LA DISTRIBUCIÓN POISSON-BETA: APLICACIONES Y PROPIEDADES EN LA TEORÍA DEL RIESGO COLECTIVO

Emilio Gómez Déniz[1], *José María Sarabia*[2] *y Faustino Prieto*[2]

Resumen

En el presente trabajo se estudia la distribución Poisson-Beta, tanto en los seguros individuales como en la teoría del riesgo colectivo. Se comienza revisando las propiedades básicas de la distribución. Estas propiedades incluyen los momentos ordinarios y factoriales, relaciones de recurrencia, así como las primas de riesgo, colectiva y Bayes. Se estudian diversas propiedades del modelo colectivo y se obtiene una relación de recurrencia para la distribución de la cantidad total reclamada, suponiendo que la cuantía se distribuye según una distribución discreta arbitraria. Se proponen métodos de estimación de momentos y de máxima verosimilitud para la distribución primaria Poisson-Beta. Finalmente, se incluyen varias aplicaciones con datos reales.

Palabras Clave: Distribución Poisson-Beta, Modelo de Riesgo Colectivo, Prima, Sobredispersión.

Abstract

In this paper the Poisson-Beta distribution is studied, both in individual and collective risk models. Basic properties are studied, including raw and factorial moments, recursive relations and risk, collective and Bayes premium. For the collective risk model, several properties are given and a recursive formula for calculating the total amount claim is obtained, assuming an arbitrary discrete distribution for the secondary distribution. Estimation methods based on moments and maximum likelihood are proposed, where the primary distribution is Poisson-Beta. Finally, several applications with real count data are included.

Key words: Poisson-Beta Distribution, collective risk model, premium, overdispersion.

---

[1] Departamento de Métodos Cuantitativos en Economía y Gestión. Universidad de Las Palmas de Gran Canaria, 35017-Las Palmas de Gran Canaria, España. E-mail: egomez@dmc.ulpgc.es

[2] Departamento de Economía. Universidad de Cantabria, Avda. de los Castros s/n, 39005-Santander, España. E-mail: sarabiaj@unican.es ; faustino.prieto@unican.es

1. Introducción

Los datos de conteo aparecen en diversos problemas prácticos en estadística actuarial, como por ejemplo, el número de reclamaciones recibidas por una compañía de seguros o el número de hospitalizaciones registradas en un servicio médico. Tradicionalmente, la distribución de Poisson ha sido la más utilizada para modelizar este tipo de datos, tanto por su simplicidad como por sus satisfactorias propiedades teóricas. Sin embargo, resulta bien conocido que dicha distribución subestima la varianza debido al fenómeno de sobredispersión. La sobredispersión aparece en los ejemplos antes mencionados, y dicha distribución de Poisson no captura esta situación. Esto sugiere que es necesario más de un parámetro para describir las propiedades empíricas de los datos, proponiéndose modelos compuestos obtenidos a partir de mezclas de distribuciones de Poisson. Entre los modelos compuestos de Poisson, destacamos el modelo Poisson-gamma, que da lugar a la distribución binomial negativa, y que ha sido ampliamente estudiado en el ámbito actuarial por Lemaire (1979, 1985 y 1995), entre otros. Otros modelos compuestos incluyen la distribución Poisson-lognormal (Aitchison y Ho (1989)), la distribución Poisson-inversa Gaussiana (Tremblay, 1992 y Willmot, 1987), la distribución Poisson-uniforme (Bhattacharya y Holla, 1965), la distribución Poisson-gamma-gamma (Gómez-Déniz et al. (2008b)) y la distribución Poisson-gamma general basada en especificación condicional (Sarabia et al. (2004)). Una revisión detallada de modelos compuestos, también llamados mixturas de Poisson, aparece en Karlis y Xekalaki (2005) y en Willmot (1986, 1993) donde se proponen diversas aplicaciones en estadística actuarial. Propiedades teóricas de algunos de estos modelos pueden encontrarse en el libro de Johnson et al. (2005).

Una de las ventajas de las distribuciones compuestas o mezclas de distribuciones es que suelen ser versiones sobredispersas (donde la varianza es mayor que la media) con colas más pesadas que la distribuciones de origen, y a menudo proporcionan mejores ajustes que las distribuciones de partida.

Basándonos en el proceso de mezcla o mixtura, presentamos en este trabajo la distribución Poisson-Beta, que ha sido propuesta por Gurland (1958) y Katti (1966) en problemas de Ecología y por Willmot (1986, 1993) en el ámbito actuarial. Este último autor únicamente estudia algunos aspectos teóricos. Demostraremos la utilidad de dicha distribución para modelizar datos de reclamaciones en seguros de automóviles.

