}{(a+x)^{b+J}},$$

que resulta ser una distribución generalizada de Pareto, GPar(a,b, J).

La prima neta de riesgo se obtiene utilizando (2), y es,

$$P(q)=\frac{J}{q}.$$

La prima de riesgo colectiva o a priori es, de (3),

$$P^*_{p_0}=\frac{aJ}{b-1},\quad b>1,$$

ya que se trata de la esperanza de una distribución generalizada de Pareto, GPar(a,b, J).

Ahora, si en un período de tiempo t se observan las indemnizaciones $x_1,x_2,\ldots,x_t$, la probabilidad de este suceso (la verosimilitud) es

$$f(m|è)=\left(\frac{è^{\vartheta}}{\Gamma(\vartheta)}\right)^t\left(x_1x_2\cdots x_t\right)^{\vartheta}e^{-tmθ},\qquad m=\frac{1}{t}\sum_{i=1}^t x_i\,,$$

mientras que la distribución a posteriori de q dada la muestra es, de (4),

$$\pi_0(\theta|\mathrm{m}) \propto \theta^{b+t\vartheta-1}e^{-(a+\mathrm{m}t)\theta},$$

que resulta ser una distribución, también, gamma con parámetros $a+t\mathrm{m}$ y $b+t\vartheta$. La prima bayesiana se calcula de la misma manera que la prima colectiva, sustituyendo $p_0(q)$ por $\pi_0(è|\mathrm{m})$, resultando

$$\mathrm{P}^*_{\pi_0}(\mathrm{m}) = \frac{(a+t\mathrm{m})\vartheta}{b+t\vartheta-1} \qquad \textbf{(7)}$$

En nuestro caso, donde $X \sim G(q,J)$ y $q \sim G(a,b)$, se tiene,

$$\mathrm{E}[X|\theta] = \vartheta/\theta,\quad \mathrm{V}[X|q] = J/q^2,\quad \mathrm{E}[\mathrm{V}(X|q)] = \mathrm{E}(J/q) = \frac{Ja^2}{(b-1)(b-2)},$$

$$\mathrm{V}[\mathrm{E}(X|q)] = \mathrm{V}(1/q) = \frac{J^2a^2}{(b-1)^2(b-2)}.$$

Luego, usando las expresiones de (6),

$$\mathrm{Z}_t = \frac{t}{t+(b-1)/\vartheta}.$$

Obsérvese que (7) puede reescribirse como

$$\mathrm{P}^*_{\pi_0}(\mathrm{m}) = \frac{(a+t\mathrm{m})\vartheta}{b+t\vartheta-1} = \frac{t\vartheta}{b+t\vartheta-1}\mathrm{m} + \frac{b-1}{b+t\vartheta-1}\frac{a\vartheta}{b-1}$$

$$= \mathrm{Z}_t\mathrm{m} + (1-\mathrm{Z}_t)\mathrm{P}^*_{\pi_0}.$$

Teniendo en cuenta que si $X \sim \tilde{A}(\theta,\vartheta)$ con J conocido, entonces el estimador de máxima verosimilitud de q, para una muestra de tamaño t es $q^* = tJ/\left(\sum x_i\right)$ y de aquí se deduce, teniendo en cuenta que $\mathrm{P}(q) = J/q$, que $\mathrm{P}(g(\mathrm{m})) = \mathrm{P}(è^*) = \vartheta/è^* = \mathrm{m}$.

La prima bayesiana, pues, en nuestro modelo adopta la forma de una fórmula de credibilidad con factor de credibilidad como en (6).

4. Consideraciones más flexibles para el cálculo de la prima bayesiana

A pesar de los avances en el campo de la estadística bayesiana, varios son los problemas que suelen achacarse a esta aproximación. La elección de la distribución a priori, que tiene un carácter subjetivo, al estar basada en la información previa acumulada por el investigador. Sin embargo, en la práctica puede ser muy difícil discernir entre ciertas distribuciones con características similares. Por ejemplo, podemos estar convencidos de que la distribución a priori es unimodal y simétrica, pero distinguir entre las distribuciones normal y Cauchy puede resultar muy difícil. En otras ocasiones, la especificación de la distribución a priori se hace difícil porque la decisión ha de ser tomada por un grupo de personas que pudieran tener opiniones a priori diferentes.

