

Aspectos básicos del modelo de riesgo colectivo*

César Escalante C.

Gerardo Arango O.

Recibido Dic. 18, 2003

Aceptado Jun. 12, 2004

Resumen

El *modelo de riesgo colectivo* estudia la suma de un número aleatorio de variables aleatorias, en el contexto actuarial. Aunque no se conoce una solución analítica general para la distribución de probabilidad de la variable aleatoria correspondiente, es posible caracterizarla a través del cálculo de sus tres primeros momentos centrales y del análisis de algunos casos especiales que permiten soluciones analíticas. El propósito de este artículo es presentar en forma integral, que hilvana y relaciona resultados dispersos en la literatura, un modelo matemático de gran interés tanto por su contenido teórico como por sus posibilidades de aplicación al manejo de riesgos en carteras de pólizas de compañías de seguros y en la administración de riesgos de todo tipo de empresa.

Palabras y frases claves: Riesgo colectivo, distribución de pérdidas, teoría de riesgos, pérdidas agregadas.

Abstract

The *collective risk model* studies the sum of a random number of random variables, in the actuarial context. Although there is no known analytic general solution for the probability distribution of the corresponding random variable, it is possible to characterize it in general through the calculation of its first three central moments and in some particular cases by analytic solution. The purpose of this article is to describe, in a comprehensive manner that interweaves and relates results which are widespread in the specialized literature, a mathematical model of great interest not only for its theoretical content but also for its applications to the management of risk portfolios in insurance companies and different other risks in almost any company.

Keywords: Collective risk model, loss distributions, risk theory, aggregate losses.

AMSC(2000): Primary: 62P05, Secondary: 60E05

1 Introducción

“Una buena parte de la teoría de probabilidades se relaciona con la suma de variables aleatorias independientes, y hay muchas situaciones en las que el número de términos es en sí mismo una variable aleatoria”[5]. Un ejemplo actuarial de la realidad indicada por el profesor Feller es el modelo de riesgo colectivo, variable aleatoria (v.a.) de génesis natural en el quehacer de los seguros, que no deja de ser estimulante tanto para el investigador teórico como para el aplicado, pues a pocos pasos de iniciado su estudio manifiesta su complejidad analítica que conlleva al uso general de los métodos numéricos de la simulación para su aproximación.

A pesar de no existir solución analítica general para su distribución de probabilidad, es posible calcular algunos casos particulares basados en la

*El presente artículo es parte de la tesis de Maestría en Matemática Aplicada presentada por César Escalante C. bajo la dirección del Profesor Gerardo Arango O.

clausura bajo convolución de algunas familias de distribuciones. Esto se hace en la última parte del artículo, luego de la caracterización teórica general del fenómeno en estudio.

Desde el punto de vista de una empresa aseguradora, considérese una cartera conformada por un número no especificado de pólizas observadas en conjunto, sin distinguir una de otra. Desde el punto de vista de una industria, un número no especificado de unidades homogéneas expuestas a una misma amenaza o peligro (daños de maquinaria, por ejemplo).

Sea la v.a. no negativa N , número de reclamos o pérdidas acaecidos en un lapso de estudio $[0, t]$; un año, por ejemplo. Sea la sucesión de variables aleatorias (v.as.) no negativas $X_1, X_2, X_3, \dots$ independientes e idénticamente distribuidas con función de distribución (f.d.) común $F(x) = \Pr(X \leq x)$ y función de densidad de probabilidad (f.d.p.) $f(x)$. La v.a. X_j ($j = 1, 2, \dots$) representa la cuantía o severidad de la j -ésima pérdida individual. Todas las v.as. X_j son a la vez independientes de N ([2], [5], [6], [7], [9], [10]).