Con objeto de hacer el trabajo autocontenido, presentamos a continuación algunos resultados que serán utilizados más adelante. La función de masa de probabilidad de una variable aleatoria X que sigue la distribución de Poisson con parámetro $\phi\theta,\ \phi > 0,\ 0 < \theta < 1$ viene dada por:

$$f(x \mid \phi, \theta) = \frac{e^{-\phi\theta}(\phi\theta)^x}{x!}, \qquad x = 0, 1, 2, \ldots,$$

con función generatriz de probabilidad dada por $G_X(z \mid \phi, \theta) = e^{\phi\theta(z-1)}$.

La función de densidad de probabilidad de una variable aleatoria X de tipo Beta clásica con parámetros $a, b > 0$ viene dada por:

$$f(x \mid a, b) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}, \qquad 0 < x < 1,$$

donde $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,dt$ es la función beta.

Finalmente, la función confluente hipergeométrica, denotada mediante ${}_1F_1(a;c;x)$, y que se conoce también como función de Kummer, se define por medio de la serie numérica:

$${}_1F_1(a;c;x) = \sum_{j=0}^{\infty} \frac{(a)_j\, x^j}{(c)_j\, j!}, \quad c \neq 0, -1, -2, \ldots$$

donde $(a)_j$ es el símbolo de Pochhammer, definido por $(a)_j = \Gamma(a+j)/\Gamma(a)$, siendo $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$ la función gamma. La función confluente hipergeométrica (ver Johnson et al., 2005) admite también la siguiente representación integral:

$${}_1F_1(a;c;x) = \frac{\Gamma(c)}{\Gamma(a)\,\Gamma(c-a)} \int_0^1 z^{a-1}(1-z)^{c-a-1}e^{xz}\,dz,$$

donde $c > a > 0$.

El resto del trabajo está organizado de la siguiente manera. La Sección 2 presenta las propiedades básicas de la distribución Poisson-Beta. Se incluyen las expresiones de los momentos ordinarios, los momentos factoriales, relaciones de recurrencia y las primas netas de riesgo, colectiva y Bayes, donde se supone un modelo Poisson para el riesgo y un modelo Beta para la distribución estructura (o distribución a priori). En la Sección 3 se estudian diversos métodos de estimación. Se proponen estimadores basados en la frecuencia de ceros junto con los dos primeros momentos, así como estimadores clásicos obtenidos a partir de los momentos factoriales. Se obtienen expresiones de los estimadores de máxima verosimilitud para la distribución primaria Poisson-Beta. En la Sección 4 se estudia el modelo de riesgo colectivo Poisson-Beta, donde la distribución primaria es de tipo Poisson-Beta, y la distribución secundaria de tipo discreto. Se estudian diversas propiedades del modelo colectivo y se obtiene una relación de recurrencia para la distribución de la cantidad total reclamada, suponiendo que la cuantía se distribuye según una distribución discreta arbitraria. En la Sección 5 se incluyen aplicaciones con datos reales del número de reclamaciones de seguro de automóviles, que ponen de manifiesto la importancia de esta distribución en seguros generales. El trabajo finaliza con una sección de conclusiones.

2 La distribución Poisson-Beta

Comenzamos presentando la definición de la distribución Poisson-Beta propuesta por Gurland (1958) y estudiada por Katti (1966) y Willmot (1986, 1993).

Definición 1 *Se dice que una variable aleatoria X sigue una distribución Poisson-Beta si admite la siguiente representación estocástica:*

$$X \mid \theta \approx Po(\phi\theta) \qquad (1)$$

$$\theta \approx Be(a,b) \qquad (2)$$

donde $Po(\phi\theta)$ representa una distribución de Poisson clásica con parámetro $\phi\theta,\ \phi > 0,\quad 0 < \theta < 1$ y $Be(a,b)$ representa una distribución Beta clásica con parámetros $a,b > 0.$

Una variable aleatoria X con la representación estocástica (1)-(2) se representará en adelante por $X \approx PB(a,b,\phi)$ y se dice que la variable aleatoria X sigue una distribución Poisson-Beta.

Incluimos a continuación propiedades básicas de esta distribución. Algunas de estas propiedades son conocidas y otras nuevas.

La función generatriz de probabilidad de la variable aleatoria Poisson-Beta viene dada por:

$$G_X(z) = {}_1F_1(a; a+b; \phi(z-1)), \tag{3}$$

donde ${}_1F_1(a;c;x)$ representa la función confluente hipergeométrica definida en la sección anterior.

