Algoritmo de Panjer

Autor: Lic. Emma M. Mancilla Flores

Resumen

En el presente artículo, se realiza la revisión así como el análisis de la aproximación numérica de las algunas distribuciones compuestas, conocidas como las distribuciones de la suma de variables aleatorias (variables compuestas), teniendo en cuenta el grado de dificultad para la respectiva evaluación de manera exacta. Para ello se toma de referencia la propuesta realizada por Harry Panjer (1981)[4], mediante el denominado Algoritmo de Panjer, para lo cual previamente se realiza la revisión de la definición, así como las características de las distribuciones clase (a, b). Y se concluye con la verificación de las expresiones simplificadas de recursividad, al aplicar el algoritmo, considerando el caso particular de las distribuciones Binomial y Poisson.

1. Introducción.

En estas últimas décadas, cobró gran importancia las distribuciones de variables aleatorias discretas, debido a la aplicación especialmente en cuanto a la Teoría del Riesgo y la naturalidad con la que se adecúan a los diferentes fenómenos a modelarse; especialmente las distribuciones Binomial, Poisson, Binomial Negativa y las derivadas de ellas. Por eso, es que se realiza una breve introducción acerca de las variables aleatorias discretas.

Definición. 1. Sea N una v.a.d., si existe un subconjunto numerable de reales

$$A = \{n_0, n_1, ...\},\ A \subseteq R \quad \text{tal que } P(N \in A) = 1$$

Se define la función de probabilidad:

$$p : A \to [0,1]$$

tal que: $p(n_k) = P(N = n_k)$

Si en particular, se consideran variables aleatorias discretas no negativas, con $A \subseteq \mathbb{N}_0$, para algún

$h > 0$, es decir:

$$n_k = hk; \qquad (1)$$

donde $p_k = P(N = n_k)$; se dice que es una *distribución aritmética*[1].

2. Distribuciones clase (a, b).

Definición. 2. Una distribución de frecuencias $\{p_k\}$ es un miembro de la

clase (a, b; 0) si existen constantes a, b tal que:

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k}; \quad k = 1, 2, 3, \ldots \qquad (2)$$

Esta clase puede ser también ampliada, mediante la siguiente definición.

Definición. 3. Una distribución de frecuencias $\{p_k\}$ es un miembro de la

clase (a, b; 1) si existen constantes a y b tal que:

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k}; \quad k = 1, 2, 3, \ldots \qquad (3)$$

Observación 1.

Si la recursión inicia en k =1, se tiene la clase (a, b; 0), y si inicia en k =2, se tiene la clase (a, b; 1).

1. Todo miembro de la clase (a, b; 0), lo es también de la clase (a, b; 1).
2. En general las distribuciones clase (a, b), al especificarse las constantes a, b y el valor inicial de p_0, podrán ser definidas de forma recursiva.

1 - Conocida también como látice.

Teorema.1. Las únicas distribuciones no degeneradas cuyas funciones de probabilidad verifican la fórmula recursiva (2) son:

- *P*oisson(λ)
- *B*inomial *N*egativa(r, p)
- *B*inomial(n, p).

Una demostración bastante práctica del teorema se puede encontrar en [2].

Nota 1. En las distribuciones mencionadas en el teorema, las constantes a, b y el valor inicial de p_0 adecuadas, son las siguientes:

i) $N \sim Binomial(n, p)$

$$a = \frac{-p}{1-p}; \quad b = \frac{(n+1)p}{1-p} \qquad y$$

$$p_0 = (1-p)^n$$

ii) $N \sim Poisson(\lambda)$

$$a = 0; \quad b = \lambda \qquad y$$

$$p_0 = e^{-\lambda}$$

iii) $N \sim Binomial \quad Negativa(r, p)$

$$a = 1-p; \quad b = (r-1)(1-p) \qquad y$$

$$p_0 = p^r$$

De acuerdo a las condiciones restrictivas del teorema, existen muchas distribuciones, en las cuales no es posible emplear la forma recursiva para definirlas. Un proceso muy útil para resolver este problema es el *truncamiento de distribuciones* el cual permite básicamente redefinir la secuencia de probabilidades de tal manera que éstas se adecúen de manera más natural al problema que se desea modelar dada una condición inicial; una de ellas son las denominadas *distribuciones Cero-Truncadas*.

