

Propiedades del principio de prima neta con informacion a priori imprecisa.

Emilio Gómez Dénizl¹

Universidad de Las Palmas de G.C.

F'rancisco J. Vázquez Polo²

Universidad de Las Palmas de G.C.

Resumen

Son cinco las propiedades que un buen principio de cálculo de prima debe satisfacer: sobreprima de seguridad no negativa, no estafa, consistencia, aditividad e iteratividad, todas ellas satisfechas por el principio de prima neta de la Teoría del Riesgo Colectivo en un modelo clásico de estadística, y todas menos la quinta en un modelo bayesiano clásico. Nuestro trabajo plantea la pregunta de si la prima a posteriori en un modelo contaminado puede adoptar la forma de una fórmula *de credibilidad*, que verifica las cuatro primeras propiedades deseables para un principio de cálculo de prima, en el contexto de la Teoría del Riesgo Colectivo y para el modelo compuesto.

1 Introducción.

Es evidente la necesidad que tiene el actuario de asignar una prima justa y razonable (azar moral y competitividad) y la importancia que puede tener el disponer de un rango de valores para la misma con el objetivo de cobrar la que considere. en el momento oportuno; el análisis bayesiano robusto presenta una de las vías para alcanzar este objetivo.

Una de las maneras de estudiar la robustez consiste en plantear las diferentes creencias a

¹Investigación parcialmente financiada por la Pundación Universitaxia de Las Palmas.

²Investigación financiada por el proyecto número 2705 de la Dir.Gral.Univ. e Inv. Gob. Autónomo de Canarias.

priori (expresadas en densidades a priori) y estudiar sus diferencias respecto a ciertas medidas de interés. La metodología consiste en asignar una clase Γ de densidades a priori plausibles y realizar un análisis bayesiano estándar sobre cada uno de los elementos de la clase. El análisis será robusto cuando las conclusiones sean aproximadamente las mismas para cada elemento de la clase.

En definitiva, lo que nos proponemos es incorporar la imprecisión del actuario sobre sus creencias a priori mediante la consideración de clases de densidades a priori que sean coherentes con sus creencias.

El objetivo de este artículo, sin embargo, no es estudiar la robustez o carencia de la misma del principio de prima neta sino preguntarnos sobre el comportamiento de una prima a posteriori con información a priori imprecisa.

2 Modelo colectivo de la Teoría del Riesgo.

En el modelo colectivo de la Teoría del Riesgo, a partir de ahora T.R., un riesgo es una secuencia N, X_1, X_2, \dots de variables aleatorias con la siguiente interpretación:

- * N es la variable aleatoria *número de siniestros*.
- * $X_i, (i=1, 2, N)$ es la variable aleatoria *coste del i-ésimo siniestro*. Estas v.a. son entre sí independientes y equidistribuidas.
- * $X = \sum_{i=1}^N X_i$ es la variable aleatoria *coste total*.

Cuando se trabaja en T.R. el camino óptimo es hacerlo con la distribución del daño total obtenida componiendo los modelos de número de siniestros y cantidad monetaria. En este caso la distribución del daño total es:

$$F(x) = \sum_{N=0}^{\infty} P_N \cdot V^{(N*)}(x_i), \quad (1)$$

donde $V^{(N*)}(x_i)$ es la convolución N-ésima de la variable aleatoria X_i .

La forma que toma $F(x)$ depende de la distribución del número de siniestros P_N , siendo el caso más extendido aquél en el que P_N es Poisson, para el que $F(x)$ se denomina *distribución de Poisson compuesta*.

3 Principios de cálculo de primas.

La prima es el precio para el seguro (o reaseguro) vendido por la compañía aseguradora. El precio correcto, que es llamado *rating*, es vital, pues si es demasiado bajo representa una pérdida para la compañía y si es demasiado alto se pierde competitividad frente a otras. 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*.

