Nuestro objetivo ahora es encontrar probabilidades exactas de S. Para eso, haremos supuesto que puedan facilitar los cálculos.

Bajo ciertos supuestos, existe una metodología recursiva muy conocida para calcular probabilidades exactas de S. Nosotros nos enfocaremos principalmente en su aplicación

Formula de Pril

Presentaremos a continuación la formula de Pril. Este resultado fue mostrado por Nelson De Pril en 1986 y proporciona una expresión exacta, aunque recursiva de la distribución de probabilidad de un riesgo en el modelo individual. La formula es bastante general aunque presupone que las reclamaciones toman los valores en el conjunto $\{1, 2, ...\}$. Este supuesto no es realmente una restricción fuerte pues en la práctica el pago de siniestros se realiza siempre usando alguna unidad monetaria. Para establecer la fórmula De Pril es necesario dividir el portafolio de n asegurados de acuerdo a la tasa de mortalidad y la suma asegurada. Denotaremos por n_{ij} al número de asegurados que tienen probabilidad de reclamación q_j y monto de reclamación i, en donde i toma valores en $\{1, 2, ..., I\}$ y j en $\{1, 2, ..., J\}$. De esta forma se tiene la tabla 1 en donde la suma de las entradas es n_i, es decir,

$$\boxed{n = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij}}$$

Denotaremos por Y_{ij} el monto real reclamado por un asegurado cuya probabilidad de reclamación es q_j y posible monto reclamado i, es ,decir,

$$\boxed{Y_{ij} = \begin{cases} 0 & \text{con probabilidad} \quad 1 - q_j \\ i & \text{con probabilidad} \quad q_j \end{cases}}$$

		Probabilidad de reclamación					
		q_1	q_2	...	q_j	...	q_J
Monto de la reclamación	1	n_{11}	n_{12}	...	...	...	n_{1J}
	2	n_{21}	n_{22}	...	...	...	n_{2J}
	...	...	...	...	...	...	...
	i	...	...	...	n_{11}	...	n_{iJ}
	...	...	...	...	...	...	...
	I	...	...	...	n_{Ij}	...	n_{IJ}

Cuadro 1: Probabilidad de reclamación - Monto de la reclamación

Teorema: Formula De Pril (i)

Sea n_{ij} el número de pólizas cuyos asegurados tienen tasa de mortalidad q_j y suma asegurada i. Suponga que $j = 1, 2, ..., J$ e $i = 1, 2, ..., I$. Entonces las probabilidades $g_x = \mathbb{P}(S = x)$ están dadas por:

$$
\begin{aligned}
g_x &= \frac{1}{x} \sum_{i=1}^{x \wedge I} \sum_{k=1}^{\lfloor x/i \rfloor} g_{x-ik}\, h(i,k), \quad \text{para } x \geq 1 \\
g_0 &= \prod_{i=1}^{I} \prod_{j=1}^{J} (1 - q_j)^{n_{ij}}
\end{aligned}
$$

En donde

$$
h(i,k) = i(-1)^{k-1} \sum_{j=1}^{J} n_{ij} \left(\frac{q_j}{1 - q_j} \right)^k
$$

Nota: $x \wedge I \doteq \text{mín}\{x, I\}$

Demostración

La función generadora de probabilidad del monto reclamado Y_{ij} por un asegurado con probabilidad de reclamación q_j, y monto reclamado i es:

$$
\mathbb{E}(t^{Y_{ij}}) = (1 - q_j) + q_j t^i
$$

Por lo tanto, usando la hipótesis de independencia, la función generadora de probabilidad de la cartera completa es:

$$
\begin{aligned}
G(t) &= \mathbb{E}(t^S) \\
&= \sum_{r=0}^{\infty} t^r g_r \\
&= \prod_{i=1}^{I} \prod_{j=1}^{J} (1 - q_j + q_j t^i)^{n_{ij}}
\end{aligned}
$$

En donde $q_r = \mathbb{P}(S = r)$. Tomando logaritmo y después derivando.

$$\begin{aligned}
\ln(G(t)) &= \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} \quad \ln\big(1 - q_j + q_j t^i\big) \\
\frac{d}{dt}\ln(G(t)) &= \frac{G'(t)}{G(t)} \\
&= \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} \frac{i q_j t^{i-1}}{1 - q_j + q_j t^i}
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
tG'(t) &= G(t) \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} \frac{i q_j t^{i}}{1 - q_j + q_j t^i} \\
&= G(t) \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} i \frac{q_j t^i}{1 - q_j} \left(1 + \frac{q_j t^i}{1 - q_j}\right)^{-1} \\
&= G(t) \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} i \frac{q_j t^i}{1 - q_j} \sum_{k=1}^{\infty} (-1)^{k-1} \left(\frac{q_j t^i}{1 - q_j}\right)^{k-1}
\end{aligned}$$

En donde hemos usado la expansión $(1-x)^{-1} = \sum_{k=0}^{\infty} x^k$, válida para $|x| < 1$. Por lo tanto, para valores suficientemente pequeños de t.

$$tG'(t) = G(t) \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} i \sum_{k=1}^{\infty} (-1)^{k-1} \left(\frac{q_j}{1 - q_j}\right)^{k} t^{ik}.$$