Para salvar esta dificultad y desde hace algún tiempo se trabaja en análisis bayesiano con una metodología que consiste en procesar información a priori más flexible que la que se exige en un análisis bayesiano clásico. Bajo estas ideas el problema a tratar consiste en realizar un análisis de sensibilidad bayesiano en el proceso de tarificación de primas de seguros. Por ejemplo, si las creencias a priori del actuario tienen una forma menos elaborada que una distribución a priori, nos planteamos si se podrían ampliar las entradas del análisis bayesiano permitiendo que la especificación a priori fuera una clase o familia de distribuciones en lugar de una sola. Para concretar esta clase podríamos incorporar características que pudieran ser muy claras para un actuario, como la unimodalidad, conocimiento de algunos cuantiles, etc. Sobre esta familia el actuario calcularía los extremos inferior y superior de la prima a cobrar, de modo que si la diferencia entre esos dos valores, denominada rango de variación, es grande se habla de carencia de robustez; por el contrario, si la diferencia es pequeña se dice entonces que el modelo es robusto. La carencia de robustez debe interpretarse de la siguiente forma: densidades muy parecidas no producen cantidades próximas, y de ahí que el actuario deberá tomar sus decisiones con mucha precaución. Por contra, un modelo robusto debe interpretarse de esta otra forma: las decisiones del actuario no se verán sustancialmente modificadas con un elemento u otro de la clase.

Una familia que permite realizar un análisis de sensibilidad como el anteriormente señalado es la clase de contaminación (Berger, J. (1994); Moreno, E. y Cano, J. (1991); Sivaganesan, S. y Berger, J. (1989); entre otros). En ella la distribución a priori del parámetro está dentro de una clase de distribuciones a priori de la forma

$$G_e = \{p = (1-e)p_0 + e\,q : q \in \mathrm{Q}\}, \qquad \textbf{(8)}$$

en la que $e \in [0,1]$ determina la cantidad de incertidumbre en la distribución a priori inicial y Q es una clase de posibles contaminaciones.

ÍNDICE

Un interés natural del análisis de robustez consiste en encontrar el rango de variación de una magnitud a posteriori de interés, en nuestro caso la prima neta bayesiana; luego, estamos interesados en calcular el inf $\left\{ P^*_\pi(m); \pi \in \Gamma_\varepsilon \right\}$ y el sup $\left\{ P^*_\pi(m); \pi \in \Gamma_\varepsilon \right\}$. La metodología a seguir para ello se expone a continuación.

Para la clase de contaminación en (8) se tiene que

$$\pi(è|m) = \frac{f(m|è)\pi(è)}{\int_\Theta f(m|è)\pi(è)dè}$$

$$= \frac{(1-e)p_0(q|m)\int_Q f(m|q)p_0(q)dq + e\, q(q|m)\int_Q f(m|q)q(q)dè}{(1-e)\int_Q f(m|q)p_0(q)dq + e\int_Q f(m|q)q(q)dè},$$

la prima bayesiana, para el modelo de la sección 3, puede escribirse como,

$$P^*_p(m) = \int_Q \frac{J}{q} p(q|m)dq$$

$$= \frac{(1-e)P^*_{p_0}(m)\int_Q f(m|q)p_0(q)dq + eP^*_q(m)\int_Q f(m|q)q(q)d(q)}{(1-e)\int_Q f(m|q)p_0(q)dq + e\int_Q f(m|q)q(q)dq}. \qquad \textbf{(9)}$$

En este artículo desarrollaremos el análisis de sensibilidad de la prima bayesiana para las clases

$$G^i_e = \{p = (1-e)p_0 + e\,q : q \in Q_i\},\ \ i = 1,2,$$

tomando como clases contaminantes,

Q_1 = {Todas las distribuciones de probabilidad},

y

Q_2 ={Distribuciones unimodales con la misma moda, q_0, que $p_0(q)$ }.

Los siguientes resultados muestran cómo calcular el rango de variación de la prima bayesiana, que pasa por calcular los extremos de una función de una variable.

