TEMAS DE MATEMÁTICAS

Introducción a la teoría del riesgo

Luis Rincón

las prensas de ciencias

Luis Rincón

Introducción a la teoría del riesgo

Facultad de Ciencias, UNAM
2012

Rincón, Luis
Introducción a la teoría del riesgo / Luis Rincón. -- 1ª edición.
-- México : UNAM, Facultad de Ciencias, 2012.
x, 278 páginas : ilustraciones ; 22 cm. -- (Temas de matemáticas)

Incluye índice
Bibliografía: páginas 271-273
ISBN 978-607-02-3773-7

1. Seguros – Matemáticas. 2. Estadística matemática. 3. Probabilidades. I. Universidad Nacional Autónoma de México. Facultad de Ciencias. II. título. III. Serie.

519.50727-scdd21 Biblioteca Nacional de México

Esta obra fue elaborada con apoyo PAPIME: PE103111

Introducción a la teoría del riesgo
1ª edición, 9 de octubre de 2012.

Facultad de Ciencias.
Ciudad Universitaria. Delegación Coyoacán,
C. P. 04510, México, Distrito Federal.
editoriales@ciencias.unam.mx

ISBN: 978-607-02-3773-7

Diseño de portada: Laura Uribe Hernández.

Impreso y hecho en México.

Prólogo

El presente texto contiene material básico para un curso introductorio a ciertos temas de la teoría del riesgo aplicada a seguros. Está basado en el curso semestral de Teoría del Riesgo que el autor ha impartido a estudiantes de último semestre de la carrera de Actuaría en la Facultad de Ciencias de la UNAM. La intención es que este libro sea de alguna ayuda para los numerosos alumnos de las distintas escuelas de actuaría y matemáticas aplicadas de países de habla hispana, y contribuya también a apoyar el trabajo docente de sus profesores. El texto fue escrito en el sistema LaTeX y las ilustraciones fueron elaboradas usando el paquete *pstricks*. El material que se presenta fue compilado enteramente de las fuentes bibliográficas que aparecen al final del texto. La última versión disponible de este texto en su formato digital puede encontrarse en la página web

http://www.matematicas.unam.mx/lars

Agradezco sinceramente los comentarios, correcciones y sugerencias que he recibido por parte de alumnos y profesores para mejorar este material. En particular agradezco a uno de los revisores anónimos, quien dedicó bastante tiempo en la lectura cuidadosa del texto y quien sugirió importantes mejoras para la versión final. Agradezco también el apoyo otorgado por parte de la UNAM DGAPA a través del proyecto PAPIME PE103111, con el cual pudo ser posible el desarrollo y la edición de este trabajo. Toda comunicación dirigida al autor puede enviarse a la cuenta de correo que aparece abajo.

Luis Rincón
Agosto 2012
Ciudad Universitaria UNAM
lars@ciencias.unam.mx

Contenido

Capítulo 1

El modelo individual y el modelo colectivo

En este capítulo se presenta una introducción al esquema del seguro y al concepto general de riesgo. Se presentan además las perspectivas individual y colectiva para modelar el riesgo correspondiente al conjunto de reclamaciones que afronta una compañía aseguradora. Se estudian también algunas propiedades y relaciones entre estas dos perspectivas. En el resto del texto se adopta el modelo colectivo como modelo fundamental.

1.1. Introducción

El término riesgo tiene muchas acepciones dependiendo del área de estudio que se trate y en términos imprecisos puede definirse como la posibilidad de experimentar ciertos eventos de interés y las consecuencias derivadas de dichos eventos. Los riesgos pueden tener un sentido positivo o negativo, pero en general tienen una connotación de pérdida. El objetivo es identificar los riesgos, ponderarlos con base en sus consecuencias, decidir la aceptación o no de los mismos y tomar provecho de su existencia. Por lo tanto no se trata necesariamente de evitarlos o de protegerse contra ellos. El quehacer cotidiano del hombre, ya sea en el ámbito personal o profesional, implica necesariamente y a cada momento hacer frente a ciertos riesgos, y ello puede tener consecuencias no deseadas, pero también puede ofrecer oportunidades. Por ejemplo, el comprar un boleto de lotería conlleva el riesgo de perder el importe pagado por el boleto, pero al mismo tiempo la posibilidad de ganar una gran cantidad de dinero. Cada uno de nosotros pondera estas dos posibilidades de manera distinta y toma una decisión al respecto. Otro ejemplo en donde es evidente la evaluación (a veces inconsciente) de los riesgos es cuando una persona decide viajar en avión. En este caso se considera primordial la rapidez y comodidad del viaje, y se desdeña convenientemente cualquier posibilidad de accidente.

Como hemos mencionado, el término riesgo se define de manera distinta dependiendo de la disciplina de estudio. En ingeniería, por ejemplo, puede definirse el riesgo como el producto de la probabilidad de que un evento no deseable ocurra y el daño esperado debido a la ocurrencia del evento, es decir, Riesgo=(Probabilidad de un accidente)×(Daños como consecuencia del accidente). En finanzas puede definirse el riesgo en términos de la variación o volatilidad de una inversión, o también como la posible pérdida en una inversión; en general, se considera que una inversión en la bolsa de valores (tasa de interés variable) es más riesgosa comparada con una inversión en un banco (tasa de interés fija). Finalmente, en seguros, el riesgo puede definirse como el monto de las reclamaciones totales de los asegurados. Veremos a continuación con más detalle este último caso pues es al que están dirigidos principalmente los modelos matemáticos que estudiaremos.

A grandes rasgos, la forma en la que opera un seguro es la siguiente: un grupo de personas reconocen que están expuestas a sufrir algún tipo de siniestro en sus bienes o en sus personas, y que dichos siniestros pueden causarles consecuencias irreparables, como la pérdida de sus vidas, o bien, pérdidas económicas considerables. Al contratar un seguro (es decir, firmar una póliza de seguro), cada una de estas personas paga por adelantado una cantidad de dinero (generalmente pequeña) llamada *prima* a una compañía aseguradora, quien se compromete a resarcir monetariamente a todas aquellas personas aseguradas que sufrieron algún siniestro durante el tiempo de vigencia del seguro y según las condiciones especificadas en la póliza del seguro. De esta manera, aunque no se conozca de manera individual a las personas que sufrirán un siniestro, el capital obtenido de manera colectiva debe ser suficiente para solventar los gastos de los siniestros individuales que se presenten. Es claro que bajo este mecanismo las pérdidas económicas individuales se distribuyen en la población entera de asegurados logrando así garantizar la sobrevivencia financiera de todos ellos. Naturalmente, para que tal mecanismo de asistencia colectiva sea factible, se deben cumplir varias condiciones, entre ellas, es necesario que el número de asegurados sea suficientemente grande, que se establezcan con precisión las características de los siniestros a considerar y que exista buena fe de ambas partes para respetar los acuerdos pactados. Es claro también que el número de siniestros, los momentos en los que éstos se presentan, así como el monto de las reclamaciones son variables desconocidas dependientes del azar, y que los modelos de la teoría de la probabilidad podrían ser de alguna ayuda en su estudio. En efecto, en las siguientes páginas estudiaremos algunos modelos matemáticos que han ayudado a entender y controlar en alguna medida el aspecto aleatorio de ciertas variables relevantes en los seguros.

1.2. Modelo individual

Suponga que se tiene un portafolio de n pólizas individuales de seguros válidas por un año como se muestra en la Figura 1.1.

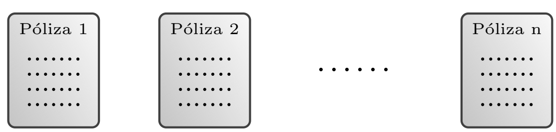


Figura 1.1

Sea p_j la probabilidad de que el j-ésimo asegurado no efectúe ninguna reclamación durante el tiempo de vigencia del seguro, y sea q_j la probabilidad de que se observe exactamente una reclamación. Suponga que la igualdad $p_j + q_j = 1$ se cumple, ello significa que no puede haber más de una reclamación por cada asegurado. Tal situación puede corresponder, por ejemplo, a los seguros de vida. Defina la variable aleatoria

$$D_j = \begin{cases} 1 & \text{si hay reclamación en la póliza } j, \\ 0 & \text{si no hay reclamación en la póliza } j. \end{cases}$$

Claramente D_j tiene distribución Bernoulli con parámetro q_j. El uso de la letra D viene del término en inglés *Death*. Observe que el número total de reclamaciones está dado por la variable aleatoria $N = D_1 + \cdots + D_n$. Ahora suponga artificialmente que cada póliza efectúa una reclamación, y sea la variable aleatoria $C_j > 0$ el monto de la reclamación efectuada por la póliza j. Debido a que los siniestros pueden presentarse con características distintas y ello puede derivar en distintos montos de reclamación, consideraremos de manera general a C_j no como una constante sino como una variable aleatoria. La letra C proviene del término en inglés *Claim*, que se traduce en español como reclamación. La verdadera reclamación de la póliza j está dada por el producto

$$D_j C_j = \begin{cases} C_j & \text{si } D_j = 1, \\ 0 & \text{si } D_j = 0. \end{cases}$$

Observe que esta variable aleatoria puede ser mixta, es decir, no ser continua ni discreta. Véase la Figura 1.2(b) en donde se muestra una posible gráfica de la función de distribución de esta variable aleatoria. De esta forma se considera como datos en este modelo la colección de vectores aleatorios:

$$(D_1, C_1), \ldots, (D_n, C_n),$$

los cuales supondremos independientes. Consideraremos además que las variables D_j y C_j también son independientes entre sí.

Definición 1.1 *El monto de reclamaciones agregadas, o también llamado agregado de reclamaciones, en el modelo individual, es la variable aleatoria*

$$S = \sum_{j=1}^{n} D_j C_j. \tag{1.1}$$

en donde $D_1, \ldots, D_n, C_1, \ldots, C_n$ son variables aleatorias independientes con $C_j > 0$ y D_j con distribución Ber(q_j).

La variable aleatoria S es el monto que afronta una compañía aseguradora por concepto de reclamaciones durante el periodo completo del seguro. La ecuación (1.1) representa el modelo individual para una póliza de seguros de las características señaladas. El modelo tiene este nombre pues supone conocer las probabilidades de reclamación y posible monto de reclamación de todos y cada uno de los asegurados de manera individual. Una posible desventaja de este modelo es que presupone que el número de asegurados en la cartera se mantiene constante durante todo el tiempo de vigencia del seguro. Desde el punto de vista matemático, y también desde la perspectiva del negocio del seguro, nuestro objetivo es conocer las características de la variable S, a la cual llamaremos riesgo.

Notación

$F_j(x)$	Función de distribución de D_jC_j
$G_j(x)$	Función de distribución de C_j

Usando esta notación, la función de distribución $F(x)$ del riesgo S adquiere la siguiente expresión en términos de convoluciones: $F(x) = (F_1 * \cdots * F_n)(x)$, sin embargo esta expresión es un tanto difícil de calcular justamente debido a las convoluciones y no la utilizaremos con frecuencia. Por otro lado, es inmediato comprobar que las funciones $F_j(x)$ y $G_j(x)$ guardan la siguiente relación.

Proposición 1.1 $$F_j(x) = \begin{cases} 1 - q_j\,(1 - G_j(x)) & si \;\; x \geqslant 0, \\ 0 & si \;\; x < 0. \end{cases}$$

Demostración. Para cualquier número real $x \geqslant 0$,

$$\begin{aligned} F_j(x) &= P(D_jC_j \leqslant x) \\ &= P(D_jC_j \leqslant x \,|\, D_j = 0)\, P(D_j = 0) \\ &\quad + P(D_jC_j \leqslant x \,|\, D_j = 1)\, P(D_j = 1) \\ &= P(0 \leqslant x \,|\, D_j = 0)\, p_j + P(C_j \leqslant x \,|\, D_j = 1)\, q_j \\ &= p_j + q_j\, G_j(x) \\ &= 1 - q_j(1 - G_j(x)). \end{aligned}$$

■

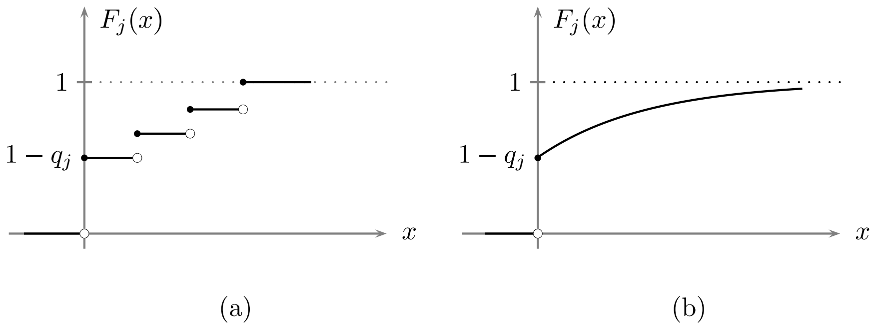


Figura 1.2

En la Figura 1.2 se muestran gráficas de la función de distribución $F_j(x)$ en dos casos: uno cuando $G_j(x)$ es discreta y otro cuando $G_j(x)$ es continua. Denotaremos también por $M_X(t)$ a la función generadora de momentos de una variable X, cuando exista. Como primeros resultados generales presentamos a continuación algunas características de S.

Proposición 1.2 *Bajo la notación e hipótesis del modelo individual se tienen los siguientes resultados.*

1. $E(S) = \sum_{j=1}^{n} q_j E(C_j)$.

2. $Var(S) = \sum_{j=1}^{n} q_j \,[\, Var(C_j) + p_j E^2(C_j)\,]$.

3. $M_{D_j C_j}(t) = 1 - q_j\,(1 - M_{C_j}(t))$.

4. $M_S(t) = \prod_{j=1}^{n} [\, 1 - q_j(1 - M_{C_j}(t))\,]$.

Demostración.

1. Por la hipótesis de independencia,
$$E(S) = \sum_{j=1}^{n} E(D_jC_j) = \sum_{j=1}^{n} E(D_j)E(C_j) = \sum_{j=1}^{n} q_jE(C_j).$$

2. Primeramente tenemos que
$$\begin{aligned} \mathrm{Var}(D_jC_j) &= E(D_j^2C_j^2) - E^2(D_jC_j) \\ &= q_jE(C_j^2) - q_j^2E^2(C_j) \\ &= q_j\left[\mathrm{Var}(C_j) + E^2(C_j)\right] - q_j^2E^2(C_j) \\ &= q_j\mathrm{Var}(C_j) + q_jp_jE^2(C_j). \end{aligned}$$
Por lo tanto,
$$\mathrm{Var}(S) = \sum_{j=1}^{n} \mathrm{Var}(D_jC_j) = \sum_{j=1}^{n} q_j\left[\,\mathrm{Var}(C_j) + p_jE^2(C_j)\,\right].$$

3. Nuevamente condicionando sobre el valor de D_j,
$$\begin{aligned} M_{D_jC_j}(t) &= E(e^{tD_jC_j}) \\ &= E(e^{tD_jC_j} \mid D_j = 0)P(D_j = 0) \\ &\quad + E(e^{tD_jC_j} \mid D_j = 1)P(D_j = 1) \\ &= p_j + q_jM_{C_j}(t) \\ &= 1 - q_j(1 - M_{C_j}(t)). \end{aligned}$$

4. Esta igualdad se sigue directamente de la anterior usando la hipótesis de independencia.

■

Puede considerarse que la variable S tiene una distribución binomial generalizada en el siguiente sentido: se tienen n ensayos independientes Bernoulli, en donde la probabilidad de éxito en el ensayo j es q_j y el valor asociado al resultado éxito no es 1 como en el esquema usual sino el monto $C_j > 0$.

Puede comprobarse que cuando q_j es constante q y los montos C_j son todos iguales a 1, la variable S tiene distribución $\text{bin}(n,q)$, y las fórmula generales demostradas para S se reducen a las de esta distribución. Al respecto véase el ejercicio 15.

Es también interesante observar que aunque inicialmente el modelo individual de riesgo que hemos presentado puede aplicarse a esquemas de seguros en donde hay como máximo una reclamación por póliza, esta única reclamación puede considerarse como el monto total conformado por la suma de varias posibles reclamaciones efectuadas por una póliza a lo largo del periodo de vigencia del seguro. De este modo el modelo individual puede también aplicarse al caso de reclamaciones múltiples. En cualquier caso, los datos que deben tenerse o estimarse estadísticamente para aplicar el modelo individual a una situación real son el número de asegurados n, las probabilidades de reclamación $q_1, q_2, \ldots, q_n$, y las distribuciones de probabilidad de los montos $C_1, C_2, \ldots, C_n$.

Aproximación normal

Cuando n es grande y el portafolio de asegurados es homogéneo en el sentido de que las variables $D_j C_j$ son idénticamente distribuidas con segundo momento finito, puede usarse el teorema central del límite para aproximar la distribución de S mediante la distribución normal, es decir,

$$
\begin{aligned}
P(S \leqslant x) &= P\Big(\frac{S - E(S)}{\sqrt{\text{Var}(S)}} \leqslant \frac{x - E(S)}{\sqrt{\text{Var}(S)}}\Big) \\
&\approx \Phi\Big(\frac{x - E(S)}{\sqrt{\text{Var}(S)}}\Big).
\end{aligned}
$$

Esta aproximación puede ser adecuada para ciertos riesgos pero tiene la desventaja de que asigna una probabilidad positiva al intervalo $(-\infty, 0)$, lo cual no es consistente con el hecho de que $S \geqslant 0$. Sin embargo, dado que la distribución $\text{N}(\mu, \sigma^2)$ se concentra principalmente en el intervalo $(\mu - 4\sigma, \mu + 4\sigma)$, cuando la esperanza y la varianza de S son tales que $E(S) - 4\sqrt{\text{Var}(S)} \geqslant 0$, la probabilidad asignada a la parte negativa del eje es realmente pequeña, ello alivia un poco el hecho de que esta distribución

no tenga soporte en el intervalo $[0, \infty)$. Tal vez la situación más comprometedora sea que la función de densidad normal decae muy rápidamente pues existen riesgos cuyas funciones de densidad no cumplen con tal característica. Más adelante mencionaremos esta propiedad de las distribuciones de los riesgos en términos de colas pesadas y ligeras.
En la siguiente sección encontraremos una forma recursiva para calcular la función de probabilidad de S cuando el monto de las reclamaciones se modela mediante una variable aleatoria discreta.

1.3. Fórmula de De Pril

Presentamos a continuación la fórmula de De Pril. Este resultado fue demostrado por Nelson De Pril [12] en 1986 y proporciona una expresión exacta, aunque recursiva, de la distribución de probabilidad de un riesgo en el modelo individual. La fórmula es bastante general aunque presupone que las reclamaciones toman los valores en el conjunto $\{1, 2, \ldots\}$. Este supuesto no es realmente una restricción fuerte, pues en la práctica el pago de siniestros se realiza siempre usando alguna unidad monetaria. Por otro lado, la fórmula que encontraremos permite que los riesgos a asegurar no sean homogéneos, sin embargo, estarán determinados *a priori*, es decir, serán deterministas y fijos. Para establecer la fórmula de De Pril es necesario dividir el portafolio de n asegurados de acuerdo con 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 $\{1, 2, \ldots, I\}$ y j en $\{1, 2, \ldots, J\}$. De esta forma se tiene la tabla de la Figura 1.3 en donde la suma de las entradas es n, es decir,

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

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,

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

En los siguientes cálculos haremos uso de la función generadora de probabilidad. El lector encontrará en el Apéndice (página 249) un recordatorio de

Probabilidades de reclamación

		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}
Monto de la reclamación	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	i	$\cdots$	$\cdots$	$\cdots$	n_{ij}	$\cdots$	n_{iJ}
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	I	n_{I1}	n_{I2}	$\cdots$	n_{Ij}	$\cdots$	n_{IJ}

Figura 1.3

esta función y de algunas de sus propiedades.

Teorema 1.1 (Fórmula de De Pril [i]) *Sea n_{ij} el número de pólizas cuyos asegurados tienen tasa de mortalidad q_j y suma asegurada i. Suponga que $j = 1, 2, \ldots, J$, e $i = 1, 2, \ldots, I$. Entonces para el riesgo S que sigue el modelo individual, las probabilidades $g_x = P(S = x)$ están dadas por*

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

en donde

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

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

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

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

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

en donde $g_r = P(S=r)$. Tomando logaritmo y después derivando,

$$\ln G(t) = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \ln(1-q_j+q_j t^i).$$

$$\frac{d}{dt} \ln G(t) = \frac{G'(t)}{G(t)} = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \, \frac{i q_j t^{i-1}}{1-q_j+q_j t^i}.$$

Por lo tanto,

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

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,

$$tG'(t) = G(t) \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \, i \sum_{k=1}^{\infty} (-1)^{k-1} \Big(\frac{q_j}{1-q_j}\Big)^k t^{ik}. \tag{1.2}$$

Defina ahora la función

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

La doble suma respecto de los índices j y k que aparece en la expresión (1.2) es absolutamente convergente en cualquiera de los dos órdenes que se efectúen estas sumas y el resultado es el mismo. Por lo tanto, es válido el intercambio en el orden de estas 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 las expresiones para $G'(t)$ y $G(t)$ en sus correspondientes series de potencias se obtiene

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

Para $x \geqslant 1$, 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 $1 \leqslant ik \leqslant x$. Se pueden primero establecer los posibles valores para i de la siguiente forma: $i = 1, \ldots, x \wedge I$ y por lo tanto los valores para k son $k = 1, \ldots, \lfloor x/i \rfloor$, 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 $x \geqslant 1$,

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

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

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

■

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

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

Ejemplo 1.1 *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 se encuentra el código en R de una implementación de la fórmula de De Pril [i]. Debe tenerse cuidado en la implementación numérica de esta fórmula, pues dado su carácter 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

$i \setminus q$	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

La fórmula que hemos denominado de De Pril [i] y que se encuentra expresada en el contexto de un portafolio de asegurados individuales, puede

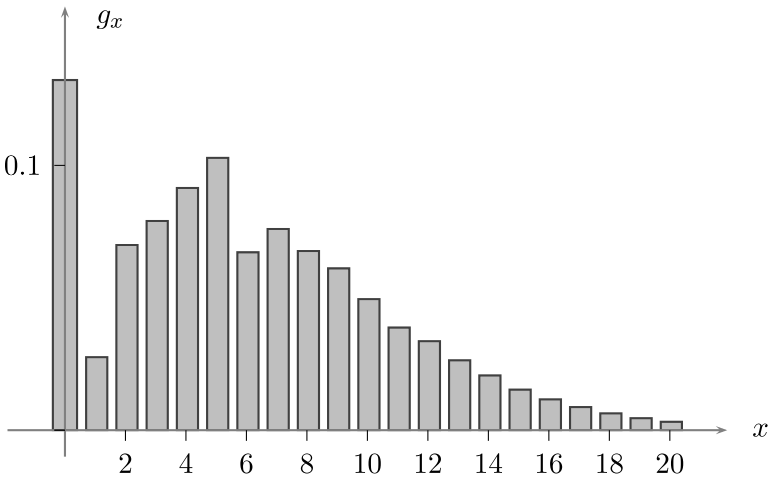


Figura 1.5

escribirse como un resultado teórico de la teoría de la probabilidad. Este resultado es el contenido de la siguiente proposición. La fórmula tiene una expresión más simple y la demostración sigue los mismos lineamientos que la que hemos presentado, sin embargo, escribiremos nuevamente los detalles de la prueba en esta versión simplificada.

Proposición 1.3 (Fórmula de De Pril [ii]) *Sean* $X_1, X_2, \ldots, X_n$ *v.a.i.i.d. con valores en el conjunto* $\{0, 1, 2, \ldots\}$. *Para cada entero* $j \geqslant 0$, *defina la probabilidad* $f_j = P(X = j)$, *y suponga* $f_0 \neq 0$. *Sea* $S = X_1 + \cdots + X_n$. *Entonces las probabilidades* $g_x = P(S = x)$ *se pueden calcular recursivamente mediante la siguiente fórmula:*

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

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,

$$P_X(t) = E(t^X) = \sum_{k=0}^{\infty} t^k f_k,$$
$$P_S(t) = E(t^S) = \sum_{k=0}^{\infty} t^k g_k.$$

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 $t\, P_X(t)$,

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

que en términos de sumas se escribe como sigue

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

El siguiente paso es identificar el coeficiente del término t^x para $x \geqslant 1$ en cada lado de la ecuación. Por ejemplo, para el lado izquierdo el coeficiente es el término $f_j\, k\, g_k$ para todos aquellos valores de $j \geqslant 0$ y $k \geqslant 1$ 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

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

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

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

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

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

■

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

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

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 1 y el resto toma el valor 0, y la segunda situación cuando uno de los sumandos toma el valor 2 y el resto es 0. Los coeficientes binomiales dan cuenta de las distintas formas en las que se pueden presentar estos arreglos.

Ejemplo 1.2 *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*

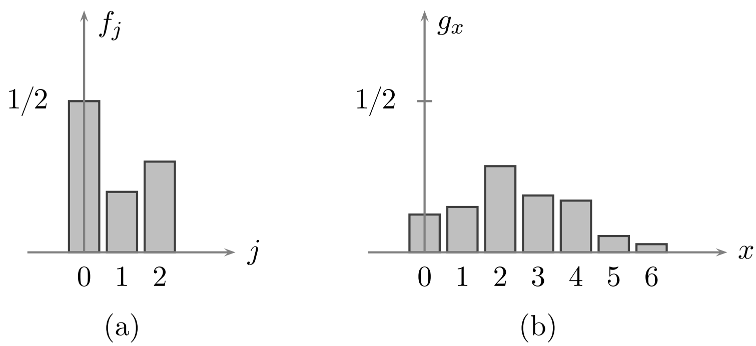


Figura 1.6

Observe que la variable suma puede tomar cualquiera de los valores $0, 1, \ldots, 6$. *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).*

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

En la sección de ejercicios el lector puede encontrar algunas aplicaciones de la fórmula de DePril [ii] para obtener la función de probabilidad de variables aleatorias que pueden construirse como sumas de variables discretas con distribución conocida, por ejemplo, las distribuciones binomial y Poisson.

1.4. Modelo colectivo

Considere un conjunto de un número no determinado de contratos de seguros con vigencia en un periodo de tiempo $[0,T]$. Este periodo puede corresponder a un año por ejemplo. Sea N la variable aleatoria que denota el número de reclamaciones ocurridas en este intervalo y sean las variables positivas $Y_1,\ldots,Y_N$ los montos de estas reclamaciones. Gráficamente una posible realización de tal esquema se muestra en la Figura 1.7.

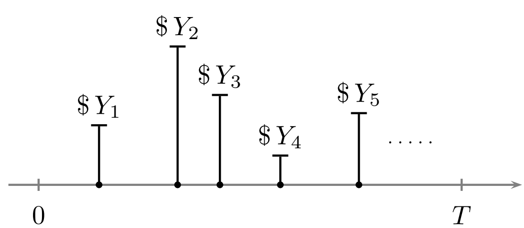


Figura 1.7

Consideraremos que el número de reclamaciones y los montos de éstas son variables aleatorias independientes. Más aún, supondremos que las reclamaciones mismas son independientes entre sí y que comparten la misma distribución de probabilidad.

Definición 1.2 *El monto agregado o monto acumulado de todas las reclamaciones efectuadas es la variable aleatoria S, llamada riesgo, y definida como sigue*

$$S = \sum_{j=1}^{N} Y_j, \tag{1.3}$$

en donde $Y_1, Y_2, \ldots$ es una colección de variables aleatorias independientes positivas idénticamente distribuidas e independientes de la variable aleatoria N con valores en el conjunto $\{0, 1, 2, \ldots\}$. Cuando $N = 0$ se define S como cero.

Observe que cada sumando es una variable aleatoria y que el número de sumandos es también aleatorio. Observe además que S puede ser una variable aleatoria mixta, es decir, no ser discreta ni continua, pues cuando los montos de las reclamaciones Y son variables continuas estrictamente positivas, la variable S puede tomar el valor 0 con probabilidad $P(S = 0) = P(N = 0) > 0$, y puede además tomar cualquier valor en el intervalo $(0, \infty)$. La ecuación (1.3) representa el modelo colectivo para un contrato de seguros, cuyas posibles realizaciones como función del tiempo tienen la forma de la gráfica de la Figura 1.8.