Defina ahora la función

$$h(i,k) = i(-1)^{k-1} \sum_{j=1}^{J} n_{ij} \left(\frac{q_j}{1 - q_j}\right)^{k}$$

La doble suma respecto de los índices j y k es absolutamente convergente en cualquiera de los dos órdenes que se efectúen en estas sumas y el resultado es el mismo. Por lo tanto es válido el intercambio en el orden de las sumas y la expresión anterior puede escribirse como sigue:

$$tG'(t) \;=\; G(t)\sum_{i=1}^{I}\sum_{k=1}^{\infty} t^{ik}h(i,k)$$

Sustituyendo de las expresiones $G'(t)$ y $G(t)$ en sus correspondientes series de potencias se obtiene:

$$\sum_{r=1}^{\infty} rt^r g_r \;=\; \sum_{r=0}^{\infty} t^r g_r \sum_{i=1}^{I}\sum_{k=1}^{\infty} t^{ik}h(i,k)$$

Para $x \geq 1$, el coeficiente de t^x en el lado izquierdo es xg_x, mientras que en el lado derecho es la suma de los términos $g_{x-ik}h(i,k)$ para aquellos valores de i y k tales que $I \leq ik \leq x$. Se pueden establecer primero los posibles valores para i de la siguiente forma $i = 1, ..., x \wedge I$, y por lo tanto los valores para k son $k = 1, ..., [x/i]$, en donde $x \wedge I$ es el valor mínimo entre x e I, y $\lfloor x/i \rfloor$ es la parte entera del cociente x/i. Igualando estos coeficientes se tiene que:

$$xg_x \;=\; \sum_{i=1}^{x\wedge I}\sum_{k=1}^{\lfloor x/i \rfloor} g_{x-ik}h(i,k)$$

De esta forma se llega a la siguiente expresión, para $x \geq 1$

$$g_x \;=\; \frac{1}{x}\sum_{i=1}^{x\wedge I}\sum_{k=1}^{\lfloor x/i \rfloor} g_{x-ik}h(i,k)$$

Por otro lado, como $S = 0$ sólo cuando ningún asegurado efectúa ninguna reclamación, para $x = 0$ se tiene que:

$$g_0 \;=\; \prod_{i=1}^{I}\prod_{j=1}^{J}(1-q_j)^{n_{ij}}$$

Notemos que lo que hace este modelo, es agrupar a los asegurados por monto de reclamación (suma asegurada) i y probabilidad de reclamación q_j (grupo (i,j)). Observe que la v.a Y_{ij} indica el monto que reclamó un individuo del grupo (i,j), sin embargo hay n_{ij} de ellos; ergo, necesitamos denotar a Y_{ijk} como el monto real que puede reclamar el k-ésimo asegurado del grupo (i,j) para poder ser más específicos. De esta manera, podemos ver al riesgo S como:

$$
\begin{aligned}
S &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} Y_{ijk} \\
&= \underbrace{\sum_{k=1}^{n_{11}} Y_{11k}}_{\text{Grupo (1,1)}} + \underbrace{\sum_{k=1}^{n_{21}} Y_{21k}}_{\text{Grupo (2,1)}} + \ldots + \underbrace{\sum_{k=1}^{n_{I1}} Y_{I1k}}_{\text{Grupo (I,1)}} + \underbrace{\sum_{k=1}^{n_{12}} Y_{12k}}_{\text{Grupo (1,2)}} + \ldots + \underbrace{\sum_{k=1}^{n_{I2}} Y_{I2k}}_{\text{Grupo (I,2)}} + \ldots + \underbrace{\sum_{k=1}^{n_{IJ}} Y_{IJk}}_{\text{Grupo (I,J)}}
\end{aligned}
$$

No olvidemos que estamos hablando del modelo individual. Observemos que $Y_{ij} = B_j(i)$ donde $B_j \sim Ber(q_j)$. Ahora vamos a desagrupar a los individuos; lo que tenemos es un portafolio con n individuos donde algunos tienen características diferentes entre si.

Sea $X_t \doteq D_t C_t$ con $D_t \sim Ber(\gamma_t)$ y C_t una constante que indica el monto que reclama el t-ésimo asegurado del portafolio con n-individuos. Entonces existen t y k tales que $X_t = Y_{ijk}$ en cuyo caso sucede que $\gamma_t = q_j$ y $C_t = i$.

Observe que esto no es más que renombrar variables por grupo y sin grupos. Entonces podríamos escribir también al riesgo S como:

$$
\boxed{S = \sum_{t=1}^{n} C_t D_t = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} Y_{ijk}}
$$

Con lo que podemos notar que si seguimos hablando del modelo individual, solo que con ciertas etiquetas. En consecuencia cada Y_{ijk} es independiente una de otra (esto puesto que los individuos son independientes entre si) y son idénticamente distribuidos por grupo(i, j).

