

LA DISTRIBUCIÓN BINOMIAL-EXPONENCIAL TRUNCADA CON APLICACIONES EN EL SECTOR DEL SEGURO DE AUTOMÓVILES

Emilio Gómez-Déniz¹ y José María Sarabia²

Abstract

In this paper we present a new claim count distribution with overdispersion. The new distribution is obtained by mixing the p parameter in the binomial distribution with a $(0,1)$ truncated exponential distribution. The new distribution can be considered as an alternative to the classical Poisson, binomial and geometric distributions. Some properties of the new distribution and some applications in the field of automobile insurance are given, by fitting empirical claim data and obtaining bonus-malus premiums.

Key words: Mixture distributions, Bonus-malus Systems, Binomial distribution, Truncated Exponential Distribution, Confluent Hypergeometric Function, Insurance.

Resumen

En este trabajo se presenta una nueva distribución de conteo que puede presentar sobredispersión. La distribución se obtiene como mezcla del parámetro p de la distribución binomial con la distribución exponencial truncada en el intervalo $(0,1)$. La nueva distribución puede considerarse una alternativa a las distribuciones clásicas de Poisson, binomial y geométrica. Se muestran algunas propiedades de esta nueva distribución, así como aplicaciones en el ramo de seguro de automóviles mediante el ajuste de datos empíricos de reclamaciones y cálculo de primas bonus-malus.

Palabras Clave: Mezcla de distribuciones, Sistemas Bonus-Malus, Distribución Binomial, Distribución Exponencial Truncada, Función Confluyente Hipergeométrica, Seguro.

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

1 Introducción

Las distribuciones de conteo aparecen de manera natural en diversos problemas prácticos en el mundo de los seguros. En el ramo de seguro de automóviles aparecen como el número de reclamaciones en una cartera de seguros.

Tradicionalmente los modelos de esta naturaleza se comienzan modelando por medio de la distribución clásica de Poisson. Sin embargo, este modelo generalmente infraestima la varianza de los datos debido a la presencia de sobredispersión (varianza mayor que la media). De ahí que se propongan otros modelos que sean capaces de recoger este fenómeno. Tomando como punto de partida la distribución de Poisson, numerosos autores han propuesto alternativas a la misma mezclando dicha distribución con alguna otra. En este sentido destacamos los modelos Poisson-gamma (Lemaire, 1979, 1985, 1995) que genera la distribución binomial negativa; la distribución Poisson-lognormal (Aitchison y Ho, 1989); la distribución Poisson-inversa Gaussiana (Tremblay, 1992 y Willmot, 1987); la distribución Poisson-beta (Gurland, 1958; Katti, 1966 y Sarabia y Gómez-Déniz, 2007); 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); entre otros. Una revisión detallada sobre los distintos modelos basados en mezclas de distribuciones partiendo de la distribución de Poisson aparece en Karlis y Xekalaki (2005) y con aplicaciones en estadística actuarial en Willmot (1993).

Existen otros modelos construídos mediante mezclas de distribuciones que toman como punto de partida otras distribuciones discretas distintas de la distribución de Poisson. Algunos modelos importantes son: la distribución binomial-beta (Griffiths, 1973); la distribución binomial-gamma (Alanko y Duffy, 1996); la distribución binomial-inversa Gaussiana (Alanko y Duffy, 1996); la distribución binomial negativa-inversa Gaussiana (Gómez-Déniz et al., 2008a) y la distribución binomial negativa-Pareto (distribución binomial negativa-beta, Meng et al., 1999); entre otros.

Una de las ventajas de estas nuevas distribuciones de probabilidad obtenidas mediante mezclas es que las distribuciones obtenidas suelen ser sobredispersas y con colas más largas que la distribución de partida por lo que dan lugar a mejores ajustes para distribuciones empíricas de frecuencias. Diferentes propiedades sobre mezclas de distribuciones y su uso en diferentes disciplinas pueden ser consultadas también en Gómez-Déniz y Sarabia (2007),

Johnson et al (2005), Klugman et al. (1998), Sarabia et al. (2004) y Willmot (1987).

Basándose en el proceso de mezclas de distribuciones de probabilidad, en este trabajo se presenta una nueva distribución de probabilidad discreta cuya función de cuantía aparecerá expresada en términos de la función confluyente hipergeométrica y que depende de dos parámetros. Esta distribución puede ser utilizada como una alternativa a las distribuciones clásicas de Poisson, binomial y geométrica. La nueva distribución se obtiene mezclando la distribución binomial clásica con la distribución exponencial truncada en $(0,1)$ (Bain y Weeks, 1964 y Gómez-Déniz y Sarabia, 2007).

