ACT-11302: Cálculo Actuarial III **ITAM**

Modelos agregados (Distribución de severidades)

Prof: Juan Carlos Martínez Ovando *13 de octubre de 2015*

Las distribuciones compuestas se emplean para modelar el monto agregado de reclamos de un portafolio de seguros. Sea N el número de pólizas (o reclamaciones) que integran el portafolio, y X_i el monto del siniestro.

Nota: Recuerde que cuando trabajamos con portafolios de seguros, éstos se segmentan en grupos, de manera que al interior de cada grupo la exposición al riesgo de ocurrir un siniestro es homogéneo. De igual forma, las condiciones de cobertura de las pólizas en cada grupo son homogéneas. Así, podemos referirnos al monto del siniestro, X_i, o al monto del reclamo a la aseguradora, Y_i casi indistintamente.

Así, la variable de interés se describe como:

$$S = \sum_{i=1}^{N} X_i. \tag{1}$$

La distribución que nos interesa describir es la de S. Ahora bien, tal distrbución estará en función de la forma de modelación de N, distinguiendo entre dos tipos de modelos agregados de siniestros: i) modelos de riesgo individuales, y ii) modelos de riesgo colectivo.

Modelo de riesgo individual

En el modelo de riesgo individual, se supone que N corresponde al número de pólizas en el portafolio de seguros. Así, la distribución de S estará definida solamente por la distribución conjunta de los siniestros individuales,

$$\Pr(X_1, ..., X_N). \tag{2}$$

El supuesto común respecto a los siniestros individuales los considera estocásticamente independientes e idénticamente distribuidos,

$$\Pr(X_1, ..., X_N) = \prod_{i=1}^{N} \Pr(X_i). \tag{3}$$

Así, la distribución de S quesdaría expresada como la convolución de N variables aleatorias. Tal distribución puede calcularse empleando directamente la fórmula de convolución o simplificando su cálculo empleando la función generadora de momentos o función característica de X.

Respecto a $\Pr(X)$, debe notarse que tal distribución estar'a en función del concepto de N:

- Si N representa el **número de pólizas**, entonces X_i debe considerar la posibilidad de no siniestro , i.e. $X_i = 0$ (debido a la ausencia de siniestro). De esta forma,

  $$\Pr(X_i \leq x) = q(X_i = 0) + (1 - q)F_{X_i}(x), \tag{4}$$

  para $x \in \mathscr{X} = \Re_+$. En esta expresión, q representa la probabilidad de que NO haya siniestro, y $(1 - q)$ su complemento. También, $F_{X_i}(x)$ representa la distribución de una variable aleatoria positiva, la cual representa el monto del siniestro.

  Siendo esta distribución mezclada, el cálculo de S mediante la fórmula de convolución o método de momentos es complicado. El empleo de simulación estocástica puede ser la mejor opción en la práctica.

- Si N representa la **frecuencia de siniestros**, entonces X_i representaría el monto del siniestro, con

  $$\Pr(X_i \leq x) = F_{X_i}(x), \tag{5}$$

la distribución para el monto del siniestro. En este caso, la fórmula de convolución o el método de momentos puede ser la mejor alternativa de uso en la práctica, si F_{X_i} pertenece a una clase conocida de distribuciones.

Modelo de riesgo colectivo

En el modelo de riesgo colectivo, se supone que N corresponde al número (frecuencia) de siniestros en el portafolio de seguros, suponiendo que tal frecuencia *desconocida* y *aleatoria*. En este caso, la distribución de S estará definida por la distribución conjunta de la frecuencia de los siniestros y del monto (o severidad) los mismos a nivel individual,

$$\Pr(N, X_1, ..., X_N). \tag{6}$$

En este caso, las variables X_i representan el monto del siniestro, condicional en que el siniestro individual haya ocurrido.