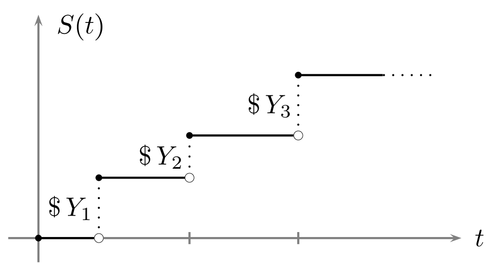


Figura 1.8

A la función de distribución de cada reclamación Y la denotaremos por la letra G. Se asume naturalmente que $G(0) = 0$, ello equivale a decir que la variable Y es positiva. Adicionalmente usaremos la abreviación:

Notación $\quad \mu_n := E(Y^n)$

En particular se escribe μ en lugar de $\mu_1 = E(Y)$. Nuevamente el problema central es encontrar la distribución de probabilidad de S, la cual depende de la distribución de Y y de N. Un primer resultado general al respecto es el que aparece a continuación. Antes de enunciarlo recordemos que la 0-convolución de una función de distribución G se define como

$$G^{*0}(x) = \begin{cases} 1 & \text{si } x \geqslant 0, \\ 0 & \text{si } x < 0. \end{cases}$$

Proposición 1.4 *La función de distribución del riesgo S en el modelo colectivo es*

$$F(x) = \sum_{n=0}^{\infty} G^{*n}(x)\, P(N=n).$$

Demostración.

$$\begin{aligned} F(x) &= \sum_{n=0}^{\infty} P(S \leqslant x \,|\, N=n)\, P(N=n) \\ &= P(S \leqslant x \,|\, N=0)\, P(N=0) + \sum_{n=1}^{\infty} P(Y_1 + \cdots + Y_n \leqslant x)\, P(N=n) \\ &= G^{*0}(x)\, P(N=0) + \sum_{n=1}^{\infty} G^{*n}(x)\, P(N=n) \\ &= \sum_{n=0}^{\infty} G^{*n}(x)\, P(N=n). \end{aligned}$$

■

Algunas características numéricas de la variable S se muestran a continuación.

Proposición 1.5 *Suponiendo que las cantidades y funciones indicadas existen, el riesgo S en el modelo colectivo cumple las siguientes propiedades.*

1. $E(S) = E(N)E(Y)$.
2. $E(S^2) = E(N)E(Y^2) + E(N(N-1))E^2(Y)$.
3. $Var(S) = Var(N)E^2(Y) + Var(Y)E(N)$.
4. $M_S(t) = M_N(\ln(M_Y(t)))$.

Demostración.

1. Condicionaremos sobre el valor de N y después usaremos la hipótesis de independencia. El resultado del cálculo es el mismo cuando la variable N inicia en el valor 0 o en el valor 1.

$$\begin{aligned}
E(S) &= \sum_{n=0}^{\infty} E(\sum_{j=1}^{N} Y_j \,|\, N=n)\, P(N=n) \\
&= \sum_{n=0}^{\infty} E(\sum_{j=1}^{n} Y_j \,|\, N=n)\, P(N=n) \\
&= \sum_{n=0}^{\infty} nE(Y)\, P(N=n) \\
&= E(N)\, E(Y).
\end{aligned}$$

2. Nuevamente condicionando sobre el valor de N,

$$
\begin{aligned}
E(S^2) &= \sum_{n=0}^{\infty} E((\sum_{j=1}^{N} Y_j)^2 \,|\, N=n)\, P(N=n) \\
&= \sum_{n=0}^{\infty} E((\sum_{j=1}^{n} Y_j)^2 \,|\, N=n)\, P(N=n) \\
&= \sum_{n=0}^{\infty} E((\sum_{j=1}^{n} Y_j)^2)\, P(N=n) \\
&= \sum_{n=0}^{\infty} \Big[\, \sum_{j=1}^{n} E(Y_j^2) + \sum_{\substack{j,k=1 \\ j\neq k}}^{n} E(Y_j Y_k) \Big]\, P(N=n).
\end{aligned}
$$

Observe que segunda suma es nula cuando $n=0$ y a su vez la tercera suma se anula cuando $n=0$ o 1. Así, por la idéntica distribución tenemos que

$$
\begin{aligned}
E(S^2) &= \sum_{n=0}^{\infty} nE(Y^2)\, P(N=n) + \sum_{n=0}^{\infty} n(n-1)\, E^2(Y)\, P(N=n) \\
&= E(N)\, E(Y^2) + E(N(N-1))\, E^2(Y).
\end{aligned}
$$

3. Por las fórmulas anteriores,

$$
\begin{aligned}
\mathrm{Var}(S) &= E(S^2) - E^2(S) \\
&= E(N)\, E(Y^2) + E(N(N-1))\, E^2(Y) - E^2(N)\, E^2(Y) \\
&= E(N)\,[\, E(Y^2) - E^2(Y)\,] + [\, E(N^2) - E^2(N)\,]\, E^2(Y) \\
&= E(N)\, \mathrm{Var}(Y) + \mathrm{Var}\,(N)\, E^2(Y).
\end{aligned}
$$

4. De manera análoga a los dos primeros incisos,

$$
\begin{aligned}
M_S(t) &= \sum_{n=0}^{\infty} E(e^{r(Y_1+\cdots+Y_N)} \,|\, N=n)\, P(N=n) \\
&= \sum_{n=0}^{\infty} E(e^{r(Y_1+\cdots+Y_n)})\, P(N=n) \\
&= \sum_{n=0}^{\infty} (M_Y(t))^n\, P(N=n) \\
&= E((M_Y(t))^N) \\
&= E(e^{N\ln(M_Y(t))}) \\
&= M_N(\ln(M_Y(t))).
\end{aligned}
$$

■

Véase la sección de ejercicios para los terceros momentos de S. Consideraremos a continuación algunos casos particulares del modelo colectivo.

Modelo binomial compuesto

Cuando el número de reclamaciones N tiene una distribución $\text{bin}(n,p)$ se dice que el riesgo S tiene una distribución binomial compuesta, y se escribe $S \sim \text{bin comp}(n,p,G)$, en donde G es la función de distribución de cada sumando en la definición de S. Bajo esta hipótesis se tienen los siguientes resultados.

Proposición 1.6 *Si N tiene distribución bin(n,p), entonces*

a) $E(S) = np\mu$.

b) $E(S^2) = np(\mu_2 + (n-1)p\mu^2)$.

c) $Var(S) = np(\mu_2 - p\mu^2)$.

d) $M_S(t) = (1-p+pM_Y(t))^n$.

Estas expresiones se siguen fácilmente de las fórmulas generales demostradas antes, basta recordar que si N tiene distribución bin(n,p), entonces $E(N) = np$, $\text{Var}(N) = np(1-p)$, y $M_N(t) = (1-p+pe^t)^n$. Véase la sección de ejercicios para los terceros momentos de este modelo. Observe que en este caso se tiene una cota superior para el número de reclamaciones que pueden efectuarse.

Modelo binomial negativo compuesto

Cuando el número de reclamaciones N tiene una distribución binomial negativa se dice que el riesgo S tiene una distribución binomial negativa compuesta. Esto es, si $N \sim$ bin neg(k,p), entonces $S \sim$ bin neg comp(k,p,G), donde nuevamente G hace referencia a la función de distribución de cada sumando de S. En este caso se cumple lo siguiente.

Proposición 1.7 *Si N tiene distribución bin neg(k,p), entonces*

a) $E(S) = k(1/p-1)\mu$.

b) $E(S^2) = k(1/p-1)[\mu_2 + (k+1)(1/p-1)\mu^2]$.

c) $Var(S) = k(1/p-1)[\mu_2 + (1/p-1)\mu^2]$.

d) $M_S(t) = \left(\dfrac{p}{1-(1-p)M_Y(t)}\right)^k$.

Para encontrar estas fórmulas es suficiente recordar que si N tiene distribución bin neg(k,p), entonces $E(N) = k(1-p)/p$, $\text{Var}(N) = k(1-p)/p^2$, y $M_N(t) = [p/(1-(1-p)e^t)]^k$. En la sección de ejercicios se presenta el tercer momento de este modelo. En el caso particular cuando $k=1$, la distribución de N se reduce a la distribución geométrica de parámetro p y se dice que S tiene distribución geométrica compuesta.

Modelo Poisson compuesto

Cuando el número de reclamaciones N tiene una distribución Poisson se dice que el riesgo S tiene una distribución Poisson compuesta, y se escribe $S \sim$ Poisson comp(λ, G), en donde λ es el parámetro de la distribución Poisson y G es la función de distribución de cada sumando de S. Para este modelo se tienen los siguientes resultados.

Proposición 1.8 *Si N tiene distribución Poisson(λ), entonces*

a) $E(S) = \lambda\mu$.

b) $E(S^2) = \lambda\mu_2 + \lambda^2\mu^2$.

c) $Var(S) = \lambda\mu_2$.

d) $M_S(t) = \exp[\,\lambda(M_Y(t) - 1)\,]$.

Nuevamente estas expresiones son consecuencia de las fórmulas generales demostradas antes, y del hecho de que si N tiene distribución Poisson(λ), entonces $E(N) = \lambda$, $\text{Var}(N) = \lambda$, y $M_N(t) = \exp(\lambda(e^t - 1))$. Véase la sección de ejercicios para los terceros momentos de este modelo. Observe que el parámetro λ y la distribución de la variable Y determinan por completo al modelo Poisson compuesto. Estudiaremos con más detalle este modelo en la siguiente sección.

1.5. Modelo colectivo Poisson

En esta sección retomamos el caso cuando el número de reclamaciones en el modelo colectivo sigue una distribución Poisson. Primeramente explicaremos la forma en la que se puede obtener un modelo colectivo Poisson a partir del modelo individual. Después mostraremos cómo este modelo Poisson compuesto aproxima al modelo individual. Finalmente estudiaremos algunas propiedades interesantes y útiles del modelo colectivo Poisson.

Modelo Poisson compuesto asociado al modelo individual

Considere el modelo individual de riesgo $S^i = \sum_{j=1}^{n} D_j C_j$, junto con la notación e hipótesis correspondientes. El superíndice i indica que se trata de un modelo individual. A partir de este modelo se construye a continuación un modelo colectivo con distribución Poisson compuesta. Para ello recuerde que únicamente se necesita establecer el valor del parámetro λ de la distribución Poisson y la función de distribución $G(x)$ del monto de las reclamaciones. Sean entonces

$$\lambda = \sum_{j=1}^{n} q_j, \tag{1.4}$$

$$\text{y} \qquad G(x) = \sum_{j=1}^{n} \frac{q_j}{\lambda} G_j(x), \tag{1.5}$$

en donde $G_j(x)$ es la función de distribución de la variable C_j. Haremos algunas observaciones sobre estas definiciones.

a) Mediante la primera igualdad se establece que el número esperado de reclamaciones en ambos modelos es el mismo.

b) En la segunda ecuación se define a la función de distribución de una reclamación en el modelo colectivo como el promedio ponderado de las funciones de distribución del monto de todas las reclamaciones en el modelo individual, siendo las ponderaciones los factores $q_1/\lambda, \ldots, q_n/\lambda$. Como estos números conforman una distribución de probabilidad, el promedio resulta ser una función de distribución. Observamos que al agregar los distintos tipos de riesgo C_j del modelo individual para formar el modelo colectivo con reclamaciones Y_j independientes e idénticamente distribuidas, se pierde el comportamiento individual de los riesgos. Así, en la proporción en la que el j-ésimo riesgo individual se agrega, esto es q_j/λ, así es su contribución en la función de distribución de las reclamaciones en el modelo colectivo.

De esta forma se construye el modelo colectivo $S^c = \sum_{j=1}^{N} Y_j$, al cual llamaremos modelo colectivo Poisson compuesto asociado al modelo individual, en donde el superíndice c indica que se trata de un modelo colectivo. Para este modelo particular se cumplen las siguientes igualdades:

a) $E(S^c) = \sum_{j=1}^{n} q_j \, E(C_j)$.

b) $\text{Var}(S^c) = \sum_{j=1}^{n} q_j \, E(C_j^2)$.

c) $E(Y^k) = \sum_{j=1}^{n} \frac{q_j}{\lambda} \, E(C_j^k)$.

d) $M_Y(t) = \sum_{j=1}^{n} \frac{q_j}{\lambda} \, M_{C_j}(t)$.

Estas expresiones se siguen directamente de resultados previos acerca del modelo Poisson compuesto y de las igualdades (1.4) y (1.5). Por ejemplo, usando integrales de Riemann-Stieltjes (véase el Apéndice que aparece al final del texto), el k-ésimo momento de una reclamación del modelo S^c es,

$$E(Y^k) = \int_0^\infty y^k dG(y) = \sum_{j=1}^{n} \frac{q_j}{\lambda} \int_0^\infty y^k dG_j(y) = \sum_{j=1}^{n} \frac{q_j}{\lambda} \, E(C_j^k).$$

De manera análoga se verifican las otras identidades. A modo de comparación se tiene que $E(S^i) = E(S^c)$, mientras que $\text{Var}(S^i) \leqslant \text{Var}(S^c)$.

El modelo Poisson compuesto asociado como límite del modelo individual: primera argumentación

Sea S^i el riesgo en un modelo individual y sea S^c el riesgo del modelo colectivo Poisson compuesto asociado. Demostraremos que este modelo colectivo puede ser obtenido como un proceso límite en el modelo individual. Consideremos entonces el modelo individual junto con la notación e hipótesis usuales. Por resultados previos sabemos que

$$M_{S^i}(t) = \prod_{j=1}^{n} [1 + q_j(M_{C_j}(t) - 1)].$$

Sea k un entero positivo cualquiera. De manera artificial y teórica se construye ahora un nuevo portafolio de asegurados con las siguientes características:

cada póliza j de las n originales se reemplaza por k subpólizas idénticas, en cada una de las cuales la probabilidad de reclamación se define como q_j/k, y la función de distribución del monto de una reclamación es la misma $G_j(y)$. Véase la Figura 1.9. En consecuencia, el portafolio consiste ahora de kn subpólizas en donde la probabilidad de reclamación en cada una de ellas es menor. No es de sorprenderse entonces que al hacer k tender a infinito se obtenga un modelo Poisson.

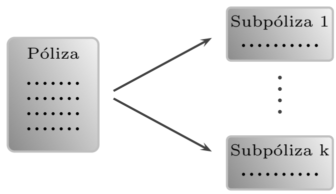


Figura 1.9

De este modo, si S_k^i denota el riesgo asociado al portafolio modificado, entonces se tiene que

$$M_{S_k^i}(t) = \prod_{j=1}^{n} [1 + \frac{q_j}{k}(M_{C_j}(t) - 1)]^k.$$

Se hace ahora tender k a infinito y usando el resultado $\lim_{k\to\infty} (1 + x/k)^k = e^x$ se obtiene

$$\begin{aligned} \lim_{k\to\infty} M_{S_k^i}(t) &= \prod_{j=1}^{n} \exp\left[\, q_j(M_{C_j}(t) - 1)\,\right] \\ &= \exp\left[\, \sum_{j=1}^{n} q_j M_{C_j}(t) - \lambda \,\right] \\ &= \exp\left[\, \lambda(M_Y(t) - 1)\,\right] \\ &= M_{S^c}(t). \end{aligned}$$

Puesto que la convergencia de funciones generadoras de momentos es equivalente a la convergencia en distribución de las correspondientes variables aleatorias, se obtiene entonces que el modelo colectivo Poisson compuesto asociado es el límite en distribución del modelo individual cuando el número de pólizas crece y las probabilidades de reclamación son cada vez más pequeñas. Así, el argumento presentado en esta sección justifica el uso del modelo Poisson compuesto en el caso cuando el portafolio es grande y las probabilidades de reclamación son pequeñas. Para reforzar esta idea, el argumento informal que presentamos en el siguiente párrafo también favorece al modelo Poisson compuesto bajo las condiciones límites mencionadas.

El modelo Poisson compuesto asociado como límite del modelo individual: segunda argumentación

Recordemos nuevamente que la función generadora de momentos del riesgo S^i es

$$M_{S^i}(t) = \prod_{j=1}^{n} [1 + q_j(M_{C_j}(t) - 1)].$$

Observemos que el término $q_j(M_{C_j}(t)-1)$ es pequeño para valores pequeños de t. Usando la fórmula $\ln(1+x) = x - x^2/2 + x^3/3 + \cdots$, se puede escribir la aproximación $\ln(1+x) \approx x$. De modo que

$$\begin{aligned}
\ln(M_{S^i}(t)) &= \sum_{j=1}^{n} \ln[1 + q_j(M_{C_j}(t) - 1)] \\
&\approx \sum_{j=1}^{n} q_j(M_{C_j}(t) - 1) \\
&= \lambda \, [\sum_{j=1}^{n} \frac{q_j}{\lambda} M_{C_j}(t) - 1 \,],
\end{aligned}$$

en donde $\lambda = q_1 + \cdots + q_n$. Por lo tanto,

$$M_{S^i}(t) \approx \exp\,(\,\lambda\,[\sum_{j=1}^{n} \frac{q_j}{\lambda} M_{C_j}(t) - 1\,]\,).$$

Esta fórmula corresponde a la función generadora de momentos del riesgo con distribución Poisson compuesta, en donde los montos de las reclamaciones tienen función generadora de momentos $\sum_{j=1}^{n} \frac{q_j}{\lambda} M_{C_j}(t)$.

Modelo Poisson compuesto con varios tipos de riesgos

Demostraremos ahora que la suma de riesgos independientes que siguen el modelo Poisson compuesto también es Poisson compuesto. Esta es una propiedad bastante útil y generaliza el resultado de probabilidad que establece que la suma de dos variables aleatorias independientes con distribución Poisson tiene nuevamente distribución Poisson.

Proposición 1.9 *Sean S_1 y S_2 dos riesgos independientes con distribución Poisson compuesta con parámetros λ_1 y λ_2, y reclamaciones $Y^{(1)}$ y $Y^{(2)}$ con función de distribución $G_1(x)$ y $G_2(x)$ respectivamente. Entonces el riesgo $S = S_1 + S_2$ también sigue una distribución Poisson compuesta con parámetro $\lambda = \lambda_1 + \lambda_2$, y las reclamaciones tienen función de distribución*

$$G(x) = \frac{\lambda_1}{\lambda} G_1(x) + \frac{\lambda_2}{\lambda} G_2(x).$$

Demostración. Por independencia tenemos que

$$\begin{aligned} M_{S_1+S_2}(t) &= M_{S_1}(t) M_{S_2}(t) \\ &= \exp[\lambda_1(M_{Y^{(1)}}(t) - 1)] \exp[\lambda_2(M_{Y^{(2)}}(t) - 1)] \\ &= \exp[\lambda(\frac{\lambda_1}{\lambda} M_{Y^{(1)}}(t) + \frac{\lambda_2}{\lambda} M_{Y^2}(t) - 1)], \end{aligned}$$

en donde $\frac{\lambda_1}{\lambda} M_{Y^{(1)}}(t) + \frac{\lambda_2}{\lambda} M_{Y^{(2)}}(t)$ es la función generadora de momentos de la función de distribución $G(x) = \frac{\lambda_1}{\lambda} G_1(x) + \frac{\lambda_2}{\lambda} G_2(x)$. ■

El resultado anterior puede extenderse fácilmente al caso $S = S_1 + \cdots + S_n$. Véase el enunciado del ejercicio 40 en la página 41.

Modelo Poisson compuesto con reclamaciones clasificadas

Sea S un riesgo con distribución Poisson compuesta de parámetro λ. Suponga que los montos de las reclamaciones pueden ser clasificadas en m categorías excluyentes y exhaustivas denotadas por $A_1, \ldots, A_m$. Típicamente estas categorías pueden ser intervalos de valores para las reclamaciones. Sea

$p_k = P(Y \in A_k) > 0$ tal que $p_1 + \cdots + p_m = 1$. Sea N_k el número de reclamaciones del tipo k. Entonces $N = N_1 + \cdots + N_m$ y debido a la independencia de los montos en las reclamaciones, el vector $(N_1, \ldots, N_m)$ tiene una distribución condicional multinomial$(p_1, \ldots, p_m; n)$ cuando $N = n$, es decir, para enteros no negativos $n_1, \ldots, n_m$ tales que $n_1 + \cdots + n_m = n$, la distribución multinomial establece que

$$\begin{aligned} P(N_1 = n_1, \ldots, N_m = n_m \mid N = n) &= \binom{n}{n_1 \; \cdots \; n_m} p_1^{n_1} \cdots p_m^{n_m} \\ &= \frac{n!}{n_1! \cdots n_m!}\, p_1^{n_1} \cdots p_m^{n_m}. \end{aligned}$$

La distribución no condicional del vector $(N_1, \ldots, N_m)$ es el contenido del siguiente resultado.

Proposición 1.10 *Las variables aleatorias $N_1, \ldots, N_m$ son independientes y cada variable N_k tiene distribución Poisson(λp_k).*

Demostración. Sean $n_1, \ldots, n_m$ enteros no negativos cualesquiera y sea n la suma de todos estos números, es decir, $n_1 + \cdots + n_m = n$. Entonces

$$\begin{aligned} P(N_1 = n_1, \ldots, N_m = n_m) &= P(N_1 = n_1, \ldots, N_m = n_m, N = n) \\ &= P(N_1 = n_1, \ldots, N_m = n_m \mid N = n) P(N = n) \\ &= \frac{n!}{n_1! \cdots n_m!}\, p_1^{n_1} \cdots p_m^{n_m} \, \frac{\lambda^n}{n!} e^{-\lambda} \\ &= \prod_{k=1}^{n} \frac{(\lambda p_k)^{n_k}}{n_k!} e^{-\lambda p_k}. \end{aligned}$$

Se desprende de esta igualdad que la variable N_k tiene distribución marginal Poisson(λp_k). De esta identidad se verifica también la independencia. ∎

Observe que, condicionadas al evento $(N = n)$, las variables $N_1, \ldots, N_m$ no son independientes, mientras que sin tal condición, lo son. Por otro lado, como los montos de las reclamaciones son independientes de N, el riesgo de tipo k está dado por la variable

$$S_k = \sum_{j=1}^{N_k} Y_j^{(k)},$$

en donde $Y_j^{(k)}$ es una variable aleatoria con función de distribución

$$G_k(x) = P(Y_j \leqslant x \,|\, Y_j \in A_k) = \frac{P(Y_j \leqslant x, Y_j \in A_k)}{P(Y_j \in A_k)}.$$

Por lo anterior, el riesgo S_k tiene distribución Poisson compuesta con parámetros λp_k y $G_k(x)$. En particular, cuando $A_k = (x_{k-1}, x_k]$, con $0 \leqslant x_0 < x_1 < \cdots < x_m$, la función de distribución $G_k(x)$ tiene la siguiente forma:

$$G_k(x) = \begin{cases} 0 & \text{si } \; x < x_k, \\ \dfrac{G(x) - G(x_{k-1})}{G(x_k) - G(x_{k-1})} & \text{si } \; x_{k-1} \leqslant x < x_k, \\ 1 & \text{si } \; x \geqslant x_k. \end{cases}$$

Modelo Poisson compuesto mixto

Cuando el número de reclamaciones N tiene una distribución Poisson(λ) y el parámetro λ es a su vez una variable aleatoria, se dice que el riesgo S tiene una distribución Poisson compuesta mixta. Algunas características de este modelo se muestran a continuación.

Proposición 1.11 *Si N tiene distribución Poisson(Λ) en donde Λ es una variable aleatoria con función de distribución $F_\Lambda(\lambda)$, entonces*

a) $P(N = n) = \displaystyle\int_0^\infty e^{-\lambda} \frac{\lambda^n}{n!}\, dF_\Lambda(\lambda).$

b) $E(S) = E(\Lambda)\mu.$

c) $E(S^2) = E(\Lambda)\mu_2 + E(\Lambda^2)\mu^2.$

d) $Var(S) = Var(\Lambda)\mu^2 + E(\Lambda)\mu_2.$

e) $M_S(t) = M_\Lambda(M_Y(t) - 1).$

Para obtener las identidades que aparecen en esta proposición es suficiente condicionar sobre el valor de Λ. Véase la sección de ejercicios para los terceros momentos de este modelo.

Comentarios y referencias

Hemos presentado dos modelos de riesgo con características distintas: el modelo individual y el modelo colectivo. En ambos casos se trata de una variable aleatoria a la cual le hemos llamado riesgo y que engloba el total de reclamaciones de un conjunto de asegurados durante un intervalo de tiempo arbitrario pero fijo. Con el objetivo de cuantificar y tener control de estos riesgos, el problema central ha sido encontrar las características numéricas de estas variables aleatorias y mejor aún, la distribución de probabilidad de ellas. Así, hemos encontrado varias expresiones para algunas características numéricas de estos dos modelos de riesgo. Hemos también encontrado que, bajo ciertas condiciones, las fórmulas de De Pril pueden aplicarse para calcular la distribución de probabilidad exacta de un riesgo que sigue un modelo individual. Para completar estos primeros resultados, en el siguiente capítulo veremos la fórmula de Panjer que nos permitirá aplicar un mecanismo recursivo para calcular la distribución de probabilidad exacta de un riesgo que sigue el modelo colectivo. El lector puede encontrar otras exposiciones sobre el modelo individual y colectivo en las siguientes referencias: Bowers *et al.* [7], Gerber [15], Klugman *et al.* [23].

1.6. Ejercicios

Modelo individual

1. Considere el modelo individual para un portafolio de n pólizas de seguros. Bajo la notación e hipótesis usuales, demuestre que el número esperado de reclamaciones es $q_1 + \cdots + q_n$.

2. Para un modelo individual de riesgo, encuentre la distribución de probabilidad del número total de reclamaciones en una cartera de n asegurados cuando D_j tiene distribución $\text{Ber}(q)$, es decir, $q_j = q > 0$ es constante.

3. Considere el modelo individual de riesgo en donde D_j tiene distribución Ber(q), es decir, $q_j = q > 0$ es constante. Suponga además que cada reclamación C_j es constante $c > 0$, es decir, se trata de una misma suma asegurada para todos. Encuentre una expresión para la esperanza y la varianza de S.

4. Considere el modelo individual para un portafolio de n pólizas de seguros de vida. Suponga que el j-ésimo asegurado tiene una suma asegurada constante c_j. Demuestre que:

   a) $E(S) = \sum_{j=1}^{n} q_j c_j$.

   b) $\text{Var}(S) = \sum_{j=1}^{n} q_j p_j c_j^2$.

5. Demuestre que si $f_1(x), \ldots, f_n(x)$ son funciones diferenciables que no se anulan, entonces

$$\frac{d}{dx}\left[\,\prod_{j=1}^{n} f_j(x)\,\right] = \left[\,\prod_{j=1}^{n} f_j(x)\,\right]\left[\,\sum_{j=1}^{n} \frac{f_j'(x)}{f_j(x)}\,\right].$$

   Use esta fórmula y la expresión encontrada para $M_S(t)$ en el modelo individual de riesgo para encontrar nuevamente $E(S)$ y $\text{Var}(S)$.

6. Para un riesgo S que sigue el modelo individual demuestre que:

   a) $E(S^2) = \sum_{j=1}^{n} q_j E(C_j^2) + \sum_{i \neq j} q_i q_j E(C_i) E(C_j)$.

   b) $$E(S^3) = \sum_{j=1}^{n} q_j E(C_j^3) + 3 \sum_{i \neq j} q_i q_j E(C_i) E(C_j^2) + \sum_{\substack{i,j,k \\ \text{distintos}}} q_i q_j q_k E(C_i) E(C_j) E(C_k).$$

7. Suponga que D y C son dos variables aleatorias independientes tales que D tiene distribución Ber(q) y C se distribuye $\exp(\lambda)$. Calcule y grafique la función de distribución de la variable aleatoria mixta DC.