El trabajo está organizado del siguiente modo. La Sección 2 presenta las distribuciones binomial clásica y la distribución exponencial truncada, así como algunas de sus propiedades. En la Sección 3 se construye la distribución binomial-exponencial truncada. Se obtiene la función de cuantía así como las principales características incluyendo la media y la varianza y se prueba su carácter unimodal. En la Sección 4 se presentan los métodos de estimación de momentos y de máxima verosimilitud. La Sección 5 está dedicada a estudiar el modelo Bayesiano binomial-exponencial truncada en $(0,1)$. Se calcula la distribución a posteriori para el modelo binomial-exponencial truncada en $(0,1)$, las primas netas de riesgo, colectiva y Bayes, así como la prima neta bonus-malus bajo el modelo considerado. Diferentes aplicaciones a carteras de seguros que aparecen en numerosas ocasiones en la literatura actuarial son presentadas en la Sección 6. Finalmente, en la Sección 7 se muestran las conclusiones del trabajo.

2 Las distribuciones binomial y exponencial truncada

Con objeto de hacer este trabajo autocontenido se presentan a continuación las distribuciones binomial clásica y exponencial truncada, así como algunas características de las mismas que serán de utilidad más adelante.

2.1 La distribución binomial

La distribución binomial permite obtener la probabilidad del número de éxitos en una secuencia de n experimentos independientes, con una probabilidad constante p de ocurrencia de éxito en cada experimento. Si la variable aleatoria X sigue una distribución binomial, su función de masa de probabilidad está dada por

La distribución binomial-exponencial truncada con aplicaciones en el sector ...

$$f(x; p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad 0 < p < 1, x = 0, 1, \dots, n, \quad (1)$$

donde p es el parámetro a estimar y supondremos a lo largo de la investigación que n es conocido. Representaremos $X \approx B(n, p)$ para indicar que X sigue la distribución binomial en (1). La función generatriz de probabilidad viene dada por

$$G_X(s) = [1 + p(s-1)]^n,$$

y la media y la varianza son

$$\begin{aligned} E(X; p) &= np, \\ \text{Var}(X; p) &= np(1-p), \end{aligned}$$

respectivamente. Finalmente, los momentos factoriales (ver Balakrishnan y Nevzorov, 2003) vienen dados por:

$$\mu_{[k]}(X) = E[X(X-1)\cdots(X-k+1)] = n(n-1)\cdots(n-k+1)p^k, \quad (2)$$

con $k = 1, 2, \dots$

2.2 La distribución exponencial truncada

La función de densidad de probabilidad de la distribución exponencial truncada en el intervalo $(0, 1)$ viene dada por

$$f(x; \lambda) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda}}, \quad \lambda > 0, 0 < x < 1. \quad (3)$$

Representaremos $X \approx ET(\lambda)$ para indicar que X sigue la distribución exponencial truncada en (3). Puesto que $f'(x) < 0, \forall x > 0$, se deduce que la distribución (3) es unimodal con moda en $x=0$.

La figura 1 presenta algunos gráficos de la función de densidad de la distribución exponencial truncada (3) para diferentes valores del parámetro λ .

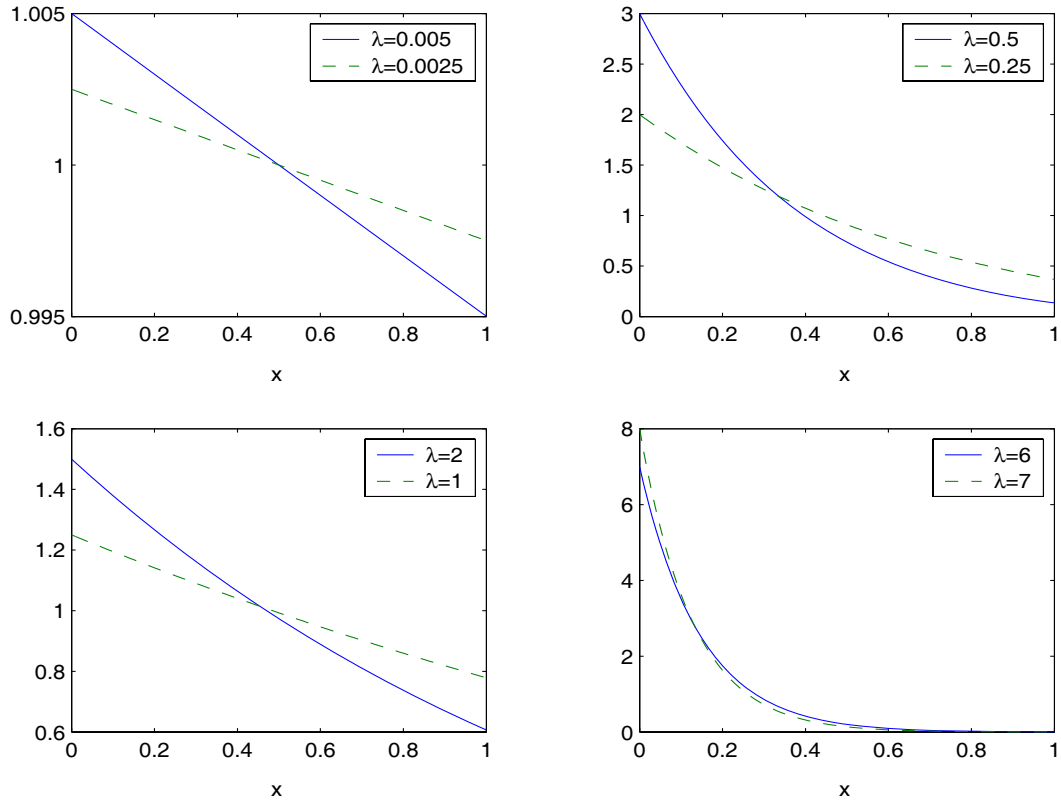


Figura 1. Función de densidad de probabilidad de la distribución exponencial truncada (3) para diferentes valores del parámetro λ

Esta distribución ha sido propuesta recientemente por Gómez-Déniz y Sarabia (2007) a partir de una mezcla una distribución de tipo potencial con la distribución Poisson truncada en cero.