Lema 1. Si $X \sim G(q,J)$ y $q \sim G(a,b)$, entonces:

$$\inf_{p \in G^1_e} P^*_p(m) = \inf_{q \in Q} \frac{R_1 P^*_{p_0}(m) + R_2(q)}{R_1 + R_3(q)},$$

donde $\mathrm{R}_1 = (1-e)\int_Q f(\mathrm{m}|q)p_0(q)dq$, $\mathrm{R}_2(q) = e(J/q)f(\mathrm{m}|q)$, $\mathrm{R}_3(q) = q\,\mathrm{R}_2(q)$ y $\mathrm{P}^*_{p_0}(\mathrm{m})$ como en (7).

El supremo se obtiene reemplazando inf por sup.

DEMOSTRACIÓN.
Basta aplicar el lema A.1. de Sivaganesan, S. y Berger, J. (1989) a la expresión (9).

Lema 2. Si $X \sim G(q, J)$ y $q \sim G(a, b)$, entonces $\inf_{p \in G_e^2} \mathrm{P}^*_p(\mathrm{m}) = \inf_{z \geq 0} \mathrm{R}(z)$, donde

$$\mathrm{R}(z) = \frac{\mathrm{R}_1 \mathrm{P}^*_{p_0}(\mathrm{m}) + (1/z)\int_{q_0}^{q_0+z} \mathrm{R}_2(q)dq}{\mathrm{R}_1 + (1/z)\int_{q_0}^{q_0+z} \mathrm{R}_3(q)dq} \quad \text{si } z > 0,$$

y

$$\mathrm{R}(0) = \frac{\mathrm{R}_1 \mathrm{P}^*_{p_0}(\mathrm{m}) + \mathrm{R}_2(q_0)}{\mathrm{R}_1 + \mathrm{R}_3(q_0)},$$

con R_1, $\mathrm{R}_2(q)$, $\mathrm{R}_3(q)$ y $\mathrm{P}^*_{p_0}(\mathrm{m})$ como en el lema 1.

El supremo se obtiene reemplazando inf por sup.

DEMOSTRACIÓN.
Basta aplicar el lema 3.2.1 y el lema A.1. de Sivaganesan, S. y Berger, J. (1989) a la expresión (9).

Como medida de la robustez utilizaremos el factor de sensibilidad relativa, introducido por Sivaganesan (1991), y que, adaptado a nuestro problema, viene dado por

$$\mathrm{RS}^i = \frac{\sup\{\mathrm{P}^*_p(\mathrm{m}); p \in G_e^i\} - \inf\{\mathrm{P}^*_p(\mathrm{m}); p \in G_e^i\}}{2\mathrm{P}^*_{p_0}(\mathrm{m})} \times 100\%, \quad (i = 1, 2)$$

que puede interpretarse como el porcentaje de variación de la prima calculada para la distribución π alrededor de la prima calculada para la distribución a priori inicial π_0. En la sección 5 se aplican los resultados de esta sección a casos concretos.

5. Aplicación

Para ilustrar las ideas expuestas anteriormente desarrollaremos tres ejemplos numéricos. Supongamos que el actuario confía que la variable indemnización de reclamaciones tiene una distribución gamma con media en torno a 25 unidades monetarias (u.m.), y que la indemnización media más frecuente, la moda, está en torno a 12.5 u.m. Un modelo que refleje lo anterior puede ser el siguiente

$$X \sim G(q,2), \quad q \sim G(2,16),$$

donde, por comodidad computacional trabajaremos con los datos divididos por 10.