Definición 4. Una ***distribución Cero-Truncada*** con función de probabilidad p_k^T surge de la secuencia $p_k, k \in N_0$ cuando se asigna a la probabilidad p_0 el valor cero, de esta forma la nueva secuencia de probabilidades definida a partir de la anterior está dada por:

$$p_k^T = \frac{p_k}{1-p_0}; \qquad k = 1, 2, 3, \ldots$$

Así, la clase (a, b) de distribuciones en las cuales se pueda aplicar las definiciones recursivas se ampliará.

4. Variables Aleatorias Compuestas.

Definición 5. Una ***variable aleatoria compuesta*** $S = S_N$, se define como la suma de variables aleatorias X_j:

$$S_0 = 0; \qquad S_N = X_1 + X_2 + \ldots + X_N; \qquad N \geq 1 \qquad (4)$$

las cuales son i.i.d. con función de distribución común $F_X(x) = P(X_j = x)$ y sea la variable de conteo N independiente de las variables aleatorias Xj.

La función generadora de probabilidad de S será:

$$P_s(z) = P_N\,[P_X\,(z)] \qquad (5)$$

siempre que P_X exista.

Además, la función de distribución de la variable aleatoria compuesta S es:

$$F_S(x) = P\,(S \leq x) = \sum_{n=0}^{\infty} p(x)\, F_X^{*n}(x) \qquad (6)$$

donde

$$F_X^{*n}(x) = P\,(X_1 + X_2 + \ldots + X_n \leq x)$$

es la n-ésima convolución de F_X.

Finalmente, la función de probabilidad de la variable aleatoria compuesta S es $\{g_k\}$, donde $g_k = P(S \leq k)$.

Nota 2. Cabe notar que, para la variable aleatoria definida en (4); la función de probabilidad $\{p_k\}$ de N se denomina *distribución primaria*:

$$p_k = P(N = k)$$

y $\{f_k\}$ de X se denomina *distribución secundaria*:

$$f_k = P(X = k)$$

en el caso de X ser una variable discreta o puede ser discretizada.[2]

4. Algoritmo de Panjer.

Un problema que se nos presenta, es que al aplicar (6) en busca de la función de distribución de una variable aleatoria compuesta, el trabajo sería bastante tedioso al depender de múltiples convoluciones y además no se puede asegurar que exista una solución general a este problema; para facilitar tan penoso trabajo existen diferentes métodos; entre ellos la diferencia es al realizar consideraciones acerca de las variables aleatorias N y X_j. Uno de los más representativos y aplicados en el caso de las variables aleatorias compuestas es el conocido *Algoritmo de Panjer.*

Teorema 2. (*Algoritmo de Panjer*.)

Si la distribución primaria de la v.a. compuesta es miembro de la clase (a, b; 1), entonces:

$$g_0 = P_N(f_0)$$

$$g_k = \frac{[p_1 - (a+b)p_0]f_k}{1-af_0} + \frac{\sum_{j=1}^{k}\left(a + \frac{bj}{k}\right) f_j g_{k-j}}{1-af_0}; \quad (7)$$

para k =1, 2, ... y f_k denota la distribución secundaria.

Como todo miembro de la clase (a, b; 0), lo es también de la clase (a, b; 1), se tiene el siguiente resultado.

Corolario 1. (*Algoritmo de Panjer*.) Si la distribución primaria de la v.a. compuesta (4) es miembro de la clase (a, b; 0), entonces:

$$g_0 = P_N(f_0)$$

$$g_k = \frac{1}{1-af_0} \sum_{j=1}^{k}\left(a + \frac{bj}{k}\right) f_j g_{k-j}; \quad (8)$$

para k =1, 2, ...