El total de pérdidas agregadas (anuales) se obtiene entonces a través de la v.a.

$$S = S^{\text{col}} = X_1 + X_2 + \dots + X_N = \sum_{j=1}^N X_j \quad (1)$$

y corresponde al denominado *modelo de riesgo colectivo*. En contraste, en el *modelo de riesgo individual* se estudia la v.a.

$$S^{\text{ind}} = X_1 + X_2 + \dots + X_n = \sum_{j=1}^n X_j, \quad (2)$$

donde el número n es fijo y corresponde al número de pólizas de la cartera en estudio, o bien, el número de máquinas de una industria sometidas a riesgos más o menos homogéneos (desde el punto de vista de la causalidad) y donde no es necesario que las v.as. X_j sean iguales en distribución, y si lo son, se denomina modelo de riesgo individual *homogéneo*. Otra posibilidad de interpretación es

$$S^{\text{ind}} = S^{\text{col } 1} + S^{\text{col } 2} + \dots + S^{\text{col } n} = \sum_{j=1}^n S^{\text{col } j}.$$

En lo sucesivo, el artículo se refiere al modelo de riesgo colectivo, a no ser que se especifique algo distinto.

2 Distribución de probabilidad de $S = S^{\text{col}}$

En esta sección se estudia la distribución de las pérdidas agregadas asociadas al modelo de riesgo colectivo (1).

La f.d. de S es

$$\begin{aligned}
 F_S(x) &= \Pr(S \leq x) \\
 &= \sum_{n=0}^{\infty} \underbrace{\Pr(N=n)}_{p_n} \Pr(S \leq x | N=n) \\
 &= \sum_{n=0}^{\infty} p_n \Pr(X_1 + X_2 + \cdots + X_n \leq x | N=n) \\
 &= \sum_{n=0}^{\infty} p_n F^{*n}(x),
 \end{aligned} \tag{3}$$

donde $F(x) = \Pr(X \leq x)$ es la f.d. común de las v.as. X_j y

$$F^{*n}(x) = \Pr(X_1 + \cdots + X_n \leq x | N=n)$$

es la n -ésima convolución de F ([2], [7], [8]).

Cuando N sigue un modelo Poisson, se dice que S tiene distribución compuesta de Poisson¹. Este modelo particular es de capital importancia en el estudio del problema clásico de la ruina de un asegurador ([6], [7], [12]).

En general la n -ésima convolución de F se obtiene como sigue: Para $n=0$

$$F^{*0}(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}.$$

Para $n=1, 2, \dots$

$$F^{*n}(x) = \int_{-\infty}^{\infty} F^{*(n-1)}(x-y) dF(y).$$

Si X es una v.a. continua no negativa (como en el caso en estudio), entonces: Para $n=1$

$$F^{*1}(x) = F(x).$$

Para $n=2, 3, \dots$

$$F^{*n}(x) = \int_0^x F^{*(n-1)}(x-y) f(y) dy.$$

En efecto, la variable de integración y queda restringida al intervalo $[0, x]$ porque la v.a. X toma sólo valores no negativos y su f.d.p. es

$$f(x) = \begin{cases} \frac{d}{dx} F(x), & x \geq 0 \\ 0, & \text{en otros casos} \end{cases}.$$

¹Algunos autores exigen, para la distribución compuesta de Poisson, que las v.as. X_j ($j = 1, 2, \dots, N$) tengan distribución gamma. No es el caso en este artículo.

Y, por diferenciación,

$$f^{*n}(x) = \frac{d}{dx} F^{*n}(x) = \int_0^x f^{*(n-1)}(x-y) f(y) dy.$$

Así, la f.d.p. de S es

$$f_S(x) = \sum_{n=0}^{\infty} p_n f^{*n}(x). \quad (4)$$

En resumen, la f.d. de S viene dada por (3) y, si la v.a. X es continua no negativa, entonces su f.d.p. queda determinada por (4).

3 Cálculo de los tres primeros momentos centrales de S

Para una v.a. cualquiera Y se define el k -ésimo momento respecto del origen como

$$\mu'_{Yk} = E(Y^k)$$

y cuando no hay ambigüedad, la notación se simplifica a μ'_k , sin especificar la v.a. Y . Así, la media de Y corresponde a su primer momento respecto del origen: $\mu_Y = \mu'_{Y1} = E(Y)$.

De forma similar, el k -ésimo momento central de Y se define como

$$\mu_{Yk} = E[(Y - \mu_Y)^k]$$

y cuando no hay ambigüedad, la notación se simplifica a μ_k . De esta manera, la varianza de Y corresponde a su segundo momento central:

$$\text{Var}(Y) = \sigma_Y^2 = E[(Y - \mu_Y)^2] = \mu_{Y2},$$

que, con un poco de álgebra y aplicando algunas propiedades básicas de la esperanza, también se expresa en términos de los momentos respecto del origen: $\sigma_Y^2 = \mu'_{Y2} - (\mu_Y)^2$ ([7], [9]).