Ejemplo

Tomando un portafolio del estilo:

SA	q_1	q_2
1	2	1
2	3	0

En el fondo tenemos el riesgo:

$$\begin{aligned} S &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} Y_{ijk} \\ &= \sum_{k=1}^{2} Y_{11k} + \sum_{k=1}^{1} Y_{12k} + \sum_{k=1}^{3} Y_{21k} + 0 \\ &= (Y_{111} + Y_{112}) + (Y_{121}) + (Y_{211} + Y_{212} + Y_{213}) \end{aligned}$$

Definiendo:

$$\begin{array}{llllll} -D_1 \sim Ber(q_1) & \mathbf{y} & C_1 = 1 & -D_4 \sim Ber(q_2) & \mathbf{y} & C_4 = 1 \\ -D_2 \sim Ber(q_1) & \mathbf{y} & C_2 = 1 & -D_5 \sim Ber(q_2) & \mathbf{y} & C_5 = 1 \\ -D_3 \sim Ber(q_1) & \mathbf{y} & C_3 = 2 & -D_6 \sim Ber(q_2) & \mathbf{y} & C_6 = 1 \end{array}$$

$$\begin{aligned} S &= (D_1C_1 + D_2C_2) + (D_3C_3) + (D_4C_4 + D_5C_5 + D_6C_6) \\ &= \sum_{t=1}^{6} D_tC_t \end{aligned}$$

Otra observación importante es que S está acotada. Como $D_t \in \{0,1\}$ y $C_t > 0$, entonces en el mejor de los escenarios $S = 0$ que ocurre cuando nadie reclama. Pero si todos reclaman:

$$\begin{aligned} S &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} Y_{ijk} \\ &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} i \quad \textbf{Cuando reclaman vale i la suma asegurada} \\ &= \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij}(i) \\ &= \sum_{i=1}^{I} i \sum_{j=1}^{J} n_{ij} \end{aligned}$$

Que teniendo el cuadro:

	q_1	q_2	...	q_j	...	q_J
1	n_{11}	n_{12}	...	...	...	n_{1J}
2	n_{21}	n_{22}	...	...	...	n_{2J}
...	...	...	...	...	...	...
i	...	...	...	n_{i1}	...	n_{iJ}
...	...	...	...	...	...	...
I	...	...	...	n_{Ij}	...	n_{IJ}

Se traduce a simplemente tomar:

$S_{máx} =$

	q_1	q_2	...	q_j	...	q_J
+ (1) x (	$n_{11}+$	$n_{12}+$	...+	...+	...+	n_{1J})
+ (2) x (	$n_{21}+$	$n_{22}+$	...+	...+	...+	n_{2J})
+ (...)x (	... +	...+	...+	...+	...+	...)
+ (i) x (	...+	...+	...+	$n_{i1}+$	...+	n_{iJ})
+ (...)x (	...+	...+	...+	...+	...+	...)
+ (I) x (	...+	...+	...+	$n_{Ij}+$	...+	n_{IJ})

Viendo esto, haremos el cálculo de la esperanza del riesgo S.

$$\mathbb{E}[S] \doteq \mathbb{E}\left[\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} Y_{ijk}\right] = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} \mathbb{E}[Y_{ijk}]$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} iq_j \begin{cases} \mathbb{E}[Y_{ijk}] \\ = 0 \cdot (1 - q_i) + iq_j \\ = iq_j \end{cases}$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J} iq_j n_{ij} = \sum_{i=1}^{I} i \sum_{j=1}^{J} q_j n_{ij}$$

Que es muy similar al cálculo anterior solo que ahora debemos multiplicar cada n_{ij} por su correspondiente q_j

$$
\begin{aligned}
\mathbb{E}[S] = &(1) \times [q_1 n_{11} + q_2 n_{12} + ... + ... + ... + q_J n_{1J}] \\
&+(2) \times [q_1 n_{21} + q_2 n_{22} + ... + ... + ... + q_J n_{2J}] \\
&+(...) \times [... + ... + ... + ... + ... + ...] \\
&+(i) \times [... + ... + ... + q_j n_{ij} + ... + ...] \\
&+(...) \times [... + ... + ... + ... + ... + ...] \\
&+(I) \times [... + ... + ... + q_j n_{Ij} + ... + q_J n_{IJ}]
\end{aligned}
$$

Equivalentemente es multiplicar cada número dentro del cuadro, multiplicarlo por su (SA, q_j) y sumar todo.

El resultado anterior se puede deducir de la fórmula general del modelo individual simplemente agrupando por (i, j). Haremos el cálculo de la varianza de esta manera para ejemplificarlo. Con base en los resultados obtenidos para el modelo individual:

$$
\begin{aligned}
Var(S) &= \sum_{t=1}^{n} [\gamma_t Var(C_t) + \gamma_t (1-\gamma_t)\mathbb{E}^2[C_t]] \\
&= \sum_{t=1}^{n} \gamma_t (1-\gamma_t) C_t^2 \begin{cases} \textit{Como } C_t \textit{ es constante} \\ \\ Var(C_t) = 0 \ \ y \ \ \mathbb{E}[C_t] = C_t \end{cases}
\end{aligned}
$$