La función de masa de probabilidad de una variable aleatoria X que sigue la distribución Poisson-Beta, i.e. $X \approx PB(a,b,\phi)$, admite la siguiente representación.

$$\begin{aligned}\Pr(X = x) &= \frac{\phi^x}{x!}\frac{\Gamma(a+b)\,\Gamma(a+x)}{\Gamma(a)\,\Gamma(a+b+x)}\,{}_1F_1(a+x; a+b+x; -\phi) \\ &= \frac{\phi^x}{x!}\frac{a\cdots(a+x-1)}{(a+b)\cdots(a+b+x-1)}\,{}_1F_1(a+x; a+b+x; -\phi),\end{aligned} \tag{4}$$

donde $x = 0,1,2,...$

Teniendo en cuenta que la distribución Beta es unimodal para valores $a > 1, b > 1,$ se deduce que si se verifican esas condiciones la distribución Poisson-Beta es de nuevo unimodal (ver Holgate (1970) y Al-Zaid (1983)).

El momento factorial de orden k puede obtenerse a partir de (3), y viene dado por:

$$\mu_{[k]}(X) = E\left[X(X-1)\ldots(X-k+1)\right] = \frac{\Gamma(a+b)\,\Gamma(a+x)\,\phi^x}{\Gamma(a)\,\Gamma(a+b+x)},\, k = 1,2,\ldots \quad (5)$$

A partir de (5) se deducen las expresiones de la esperanza y la varianza, que vienen dadas por:

$$E(X) = \frac{a\,\phi}{a+b}, \quad (6)$$

$$Var(X) = \frac{a\,\phi}{a+b} + \frac{a\,b\,\phi^2}{(a+b)^2(1+b+1)}, \quad (7)$$

y por tanto a partir de (6) y (7) se concluye que el modelo presenta sobredispersión.

Por otro lado, las probabilidades pueden obtenerse también de manera recursiva a partir de las probabilidades $\Pr(X=0)$ y $\Pr(X=1)$ utilizando la siguiente relación (Johnson et al., 2005),

$$(x+1)\,(x+2)\;p_{x+2} = (x+1)\,(x+a+b+\phi)\;p_{x+1} - \phi\,(x+a)\;p_x, \quad (8)$$

donde $p_x = \Pr(X=x)$. Por otro lado, teniendo en cuenta el primer teorema de Kummer:

$${}_1F_1(a\,,c\,,x) = e^x\;{}_1F_1(c-a\,,c\,,-x),$$

se verifica que la función de masa de probabilidad (4) puede también ser escrita como:

$$\Pr(X=x) = \frac{\phi^x}{x!}\,\frac{\Gamma(a+b)\,\Gamma(a+x)}{\Gamma(a)\,\Gamma(a+b+x)}\,e^{-\phi}\;{}_1F_1(b,a+b+x,\phi),$$

y por tanto,

$$\frac{\Pr(X=x+1)}{\Pr(X=x)}=\frac{\phi}{x+1}\,\frac{a+x}{a+b+x}\,\frac{{}_1F_1(b,a+b+x+1,\phi)}{{}_1F_1(b,a+b+x,\phi)}, \tag{9}$$

donde

$$\Pr(X=0)=e^{-\phi}\;{}_1F_1(b,a+b,\phi). \tag{10}$$

La expresión (9) permite por tanto obtener las probabilidades del modelo de manera recursiva.

En la representación estocástica (1)–(2), si nos preguntamos sobre la distribución condicional $\theta\big|X$, los momentos posteriores se pueden obtener mediante la siguiente fórmula (Willmot y Sundt, 1989 y Karlis y Xekalaki, 2005):

$$E(\theta^r\big|X=x)=\phi^r\,\frac{\Gamma(a+x+r)\,\Gamma(a+b+x)}{\Gamma(a+b+x+r)\,\Gamma(a+x)}\,\frac{{}_1F_1(a+x+r;a+b+x+r;-\phi)}{{}_1F_1(a+x,a+b+x,-\phi)}.$$

Cuando $r=1$, tenemos que:

$$E(\theta\big|X=x)=\phi\,\frac{a+x}{a+b+x}\,\frac{{}_1F_1(a+x+1;a+b+x+1;-\phi)}{{}_1F_1(a+x,a+b+x,-\phi)}. \tag{11}$$

Si el número de reclamaciones de un asegurado perteneciente a una cartera de seguros sigue la distribución de Poisson con parámetro $\phi\theta,$ la prima neta de riesgo es $P_R=\phi\theta.$ Si se considera que la cartera es heterogénea y que θ sigue la distribución a priori (función estructura) Beta clásica, entonces la prima neta colectiva vendrá dada por:

$$P_C=\int_0^1\phi\theta\frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)}d\theta=\frac{a\,\phi}{a+b}.$$

Ahora (11) puede ser vista como la prima neta Bayes cuando se dispone de un período de tiempo de observación con información muestral x. Para una revisión del principio de prima neta, el lector puede consultar a Gómez-Déniz et al. (2008a y 2008b), Sarabia et al. (2004) y Tremblay (1992), entre otros.