La media y el momento de segundo orden son (ver Gómez-Déniz y Sarabia, 2007):

$$E(X; \lambda) = \frac{1 - e^{-\lambda} + \lambda}{\lambda(1 - e^{-\lambda})},$$

$$E(X^2; \lambda) = \frac{\lambda(2 + \lambda) + 2(1 - e^{-\lambda})}{\lambda^2(1 - e^{-\lambda})}.$$

Puesto que $E(X; \lambda)$ es una función decreciente en λ y $\lim_{\lambda \rightarrow 0} E(X; \lambda) = 0$ y $\lim_{\lambda \rightarrow \infty} E(X; \lambda) = 1/2$ se desprende que $0 < E(X; \lambda) < 1/2$.

3 La distribución discreta binomial-exponencial truncada

Se propone la siguiente definición de la distribución binomial-exponencial truncada en $(0,1)$.

Definición 1 *Se dice que una variable aleatoria discreta X sigue la distribución binomial-exponencial truncada si admite la siguiente representación estocástica:*

$$\begin{aligned} X | p &\approx B(n, p) \\ p &\approx ET(\lambda), \end{aligned} \tag{4)-(5}$$

donde $0 < p < 1$, $\lambda > 0$ y $n=1,2,\dots$. La nueva distribución será denotada por $X \approx BET(\lambda, n)$ y suponemos que n es un parámetro conocido.

El siguiente resultado proporciona una expresión cerrada para su función de masa de probabilidad expresada en términos de la función confluyente hipergeométrica.

Teorema 2 *Sea $X \approx BET(\lambda, n)$ la distribución de probabilidad discreta definida en (4)-(5). Su función de masa de probabilidad viene dada por:*

$$f(x; \lambda, n) = \frac{\lambda {}_1F_1(x+1, n+2, -\lambda)}{(n+1)(1-e^{-\lambda})}, \quad x = 0, 1, \dots, n, n > 0, \lambda > 0, \tag{6}$$

donde ${}_1F_1(a; c; x)$ representa la función confluyente hipergeométrica (también conocida como función de Kummer) definida como

$${}_1F_1(a; c; x) = \sum_{j=0}^{\infty} \frac{(a)_j x^j}{(c)_j j!} = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 z^{a-1} (1-z)^{c-a-1} e^{xz} dz, \quad c \neq 0, -1, -2, \dots \tag{7}$$

donde $c > a > 0$ y $(a)_j = a(a+1)\dots(a-j+1)$ representa el símbolo de Pochhammer.

Demostración: El resultado se obtiene calculando la distribución incondicional de X mediante la siguiente fórmula:

$$\begin{aligned} f(x; \lambda, n) &= \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} \frac{\lambda e^{-\lambda x}}{1-e^{-\lambda}} d\lambda \\ &= \frac{\lambda}{(n+1)(1-e^{-\lambda})} \int_0^1 \frac{\Gamma(n+2)}{\Gamma(n-x+1)\Gamma(x+1)} p^{(x+1)-1} (1-p)^{n+2-(x+1)-1} e^{-\lambda x} d\lambda, \end{aligned}$$

de donde utilizando (7) se obtiene el resultado.

Haciendo uso del primer Teorema de Kummer (Johnson et al., 2005, p.24) la función de masa de probabilidad en (6) puede reescribirse como:

$$f(x; \lambda, n) = \frac{\lambda}{n+1} \frac{{}_1F_1(n-x+1, n+2, -\lambda)}{e^\lambda - 1}, \quad x = 0, 1, \dots, n, \lambda > 0.$$

Otras propiedades de la función confluyente hipergeométrica pueden consultarse en Johnson et al. (2005).

Las probabilidades anteriores pueden obtenerse de manera recursiva a partir de la relación:

$$f(x+1; \lambda, n) = \frac{{}_1F_1(x+2, n+2, -\lambda)}{{}_1F_1(x+1, n+2, -\lambda)} f(x; \lambda, n), \quad x = 0, 1, \dots, n,$$

donde

$$f(0; \lambda, n) = \frac{\lambda {}_1F_1(1, n+2, -\lambda)}{(n+1)(1-e^{-\lambda})}. \quad (8)$$

A continuación presentamos algunas propiedades de la nueva distribución.