Si denotamos por X la variable aleatorio representativa del riesgo (número de siniestros o cantidad monetaria reclamada), un procedimiento de cálculo de prima se define como sigue³

Definición 1 (de principio de cálculo de prima). *Un principio de cálculo de prima es una función H que asigna a un riesgo X un número real $H(X)$, la prima asignada al riesgo X .*

Ahora sea $L: \mathfrak{R}^2 \rightarrow \mathfrak{R}$ una función de pérdida, i.e. una función que atribuye a algún $(x, P) \in \mathfrak{R}^2$ la pérdida sostenida por un decisor que toma la acción P y se encuentra con el resultado x de algún experimento aleatorio. A partir de aquí se define la verdadera *prima individual* de la siguiente manera:

Definición 2 (de verdadera prima individual). *Dados un riesgo X con función de distribución $F(x)$ y una función de pérdida $L: \mathfrak{R}^2 \rightarrow \mathfrak{R}$ la verdadera prima individual es el valor de P que minimiza la pérdida esperada*

$$\int_{\mathfrak{R}} L(x, P) \cdot dF(x) = E[L(X, P)]$$

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

Según cual sea la función de pérdida que se considera surgirán distintos principios. Así, si $L(x, P) = (x - P)^2$, entonces:

$$P = E_x[X],$$

³Para una consulta más detallada de las definiciones 1 2 y 3 ver Heilmann (1989).

constituye el principio de prima neta o de equivalencia.

La verdadera prima individual representa el precio teórico que la compañía cobraría a un cliente al suscribir éste un contrato de seguros; para obtener este precio exacto la compañía debe conocer la forma de la distribución de probabilidad de siniestralidad y los parámetros de esta distribución. Si esta información le es asequible a la compañía, la verdadera prima individual se podrá calcular. En T.R., como es obvio, se sabe que esta información no está disponible; generalmente los valores actuales de los parámetros de la distribución de siniestralidad son desconocidos. En el modelo colectivo compuesto generalmente la distribución del número de siniestros, PN, depende de un parámetro θ , la tasa promedio de siniestros, y así a partir de $(l, F(x)) \equiv F(x|\mathbf{q})$.

En este trabajo consideraremos que el proceso de aprendizaje del actuario le va a servir para especificar una distribución a priori, $\pi_0(\theta)$, al parámetro de la distribución de siniestralidad, $F(x|\mathbf{q})$. Esta distribución a priori combinada con la información proporcionada por una muestra de contratos le servirá para construir la distribución a posteriori del parámetro y a partir de ésta la prima a *posterior*, que se define como sigue:

Definición 3 (de prima a posteriori). *Dados un riesgo X con distribución $F(x|\mathbf{q})$, siendo θ un parámetro desconocido con distribución a priori $\pi_0(\theta)$ una función de pérdida $L: \mathcal{R}^2 \rightarrow \mathcal{R}$, y una muestra \mathbf{x} la prima a posteriori es el valor P^* que minimiza*

$$\int_{\Theta} L(P, P^*) p_0(\mathbf{q} | \mathbf{x}) d\mathbf{q},$$

siendo $p_0(\mathbf{q} | \mathbf{x})$ la distribución a posteriori de \mathbf{q} dada la muestra \mathbf{x} , y P la verdadera prima

De nuevo, si $L(P, P^*) = (P - P^*)^2$, entonces:

$$P^* = \int_{\Theta} \left[\int_{\mathcal{X}} x j(x | \mathbf{q}) dx \right] p_0(\mathbf{q} | \mathbf{x}) = E_{\Theta} [E_{\Theta} [X | \mathbf{q}]]$$

Por otro lado es sabido que un buen principio de cálculo de prima, $P = H[X]$, debería satisfacer el mayor número de las propiedades deseables de un principio de cálculo de prima, y que son (ver Gerber (198)): **1. Sobreprima de seguridad no** negativa (esto es, $P \geq E[X]$.), **2. No estafa** ($P \leq r_x$, siendo r_x la reclamación máxima posible), **3. Consistencia** ($H[X + c] = H[X] + c$, para $c \in \mathcal{H}$), **4. Aditividad**. ($H[X_1 + X_2] = H[X_1] + H[X_2]$, con X_1 y X_2 riesgos independientes) y **5. Iteratividad** ($H[X] = H[H[X | S]]$, con X y S riesgos

arbitrarios).