3 Estimación

En esta Sección se proponen diversos métodos de estimación de los tres parámetros del modelo. Consideremos una muestra aleatoria simple $x_1,\ldots,x_n$ procedente de la distribución Poisson-Beta con función de cuantía (4), y sean $m_{[k]}(X), k=1,2,3$ las versiones muestrales de los momentos factoriales.

3.1 Estimadores basado en la frecuencia de cero y los dos primeros momentos

Resulta bien conocido que en las distribuciones empíricas de frecuencias correspondientes a carteras de seguros de automóviles, el valor más frecuente observado es el cero. Se propone entonces un método de estimación que tenga en cuenta este importante hecho empírico. Consideramos un método de estimación basado en la frecuencia de ceros y en los dos primeros momentos. El sistema a resolver está formado por tres ecuaciones. La primera ecuación es la formada por la frecuencia teórica de ceros (10) y la correspondiente frecuencia observada p_0. Las otras dos ecuaciones las forman los dos primeros momentos teóricos junto con sus versiones muestrales. Se obtiene entonces el sistema:

$$\begin{aligned} p_0 &= e^{-\phi}\ {}_1F_1(b,a+b,\phi),\\ m_1 &= \frac{a\phi}{a+b},\\ \frac{m_2}{m_1} &= \frac{\phi(a+1)}{a+b+1}. \end{aligned}$$

Si en las dos últimas relaciones se despejan a y b en función de los momentos muestrales y de ϕ, se obtiene que:

$$a = \frac{m_1(m_2 - \phi m_1)}{\phi(m_1^2 - m_2)}, \tag{12}$$

$$b = \frac{(m_1 - \phi)(\phi m_1 - m_2)}{\phi(m_1^2 - m_2)}, \tag{13}$$

expresiones que sustituidas en la primera de las ecuaciones, dan lugar a una ecuación que sólo depende de ϕ, y puede resolverse numéricamente. Finalmente, sustituyendo este valor en (12) y (13) se obtienen los estimadores a y b.

Se propone a continuación estimadores de momentos obtenidos por medio de los momentos factoriales.

3.2 Estimadores de momentos

Los estimadores clásicos de momentos pueden obtenerse igualando los momentos factoriales teóricos a sus correspondientes muestrales, considerando el siguiente sistema de ecuaciones,

$$\mu_{[k]}(X) = m_{[k]},\ k = 1,2,3. \tag{14}$$

Resolviendo (14) en a, b y ϕ, obtenemos los estimadores de momentos en forma cerrada:

$$\hat{a} = \frac{-2\left(-m_1\, m_2^2 + m_1^2\, m_3\right)}{-m_1\, m_2^2 + 2\, m_1^2\, m_3 - m_2\, m_3},$$

$$\hat{b} = \frac{2\left(m_1^2 - m_2\right)\left(m_1\, m_2 - m_3\right)\left(m_1\, m_3 - m_2^2\right)}{\left(m_1^2\, m_2 - 2\, m_2^2 + m_1\, m_3\right)\left(2\, m_1^2\, m_3 - m_1\, m_2^2 - m_2\, m_3\right)},$$

$$\hat{\phi} = \frac{-\, m_1\, m_2^2 + 2\, m_1^2\, m_3 - m_2\, m_3}{m_1^2\, m_2 - 2\, m_2^2 + m_1\, m_3}.$$

Estos estimadores son consistentes y asintóticamente normales y pueden ser utilizados como valores iniciales en el método de estimación de máxima verosimilitud.

3.3 Estimación por máxima verosimilitud

El logaritmo de la función de verosimilitud viene dada por,

$$l(a,b,\phi) = n\log\Gamma(a+b) - n\log\Gamma(a) + n\,\overline{x}\,\log\phi + \sum_{i=1}^{n}\log\Gamma(a+x_i)$$
$$-\sum_{i=1}^{n}\log\Gamma(a+b+x_i) - \sum_{i=1}^{n}\log\,{}_1F_1(a+x_i;a+b+x_i;-\phi) - k,$$

de donde se obtienen las ecuaciones normales:

$$n\psi(a+b) - n\psi(a) + \sum_{i=1}^{n}\psi(a+x_i) - \sum_{i=1}^{n}\psi(a+b+x_i) -$$
$$-\sum_{i=1}^{n}\frac{{}_1F_1^{(0,1,0)}(a+x_i;a+b+x_i;-\phi) + {}_1F_1^{(1,0,0)}(a+x_i;a+b+x_i;-\phi)}{{}_1F_1(a+x_i;a+b+x_i;-\phi)} = 0,$$
$$n\psi(a+b) - \sum_{i=1}^{n}\psi(a+b+x_i) - \sum_{i=1}^{n}\frac{{}_1F_1^{(0,1,0)}(a+x_i;a+b+x_i;-\phi)}{{}_1F_1(a+x_i;a+b+x_i;-\phi)} = 0,$$
$$\frac{n\,\overline{x}}{\phi} + \sum_{i=1}^{n}\frac{{}_1F_1(1+a+x_i;1+a+b+x_i;-\phi)\,(a+x_i)}{{}_1F_1(a+x_i;a+b+x_i;-\phi)\,(a+b+x_i)} = 0,$$

donde $\psi(x) = \dfrac{d}{dx}\log\Gamma(x)$ y ${}_1F_1^{(i,j,k)}$ representa las derivadas parciales de la función confluente hipergeométrica. A pesar del aspecto aparentemente complejo de estas ecuaciones, el sistema puede resolverse sin dificultad por medio de los métodos numéricos disponibles en la mayoría de los paquetes comerciales de software. En este trabajo se ha utilizado el paquete de software Mathematica.

4 La distribución Poisson-Beta como distribución primaria

En esta Sección nos ocupamos del modelo de riesgo colectivo compuesto en el que la distribución primaria es la Poisson-Beta. Consideremos entonces la variable aleatoria $S = X_1 + X_2 + \ldots X_N$ donde $N, X_1, X_2, \ldots$ son variables aleatorias mutuamente independientes, no negativas y donde las variables aleatorias $X_1, X_2, \ldots$ están idénticamente distribuidas con función de cuantía común $f(x_i) = f_i$. La distribución de la variable aleatoria S se denomina Poisson-Beta compuesta. Resulta bien conocido que la función característica de S viene dada por:

$$\varphi_S(t) = G_N\left[\varphi_X(t)\right] = {}_1F_1(a; a+b; \phi(\varphi_X(t)-1)),$$

donde $\varphi_X(t)$ representa la función característica de la variable aleatoria X.

La distribución de S juega un papel importante en estadística actuarial cuando N representa el número de reclamaciones y X_i la cuantía asociada a la *i*-ésima reclamación, de modo que S representa la cantidad total reclamada. Resulta bien conocido (ver por ejemplo, Klugman et al. (2008)) que la función densidad de probabilidad de S viene dada por:

$$g_i = \sum_{n=0}^{\infty} p_n f_i^{*n}, \quad i > 0,$$

y $g_0 = G_N(f_0) = {}_1F_1(a, a+b; \phi(f_0 - 1)),$ donde f_i^{*n} denota la convolución *n*-ésima de f_i y $p_n = \Pr(N = n)$ viene definida en la expresión (4). Los avances computacionales permiten hoy en día calcular los valores de g_i con una exactitud razonable. Sin embargo, para la mayoría de las distribuciones discretas existen fórmulas recursivas que permiten su cálculo de manera exacta. A continuación demostraremos que esto resulta también posible para el caso del modelo Poisson-Beta compuesto.

A partir de la expresión (8), suponiendo que la variable aleatoria N sigue la distribución Poisson-Beta, se puede demostrar que:

$$p_n = \left(1 + \frac{a+b+\phi-2}{n}\right) p_{n-1} - \frac{\phi(n+a-2)}{n(n-1)} p_{n-2}, \quad n = 2,3,\ldots, \tag{15}$$

siendo $p_0 = e^{-\phi}\,{}_1F_1(b, a+b, \phi)$, mientras que

$$p_1 = \phi(a+1)\big/\left(a+b+1\right){}_1F_1(b, a+b+1, \phi)\big/{}_1F_1(b, a+b+1, \phi)\, p_0.$$

Si en la expresión (15) hacemos $a = 1$, se obtiene la siguiente relación de recurrencia:

$$p_n = \left(1 + \frac{b+\phi-1}{n}\right) p_{n-1} - \frac{\phi}{n} p_{n-2}, \quad n = 2,3,\ldots,$$

Una expresión similar, aparece en Schröter (1990) (ver expresión 2.3 en dicho trabajo). Finalmente, el siguiente resultado proporciona una fórmula que permite el cálculo exacto de la distribución de la cantidad total reclamada.