Teorema 3 Sea $X \approx BET(\lambda, n)$. Se verifican las siguientes propiedades:

- El momento factorial de orden k viene dado por

$$\begin{aligned}\mu_{[k]}(X; \lambda, n) &= E[X(X-1)\cdots(X-k+1)] \\ &= \frac{n(n-1)\cdots(n-k+1)}{1-e^{-\lambda}} \frac{k!-\Gamma(k+1, \lambda)}{\lambda^k},\end{aligned}$$

con $k = 1, 2, \dots$ y siendo $\Gamma(a, z) = \int_z^{\infty} t^{a-1} e^{-t} dt$ la función gamma incompleta.

• La media y la varianza vienen dadas por:

$$\begin{aligned}E(X; \lambda, n) &= \frac{n(1-e^{-\lambda} + \lambda)}{\lambda(1-e^{-\lambda})}, \\ \text{Var}(X; \lambda, n) &= n(n-1) \frac{\lambda(2+\lambda) + 2(1-e^{-\lambda})}{\lambda^2(1-e^{-\lambda})} + \frac{n(1-e^{-\lambda} + \lambda)}{\lambda(1-e^{-\lambda})} \left[1 - \frac{n(1-e^{-\lambda} + \lambda)}{\lambda(1-e^{-\lambda})} \right].\end{aligned}\tag{9}$$

Demostración: El momento factorial se obtiene de forma inmediata teniendo en cuenta el momento factorial de la distribución binomial dado en (2) y utilizando la esperanza condicionada.

La varianza se obtiene teniendo en cuenta que el momento factorial de orden dos es:

$$\mu_{[2]}(X) = E(X(X-1)) = n(n-1)E(p^2) = n(n-1) \frac{\lambda(2+\lambda) + 2(1-e^{-\lambda})}{\lambda^2(1-e^{-\lambda})}.$$

Puede probarse empíricamente que la nueva distribución es infradispersa o sobredispersa dependiendo de los valores de n y λ . El siguiente resultado prueba que la función de masa de probabilidad (6) es unimodal.

Teorema 4 La función de masa de probabilidad dada en (6) es unimodal.

Demostración: A partir de

$$\int_0^1 \binom{n}{x} p^x (1-p)^{n-x} dp = \binom{n}{x} B(x+1, n-x+1) = \frac{1}{n+1},$$

en la que $B(a,b)$ es la función beta, el resultado se desprende de forma inmediata teniendo en cuenta que la distribución exponencial truncada (3) es unimodal y haciendo uso del Teorema 3 en Al-Zaid (1983).

En la figura 2 aparece representada la función de probabilidad de la distribución binomial-exponencial truncada para diferentes valores de los parámetros n y λ . Como se aprecia en todos los casos, la moda está en cero.

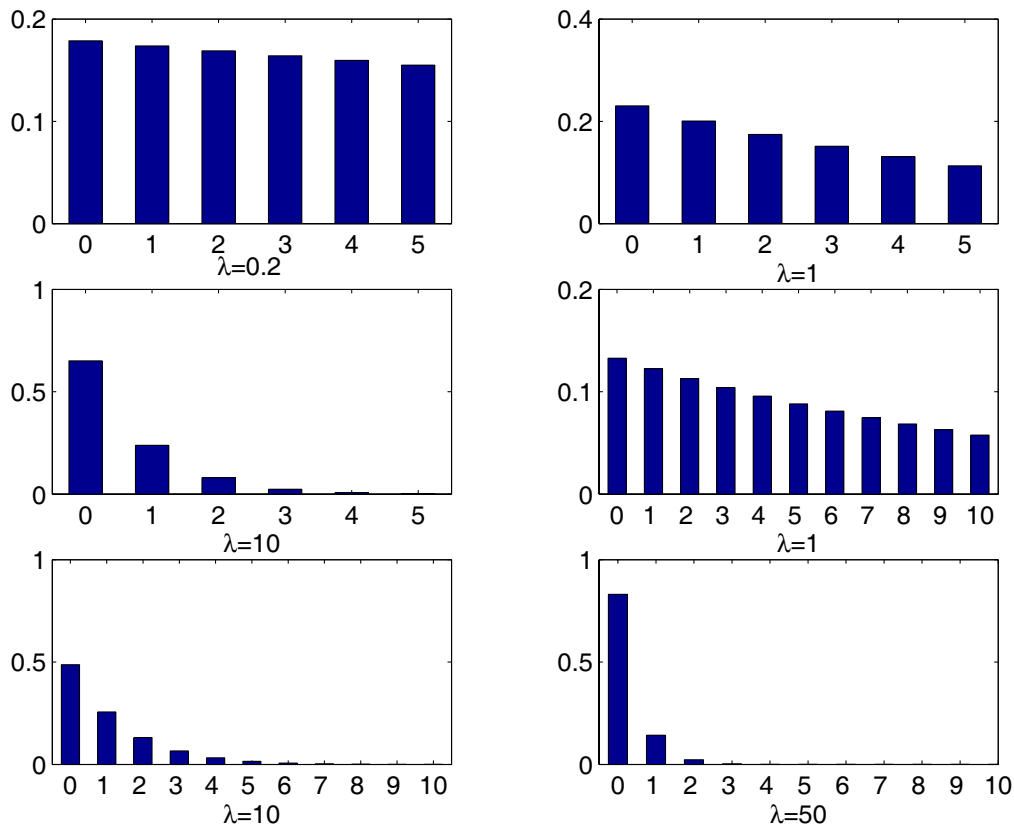


Figura 2. Función de probabilidad de la distribución binomial-exponencial truncada para diferentes valores de n y λ

4 Estimación

En esta Sección se proponen diversos métodos de estimación del parámetro λ de la distribución binomial-exponencial truncada. Se presta especial atención al método de los momentos y al de máxima verosimilitud. Se supone que el parámetro n es conocido.