8. Para el modelo individual, suponga que D_j tiene distribución Ber(q), es decir, $q_j = q > 0$ es constante. Encuentre la distribución del riesgo S cuando $n = 3$ y C_j tiene la siguiente distribución de probabilidad:

$$\begin{aligned} P(C_j = 1) &= 0.6, \\ P(C_j = 2) &= 0.3, \\ P(C_j = 3) &= 0.1. \end{aligned}$$

9. Considere un portafolio de 21 pólizas individuales de seguros de vida válidas por un año como se indica en la tabla que aparece abajo. Usando el modelo individual calcule $E(S)$ y $\text{Var}(S)$.

Tasa de mortalidad q_j	Suma asegurada			
	\$2	\$3	\$4	\$5
0.04	1	1	2	1
0.05	0	2	3	3
0.06	1	1	2	4

10. Sean $q_{j,0}$, $q_{j,1}$ y $q_{j,2}$ las probabilidades de que el j-ésimo asegurado presente 0, 1 y 2 reclamaciones respectivamente durante el tiempo de vigencia del seguro. Suponga que cada una de las posibles reclamaciones de la póliza j es constante z_j y que $q_{j,0} + q_{j,1} + q_{j,2} = 1$. Encuentre fórmulas para $E(S)$ y $\text{Var}(S)$ en el modelo individual.

11. Una compañía aseguradora tiene una cartera con pólizas de seguros de vida y diferentes sumas aseguradas como se muestra en la tabla que aparece abajo. Calcule $E(S)$ y $\text{Var}(S)$ usando el modelo individual.

Suma asegurada	Número de pólizas	Probabilidad de reclamación
\$10,000	50	0.0040
\$20,000	75	0.0035
\$30,000	100	0.0030

12. Considere el modelo individual de riesgo para una cartera de $n = 56$ asegurados divididos en cinco subgrupos de la siguiente forma: 11 asegurados con probabilidad de reclamación $q = 0.01$, 7 asegurados con $q = 0.015$, 20 asegurados con $q = 0.02$, 10 asegurados con $q = 0.025$ y 8 asegurados con $q = 0.03$. Todos ellos con suma asegurada \$100. Calcule el valor esperado del agregado de reclamaciones del riesgo correspondiente a esta cartera de asegurados.

13. Para el modelo individual, suponga que D_j tiene distribución $\text{Ber}(q)$, es decir, $q_j = q > 0$ es constante, y que cada reclamación C_j tiene distribución $\exp(\lambda)$. Encuentre la distribución de probabilidad del riesgo S.

14. Considere el modelo individual de riesgo $S = \sum_{j=1}^{n} D_j C_j$, en donde $C_1, C_2, \ldots, C_n$ son constantes posiblemente distintas todas ellas, y cada D_j tiene distribución $\text{Ber}(q_j)$ para $j = 1, 2, \ldots, n$.

    a) Demuestre que la función de probabilidad de la variable $D_j C_j$ es

    $$f_{D_j C_j}(x) = \begin{cases} 1 - q_j & \text{si } x = 0, \\ q_j & \text{si } x = C_j, \\ 0 & \text{en otro caso.} \end{cases}$$

    b) Defina $S_j = S_{j-1} + D_j C_j$ para $j = 2, 3, \ldots, n$ y $S_1 = D_1 C_1$. Demuestre que la función de probabilidad de la variable S_j puede calcularse recursivamente de la siguiente forma: para $j = 2, 3, \ldots, n$,

    $$f_{S_j}(x) = (1 - q_j) f_{S_{j-1}}(x) + q_j f_{S_{j-1}}(x - C_j).$$

15. Considere el modelo individual de riesgo en donde todas las tasas de muerte q_j son una misma probabilidad q y los montos de las reclamaciones son iguales a 1. Verfique que las fórmula generales de la Proposición 1.2 se reducen a las de la distribución binomial(n, q), es decir,

    a) $E(S) = nq$.

    b) $\text{Var}(S) = nq(1 - q)$.

c) $M_{D_jC_j}(t) = 1 - q + qe^t$.

d) $M_S(t) = (1 - q + qe^t)^n$.

Fórmula de De Pril

16. Demuestre que la fórmula recursiva de De Pril [ii] produce efectivamente una función de probabilidad.

17. A partir de la fórmula de De Pril [ii] y siguiendo la misma notación de dicho enunciado, demuestre que:

    a) $E(S) = nE(X)$.

    b) $\text{Var}(S) = n\text{Var}(X)$.

18. Sean X_1, X_2, X_3, X_4 variables aleatorias independientes con distribución común como aparece en la tabla de abajo. Encuentre la distribución de $X_1 + X_2 + X_3 + X_4$.

j	0	1	2	3
f_j	0.1	0.2	0.3	0.4

19. Sean $X_1, \ldots, X_n$ variables aleatorias independientes con distribución $\text{Ber}(p)$. Mediante la fórmula de De Pril [ii] compruebe que la variable $S = X_1 + \cdots + X_n$ tiene distribución $\text{bin}(n, p)$.

20. Sean X_1 y X_2 dos variables aleatorias independientes con idéntica distribución $\text{Poisson}(\lambda)$. Use la fórmula de De Pril [ii] para demostrar que $X_1 + X_2$ tiene distribución $\text{Poisson}(2\lambda)$.

21. Sean X_1 y X_2 dos variables aleatorias independientes con idéntica distribución $\text{bin}(n, p)$. Use la fórmula de De Pril [ii] para demostrar que $X_1 + X_2$ tiene distribución $\text{bin}(2n, p)$.
    Sugerencia: $\displaystyle\sum_{j=0}^{x} \binom{n}{j}\binom{m}{x-j} = \binom{n+m}{x}$.

22. *Variación de la fórmula de De Pril [ii].* Suponga que las variables aleatorias $X_1, \ldots, X_n$ en la fórmula de De Pril [ii] son estrictamente positivas y toman valores en el conjunto $\{1, 2, \ldots\}$, con $f_1 = P(X =$

$1) \neq 0$. Demuestre que la distribución de $S = X_1 + \cdots + X_n$ ahora es la siguiente:

a) $g_n = (f_1)^n$.

b) $g_{x+n} = \dfrac{1}{f_1} \displaystyle\sum_{j=1}^{x} \left[\frac{j(n+1)}{x} - 1\right] f_{j+1}\, g_{x+n-j}, \quad \text{para } x \geqslant 1.$

Véase el siguiente ejercicio para una extensión de esta fórmula.

23. *Extensión de la fórmula de De Pril [ii].* Suponga que las variables aleatorias $X_1, \ldots, X_n$ son independientes e idénticamente distribuidas pero ahora con valores en el conjunto $\{m, m+1, \ldots\}$ para algún entero $m \geqslant 1$. Sea nuevamente $S = X_1 + \cdots + X_n$, $f_j = P(X = j)$ para $j \geqslant m$, con $f_m \neq 0$, y $g_x = P(S = x)$ para $x \geqslant nm$. Demuestre que:

    a) $g_{nm} = (f_m)^n$.

    b) $g_x = \dfrac{1}{f_m} \displaystyle\sum_{j=1}^{x-nm} \left[\frac{j(n+1)}{x-nm} - 1\right] f_{j+m}\, g_{x-j}, \quad \text{para } x \geqslant nm + 1.$

    Sugerencia: las variables $X_i - m$, para $i = 1, \ldots, n$, ahora ya tienen soporte en el conjunto $\{0, 1, \ldots\}$, aplique entonces la fórmula de De Pril [ii] demostrada y un cambio de variable adecuado.

Modelo colectivo

24. Para el modelo colectivo de riesgo demuestre que

$$\begin{aligned} E(S^3) &= E(N)E(Y^3) + 3E[N(N-1)]E(Y)E(Y^2) \\ &\quad + E[N(N-1)(N-2)]E^3(Y). \end{aligned}$$

25. Suponga que las variables $Y_1, Y_2, \ldots$ en el modelo colectivo de riesgo son discretas con valores en el conjunto $\{1, 2, \ldots\}$. Sea $f_j = P(Y = j)$ para $j \geqslant 1$, $g_x = P(S = x)$ para $x \geqslant 0$, y $p_n = P(N = n)$ para $n \geqslant 0$. Demuestre que:

    a) $g_0 = p_0$.

b) $g_x = \sum\limits_{n=1}^{\infty} f_x^{*n}\, p_n, \quad$ para $x \geqslant 1$.

26. A partir de la fórmula encontrada para $M_S(t)$ en el modelo colectivo de riesgo, encuentre nuevamente las expresiones para $E(S)$, $E(S^2)$ y $\mathrm{Var}(S)$.

27. Considere el modelo colectivo de riesgo $S = \sum_{j=1}^{N} Y_j$, en donde N tiene distribución Poisson(λ) y Y sigue una distribución log normal(m, σ^2). Demuestre que:

    a) $E(Y) = \exp(\sigma^2/2 + m)$.

    b) $E(Y^n) = \exp(nm + n^2\sigma^2/2), \quad n \geqslant 1$.

    c) $\mathrm{Var}(Y) = (e^{\sigma^2} - 1)\exp(\sigma^2 + 2m)$.

    d) $E(S) = \lambda \exp(\sigma^2/2 + m)$.

    e) $\mathrm{Var}(S) = \lambda \exp(2\sigma^2 + 2m)$.

    f) $\alpha_3 := \dfrac{E[(S - E(S))^3]}{[\mathrm{Var}(S)]^{3/2}} = \dfrac{1}{\sqrt{\lambda}} \exp(3\sigma^2/2)$.

28. *Transformada de Laplace-Stieltjes.* La transformada de Laplace-Stieltjes de una variable aleatoria X o de su función de distribución se define como la función

$$l_X(t) = E(e^{-tX}) = \int_{-\infty}^{\infty} e^{-tx}\, dF_X(x).$$

Sea $P_N(t)$ la función generadora de probabilidad de la variable N. Demuestre que para el modelo colectivo de riesgo se cumple la identidad

$$l_S(t) = P_N(l_Y(t)).$$

29. Sea $P_X(t) = E(t^X)$ la función generadora de probabilidad de una variable aleatoria discreta X. Considere un modelo colectivo de riesgo en donde las reclamaciones son discretas con valores en $\{0, 1, \ldots\}$. Suponiendo la existencia de las funciones involucradas, demuestre que se cumple la identidad

$$P_S(t) = P_N(P_X(t)).$$

Modelo binomial compuesto

30. Verifique la validez de las fórmulas para el modelo binomial compuesto que aparecen en la Proposición 1.6 de la página 23.

31. Para el modelo binomial compuesto, demuestre las siguientes fórmulas:

    a) $E(S^3) = n(n-1)p^2\mu((n-2)p\mu^2 + 3\mu_2) + np\mu_3$.

    b) $E[(S - E(S))^3] = np(\mu_3 - 3p\mu\mu_2 + 2p^2\mu^3)$.

    c) $\alpha_3 := \dfrac{E[(S-E(S))^3]}{[\operatorname{Var}(S)]^{3/2}} > 0$.

32. Sean S_1 y S_2 dos riesgos independientes con distribución binomial compuesta con parámetros $(n_1, p; G)$ y $(n_2, p; G)$ respectivamente. Demuestre que el riesgo $S_1 + S_2$ también sigue una distribución binomial compuesta con parámetros $(n_1 + n_2, p; G)$.

Modelo binomial negativo compuesto

33. Verifique la validez de las fórmulas para el modelo binomial negativo compuesto de la Proposición 1.7 en la página 24.

34. Para el modelo binomial negativo compuesto, demuestre las siguientes fórmulas:

    a) $$E(S^3) = k(k+1)(k+2)(1/p-1)^3\mu^3 + 3k(k+1)(1/p-1)^2\mu\mu_2 + k(1/p-1)\mu_3.$$

    b) $$E[(S-E(S))^3] = k(1/p-1)\mu_3 + 3k(1/p-1)^2\mu\mu_2 + 2k(1/p-1)^3\mu^3.$$

    c) $\alpha_3 := \dfrac{E[(S-E(S))^3]}{[\operatorname{Var}(S)]^{3/2}} > 0$.

35. Sea N con distribución bin neg(k,p). Demuestre que:

    a) $P(N = 0) = p^k$.

    b) $P(N = n+1) = \dfrac{k+n}{n+1}\,(1-p)\,P(N = n), \quad \text{para } n \geqslant 0.$

Modelo Poisson compuesto

36. Verifique la validez de las fórmulas para el modelo Poisson compuesto que aparecen en la Proposición 1.8 de la página 25.

37. Para el modelo Poisson compuesto, demuestre las siguientes fórmulas:

    a) $E(S^3) = \lambda\mu_3 + 3\lambda^2\mu_2\mu + \lambda^3\mu^3$.

    b) $E[(S - E(S))^3] = \lambda\mu_3$.

    c) $\alpha_3 := \dfrac{E[(S-E(S))^3]}{[\mathrm{Var}(S)]^{3/2}} = \dfrac{\mu_3}{\sqrt{\lambda\mu_2^3}} > 0$.

38. Demuestre las fórmulas para el modelo Poisson compuesto asociado de la página 26.

39. Sean $F_1(x)$ y $F_2(x)$ dos funciones de distribución con funciones generadoras de momentos $M_1(t)$ y $M_2(t)$ respectivamente. Demuestre que para cualquier $\alpha \in [0,1]$, la función $\alpha F_1(x) + (1-\alpha)F_2(x)$ es una función de distribución cuya función generadora de momentos asociada es $\alpha M_1(t) + (1-\alpha)M_2(t)$. Este resultado fue utilizado en el análisis de la suma de dos riesgos con distribución Poisson compuesta.

40. Sean $S_1, \ldots, S_n$ riesgos independientes con distribución Poisson compuesta con parámetros $\lambda_1, \ldots, \lambda_n$, respectivamente. Suponga que los montos de las reclamaciones de estos riesgos son $Y^{(1)}, \ldots, Y^{(n)}$, con función de distribución $G_1(x), \ldots, G_n(x)$, respectivamente. Demuestre que el riesgo $S = S_1 + \cdots + S_n$ también sigue una distribución Poisson compuesta con parámetro $\lambda = \lambda_1 + \cdots + \lambda_n$, y la función de distribución de las reclamaciones es

$$G(x) = \frac{\lambda_1}{\lambda} G_1(x) + \cdots + \frac{\lambda_n}{\lambda} G_n(x).$$

41. Sean S_1 y S_2 dos riesgos independientes, el primero con distribución Poisson comp(λ_1, F_1) con $\lambda_1 = 50$, y el segundo con distribución Poisson comp(λ_2, F_2) con $\lambda_2 = 100$, en donde $F_1(x) = \min\{x, 1\}$ para $x \geqslant 0$, y $F_2(x) = 1 - e^{-x}$ para $x \geqslant 0$. Encuentre la distribución

de $S = S_1 + S_2$ y la función de distribución de las reclamaciones del riesgo S.

42. Sean S_1 y S_2 dos riesgos independientes con distribución Poisson compuesta, el primero Poisson comp(λ_1, F_1) con $\lambda_1 = 10$, y el segundo Poisson comp(λ_2, F_2) con $\lambda_2 = 20$. Suponga que las reclamaciones de ambos riesgos son de magnitud 50 o 100, y por lo tanto tienen la siguiente función de distribución:

$$F(x) = \begin{cases} 0 & \text{si } x < 50, \\ p & \text{si } 50 \leqslant x < 100, \\ 1 & \text{si } x \geqslant 100. \end{cases}$$

Suponga que el parámetro p es igual a $1/2$ para el primer riesgo y es $1/3$ para el segundo riesgo. Encuentre la distribución de $S = S_1 + S_2$ y la función de distribución de las reclamaciones del riesgo S.

43. Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias independientes cada una de ellas con distribución Ber(p) y sea $X_0 = 0$. Sea N otra variable aleatoria con distribución Poisson(λ) independiente de las anteriores. Defina la variable

$$X = \sum_{i=0}^{N} X_i.$$

Demuestre que X tiene distribución Poisson(λp). Esta variable tiene la siguiente interpretación: si N representa el total de siniestros ocurridos y cada siniestro es reportado con probabilidad p, entonces X representa el total de siniestros ocurridos y reportados.

44. Sea Y una variable aleatoria con función de distribución $F(y)$, y sean $a < b$ dos números tales que $F(a) < F(b)$. Demuestre que la función de distribución condicional de Y dado el evento $(Y \in (a, b])$ es

$$F(y \,|\, Y \in (a,b]) = \begin{cases} 0 & \text{si } y < a, \\ \dfrac{F(y) - F(a)}{F(b) - F(a)} & \text{si } a \leqslant y \leqslant b, \\ 1 & \text{si } y > b. \end{cases}$$

Aplique este resultado al caso cuando Y tiene distribución $\exp(\alpha)$. Encuentre y grafique ambas funciones de distribución: la original y la condicional.

Modelo Poisson compuesto mixto

45. Verifique la validez de las fórmulas de la Proposición 1.11 de la página 32.

46. Para el modelo Poisson compuesto mixto, demuestre que:

    a) $E(S^3) = E(\Lambda)\mu_3 + 3E(\Lambda^2)\mu_2\mu + E(\Lambda^3)\mu^3$.

    b) $E[(S - E(S))^3] = E[(\Lambda - E(\Lambda))^3]\mu^3 + 3\mathrm{Var}(\Lambda)\mu_2\mu + E(\Lambda)\mu_3$.

Esperanza condicional

47. Sea (X, Y) un vector aleatorio discreto con la función de probabilidad

$$f_{X,Y}(x, y) = \begin{cases} (x + y)/36 & \text{si } x, y = 1, 2, 3, \\ 0 & \text{en otro caso.} \end{cases}$$

    a) Encuentre la distribución de la variable aleatoria $E(X \mid Y)$.

    b) Compruebe que $E(E(X \mid Y)) = E(X) = 78/36$.

Capítulo 2

Fórmula de Panjer y algunos métodos de aproximación

En este capítulo se presenta la famosa fórmula de Panjer. Este resultado proporciona una expresión exacta, aunque recursiva, de la distribución de probabilidad de un riesgo en el modelo colectivo y es válida cuando la distribución del número de reclamaciones y los montos cumplen ciertas condiciones. Se presentan además algunos métodos de aproximación con validez general para estimar la distribución de un riesgo. Estos métodos generales de aproximación pueden ser útiles cuando no se cumplen las condiciones requeridas para aplicar la fórmula de Panjer.

2.1. Fórmula de Panjer

Primeramente se enuncia la condición que debe satisfacer el número de reclamaciones para obtener la fórmula de Panjer.

Proposición 2.1 *Sea N una variable aleatoria discreta con valores en $\{0,1,\ldots\}$ y sea $p_k = P(N = k)$ para $k = 0,1,\ldots$ Sean a y b dos constante. Entonces la igualdad*

$$p_k = (a + \frac{b}{k})\, p_{k-1}, \qquad k \geqslant 1, \tag{2.1}$$

se cumple cuando

1. *N es bin(n,p), con $a = -p/(1-p)$ y $b = (n+1)p/(1-p)$.*
2. *N es Poisson(λ), con $a = 0$ y $b = \lambda$.*
3. *N es bin neg(r,p), con $a = 1-p$ y $b = (r-1)(1-p)$.*

La demostración de este resultado es inmediata después de realizar algunos cálculos algebraicos sencillos y se dejan como ejercicio al lector.

Toda distribución de probabilidad con soporte en $\{0,1,\ldots\}$ que cumple la identidad (2.1) se le llama distribución de clase $(a,b,0)$, los términos a y b se refieren a las constantes del mismo nombre que aparecen en la fórmula (2.1) y el cero se refiere a que la probabilidad de inicio de la fórmula recursiva es aquella que tiene subíndice cero, es decir, p_0. Observe que la identidad (2.1) es muy atractiva, pues permite generar la distribución de probabilidad de estas variables aleatorias discretas de una forma recursiva: se calcula primero p_0, a partir de ella se obtiene p_1, a partir de p_1 se obtiene p_2, y así sucesivamente. Supondremos entonces que la distribución del número de reclamaciones cumple con la condición (2.1) y la proposición establece que tal condición es válida para las tres distribuciones señaladas. En el ejercicio 48 en la página 61 se pide demostrar el resultado recíproco de la proposición anterior, es decir, que las únicas distribuciones discretas de probabilidad no degeneradas que cumplen (2.1) son las tres mencionadas.

Recordando la notación e hipótesis del modelo colectivo de riesgo $S = \sum_{j=1}^{N} Y_j$, definido en el capítulo anterior, tenemos las siguientes hipótesis y notación adicionales: supondremos que las reclamaciones Y_j son tales que

$P(Y_j \in \mathbb{N}) = 1$, lo cual no es ningún problema pues puede considerarse que las reclamaciones se efectúan en unidades monetarias, cualesquiera que éstas sean. En los cálculos que haremos a continuación usaremos los siguientes símbolos:

Notación

$$
\begin{array}{ll}
p_k = P(N = k) & k = 0, 1, \ldots \\
f_r = P(Y = r) & r = 1, 2, \ldots \\
f_r^{*k} = P(Y_1 + \cdots + Y_k = r) & 1 \leqslant k \leqslant r = 1, 2, \ldots \\
g_r = P(S = r) & r = 0, 1, \ldots
\end{array}
$$

En particular, para $1 \leqslant k + 1 \leqslant r$,

$$f_r^{*(k+1)} = (f^{*k} * f)_r = \sum_{i=1}^{r-1} f_i^{*k} f_{r-i}.$$

Además, $g_0 = P(S = 0) = P(N = 0) = p_0$ y para $r \geqslant 1$,

$$g_r = P(S = r) = \sum_{k=1}^{\infty} P(S = r \,|\, N = k) P(N = k) = \sum_{k=1}^{\infty} f_r^{*k} p_k.$$

Proposición 2.2 *Bajo la notación e hipótesis anteriores, se cumplen las siguientes igualdades:*

1. $E(\ Y_1 \mid \sum_{j=1}^{k} Y_j = r\ ) = \dfrac{r}{k}$, *para* $k \geqslant 1$.

2. $p_k\, f_r^{*k} = p_{k-1} \sum_{i=1}^{r-1} (a + \dfrac{bi}{r})\, f_{r-i}^{*(k-1)}\, f_i$, *para* $k \geqslant 2$.

Demostración. Para el primer inciso, por idéntica distribución,

$$\begin{aligned} E(\, Y_1 \mid \sum_{j=1}^{k} Y_j = r \,) &= \frac{1}{k} \sum_{i=1}^{k} E(\, Y_i \mid \sum_{j=1}^{k} Y_j = r \,) \\ &= \frac{1}{k} E(\, \sum_{i=1}^{k} Y_i \mid \sum_{j=1}^{k} Y_j = r \,) \\ &= \frac{r}{k}. \end{aligned}$$

Para el segundo inciso desarrollamos el lado derecho,

$$\begin{aligned} p_{k-1} \sum_{i=1}^{r-1} (a + \frac{bi}{r}) f_{r-i}^{*(k-1)} f_i &= p_{k-1} \sum_{i=1}^{r} (a + \frac{bi}{r}) P(Y_2 + \cdots + Y_k = r - i) P(Y_1 = i) \\ &= p_{k-1} \sum_{i=1}^{r} (a + \frac{bi}{r}) P(Y_1 = i, Y_2 + \cdots + Y_k = r - i) \\ &= p_{k-1} \sum_{i=1}^{r} (a + \frac{bi}{r}) P(Y_1 = i, \sum_{j=1}^{k} Y_j = r) \\ &= p_{k-1} \sum_{i=1}^{r} (a + \frac{bi}{r}) P(Y_1 = i \mid \sum_{j=1}^{k} Y_j = r) f_r^{*k} \\ &= p_{k-1} E(a + \frac{bY_1}{r} \mid \sum_{j=1}^{k} Y_j = r) f_r^{*k} \\ &= p_{k-1} (a + \frac{b}{k}) f_r^{*k} \\ &= p_k f_r^{*k} \end{aligned}$$

■

Ahora estamos listos para enunciar y demostrar la fórmula de Harry Panjer [27], publicada en 1981.

Teorema 2.1 (Fórmula de Panjer) *Para un riesgo S que sigue el modelo colectivo en donde las reclamaciones Y son de la clase $(a,b,0)$, la probabilidad $g_r = P(S=r)$ está dada por*

$$\begin{aligned} g_r &= \sum_{i=1}^{r} (a+\frac{bi}{r})\, f_i\, g_{r-i}, \qquad \textit{para } r \geqslant 1. \\ g_0 &= p_0, \end{aligned}$$

en donde $p_0 = P(N=0)$ y $f_i = P(Y=i)$ para $i \geqslant 1$.

Demostración. Hemos observado antes que $g_0 = P(N=0) = p_0$. Para el caso $r \geqslant 1$,

$$\begin{aligned} g_r &= \sum_{k=1}^{\infty} P(S=r \,|\, N=k) P(N=k) \\ &= \sum_{k=1}^{\infty} p_k\, f_r^{*k} \\ &= p_1\, f_r + \sum_{k=2}^{\infty} p_k\, f_r^{*k} \\ &= (a+b)\, p_0\, f_r + \sum_{k=2}^{\infty} \sum_{i=1}^{r-1} (a+\frac{bi}{r})\, p_{k-1}\, f_{r-i}^{*(k-1)}\, f_i \\ &= (a+b)\, p_0\, f_r + \sum_{i=1}^{r-1} (a+\frac{bi}{r}) f_i \sum_{k=2}^{\infty} p_{k-1}\, f_{r-i}^{*(k-1)} \\ &= (a+b)\, p_0\, f_r + \sum_{i=1}^{r-1} (a+\frac{bi}{r})\, f_i\, g_{r-i} \\ &= \sum_{i=1}^{r} (a+\frac{bi}{r})\, f_i\, g_{r-i}. \end{aligned}$$

■

Observe que cuando las reclamaciones son constantes y unitarias, es decir, $f_1 = P(Y = 1) = 1$, entonces $S = N$ y la fórmula de Panjer se reduce a la fórmula recursiva (2.1) de las distribuciones $(a, b, 0)$. A continuación escribimos explícitamente los primeros términos de la fórmula recursiva de Panjer:

$$\begin{aligned} g_0 &= p_0 = P(N=0) \\ g_1 &= (a + \frac{b}{1}) f_1 g_0 \\ g_2 &= (a + \frac{b}{2}) f_1 g_1 + (a + \frac{2b}{2}) f_2 g_0 \\ g_3 &= (a + \frac{b}{3}) f_1 g_2 + (a + \frac{2b}{3}) f_2 g_1 + (a + \frac{3b}{3}) f_3 g_0 \\ &\vdots \end{aligned}$$

Ejemplo 2.1 *Consideremos el caso cuando N sigue una distribución Poisson de parámetro $\lambda = 3.5$ y el monto de las reclamaciones tiene la siguiente función de densidad:*

r	*1*	*2*	*3*	*4*	*5*
f_r	*0.1*	*0.1*	*0.2*	*0.3*	*0.3*

Entonces la fórmula de Panjer produce la función de probabilidad para S que se muestra en la Figura 2.1. El código en R correspondiente se encuentra en el apéndice.