El cálculo de los tres momentos centrales de S se hará a través de su función generadora de probabilidades.

Definición 1 (Función generadora de probabilidades). La función generadora de probabilidades (f.g.p.) de una variable aleatoria Y con f.d. F , se define por

$$P_Y(z) = E(z^Y) = \int_{-\infty}^{\infty} z^y dF(y), \quad z > 0$$

siempre que la integral exista ([9]).

Si una v.a. tiene f.g.p., ésta es única, y dos v.as. con igual f.g.p. tienen la misma distribución de probabilidad. Para una variable aleatoria cualquiera Y con esperanza finita, se verifica

$$P'_Y(1) = \left. \frac{d}{dz} P_Y(z) \right|_{z=1} = E(Y z^{Y-1})|_{z=1} = E(Y) \quad (5)$$

y de manera análoga,

$$P''_Y(1) = E[Y(Y-1)z^{Y-2}]|_{z=1} = E(Y^2) - E(Y). \quad (6)$$

En particular, por ser N una v.a. discreta, su f.g.p. es

$$P_N(z) = \sum_{n=0}^{\infty} \Pr(N=n) z^n = \sum_{n=0}^{\infty} p_n z^n.$$

Más específico aún, cuando N se distribuye Poisson, su f.g.p. es

$$P_N(z) = \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} z^n = e^{\lambda(z-1)}. \quad (7)$$

Proposición 2. La f.g.p. de S , denominada función compuesta ([7], [9]), es

$$P_S(z) = P_N[P_X(z)]. \quad (8)$$

Demostración. Debido a la independencia de las v.as. X_j , se verifica

$$\begin{aligned} P_S(z) &= E(z^S) = E(z^{X_1 + \dots + X_N}) \\ &= \sum_{n=0}^{\infty} \Pr(N=n) E(z^{X_1 + \dots + X_n} | N=n) \\ &= \sum_{n=0}^{\infty} p_n E\left(\prod_{j=1}^n z^{X_j}\right) = \sum_{n=0}^{\infty} p_n \underbrace{P_X(z) \cdots P_X(z)}_{n \text{ factores}} \\ &= \sum_{n=0}^{\infty} p_n [P_X(z)]^n = E\left\{[P_X(z)]^N\right\} \\ &= P_N[P_X(z)]. \end{aligned}$$

Lo que justifica el nombre de *función compuesta* para la f.g.p. de S . $\square$

Proposición 3. Los tres primeros momentos centrales de S son

$$E(S) = \mu_S = \mu_N \cdot \mu_X = E(N) \cdot E(X) \quad (9)$$

$$Var(S) = \mu_N \cdot \mu_{X^2} + \mu_{N^2} \cdot (\mu_X)^2 \quad (10)$$

$$\mu_{S^3} = E[(S - \mu_S)^3] = \mu_N \cdot \mu_{X^3} + 3\mu_{N^2} \cdot \mu_X \cdot \mu_{X^2} + \mu_{N^3} \cdot (\mu_X)^3, \quad (11)$$

donde el primer subíndice indica la variable y el segundo (de los que lo tienen) el orden del momento centrado ([7]).

Demostración. Las dos primeras derivadas de (8) son

$$P'_S(z) = P'_N[P_X(z)] \cdot P'_X(z)$$

y

$$P''_S(z) = P'_N[P_X(z)] \cdot P''_X(z) + [P'_X(z)]^2 \cdot P''_N[P_X(z)].$$

Por tanto, aplicando (5) y la primera derivada de (8), se obtiene

$$E(S) = P'_S(1) = \underbrace{P'_N[P_X(1)]}_{1} \cdot \underbrace{P'_X(1)}_{E(X)} = E(N) \cdot E(X)$$

con lo cual se justifica (9). Para probar (10), primero se calcula $E(S^2)$ a partir de (6) y la segunda derivada de (8) como sigue