La suma anterior la podemos separar en los grupos (i, j) pues $n = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij}$, al separarla, tendremos una suma de sumas de grupos y al mismo tiempo γ_t coincidirá con q_j y C_t con i para cada grupo (i, j):

$$
= \underbrace{\sum_{t=1}^{n_{11}} q_1(1-q_1)(1)^2}_{\text{Grupo(1,1)}} + \underbrace{\sum_{t=1}^{n_{12}} q_2(1-q_2)(1)^2}_{\text{Grupo(1,2)}} + ... + \underbrace{\sum_{t=1}^{n_{1J}} q_J(1-q_J)(1)^2}_{\text{Grupo(1,J)}}
$$

$$
+ \underbrace{\sum_{t=1}^{n_{21}} q_1(1-q_1)(2)^2}_{\text{Grupo(2,1)}} + \underbrace{\sum_{t=1}^{n_{22}} q_2(1-q_2)(2)^2}_{\text{Grupo(2,2)}} + ... + \underbrace{\sum_{t=1}^{n_{2J}} q_J(1-q_J)(2)^2}_{\text{Grupo(2,J)}}
$$

$$
+ ... + ... + ...
$$

$$+\underbrace{\sum_{t=1}^{n_{I1}} q_1(1-q_1)(I)^2}_{\text{Grupo(I,1)}} + \underbrace{\sum_{t=1}^{n_{I2}} q_2(1-q_2)(I)^2}_{\text{Grupo(I,2)}} + ... + \underbrace{\sum_{t=1}^{n_{IJ}} q_J(1-q_J)(I)^2}_{\text{Grupo(I,J)}}$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n_{ij}} q_j(1-q_j)i^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} q_j(1-q_j)i^2 n_{ij}$$

$$= \sum_{i=1}^{I} i^2 \sum_{j=1}^{J} q_j(1-q_j)n_{ij}$$

Que análogo al último cálculo no es más que, tomando como referencia la tabla con frecuencias:

$$\begin{aligned} Var(S) = & (1)^2 \times [(1-q_1)q_1 n_{11} + (1-q_2)q_2 n_{12} + ... + ... + ... + (1-q_J)q_J n_{1J}] \\ & +(2)^2 \times [(1-q_1)q_1 n_{21} + (1-q_2)q_2 n_{22} + ... + ... + ... + (1-q_J)q_J n_{2J}] \\ & +(...) \times [... + ... + ... + ... + ... + ...] \\ & +(i)^2 \times [... + ... + ... + (1-q_j)q_j n_{ij} + ... + ...] \\ & +(...) \times [... + ... + ... + ... + ... + ...] \\ & +(I)^2 \times [... + ... + ... + (1-q_j)q_j n_{Ij} + ... + (1-q_J)q_J n_{IJ}] \end{aligned}$$

Que es lo mismo que la esperanza pero multiplicando $(1-q_j)q_j$ y no solo q_j para cada j Con lo anterior ya tenemos maneras para calcular probabilidades de S, así como varianza y esperanza. Los ejemplos los veremos en R.

Para un mejor entendimiento de la fórmula recursiva de DePril [i] escribiremos a continuación de manera explícita los primeros términos de este desarrollo

$$\begin{aligned} g_0 &= \prod_{i=1}^{I} \prod_{j=1}^{J} (1-q_j)^{n_{ij}} \\ g_1 &= g_0 h(1,1) \\ g_2 &= \frac{1}{2}\{g_0[h(1,2) + h(2,1)] + g_1 h(1,1)\} \\ g_3 &= \frac{1}{3}\{g_0[h(1,3) + h(3,1)] + g_1[h(1,2) + h(2,1)] + g_2 h(1,1)\} \\ & ... \end{aligned}$$

Ejemplo Para los datos que se muestran en la tabla de la Figura 1.4 en donde se tienen 48 pólizas de seguros con las probabilidades de reclamación y los montos de reclamaciones indicados, la correspondiente función de densidad para este riesgo es la que se muestra en la Figura 1.5 En el apéndice A se encuentra el código en R de una implementación de la fórmula de De Pr
Debe tenerse cuidado en la implementación numérica de esta fórmula
pues dado su caracter recursivo y que algunas de las probabilidades involucradas pueden ser muy pequeñas, pueden generarse resultados incorrectos debido al inevitable redondeo de cifras en una computadora

Probabilidades de reclamación

Monto de la reclamación $i \backslash q$	0.03	0.04	0.05
1	1	3	1
2	3	5	4
3	5	3	4
4	2	2	6
5	2	3	4

Figura 1.4: Cartera de 48 pólizas individuales.

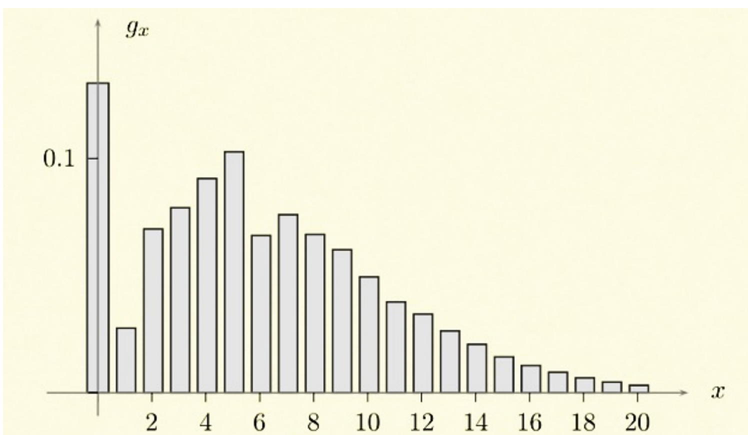


Figura 1.5

La fórmula que hemos denominado de De Pril [i] y que se encuentra expresada en el contexto de un portafolio de asegurados individuales puede escribirse como un resultado teórico de la teoría de probabilidad. Este es el contenido de la siguiente proposición. La fórmula tiene una expresión más simple y la demostración sigue los mismo lineamientos que la que hemos presentado, sin embargo, escribiremos nuevamente los detalles de la prueba en esta versión simplificada