Usualmente, se supone que la frecuencia de los siniestros es estocásticamente independiente de los montos de los siniestros individuales, i.e.

$$\Pr(N, X_1, ..., X_N) = \Pr(N) \Pr(X_1, ..., X_N | N). \tag{7}$$

Adicionalmente, se supone que los montos de los siniestros individuales son condicionalmente independientes, dado N, i.e.

$$\Pr(X_1, ..., X_N | N) = \prod_{i=1}^{N} \Pr(X_i). \tag{8}$$

Suponiendo que N es una variable aleatoria, la distribución agregada de los siniestros en el portafolio, S, se conoce como una distribución compuesta. En el argot de estadística actuarial, la distribución de N, $\Pr(N)$, se conoce como **distriución primaria**, mientras que la distribución del monto individual del siniestro, $\Pr(X_i)$, se conoce como **distribución secundaria** de la distribución de S.

El cálculo de la distribución de S en este caso puede realizarse empleando la función generadora de momentos. Supongamos:

- $\Pr(N)$ es una distribución con soporte en $\mathscr{N} = \{1,2,3,\ldots\}$ (enteros positivos), con función generadora de momentos $M_N(t)$, y
- $\Pr(X_i \leq x) = F_{X_i}(x)$ es una función de distribución (continua o absolutamente continua), con soporte en $\mathscr{X} = (0,\infty)$, y función generadora de momentos $M_X(t)$,

entonces, la función generadora de momentos del monto agregado de siniestros sería,

$$M_S(t) = M_n\left(\log M_X(t)\right). \tag{9}$$

De esta expresión se toma el término de que $\Pr(S \leq x)$ sea una distribución compuesta. Este cálculo descansa en el supuesto de que los montos individuales de siniestros, X_i's, sean condicionalmente i.i.d. dado N. Note también que el supuesto de que las X_i's sean intermabiables dado N también aplica, para poder hacer uso del enfoque bayesiano de inferencia.

Ejemplo 1: Distribución compuesta Poisson-gamma. Suponga que N es una variable aleatoria con distribución $\mathrm{Po}(N|\lambda)$ y que las X_i's son variables aleatorias con distribución $\mathrm{Ga}(x|\alpha,\beta)$. Entonces, la distribución de S se define como la distribución Poisson compuesta con respecto a la distribución gamma. La función generadora de probabilidades de S estaría expresada como

$$P_S(t) = P_N\left(P_X(t)\right) = \exp\left\{\lambda\left(P_X(t)-1\right)\right\}, \tag{10}$$

con $P_X(t)$ la distribución generadora de probabilidades del monto individual de siniestros.

Note que el cálculo de la expresión anał'itica para F_S es complicado. Su cálculo en la práctica descansa en métodos recursivos de cálculo, como el método de Pánjer (1981). Trabajando con recursiones, será más conveniente definir una expresión general en términos de la distribución $(\alpha, \beta, 0)$.

Ejemplo 2: Distribución compuesta $(\alpha, \beta, 0) - F_X$. Supongamos que N sigue una distribución $(\alpha, \beta, 0)$, siendo las X_i's variables aleatorias con ditribución F_X soportadas en los enteros no negativos, $\mathscr{X} = \{1, 2, 3, \ldots\}$ (este supuesto es fundamental). Entonces la distribucuón de S puede expresarse como

$$f_S(s) = \frac{1}{1 - \alpha f_X(0)} \sum_{x=1}^{s} \left(\alpha + \frac{\beta x}{s} \right) f_X(x) f_S(s - x), \qquad (11)$$

para $s \in \mathscr{S} = \{1, 2, \ldots\}$, con valor inicial de la recursión dado por

$$f_S(0) = P_S(0) = P_N\left(P_X(0)\right), \qquad (12)$$

donde P_N y P_X son las funciones generadoras de probabilidad de N y X respectivamente.

Note que la fórmula de recursión anterior requiere suponer que el monto individual de siniestros tiene soporte discreto no negativo. Cuando el monto de los siniestros individuales se supone continuo, tal recursión no es válida, pero puede realizarse una sustitución ad-hoc de la suma en X como una integral para definir la convolución de X parcial con S parcial.