Aproximación en el caso de reclamaciones continuas

Cuando los montos de las reclamaciones toman valores continuos, puede usarse el siguiente método de discretización de estos montos para poder aplicar la fórmula de Panjer. Se toma cualquier unidad monetaria $\rho > 0$ y se definen las variables aleatorias enteras:

$$\begin{aligned} \overline{Y}_j &= \text{ínf}\,\{ n \in \mathbb{N} : Y_j \leqslant n\rho \}, \\ \text{y} \quad \underline{Y}_j &= \sup\,\{ n \in \mathbb{N} : Y_j \geqslant n\rho \}, \end{aligned}$$

en donde también se define $\text{ínf}\,\emptyset = 0$. Entonces,

$$\rho \underline{Y}_j \leqslant Y_j \leqslant \rho \overline{Y}_j.$$

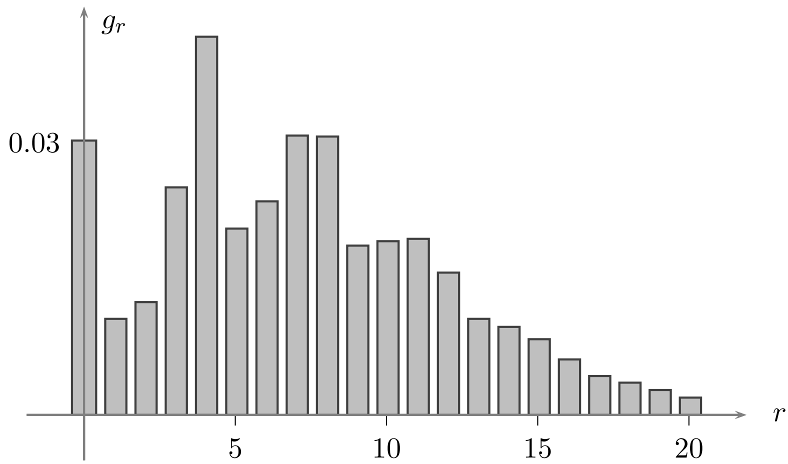


Figura 2.1

Por ejemplo, para la situación que se muestra en la Figura 2.2 se tiene que $\overline{Y}_j = 4$ y $\underline{Y}_j = 3$, y efectivamente se cumple que $3\rho \leqslant Y_j \leqslant 4\rho$.

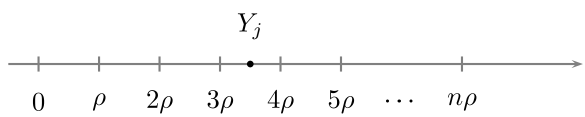


Figura 2.2

Se definen entonces los siguientes riesgos cuyas reclamaciones ahora son enteras:

$$\overline{S} = \sum_{j=1}^{N} \overline{Y}_j, \qquad \text{y} \qquad \underline{S} = \sum_{j=1}^{N} \underline{Y}_j.$$

Entonces se cumple que

$$\rho \underline{S} \leqslant S \leqslant \rho \overline{S}, \tag{2.2}$$

de donde obtendremos el siguiente resultado.

Proposición 2.3 *Para cualquier $x > 0$,*

$$P(\overline{S} \leqslant x/\rho) \leqslant P(S \leqslant x) \leqslant P(\underline{S} \leqslant x/\rho).$$

Demostración. Por (2.2) se cumple la contención de eventos

$$(S \leqslant x) \subseteq (\rho\underline{S} \leqslant x),$$

entonces se tiene que $P(S \leqslant x) \leqslant P(\rho\underline{S} \leqslant x)$ y por lo tanto,

$$P(S \leqslant x) \leqslant P(\underline{S} \leqslant x/\rho).$$

Para la segunda desigualdad, se observa la contención de los eventos

$$(\rho\overline{S} \leqslant x) \subseteq (S \leqslant x),$$

y se procede de manera análoga. ■

Esto provee de cotas superior e inferior, calculadas usando la fórmula de Panjer, para la función de distribución del riesgo. Conforme más pequeña sea la unidad monetaria ρ mejor es la aproximación. Debe observarse, sin embargo, que surge una dificultad técnica para aquellas reclamaciones Y_j con valores en $(0, \rho)$, pues estas reclamaciones llevan a la definición $\underline{Y}_j = 0$, lo cual no es un caso contemplado en el esquema de la fórmula de Panjer. En una situación real, el monto de las reclamaciones es grande comparado con el valor del parámetro ρ, de modo que la probabilidad de que una reclamación tome un valor entre 0 y ρ es realmente muy pequeña. Existe también un caso particular simple que es cuando $\rho = 1$, es decir, se aproximan las reclamaciones mediante valores enteros.

Ejemplo 2.2 (Aproximación con reclamaciones enteras) *Para cada valor de una reclamación Y_j existe un entero $n \geqslant 0$ tal que $n \leqslant Y_j < n + 1$, y por lo tanto,*

$$\underline{Y}_j = n, \qquad y \qquad \overline{Y}_j = n + 1.$$

Entonces se tienen nuevamente las relaciones $\underline{S} \leqslant S \leqslant \overline{S}$ y en consecuencia para cualquier $x > 0$,

$$P(\overline{S} \leqslant x) \leqslant P(S \leqslant x) \leqslant P(\underline{S} \leqslant x),$$

en donde $\overline{S}$ y $\underline{S}$ son riesgos con reclamaciones enteras para los cuales puede aplicarse la fórmula de Panjer y obtener su distribución (exceptuando un pequeño error obtenido por el caso $\underline{Y}_j = 0$). En cualquier caso podemos ser prudentes y tomar las reclamaciones de manera sobreestimada: para $n \geqslant 0$,

$$\overline{Y}_j = n + 1 \quad \textit{cuando} \quad Y \in (n, n+1].$$

De esta forma cualquier valor continuo de una reclamación en el intervalo $(n, n+1]$ se considera como si fuera de magnitud $n+1$. Por lo tanto, los montos de las reclamaciones están siendo ligeramente sobrevaluadas: si $G(x)$ denota la función de distribución de una reclamación Y cualquiera, entonces

$$P(\overline{Y}_j = n + 1) = G(n + 1) - G(n), \quad n = 0, 1, 2, \ldots$$

En las siguientes secciones estudiaremos algunos métodos generales para aproximar la distribución de probabilidad de un riesgo en el modelo colectivo. Estas aproximaciones son muy generales y no presuponen el cumplimiento de las hipótesis para la validez de la fórmula de Panjer, es decir, el número de reclamaciones no necesariamente tiene una distribución en la clase $(a, b, 0)$, ni el monto de las reclamaciones es necesariamente discreto. Por otro lado, el problema de estimar el error en estas aproximaciones es muy general y no nos ocuparemos de ello.

2.2. Aproximación normal

Si la distribución de probabilidad del número de reclamaciones N se concentra mayormente en valores grandes, entonces el teorema central del límite sugiere aproximar la distribución del riesgo S mediante la distribución normal. Suponga que la esperanza de S es m y la varianza es σ^2. Entonces, para $x > 0$,

$$\begin{aligned} P(S \leqslant x) &= P\Big(\frac{S-m}{\sigma} \leqslant \frac{x-m}{\sigma}\Big) \\ &\approx \Phi\Big(\frac{x-m}{\sigma}\Big). \end{aligned}$$

Derivando esta expresión se encuentra una fórmula aproximada para la función de densidad de S.

Proposición 2.4 (Aproximación normal) *Sea S es una riesgo con media m y varianza finita σ^2. Para cualquier $x > 0$,*

$$f_S(x) \approx \frac{1}{\sigma}\,\phi(\frac{x-m}{\sigma}). \tag{2.3}$$

Esta aproximación hace uso únicamente de la media y la varianza del riesgo, y en general no es una buena aproximación a menos que la densidad del riesgo presente la forma de campana o bien el número de reclamaciones tenga una distribución tal que sea razonable aplicar el teorema central del límite, esto es, cuando la distribución de probabilidad se concentre en un conjunto de valores grandes de la variable aleatoria. Por ejemplo, dada cualquier $\epsilon > 0$ suficientemente pequeña, la aproximación podría ser buena si los parámetros de la distribución de N son tales que

$$P(N \geqslant 30) \geqslant 1 - \epsilon.$$

Se puede particularizar la aproximación normal cuando la distribución de N es conocida, por ejemplo, cuando el número de reclamaciones N sigue una distribución Poisson con parámetro λ, la aproximación (2.3) adquiere la expresión

$$f_S(x) \approx \frac{1}{\sqrt{\lambda\mu_2}}\,\phi(\frac{x-\lambda\mu}{\sqrt{\lambda\mu_2}}).$$

Otro caso particular se obtiene cuando N es bin(n,p), entonces la expresión (2.3) se reduce a la fórmula siguiente

$$P(S \leqslant x) \approx \Phi(\frac{x - np\mu}{\sqrt{np(\mu_2 - \mu^2 p)}}).$$

En el caso cuando N es bin neg(r,p) se tiene que (2.3) es

$$P(S \leqslant x) \approx \Phi(\frac{x - r(1-p)\mu/p}{\sqrt{r(1-p)[\mu_2/p + (1-p)\mu^2/p^2]}}).$$

2.3. Aproximación gama trasladada

En algunos casos el histograma creado a partir de las observaciones históricas de un riesgo puede presentar un aspecto semejante a la forma de la distribución gama. Esta semejanza sugiere aproximar la distribución del riesgo S por la distribución de la variable aleatoria

$$k + Z,$$

en donde k es una constante y Z es una variable aleatoria con distribución gama(γ, α). En los ejemplos concretos que hemos presentado en donde se ha calculado la distribución exacta del riesgo, usando la fórmula de De Pril o la fórmula de Panjer, puede constatarse que la distribución de algunos riesgos tienen algún parecido con la distribución gama. Es por ello que se propone este método de aproximación. Para llevar a cabo este procedimiento se deben escoger adecuadamente valores para los tres parámetros k, γ y α, que determinan la distribución de $k + Z$. Supongamos entonces conocidas o estimadas las siguientes cantidades:

a) $E(S) = m$.

b) $\text{Var}(S) = \sigma^2$.

c) $\dfrac{E[(S - E(S))^3]}{[\text{Var}(S)]^{3/2}} = \alpha_3$.

La correspondiente media, varianza y coeficiente de asimetría (de Fisher) de la variable aleatoria aproximante $k + Z$ son:

a) $E(k + Z) = k + \gamma/\alpha$.

b) $\text{Var}(k + Z) = \gamma/\alpha^2$.

c) $\dfrac{E[(k + Z - E(k + Z))^3]}{[\text{Var}(k + Z)]^{3/2}} = 2/\sqrt{\gamma}$.

La idea es hacer que las distribuciones de S y $k + Z$ coincidan en el sentido de que las tres cantidades mencionadas sean las mismas para las dos distribuciones. Haciendo coincidir estas cantidades se obtiene el sistema de ecuaciones

$$k + \frac{\gamma}{\alpha} = m, \qquad \frac{\gamma}{\alpha^2} = \sigma^2, \qquad \frac{2}{\sqrt{\gamma}} = \alpha_3,$$

cuya solución es

$$k = m - \frac{2\sigma}{\alpha_3}, \qquad \gamma = \frac{4}{\alpha_3^2}, \qquad \alpha = \frac{2}{\sigma\alpha_3}.$$

De esta forma se tiene la siguiente aproximación.

Proposición 2.5 (Aproximación gama trasladada) *La distribución del riesgo S en el modelo colectivo puede aproximarse mediante la distribución de la variable aleatoria*

$$m - \frac{2\sigma}{\alpha_3} + Z,$$

en donde Z se distribuye gama$(\frac{4}{\alpha_3^2}, \frac{2}{\sigma\alpha_3})$.

Pueden sustituirse las expresiones generales para la media, varianza y coeficiente de asimetría de un riesgo que sigue el modelo colectivo para obtener fórmulas un poco más particulares de esta aproximación. Debe hacerse notar que la aproximación gama es en esencia la aplicación del método de momentos para la estimación de parámetros y naturalmente el método puede aplicarse a cualquiera otra distribución de probabilidad conocida que tenga alguna semejanza o parecido con la distribución del riesgo en estudio.

2.4. Aproximación de Edgeworth

Considere un cierto riesgo S modelado mediante una variable aleatoria con esperanza m, varianza σ^2 y tal que su función generadora de momentos existe. Defina la variable $Z = (S - m)/\sigma$, cuya esperanza es cero y varianza es 1. Sea $M_Z(r)$ la función generadora de momentos de Z. La serie de Taylor de la función $\ln M_Z(r)$ alrededor de cero es

$$\ln M_Z(r) = a_0 + a_1 r + \frac{a_2}{2!} r^2 + \frac{a_3}{3!} r^3 + \frac{a_4}{4!} r^4 + \cdots ,$$

en donde los coeficientes son $a_k = \frac{d^k}{dr^k} \ln M_Z(r)\Big|_{r=0}$. Calculando las derivadas y evaluando en cero se encuentra que los primeros cinco coeficientes son:

$$\begin{aligned} a_0 &= 0, \\ a_1 &= E(Z) = 0, \\ a_2 &= E(Z^2) = 1, \\ a_3 &= E(Z^3), \\ a_4 &= E(Z^4) - 3, \\ &\vdots \end{aligned}$$

La aproximación de Edgeworth consiste en truncar la serie de Taylor de la función $\ln M_Z(r)$ hasta algún término adecuado. Por ejemplo, la aproximación hasta la cuarta potencia de r es

$$\ln M_Z(r) \approx \frac{1}{2!}r^2 + \frac{a_3}{3!}r^3 + \frac{a_4}{4!}r^4.$$

Entonces

$$\begin{aligned} M_Z(r) &\approx \exp(\frac{1}{2!}r^2 + \frac{a_3}{3!}r^3 + \frac{a_4}{4!}r^4) \\ &= e^{r^2/2} \exp(\frac{a_3}{6}r^3 + \frac{a_4}{24}r^4). \end{aligned}$$

Ahora se usa la serie $e^x = 1 + x + x^2/2! + x^3/3! + \cdots$ en el segundo factor y se obtiene la aproximación

$$\begin{aligned} M_Z(r) &\approx e^{r^2/2}(1 + \frac{a_3}{6}r^3 + \frac{a_4}{24}r^4 + \frac{a_3^2}{72}r^6) \\ &= e^{r^2/2} + \frac{a_3}{6}r^3 e^{r^2/2} + \frac{a_4}{24}r^4 e^{r^2/2} + \frac{a_3^2}{72}r^6 e^{r^2/2}. \end{aligned} \tag{2.4}$$

El siguiente paso es *invertir* cada término de esta ecuación encontrando una distribución aproximada para Z. El resultado que utilizaremos es el siguiente.

Proposición 2.6 *Si $\phi(x)$ es la función de densidad de la distribución normal estándar, entonces para cualquier entero $n \geqslant 0$,*

$$\int_{-\infty}^{\infty} e^{rx}(-1)^n \phi^{(n)}(x)\, dx = r^n e^{r^2/2}.$$

Demostración. Primeramente tenemos que para $n = 0$,

$$\int_{-\infty}^{\infty} e^{rx} \phi(x)\, dx = e^{r^2/2},$$

es decir, $e^{r^2/2}$ es la función generadora de momentos de la distribución normal estándar. Multiplicando por r tenemos que

$$\begin{aligned} re^{r^2/2} &= \int_{-\infty}^{\infty} re^{rx}\phi(x)\, dx \\ &= \int_{-\infty}^{\infty} (\frac{d}{dx}e^{rx})\phi(x)\, dx \\ &= -\int_{-\infty}^{\infty} e^{rx}\phi'(x)\, dx. \end{aligned}$$

Es decir, $re^{r^2/2}$ es la transformada de Laplace de $-\phi'(x)$. Procediendo de manera análoga, multiplicando sucesivamente por r, se llega a la fórmula anunciada. ■

En particular, se ha demostrado que $re^{r^2/2}$ es la transformada de Laplace de $-\phi'(x)$. En este caso usamos el término transformada de Laplace y no función generadora de momentos, pues la función $-\phi'(x)$ no es una función de densidad. Entonces el resultado anterior establece que la función $r^n e^{r^2/2}$ es la transformada de Laplace de $(-1)^n\phi^{(n)}(x)$. El siguiente paso es *invertir* cada uno de los términos de la igualdad (2.4), aunque realmente que no se está calculando de manera formal la inversa de la función generadora de momentos (o transformada de Laplace) sino que se está usando el hecho de que si dos distribuciones de probabilidad tienen la misma función generadora de momentos, entonces las distribuciones coinciden. Así, *invirtiendo* término a término la igualdad (2.4), se obtiene

$$f_Z(z) \approx \phi(z) - \frac{a_3}{6}\,\phi^{(3)}(z) + \frac{a_4}{24}\,\phi^{(4)}(z) + \frac{a_3^2}{72}\,\phi^{(6)}(z). \tag{2.5}$$

Recordemos que hemos definido $Z = (S - m)/\sigma$, por lo tanto la función de densidad de $S = m + \sigma Z$ está dada por

$$f_S(x) = \frac{1}{\sigma}\,f_Z(\frac{x-m}{\sigma}),$$

y de esta forma se llega al siguiente resultado:

Proposición 2.7 (Aproximación de Edgeworth) *Sea S un riesgo con media m, varianza σ^2 y cuya función generadora de momentos existe. Entonces la función de densidad de S puede aproximarse de la siguiente forma:*

$$f_S(x) \approx \frac{1}{\sigma}\,[\,\phi(\frac{x-m}{\sigma}) - \frac{a_3}{6}\,\phi^{(3)}(\frac{x-m}{\sigma}) + \frac{a_4}{24}\,\phi^{(4)}(\frac{x-m}{\sigma}) + \frac{a_3^2}{72}\,\phi^{(6)}(\frac{x-m}{\sigma})\,].$$

Derivando directamente la función de densidad $\phi(x)$ de la distribución normal estándar puede demostrarse que

$$\begin{aligned} \phi^{(3)}(x) &= (3x - x^3)\,\phi(x), \\ \phi^{(4)}(x) &= (3 - 6x^2 + x^4)\,\phi(x), \\ \phi^{(6)}(x) &= (-15 + 45x^2 - 15x^4 + x^6)\,\phi(x). \end{aligned}$$

Estas expresiones pueden sustituirse en la aproximación de Edgeworth para obtener una expresión en términos de únicamente la función $\phi(x)$. Observe que el primer sumando en la aproximación de Edgeworth corresponde a la función de densidad normal con media m y varianza σ^2. No es difícil verificar además que cuando el riesgo S sigue una distribución normal, la aproximación de Edgeworth es exacta, pues produce como resultado la misma distribución con los mismos parámetros.

La aproximación de Edgeworth puede expresarse también en términos de la función de distribución de la siguiente forma: integrando (2.5) se obtiene

$$F_Z(z) \approx \Phi(z) - \frac{a_3}{6}\Phi^{(3)}(z) + \frac{a_4}{24}\Phi^{(4)}(z) + \frac{a_3^2}{72}\Phi^{(6)}(z),$$

considerando la identidad $F_S(x) = F_Z((x-m)/\sigma)$, se tiene que la aproximación de Edgeworth para un riesgo S en términos de la función de distribución es

$$F_S(x) \approx \Phi(\frac{x-m}{\sigma}) - \frac{a_3}{6}\Phi^{(3)}(\frac{x-m}{\sigma}) + \frac{a_4}{24}\Phi^{(4)}(\frac{x-m}{\sigma}) + \frac{a_3^2}{72}\Phi^{(6)}(\frac{x-m}{\sigma}),$$

en donde derivando directamente la función $\Phi(z)$ puede demostrarse que

$$\begin{aligned} \Phi^{(3)}(z) &= (z^2-1)\,\phi(z), \\ \Phi^{(4)}(z) &= (-z^3+3z)\,\phi(z), \\ \Phi^{(6)}(z) &= (-z^5+10z^3-15z)\,\phi(z). \end{aligned}$$

Comentarios y referencias

Hemos dedicado este pequeño capítulo a la presentación de la fórmula de Panjer. Dentro de su ámbito de aplicación, esta importante fórmula proporciona un mecanismo recursivo para calcular de manera exacta la distribución de probabilidad de un riesgo que sigue el modelo colectivo. Debe tenerse cuidado en la implementación en computadora de esta fórmula, pues podrían presentarse problemas numéricos por cuestiones de redondeo o baja precisión en los cálculos. Cuando las hipótesis para aplicar la fórmula de Panjer no se cumplen, se cuenta con algunos métodos para aproximar la distribución del riesgo y hemos presentado sólo unos pocos de ellos, dejando de lado el problema general de establecer condiciones bajo las cuales los métodos de aproximación mencionados son adecuados. El material de este capítulo está basado en las notas de Schmidli [34]. Otras referencias en el tema son: Dickson [13] y Rolski *et al.* [32].

2.5. Ejercicios

Distribuciones de la clase $(a, b, 0)$

48. *Las cuatro distribuciones de la clase* $(a, b, 0)$. Sea $\{p_k : k \geqslant 0\}$ una distribución de probabilidad en la clase $(a, b, 0)$, es decir, se trata de una distribución discreta con soporte el conjunto $\{0, 1, \ldots\}$ y tal que cumple la relación $p_k = (a + b/k)\, p_{k-1}$ para $k \geqslant 1$, en donde a y b son dos constantes. Observe que tomando el caso particular cuando $k = 1$, se llega a la conclusión de que las constantes deben satisfacer la desigualdad $a + b \geqslant 0$.

    a) Demuestre que en el caso $a + b = 0$, la distribución se concentra en cero, es decir, $p_0 = 1$.

    b) Demuestre que si $a + b > 0$ y $a = 0$ entonces $\{p_k : k \geqslant 0\}$ es la distribución Poisson(λ) con $\lambda = b > 0$.

    c) Demuestre que si $a + b > 0$ y $a > 0$ entonces para cualquier $k \geqslant 1$,

    $$p_k = \frac{a^k}{k!}(k + \frac{b}{a})(k - 1 + \frac{b}{a}) \cdots (2 + \frac{b}{a})(1 + \frac{b}{a})p_0. \qquad (2.6)$$

    Defina $r = 1 + b/a$ e incorpore este valor en (2.6). Concluya que $\{p_k : k \geqslant 0\}$ es la distribución bin neg(r, p) con $p = 1 - a$.

    d) Observe que si $a + b > 0$ y $a < 0$, y si la relación iterativa es válida para cualquier $k \geqslant 1$ como hemos supuesto, entonces necesariamente los valores de las constantes a y b deben ser tales que exista un entero $n \geqslant 1$ tal que

    $$a + \frac{b}{n} > 0 \qquad \text{y} \qquad a + \frac{b}{n+1} = 0.$$

    Y por lo tanto se debe tener que $p_k = 0$ para $k = n+1, n+2, \ldots$ De la igualdad anterior obtenga $n = -b/a - 1$. Compruebe la validez de la ecuación (2.6) para $0 \leqslant k \leqslant n$ e incorpore allí el valor de n. Concluya que $\{p_k : k \geqslant 0\}$ es la distribución binomial(n, p) con $p = a/(a - 1)$.

49. *Función generadora de probabilidad de una distribución en la clase $(a, b, 0)$.* Sea N una variable aleatoria con distribución de probabilidad en la clase $(a, b, 0)$.

    a) Demuestre que la correspondiente función generadora de probabilidad $P_N(t) = E(t^N)$ satisface la ecuación diferencial

    $$(1 - at)\, P_N'(t) = (a + b)\, P_N(t). \tag{2.7}$$

    b) Demuestre que la solución a la ecuación (2.7) con condición inicial $P_N(1) = 1$ y para $a \notin \{0, 1\}$ es

    $$P_N(t) = \left(\frac{1 - at}{1 - a} \right)^{-(a+b)/a}.$$

    c) En particular, demuestre que

    i) $P(N = 0) = (1 - a)^{(a+b)/a}$ para $a \neq 0$.

    ii) $\lim\limits_{a \to 0} P(N = 0) = e^{-b}$.

50. *Fórmula para la esperanza de una distribución en la clase $(a, b, 0)$.* Sea N una variable aleatoria con distribución de probabilidad en la clase $(a, b, 0)$. A partir de la definición elemental de esperanza, o bien, usando la fórmula (2.7) del ejercicio anterior, demuestre que para $a \neq 1$,

    $$E(N) = \frac{a + b}{1 - a}.$$

Fórmula de Panjer

51. Compruebe que la fórmula recursiva de Panjer efectivamente produce una función de probabilidad.

52. A partir de la fórmula recursiva de Panjer compruebe nuevamente que

    $$E(S) = E(N)E(Y).$$

53. *Fórmula recursiva para los momentos de S.* Use la fórmula de Panjer y la misma notación e hipótesis de dicho resultado para demostrar que para $a \neq 1$ y para $n \geqslant 1$,

$$E(S^n) = \frac{1}{1-a} \sum_{i=0}^{n-1} \left[a\binom{n}{i} + b\binom{n-1}{i} \right] E(S^i)\, E(Y^{n-i}).$$

54. *Fórmula recursiva para la función de distribución de S en un caso particular.* En general es difícil encontrar una expresión recursiva para la función de distribución $F(x)$ de un riesgo S en el modelo colectivo. Sin embargo, para el siguiente caso particular es posible encontrar dicha relación: usando la notación e hipótesis de la fórmula de Panjer y en el caso cuando N tiene una función de probabilidad geo(p), es decir, $p_k = P(N = k) = (1-p)^k p$, para $k \geqslant 0$, demuestre que para $x \geqslant 1$,

$$F(x) = p + (1-p) \sum_{j=1}^{x} f_j\, F(x-j).$$

55. Suponga que un riesgo S sigue un modelo colectivo en donde N tiene una distribución Poisson de parámetro $\lambda = 2$ y el monto de las reclamaciones tiene la función de probabilidad que aparece en la tabla de abajo. Use la fórmula de Panjer para encontrar la función de probabilidad de S.

r	1	2
f_r	1/2	1/2

56. Suponga que un riesgo S sigue un modelo colectivo en donde N tiene una distribución geométrica de parámetro $p = 1/2$, y el monto de las reclamaciones tiene la función de probabilidad que aparece en la tabla de abajo. Use la fórmula de Panjer para encontrar la función de probabilidad de S.

r	1	2
f_r	1/2	1/2

Aproximación normal

57. Sea S un riesgo con distribución gama(n, λ) en donde n es un entero.

    a) Asigne valores a los parámetros n y λ y calcule la aproximación normal para S.

    b) Para fines comparativos y con ayuda de algún paquete computacional, grafique tanto la función de densidad gama como la densidad normal aproximante.

    c) Duplique el valor de n y grafique nuevamente las dos funciones. ¿Mejoró la aproximación?

58. Suponga que un cierto riesgo S tiene una distribución Poisson compuesta con parámetro $\lambda = 50$, en donde los montos de las reclamaciones siguen una distribución uniforme$(0, 10)$. Use la aproximación normal para encontrar el valor de la prima p tal que:

    a) $P(S > p) \leqslant 0.05$.

    b) $P(S > p) \leqslant 0.01$.

59. Suponga que un riesgo S sigue una distribución Poisson compuesta con parámetro $\lambda = 40$ y los montos de las reclamaciones tienen distribución exp(α) con $\alpha = 10$. Use la aproximación normal para estimar la probabilidad $P(S > E(S))$.

60. Suponga que un riesgo S sigue una distribución Poisson compuesta con parámetro $\lambda = 30$ y los montos de las reclamaciones tienen distribución Pareto$(4, 3)$. Use la aproximación normal para comprobar que el valor de la prima p que cumple la condición $P(S > p) \leqslant 0.01$ es $p = 38.0194$.

61. Suponga que un riesgo S sigue un modelo colectivo Poisson de parámetro λ y las reclamaciones tiene la función de probabilidad que aparece en la siguiente tabla:

r	1	2	3
f_r	1/3	1/3	1/3

a) Suponga $\lambda = 30$. Use la fórmula de Panjer para encontrar la función de probabilidad de S.

b) Calcule la aproximación normal para S.

c) Con ayuda de un paquete computacional grafique la función de probabilidad de S y la densidad normal aproximante.

d) Suponga ahora $\lambda = 60$ y repita los incisos anteriores. ¿Mejoró la aproximación?

Aproximación gama trasladada

62. Sea S un riesgo con distribución gama(γ, α). Demuestre que:

    a) $E[(S - \gamma/\alpha)^3] = 2\gamma/\alpha^3$.

    b) la aproximación gama trasladada para S es exacta.

63. Compruebe que la aproximación gama trasladada para un riesgo con distribución exponencial es exacta.

64. Durante la derivación de la aproximación gama trasladada se usa el hecho de que la variable aleatoria $k + Z$, en donde Z tiene una distribución gama(γ, α), tiene media, varianza y coeficiente de asimetría $k + \gamma/\alpha$, γ/α^2 y $2/\sqrt{\gamma}$ respectivamente. Demuestre estas fórmulas.