Teorema 1 Considérese la clase de distribuciones de masa de probabilidad que verifican la siguiente fórmula recursiva

$$p_n = \left(1 + \frac{m}{n}\right) p_{n-1} + \frac{c}{n} p_{n-2}, \quad n = 2,3,\ldots,$$

Supongamos que la variable aleatoria asociada a la *i*-ésima reclamación, con función de masa de probabilidad f_i, es de tipo discreto. Entonces, la función de masa de probabilidad de la cantidad total reclamada, $g_S(x_i) = g_i$ satisface la siguiente relación de recurrencia:

$$g_i = p_o + \frac{1}{1-f_0}\left\{p_1 f_i + \sum_{j=1}^{i}\left[\left(1+\frac{m\,j}{i}\right)f_j + \frac{c\,j}{2\,i}f_j^{*2}\right]g_{i-j} - p_0\sum_{j=1}^{i}\left(1+\frac{m\,j}{i}\right)f_j,\quad i\geq 1,\right.$$

siendo $f_j^{*2} = \sum_{j=0}^{i} f_j f_{i-j}$, mientras que $g_0 = p_0$.

Demostración: Partimos de la expresión:

$$g_i = \sum_{n=0}^{\infty} p_n f_i^{*n} = p_0 + p_1 f_i + \sum_{n=2}^{\infty} p_n f_i^{*n}.$$

A continuación, utilizando las siguientes expresiones (ver por ejemplo, Schröter, 1990)

$$f_j^{*n} = \sum_{j=0}^{i} f_j f_{i-j}^{*(n-1)},$$

$$f_j^{*n} = \frac{n}{i}\sum_{j=1}^{i} j f_j f_{i-j}^{*(n-1)},$$

$$f_j^{*n} = \frac{n}{k\,i}\sum_{j=1}^{i} j f_j^{*k} f_{i-j}^{*(n-k)},\quad i=1,2,...,\ n=1,2,...$$

se obtiene que:

$$
\begin{aligned}
\sum_{n=2}^{\infty} p_n f_i^{*n} &= \sum_{n=2}^{\infty}\left[\left(1+\frac{m}{n}\right)p_{n-1}+\frac{c}{n}p_{n-2}\right]f_i^{*n} \\
&= \sum_{n=2}^{\infty} p_{n-1} f_i^{*n} + m\sum_{n=2}^{\infty}\frac{1}{n}p_{n-1}f_i^{*n} + c\sum_{n=2}^{\infty}\frac{1}{n}p_{n-2}f_i^{*n} \\
&= \sum_{n=2}^{\infty} p_{n-1}\sum_{j=0}^{i} f_{i-j}^{*(n-1)} f_j + m\sum_{n=2}^{\infty}\frac{1}{n}p_{n-1}\frac{n}{i}\sum_{j=1}^{i} jf_j f_{i-j}^{*(n-1)} + c\sum_{n=2}^{\infty}\frac{1}{n}p_{n-2}\sum_{j=1}^{i} jf_j^{*2} f_{i-j}^{*(n-2)} \\
&= \sum_{j=0}^{i} f_j \sum_{n=2}^{\infty} p_{n-1} f_{i-j}^{*(n-1)} + \frac{m}{i}\sum_{j=1}^{i} jf_j \sum_{n=2}^{\infty} p_{n-1} f_{i-j}^{*(n-1)} + \frac{c}{2i}\sum_{j=1}^{i} jf_{i-j}^{*(n-2)} \sum_{n=2}^{\infty} p_{n-1} f_{i-j}^{*(n-1)} \\
&= \sum_{j=0}^{i} f_j\left(g_{i-j}-p_0\right) + \frac{m}{i}\sum_{j=1}^{i} jf_j\left(g_{i-j}-p_0\right) + \frac{c}{2i}\sum_{j=1}^{i} jf_j^{*2} g_{i-j}.
\end{aligned}
$$

Por tanto:

$$
\begin{aligned}
g_i &= p_0 + p_1 f_i + f_0\left(g_i - p_0\right) \\
&\quad + \sum_{j=1}^{i}\left[\left(1+\frac{m\,j}{i}\right)f_j + \frac{c\,j}{2i}f_j^{*2}\right]g_{i-j} - p_0\sum_{j=1}^{i}\left(1+\frac{m\,j}{i}\right)f_j,
\end{aligned}
$$

de donde se sigue el resultado.□

La fórmula recursiva para la cantidad total reclamada para el modelo Poisson-Beta compuesto se obtiene haciendo $m = b + \phi - 1$ y $c = -\phi$ en el teorema anterior. Finalmente, se puede obtener una fórmula alternativa para el caso en que f_i sea continua, cambiando los correspondientes sumatorios por integrales.