Teniendo en cuenta que en las distribuciones empíricas de frecuencias correspondientes a carteras de seguros el valor más frecuente observado es el cero, un primer método que se propone es el basado en este hecho. Para ello igualamos (8) a p_0 , siendo p_0 la frecuencia relativa de ceros. Por tanto se necesitará resolver la ecuación

$$(n+1)(1-e^{-\lambda})p_0 - \lambda {}_1F_1(1, n+2, -\lambda) = 0.$$

Para n fijo, la solución de la ecuación anterior en λ da lugar al estimador basado en la frecuencia de ceros.

A continuación se propone un estimador por el método de los momentos.

4.1 Estimador del parámetro por el método de los momentos

Si \bar{x} es la media muestral de los datos observados, el parámetro λ puede estimarse fácilmente por el método de los momentos igualando la media muestral a la media poblacional dada en (9). Resulta la ecuación

$$\frac{n(1-e^{-\lambda} + \lambda)}{\lambda(1-e^{-\lambda})} = \bar{x}.$$

Puesto que el lado izquierdo de la ecuación anterior es una función decreciente en λ , siempre disponemos de solución única. La ecuación puede resolverse fácilmente mediante algún programa estándar que incorpore cálculo numérico. La figura 3 muestra el lado izquierdo de la ecuación anterior para algunos valores de n .

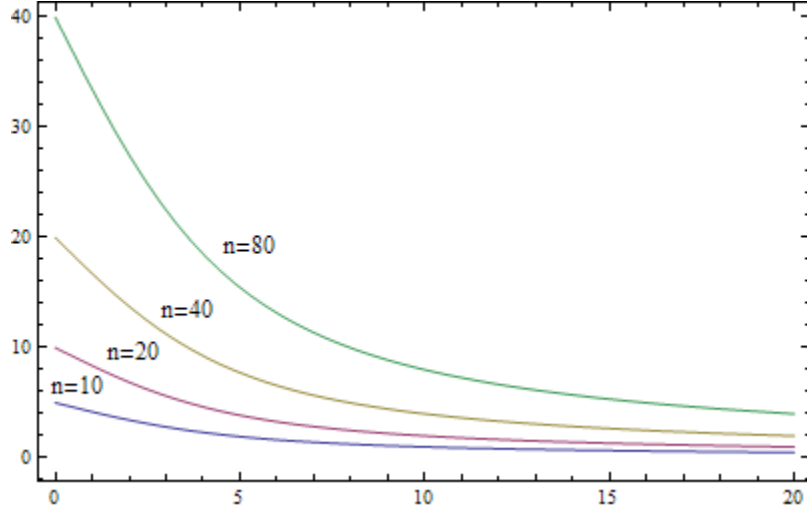


Figura 3. Media de la distribución binomial-exponencial truncada como función de λ , para diferentes valores de n .

4.2 Estimador del parámetro por máxima verosimilitud

Supongamos que disponemos de una muestra de tamaño t procedente de la población (6). El logaritmo de la función de verosimilitud es:

$$\ell = t[\log \lambda - \log(n+1) - \log(1 - e^{-\lambda})] + \sum_{i=1}^t \log {}_1F_1(x_i + 1, n + 2, -\lambda). \quad (10)$$

Derivando (10) respecto a λ e igualando a cero obtenemos la ecuación:

$$\frac{d\ell}{d\lambda} = \frac{t}{\lambda} - \frac{1}{e^{\lambda} - 1} - \sum_{i=1}^t \frac{(x_i + 1) {}_1F_1(x_i + 2, n + 3, -\lambda)}{(n + 2) {}_1F_1(x_i + 1, n + 3, -\lambda)} = 0, \quad (11)$$

para lo que se ha tenido en cuenta (ver Johnson et al., 2005) que

$$\frac{d}{dx} {}_1F_1(a, c, x) = \frac{a}{c} {}_1F_1(a + 1, c + 1, x).$$

La derivada segunda de la función de verosimilitud viene dada por,

$$\frac{d^2 \ell}{d\lambda^2} = -\frac{t}{\lambda^2} - \frac{e^\lambda}{(e^\lambda - 1)^2} - \frac{1}{(n+2)(n+3)} \sum_{i=1}^t \left\{ \frac{(x_i + 2) {}_1F_1(x_i + 3, n + 4, -\lambda)}{{}_1F_1(x_i + 1, n + 3, -\lambda)} \right. \\ \left. - \frac{(x_i + 1) {}_1F_1(x_i + 2, n + 3, -\lambda) {}_1F_1(x_i + 2, n + 4, -\lambda)}{{}_1F_1(x_i + 1, n + 3, -\lambda)} \right\}.$$

La raíz cuadrada de la inversa de la expresión anterior cambiada de signo permite obtener una estimación de la desviación estándar del estimador de máxima verosimilitud.