65. Durante la derivación de la aproximación gama trasladada se llega al sistema de ecuaciones $k + \gamma/\alpha = m$, $\gamma/\alpha^2 = \sigma^2$ y $2/\sqrt{\gamma} = \alpha_3$, en donde k, γ y α son las incógnitas. Demuestre que la solución a este sistema es efectivamente $k = m - 2\sigma/\alpha_3$, $\gamma = 4/\alpha_3^2$ y $\alpha = 2/\sigma\alpha_3$.

66. Encuentre una expresión para la aproximación gama trasladada cuando el riesgo sigue una distribución:

    a) Poisson compuesta.

    b) binomial compuesta.

    c) binomial negativa compuesta.

67. Suponga que Z tiene una distribución gama(γ, α). Demuestre que si 2γ es un número entero natural, entonces $2\alpha Z$ tiene una distribución $\chi^2(2\gamma)$.

68. Suponga que un riesgo S sigue un modelo colectivo Poisson de parámetro $\lambda = 10$ y las reclamaciones tienes la función de probabilidad que aparece en la siguiente tabla:

r	1	2	3	4	5
f_r	1/8	1/4	1/4	1/4	1/8

a) Use la fórmula de Panjer para encontrar la función de probabilidad exacta de S.

b) Calcule la aproximación gama trasladada para S.

c) Con ayuda de un paquete computacional grafique la función de probabilidad exacta de S y la densidad aproximante.

Aproximación de Edgeworth

69. Demuestre que la aproximación de Edgeworth para un riesgo con distribución normal con media m y varianza σ^2 produce esta misma distribución con los mismos parámetros. Recuerde que si Z tiene una distribución normal estándar, entonces los momentos impares de esta variable aleatoria se anulan y para cualquie número natural n,

$$E(Z^{2n}) = \frac{(2n)!}{2^n\, n!}.$$

70. Suponga que S tiene una distribución Poisson compuesta de parámetro $\lambda > 0$ y que se desea usar la aproximación de Edgeworth para S. Demuestre que

$$a_k = \frac{d^k}{dr^k} \ln M_Z(r)\bigg|_{r=0} = \lambda\mu_k(\lambda\mu_2)^{-k/2}, \quad \text{para } k = 2, 3, 4.$$

En consecuencia,

$$\begin{aligned} P(S \leqslant x) \quad \approx \quad & \Phi(\frac{x-\lambda\mu}{\sqrt{\lambda\mu_2}}) - \frac{\lambda\mu_3(\lambda\mu_2)^{-3/2}}{6}\Phi^{(3)}(\frac{x-\lambda\mu}{\sqrt{\lambda\mu_2}}) \\ & + \frac{\lambda\mu_4(\lambda\mu_2)^{-2}}{24}\Phi^{(4)}(\frac{x-\lambda\mu}{\sqrt{\lambda\mu_2}}) \\ & + \frac{\lambda^2\mu_3^2(\lambda\mu_2)^{-3}}{72}\Phi^{(6)}(\frac{x-\lambda\mu}{\sqrt{\lambda\mu_2}}). \end{aligned}$$

71. Suponga que S tiene una distribución Poisson compuesta de parámetro $\lambda > 0$ y que se desea usar la aproximación de Edgeworth para S. Suponga adicionalmente que el monto de las reclamaciones siguen una distribución Pareto$(4,3)$. Demuestre que $\mu_4 = \infty$ y por lo tanto la fórmula del ejercicio anterior no puede aplicarse.

72. Sea S un riesgo con distribución $\exp(\lambda)$.

    a) Calcule la aproximación de Edgeworth de S.

    b) Asigne un valor al parámetro λ y con ayuda de un paquete computacional grafique tanto la función densidad exponencial como la función aproximante.

73. Suponga que un riesgo S sigue un modelo colectivo Poisson de parámetro $\lambda = 10$ y las reclamaciones tienes la función de probabilidad que aparece en la siguiente tabla:

r	1	2	3	4	5
f_r	1/2	1/4	1/8	1/16	1/16

    a) Use la fórmula de Panjer para encontrar la función de probabilidad exacta de S.

    b) Calcule la aproximación de Edgeworth para S.

    c) Con ayuda de un paquete computacional grafique la función de probabilidad exacta de S y la densidad aproximante.

Capítulo 3

Principios para el cálculo de primas

Hemos mencionado antes que una prima es un pago por adelantado que un asegurado realiza a una compañía aseguradora para obtener una cobertura parcial o completa contra un riesgo determinado, en los términos y condiciones que establece la póliza del seguro. En este capítulo vamos a estudiar algunas reglas generales para calcular el valor de una prima tomando en consideración únicamente los aspectos matemáticos del riesgo, es decir, no consideraremos cuestiones administrativas o mercadológicas del negocio del seguro, las cuales en situaciones prácticas son indispensables de considerar. Denotaremos por p, p_S o $p(S)$ a la prima para cubrir un riesgo S. De esta manera, a la fórmula para calcular una prima se le puede considerar como una función numérica de la variable aleatoria S o de su distribución.

3.1. Propiedades

¿Qué propiedades es razonable que cumpla una función $p(S)$ para el cálculo de las primas? Vamos a enunciar a continuación algunas propiedades generales que son deseables que posea cualquier método para calcular primas.

Simplicidad

El cálculo de la prima debe ser fácil de calcular. La simplicidad en el cálculo de la prima es deseable que se cumpla por varias razones, entre ellas está el aspecto práctico del cálculo mismo, así como el de lograr una cabal comprensión del cálculo de la prima por parte del asegurado y del resto de las personas involucradas en los procesos técnicos, administrativos y legales del seguro.

Consistencia

Si un riesgo se incrementa en una constante, entonces la prima debe reflejar ese cambio incrementándose en la misma cantidad, es decir, si $c > 0$ es una constante, entonces

$$p(S + c) = p(S) + c.$$

Aditividad

La prima de un portafolio consistente en dos riesgos independientes debe ser la suma de las primas individuales, es decir,

$$p(S_1 + S_2) = p(S_1) + p(S_2),$$

cuando S_1 y S_2 son dos riesgos independientes. Es claro que cuando se cumple esta propiedad, el intentar combinar o separar los riesgos no resulta en ninguna ventaja o provecho ni para el asegurado ni para el asegurador.

Invarianza de escala

Si $a > 0$ es una constante, entonces

$$p(aS) = a\,p(S),$$

es decir, si la cuantificación del riesgo S cambia de escala y se considera ahora el riesgo aS, la prima para este nuevo riesgo debe ser $a\,p(S)$, esto equivale a la prima original modificada con la misma escala.

Cota inferior

La prima debe tener siempre como cota inferior la prima pura de riesgo, es decir,

$$p \geqslant E(S).$$

Sin embargo, en algunas situaciones es necesario suponer que las primas deben tener siempre un recargo positivo y se considera la condición más restrictiva $p > E(S)$. A menos que se establezca lo contrario, la propiedad de cota inferior se entenderá en el sentido $p \geqslant E(S)$, la cual es más fácil de verificar en los métodos de cálculos de primas que estudiaremos.

Cota superior

Si un riesgo está acotado superiormente, entonces la prima para cubrir este riesgo también debe tener la misma cota superior, es decir, si $S \leqslant M$ para alguna constante $M > 0$, entonces

$$p(S) \leqslant M.$$

3.2. Principios generales

Recordemos que la prima pura de riesgo está dada por $p = E(S)$. Esta es la prima destinada a solventar exclusivamente la reclamación del riesgo, sin embargo veremos a continuación la posible situación catastrófica que podría presentarse cuando se toma $p = E(S)$. Considere un portafolio homogéneo de n pólizas de seguro de un mismo riesgo y válidas por un tiempo determinado. Suponga que se cobra una misma prima p por cada póliza y que S_j representa el monto de las reclamaciones efectuadas por la póliza j, las cuales se presuponen independientes y con idéntica distribución. Si u es el capital inicial de la aseguradora, entonces el capital de la misma al término de la vigencia de las pólizas es

$$\begin{aligned} X_n &= u + np - \sum_{j=1}^{n} S_j \\ &= u + \sum_{j=1}^{n} (p - S_j). \end{aligned}$$

Tenemos entonces las siguientes dos situaciones:

a) Cuando $p = E(S)$, al tomar esperanza en la ecuación anterior se obtiene $E(X_n) = u+n\,(p-E(S)) = u$. Es decir, en promedio la compañía aseguradora permanece con su capital inicial, sin embargo puede demostrarse que cuando $n \to \infty$, casi seguramente,

$$\limsup_{n\to\infty} X_n = -\liminf_{n\to\infty} X_n = \infty.$$

Esto quiere decir que el capital X_n puede oscilar y tomar valores grandes, tanto negativa como positivamente. Este resultado es parte del Teorema 6.3.1 del texto de Rolski *et al* [32] y su demostración hace uso de algunos elementos de caminatas aleatorias.

b) Cuando $p \neq E(S)$, por la ley de los grandes números,

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} X_n &= \lim_{n\to\infty} \frac{1}{n}\Big(u + \sum_{j=1}^{n} (p - S_j)\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} (p - S_j) \\
&= E(p - S) \\
&= p - E(S).
\end{aligned}$$

Así, para que este límite sea el indicado la variable X_n tiene que diverger a infinito o menos infinito dependiendo del signo de $p-E(S)$. Por lo tanto X_n tiene el siguiente comportamiento límite en el sentido casi seguro,

$$\lim_{n\to\infty} X_n = \begin{cases} +\infty & \text{si } p > E(S), \\ -\infty & \text{si } p < E(S). \end{cases}$$

En vista de estos resultados, es natural y deseable suponer $p > E(S)$. Esta condición se conoce con el nombre de condición de ganancia neta (*net profit condition*) y debe prevalecer en cualquier método para calcular p.

En general no existe un mecanismo de cálculo para la prima que sea el mejor pues existen varias condiciones que afectan la forma de calcular primas, entre ellas, las restricciones legales y financieras, las condiciones del asegurado, las condiciones de la propia aseguradora y de las otras aseguradoras, así como las condiciones del mercado del seguro. Todos estos son factores que determinan, directa o indirectamente, el valor de una prima para cubrir un riesgo particular en una situación real. Estudiaremos a continuación algunas formas particulares para el cálculo de primas. A estos procedimientos se les denomina con el término de prinicipios.

Principio del valor esperado

Este principio es uno de los más sencillos y establece que la prima puede calcularse de la siguiente forma:

$$p = (1 + \theta)E(S),$$

en donde $\theta > 0$ es una constante llamada factor de recargo (*safety loading*). Es decir, se trata de la reclamación promedio mas un porcentaje de ésta. En el factor de recargo se encuentran inmersos los costos administrativos y comerciales del seguro, así como los márgenes de utilidad de la aseguradora. La forma simple en la se expresa este principio es una de sus características principales, sin embargo puede observarse que una desventaja de esta fórmula es que asigna la misma prima a dos riesgos con distinta distribución pero con media común, y no toma en cuenta otro aspectos, por ejemplo, si las varianzas de los riesgos fueran distintas, entonces las primas tal vez deberían ser distintas.

Principio de la varianza

Este principio hace uso de la esperanza y la varianza del riesgo. En este caso el factor de recargo $\theta > 0$ se aplica sobre el valor de la varianza de la siguiente forma:

$$p = E(S) + \theta \operatorname{Var}(S).$$

Principio de la desviación estándar

Sea nuevamente $\theta > 0$ una constante. En este principio el factor de recargo se aplica sobre la desviación estándar del riesgo como indica la fórmula que aparece abajo. A diferencia del principio de la varianza, en este caso las unidades de medición del riesgo y de la prima coinciden. Y es evidente que la prima calculada mediante este principio produce una prima menor o igual a aquella calculada mediante el principio de la varianza.

$$p = E(S) + \theta\sqrt{\mathrm{Var}(S)}.$$

Principio de utilidad cero

Este principio hace uso de una función de utilidad, esto es, una función $v(x)$ definida sobre $[0, \infty)$ o un subconjunto de este intervalo y con valores en $\mathbb{R}$, que cumple las propiedades que se mencionan a continuación, y cuya gráfica en términos generales se muestra en la Figura 3.1.

a) Es estrictamente creciente.

b) Es cóncava.

Una función con estas características puede usarse para modelar el valor o utilidad que una persona o institución asocia a un bien monetario o material. En el apéndice el lector encontrará un exposición breve sobre algunas propiedades de este tipo de funciones. Suponiendo diferenciabilidad, la primera condición se escribe $v'(x) > 0$, y la segunda condición significa que $v''(x) \leqslant 0$. A veces se añade la condición $v(0) = 0$ pues toda función de utilidad (definida en $x = 0$) puede modificarse de tal forma que cumpla esa condición sin afectar el resultado en los procesos de decisión que se llevan a cabo usando estas funciones. La nueva función de utilidad sería $v(x) - v(0)$. Véase la sección sobre este tema en el apéndice.

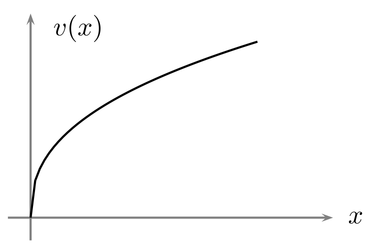


Figura 3.1: Función cóncava.

El principio de utilidad cero establece que la prima para cubrir un cierto riesgo S es aquel número p que satisface la ecuación

$$v(u) = E[v(u + p - S)], \tag{3.1}$$

en donde u es el capital inicial de la aseguradora. Es decir, la utilidad que representa para la aseguradora el capital inicial u debe ser idéntica a la utilidad esperada al cubrir el riesgo. Así, el cálculo de p está dado implícitamente por la ecuación (3.1) y para que la prima esté bien definida supondremos el caso cuando esta ecuación tiene una única solución p. Debemos mencionar, sin embargo, que no es sencillo resolver de manera exacta ecuaciones de la forma (3.1), pero pueden usarse métodos numéricos para conocer p de manera aproximada. El siguiente ejemplo es un caso muy particular y atípico en donde se puede calcular con facilidad la solución p de (3.1).

Ejemplo 3.1 *Considere la función de utilidad $u(x) = 1 - e^{-\alpha x}$, con $\alpha > 0$. La prima se calcula como aquel valor de p que es solución de la ecuación*

$$1 - e^{-\alpha u} = E[1 - e^{-\alpha(u+p-S)}].$$

Después de algunos cálculos sencillos, de la identidad anterior se obtiene la expresión

$$p = \frac{1}{\alpha} \ln M_S(\alpha). \tag{3.2}$$

Se presentan a continuación algunos ejemplos de funciones de utilidad.

a) Función de utilidad exponencial.

   $v(x) = 1 - e^{-\alpha x}, \quad \alpha > 0.$

b) Función de utilidad cuadrática.

   $v(x) = x - \alpha x^2, \quad \alpha > 0, \quad \text{para } 0 \leqslant x \leqslant 1/(2\alpha).$

c) Función de utilidad logarítmica.

   $v(x) = \alpha \ln x, \quad \alpha > 0.$

d) Función de utilidad de potencia fraccional.

   $v(x) = x^{\alpha}, \quad 0 < \alpha \leqslant 1.$

Demostraremos a continuación que el principio de utilidad cero produce primas que cumplen la condición $p \geqslant E(S)$. Por la desigualdad de Jensen en el caso de funciones cóncavas,

$$\begin{aligned} v(u) &= E[v(u+p-S)] \\ &\leqslant v(E(u+p-S)) \\ &= v(u+p-E(S)). \end{aligned}$$

Como v es una función estrictamente creciente, es uno a uno, y por lo tanto su inversa v^{-1} existe y también es estrictamente creciente. Al aplicar entonces la inversa se preserva la desigualdad anterior y se obtiene $p \geqslant E(S)$. La igualdad se logra, por ejemplo, cuando S es constante.

Principio del valor medio

Este principio hace uso de una función de valor, esto es, una función $v(x)$ que cumple las propiedades que aparecen abajo y cuya gráfica general se muestra en la Figura 3.2.

a) $v(0) = 0$.

b) Es estrictamente creciente.

c) Es estrictamente convexa.

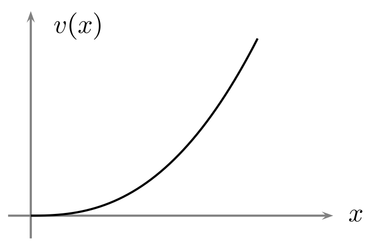


Figura 3.2: Función convexa.

El principio del valor medio establece que la prima p debe calcularse a partir de la igualdad

$$v(p) = E[v(S)]. \tag{3.3}$$

Esta identidad significa que la compañía aseguradora asigna el mismo valor a la prima que al promedio del valor de la reclamación y por lo tanto es indiferente a cualquiera de las dos situaciones. Como la función $v(x)$ es estrictamente creciente, es uno a uno, su inversa por lo tanto existe y es también estrictamente creciente. De hecho, la inversa de cualquier función de utilidad que se anula en cero es un ejemplo de una función de valor. Así, la prima mediante este principio se puede escribir de la siguiente forma:

$$p = v^{-1}(E(v(S))).$$

Por la desigualdad de Jensen para la función convexa v, $E(v(S)) \geqslant v(E(S))$, o bien por la misma desigualdad para la función cóncava v^{-1}, $v^{-1}(E(X)) \geqslant E(v^{-1}(X))$. Ambos caminos llevan a la desigualdad

$$p \geqslant E(S).$$

Ejemplo 3.2 *Considere la función de valor $v(x) = e^{\alpha x} - 1$, con $\alpha > 0$. Bajo este principio, la igualdad (3.3) se escribe $e^{\alpha p} - 1 = E(e^{\alpha S} - 1)$, lo que lleva a la siguiente solución, que es idéntica a (3.2),*

$$p = \frac{1}{\alpha} \ln M_S(\alpha). \tag{3.4}$$

Principio exponencial

Este es el principio de utilidad cero aplicado a la función de utilidad $v(x) = 1 - e^{-\alpha x}$, con $\alpha > 0$. Y coincide también con el principio del valor medio aplicado a la función de valor $v(x) = e^{\alpha x} - 1$, con $\alpha > 0$. Hemos visto que la prima calculada bajo estos principios es

$$p = \frac{1}{\alpha} \ln M_S(\alpha).$$

Observe que en este caso la prima no depende del capital inicial u. Puede verificarse directamente que $p \geqslant E(S)$, lo cual hemos demostrado antes de manera general.

Principio del porcentaje

Sea $\epsilon > 0$ una constante. El principio del porcentaje sugiere que la prima p puede calcularse mediante la expresión que aparece abajo. El significado geométrico de esta fórmula se muestra en la Figura 3.3.

$$p = \operatorname{ínf} \{x > 0 : P(S > x) \leqslant \epsilon\}.$$

De esta forma la probabilidad de que el riesgo exceda el monto de la prima debe ser pequeño o ajustable mediante el parámetro ϵ. A este principio también se le conoce también como principio de pérdida máxima. Por ejemplo, si S sigue una distribución exponencial de parámetro λ, entonces $P(S > x) = e^{-\lambda x}$. Y por lo tanto p es aquel valor numérico tal que $e^{-\lambda p} = \epsilon$,

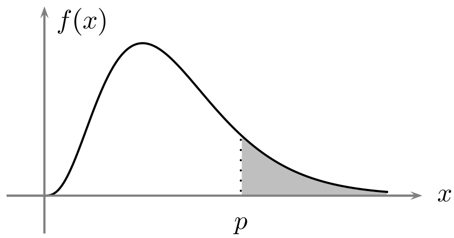


Figura 3.3

es decir, $p = -\frac{1}{\lambda}\ln \epsilon$. Así, para este ejemplo particular, se cumple la condición $p \geqslant E(S)$ si, y sólo si, $-\frac{1}{\lambda}\ln\epsilon \geqslant \frac{1}{\lambda}$, es decir, $\epsilon \leqslant e^{-1}$. Esto muestra que el principio del porcentaje no produce en general primas que cumplen la condición de ganancia neta.

Principio de Esscher

Antes de establecer este principio es necesario definir primero la transformada de Esscher de una distribución de probabilidad para la cual existe la función generadora de momentos.

Transformada de Esscher. Sea S un riesgo con función de densidad $f(x)$, función de distribución $F(x)$ y para la cual existe la función generadora de momentos $M_S(h)$, para algunos valores de $h \geqslant 0$. La transformada de Esscher con parámetro h de $f(x)$ es la función

$$x \mapsto g(x) = \frac{1}{M_S(h)}\, e^{hx} f(x). \tag{3.5}$$

Es inmediato comprobar que esta función es efectivamente de densidad. Por ejemplo, puede demostrarse que la transformada de Esscher de la distribución exponencial es nuevamente la distribución exponencial pero con parámetro distinto (ejercicio 87). Cuando el parámetro h es cero se obtiene la función de densidad original. La definición de transformada de Esscher puede definirse de manera análoga para variables aleatorias discretas.

El principio de Esscher establece que la prima para cubrir el riesgo S es la esperanza de esta nueva función de densidad, es decir,

$$\begin{aligned} p &= \frac{1}{M_S(h)} \int_0^\infty x\, e^{hx} f(x)\, dx \\ &= \frac{E(Se^{hS})}{E(e^{hS})}. \end{aligned}$$

Denotemos por $p(h)$ a esta función. Es claro que $p(0) = E(S)$ y puede demostrarse que $p(h)$ es una función creciente de h, véase el ejercicio 90. Por lo tanto, $p(h) \geqslant p(0) = E(S)$. Esto demuestra que se cumple la condición de ganancia neta y que mientras mayor es el parámetro h mayor es la prima. Habiendo definido la forma de calcular primas bajo este principio, vamos a hacer algunas observaciones acerca de la función de densidad (3.5), la cual es la función de densidad original ponderada por la función creciente $x \mapsto e^{hx}/M_S(h)$. La correspondiente función de distribución de (3.5) es

$$G(x) = \frac{1}{M_S(h)} \int_0^x e^{hy} f(y)\, dy.$$

A esta función también se le llama la transformada de Esscher de la función de distribución $F(x)$. Sea $\tilde{S}$ una variable aleatoria asociada a esta función de distribución. Algunos cálculos sencillos muestran que la función generadora de momentos de esta nueva variable aleatoria está dada por

$$M_{\tilde{S}}(t) = \frac{M_S(t+h)}{M_S(h)}.$$

Principio del riesgo ajustado

Este principio, así como el de Esscher, está basado en una transformación de la distribución del riesgo. Para un riesgo S con función de distribución $F(x)$ se define una nueva función de distribución de la siguiente forma

$$G(x) = 1 - (1 - F(x))^{1/\rho},$$

en donde $\rho \geqslant 1$ es un parámetro conocido como el índice del riesgo. Puesto que $1 - F(x)$ es un número entre 0 y 1, y $\rho \geqslant 1$, se cumple que

$$\begin{aligned} 1 - G(x) &= (1 - F(x))^{1/\rho} \\ &\geqslant 1 - F(x). \end{aligned} \tag{3.6}$$

Esto significa que la cola de la distribución del riesgo está siendo sobre estimada por la cola de la nueva distribución. Esta sobre estimación es usada para definir la prima para cubrir S. En la Figura 3.4 se muestra el orden que guardan las funciones $G(x)$ y $F(x)$.

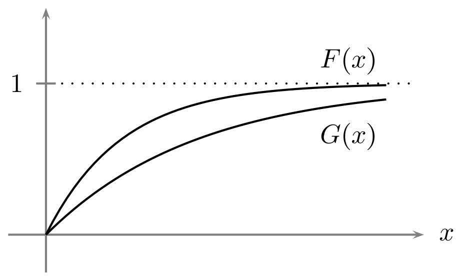


Figura 3.4

Así, la prima por el principio del riesgo ajustado para el riesgo S se define como la esperanza de la nueva función de distribución, es decir,

$$p = \int_0^\infty (1 - G(x))\,dx = \int_0^\infty (1 - F(x))^{1/\rho}\,dx.$$

Se verifica entonces la condición $p \geqslant E(S)$, pues

$$p = \int_0^\infty (1 - F(x))^{1/\rho}\,dx \geqslant \int_0^\infty (1 - F(x))\,dx = E(S).$$

Puesto que hemos revisado algunos métodos particulares para calcular primas y que contamos con una lista de propiedades deseables que deben cumplir los métodos, surge el problema de determinar si se cumplen o no se cumplen las propiedades para cada uno de los métodos enunciados. Tal tarea es el contenido del ejercicio 94. Algunas de estas comprobaciones son inmediatas, algunas otras requieren un poco mas de trabajo. En la tabla que hemos colocado en este ejercicio se ha omitido la propiedad de simplicidad, pues tal característica es subjetiva.

3.3. Primas y funciones de utilidad

Hemos mencionado que el principio de utilidad cero establece que la prima que una aseguradora está dispuesta a cobrar a un asegurado para cubrir un cierto riesgo S es aquel número p que satisface la ecuación

$$v_1(u_1) = E[v_1(u_1 + p - S)], \tag{3.7}$$

en donde u_1 es el capital inicial de la aseguradora y $v_1(x)$ es una función de utilidad . Denotemos por p^- a esta prima puesto que en realidad la aseguradora estaría contenta en cobrar una prima p que sea mayor o igual a p^-, es decir, desde el punto de vista de la aseguradora y bajo el criterio de utilidad cero, la prima p^- es la mínima prima a cobrar, y por lo tanto $p \geqslant p^-$.

En contraparte, un asegurado con capital inicial o riqueza u_2 y con función de utilidad $v_2(x)$, considera que puede aceptar contratar un seguro para cubrirse contra el riesgo S cuando, de acuerdo al principio de utilidad cero, la prima p está dada por

$$v_2(u_2 - p) = E[v_2(u_2 - S)]. \tag{3.8}$$

El posible valor de p solución de esta ecuación representa el punto de balance (misma utilidad) para el asegurado entre la decisión de contratar el seguro o no contratarlo. Denotemos ahora por p^+ a la solución de (3.8). Nuevamente ese valor es en realidad la prima máxima que el asegurado está dispuesto a pagar para cubrirse contra S, pues es claro que una prima menor o igual a tal valor es conveniente para él. De esta manera el principio de utilidad cero establece las condiciones de ambas partes para tomar un decisión respecto a firmar o no firmar el contrato del seguro. Es claro que habrá un acuerdo entre ambas partes si existe un valor de p tal que

$$p^- \leqslant p \leqslant p^+.$$

En tal caso se dice que el riesgo es asegurable bajo el criterio y condiciones mencionados. La situación se ilustra en la Figura 3.5.

Ejemplo 3.3 *Suponga que una compañía aseguradora con capital inicial $u_1 > 0$ decide asegurar un riesgo S y el cálculo de la prima se determina de acuerdo al principio de utilidad cero usando la función de utilidad $v_1(x) = 1 - e^{-\alpha_1 x}$, con $\alpha_1 > 0$. De este modo, la prima mínima que la aseguradora está dispuesta a cobrar es p^- dada por la solución de la ecuación*

$$v_1(u_1) = E[v_1(u_1 + p^- - S)].$$

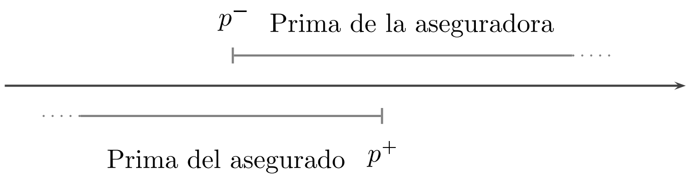


Figura 3.5

Resolviendo esta ecuación como se ha hecho antes se encuentra que $p^- = \frac{1}{\alpha_1} \ln M_S(\alpha_1)$. *Por otro lado, un asegurado con capital inicial* u_2 *está dispuesto a pagar una prima máxima* p^+ *para asegurarse contra este riesgo, determinada por la ecuación* $v_2(u_2 - p^+) = E[v_2(u_2 - S)]$, *en donde* $v_2(x)$ *es la función de utilidad particular* $v_2(x) = 1 - e^{-\alpha_2 x}$, *con* $\alpha_2 > 0$. *La solución de esta ecuación es* $p^+ = \frac{1}{\alpha_2} \ln M_S(\alpha_2)$. *Así, el riesgo mencionado es asegurable bajo las condiciones mencionadas si, y sólo si, las constantes* α_1 *y* α_2 *satisfacen la relación*

$$\frac{1}{\alpha_1} \ln M_S(\alpha_1) \leqslant \frac{1}{\alpha_2} \ln M_S(\alpha_2).$$

Comentarios y referencias

Uno de los problemas fundamentales de las compañías aseguradoras para operar un seguro es determinar la forma de calcular la prima a cobrar para asegurar un riesgo. En este capítulo hemos estudiado los primeros elementos para resolver este problema desde el punto de vista matemático y considerando exclusivamente la naturaleza estocástica del riesgo. Así, suponiendo que hemos adoptado un modelo matemático para representar un riesgo proveniente de una cartera de asegurados y suponiendo que podemos calcular algunas de sus características, en este capítulo hemos visto que no hay una forma única de calcular la prima correspondiente, pues en efecto, hemos mostrado varios métodos del cálculo de primas.