Es fácil demostrar (ver Gerber (1980)) que la verdadera prima individual para el principio de prima neta satisface las cinco propiedades anteriores y que la prima a posteriori verifica todas menos la quinta.

4 Cálculo de la prima en un modelo de clases de contaminaciones.

Nos proponemos ahora, como adelantamos en la introducción, incorporar la imprecisión del actuario sobre sus creencias a priori mediante la consideración de clases de densidades a priori, llamémoslas Γ , que sean coherentes con sus creencias. Ahora bien, ¿cómo asignamos esta clase Γ de densidades a priori?. Una de las técnicas habituales en el sentido anterior es denominada *clases de contaminaciones* (Sivaganesan y Berger(1989), entre otros).

La metodología consiste en considerar que la distribución a priori de θ pertenece a la clase:

$$\Gamma_{\varepsilon} = \{\pi(\theta) = (1-\varepsilon) \cdot \pi_0(\theta) + \varepsilon \cdot q(\theta) \mid q \in Q; \varepsilon \in [0,1]\}.$$

Esto significa que se asume cierta confianza sobre π_0 , tanta como $(1 - \varepsilon) \cdot 100\%$ pero algo perturbada (mediante q , usualmente llamada contaminación, perteneciente a la denominada clase contaminante Q).

Para el estudio de la robustez sobre esta clase, una vía natural consiste en estudiar la variación de ciertas cantidades $g(x, \theta)$, a posteriori sobre el rango $\pi \in \Gamma_{\varepsilon}$ y que en nuestro caso será el $\inf_{g \in \Gamma_{\varepsilon}} P^*$ y el $\sup_{g \in \Gamma_{\varepsilon}} P^*$. Si el rango de $\left[\inf_{g \in \Gamma_{\varepsilon}} P^*, \sup_{g \in \Gamma_{\varepsilon}} P^* \right]$ es pequeño, entonces podemos asegurar robustez en la asignación de π_0 ; si el rango es grande, no tenemos robustez respecto a Γ_{ε} . Sobre la clase contaminante Q es usual considerar:

$$Q_1 = \{\text{Todas las distribuciones}\},$$

o bien,

$$Q_2 = \{\text{Distrib. Unimodales y con la misma moda que } \pi_0\}$$

Una vez planteada la necesidad de utilizar una metodología de robustez bayesiana en cualquier análisis que incorpore información a priori pasaremos a continuación a exponer las expresiones analíticas que nos permiten obtener los extremos inferior y superior para las cantidades de interés. No es nuestro objetivo en este trabajo realizar un análisis de robustez como el indicado. En este trabajo proponemos dos aspectos que son esencialmente dos preguntas. La primera de ellas se expone inmediatamente y la segunda a continuación.

Tenemos alguna constancia de que los actuarios tienden a interpretar el análisis antes esbozado como una nueva definición de prima a posteriori; en el siguiente sentido.

Definición 4.(de prima a posteriori en el modelo contaminado.) *Sea un riesgo X con distribución $F(x \mid q)$ especificada bajo un parámetro desconocido q sobre el que se tiene incertidumbre respecto a una única distribución a priori $p_0(q)$ y G una clase de posibles distribuciones a priori. Se define la prima a posteriori del riesgo X en el modelo G , como cualquier número real perteneciente al intervalo*

$$\left[\inf_{p \in \Gamma} P^*, \sup_{p \in \Gamma} P^* \right],$$

siendo P^ la prima a posteriori, dado algún principio de cálculo de prima y dada la verosimilitud $f(x \mid q)$ mientras que $p \in G$ indica una densidad a priori cualquiera dentro de la clase G .⁴*

Es evidente desde un punto de vista estadístico que la definición anterior tiene poco sentido, de hecho una mayor imprecisión lleva a un intervalo más amplio que facilita al actuario una elección más flexible. No obstante si nos restringimos a situaciones donde previamente se haya detectado robustez, ¿la definición anterior puede ser útil para el actuario?

El siguiente teorema muestra que el rango del valor a posteriori de $g(\theta)$ sobre la clase Γ_ε , se puede obtener encontrando los extremos de una función de una variable. La demostración del mismo es similar a un teorema que se encuentra en Sivaganesan (1988), y para un seguimiento detallada de la misma puede consultarse Gómez (1996).