Tomaremos como observación muestral m la cantidad media de indemnización observada en los últimos t=10 años de vigencia de la póliza, que supondremos es de 10, 20 y 25 u.m., para tres ejemplos considerados. Después de los cálculos oportunos se obtienen los rangos de variación y los valores del factor RS que aparecen en las figuras 1, 2 y 3. Se ha tomado como grado de contaminación valores desde 0 hasta el 0.35 con pasos de 0.05, observándose que el factor RS no resulta particularmente alto en ninguno de los casos aquí considerados; luego, los resultados son razonablemente robustos. Por ejemplo, para el caso en que la media muestral es de 25 u.m., se obtiene como valor de la prima bayesiana 25.71 u.m., tomando el factor RS valores desde 1.03 hasta el 7.69% para contaminaciones con Q_1. Este factor se reduce hasta tomar los valores desde 0.47 hasta el 3.16% para contaminaciones con Q_2. Luego, desde nuestro punto de vista la compañía aseguradora puede sentirse tranquila cobrando, en este caso, como valor de la prima cualquiera de los que se encuentren en el intervalo de variación. Por ejemplo, si el actuario confía en una distribución a priori gamma en un 95% (ε=0.05), en el caso de contaminaciones unimodales la prima bayesiana toma valores en el intervalo [25.53, 25.78].

La prima colectiva, también en este ejemplo, toma el valor 26.66 u.m. de lo que resulta Z_t=0.5714. Esto quiere decir que la información muestral o datos observados pondera un 57%, mientras la información del colectivo o información a priori lo hace un 43%. En términos actuariales se comenta que la información muestral es creíble al 57% y la del colectivo al 43%.

Para un asegurado que se incorpore a la compañía de seguros, y del que no dispongamos de historial anterior, la compañía aseguradora le cobraría la prima colectiva con una credibilidad del 43%; en la medida que se vaya disponiendo de experiencia de reclamaciones para este sujeto, la prima que se le cobraría se le iría ajustando adecuadamente.

Para finalizar con este apartado digamos que todos los cálculos han sido elaborados con el software Mathematica y los gráficos con el software Matlab en un Pentium II a 550 Mh. y un tiempo de cómputo relativamente rápido.

Figura 1. Rangos de variación de la prima bayesiana con media 10

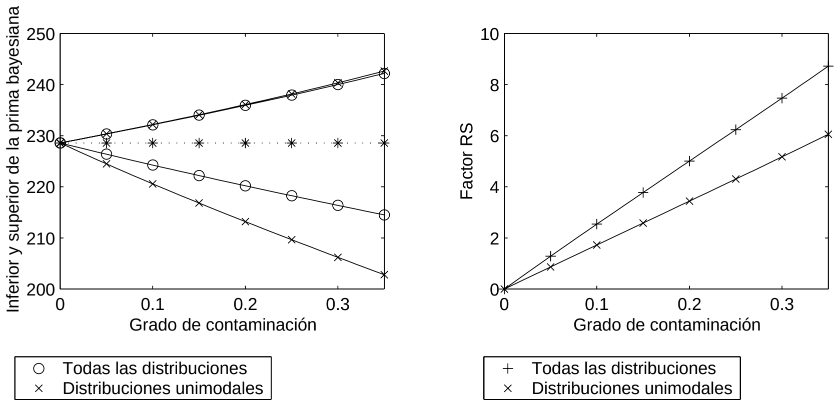


Figura 2. Rangos de variación de la prima bayesiana con media 20

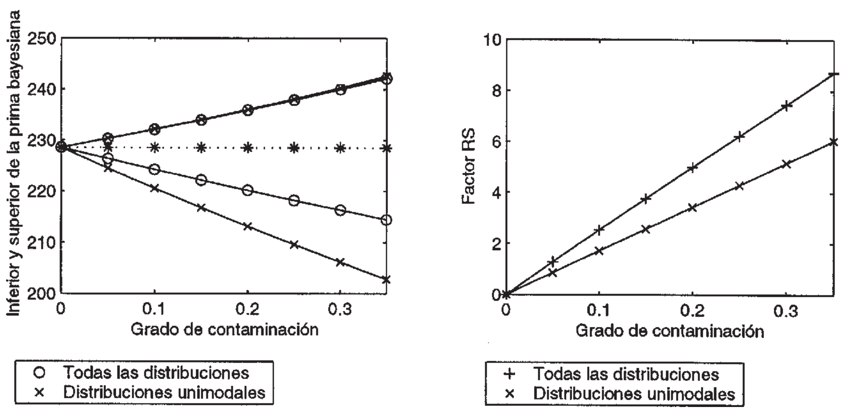
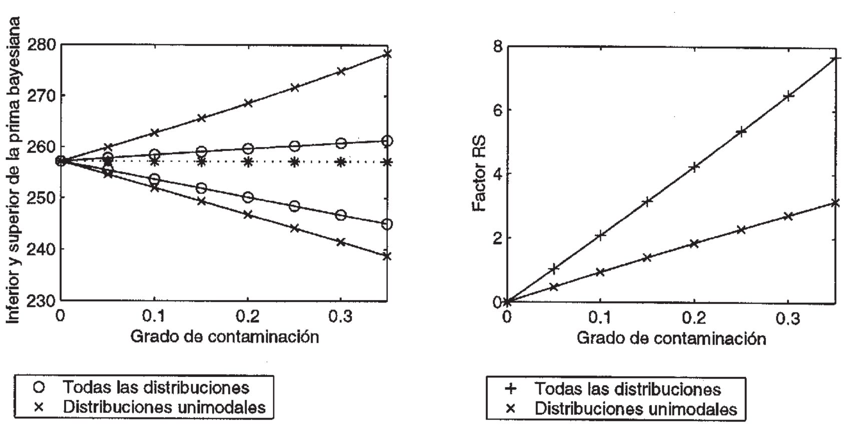