3.4. Ejercicios

Desigualdad de Jensen

74. *Convexidad y desigualdad de Jensen.* Una función $u : (a, b) \to \mathbb{R}$ es convexa si para cualesquiera números $x < y$ en (a, b), y para cualquier $t \in [0, 1]$ se cumple la desigualdad

$$u(tx + (1 - t)y) \leqslant tu(x) + (1 - t)u(y).$$

Geométricamente esta desigualdad significa que la recta que une los puntos $(x, u(x))$ y $(y, u(y))$ se encuentra por arriba de la función en el intervalo $[x, y]$. Cuando u es dos veces diferenciable, la condición de convexidad se escribe $u''(x) > 0$. La desigualdad de Jensen establece que si u es una función convexa y X es una variable aleatoria tal que tanto X como $u(X)$ tienen esperanza finita, entonces

$$u(E(X)) \leqslant E(u(X)).$$

Demuestre esta desigualdad en el caso cuando u es dos veces diferenciable siguiendo los siguientes pasos:

a) Escriba la serie de Taylor de la función u alrededor de un punto x_0 hasta la segunda derivada.

b) Use la condición $u''(x) > 0$ para concluir que $u(x) \geqslant u(x_0) + u'(x_0)(x - x_0)$.

c) Sustituya x por X, x_0 por $E(X)$ y después tome esperanza en ambos lados.

Funciones de utilidad y el principio de utilidad cero

75. Sean $a > 0$ y $\alpha > 0$ constantes. Demuestre que la función $v(x) = a(1 - e^{-\alpha x})$ definida para $x \geqslant 0$, es una función de utilidad y que usada bajo el principio de utilidad cero determina que la prima para cubrir un riesgo S debe ser $p = \frac{1}{\alpha} \ln M_S(\alpha)$.

76. Suponga que una persona con capital u tiene dos opciones de inversión, las cuales le reportan ganancias o pérdidas dadas por las

variables aleatorias I_1 e I_2. Es decir, al cabo de cada una de las inversiones su capital será $u + I_1$ o $u + I_2$. Suponga que la persona decide tomar la inversión que le otorga una utilidad esperada mayor, en donde su función de utilidad es exponencial. Demuestre que su decisión es independiente del capital inicial u.

Funciones de valor y el principio del valor medio

77. Suponga que un riesgo S tiene distribución $\exp(\lambda)$. Use el principio del valor medio para calcular la prima para cubrir este riesgo usando la función de valor $v(x) = x^2$.

Principio del valor esperado

78. Suponga que se tiene un portafolio de 350 pólizas de seguros contra robo a casahabitación con validez por un año, como se muestra en la siguiente tabla:

Probabilidad de siniestro	Reclamación	
	unif$(0, 10]$	unif$(10, 20]$
0.01	100	60
0.02	150	40

Esto quiere decir, por ejemplo, que se tienen 100 asegurados cuya probabilidad de reclamación es 0.01 y cuyo monto de reclamación se distribuye unif$(0, 10]$. Suponga que nos interesa modelar el agregado de reclamaciones mediante la variable S^i usando modelo individual, pero también con la variable S^c usando el modelo colectivo Poisson.

a) Encuentre la media y la varianza de las variables S^i y S^c.

b) Suponga que para el cálculo de las primas se usa el principio del valor esperado. Calcule el factor de recargo θ^i y θ^c, correspondiente a cada modelo, de tal forma que con probabilidad de 0.95 las primas sean suficientes para hacer frente a las reclamaciones.

c) Suponga que las probabilidades de reclamación se duplican. Encuentre nuevamente los factores de recargo θ^i y θ^c para que con probabilidad de 0.95 las primas sean suficientes para hacer frente a las reclamaciones.

Principio de la varianza

79. Usando el principio de la varianza calcule la prima p para cubrir un riesgo S con distribución unif$(0,1)$, y encuentre el valor del factor θ tal que $P(S > p) = 0.1$.

Principio exponencial

80. Use la desigualdad de Jensen para demostrar directamente que la prima calculada mediante el principio exponencial cumple la desigualdad $p \geqslant E(S)$.

Principio del porcentaje

81. Usando el principio del porcentaje, calcule la prima para cubrir un riesgo S con distribución exp(λ).

82. Mediante el principio del porcentaje calcule la prima para cubrir un riesgo con distribución Pareto(a, b).

83. Mediante el principio del porcentaje calcule la prima para cubrir un riesgo con distribución Weibull(r, α).

84. Considere un riesgo S con función de densidad como aparece abajo. Calcule la prima para cubrir S mediante el principio del porcentaje.

$$f(x) = \frac{\sqrt{2}}{\sqrt{\pi}}\, e^{-x^2/2}, \quad x > 0.$$

Transformada de Esscher

85. Calcule la transformada de Esscher de parámetro $h > 0$ de la distribución:

a) geo(p).
b) Poisson(λ).
c) unif(a, b).
d) N(μ, σ^2).

86. *Doble transformada.* Sea S un riesgo con función de densidad $f(x)$ y función generadora de momentos finita.

    a) Demuestre que la doble transformada de Esscher con parámetros h_1 y h_2 de la función $f(x)$ es
    $$g(x) = \frac{1}{M_S(h_1 + h_2)} \, e^{(h_1+h_2)x} f(x).$$
    b) Compruebe que la función generadora de momentos de la doble transformada es
    $$M(t) = \frac{M_S(t + h_1 + h_2)}{M_S(h_1 + h_2)}.$$

Principio de Esscher

87. Sea S un riesgo con distribución exp(λ).

    a) Demuestre que la transformada de Esscher de parámetro h es la distribución exp($\lambda - h$), para $0 \leqslant h < \lambda$.
    b) Encuentre la prima mediante el principio de Esscher.
    c) Verifique la condición de ganancia neta.

88. Suponga que la reclamación de un cierto riesgo S tiene distribución uniforme en el intervalo $(0, 1)$.

    a) Calcule la transformada de Esscher de parámetro $h = 1$ de esta distribución. Grafique y compare las dos funciones de densidad.
    b) Calcule la prima para cubrir este riesgo usando el principio de Esscher de parámetro $h = 1$.
    c) Calcule la probabilidad de que la prima sea suficiente para cubrir una reclamación del riesgo.

89. Sea S un riesgo con distribución gama(γ, λ).

    a) Demuestre que la transformada de Esscher de parámetro h es la distribución gama$(\gamma, \lambda - h)$, para $0 \leqslant h < \lambda$.
    b) Encuentre la prima mediante el principio de Esscher.
    c) Verifique la condición de ganancia neta.

90. *Condición de ganancia neta.* Sea $p(h) = E(Se^{hS})/E(e^{hS})$ la prima para cubrir un riesgo S calculada mediante el principio de Esscher. Demuestre que la función diferenciable $p(h)$ es creciente. En consecuencia, $p(h) \geqslant p(0) = E(S)$.
    Sugerencia: la desigualdad de Cauchy-Schwarz establece que para dos variables aleatorias con segundo momento finito se cumple que $E^2(XY) \leqslant E(X^2)\,E(Y^2)$.

Principio del riesgo ajustado

91. Sea S un riesgo con distribución exp(λ).

    a) Demuestre que la transformada del principio del riesgo ajustado de parámetro ρ es la distribución exp(λ/ρ).
    b) Encuentre la prima mediante el principio de riesgo ajustado para cubrir el riesgo S.
    c) Verifique la condición de ganancia neta para $\rho > 1$.

92. Sea S un riesgo con distribución Pareto(a, b).

    a) Demuestre que la transformada del principio del riesgo ajustado de parámetro ρ es la distribución Pareto$(a/\rho, b)$.
    b) Encuentre la prima mediante el principio de riesgo ajustado para cubrir el riesgo S.
    c) Verifique la condición de ganancia neta para $\rho > 1$.

93. Calcule la prima para cubrir un riesgo S con distribución unif$(0, 1)$ usando el principio del riesgo ajustado con índice de riesgo ρ y verifique la condición de ganancia neta.

Propiedades

94. Complete la siguiente tabla determinando si cada uno de los principios mencionados para calcular primas cumple la propiedad correspondiente. En cada caso demuestre su afirmación o proporcione un contraejemplo. Para simplificar las cosas considere que la propiedad de cota inferior se refiere a la desigualdad no estricta $p \geqslant E(S)$.

	Consistencia	Aditividad	Invarianza	Cota inferior	Cota superior
Valor esperado					
Varianza					
Desv. estándar					
Utilidad cero					
Valor medio					
Exponencial					
Porcentaje					
Esscher					
Riesgo ajustado					

Primas y funciones de utilidad

95. Calcule la prima máxima que un asegurado está dispuesto a pagar para cubrirse contra un riesgo S que puede tomar los valores 0 y 100, con idéntica probabilidad 1/2. Suponga que como función de utilidad se toma la función identidad.

96. Considere un riesgo S con distribución Poisson(λ) con $\lambda = 10$. Sea $v_1(x) = 1 - \alpha e^{-\alpha x}$ con $\alpha = 1/2$ la función de utilidad de la aseguradora. Sea $v_2(x) = x + 1$ la función de utilidad del solicitante del seguro. Determine si el riesgo S es asegurable. Observe que en este caso no es necesario conocer el capital inicial del asegurado ni del asegurador.

97. Sea $v(x)$ una función de utilidad, y sean a y b dos constantes con $a > 0$. Demuestre que $a\,v(x) + b$ es una función de utilidad.

98. Suponiendo diferenciabilidad, demuestre que la composición de dos funciones de utilidad es una función de utilidad.

99. *Aversión al riesgo.* Suponga que una persona con capital inicial u tiene la posibilidad de participar en un juego justo en el que recibirá una cantidad positiva x con probabilidad $1/2$, o perderá dicha cantidad con probabilidad $1/2$. En caso de no desear participar en este juego, el capital de la persona no se modifica y permanece con el valor u. Suponga que la persona toma la decisión de participar o no participar en el juego usando el criterio de la utilidad esperada máxima, es decir, tomará aquella decisión que le reditúe una utilidad promedio máxima. Demuestre que si la función de utilidad usada es estrictamente cóncava, es decir, $v''(x) < 0$, entonces la decisión será siempre no participar en el juego, a pesar de que éste es justo. Esto ilustra la interpretación de que las funciones de utilidad estrictamente cóncavas representan la utilidad de personas con aversión al riesgo.

100. *Coeficiente de aversión al riesgo.* Se define el coeficiente de aversión al riesgo de una función de utilidad $v(x)$ como la función $-v''(x)/v'(x) \geqslant 0$. Calcule este coeficiente en el caso de la función de utilidad exponencial, cuadrática, logarítmica y potencia fraccional.

Capítulo 4

Reaseguro

En este capítulo estudiaremos algunas ideas simples y generales del reaseguro. El reaseguro se presenta cuando una aseguradora firma un contrato para cubrir ciertos riesgos con otra compañía aseguradora llamada reaseguradora. De esta manera ambas aseguradoras adquieren la obligación de solventar las posibles reclamaciones del riesgo en cuestión y nos encontramos nuevamente en el esquema general de un asegurado y una aseguradora, en donde debe existir un acuerdo entre ambas partes acerca de las condiciones del contrato, las características del riesgo, las condiciones para el pago de la suma asegurada y, por supuesto, el cálculo de la prima correspondiente. Desde el punto de vista de la aseguradora, el reaseguro le ayuda a evitar posibles fuertes montos en las reclamaciones, aunque naturalmente disminuyen sus ingresos por primas pues tiene que compartir éstas con la reaseguradora.

Consideremos entonces un riesgo S y denotemos por S^A la parte del riesgo asumido por el asegurador y sea S^R la parte asumida por el reasegurador. Las letras A y R indican los términos Asegurador y Reasegurador, respectivamente. Debe entonces cumplirse la igualdad

$$S = S^A + S^R.$$

Los distintos tipos de reaseguro corresponden a las distintas formas en las que esta descomposición puede llevarse a cabo y tales descomposiciones pueden ser deterministas o aleatorias. El objetivo ahora es estudiar las características probabilísticas de las variables aleatorias S^A y S^R, y si se cumplen las condiciones adecuadas, pueden aplicarse los resultados encontrados antes para lograr dicho objetivo. El reaseguro puede tomar por lo menos dos perspectivas: actuar sobre las reclamaciones individuales o sobre el total del riesgo.

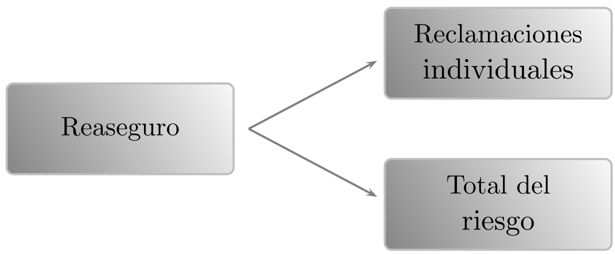


Figura 4.1

En ambos esquemas de reaseguro se aplica una función continua h del intervalo $[0,\infty)$ en sí mismo tal que:

a) $h(0) = 0$.

b) $h(x) \leqslant x$.

Las dos funciones de este tipo que consideraremos son $h(x) = ax$, para $a \in (0,1)$, y $h(x) = \text{mín}\{x, M\}$, para alguna constante $M > 0$. Gráficamente estas funciones se muestran a la Figura 4.2.

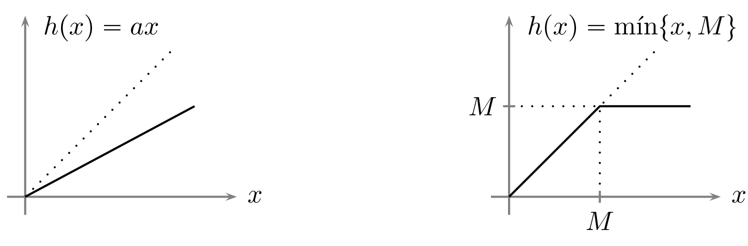


Figura 4.2

Bajo el esquema de reaseguro, la aseguradora afronta únicamente el riesgo resultante de aplicar la función h al riesgo original o a cada una de sus reclamaciones. Cuando se utiliza la primera de las funciones h mencionadas, el reaseguro se llama proporcional y en el segundo caso se le llama no proporcional. Estudiaremos a continuación estos dos tipos de reaseguro aunque debemos mencionar que a pesar de que los nombres de estos dos esquemas de reaseguro podrían sugerir que son los únicos, esto no es así, pues existen otras formas en las que los riesgos S^A y S^R pueden determinarse.

4.1. Reaseguro proporcional

Suponga que una compañía aseguradora reasegura un riesgo de la forma $S = \sum_{j=1}^{N} Y_j$. En el reaseguro proporcional se usa la función $h(x) = ax$, para algún valor de la constante a en el intervalo $(0,1)$. Usando la linealidad de esta función, es claro que es equivalente aplicar la función a cada reclamación por separado o al riesgo completo. Dada una reclamación Y, la aseguradora hace frente a una proporción de ésta, aY, y la reaseguradora cubre el porcentaje restante, $(1-a)Y$. El riesgo total asumido por la aseguradora es aS, y la reaseguradora cubre $(1-a)S$, es decir,

$$S^A = aS = \sum_{j=1}^{N} a\,Y_j\,,$$

$$S^R = (1-a)S = \sum_{j=1}^{N} (1-a)\,Y_j.$$

En este caso, las características probabilísticas de S^A y de S^R se encuentran fácilmente de las de S, pues no es difícil comprobar los siguientes resultados para S^A con análogos resultados para S^R.

a) $F_{S^A}(x) = F_S(x/a)$.

b) $f_{S^A}(x) = \frac{1}{a} f_S(x/a)$, cuando S es absolutamente continua.

c) $M_{S^A}(r) = M_S(ar)$.

d) $E(S^A) = a\,E(S) \leqslant E(S)$.

e) $\mathrm{Var}(S^A) = a^2\,\mathrm{Var}(S) \leqslant \mathrm{Var}(S)$.

Probabilidad de insolvencia

Comprobaremos a continuación que el reaseguro proporcional es conveniente para la aseguradora en el sentido de que la probabilidad de insolvencia bajo este tipo de reaseguro es menor o igual a la probabilidad de insolvencia cuando no hay reaseguro. Suponga que una compañía aseguradora adquiere un riesgo S durante un cierto periodo y por el cual cobra una prima p. Suponga que esta aseguradora tiene un capital inicial u. La probabilidad de que la compañía aseguradora no pueda solventar el riesgo es

$$P(S > p + u).$$

Tal probabilidad se muestra gráficamente en la Figura 4.3. Ahora suponga que la compañía decide reasegurarse mediante el esquema de reaseguro proporcional. Suponga además que el asegurador retiene una parte de la prima original p dada por ap, con $0 < a < 1$, y el reasegurador obtiene la cantidad $(1 - a)p$ al aceptar reasegurar este riesgo. La probabilidad de que la aseguradora no pueda solventar el nuevo riesgo bajo este esquema es

$$\begin{aligned} P(aS > ap + u) &= P(S > p + \frac{u}{a}) \\ &\leqslant P(S > p + u). \end{aligned}$$

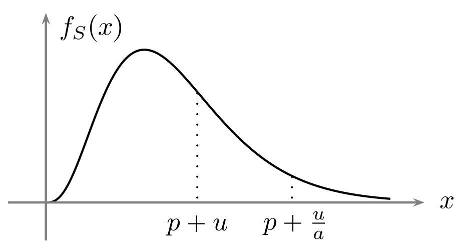


Figura 4.3

De esta forma hemos comprobado que la probabilidad de insolvencia bajo el esquema de reaseguro proporcional es menor o igual a la probabilidad del mismo evento cuando no hay reaseguro. La ventaja para la aseguradora al contratar el reaseguro es que la incertidumbre disminuye pero sus ganancias también disminuyen al ceder un porcentaje de sus ingresos por primas. Esta propiedad de disminución de insolvencia mediante el reaseguro es un tanto más difícil de verificar en otros tipos de reaseguro y dependerá, en general, del cálculo de la prima para el reaseguro.

4.2. Reaseguro no proporcional

En el reaseguro no proporcional se toma la función $h(x) = \text{mín}\{x, M\}$ para alguna constante $M > 0$ llamada nivel de retención. Distinguiremos dos casos: cuando se aplica esta función a cada reclamación y cuando se aplica sobre el total del riesgo.

Reaseguro en el riesgo completo (*stop loss*)

En este caso se aplica la función $h(x) = \text{mín}\{x, M\}$ sobre la totalidad del riesgo para obtener la parte que cubre la aseguradora y el riesgo excedente lo cubre la reaseguradora. De esta manera cada una de las aseguradoras cubren los riesgos:

$$\begin{aligned} S^A &= \text{mín}\,\{S, M\}, \\ S^R &= \text{máx}\,\{0, S - M\}. \end{aligned}$$

Tras un análisis puntual de estas funciones puede verificarse la identidad $S = S^A + S^R$, la cual no es evidente a primera vista. Así, mediante este mecanismo la aseguradora tiene la certidumbre de que cubrirá un monto máximo de M para el riesgo S. Es por ello que a este tipo de reaseguro se le llama reaseguro de pérdida máxima (*stop loss*).

Tanto la variable S^A como S^R son, en general, variables aleatorias mixtas, es decir, no son discretas ni continuas. Por ejemplo, la variable S^A tiene una masa de probabilidad en el punto M de valor $P(S \geqslant M)$, es decir, $P(S^A = M) = P(S \geqslant M)$. Puede demostrarse que su función de distribución toma la expresión

$$F_{S^A}(x) = \begin{cases} F_S(x) & \text{si } x < M, \\ 1 & \text{si } x \geqslant M, \end{cases}$$

la cual se muestra en la Figura 4.4 (a) en el caso cuando $F_S(x)$ es continua. Debido a su condición de variable aleatoria mixta, la función de densidad del riesgo S^A no puede expresarse como una función tradicional sino que es necesario hacer uso de las funciones generalizadas si se desea escribir tal función. En estos casos preferiremos trabajar únicamente con la función de distribución. Suponiendo que S es absolutamente continua, el n-ésimo momento de S^A puede expresarse de la siguiente forma:

$$E[(S^A)^n] = \int_0^M x^n f_S(x)\,dx + M^n P(S \geqslant M).$$

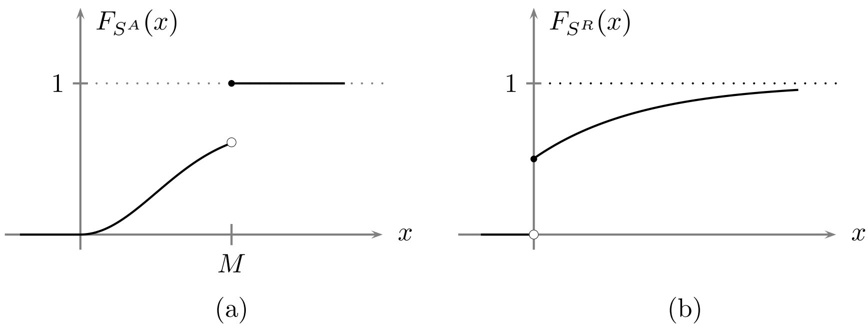


Figura 4.4

Por su parte, la variable S^R tiene una masa de probabilidad en el punto 0 de valor $F_S(M)$, es decir, $P(S^R = 0) = F_S(M)$ y su función de distribución es:

$$F_{S^R}(x) = \begin{cases} 0 & \text{si } x < 0, \\ F_S(M+x) & \text{si } x \geqslant 0, \end{cases}$$

la cual se muestra en la Figura 4.4 (b) en el caso cuando $F_S(x)$ es continua. Nuevamente, no es posible escribir la función de densidad de S^R en términos tradicionales y matemáticamente es preferible trabajar con la función de distribución. En el caso absolutamente continuo el n-ésimo momento de S^R adquiere la siguiente expresión:

$$E[(S^R)^n] = \int_0^\infty x^n f_S(M+x)\,dx.$$

Es interesante notar que la variable S^R puede tomar el valor 0 con probabilidad positiva y tal situación correspondería a una reclamación nula para la reaseguradora. No es difícil comprobar que la función de distribución de S^R condicionada a tomar valores positivos es, para $x > 0$,

$$F_{S^R \,|\, S^R > 0}(x) = P(S^R \leqslant x \,|\, S^R > 0) = \frac{F_S(M + x) - F_S(M)}{1 - F_S(M)}.$$

En el caso absolutamente continuo, diferenciando la identidad anterior se encuentra la función de densidad correspondiente.

$$f_{S^R \,|\, S^R > 0}(x) = \frac{f_S(M + x)}{1 - F_S(M)}, \quad \text{para } x > 0.$$

Seguros con deducible fijo

Algunos tipos de seguros, como el seguro contra accidentes de automóviles, contemplan el pago de una cantidad llamada deducible cada vez que el asegurado presenta una reclamación ante la compañía aseguradora. Supongamos que dicho deducible es una cantidad fija de d unidades monetarias y que una reclamación se modela con la variable aleatoria Y. Si el monto de la reclamación es menor a d, el asegurado cubre la totalidad de la reclamación, es decir, no hay reclamación para la aseguradora. En cambio, si el monto de la reclamación excede el valor de d, entonces el asegurado paga el deducible d y la aseguradora cubre la cantidad restante, es decir, $\max\{0, Y - d\}$. De esta forma, los riesgos para este tipo de pólizas de seguros se modelan con las herramientas que hemos mencionado para el reaseguro de pérdida máxima con nivel de retención d.

Reaseguro en cada reclamación (*excess of loss*)

Si se aplica la función $h(x) = \min\{x, M\}$ a cada reclamación Y de un riesgo que sigue el modelo colectivo, entonces el asegurador cubre

$$Y^A = \min\{Y, M\}$$

y el reasegurador cubre el monto restante

$$\begin{aligned} Y^R &= Y - \min\{Y, M\} \\ &= \max\{0, Y - M\}. \end{aligned}$$

Por lo tanto, cada una de las aseguradoras asume los siguientes riesgos:

$$\begin{aligned} S^A &= \sum_{j=1}^{N} \text{mín}\,\{Y_j, M\}, \\ S^R &= \sum_{j=1}^{N} \text{máx}\,\{0, Y_j - M\}. \end{aligned}$$

Analizando sumando a sumando estas dos expresiones es inmediato comprobar que

$$S = S^A + S^R.$$

A este esquema de reaseguro se le conoce también con el nombre de reaseguro por exceso de pérdida (*excess of loss*). Observe nuevamente que si se supone una distribución continua para la variable aleatoria Y_j con soporte el intervalo $(0, \infty)$, entonces las variables $\text{mín}\{Y_j, M\}$ y $\text{máx}\{0, Y_j - M\}$ serán mixtas. La primera de ellas tendrá una masa de probabilidad en el punto M y la segunda en el origen. Observe además que las expresiones para S^A y S^R corresponden exactamente a lo que hemos llamado modelo colectivo de riesgo, es decir, se trata de sumas aleatorias de variables aleatorias no negativas. En la parte inicial del texto hemos encontrado fórmulas y procedimientos que pueden aplicarse para conocer las características de estas variables aleatorias.

Número de reclamaciones

Sea N^A el número de reclamaciones que un asegurador afronta bajo un reaseguro por exceso de pérdida (*excess of loss*) con nivel de retención M sin recurrir a la reaseguradora, es decir,

$$N^A = \sum_{j=1}^{N} 1_{(Y_j \leqslant M)}.$$

Por otro lado, sea N^R el número de reclamaciones que el reasegurador atiende para este tipo de reaseguro, es decir,

$$N^R = \sum_{j=1}^{N} 1_{(Y_j > M)}.$$

Es claro que se cumple la identidad $N = N^A + N^R$. Observamos además que, condicionado al evento $(N = n)$, cada una de estas variables aleatorias tiene una distribución binomial en donde n es el número de ensayos y las probabilidades de éxito son $p = P(Y_j \leqslant M)$ en el primer caso y $1 - p = P(Y_j > M)$ en el segundo caso. La distribución no condicional de estas variables es el contenido del siguiente resultado en donde supondremos que la distribución de N es conocida y corresponde a uno de tres casos.

Proposición 4.1 *Sea* $a = P(Y_j \leqslant M)$.

1. *Si* N *tiene distribución* $bin(n, p)$, *entonces*
   a) $N^A \sim bin(n, ap)$.
   b) $N^R \sim bin(n, (1 - a)p)$.
2. *Si* N *tiene distribución* $Poisson(\lambda)$, *entonces*
   a) $N^A \sim Poisson(\lambda a)$.
   b) $N^R \sim Poisson(\lambda(1 - a))$.
3. *Si* N *tiene distribución* $bin\ neg(k, p)$, *entonces*
   a) $N^A \sim bin\ neg(k, p/(p + a(1 - p)))$.
   b) $N^R \sim bin\ neg(k, p/(p + (1 - a)(1 - p)))$.