5 Análisis Bayesiano

Las propiedades que satisface la distribución binomial-exponencial truncada la hacen candidata para ser utilizada en datos de seguros, en los que empíricamente está probado que se manifiesta el fenómeno de la sobredispersión así como la alta presencia de ceros observados. En esta Sección se calculará la prima Bayes, para lo que se requiere de las primas de riesgo y colectiva y, por supuesto, de la distribución a posteriori del parámetro λ .

Suponiendo que los datos provienen de la distribución binomial (1) en la que se admite que el parámetro p es desconocido y aleatorio, elegimos como distribución a priori del parámetro p la distribución exponencial truncada en $(0,1)$ esto es:

$$\pi(p) = \frac{\lambda e^{-\lambda p}}{1 - e^{-\lambda}}, \quad \lambda > 0, 0 < p < 1.$$

El siguiente resultado proporciona la distribución a posteriori del parámetro p dada la información muestral.

Teorema 5 *Supongamos una muestra $\underline{X} = (X_1, X_2, \dots, X_t)$, de tamaño t procedente de una distribución binomial, y una distribución a priori de tipo exponencial truncada. Entonces, la distribución a posteriori viene dada por,*

$$\pi(p | \underline{X}) = \frac{\Gamma(nt + 2)(1 - p)^{t\bar{X}} p^{t(n-\bar{X})} e^{-\lambda p}}{\Gamma(t\bar{X} + 1)\Gamma(t(n - \bar{X}) + 1) {}_1F_1(t(n - \bar{X}) + 1, nt + 2, -\lambda)}. \quad (12)$$

Demostración: Haciendo uso del Teorema de Bayes, la distribución a posteriori es proporcional a:

$$\pi(p | \underline{X}) \propto p^{t\bar{X}} (1 - p)^{t(n-\bar{X})} e^{-\lambda p}.$$

Para la obtención de la constante de normalización hacemos:

$$\begin{aligned} \pi(p | \underline{X}) &= \frac{p^{t\bar{X}} (1 - p)^{t(n-\bar{X})} e^{-\lambda p}}{\int_0^1 p^{t\bar{X}} (1 - p)^{t(n-\bar{X})} e^{-\lambda p} dp} = \frac{p^{t\bar{X}} (1 - p)^{t(n-\bar{X})} e^{-\lambda p}}{\int_0^1 p^{(t\bar{X}+1)-1} (1 - p)^{nt+2-(t\bar{X}+1)-1} e^{-\lambda p} dp} \\ &= \frac{\Gamma(nt + 2)(1 - p)^{t\bar{X}} p^{t(n-\bar{X})} e^{-\lambda p}}{\Gamma(t\bar{X} + 1)\Gamma(t(n - \bar{X}) + 1) {}_1F_1(t(n - \bar{X}) + 1, nt + 2, -\lambda)}, \end{aligned}$$

de donde se obtiene el resultado.

Si adoptamos el principio de prima neta (Lemaire, 1979 y Gómez-Déniz et al., 2002 y 2003; Sarabia et al., 2006; entre otros) como método de cálculo de prima interesará obtener las primas netas de riesgo, colectiva y Bayes que se muestran en el siguiente resultado.

Proposición 6 *Bajo el modelo binomial-exponencial truncada, las primas netas de riesgo, colectiva y Bayes vienen dadas por las expresiones,*

$$\begin{aligned} P_R &= np, \\ P_C &= \frac{n(\lambda - e^\lambda + 1)}{\lambda(1 - e^\lambda)}, \end{aligned} \quad (13)$$

$$P_B = \frac{n(t\bar{X} + 1)}{nt + 2} \frac{{}_1F_1(t\bar{X} + 2, nt + 3, -\lambda)}{{}_1F_1(t\bar{X} + 1, nt + 2, -\lambda)}, \quad (14)$$

respectivamente.

Demostración: La prima de riesgo es $P_R = E(X | \lambda)$, y por tanto será la media de la distribución binomial. Ahora la prima de riesgo es $P_R = E(E(X | \lambda))$ que es como en (9). Finalmente la prima Bayes es la media a posteriori de la prima de riesgo. Por tanto se calculará del siguiente modo.

$$P_B = \int_0^1 np\pi(p | \bar{X})dp$$

$$= \frac{n\Gamma(nt + 2)}{\Gamma(t\bar{X} + 1)\Gamma(t(n - \bar{X}) + 1)} \frac{1}{{}_1F_1(t\bar{X} + 1, nt + 2, -\lambda)} \int_0^1 p^{t\bar{X}+1} (1-p)^{t(n-\bar{X})} e^{-\lambda p} dp,$$

y después de algunos cálculos se obtiene el resultado.