Proposición (Fórmula de De Pril [ii])

Sean $X_1, X_2, ..., X_n$ v.a.i.i.d. con valores en el conjunto $\{0, 1, 2, ...\}$. Para cada entero $j \geq 0$, defina la probabilidad $f_j = P(X = j)$, y suponga $f_0 \neq 0$. Sea $S = X_1 + ... + X_n$. Entonces las probabilidades $g_x = P(S = x)$ se pueden calcular recursivamente mediante la siguiente fórmula

$$\begin{aligned} g_0 &= (f_0)^n, \\ g_x &= \frac{1}{f_0} \sum_{j=1}^{x} \left[\frac{j(n+1)}{x} - 1 \right] f_j g_{x-j}, \;\; para \;\; x \geq 1. \end{aligned}$$

Demostración. Primeramente observemos que el evento $(S = 0)$ ocurre si y sólo si todos los sumandos de S son cero, de modo que por independencia, $g_0 = (f_0)^n$. Ahora veamos la forma de obtener la fórmula recursiva. Sean $P_X(t)$ y $P_S(t)$ las funciones generadoras de probabilidad de las variables discretas X y S respectivamente, es decir,

$$P_X(t) = E(t^X) = \textstyle\sum_{k=0}^{\infty} t^k f_k,$$

$$P_S(t) = E(t^S) = \textstyle\sum_{k=0}^{\infty} t^k g_k.$$

Por independencia e idéntica distribución, $P_S(t) = [P_X(t)]^n$. Derivando respecto de t,

$$P_S'(t) = n[P_X(t)]^{n-1} P_X'(t)$$

Multiplicando ambos lados por $tP_X(t)$,

$$P_X(t) t P_S'(t) = n P_S(t) t P_X'(t),$$

que en términos de sumas se escribe como sigue

$$\textstyle\sum_{j=0}^{\infty} t^j f_j \sum_{k=1}^{\infty} k t^k g_k = n \sum_{k=0}^{\infty} t^k g_k \sum_{j=1}^{\infty} j t^j f_j.$$

El siguiente paso es identificar el coeficiente del término t^x en cada lado de la ecuación, para $x \geq 1$. Por ejemplo, para el lado izquierdo el coeficiente es el término $f_j k g_k$ para todos aquellos valores de $j \geq 0$ y $k \geq 1$ tales que $j + k = x$. Esta doble suma puede escribirse como $\sum_{j=0}^{x-1} f_j (x-j) g_{x-j}$. De manera similar se encuentra el coeficiente del lado derecho. Igualando estos coeficientes se llega a la identidad

$$\sum_{j=0}^{x-1} (x-j) f_j g_{x-j} = n \sum_{j=1}^{x} j f_j g_{x-l}.$$

Separando el primer sumando del lado izquierdo y añadiendo en esa misma suma el término correspondiente a $j = x$, es es cero, se obtiene

$$x f_0 g_x + \sum_{j=1}^{x} (x-j) f_j g_{x-j} = n \sum_{j=1}^{x} j f_j g_{x-l}.$$

Finalmente se despeja el término g_x para llegar a la fórmula anunciada,

$$g_x = \frac{1}{f_0} \sum_{j=1}^{x} \left[\frac{j(n+1)}{x} - 1 \right] f_j g_{x-j},\ x \geq 1.$$

Los primeros términos de la fórmula de De Pril [ii] se muestran a continuación.

$$\begin{aligned}
g_0 &= (f_0)^n, \\
g_1 &= \frac{1}{f_0}(n f_1 g_0) = \binom{n}{1} f_1 (f_0)^{n-1}, \\
g_2 &= \frac{1}{f_0}\left(\frac{n-1}{2} f_1 g_1 + n f_2 g_0 \right) = \binom{n}{2} (f_1)^2 (f_0)^{n-2} + \binom{n}{1} f_2 (f_0)^{n-1}, \\
g_3 &= \frac{1}{f_0}\left(\frac{n-1}{3} f_1 g_2 + \frac{2n-1}{3} f_2 g_1 + n f_3 g_0 \right) \\
&= \binom{n}{3} (f_1)^3 (f_0)^{n-3} + 2! \binom{n}{2} f_2 f_1 (f_0)^{n-2} + \binom{n}{1} f_3 (f_0)^{n-1}.
\end{aligned}$$

Observe que las expresiones simplificadas tienen una interpretación natural en términos combinatoriales. Por ejemplo, la expresión para g_2 involucra dos situaciones: la primera cuando dos sumandos distintos de S toman cada uno el valor uno y el resto toma el valor cero, y la segunda situación cuando uno de los sumandos toma el valor dos y el resto es cero. Los coeficientes binomiales dan cuenta de las distintas formas en las que se pueden presentar estos arreglos.