Figura 3. Rangos de variación de la prima bayesiana con media 25



6. Tarificación alternativa del riesgo

En la literatura actual no existe 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, H. (1979), Heilmann, W. (1989), Hürlimann, W. (1994), Kuen, S. y Yang, H. (1999) y Young, V. (1999), entre otros.

Generalmente se acepta que las propiedades que un principio de cálculo de prima $\mathsf{P} = \mathsf{H}[X]$ debería satisfacer son:

i. *Sobreprima de seguridad no negativa:* $\mathsf{P} \geq \mathsf{E}[X]$. Esto significa que para evitar la ruina técnica la ganancia esperada será no negativa.

ii. *No estafa:* La prima no excederá a la reclamación máxima posible r_X esto es, $\mathsf{P} \leq \mathsf{r}_X$.

iii. *Consistencia:* Para cada riesgo X y cada constante c, $\mathsf{H}[X+c] = \mathsf{H}[X] + c$. Esto significa que si el beneficio se incrementa en una constante ésta tiene que ser añadida a la prima.

iv. *Aditividad:* Si X_1 y X_2 son riesgos independientes, siempre se tendrá que cumplir $\mathsf{H}[X_1 + X_2] = \mathsf{H}[X_1] + \mathsf{H}[X_2]$. Esto quiere decir que la incorporación de riesgos independientes no afectan a la prima total.

v. $\mathsf{H}[c] = c$, para toda constante $c \geq 0$. Esto significa que para un riesgo no aleatorio X=c, con Prob[X=c]=1, la prima a cobrar será c.

vi. *Homogeneidad positiva:* $\mathsf{H}[cX] = c\,\mathsf{H}[X]$, para todo $c \geq 0$, que resulta conveniente para corregir efectos inflacionarios.

vii. $\mathsf{H}[pX + qY] \leq p\,\mathsf{H}[X] + q\,\mathsf{H}[Y]$, $\forall p > 0, q > 0$, tales que $p + q = 1$.

Es fácil probar que la prima neta de riesgo, colectiva y bayesiana verifican todas las propiedades anteriores. La cuestión es averiguar si cualquier valor de la prima bayesiana contenida en el intervalo de variación $\left[\inf\left\{\mathsf{P}_p^*(\mathsf{m}); p \in G_e^i\right\}, sup\left\{\mathsf{P}_p^*(\mathsf{m}); p \in G_e^i\right\}\right]$, sigue conservando estas propiedades. Para estudiar esto, obsérvese que de (9) se tiene que

$$\mathsf{P}_p^*(\mathsf{m}) = \frac{\mathsf{R}_1 \mathsf{P}_{p_0}^*(\mathsf{m}) + e\int \mathsf{P}(q) f(\mathsf{m}|q) q(q)\,dq}{\mathsf{R}_1 + e\int f(\mathsf{m}|q) q(q)\,dq},$$