Resulta bien conocido que un sistema de tarificación bonus-malus es un método en el que la prima inicial se va modelando a medida que se incorpora la experiencia de siniestralidad.

Existen diversas metodologías desarrolladas para construir un sistema de esta naturaleza: métodos Bayesianos, métodos markovianos y métodos basados en programación matemática. El lector interesado en profundizar acerca de las diferentes metodologías para construir un sistema de tarificación bonus-malus puede consultar a Lemaire (1985, 1995) y Sarabia et al. (2006). En este trabajo adoptaremos la metodología Bayesiana. Bajo este procedimiento la prima se construye como la relación entre la prima Bayes y la prima colectiva calculadas ambas bajo el mismo principio de cálculo de prima. Escribiendo $t\bar{X} = k$, tenemos que la prima neta bonus-malus bajo el modelo binomial-exponencial truncada en (0,1) será igual a (14) entre (13), de donde:

$$P_{BM} = \frac{\lambda(1 - e^{-\lambda})}{\lambda - e^{-\lambda} + 1} \frac{(k + 1)}{nt + 2} \frac{{}_1F_1(k + 2, nt + 3, -\lambda)}{{}_1F_1(k + 1, nt + 2, -\lambda)}. \quad (15)$$

De este modo se asegura que partiendo de un nivel neutro ($P_{BM} = 1$), si el asegurado no experimenta reclamación la prima a pagar se verá reducida en el siguiente período. Por el contrario, si el asegurado experimenta reclamación se verá penalizado mediante un incremento en la prima a pagar.

6 Aplicaciones numéricas

Presentamos en esta Sección algunas aplicaciones del modelo binomial-exponencial truncada. En particular, se prestará atención, primero al ajuste de datos de seguros de distribuciones de frecuencias de reclamaciones, y en segundo lugar, al cálculo de primas netas bonus-malus para dichos datos. La aplicación pretende ser únicamente ilustrativa y no presentaremos comparaciones con otros modelos existentes en la literatura. Se ajustará el nuevo modelo propuesto y se obtendrán las correspondientes medidas de bondad de ajuste.

Hemos seleccionado tres conjuntos de datos, correspondientes a las siguientes carteras de seguros:

- Bélgica (1975-76), que aparece en Denuit (1997); Gómez-Déniz et al., (2008b); Kestemont y Paris (1985); Klugman et al. (1998); Lemaire (1979), (1985), (1995); Meng et al. (1999); Sarabia et al. (2006) y Willmot (1987) entre otros.
- Alemania (1960), que aparece en Denuit (1997); Gómez-Déniz et al. (2008b); Kokonendji y Khoudar (2004); Willmot (1987) entre otros.
- Bélgica (1993), que aparece en Denuit (1997).

Se han ajustado dichos datos mediante la distribución discreta binomial-exponencial truncada en $(0,1)$ que aparece en (6) utilizando el método de máxima verosimilitud. Para ello se ha hecho uso de la ecuación (11) que permite estimar el valor del parámetro λ y se ha comprobado que la solución corresponde a un máximo. Los valores observados y predichos bajo el nuevo modelo aparecen en las Tablas 1, 2 y 3.

Número de Reclamaciones	Observadas	Ajustadas
0	96978	96984.50
1	9240	9214.95
2	704	727.72
3	43	44.88
4	9	1.92
5	0	0.04

Tabla 1. Ajuste de datos sobre reclamaciones en la cartera de seguros correspondiente a Bélgica, 1975-76

Número de Reclamaciones	Observadas	Ajustadas
0	20592	20556.70
1	2651	2693.49
2	297	307.12
3	41	29.34
4	7	2.21
5	0	0.11
6	1	0.00

Tabla 2. Ajuste de datos sobre reclamaciones en la cartera de seguros correspondiente a Alemania, 1960

Número de Reclamaciones	Observadas	Ajustadas
0	57178	57170.20
1	5617	5606.17
2	446	484.45
3	50	35.90
4	8	2.19
5	0	0.10

Tabla 3. Ajuste de datos sobre reclamaciones en la cartera de seguros correspondiente a Bélgica, 1993

La Tabla 4 muestra entre otros datos de interés, el estimador de máxima verosimilitud del parámetro, el valor del estadístico χ^2 , el logaritmo de la función de verosimilitud. Como se aprecia, el modelo ajusta correctamente los datos. Merece destacarse que el ajuste resulta satisfactorio teniendo en cuenta que únicamente se ha ajustado un solo parámetro. De hecho el modelo propuesto supera a la distribución de Poisson en bondad de ajuste.