Demostración. Verificaremos únicamente el primer resultado aplicando la fórmula general $M_S(t) = M_N(\ln(M_Y(t))$ cuando el riesgo S es la variable N^A y una reclamación Y tiene la forma particular $1_{(Y_j \leqslant M)}$, lo cual es una variable aleatoria con distribución Bernoulli. Observemos primeramente que

$$M_Y(t) = E(e^{tY}) = 1 - a + ae^t.$$

Entonces $M_{N^A}(t) = M_N(\ln(1 - a + ae^t))$. Cuando N tiene distribución $\text{bin}(n, p)$ tenemos que $M_N(t) = (1 - p + pe^t)^n$. Por lo tanto,

$$M_{N^A}(t) = (1 - p + p(1 - a + ae^t))^n = (1 - ap + ape^t)^n.$$

Es decir N^A tiene distribución $\text{bin}(n, ap)$. Análogamente puede verse que N^R tiene distribución $\text{bin}(n, (1-a)p)$. De la misma forma se demuestran los otros casos. Para ello debe recordarse que si N tiene distribución Poisson(λ), entonces

$$M_N(t) = \exp(\lambda(e^t - 1)),$$

y si N es bin neg(k, p), entonces

$$M_N(t) = (p/(1 - (1 - p)e^t))^k.$$

■

A manera de resumen de los tipos de reaseguro mencionados se tiene el diagrama de la Figura 4.5.

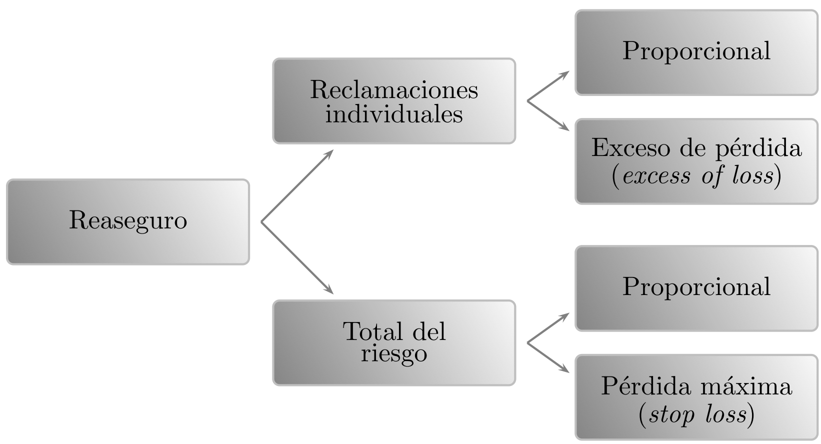


Figura 4.5

Comentarios y referencias

Mediante el reaseguro la compañía aseguradora se permite ceder parte de un riesgo a otra compañía llamada reaseguradora. Ante tal situación, el esquema de asegurado y asegurador cobra vigencia y se replantean los problemas originales ahora para el nuevo asegurado y la nueva aseguradora. En este capítulo hemos estudiado dos formas de llevar a cabo el reaseguro: el reaseguro proporcional y el reaseguro no proporcional. A pesar de que los nombres reaseguro proporcional y reaseguro no proporcional podrían parecer complementarios sugiriendo que son las dos únicas formas de reaseguro, tal afirmacion no es correcta, pues existen otros mecanismos mediante los cuales el reaseguro puede efectuarse. Por ejemplo, en el texto de Rolski *et al.* [32] se explica el reaseguro ECOMOR. Así, desde el punto de vista matemático el esquema de reaseguro genera dos nuevas variables aleatorias susceptibles de estudiarse probabilísticamente: el riesgo que conserva la aseguradora y el riesgo cedido a la reaseguradora. Y por supuesto se renueva el problema del cálculo de la prima para asegurar el riesgo cedido. El tema de reaseguro y varios de sus aspectos es tratado en mayor o menor medida en los textos de teoría de riesgo que se mencionan en la bibliografía, en particular Bowers *et al.* [7], Dickson [13] y Kaas *et al.* [22].

4.3. Ejercicios

Reaseguro proporcional

101. Para un reaseguro proporcional demuestre que:

   a) $E(S^A) = a\,E(S)$.

   b) $\mathrm{Var}(S^A) = a^2\,\mathrm{Var}(S)$.

   c) $\dfrac{E[(S^A - E(S^A))^3]}{[\mathrm{Var}(S^A)]^{3/2}} = \dfrac{E[(S - E(S))^3]}{[\mathrm{Var}(S)]^{3/2}}$.

102. Considere un riesgo S bajo un esquema de reaseguro proporcional. Es claro que, en general, las variables aleatorias $S^A = aS$ y $S^R = (1-a)S$ no son independientes, de hecho hay una dependencia lineal entre ellas. Demuestre que:

   a) $\mathrm{Cov}(S^A, S^R) = a(1-a)\mathrm{Var}(S)$.

   b) $\rho(S^A, S^R) = 1$.

   c) $S^R = \frac{1-a}{a}\,S^A$.

103. Suponga que un riesgo S sigue una distribución Poisson compuesta con parámetros $\lambda = 100$ y $F(x)$ dada por la distribución exponencial de media 500. Bajo un esquema de reaseguro proporcional con

$a = 0.7$, encuentre la distribución y los parámetros de los riesgos $S^A = aS$, y $S^R = (1-a)S$.

104. Suponga que un riesgo S sigue una distribución binomial compuesta con parámetros $n = 100$, $p = 2/3$, y $F(x)$ dada por

$$F(x) = \begin{cases} 0 & \text{si } x < 50, \\ 1/2 & \text{si } 50 \leqslant x < 100, \\ 1 & \text{si } x \geqslant 100. \end{cases}$$

Bajo un esquema de reaseguro proporcional con $a = 0.9$, encuentre la esperanza y varianza de los riesgos $S^A = aS$, y $S^R = (1-a)S$.

105. Bajo el esquema de reaseguro proporcional de un cierto riesgo S, el riesgo asumido por el asegurador es $S^A = aS$, y la prima recibida es ap. Suponga que el capital inicial de la aseguradora para afrontar dicho riesgo es u. Demuestre que la probabilidad de insolvencia $P(aS > ap + u)$ es una función creciente de $a \in (0, 1)$.

106. Suponga que un riesgo S se modela mediante la distribución $\exp(\lambda)$. Demuestre que bajo un reaseguro proporcional el riesgo asumido por la aseguradora $S^A = aS$ tiene distribución $\exp(\lambda/a)$. ¿Cuál es la distribución de S^R?

107. Suponga que un riesgo S se modela mediante la distribución gamma(γ, λ). Demuestre que bajo un reaseguro proporcional el riesgo asumido por la aseguradora $S^A = aS$ tiene distribución gama$(\gamma, \lambda/a)$. ¿Cuál es la distribución de S^R?

108. Suponga que un riesgo S sigue una distribución Poisson compuesta con parámetros $\lambda = 100$, y $F(x)$ dada por la distribución gama(γ, λ_0). Bajo un esquema de reaseguro proporcional con $a = 0.9$, encuentre la distribución y los parámetros de los riesgos $S^A = aS$, y $S^R = (1-a)S$.

109. Suponga que un riesgo S se modela mediante la distribución log normal(μ, σ^2). Demuestre que bajo un reaseguro proporcional el riesgo asumido por la aseguradora $S^A = aS$ tiene distribución log normal$(\mu + \ln a, \sigma^2)$. ¿Cuál es la distribución de S^R?

110. Suponga que un riesgo S sigue una distribución Poisson compuesta con parámetros $\lambda = 100$, y $F(x)$ dada por la distribución log normal(μ, σ^2). Bajo un esquema de reaseguro proporcional con $a = 0.9$, encuentre la distribución y los parámetros de los riesgos $S^A = aS$, y $S^R = (1-a)S$.

Reaseguro de pérdida máxima (*stop loss*)

111. Demuestre que bajo un reaseguro de pérdida máxima con nivel de retención M, los riesgos S^A y S^R tienen las siguientes funciones de distribución:

 a) $F_{S^A}(x) = \begin{cases} F_S(x) & \text{si } x < M, \\ 1 & \text{si } x \geqslant M. \end{cases}$

 b) $F_{S^R}(x) = \begin{cases} 0 & \text{si } x < 0, \\ F_S(M+x) & \text{si } x \geqslant 0. \end{cases}$

112. Para un reaseguro de pérdida máxima con nivel de retención M y utilizando los resultados del inciso anterior, demuestre que para cualquier entero $n \geqslant 0$:

 a) $E[(S^A)^n] = \int_0^M x^n f_S(x)\, dx + M^n P(S > M)$.

 b) $E[(S^R)^n] = \int_0^\infty x^n f_S(M+x)\, dx$.

113. *El n-ésimo momento limitado.* Sea S un riesgo con función de supervivencia continua $\overline{F}(x) = P(S > x)$. Sea M cualquier constante positiva. Demuestre que para cualquier entero $n \geqslant 1$:

$$E(S \wedge M)^n = \int_0^M n x^{n-1} \overline{F}(x)\, dx.$$

114. Considere un reaseguro de pérdida máxima con nivel de retención M para un riesgo S que sigue una distribución $\exp(\lambda)$. Demuestre que las funciones generadoras de momentos de los riesgos S^A y S^R están dadas por las siguientes expresiones:

a) $M_{S^A}(t) = \dfrac{1}{\lambda - t}\,(\lambda - te^{-(\lambda-t)M})$, para $t < \lambda$.

b) $M_{S^R}(t) = 1 + \dfrac{t}{\lambda - t}\,e^{-\lambda M}$, para $t < \lambda$.

115. Suponga que un riesgo S se modela mediante la distribución $\exp(\lambda)$. Demuestre que bajo un reaseguro de pérdida máxima con nivel de retención M, se tiene que:

a) $E(S^A) = \dfrac{1}{\lambda}\left(1 - e^{-\lambda M}\right)$.

b) $E(S^R) = \dfrac{1}{\lambda}\,e^{-\lambda M}$.

116. Suponga que el riesgo S sigue una distribución lognormal(μ, σ^2). Bajo un reaseguro de pérdida máxima con nivel de retención M, demuestre que

$$
\begin{aligned}
E[(S^A)^n] &= e^{n\mu + n^2\sigma^2/2}\,\Phi((\ln M - \mu - n\sigma^2)/\sigma) \\
&\quad + M^n(1 - \Phi((\ln M - \mu)/\sigma)).
\end{aligned}
$$

117. Suponga que un riesgo S tiene la distribución de probabilidad que aparece en la tabla de abajo. Calcule la distribución de probabilidad de los pagos positivos que efectúa una reaseguradora en un reaseguro de pérdida máxima con nivel de retención $M = 350$.

x	100	200	300	400	500
$P(S = x)$	0.1	0.2	0.3	0.2	0.2

118. Suponga que se establece como función de probabilidad para un riesgo S la que aparece en la tabla de abajo. Calcule la distribución de probabilidad de los pagos positivos que efectúa una reaseguradora en un reaseguro de pérdida máxima con nivel de retención $M = 30$.

x	10	20	30	40	50
$P(S = x)$	0.2	0.3	0.1	0.1	0.3

119. Suponga que el riesgo S sigue una distribución Pareto(a, b). Bajo el reaseguro de pérdida máxima, demuestre que la distribución del riesgo S^R condicionada al evento $(S^R > 0)$ es nuevamente Pareto$(a, b + M)$.

120. Suponga que el riesgo S sigue una distribución $\exp(\lambda)$. Bajo un reaseguro de pérdida máxima con nivel de retención M, demuestre que la distribución del riesgo S^R condicionado a ser positivo es nuevamente $\exp(\lambda)$.

Reaseguro en cada reclamación (*excess of loss*)

121. Suponga que se establece como función de probabilidad para la variable aleatoria Y la que aparece en la tabla de abajo. Esta variable aleatoria representa el monto de una reclamación en un seguro. Considere una reaseguro *excess of loss* para Y, con nivel de retención $M = 100$.

y	50	100	150	200
$P(Y = y)$	0.2	0.3	0.4	0.1

a) Calcule la función de probabilidad para las variables $Y^A = \text{mín}\{Y, M\}$ y $Y^R = \text{máx}\{0, Y - M\}$.

b) Calcule las esperanzas de estas variables aleatorias y compruebe que $E(Y) = E(Y^A) + E(Y^R)$.

122. Suponga que el monto de una reclamación Y se modela mediante una variable aleatoria con distribución $\exp(\lambda)$. Para un valor de retención $M > 0$ fijo, calcule la función de distribución y la esperanza de las variables aleatorias $\text{mín}\{Y, M\}$ y $\text{máx}\{0, Y - M\}$.

123. Demuestre que la función generadora de probabilidad de la variable N^R (número de reclamaciones que la reaseguradora afronta) en un reaseguro por exceso de pérdida con nivel de retención M es

$$P_{N^R}(t) = P_N(1 - p + pt),$$

en donde $p = P(Y_j > M)$ y $P_N(t)$ es la función generadora de probabilidad del número total de reclamaciones N.

124. Considere un reaseguro por exceso de pérdida y sean N^A y N^R el número de reclamaciones afrontadas por el segurador y por el reasegurador respectivamente. El total de reclamaciones es

$$N = N^A + N^R.$$

Suponga que N tiene distribución Poisson(λ).

a) Condicionando sobre el valor de N, demuestre que las variables N^A y N^R son independientes.

b) Compruebe que N^A y N no son independientes. Verifique, por ejemplo, que $P(N^A = 0, N = 0) \neq P(N^A = 0)\, P(N = 0)$.

Suponga ahora que N tiene distribución bin(n, p). Evaluando la probabilidad, por ejemplo, en el valor cero para ambas variables aleatorias, demuestre que:

c) N^A y N^R no son independientes.

d) N y N^A no son independientes.

125. Suponga que un cierto riesgo S se representa mediante un modelo colectivo. Condicionando sobre los posibles valores de la variable aleatoria N, demuestre que las variables N^A y N^R tienen la misma distribución que N pero con distintos parámetros en los casos cuando N es Poisson, binomial y binomial negativa. Este es un método alternativo de la demostración de la Proposición 4.1 en donde se usó la función generadora de momentos.

126. Considere un riesgo S con distribución Poisson compuesta de parámetro λ. Suponga que cada reclamación Y tiene distribución Pareto(a, b). Se adquiere un reaseguro por exceso de pérdida con nivel de retención M y por lo tanto la reclamación afrontada por la aseguradora es $Y^A = \text{mín}\{Y, M\}$. Demuestre que:

a) $E(Y^A) = E(Y) - \left(\dfrac{b}{b+M}\right)^{a-1} E(Y)$.

b) $E(S^A) = E(S) - \left(\dfrac{b}{b+M}\right)^{a-1} E(S)$.

127. Suponga se tiene un riesgo de la forma $S = \sum_{j=1}^{N} Y_j$, según el modelo colectivo, en donde cada reclamación se modela mediante una variable aleatoria con distribución $\exp(\alpha)$. Bajo un reaseguro por exceso de pérdida con nivel de retención M, el monto a pagar por la aseguradora por cada reclamación es $Y^A = \text{mín}\{Y, M\}$. Demuestre que:

a) $E(Y^A) = \dfrac{1}{\alpha}\,(1 - e^{-\alpha M})$.

b) $E((Y^A)^2) = \dfrac{2}{\alpha^2} - \dfrac{2}{\alpha^2}\,e^{-\alpha M} - \dfrac{2M}{\alpha}\,e^{-\alpha M}$.

c) $\mathrm{Var}(Y^A) = \dfrac{1}{\alpha^2} - \dfrac{2M}{\alpha}\,e^{-\alpha M} - \dfrac{1}{\alpha^2}\,e^{-2\alpha M}$.

d) $E(Y^A) \leqslant E(Y)$.

e) $\mathrm{Var}(Y^A) \leqslant \mathrm{Var}(Y)$.

Usando las fórmulas generales encuentre ahora la esperanza y varianza de $S^A = \sum_{j=1}^{N} \text{mín}\{Y_j, M\}$.

Capítulo 5

Teoría de la credibilidad

Considere un cierto riesgo proveniente de un conjunto de asegurados vigentes por un periodo determinado. Si este grupo de asegurados es homogéneo en el sentido de que todos sus miembros tienen las mismas condiciones estocásticas de realizar una reclamación, entonces es razonable aplicar una misma prima a todos ellos. Sin embargo, cuando el grupo no es homogéneo, o bien, surge con el paso del tiempo cierta heterogeneidad entre ellos, habrá subgrupos de bajo riesgo y otros de alto riesgo. Cobrar una misma prima a todos sería injusto, y no sería bueno para la aseguradora, pues eventualmente los asegurados de bajo riesgo buscarían un mejor trato con otra aseguradora. La idea fundamental es aplicar primas menores a los asegurados de bajo riesgo y primas mayores a los de alto riesgo, con base en el historial de reclamaciones que cada uno de los asegurados o subgrupos hayan realizado durante los periodos anteriores. En la teoría de la credibilidad se estudian métodos para el cálculo de primas a través de la combinación de la experiencia individual (historial de reclamaciones) y la experiencia de grupo (comportamiento teórico).
Este capítulo es breve y únicamente se exponen algunas ideas introductorias a la credibilidad clásica y Bayesiana. Una exposición más completa de este tema puede encontrarse en los textos [6], [9] y [23].

5.1. Credibilidad clásica

Esta perspectiva de la teoría de la credibilidad constituye el primer intento de definir la credibilidad para un conjunto de observaciones de un riesgo. Las ideas que presentaremos en esta primera sección fueron propuestas inicialmente por A. H. Mowbray en 1914 y desarrolladas después por varios actuarios estadounidenses durante la primera mitad del siglo XX. Debido a ello es que a la credibilidad clásica se le conoce también como credibilidad americana.

Credibilidad completa

Suponga que S representa el riesgo para una aseguradora correspondiente a un asegurado o un conjunto de asegurados con ciertas características particulares, y válido por un periodo fijo, por ejemplo, un año. Sean $S_1, \ldots, S_m$ los montos de las reclamaciones efectuadas por este asegurado o grupo de asegurados durante m periodos consecutivos, y sea $\bar{S}$ la media muestral o promedio de reclamaciones, es decir,

$$\bar{S} = \frac{1}{m}(S_1 + \cdots + S_m).$$

Nos interesa estudiar el comportamiento de $\bar{S}$ a lo largo del tiempo para un conjunto de asegurados en particular, pues deseamos determinar si la prima que se les cobra a cada uno de ellos es la adecuada. Si se considera que las variables $S_1, \ldots, S_m$ son independientes, idénticamente distribuidas y con esperanza finita, entonces la ley de los grandes números garantiza que la media muestral $\bar{S}$ converge a la constante desconocida $E(S)$, conforme el número de sumandos crece a infinito, véase la Figura 5.1.
Es decir, el comportamiento de $\bar{S}$ como función de m es posiblemente oscilatorio alrededor de $E(S)$ pero eventualmente va a estabilizarse en ese valor. La pregunta es, ¿qué tan grande debe ser m para que $\bar{S}$ esté razonablemente cercano al valor desconocido $E(S)$? La intención es usar $\bar{S}$ para calcular la prima pura de riesgo de un asegurado o grupo de asegurados, siempre y cuando se tenga suficiente historial para dar credibilidad a tal estadística. El siguiente criterio es una posible respuesta a la pregunta planteada.

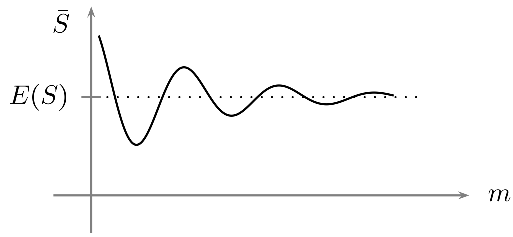


Figura 5.1

Definición 5.1 *Sea $S_1, S_2, \ldots, S_m$ una colección de v.a.i.i.d. con esperanza finita. Sean $k \in (0,1)$ y $p \in (0,1)$ dos números fijos. Se dice que la media muestral $\overline{S} = (S_1 + \cdots + S_m)/m$ tiene credibilidad completa (k,p) si*

$$P(\ |\bar{S} - E(S)| \leqslant kE(S)\ ) \geqslant p. \tag{5.1}$$

La condición anterior establece que $\bar{S}$ tiene credibilidad completa si dista de $E(S)$, en menos de $kE(S)$ con probabilidad mayor o igual a p. Observe que la definición tiene sentido cuando $E(S)$ es distinta de cero. Naturalmente se toman valores de k cercanos a cero y valores de p cercanos a 1, típicamente $k = 0.05$ y $p = 0.9$. El problema es entonces encontrar el valor del número de periodos de observación m para que se cumpla el criterio (5.1).

Credibilidad completa bajo hipótesis de normalidad

Encontraremos una condición sobre el número de periodos de observación m para obtener credibilidad completa cuando $\bar{S}$ tiene una distribución aproximada normal. Observe que $E(\bar{S}) = E(S)$, $\text{Var}(\bar{S}) = \text{Var}(S)/m$ y que estos valores son desconocidos. Tenemos entonces que, por el teorema central del límite, la parte izquierda de (5.1) es

$$
\begin{aligned}
P(\,|\bar{S}-E(S)| \leqslant kE(S)\,) &= P(\,\frac{|\bar{S}-E(S)|}{\sqrt{\mathrm{Var}(S)/m}} \leqslant \frac{kE(S)}{\sqrt{\mathrm{Var}(S)/m}}\,) \\
&\approx P(-\frac{kE(S)}{\sqrt{\mathrm{Var}(S)/m}} \leqslant Z \leqslant \frac{kE(S)}{\sqrt{\mathrm{Var}(S)/m}}) \\
&= \Phi(\frac{kE(S)}{\sqrt{\mathrm{Var}(S)/m}}) - \Phi(-\frac{kE(S)}{\sqrt{\mathrm{Var}(S)/m}}) \\
&= 2\Phi(\frac{k\sqrt{m}E(S)}{\sqrt{\mathrm{Var}(S)}}) - 1.
\end{aligned}
$$

Como esta probabilidad debe ser mayor o igual a p según el criterio (5.1), se obtiene la desigualdad

$$\Phi(\frac{k\sqrt{m}\,E(S)}{\sqrt{\mathrm{Var}(S)}}) \geqslant \frac{1+p}{2}.$$

Deseamos encontrar el valor de m más pequeño que cumpla esta desigualdad. Para ello consideraremos que se tiene la igualdad en esta ecuación. Sea u_q el q-cuantil de la distribución normal, es decir $\Phi(u_q) = q$. Entonces

$$\frac{k\sqrt{m}\,E(S)}{\sqrt{\mathrm{Var}(S)}} \geqslant u_{(1+p)/2}.$$

Despejando el parámetro m se obtiene el siguiente resultado.

Proposición 5.1 *Sea $S_1, S_2, \ldots, S_m$ una colección de v.a.i.i.d. con media y varianza finitas. La media muestral $\overline{S} = (S_1 + \cdots + S_m)/m$ alcanza credibilidad completa (k, p) de manera aproximada cuando*

$$m \geqslant \frac{u^2_{(1+p)/2}\,Var(S)}{k^2\,E^2(S)}. \tag{5.2}$$

Tomando k =0.05 y $p = 0.9$ tenemos que $u_{(1+p)/2} = u_{0.95} = 1.6449$. Por lo tanto la desigualdad (5.2) se reduce a

$$m \geqslant 1082 \, \frac{\mathrm{Var}(S)}{E^2(S)}.$$

El problema aún prevalece puesto que los términos $E(S)$ y $\mathrm{Var}(S)$ no son conocidos. En este punto se introduce otro factor de inexactitud en el análisis al sustituir $E(S)$ y $\mathrm{Var}(S)$ por la media y varianza muestrales usando la información que se tenga a disposición al momento de hacer la estimación:

$$\begin{aligned} E(S) &\approx \bar{s} = \frac{1}{n}\sum_{j=1}^{n} s_j, \\ \mathrm{Var}(S) &\approx \frac{1}{n-1}\sum_{j=1}^{n} (s_j - \bar{s})^2. \end{aligned}$$

Sustituyendo estos valores en la fórmula se puede conocer una aproximación del número de periodos m de historial para que $\bar{S}$ tenga credibilidad completa. Observe que cuando p crece, es decir, cuando se desea tener una mayor confianza en la estabilización de $\bar{S}$, entonces el número de periodos de observación m también crece. Por otro lado, si el parámetro k crece, es decir, si se pide que la distancia entre $\bar{S}$ y $E(S)$ tenga mayor amplitud, entonces m decrece. Observemos finalmente que, después de algunos cálculos sencillos, la aproximación para la credibilidad completa (5.2) puede expresarse de la siguiente forma:

$$\mathrm{Var}(\bar{S}) \leqslant \frac{k^2 E^2(S)}{u^2_{(1+p)/2}}. \tag{5.3}$$

Esto significa que la credibilidad completa se logra cuando la varianza o variabilidad de $\bar{S}$ ha sido suficientemente reducida. En esta expresión el término desconocido $E(S)$ debe ser sustituido por la media muestral. En general se necesitan muchos periodos de observación de un riesgo para obtener la credibilidad completa de la media muestral, y ello no es factible de tenerse en muchos casos.

Ejemplo 5.1 *Suponga que un riesgo para un grupo de asegurados sigue el modelo colectivo*

$$S = \sum_{j=1}^{N} Y_j,$$

en donde N tiene distribución Poisson(λ) con λ desconocido. Suponga adicionalmente que cada reclamación individual Y tiene distribución exp(α). De acuerdo con la notación previa, el primer y segundo momentos de Y son $\mu_1 = 1/\alpha$ y $\mu_2 = 2/\alpha^2$. Entonces

$$\begin{aligned} E(S) &= \lambda\mu_1 = \lambda/\alpha, \\ Var(S) &= \lambda\mu_2 = 2\lambda/\alpha^2. \end{aligned}$$

Por lo tanto, tomando nuevamente k =0.05 y p =0.9, la aproximación (5.2) es

$$m \geqslant 1082\,\frac{2\lambda/\alpha^2}{\lambda^2/\alpha^2},$$

o bien,

$$\lambda m \geqslant 2164.$$

Observe que λm representa el total de reclamaciones promedio durante m periodos y tal cantidad es desconocida pues no conocemos λ. Sin embargo, aproximando λm por el estimador $N_1 + N_2 + \cdots + N_m$, la condición anterior establece que después de 2164 reclamaciones se obtiene credibilidad completa para $\bar{S}$, suponiendo disponibles tales estadísticas. Así, en este caso el criterio de credibilidad completa para $\overline{S}$ ha quedado expresado en términos del número de reclamaciones promedio λm y no del número de periodos m.

Credibilidad parcial

En lugar del estimador $\bar{S}$ para $E(S)$ se propone la combinación lineal convexa

$$z\bar{S} + (1-z)E(S), \tag{5.4}$$

en donde $z \in [0,1]$ es llamado factor de credibilidad. Mediante una expresión como la propuesta se le otorga credibilidad parcial a la media muestral $\bar{S}$. Sin embargo, es muy importante notar que la expresión (5.4) no es un estimador de la media desconocida $E(S)$ puesto que depende de ella misma. A pesar de esto, aplicaremos la condición de credibilidad completa (5.1) a esta expresión y encontraremos una fórmula para el factor de credibilidad z. Así, la condición (5.1) aplicada a la expresión (5.4) se reduce a

$$P(\ |z(\bar{S} - E(S))| \leqslant kE(S)\ ) \geqslant p.$$

Observe que esta fórmula hace referencia únicamente al primer sumando de (5.4). Reescribimos la expresión anterior de la forma siguiente:

$$P(\ |\bar{S} - E(S)| \leqslant \frac{k}{z} E(S)\ ) \geqslant p.$$

Esta es la misma condición para la credibilidad completa sólo que en lugar del parámetro k tenemos ahora k/z, es decir, la credibilidad completa (k, p) para $z\bar{S} + (1-z)E(S)$ es equivalente a la credibilidad completa $(k/z, p)$ para $\bar{S}$.