$$\begin{aligned} E(S^2) &= P''_S(1) + E(S) \\ &= \underbrace{P'_N[P_X(1)] \cdot P''_X(1)}_{E(N)} + \underbrace{[P'_X(1)]^2}_{[E(X)]^2} \cdot \underbrace{P''_N[P_X(1)]}_1 + E(N) \cdot E(X) \\ &= E(N) [E(X^2) - E(X)] + [E(X)]^2 [E(N^2) - E(N)] \\ &\quad + E(N) \cdot E(X) \end{aligned}$$

y en consecuencia

$$\begin{aligned} Var(S) &= E(S^2) - [E(S)]^2 \\ &= E(N) \cdot Var(X) + Var(N) \cdot [E(X)]^2 \\ &= \mu_N \mu_{X2} + \mu_{N2} (\mu_X)^2 \end{aligned}$$

con lo cual se prueba (10). Una demostración diferente de los dos primeros momentos aparece en ([11]). Por último, cálculos similares conducen a la demostración de (11). $\square$

En particular, si N tiene distribución de Poisson con parámetro λ , entonces ([5], [7], [8]) los tres primeros momentos de $S = S^{\text{col}}$ son

$$\begin{aligned} E(S) &= \lambda \cdot E(X) \\ Var(S) &= \lambda \{ Var(X) + [E(X)]^2 \} = \lambda \cdot E(X^2) \\ \mu_{S3} &= E[(S - \mu_S)^3] = \lambda [\mu_{X3} + 3\mu_X \cdot \mu_{X2} + (\mu_X)^3]. \end{aligned}$$

Observación. Al presentar el modelo general estudiado en esta sección para la distribución de S^{col} correspondiente al modelo de riesgo colectivo (1) y al

mentar la posibilidad de estudiar el modelo de riesgo individual *homogéneo* correspondiente a S^{ind} , se afirmó que se consideraba que las cuantías de las pérdidas individuales X (en ambos modelos) eran independientes e idénticamente distribuidas. Pues bien, cuando se renuncia a la hipótesis de independencia, la matemática necesaria para el estudio de S^{ind} se vuelve más compleja, y se incursiona en el terreno de los órdenes estocásticos. Para hacerse una idea más precisa, el lector puede consultar el artículo ([3]) donde, en particular, se demuestra que cuando hay dependencia entre parejas de riesgos, la prima *stop-loss*² aumenta con respecto a cuando hay independencia. No obstante, a pesar de tal dependencia, la prima *stop-loss* está acotada superiormente por un caso extremo de dependencia denominada *comonotonía*. El estudio de estos modelos con dependencia está por fuera del alcance del presente artículo.

4 Algunos resultados analíticos para la distribución de S y otras propiedades particulares

Como se observa, la distribución compuesta de probabilidad de $S = S^{\text{col}}$ es de alta complejidad para el manejo analítico general, lo que conduce a usar métodos numéricos para su tratamiento. No obstante, se dispone de ciertos resultados analíticos.

En la primera parte de esta sección se presenta el concepto de familias de distribuciones cerradas bajo convolución y se aplica a la distribución de S . En las tres siguientes partes se estudia la f.d. de S , con supuestos sobre las distribuciones de N o X . Se finaliza presentando la clausura de la suma de v.as. con distribuciones independientes compuestas de Poisson.

4.1 Familias de distribuciones cerradas bajo convolución

Una familia de distribuciones es *cerrada bajo convolución* cuando al sumar miembros independientes e idénticamente distribuidos, se produce un miembro de la misma familia. Adicionalmente, en ocasiones, para una familia paramétrica de distribuciones, la parametrización cumple la siguiente propiedad: al sumar n miembros de la familia se obtiene un miembro de la misma tal que uno de los parámetros no cambia, y el resto quedan multiplicados por n .

Para determinar la distribución de $S = \sum_{j=1}^N X_j$ cuando las cuantías de las pérdidas X_j tienen esta propiedad, se procede como sigue:

La condición significa que si $f_X(x; a)$ es la función de probabilidad o de densidad de cada X_j , entonces la función de probabilidad de $\sum_{j=1}^N X_j$ es $f_X(x; na)$. En consecuencia, de (4),

²La prima *stop-loss* para los riesgos (v.as. no negativas) X_1 y X_2 se define por $E[\max\{X_1 + X_2 - d, 0\}] = E[(X_1 + X_2 - d)^+]$ para toda retención o deducible $d \geq 0$.

$$f_S(x) = \sum_{n=0}^{\infty} p_n f_X^{*n}(x; a) = \sum_{n=0}^{\infty} p_n f_X(x; na) \quad (12)$$

y se elimina la necesidad de evaluar la convolución. La distribución gamma es cerrada bajo convolución y goza de la propiedad mencionada ([2], [7], [9]).