Ejemplo Sean $X_1.X_2, X_3$ variables aleatorias independientes con idéntica distribución dada por la tabla que aparece abajo y cuya gráfica se muestra en la Figura 1.6(a).
Usando la fórmula de De Pril [ii] encontraremos la distribución de $S = X_1 + X_2 + X_3$.

j	0	1	2
f_j	0.5	0.2	0.3

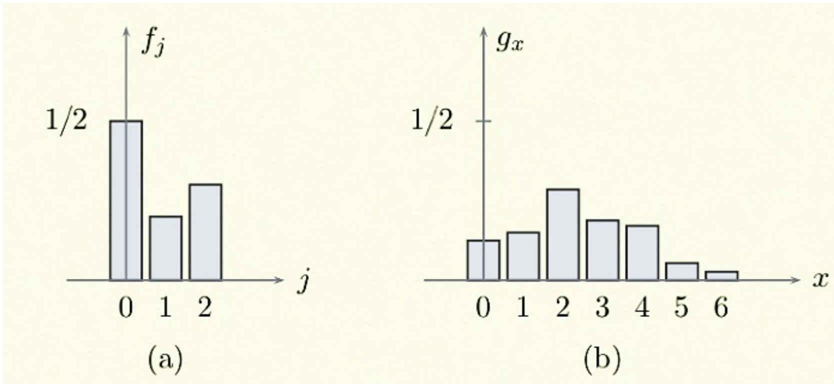


Figura 1.6

Observe que la variable suma puede tomar cualquiera de los valores $0, 1, ..., 6$. Usando la misma notación que en la fórmula de De Pril se muestran a continuación los cálculos para encontrar la función de probabilidad de S y la gráfica correspondiente aparece en la Figura 1.6(b)

$$
\begin{aligned}
g_0 &= (f_0)^3 = \mathit{0.125}, \\
g_1 &= \frac{1}{f_0}(3f_1g_0) = \mathit{0.15}, \\
g_2 &= \frac{1}{f_0}(f_1g_1 + 3f_2g_0) = \mathit{0.285}, \\
g_3 &= \frac{1}{f_0}\left(\frac{1}{3}f_1g_2 + \frac{8}{3}f_2g_1\right) = \mathit{0.188}, \\
g_4 &= \frac{1}{f_0}(f_2g_2) = \mathit{0.171}, \\
g_5 &= \frac{1}{f_0}\left(-\frac{1}{5}f_1g_4 + \frac{3}{5}f_2g_3\right) = \mathit{0.054}, \\
g_6 &= \frac{1}{f_0}\left(-\frac{1}{3}f_1g_5 + \frac{1}{3}f_2g_4\right) = \mathit{0.027}.
\end{aligned}
$$

Nosotros mostraremos la implementación de esta fórmula en los scripts de R.

Debemos observar la diferencia que existe entre los supuestos que se manejan en la fórmula de De Pril (i) y (ii). En el caso de De Pril (i) el portafolio se asume como:

$$S = \sum_{t=1}^{n} C_t D_t = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n_{ij}} Y_{ijk}$$

Donde las $C_t D_t$ no tienen la misma distribución $\forall t$ necesariamente. Esto se ve más claro cuando vemos el cuadro que nos muestra el número de individuos por grupo (i, j).

	q_1	q_2	$\cdots$	q_j	$\cdots$	q_J
1	n_{11}	n_{12}	$\cdots$	$\cdots$	$\cdots$	n_{1J}
2	n_{21}	n_{22}	$\cdots$	$\cdots$	$\cdots$	n_{2J}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
i	$\cdots$	$\cdots$	$\cdots$	n_{ij}	$\cdots$	n_{iJ}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
I	$\cdots$	$\cdots$	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}

De aquí podemos observar que, sin pérdida de generalidad, uno de los n_{11} miembros del grupo (1,1), tiene una probabilidad $q_1 \neq q_2$ para alguno de los miembros del grupo (1,2) por ejemplo.

Por otro lado, los supuestos que maneja la versión simplificada (De Pril (ii)) son que, asumiendo un portafolio con n-pólizas y tomando $\{X_t\}_{t=1}^{n}$ v.a.i.i.d. entonces el riesgo será:

$$S \sum_{t=1}^{n} X_t$$

Si meditamos un poco la relación, observemos que esta puede modelar uno de los grupos que están en los supuestos de De Pril (i). Si tomamos a $n = n_{ij}$ y $X_t = Y_{ij} = iB_j$ estaríamos modelando todo el grupo (i, j) con esta v.a.

La simplificación en los supuestos de la fórmula de De Pril(ii) nos abre un mundo de posibilidades pues ahora, la suma asegurado no es constante necesariamente y no pide más que el soporte de su severidad (X_t) esté contenido en el conjunto $\mathbb{N} \cup \{0\}$. Ahora el asegurado número t puede reclamar $X_t \in \{0, 1, ...\}$.