5 Aplicaciones

En este apartado probaremos a través de diversos conjuntos de datos la utilidad del modelo Poisson-Beta propuesto. Los ejemplos consisten en tres conjuntos de datos utilizados en la literatura actuarial. El primer conjunto de datos aparece en Klugman et al. (2008, p.465), Simon (1961) y Gómez-

Déniz et al. (2008b) entre otros. Estos datos representan el número de contratos de pólizas de seguros que han experimentado desde 0 hasta más de 12 reclamaciones. El segundo conjunto de datos ha sido utilizado por Willmot (1987) y Gómez-Déniz et al. (2008a), y corresponde a reclamaciones de una póliza de seguro de automóviles recogida en Zaire en el año 1974. Estos datos han sido ajustados previamente utilizando, entre otras, las distribuciones Poisson, binomial negativa, Poisson-inversa Gaussiana y la distribución Poisson-Gamma-Gamma. Finalmente, el tercer conjunto de datos se refiere al número de hospitalizaciones de un determinado grupo de empleados de una empresa, que aparece en Klugman et al. (2008, p.513).

Todos los conjuntos de datos presentan sobredispersión, puesto que las varianzas muestrales son mayores que las respectivas medias. Por tanto, la distribución Poisson-Beta propuesta en este trabajo parece apropiada para ajustar dichos datos. Para el primer conjunto de datos, la Tabla 1 muestra los datos originales junto con los valores ajustados obtenidos mediante los tres métodos de estimación descritos en la Sección 3 (método de la frecuencia de ceros y los dos primeros momentos factoriales (ZM), método de los momentos (MM) y método de máxima verosimilitud (ML)). Para el segundo conjunto de datos, la Tabla 3 incluye los valores observados y los ajustados mediante máxima verosimilitud. Hay que señalar que para este conjunto de datos los métodos basados en momentos no proporcionan estimadores admisibles. Las Tablas 2 y 4 muestran diversos estadísticos, incluyendo el valor del estadístico χ^2 junto con los correspondientes *p*-valores, así como el valor del logaritmo de la función de verosimilitud. Todos los ajustes son bastante satisfactorios, de modo que no se rechaza la hipótesis nula y se mejoran los ajustes proporcionados por otros modelos clásicos. Finalmente, las Tablas 5 y 6 recogen, respectivamente, los datos y valores ajustados, así como el resumen de los estadísticos de ajuste para el tercer conjunto de datos. De nuevo los ajustes son bastante satisfactorios.

6 Conclusiones

En este trabajo se ha propuesto la distribución Poisson-Beta para su aplicación en seguros individuales y en la teoría del riesgo colectivo. La mayor parte de las características de dicha distribución están disponibles en forma cerrada y son fáciles de calcular. Se han obtenido las primas de riesgo, colectiva y Bayes. Se han estudiado propiedades del modelo colectivo, y se ha obtenido una relación de recurrencia para la distribución de la cantidad total reclamada, suponiendo que la cuantía es de tipo discreto. Este resultado se puede extender al caso continuo. Se han propuesto métodos de estimación de momentos y de máxima verosimilitud. Los ajustes a partir de conjuntos de datos reales son bastante satisfactorios.

Tabla 1: Frecuencias observadas y estimadas con el modelo Poisson-Beta mediante los tres métodos de estimación propuestos.

Número de reclamaciones	Observadas	Estimadas (ZM)	Estimadas (MM)	Estimadas (ML)
0	99	99.00	95.25	96.98
1	65	70.54	74.41	72.75
2	57	49.89	51.81	50.92
3	35	34.16	34.13	34.08
4	20	22.43	21.59	21.91
5	10	14.03	13.20	13.54
6	4	8.34	7.82	8.05
7	0	4.70	4.49	4.60
8	3	2.51	2.50	2.52
9	4	1.27	1.36	1.33
10	0	0.60	0.71	0.67
11	1	0.27	0.36	0.33
≥ 12	0	0.12	0.31	0.57

Tabla 2: Resumen los ajustes para los datos de la Tabla 1.

Método	ZM	MM	ML
$\hat{a}$	0.942	1.218	1.086
$\hat{b}$	2.802	7.145	4.476
$\hat{\phi}$	7.208	12.459	9.291
χ^2	4.78	4.52	4.75
g.l.	3	3	3
p-valor	18.86 %	21.05 %	19.10 %
L_{max}	-560.863	-560.826	-560.764

Tabla 3: Frecuencias observadas y estimadas con el modelo Poisson-Beta mediante el método de máxima verosimilitud.