Es bien conocido que la *función generadora de momentos* (f.g.m.) de una v.a. X es $M_X(z) = E(e^{Xz})$. La relación entre la f.g.p. con la f.g.m. es inmediata ([7], [9]): $P_X(e^z) = M_X(z)$.

Sea $Y = X_1 + X_2 + \dots + X_n$ la suma de n v.as. independientes e idénticamente distribuidas con f.d.p. común gamma (θ, α)

$$f_X(x; \theta, \alpha) = f_X(x) = \frac{(x/\theta)^\alpha e^{-x/\theta}}{x \Gamma(\alpha)}$$

y f.g.m. $M_X(z) = (1 - \theta z)^{-\alpha}$. Entonces la f.g.m. de Y es

$$\begin{aligned} M_Y(z) &= E[e^{(X_1 + \dots + X_n)z}] = E(e^{X_1 z}) \dots E(e^{X_n z}) \\ &= [M_X(z)]^n = (1 - \theta z)^{-n\alpha} \end{aligned}$$

que corresponde a la f.g.m. de una v.a. con distribución gamma $(\theta, n\alpha)$, luego la familia es cerrada bajo convolución y goza de la propiedad mencionada.

En particular, como la distribución exponencial es un caso especial de la gamma con $\alpha = 1$, entonces se obtiene que la suma de n v.as. exponenciales independientes con media θ es una v.a. gamma (θ, n) . Este resultado se usa en la siguiente sección.

4.2 Cuantía de pérdidas individuales exponencial

Bajo los supuestos del modelo de riesgo colectivo, supóngase que las cuantías de las pérdidas individuales X_j siguen el modelo exponencial con media θ .

Del último resultado de la sección anterior se sigue que la f.d. de $X_1 + \dots + X_n$ es

$$\begin{aligned} \Pr(X_1 + \dots + X_n \leq x) &= F^{*n}(x) = \frac{1}{\Gamma(n)} \int_0^{x/\theta} t^{n-1} e^{-t} dt \\ &= \Gamma\left(n; \frac{x}{\theta}\right), \end{aligned} \quad (13)$$

esto es, la *función gamma incompleta*. En ([7]) se prueba que para valores enteros de n (como ocurre en el presente caso), el valor de $\Gamma(n; t)$ se obtiene en forma exacta según la ecuación

$$\Gamma(n; t) = 1 - \sum_{j=0}^{n-1} \frac{t^j e^{-t}}{j!}, \quad n = 1, 2, 3, \dots \quad (14)$$

es decir, la f.d. de $X_1 + \dots + X_n$ es la *cola* de una distribución de Poisson de parámetro x/θ . En consecuencia, de (3), (13) y (14),

$$\begin{aligned}
 F_S(x) &= \sum_{n=0}^{\infty} p_n F^{*n}(x) = p_0 + \sum_{n=1}^{\infty} p_n F^{*n}(x), \quad x \geq 0 \\
 &= p_0 + \sum_{n=1}^{\infty} p_n \Gamma\left(n; \frac{x}{\theta}\right) = p_0 + \sum_{n=1}^{\infty} p_n \left[1 - \sum_{j=0}^{n-1} \frac{(x/\theta)^j e^{-x/\theta}}{j!}\right] \\
 &= 1 - e^{-x/\theta} \sum_{n=1}^{\infty} p_n \sum_{j=0}^{n-1} \frac{(x/\theta)^j}{j!}, \quad x \geq 0 \\
 &= 1 - e^{-x/\theta} \sum_{j=0}^{\infty} \frac{(x/\theta)^j}{j!} \sum_{n=j+1}^{\infty} p_n \\
 &= 1 - e^{-x/\theta} \sum_{j=0}^{\infty} \bar{P}_j \frac{(x/\theta)^j}{j!}, \quad x \geq 0,
 \end{aligned} \tag{15}$$

donde $\bar{P}_j = \sum_{n=j+1}^{\infty} p_n$, $j = 0, 1, 2, \dots$

En particular, si N tiene distribución binomial con parámetros (m, q) , entonces, de (15) se obtiene

$$F_S(x) = 1 - e^{-x/\theta} \sum_{n=1}^m \binom{m}{n} q^n (1-q)^{m-n} \sum_{j=0}^{n-1} \frac{(x/\theta)^j}{j!}, \quad x \geq 0 \tag{16}$$