⁴Es frecuente considerar r , y nosotros lo haremos así $\Gamma_\varepsilon = \{\pi = (1-\varepsilon) \cdot \pi_0 + \varepsilon \cdot q \mid q \in Q\}$ denominada clase de E-contaminaciones.

Teorema 1. Sea $\pi(x)$ el valor esperado a posteriori de $g(\theta)$ con respecto a la priori π ,

entonces para Γ_e , Q_1 y siempre que $\int_{\Theta} g(\mathbf{q}) f(x | \mathbf{q}) q(\mathbf{q}) d\mathbf{q} < \infty$:

$$\sup_{p \in \Gamma_e} \left(s \inf_{p \in \Gamma_e} \right) j^p(x) = \sup_q \left(\inf_q \right) R(\mathbf{q}),$$

donde $R(\mathbf{q}) = \frac{A_0 + g(\mathbf{q}) f(x | \mathbf{q})}{A + f(x | \mathbf{q})}$, $A = \frac{1-e}{e} m(x | \mathbf{p}_0)$, $A_0 = A \cdot j^{p_0}(x)$ y $m(x | \mathbf{p}_0)$ es la densidad predictiva de x dada \mathbf{p}_0 .

Obsérvese que el Teorema 1 nos asegura que la búsqueda de extremos puede hacerse en la expresión

$$R(\mathbf{q}) = \frac{A_0 + g(\mathbf{q}) f(x | \mathbf{q})}{A + f(x | \mathbf{q})}.$$

Esta expresión es un mero instrumento analítico para el cálculo, no obstante en nuestro caso esta expresión adopta la forma siguiente,

$$P_c^* = \frac{A \cdot P^* + g(\mathbf{q}) l(m | \mathbf{q})}{A + l(m | \mathbf{q})}$$

Si llamamos

$$Z = \frac{A}{A + l(m | \mathbf{q})}$$

entonces la prima a posteriori en este modelo de contaminaciones puede escribirse de la siguiente forma:

$$\begin{aligned} P_c^* &= Z \cdot P^* + (1 - Z) \cdot P \\ &= Z \cdot E_{\Theta} [E_X [X | \mathbf{q}]] + (1 - Z) \cdot E_X [X | \mathbf{q}] \end{aligned}$$

Llamaremos a Z *factor de credibilidad* y la fórmula anterior adopta una expresión similar a las fórmulas de credibilidad de la Matemática Actuarial. Como se puede observar esta expresión es una suma ponderada de la prima a posteriori y de la verdadera prima individual. Además $Z \in [0, 1]$ y $Z \rightarrow 1$ cuando $t \rightarrow \infty$, lo que implica que $P_c^* \rightarrow P^*$ cuando $t \rightarrow \infty$. Esto significa que a medida que el actuario posee más información sobre el riesgo mayor es la ponderación de la prima a posteriori; si tiene información completa el actuario puede tener

confianza total en la prima a posteriori y no tendría sentido contaminar el modelo. Por último, si $Z \rightarrow 0$, lo que ocurre cuando $\varepsilon \rightarrow 1$, la verdadera prima individual adquiere mayor ponderación (ya comentamos que en T.R. se supone que ésta no es asequible).

El desarrollo anterior y la forma que adopta la expresión que en un principio no era otra cosa que un instrumento analítico nos lleva a la segunda pregunta objeto de esta comunicación, ¿puede considerarse la expresión indicada como una formulación de prima a posteriori con información a priori imprecisa?

A mayor abundamiento en la pregunta anterior en el siguiente epígrafe se comprueba que la expresión indicada verifica un gran número de las propiedades deseables para un principio de cálculo de prima.

5 Propiedades.

La expresión

$$P_c^* = \frac{A \cdot P^* + g(\mathbf{q})l(m \mid \mathbf{q})}{A + l(m \mid \mathbf{q})}$$

verifica las propiedades 1, 2, 3 y 4 de las propiedades deseables para un principio de cálculo de prima y para cualquier distribución del riesgo $F(x \mid \theta)$.