que puede reescribirse de la forma,

$$\mathsf{P}_p^*(\mathsf{m}) = \frac{\mathsf{R}_1}{\mathsf{R}_1 + e\int f(\mathsf{m}|q) q(q)\,dq} \mathsf{P}_{p_0}^*(\mathsf{m}) + \frac{e\int \mathsf{P}(q) f(\mathsf{m}|q) q(q)\,dq}{\mathsf{R}_1 + e\int f(\mathsf{m}|q) q(q)\,dq}$$

$$=\frac{R_1}{R_1+\varepsilon\int f(\mu|\theta)q(\theta)d\theta}\Pi^*_{\pi_0}(\mu)+\frac{\varepsilon\int f(\mu|\theta)q(\theta)d\theta}{R_1+\varepsilon\int f(\mu|\theta)q(\theta)d\theta}\frac{\int\Pi(\theta)f(\mu|\theta)q(\theta)d\theta}{\int f(\mu|\theta)q(\theta)d\theta}$$

$$=\gamma\Pi^*_{\pi_0}(\mu)+(1-\gamma)\Pi^*_q(\mu),$$

con $\gamma=\frac{R_1}{R_1+\varepsilon\int f(\mu|\theta)q(\theta)d\theta}\in[0,1]$, ya que $\int f(\mu|\theta)q(\theta)d\theta>0$.

Luego, la prima bayesiana obtenida en el modelo de contaminaciones puede escribirse como una combinación convexa de dos primas bayesianas, la calculada para la distribución a priori inicial y la calculada para la distribución contaminante (esta es una prima teórica no calculable). A partir de aquí es fácil comprobar que las propiedades se conservan al pasar al modelo de contaminaciones con Θ_1. Ahora, puesto que $\Theta_2\subset\Theta_1$, se desprende que las propiedades son también válidas en el modelo de contaminaciones unimodales.

La propuesta que hacemos nosotros es la siguiente. Si el modelo es robusto, dado un ε y por tanto para una confianza en la distribución a priori de $(1-\varepsilon)100\%$ la compañía aseguradora puede cobrar como valor de la prima cualquiera del intervalo de variación de la prima bayesiana. Esta política de tarificación, novedosa en los escenarios actuariales, le puede reportar a la compañía un elenco de valores a cobrar que pudiera solucionarle problemas de competitividad, acaparando mayor cuota de mercado.

7. Conclusiones y posibles extensiones

En este artículo se ha ilustrado el uso del análisis bayesiano robusto en teoría de la credibilidad para medir la sensibilidad de la prima neta bayesiana en el modelo gamma-gamma. El análisis bayesiano robusto asume incertidumbre en la función estructura π_0, e intercambia ésta por clases de distribuciones a priori plausibles, lo que proporciona un intervalo de variación de la prima bayesiana.

Los resultados obtenidos indican que se obtiene una reducción significativa del factor RS cuando se considera la unimodalidad. Esta aproximación permite ofrecer resultados más flexibles que con los métodos bayesianos clásicos, lo que puede ser aprovechado por las compañías aseguradoras para resolverles problemas de competitividad, acaparando mayor cuota de mercado.

Una pequeña modificación en el lema 2 podría llevarse a cabo si el actuario considera conveniente añadir la propiedad de simetría a la de unimodalidad. Esto es posible reemplazando $(1/z)\int_{q_0}^{q_0+z}(\cdot)dq$ por $(1/(2z))\int_{q_0-z}^{q_0+z}(\cdot)dq$ (Sivaganesan, S. y Berger, J. (1989)).

Otra posible modificación sería considerar mixturas de distribuciones si el actuario desea considerar distribuciones bimodales (o multimodales) como función estructura, que es razonable en muchos ramos de seguros. Lo fundamental es que la clase de contaminación es muy flexible y se puede trabajar con ella sin complicar demasiado los procedimientos matemáticos.