Por otra parte, a partir de (12) y suponiendo que N tiene distribución logarítmica³ con parámetro β , a continuación se presenta el cálculo de la función de densidad de S :

$$\begin{aligned}
 f_S(x) &= \sum_{n=1}^{\infty} p_n f_X(x; n) \\
 &= \sum_{n=1}^{\infty} \frac{\beta^n}{n(1+\beta)^n \ln(1+\beta)} \cdot \frac{x^{n-1} e^{-x/\theta}}{\theta^n (n-1)!} \\
 &= \frac{e^{-x/\theta}}{x \ln(1+\beta)} \sum_{n=1}^{\infty} \frac{1}{n!} \left[\frac{\beta x}{\theta(1+\beta)} \right]^n.
 \end{aligned}$$

4.3 Distribución compuesta geométrica-exponencial

Antes de entrar en el tema central de esta sección se requiere introducir el concepto de mixtura n -puntos y presentar una mixtura 2-puntos especial.

³ $\Pr(N = k) = \frac{\beta^k}{k(1+\beta)^k \ln(1+\beta)}$, $k = 1, 2, \dots$

Definición 4 (Mixtura n -puntos). Una variable aleatoria M es una mixtura n -puntos de las variables aleatorias X_j con funciones de distribución F_{X_j} , ($j = 1, 2, \dots, n$) si su f.d. es $F_M(y) = a_1 F_{X_1}(y) + a_2 F_{X_2}(y) + \dots + a_n F_{X_n}(y)$, donde $a_j \geq 0$, ($j = 1, 2, \dots, n$) y $a_1 + a_2 + \dots + a_n = 1$ ([7]).

$F_M(y)$ es una función de distribución. En efecto, cuando $y \rightarrow \infty$, entonces $F_M(y) \rightarrow 1$, pues $a_1 + a_2 + \dots + a_n = 1$. Cuando $y \rightarrow -\infty$, entonces $F_M(y) \rightarrow 0$, pues para $j = 1, 2, \dots, n$, $F_{X_j}(y) \rightarrow 0$. Por tratarse de una combinación lineal de funciones continuas por la derecha, entonces $F_M(y)$ es continua por la derecha. Por último, $F_M(y)$ es no de creciente, pues $a_j \geq 0$, ($j = 1, 2, \dots, n$).

La f.g.m. de la mixtura 2-puntos $F(x) = \delta F_1(x) + (1 - \delta) F_2(x)$, ($\delta \geq 0$) de una v.a. degenerada X_1 con probabilidad 1 en $X_1 = k$ y f.d. F_1 , y una variable aleatoria X_2 con media ν y f.d. exponencial F_2 es ([4],[7]):

$$M(z) = \delta e^{kz} + (1 - \delta) \frac{1}{1 - \nu z}$$

En efecto, la f.g.m. de X_1 es $M_1(z) = e^{kz}$ ([1]), y como la f.g.m de X_2 es $M_2(z) = 1/(1 - \nu z)$, en consecuencia, la f.g.m. de la mixtura 2-puntos es entonces

$$\begin{aligned} M(z) &= \int_{-\infty}^{\infty} e^{zx} dF(x) = \int_{-\infty}^{\infty} e^{zx} d[\delta F_1(x) + (1 - \delta) F_2(x)] \\ &= \delta \int_{-\infty}^{\infty} e^{zx} dF_1(x) + (1 - \delta) \int_{-\infty}^{\infty} e^{zx} dF_2(x) \\ &= \delta M_1(z) + (1 - \delta) M_2(z) \\ &= \delta e^{kz} + (1 - \delta) \frac{1}{1 - \nu z} \end{aligned}$$