Nota: $X_t = 0 \Rightarrow$ no hubo reclamación

Recordemos la siguiente proposición:

PROPOSICIÓN:

Sean X y Y independientes, y sean g y h dos funciones de $\mathbb{R}$ en $\mathbb{R}$, Borel medibles. Entonces las variables aleatorias $g(X)$ y $h(Y)$ también son independientes.

Una de las ventajas más importantes que tiene este modelo es en cuando al cálculo de momentos. Notemos que, definiendo al riesgo S como en las hipótesis de De Pril(ii):

$$\begin{aligned} M_S(t) &= \mathbb{E}\left[e^{St}\right] = \mathbb{E}\left[exp\left(t\sum_{i=1}^{n} X_i\right)\right] = \mathbb{E}\left[exp\left(\sum_{i=1}^{n} tX_i\right)\right] \\ &= \mathbb{E}\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}] \quad \left.\right\} \text{Por la independencia de las } X_i's \text{ y la proposición anterior.} \\ &= \mathbb{E}^n[e^{tX}] = \mu_X^n(t) \quad \left.\right\} \text{Por ser v.a.i.i.d.} \end{aligned}$$

Por tanto, utilizando las hipótesis de De Pril (ii) tenemos que la generadora de momentos de S es:

$$\boxed{\mu_S(t) = \mu_X^n(t)}$$

Con lo cual llegamos a los otros dos resultados importantes:

- $\mu_S'(t) = n\mu_X^{n-1}(t) \cdot \mu_X'(t) \Rightarrow \mu_S'(0) = n(1)\mu_X'(0)$

$$\boxed{\therefore \mathbb{E}[S] = n\mathbb{E}[X]}$$

- Para obtener la varianza es más fácil ver que:

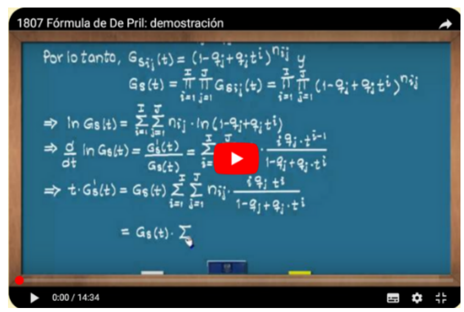
$$\begin{aligned} Var(S) &= Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) \quad \left.\right\} \text{Pues al haber independencia, la covarianza vale cero.} \\ &= nVar(X) \quad \left.\right\} \text{Pues son idénticamente distribuidas} \end{aligned}$$

$$\boxed{\therefore Var(S) = nVar(X)}$$

Los resultados anteriores se hicieron bajo el supuesto de que el riesgo S sigue las hipótesis que se manejan en la fórmula de De Pril (ii). Sin embargo, las demostraciones de estas propiedades no utilizaron el supuesto de que la severidad tenía un soporte discreto. De hecho, estas últimas propiedades son válidas en general para cualquier riesgo $S = \sum_{i=1}^{n} X_i$ donde cada X_i es v.a.i.i.d. no importando su suporte o distribución.
Pueden encontrar los vídeos con parte de la información anterior en los siguientes enlaces:

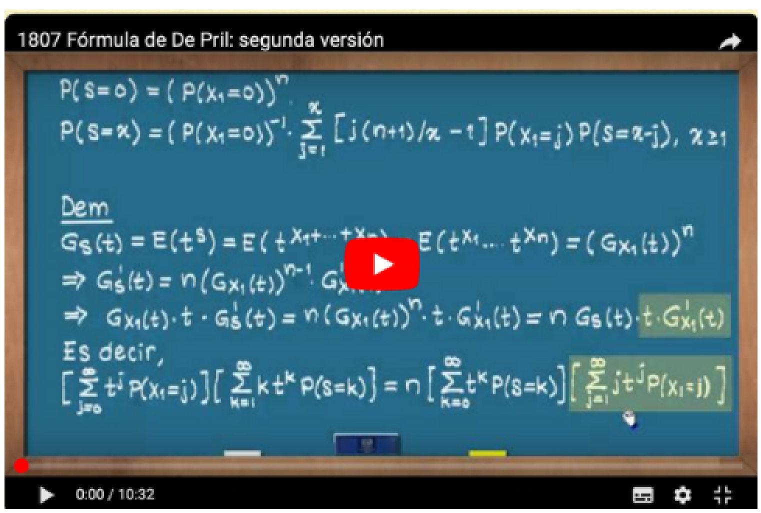
- De Pril (i)

https://youtu.be/bFPe9cODvD4



https://youtu.be/2hX1OB88BlU

- De Pril (ii)



https://youtu.be/2ZiBGxG_gFg

Inspirados en la fórmula de De Pril (ii), recordemos el siguiente ejemplo:
Supongamos que $D_j \sim Ber(0{,}1)$ y que C_j tiene la siguiente f.m.p:

$$\mathbb{P}[C_j = c] = \begin{cases} 0.8 & \text{si} \quad c = 1 \\ 0.2 & \text{si} \quad c = 2 \\ 0 & \text{e.o.c.} \end{cases} \qquad \forall j.$$

Tomando un portafolio con $n = 3$ pólizas bajo el modelo individual, busquemos probabilidades de $S_3 = \sum_{i=1}^{3} D_i C_i$