Las demostraciones se encuentran en [4,5].

5. Algunos resultados útiles.

En el caso particular de las distribuciones Binomial y Poisson, son miembros de la clase (a, b; 0), por tanto al aplicar el corolario se tiene:

5.1. Si la distribución primaria es Binomial.

Si en (4), N ~ Binomial (n, p) y al considerar las constantes necesarias:

$$a = \frac{-p}{1-p}; \quad b = \frac{(n+1)p}{1-p} \Rightarrow b = -(n+1)a \qquad y$$

$$p_0 = (1-p)^n$$

Reemplazamos en (8) para $k \in \mathbb{N}$:

2 - Existen diferentes métodos para poder discretizar una variable aleatoria, la cual dependerá de su aplicación.

$$\begin{aligned}
g_k &= \frac{1}{1\text{-af}_0}\sum_{j=1}^{k}\left(a+\frac{bj}{k}\right)f_j g_{k-j}\\
&= \frac{1}{1\text{-af}_0}\sum_{j=1}^{k}\left(a+\frac{-(n+1)j}{k}\right)f_j g_{k-j}\\
&= \frac{-a}{1\text{-af}_0}\sum_{j=1}^{k}\left(\frac{nj+j-k}{k}\right)f_j g_{k-j}\\
&= \frac{-a}{1\text{-af}_0}\sum_{j=1}^{k}\left((n+1)\frac{j}{k}-1\right)f_j g_{k-j}\\
&= \frac{-\frac{p}{1-p}}{1+\left(\frac{\text{p}}{\text{1-p}}\right)\text{f}_0}\sum_{j=1}^{k}\left((n+1)\frac{j}{k}-1\right)f_j g_{k-j}
\end{aligned}$$

Así, la forma recursiva para calcular la función de probabilidad cuando la distribución primaria es Binomial (n, p), está dada por:

$$\begin{aligned}
g_0 &= P_N(f_0)\\
g_k &= \frac{p}{1\text{ - p}}\sum_{j=1}^{k}\left((n+1)\frac{j}{k}-1\right)f_j g_{k-j}
\end{aligned}$$

Para, $k = 1, 2, 3, \ldots \Diamond$

5.2. Si la distribución primaria es Poisson.

En (4) al considerar que N ~ Poisson(λ)[3],

cuyas constantes son:

$$a=0;\quad b=\lambda \qquad y$$

$$p_0 = e^{-\lambda}$$

Al reemplazar en (8) para $k \in \mathbb{N}$, de manera inmediata, se puede obtener:

$$g_k = \sum_{j=1}^{k}\left(\frac{\lambda j}{k}\right)f_j g_{k-j}$$

Por lo tanto, la forma recursiva para g_k, está dada por:

$$\begin{aligned}
g_0 &= P_N(f_0)\\
g_k &= \sum_{j=1}^{k}\left(\frac{\lambda j}{k}\right)f_j g_{k-j}
\end{aligned}$$

Para, $k = 1, 2, 3, \ldots \Diamond$

Finalmente, se puede apreciar en ambos resultados la dependencia de la distribución secundaria.

6. Referencias.

[1] Mood A.; Graybill,F.;Boes,D.; *Introduction to the Theory of Statistics*, McGraw-Hill.

[2] Escalante,C.; *Distribuciones clase (a, b) y algoritmo de Panjer*, REDALYC, Diciembre, Año/Vol.XIV,No.002.

[3] Cruz,D.; Másmela,L.; *Poisson-Pascal Generalized Distribution using the Panjer's Algorithm;* Comunicaciones en Estadística, Junio 2010, Vol.3,No.1

[4] Panjer,H. ; *Recursive evaluation of a Family of Compound Distributions*; Astin Bulletin 12 (1982) 22-26.

[5] Panjer,H.; *The Aggregate Claims Distribution and Stop-Loss Reinsurance*; Society Acturies, XXXII.

3 - Este resultado fue presentado la primera vez por R.M.Adelson (1966).