En particular, para $k = 0$, se obtiene

$$M(z) = \delta \cdot 1 + (1 - \delta) \frac{1}{1 - \nu z} \quad (17)$$

Bajo los supuestos del modelo de riesgo colectivo, si la frecuencia de pérdidas N tiene distribución geométrica con parámetro β y f.g.p. $P_N(z) = [1 - \beta(z - 1)]^{-1}$ y las cuantías de las pérdidas individuales X_j siguen el modelo exponencial con media θ , entonces la f.g.m. de $S = \sum_{j=1}^N X_j$ es

$$\begin{aligned} M_S(z) &= P_S(e^z) = P_N[P_X(e^z)] = P_N[M_X(z)] \\ &= \frac{1}{1 - \beta \left(\frac{1}{1 - \theta z} - 1 \right)} = \frac{1}{1 + \beta} \cdot 1 + \frac{\beta}{1 + \beta} \cdot \frac{1}{1 - \theta(1 + \beta)z} \end{aligned}$$

y corresponde, según (17), a la mixtura 2-puntos de una distribución degenerada con probabilidad uno en cero y una distribución exponencial con media $\theta(1 + \beta)$. Por tanto, la f.d. de S es

$$F_S(x) = 1 - \frac{\beta}{1 + \beta} \exp \left[-\frac{x}{\theta(1 + \beta)} \right], \quad x \geq 0,$$

función con un salto en 0 y continua en el intervalo $(0, +\infty)$.

4.4 Distribución compuesta binomial negativa-exponencial

Bajo los supuestos del modelo de riesgo colectivo, si la frecuencia de pérdidas N tiene distribución binomial negativa⁴ con parámetro entero r y las cuantías de las pérdidas individuales X_j siguen el modelo exponencial con media θ , entonces ([7], [12]) la f.g.m. de $S = \sum_{j=1}^N X_j$ es

$$\begin{aligned} M_S(z) &= P_N[M_X(z)] \\ &= P_N\left(\frac{1}{1 - \theta z}\right) = \left[1 - \beta\left(\frac{1}{1 - \theta z} - 1\right)\right]^{-r} \\ &= \left\{1 + \frac{\beta}{1 + \beta} [1 - \theta(1 + \beta)z]^{-1} - 1\right\}^r \end{aligned}$$

que es de la forma $M_S(z) = P_N^*[M_X^*(z)]$, donde

$$P_N^*(z) = \left[1 + \frac{\beta}{1 + \beta}(z - 1)\right]^r$$

es la f.g.p. de una nueva v.a. N con distribución binomial con parámetros r y $\beta/(1 + \beta)$, y $M_X^*(z) = [1 - \theta(1 + \beta)z]^{-1}$ es la f.g.m. de una nueva v.a. X con distribución exponencial con media $\theta(1 + \beta)$.

En consecuencia, para las nuevas v.as. N y X con las que se calculará la f.d. de S , se verifica

$$p_n = \binom{r}{n} \left(\frac{\beta}{1 + \beta}\right)^n \left(\frac{1}{1 + \beta}\right)^{r-n}, \quad n = 0, 1, \dots, r$$

y, por (14)

$$\begin{aligned} F^{*n}(x) &= \Gamma\left(n; \frac{x}{\theta(1 + \beta)}\right) \\ &= 1 - \sum_{j=0}^{n-1} \frac{\{x/[\theta(1 + \beta)]\}^j \exp\{-x/[\theta(1 + \beta)]\}}{j!}. \end{aligned}$$

⁴Función de probabilidad $p_k = \binom{k + r - 1}{k} \left(\frac{1}{1 + \beta}\right)^r \left(\frac{\beta}{1 + \beta}\right)^k$, $k = 0, 1, 2, \dots$, $\beta > 0$ y f.g.p. $P(z) = [1 - \beta(z - 1)]^{-r}$. También es llamada distribución de Pascal.

De esta manera se obtiene, de (16),

$$F_S(x) = 1 - \exp \left[-\frac{x}{\theta(1+\beta)} \right] \cdot \sum_{n=1}^r \left\{ \binom{r}{n} \left(\frac{\beta}{1+\beta} \right)^n \left(\frac{1}{1+\beta} \right)^{r-n} \cdot \sum_{j=0}^{n-1} \frac{\{x/[\theta(1+\beta)]\}^j}{j!} \right\}$$

para $x \geq 0$.

Es importante observar que, como la distribución geométrica es un caso particular de la binomial negativa con $r = 1$, la distribución de probabilidad de la compuesta geométrica - exponencial se obtiene a partir de la compuesta Pascal - exponencial. El caso particular se presentó con el ánimo de introducir otras herramientas analíticas útiles en esta disciplina.