Por otro lado, la mayoría de los modelos bayesianos en credibilidad utilizan una aproximación bayesiana pura utilizando dos niveles, el primero introduce la verosimilitud y el segundo una distribución a priori para algún parámetro de aquélla. Esta distribución contiene nuestro conocimiento sobre las relaciones entre las distintas clases de asegurados. No es nuestra opinión a priori sobre una clase en particular. Una solución a esto lo da el análisis bayesiano jerárquico, incluyendo un tercer nivel. En el primer nivel, como anteriormente, se introduce una distribución que describe las variaciones dentro de cada grupo de asegurados; el segundo nivel incorpora una distribución que describe las variaciones entre los grupos; finalmente, el tercer nivel proporciona una distribución a priori para los parámetros desconocidos en los niveles anteriores. Desafortunadamente, en este caso, las técnicas bayesianas son complicadas desde el punto de vista analítico, y sólo se ha desarrollado el modelo con distribuciones normales para los dos primeros niveles e impropia para el tercero. Sin embargo, es obvio que la distribución normal no resulta adecuada para modelar escenarios actuariales, por lo que resultaría interesante la investigación de otros modelos en el análisis bayesiano jerárquico.

Agradecimientos

Investigación parcialmente financiada por la Dirección General de Investigación Científica y Técnica (España) mediante proyecto PB95-1194. Los autores de la Universidad de Las Palmas de G.C. desean asimismo agradecer la financiación parcial de la Dirección General de Universidades e Investigación del Gobierno Autónomo de Canaria, proyecto 2705. Los autores agradecen los valiosos comentarios de un evaluador anónimo cuyas aportaciones han mejorado sensiblemente este trabajo.

ÍNDICE

Bibliografía

BERGER, J. (1994). “An Overview of Robust Bayesian Analysis.” TEST, 3, pp. 5 -120.

BÜHLMANN, H. (1967). “Experience rating and credibility”. Astin Bulletin, 5, II, pp. 157-165.

BÜHLMANN, H. (1970). Mathematical Methods in Risk Theory. Springer. New York.

EICHENAUER, J., LEHN, J. y RETTIG, S. (1988). “A gamma-minimax result in credibility theory.”Insurance: Mathematics and Economics, 7, pp. 49-57.

GERBER, H. (1979). An introduction to mathematical risk theory. Huebner Foundation.

GÓMEZ, E., HERNÁNDEZ, A y VÁZQUEZ, F. (1998). “Un análisis de sensibilidad del proceso de tarificación en los seguros generales”. Estudios de Economía Aplicada, 9, pp.19-34.

GÓMEZ, E.; HERNÁNDEZ, A. y VÁZQUEZ, F. (1999). “The Esscher Premium Principle in Risk Theory: a Bayesian Sensitivity Study”. Insurance: Mathematics and Economics, 25 (3), pp. 287-395.

HEILMANN, W. (1989). “Decision theoretic foundations of credibility theory”. Insurance: Mathematics and Economics, 8, pp. 77-95.

HEILMANN, W. y SCHÖTER, K. (1987). “On the robustness of premium principles”. Insurance: Mathematics and Economics, 6, pp. 145-149.

HERZOG, T. (1994). Introduction to Credibility Theory. ACTEX Publications, Winsted.

HURLIMANN, W. (1994). “A note on experience rating, reinsurance and premium principles”. Insurance: Mathematics and Economics, 14, pp. 197-204.

KLUGMANN, S., PANJER, H. y WILLMOT, G. (1988). Loss Models: From Data to Decisions. John Wiley & Sond, Inc. New York.

KUEN, S. y YANG, H. (1999). “Subjetive Risk Measures: Bayesian Predictive Scenarios Analysis”. Insurance: Mathematics and Economics, 25, pp. 157-169.

MORENO, E. y CANO, J. (1991). “Robust Bayesian Analysis for contaminations partially known”. Journal of the Royal Statistical. Series B, 53, pp. 143-155.

SIVAGANESAN, S. (1991). “Sensitivity of some posterior summaries when the prior is unimodal with specified quantiles”. The Canadian Journal of Statistics, Vol., 19, 1, pp. 57-65.

SIVAGANESAN, S. y BERGER, J. (1989). “Ranges of posterior measures for priors with unimodal contaminations”. The Annals of Statistics, Vol. 17, 2, pp. 868-889.

YOUNG, V. (1999). “Optimal insurance under Wang’s premium principle”. Insurance: Mathematics and Economics, 25, pp. 109-122.