Tabla	1	2	3
$\hat{\lambda}$	49.46	41.60	66.21
\bar{x}	0.101	0.144	0.105
s^2	0.107	0.163	0.114
χ^2	1.40	9.05	13.34
g.l.	2	2	2
p -valor	49.65%	1.08%	0.12%
L_{max}	-36107.4	-10230.2	-22067.5

Tabla 4. Resumen

Finalmente, para los datos anteriores y utilizando el valor estimado del parámetro λ para cada uno de los conjuntos de datos, se han calculado de acuerdo a la expresión (15) las primas bonus-malus. Estas aparecen en las Tablas 5, 6 y 7.

	k			
t	0	1	2	3
1	0.90	1.84	2.81	3.82
2	0.82	1.67	2.55	3.46
3	0.76	1.54	2.34	3.16

Tabla 5. Primas netas bonus-malus. Bélgica, 1975-76

	k			
t	0	1	2	3
1	0.86	1.77	2.71	3.69
2	0.77	1.56	2.38	3.22
3	0.69	1.39	2.12	2.86

Tabla 6. Primas netas bonus-malus. Alemania, 1960

	k			
t	0	1	2	3
1	0.92	1.88	2.86	3.87
2	0.86	1.75	2.66	3.59
3	0.81	1.64	2.48	3.35

Tabla 7. Primas netas bonus-malus. Bélgica, 1993

7 Conclusiones

En este trabajo se ha obtenido una nueva distribución de conteo, que puede presentarse como alternativa a las distribuciones clásicas de Poisson, binomial y geométrica. La nueva distribución es unimodal, con moda en cero, sobredispersa o infradispersa, y no presenta dificultad para la estimación del parámetro de la que depende. Estas propiedades la hacen candidata, como aquí se ha mostrado en diversos ejemplos, para ser utilizada en datos de

seguros, en los que empíricamente está probado que se manifiesta el fenómeno de la sobredispersión así como la alta presencia de ceros observados.

Agradecimientos

Los autores agradecen al Ministerio de Educación y Ciencia (proyectos SEJ2006-12685 (EGD) y SEJ2007-65818 (JMS)) por la financiación parcial de este trabajo.

Referencias

- Aitchison, J., Ho, C.H. (1989). The multivariate Poisson-Log normal distribution. *Biometrika*, **76**, 643-653.
- Alanko, T., Duffy, J.C. (1996). Compound binomial distribution for modelling consumption data. *The Statistician*, **45**, 269 - 286.
- Al-Zaid, A.A. (1983). On the unimodality of mixtures. *Pakistan Journal of Statistics*, **5**, 205-209.
- Bain, L.J., Weeks, D.L. (1964). A Note on the Truncated Exponential Distribution. *The Annals of Mathematical Statistics*, **35**, 1366-1367.
- Balakrishnan, N., Nevzorov, V. (2003). *A primer on statistical distributions*. John Wiley and Sons, New York.
- 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.
- Denuit, M. (1997). A new distribution of Poisson-type for the number of claims. *Astin Bulletin*, **27**, 229-242.
- Gómez-Déniz, E., Hernández, A., Pérez, J.M., Vázquez, F. (2002). Measuring Sensitivity in a Bonus-Malus System. *Insurance: Mathematics and Economics*, **31**, 105-113.
- Gómez-Déniz, E., Sarabia, J.M. (2007). Soluciones explícitas para la función de probabilidad de ruina en el modelo clásico de riesgo. En: *Investigaciones en Seguros y Gestión de Riesgos: RIESGO 2007*, 27-41.
- 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**, 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.
- Gómez-Déniz, E., Vázquez, F. (2003). Robustness in Bayesian Model for

- Bonus-Malus Systems. In: *Intelligent and Other Computation Techniques in Insurance. Theory and Applications*. World Scientific.
- Griffiths, D.A. (1973). Maximum Likelihood Estimation for the Beta-Binomial Distribution and an Application to the Household Distribution of the Total Number of Cases of a Disease. *Biometrics*, **29**, 637-648.
 - Gurland, J. (1958). A generalized class of contagious distributions. *Biometrics*, **14**, 229-249.
 - 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.
 - Kestemont, R.M., Paris, J. (1985). Sur l'ajustement du nombre de sinistres. *Mitteilungen der Vereinigung schweiz. Versicherungsmathematiker*, Heft 2.
 - Klugman, S.; Panjer, H., Willmot, G. (1998). *Loss Models. From Data to Decisions*. John Wiley and Sons, New York.
 - Kokonendji, C., Khoudar, M. (2004). *On strict arcsine distribution*. *Communications in Statistics: Theory and Methods*, **33**, 993-1006.
 - 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.
 - Meng, S., Yuan, W., Whitmore, G.A. (1999). Accounting for individual over-dispersion in a bonus-malus automobile insurance system. *Astin Bulletin*, **29**, 327-337.
 - 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.
 - Sarabia, J.M., Gómez-Déniz, E., Vázquez, F. (2006). *Estadística Actuarial. Teoría y Aplicaciones*. Pearson. Prentice-Hall.
 - Sarabia, J.M., Gómez-Déniz, E. (2007). Distribuciones multivariantes Poisson-beta con aplicaciones a datos de seguros. En: *Investigaciones en Seguros y Gestión de Riesgos: RIESGO 2007*, 15-26.
 - Tremblay, L. (1992). Using the Poisson inverse Gaussian in bonus-malus systems. *Astin Bulletin*, **22**, 97-106.
 - Willmot, G.E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial Journal*, 113-127.

La distribución binomial-exponencial truncada con aplicaciones en el sector ...

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