4.5 Suma de distribuciones compuestas de Poisson

En la presentación del modelo de riesgo colectivo se definió la distribución compuesta de Poisson como aquella en la cual la v.a. frecuencia de pérdidas anuales N sigue la ley de Poisson y las cuantías de las pérdidas individuales X_j cualquier distribución de probabilidad no especificada.

La siguiente proposición indica que al sumar v.as. independientes compuestas de Poisson, se obtiene otra variable compuesta de Poisson ([2]).

Proposición 5. *Si $S_1, S_2, \dots, S_m$ son variables aleatorias independientes y cada una con distribución compuesta de Poisson con parámetro λ_j y f.d. F_j , ($j = 1, 2, \dots, m$), entonces $S = S_1 + S_2 + \dots + S_m$ también tiene distribución compuesta de Poisson con parámetro $\lambda = \sum_{j=1}^m \lambda_j$ y f.d. $F(x) = \sum_{j=1}^m (\lambda_j/\lambda) F_j(x)$.*

Demostración. La f.g.m. de cada S_j , ($j = 1, 2, \dots, m$) es

$$M_{S_j}(z) = P_{N_j} [M_{X_j}(z)] = \exp \{ \lambda_j [M_{X_j}(z) - 1] \}.$$

Por el supuesto de independencia de $S_1, S_2, \dots, S_m$, la f.g.m. de S es entonces

$$\begin{aligned} M_S(z) &= \prod_{j=1}^m M_{S_j}(z) = \exp \left\{ \sum_{j=1}^m \lambda_j [M_{X_j}(z) - 1] \right\} \\ &= \exp \left\{ \lambda \left[\sum_{j=1}^m \frac{\lambda_j}{\lambda} M_{X_j}(z) - 1 \right] \right\} \end{aligned}$$

que corresponde a la f.g.m. de una distribución compuesta de Poisson con parámetro $\lambda = \sum_{j=1}^m \lambda_j$ y f.d. $F(x) = \sum_{j=1}^m (\lambda_j/\lambda) F_j(x)$, correspondiente a una mezcla m -puntos. $\square$

Referencias

- [1] Apostol, T.. Análisis Matemático, 2^a. Editorial Reverté, 1988.
- [2] Bowers, N., Gerber, H. et al. Actuarial Mathematics. The Society of Actuaries, 1997.
- [3] Dhaene, Jan, Goovaerts, Marc J. Dependency of Risk and Stop-Loss Order . Astin Bulletin, Vol. 26, N° 2, 1996.
- [4] Escalante, C. Análisis Estadístico de Pérdidas en Seguros Generales. Tesis de Maestría en Matemática Aplicada. Universidad EAFIT, 2003.
- [5] Feller, William. Introducción a la Teoría de Probabilidades y sus Aplicaciones. Limusa - Wiley, S.A. Vol 1, 1973.
- [6] Gerber, Hans U. An Introduction to Mathematical Risk Theory. S. S. Huebner Foundation for Insurance Education Wharton School University of Pennsylvania Philadelphia, 1979.
- [7] Klugman, Stuart A.; Panjer, Harry H.; Willmot, Gordon E. Loss Models: From Data to Decisions. Wiley, 1998.
- [8] Latorre Llorens, Luis. Teoría del Riesgo y sus Aplicaciones a la Empresa Aseguradora. MAPFRE, 1992.
- [9] Panjer, Harry H. Willmot, Gordon E. Insurance Risk Models. Society of Actuaries, 1992.
- [10] Rolsky, Tomasz. Some Problems in the Theory of Risk. Tokio Institute of Tecnology and University of Wroclaw, 1999.
- [11] Ross, Sheldon M. Stochastic Processes. Wiley. Second Edition, 1996.
- [12] Willmot, Gordon E.; Lin, X Sheldon. Lundberg Aproximations for Compound Distributions with Insurance Applications. Lectures Notes in Statistics 156. Springer, 2001.

Dirección de los autores: Arango G., Univ. EAFIT, Medellín, garango@eafit.edu.co
— Escalante C., Delima Marsh S.A., Cesar.E.Escalante@Marsh.com