Resulta que podemos resolver esto de manera rápida ahora utilizando la fórmula de De Pril (ii). Tomando $X_i = D_i C_i$ entonces:

$$\mathbb{P}[X = x] = \begin{cases} 0.9 & \text{si} \quad x = 0 \\ (0.1)(0.8){=}0.08 & \text{si} \quad x = 1 \\ (0.1)(0.2){=}0.02 & \text{si} \quad x = 2 \\ 0 & \text{e.o.c.} \end{cases}$$

Como tenemos $X_i's$ v.a.i.i.d. con soporte contenido en $\mathbb{N} \cup \{0\}$, podemos usar la recursión.
Sin embargo, un ejemplo de desarrollar De Pril (ii) con esto ya se tocó anteriormente.Vamos a retomar el resultado al que habíamos llegado:

$$f_{S_3}(t) = \mathbb{P}[S_3 = t] \approx \begin{cases} 72.9\,\% & \text{si} \quad t = 0 \\ 19.44\,\% & \text{si} \quad t = 1 \\ 6.588\,\% & \text{si} \quad t = 2 \\ 0.9152\,\% & \text{si} \quad t = 3 \\ 0.1464\,\% & \text{si} \quad t = 4 \\ 0.0096\,\% & \text{si} \quad t = 5 \\ 0.0008\,\% & \text{si} \quad t = 6 \\ 0 & \text{e.o.c.} \end{cases}$$

Teniendo en mente lo anterior, vamos a definir ahora un modelo más general primero con un ejemplo. Sea S un riesgo tal que:

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

Donde N es v.a. con f.m.p.:

$$\mathbb{P}[N = n] = \begin{cases} 0.75 & \text{si} \quad n = 0 \\ 0.25 & \text{si} \quad n = 3 \end{cases}$$

Y además, $\{X_i\}_{i=1}^{3}$ son v.a.i.i.d. con f.m.p.:

$$\mathbb{P}[X = x] = \begin{cases} 0.9 & \text{si} \quad x = 0 \\ 0.08 & \text{si} \quad x = 1 \\ 0.02 & \text{si} \quad x = 2 \\ 0 & \text{e.o.c.} \end{cases}$$

Nota: Si $N = 0 \Rightarrow S = 0$

Observe que $S \in \{0, 1, 2, 3, 4, 5, 6\}$,además:

$$
\begin{aligned}
\mathbb{P}[S = 0] &= \overset{1}{\cancel{\mathbb{P}[S = 0|N = 0]}}\mathbb{P}[N = 0] + \mathbb{P}[S = 0|N \neq 0]\mathbb{P}[N \neq 0] \\
&= 0.75 + \mathbb{P}[S = 0|N = 3]\mathbb{P}[N = 3] \\
&= 0.75 + \underbrace{f_{S_3}(0)}_{\substack{\text{probabilidad} \\ \text{antes obtenida}}} (0.25)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}[S = 1] &= \overset{0}{\cancel{\mathbb{P}[S = 1|N = 0]}}\mathbb{P}[N = 0] + \mathbb{P}[S = 1|N \neq 0]\mathbb{P}[N \neq 0] \\
&= 0.25\mathbb{P}[S = 1|N = 3] \\
&= 0.25 f_{S_3}(1)
\end{aligned}
$$

Y haciendo esto sucesivamente, la f.m.p. de S será:

$$
f_S(t) = \mathbb{P}[S = t] = \begin{cases} 0.75{+}0.25 f_{S_3}(0) & \text{si} \quad t = 0 \\ 0.25 f_{S_3}(t) & \text{si} \quad t \in \{1, 2, 3, 4, 5, 6\} \\ 0 & \text{e.o.c.} \end{cases}
$$

Luego observamos los siguientes cálculos:

$$
\begin{aligned}
&- \quad \mathbb{E}[S] = (0.25)\sum_{t=1}^{6} t f_{S_3}(t) = 0.9 \\
&- \quad \mathbb{E}[S^2] = (0.25)\sum_{t=1}^{6} t^2 f_{S_3}(t) \approx 0.1416 \\
&\Rightarrow \quad Var(S) \approx 0.1335
\end{aligned}
$$

Pero además, se puede calcular la esperanza de S sin conocer su función de densidad:

$$
\begin{aligned}
\mathbb{E}[S] &= \mathbb{E}[\mathbb{E}[S|N]] = \underbrace{\overset{0}{\cancel{\mathbb{E}[S|N = 0]}}}_{N=0 \ \Rightarrow \ S=0}\mathbb{P}[N = 0] + \mathbb{E}[S|N = 3]\mathbb{P}[N = 3] \\
&= \mathbb{E}\left[\sum_{i=1}^{N} X_i | N = 3\right]\mathbb{P}[N = 3] = \mathbb{E}\left[\sum_{i=1}^{3} X_i\right]\mathbb{P}[N = 3] \\
&= 3\cdot \mathbb{E}[X]\mathbb{P}[N = 3] = \mathbb{E}[X](3\cdot \mathbb{P}[N = 3] + 0\cdot \mathbb{P}[N = 0]) \\
&= \mathbb{E}[X]\mathbb{E}[N]
\end{aligned}
$$

// Veremos estos cálculos en R. //