Número de reclamaciones	Observadas	Estimadas (ML)
0	3719	3719.22
1	232	229.88
2	38	39.92
3	7	8.41
4	3	1.93
5	1	0.46

Tabla 4: Resumen de los ajustes para los datos de la Tabla 3

Método	ML
$\hat{a}$	0.216
$\hat{b}$	848.403
$\hat{\phi}$	339.323
χ^2	1.43
g.l.	1
p-valor	23.17 %
L_{max}	-1183.55

Tabla 5: Hospitalizaciones observadas y ajustadas con el modelo Poisson-Beta mediante dos de los métodos de estimación propuestos.

Número de hospitalizaciones	Observadas	Estimadas (MM)	Estimadas (LM)
0	2659	2659.14	2659.07
1	244	243.45	243.62
2	19	19.80	19.68
3	2	1.50	1.50
+4	0	0.10	0.11

Tabla 6: Resumen de los ajustes para los datos de la Tabla 5

Método	MM	ML
$\hat{a}$	1.138	1.268
$\hat{b}$	14.076	60.519
$\hat{\phi}$	1.316	4.798
χ^2	0.30	0.63
g.l.	1	1
p-valor	58.38 %	42.73 %
L_{max}	-969.067	-969.065

Agradecimientos

Los autores agradecen al ministerio de Ciencia e Innovación (proyectos SEJ2004-02810 y SEJ2007-65818 (JMS) y SEJ2006-12685(EGD)) por la financiación parcial de este trabajo.

Referencias

Aitchison, J., Ho, C.H. (1989). The multivariate Poisson-Log normal distribution. *Biometrika*, **76**, 643-653.

Al-Zaid, A.A. (1983). On the unimodality of mixtures. *Pakistan Journal of Statistics*, **5**, 205-209.

Bhattacharya, S.K., Holla, M.S. (1965). On a Discrete Distribution with Special Reference to the Theory of Accident Proneness. *Journal of the American Statistical Association*, **60**, 312, 1060-1066.

Gómez-Déniz, E., Sarabia, J.M., Calderín-Ojeda, E. (2008a). Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications. *Insurance, Mathematics and Economics*, **42**, 1, 39-49.

Gómez-Déniz, E., Sarabia, J.M., Pérez, J.M., Vázquez, F. (2008b). Using a Bayesian hierarchical model for fitting automobile claim frequency data. *Communications in Statistics: Theory and Methods*, **37**, 1425-1435.

Gurland, J. (1958). A generalized class of contagious distributions. *Biometrics*, **14**, 229-249.

Holgate, P. (1970). The modality of some compound Poisson distributions. *Biometrika*, **57**, 666-667.

Johnson, N.L., Kemp, A.W., Kotz, S. (2005). *Univariate Discrete Distributions*. Third Edition. John Wiley, New York.

Karlis, D., Xekalaki, E. (2005). Mixed Poisson distributions. *International Statistical Review*, **73**, 35-58.

Katti, S.K. (1966). Interrelations among generalized distributions and their components. *Biometrics*, **22**, 44-52.

Klugman, S.; Panjer, H., Willmot, G. (2008). *Loss Models. From Data to Decisions*. John Wiley and Sons, New York. Third Edition.

Lemaire, J. (1979). How to define a bonus-malus system with an exponential utility function. *Astin Bulletin*, **10**, 274-282.

Lemaire, J. (1985). *Automobile Insurance. Actuarial Models*. Kluwer-Nijhoff Publishing, Dordrecht.

Lemaire, J. (1995). *Bonus-Malus Systems in Automobile Insurance*. Kluwer Academic Publishers, London.

Sarabia, J.M., Gómez-Déniz, E., Vázquez-Polo, F. (2004). On the use of Conditional Specification Models in Claim Count Distributions: An Application to Bonus-Malus Systems. *Astin Bulletin*, **34**, 85-98.

Simon, L. (1961). Fitting Negative Binomial Distribution by the Method of Maximum Likelihood. *Proceedings of the Casualty Actuarial Society*, **XLVIII**, 45-53 (with discussion).

Schröter, K. (1990). On a family of counting distributions and recursions for related compound distributions. Scandinavian Actuarial Journal, 161-175.

Tremblay, L. (1992). Using the Poisson inverse Gaussian in bonus-malus systems. *Astin Bulletin*, **22**, 97-106.

Willmot, G.E. (1986). Mixed compound Poisson distributions. *Astin Bulletin*, 16S, S59-S79.

Willmot, G.E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial Journal*, 113-127.

Willmot, G.E. (1993). On recursive evaluation of mixed Poisson probabilities and related quantities. *Scandinavian Actuarial Journal*, 2, 114-133.

Willmot, G.E., Sundt, B. (1989). On Posterior Probabilities and Moments in Mixed Poisson Processes. *Scandinavian Actuarial Journal*, **14**, 139-146.