Demostración.- Lo que haremos será demostrar la proposición para la clase Γ_ε , con Q_I , y con ello la proposición estaría demostrada para cualquier clase contaminante elegida, ya que estaría contenida en la anterior. Para ello observemos que:

$$H[X] = P_c^* = Z \cdot E_\Theta[E_X[X \mid \mathbf{q}]] + (1 - Z) \cdot E_X[X \mid \mathbf{q}]$$

A partir de aquí se tiene:

1.

$$\begin{aligned} E_\Theta[E_X[X \mid \mathbf{q}]] \geq E[X] &\Rightarrow Z \cdot E_\Theta[E_X[X \mid \mathbf{q}]] \geq Z \cdot E[X], \\ E_X[X \mid \mathbf{q}] \geq E[X] &\Rightarrow (1 - Z) \cdot E_X[X \mid \mathbf{q}] \geq (1 - Z) \cdot E[X], \end{aligned}$$

luego:

$$Z \cdot E_{\Theta}[E_X[X | \mathbf{q}]] + (1 - Z) \cdot E_X[X | \mathbf{q}] \geq E[X]$$

2. Puesto que $E_{\Theta}[E_X[X | \mathbf{q}]] \leq r_X$ y $E_X[X | \mathbf{q}] \leq r_X$, se tiene que:

$$Z \cdot E_{\Theta}[E_X[X | \mathbf{q}]] + (1 - Z) \cdot E_X[X | \mathbf{q}] \leq r_X.$$

3. Para probar que $H[X + c] = H[X] + c$, observemos que:

$$E_{\Theta}[E_X[X + c | \mathbf{q}]] = E_{\Theta}[E_X[X | \mathbf{q}]] + c,$$

$$E_X[X + c | \mathbf{q}] = E_X[X | \mathbf{q}] + c$$

Luego:

$$H[X + c] = Z \cdot [E_{\Theta}[E_X[X | \mathbf{q}]] + c] + (1 - Z) \cdot [E_X[X | \mathbf{q}] + c] = H[X] + c.$$

4. Si $X_1 | \theta$ y $X_2 | \theta$ son dos riesgos independientes, se trata de probar que $H[X_1 + X_2] = H[X_1 | \theta] + H[X_2 | \theta]$. En efecto:

$$\begin{aligned} H[X_1 + X_2] &= Z \cdot E_{\Theta}[E_{X_1+X_2}[X_1 + X_2 | \mathbf{q}]] \\ &\quad + (1 - Z) \cdot E[X_1 + X_2 | \mathbf{q}] \\ &= H[X_1] + H[X_2] \end{aligned}$$

Referencias Bibliográficas.

- [1] FREIFELDER, L.(1974).Statistical *Decision Theory and Credibility.Theory and Procedures. Credibility Theory and Applications.* Edited by P.M.Kahn. Academic Press,71-88.
- [2] GERBER, H.U. (1980).An introduction to mathematical risk theory. Huebner.
- [3] GÓMEZ, E. (1996). *Técnicas Estadísticas Bayesianas en Credibilidad.* Documentos de Trabajo. Universidad de Las Palmas de Gran Canaria.
- [41] GOODVAERTS, M.J. y HOOGSTAD, W.J. (1987).Credibility *Theory. Surveys of Actuarial Studies* No.4,19-52. Publication of Nationale-Nederlanden.

- [5] HEILMANN, W.R. (1989) *Decision theoretic foundations of credibility theory* Insurance: Mathematics and Economics. North Holland 8,77-95.
- [6] HICKMAN, J.C. (1974).Introduction and *Historical Overview of Credibility*. Credibility Theory and Applications. Edited by P.M.Kahn. Academic Press,181-193.
- [7] KLUGMAN, S.A. (1992).*Bayesian Statistics* in Actuarial Science. Kluwer Academic Publisher.
- [8] SIVAGANESAN, S. y BERGER, J. (1989). *Ranges of posterior measures for priors with unimodal contaminations*. The Annals of Statistics.Vol.17.No2,868-889.
- [9] VÁZQUEZ) F.J.(1992). *Técnicas estadísticas bayesianas en auditoría. Un análisis de robustez*. Tesis Doctoral. Universidad de Las Palmas de G.C.