Credibilidad parcial bajo hipótesis de normalidad [i]

Nuevamente bajo la hipótesis de normalidad para $\bar{S}$, la credibilidad completa (k, p) para $z\bar{S} + (1-z)E(S)$ se obtiene de manera aproximada cuando

$$m \geqslant \frac{z^2\, u^2_{(1+p)/2}}{k^2} \frac{\mathrm{Var}(S)}{E^2(S)}. \tag{5.5}$$

Considerando igualdad se obtiene una expresión para el factor de credibilidad

$$z = \frac{k\sqrt{m}\, E(S)}{u_{(1+p)/2}\sqrt{\mathrm{Var}(S)}}. \tag{5.6}$$

Este valor de z excede el valor de 1 para valores suficientemente grandes de m, por lo tanto se define el factor de credibilidad como

$$z = \min \left\{ \frac{k\sqrt{m}\, E(S)}{u_{(1+p)/2}\sqrt{\mathrm{Var}(S)}},\ 1 \right\}. \tag{5.7}$$

La esperanza y varianza de S se aproximan mediante la media y varianza muestrales calculados para un número determinado de periodos m. Para los valores de $k = 0.05$ y $p = 0.9$ se tiene que

$$z = \min \left\{ \frac{\sqrt{m}}{32.89} \frac{E(S)}{\sqrt{\mathrm{Var}(S)}},\ 1 \right\}.$$

Credibilidad parcial bajo hipótesis de normalidad [ii]

Otro mecanismo para obtener el factor de credibilidad (5.7) es proponiendo como estimador de $E(S)$ a la estadística

$$z\bar{S} + (1-z)\mu, \tag{5.8}$$

en donde μ es la prima que se le cobra al asegurado o grupo de asegurados antes de tener alguna experiencia real de sus reclamaciones. A esta prima se le llama "prima de manual", pues es calculada antes de tomar en cuenta el historial de reclamaciones. De acuerdo al criterio (5.3), el estimador (5.8) tiene credibilidad completa cuando su varianza cumple

$$\mathrm{Var}(z\bar{S} + (1-z)\mu) \leqslant \frac{k^2 E^2(S)}{u^2_{(1+p)/2}}$$

Considerando igualdad y despejando z se encuentra nuevamente la expresión (5.6).

Ejemplo 5.2 *Considere nuevamente un riesgo S con distribución Poisson compuesta $S = \sum_{j=1}^{N} Y_j$, en donde N tiene distribución Poisson(λ) y Y tiene distribución exp(α). Para los valores de k y p mencionados antes, la condición de credibilidad parcial (5.5) para $z\bar{S} + (1-z)E(S)$ se reduce a*

$$\lambda m \geqslant 2164 z^2,$$

Considerando igualdad y despejando z se obtiene $z = \sqrt{\lambda m}/46.52$, en donde λm se sustituye por el número de reclamaciones totales que se hayan presentado en m periodos, siempre y cuando $\sqrt{\lambda m} \leqslant 46.52$. De esta manera la prima pura de riesgo por credibilidad parcial es la combinación lineal

$$\textit{prima} = \frac{\sqrt{\lambda m}}{46.52}\,\bar{S} + \Big(1 - \frac{\sqrt{\lambda m}}{46.52}\Big)\,E(S),$$

en donde $E(S) = \lambda/\alpha$ es el valor esperado teórico supuesto para el riesgo S y $\bar{S}$ es la experiencia observada. Es interesante notar que el factor de credibilidad $z = \sqrt{\lambda m}/46.52$ es creciente conforme la expresión λm crece, dando así paulatinamente mayor credibilidad a $\bar{S}$ y reduciendo el factor para la media teórica $E(S)$. Cuando λm alcanza el valor 2164 (cuya raíz cuadrada es 46.52) la media muestral $\bar{S}$ alcanza credibilidad completa.

5.2. Credibilidad Bayesiana

La credibilidad Bayesiana es otra forma de incorporar el historial de reclamaciones de un grupo de asegurados en el cálculo de las primas. Explicaremos a continuación de manera breve la situación general en esta perspectiva de la estadística y después aplicaremos estas ideas en dos casos particulares.

Estimación Bayesiana

En la estadística tradicional se considera el problema de estimación de un parámetro θ de una distribución de probabilidad dada $f(x;\theta)$. Para ello se considera una muestra aleatoria de esta distribución y se estudian distintos métodos para estimar θ, considerando siempre que este parámetro tiene un valor desconocido y fijo. Existe, sin embargo, un punto de vista distinto llamado estimación Bayesiana. Desde esta perspectiva se considera que θ es una variable aleatoria para la cual se asume una distribución de probabilidad $h(\theta)$ llamada *a priori*. Esta distribución refleja cierta información subjetiva o cuantitativa que el observador pueda tener sobre el parámetro θ. La distribución conjunta de una muestra aleatoria de la distribución $f(x;\theta)$ y el parámetro θ es

$$f(x_1,\ldots,x_m,\theta) = f(x_1,\ldots,x_m \,|\, \theta)\, h(\theta) = \prod_{i=1}^{m} f(x_i \,|\, \theta)\, h(\theta).$$

La distribución marginal de la muestra es entonces

$$f(x_1,\ldots,x_m) = \int_\theta \prod_{i=1}^{m} f(x_i \,|\, \theta)\, h(\theta)\, d\theta.$$

Conociendo estas funciones puede ahora calcularse la distribución condicional de θ dada la muestra como sigue:

$$\begin{aligned} g(\theta \,|\, x_1,\ldots,x_m) &= \frac{f(x_1,\ldots,x_m \,|\, \theta)\, h(\theta)}{f(x_1,\ldots,x_m)} \\ &= \frac{\displaystyle\prod_{i=1}^{m} f(x_i \,|\, \theta)\, h(\theta)}{\displaystyle\int_\theta \prod_{i=1}^{m} f(x_i \,|\, \theta)\, h(\theta)\, d\theta}. \end{aligned} \tag{5.9}$$

De esta forma la distribución *a priori* $h(\theta)$ representa cierta información que el observador conoce del parámetro θ antes de tomar la muestra, y la así llamada distribución *a posteriori* $g(\theta \,|\, x_1, \ldots, x_m)$ es la distribución modificada a la luz de la muestra aleatoria. Teniendo ahora esta distribución *a posteriori*para θ, uno puede proponer varias formas de estimar este parámetro, una de ellas es simplemente tomar la esperanza de esta distribución, es decir, un estimador Bayesiano para θ es

$$\hat{\theta} = E(\theta \,|\, x_1, \ldots, x_m) = \int \theta \, g(\theta \,|\, x_1, \ldots, x_m) \, d\theta.$$

En general, la expresión que pueda encontrarse para la distribución *a posteriori* para θ usando la fórmula (5.9) puede ser muy complicada y no ser una distribución con la cual estemos familiarizados. Ilustraremos estas ideas con un ejemplo particular en donde el resultado final, sin embargo, es una distribución conocida. Aplicaremos después la estimación Bayesiana al problema de calcular primas puras de riesgo tomando en cuenta el historial de reclamaciones.

Ejemplo 5.3 *Suponga que $X_1, \ldots, X_m$ es una muestra aleatoria de la distribución Ber(p), y suponga además que el parámetro p tiene una distribución beta(a, b), con a y b conocidos. Observe que el soporte de la distribución beta es el intervalo $[0, 1]$, de modo que en este sentido tal distribución es factible para el parámetro p. La densidad conjunta de la muestra y el parámetro es*

$$\begin{aligned} f(x_1, \ldots, x_m, p) &= f(x_1, \ldots, x_m \,|\, p) \, h(p) \\ &= p^{m\bar{x}} (1-p)^{m - m\bar{x}} \, \frac{1}{B(a,b)} \, p^{a-1} \, (1-p)^{b-1} \\ &= \frac{1}{B(a,b)} \, p^{a + m\bar{x} - 1} (1-p)^{m - m\bar{x} + b - 1}. \end{aligned}$$

Integrando respecto a p se obtiene la densidad marginal de la muestra,

$$\begin{aligned} f(x_1, \ldots, x_m) &= \frac{1}{B(a,b)} \int_0^1 p^{a + m\bar{x} - 1} \, (1-p)^{m - m\bar{x} + b - 1} \, dp \\ &= \frac{B(a + m\bar{x}, m - m\bar{x} + b)}{B(a,b)}. \end{aligned}$$

Por lo tanto, la densidad a posteriori *para p es*

$$g(p\,|\,x_1,\ldots,x_m) = \frac{1}{B(a+m\bar{x}, m-m\bar{x}+b)}\, p^{a+m\bar{x}-1}(1-p)^{m-m\bar{x}+b-1}.$$

Esta es la densidad beta$(a+m\bar{x}, m-m\bar{x}+b)$, y su esperanza puede usarse como un estimador Bayesiano para p, es decir,

$$\begin{aligned}\hat{p} &= \frac{a+m\bar{x}}{m+a+b} \\ &= \frac{m}{m+a+b}\,\bar{x} + (1-\frac{m}{m+a+b})\,\frac{a}{a+b}. \end{aligned} \tag{5.10}$$

Esta última expresión para el estimador Bayesiano es interesante pues tiene la misma forma que la que se había propuesto en la credibilidad parcial. Observe que en la ecuación (5.10) conforme el tamaño de la muestra m crece, el estimador Bayesiano para la media toma esa información y le asocia un peso mayor a la media muestral $\bar{x}$ en detrimento de la media de la distribución a priori *$a/(a+b)$.*

Ahora aplicaremos estas ideas al problema del cálculo de primas tomando en cuenta la experiencia de un riesgo. Suponga que las variables $S_1,\ldots,S_m$ representan el historial de reclamaciones en m años o periodos que se han registrado de un riesgo dado. Suponga además que estas variables son independientes y todas ellas tienen una distribución común dependiente de un parámetro desconocido θ. Por lo tanto la esperanza del riesgo $E(S)$ es una función de este parámetro θ. Bajo el enfoque Bayesiano se considera que el parámetro θ es una variable aleatoria para la cual se asume una distribución de probabilidad a priori. La esperanza a posteriori de θ, es decir, $\hat{\theta} = E(\theta\,|\,S_1,\ldots,S_m)$, representa una estimación para θ tomando en cuenta el historial $S_1,\ldots,S_m$. Así, podemos incorporar este estimador Bayesiano en el cálculo de $E(S)$ y encontrar la prima pura de riesgo. En los ejemplos que consideraremos $E(S)$ es directamente el parámetro θ y por lo tanto su estimación $\hat{\theta}$ es la prima pura de riesgo para S, es decir,

$$E(S) = \hat{\theta} = E(\theta\,|\,S_1,\ldots,S_m).$$

A esta esperanza *a posteriori* se le llama prima de credibilidad Bayesiana. Los casos que consideraremos para la distribución de S son: la distribución

Poisson con media λ, y la distribución normal con media θ. Encontraremos que en estos casos la prima pura de riesgo tiene la forma de la credibilidad parcial mencionada antes.

Ejemplo 5.4 (Modelo Poisson-gama) *Este modelo adquiere su nombre a partir de las siguientes hipótesis: se postula que cada una de las variables aleatorias independientes $S_1, \ldots, S_m$ tiene distribución Poisson con parámetro λ, el cual se considera aleatorio con distribución* a priori *gama(γ, α), con γ y α parámetros conocidos. Observe que en este modelo $E(S) = \lambda$ y se considera que los montos de las reclamaciones toman valores enteros. La función de densidad* a posteriori *de λ es, para $x > 0$,*

$$\begin{aligned}
g(\lambda \,|\, S_1, \ldots, S_m) &= \frac{f(S_1, \ldots, S_m \,|\, \lambda)\, h(\lambda)}{\displaystyle\int_0^\infty f(S_1, \ldots, S_m \,|\, \lambda)\, h(\lambda)\, d\lambda} \\
&= \frac{\displaystyle\prod_{j=1}^m (\frac{\lambda^{S_j}}{S_j!} e^{-\lambda})\, \frac{\alpha^\gamma}{\Gamma(\gamma)} \lambda^{\gamma-1} e^{-\alpha\lambda}}{\displaystyle\int_0^\infty \prod_{j=1}^m (\frac{\lambda^{S_j}}{S_j!} e^{-\lambda})\, \frac{\alpha^\gamma}{\Gamma(\gamma)} \lambda^{\gamma-1} e^{-\alpha\lambda}\, d\lambda} \\
&= \frac{\lambda^{m\bar{S}+\gamma-1} e^{-(m+\alpha)\lambda}}{\displaystyle\int_0^\infty \lambda^{m\bar{S}+\gamma-1} e^{-(m+\alpha)\lambda}\, d\lambda} \\
&= \frac{(m+\alpha)^{m\bar{S}+\gamma}}{\Gamma(m\bar{S}+\gamma)}\, \lambda^{m\bar{S}+\gamma-1} e^{-(m+\alpha)\lambda}.
\end{aligned}$$

Es decir, la densidad a posteriori *es gama$(m\bar{S}+\gamma, m+\alpha)$. Por lo tanto la prima por credibilidad, esperanza de esta densidad, es*

$$\begin{aligned}
\textit{prima} &= E(S) \\
&= \hat{\lambda} \\
&= E(\lambda \,|\, S_1, \ldots, S_m) \\
&= \frac{m\bar{S}+\gamma}{m+\alpha} \\
&= \frac{m}{m+\alpha}\, \bar{S} + \frac{\alpha}{m+\alpha}\, \frac{\gamma}{\alpha} \\
&= z\, \bar{S} + (1-z)\, \frac{\gamma}{\alpha},
\end{aligned}$$

en donde $z = m/(m+\alpha)$ *es el factor de credibilidad. Esta cantidad crece monótonamente a* 1 *cuando* m *crece a infinito, dando cada vez más credibilidad a la media muestral* $\bar{S}$ *y favoreciendo cada vez menos a la media teórica* γ/α. *Observe además que cuando* m *crece a infinito, la media de la distribución* a posteriori *converge a la media muestral límite dado por el historial de reclamaciones, y que la varianza de* λ *dada por* $Var(\lambda) = (m\bar{S}+\gamma)/(m+\alpha)^2$ *converge a cero, lo cual indica que la distribución* a posteriori *se concentra cada vez más alrededor de su media.*

Ejemplo 5.5 (Modelo normal-normal) *En este modelo se postula que cada una de las reclamaciones* $S_1, \ldots, S_m$ *tiene distribución* $N(\theta, \sigma^2)$, *en donde el parámetro* σ^2 *es conocido y la media* θ *es una variable aleatoria con distribución* $N(\mu, \eta^2)$, *con* μ *y* η^2 *conocidos. La primera hipótesis puede ser justificada en el caso cuando los montos anuales se componen de un gran número de reclamaciones individuales, para ello no es necesario suponer que las reclamaciones individuales tienen la misma distribución. La segunda hipótesis podría ser razonable si es que los parámetros* μ *y* η^2 *son tales que la probabilidad asignada a la parte negativa del eje es muy pequeña. Tenemos nuevamente en este caso la situación* $E(S) = \theta$ *y nuestro objetivo es estimar mediante estadística Bayesiana este parámetro desconocido. La función de densidad a posteriori de* θ *es*

$$
\begin{aligned}
g(\theta \,|\, S_1, \ldots, S_m) &= \frac{f(S_1, \ldots, S_m \,|\, \theta)\, h(\theta)}{\displaystyle\int_{-\infty}^{\infty} f(S_1, \ldots, S_m \,|\, \theta)\, h(\theta)\, \theta\, d\theta} \\
&= \frac{\dfrac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\sum_{j=1}^m (S_j-\theta)^2/2\sigma^2}\, \dfrac{1}{\sqrt{2\pi\eta^2}}\, e^{-(\theta-\mu)^2/2\eta^2}}{\displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\sum_{j=1}^m (S_j-\theta)^2/2\sigma^2}\, \frac{1}{\sqrt{2\pi\eta^2}}\, e^{-(\theta-\mu)^2/2\eta^2}\, d\theta}.
\end{aligned}
$$

Nos concentramos en analizar únicamente el exponente. Tenemos que

$$
\begin{aligned}
-\sum_{j=1}^m \frac{(S_j-\theta)^2}{2\sigma^2} - \frac{(\theta-\mu)^2}{2\eta^2} &= -\theta^2\Big(\frac{m}{2\sigma^2} + \frac{1}{2\eta^2}\Big) + 2\theta\Big(\frac{m\bar{S}}{2\sigma^2} + \frac{\mu}{2\eta^2}\Big) \\
&\quad -\Big(\sum_{j=1}^m \frac{S_j^2}{2\sigma^2}\Big) - \frac{\mu^2}{2\eta^2}.
\end{aligned}
$$

Completando el cuadrado en θ, esta expresión se puede escribir como sigue

$$-\frac{[\theta-(\frac{m\bar{S}}{\sigma^2}+\frac{\mu}{\eta^2})/(\frac{m}{\sigma^2}+\frac{1}{\eta^2})]^2}{2(\frac{m}{\sigma^2}+\frac{1}{\eta^2})^{-1}}+\frac{(\frac{m\bar{S}}{\sigma^2}+\frac{\mu}{\eta^2})^2}{2(\frac{m}{\sigma^2}+\frac{1}{\eta^2})}-(\sum_{j=1}^{m}\frac{S_j^2}{2\sigma^2})-\frac{\mu^2}{2\eta^2}.$$

Este exponente aparece tanto en el numerador como en el denominador, y como los últimos dos sumandos no dependen de θ éstos desaparecen completamente al efectuar el cociente. Lo que resulta es nuevamente la distribución normal

$$g(\theta\,|\,S_1,\ldots,S_m)=\frac{1}{\sqrt{2\pi(\frac{m}{\sigma^2}+\frac{1}{\eta^2})^{-1}}}\exp\{-\frac{[\theta-(\frac{m\bar{S}}{\sigma^2}+\frac{\mu}{\eta^2})/(\frac{m}{\sigma^2}+\frac{1}{\eta^2})]^2}{2(\frac{m}{\sigma^2}+\frac{1}{\eta^2})^{-1}}\}.$$

La media de esta distribución es entonces la prima por credibilidad, es decir,

$$\begin{aligned}
\textit{prima} &= E(S)\\
&= \hat{\theta}\\
&= E(\theta\,|\,S_1,\ldots,S_m)\\
&= (\frac{m\bar{S}}{\sigma^2}+\frac{\mu}{\eta^2})/(\frac{m}{\sigma^2}+\frac{1}{\eta^2})\\
&= \frac{m\eta^2}{m\eta^2+\sigma^2}\,\bar{S}+\frac{\sigma^2}{m\eta^2+\sigma^2}\,\mu\\
&= z\,\bar{S}+(1-z)\,\mu,
\end{aligned}$$

en donde $z=m\eta^2/(m\eta^2+\sigma^2)$ es el factor de credibilidad, el cual tiene nuevamente comportamiento monótono creciente a 1 conforme m crece a infinito. Observe además que $Var(\theta)=(\frac{m}{\sigma^2}+\frac{1}{\eta^2})^{-1}$ y que esta cantidad converge a cero cuando el tamaño de muestra m crece a infinito, indicando nuevamente que la distribución a posteriori *se concentra cada vez más alrededor de su media.*

En los ejemplos anteriores ha sucedido que la distribución *a posteriori* $g(\theta\,|\,x_1,\ldots,x_m)$ para el parámetro de interés θ pertenece a la misma familia de la distribución *a priori* $h(\theta)$. Cuando esta situación ocurre, se dice que la distribución de la muestra y la distribución *a priori* para el parámetro son conjugadas, o que estas dos familias de distribuciones son conjugadas. En nuestros ejemplos, la distribuciones Poisson y gama son conjugadas, así como la distribución normal es conjugada consigo misma. Mediante estos dos ejemplos se ha mostrado que la propiedad de conjugación para dos familias de distribuciones tiene ventajas, pues al poder identificar la distribución *a posteriori* pueden calcularse sus características numéricas mediante las fórmulas conocidas, en particular, su esperanza.

Comentarios y referencias

En este capítulo hemos planteado el problema de incorporar el historial de reclamaciones de un asegurado o grupo de asegurados en el cálculo de la prima. Nuestra perspectiva ha sido introductoria y únicamente los elementos iniciales de algunos modelos se han mencionado. Una exposición más completa sobre la teoría de la credibilidad puede encontrarse en los textos de Klugman *et al.* [23] y Boland [6]. El libro de Bühlmann y Gisler [9] está dedicado completamente al desarrollo de este tema.

5.3. Ejercicios

Credibilidad completa

128. *Lanzamiento de una moneda.* Sean $S_1, \ldots, S_m$ los resultados de m lanzamientos sucesivos de una moneda con valores 0 y 1, dependiendo si se obtiene una cara de la moneda o la otra.

   a) Suponiendo que la moneda está equilibrada, use la aproximación normal para encontrar el valor de m para obtener la credibilidad completa (k, p) para $\bar{S} = (S_1 + \cdots + S_m)/m$. Suponga $k = 0.1$ y $p = 0.9$.

   b) Suponga ahora que no sabemos que la moneda está equilibrada pero la probabilidad θ de obtener el valor 1 es tal que $1/4 \leqslant \theta \leqslant 3/4$. Obtenga una aproximación del valor de m para la credibilidad completa de la media muestral. Suponga nuevamente $k = 0.1$ y $p = 0.9$.

129. Encuentre el valor de m para obtener la credibilidad completa (k,p) para $\bar{S} = (S_1 + \cdots + S_m)/m$ de un riesgo S con distribución $\mathrm{N}(5,1)$. Suponga $k = 0.15$ y $p = 0.9$. Observe que en esta situación conocemos $E(S)$ y simplemente se pide aplicar la condición para la credibilidad completa de la media muestral.

130. Suponga que el número de observaciones m de un riesgo es fijo y se desea que la media muestral de estas observaciones cumpla la condición (5.2) para la credibilidad completa (k,p) modificando los valores de k y p. Demuestre que debe cumplirse alguna de las siguientes dos condiciones:

    a) $k \geqslant \sqrt{\dfrac{u_{(1+p)/2}\mathrm{Var}(S)}{mE^2(S)}}$.

    b) $p \leqslant 2\Phi\Big(\sqrt{\dfrac{mk^2E^2(S)}{\mathrm{Var}(S)}}\Big) - 1$.

131. *Credibilidad para el número de reclamaciones.* Sean $N_1, \ldots, N_m$ el número de reclamaciones durante m periodos de un riesgo S con distribución Poisson compuesta, es decir,

$$S = \sum_{j=1}^{N} Y_j,$$

en donde N tiene distribución Poisson(λ) con λ desconocido. Encuentre un criterio para darle credibilidad completa (k,p) a

$$\bar{N} = \frac{1}{m}\,(N_1 + N_2 + \cdots + N_m)$$

como el verdadero valor de λ. Suponga $k = 0.05$ y $p = 0.9$.

Credibilidad parcial

132. Use la aproximación normal para encontrar la prima por credibilidad parcial, a partir de una muestra $S_1, \ldots, S_m$, de un riesgo S con distribución Pareto$(4,3)$. Suponga $k = 0.1$ y $p = 0.9$.

133. Use la aproximación normal para encontrar la prima por credibilidad parcial con $k = 0.1$ y $p = 0.9$ considerando el historial de reclamaciones $S_1, \ldots, S_m$ de un riesgo de la forma $S = \sum_{j=1}^{N} Y_j$, en donde:

   a) N se distribuye $\text{bin}(n, p)$ con $n = 100$ y $p = 1/2$, y Y_j se distribuye $\exp(\alpha)$ con $\alpha = 1/10$.

   b) N se distribuye $\text{bin neg}(\alpha, p)$ con $\alpha = 10$ y $p = 1/3$, y Y_j se distribuye $\exp(\alpha)$ con $\alpha = 1/10$.

Estimación Bayesiana

134. Suponga que el número de reclamaciones X de un cierto riesgo sigue una distribución Poisson de parámetro λ. Se considera que el parámetro λ es desconocido y se le modela como una variable aleatoria con distribución $\exp(\mu)$. Demuestre que la distribución *a posteriori* o no condicional de X es la siguiente distribución geométrica:

$$P(X = x) = \mu \left(\frac{1}{1 + \mu} \right)^{x+1} \quad \text{para} \quad x = 0, 1, 2, \ldots$$

135. Suponga que N tiene una distribución $\text{Poisson}(\lambda)$, y que condicionada al valor de N la variable X tiene distribución $\text{bin}(N, p)$. Demuestre que la distribución no condicional de X es $\text{Poisson}(\lambda p)$.

Capítulo 6

Procesos estocásticos

En este capítulo se presenta una introducción breve al tema de los procesos estocásticos. Se explican algunos conceptos y propiedades generales de este tipo de modelos matemáticos y se estudian brevemente algunos ejemplos particulares de procesos estocásticos. Este material será usado en la última parte del libro en donde consideraremos algunos modelos estocásticos dependientes del tiempo aplicados a los seguros. Supondremos como elemento base un espacio de probabilidad $(\Omega, \mathscr{F}, P)$.

6.1. Conceptos elementales

Definición 6.1 *Un proceso estocástico es una colección de variables aleatorias* $\{X_t : t \in T\}$*, parametrizada por un conjunto* T *llamado espacio parametral. A los valores de estas variables aleatorias los llamaremos estados del proceso.*

El conjunto T usualmente se interpreta como un conjunto de tiempos. Así, para cada tiempo t se tiene la variable aleatoria X_t, la cual representa el estado o valor del sistema en estudio al tiempo t. Por ejemplo, en finanzas X_t puede representar el precio de un bien o una acción al tiempo t;

en el contexto de los seguros X_t puede representar el balance al tiempo t de una aseguradora respecto de una cartera de asegurados; en física X_t puede modelar la posición o velocidad de una partícula que se mueve debido a los múltiples choques con otras partículas circundantes. Se dice que un proceso es a tiempo discreto en caso de que el conjunto de parámetros sea un conjunto discreto, por ejemplo, $T = \{0, 1, 2, \ldots\}$. Así, el sistema se observa o mide en tiempos discretos. En este caso el proceso consiste de una sucesión infinita de variables aleatorias que se denota por $\{X_n : n \geqslant 0\}$ o $\{X_n : n = 0, 1, 2, \ldots\}$ y consiste explícitamente de la sucesión

$$X_0, X_1, X_2, \ldots$$

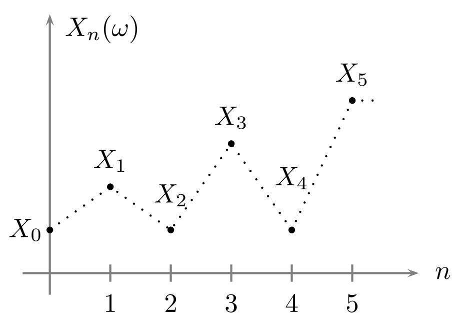


Figura 6.1

Se dice en cambio que el proceso es a tiempo continuo cuando el conjunto de parámetros consiste de un subintervalo de $\mathbb{R}$, por ejemplo, $T = [0, \infty)$, y se considera entonces que el sistema se observa continuamente en el tiempo. En este caso el proceso se puede escribir como sigue:

$$\{X_t : t \geqslant 0\}.$$

Así, tomaremos como convención que si el subíndice es n, el proceso es a tiempo discreto y si es t, el tiempo es continuo. Un proceso estocástico también puede considerarse como una función de dos variables:

$$X : T \times \Omega \to \mathbb{R},$$

tal que para cada $t \in T$, la función $\omega \mapsto X_t(\omega)$ es una variable aleatoria, mientras que para cada ω en Ω, la función $t \mapsto X_t(\omega)$ es una trayectoria o realización del proceso. Con este modelo se pretende representar la evolución aleatoria de un sistema a lo largo del tiempo. En la Figura 6.1 se muestra una trayectoria de un proceso a tiempo discreto y en la Figura 6.2 se ilustra el caso cuando el tiempo es continuo.

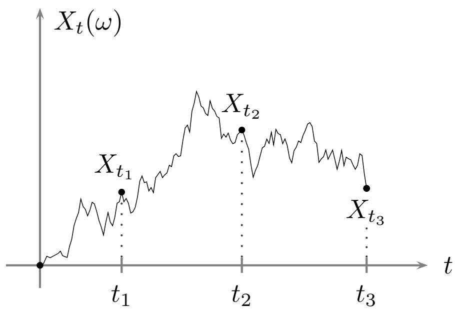


Figura 6.2

Una de las características más importantes que distingue a los distintos tipos de procesos estocásticos es el grado de dependencia probabilística que pudiera establecerse entre las variables aleatorias que conforman el proceso. La especificación de estas dependencias nos permitirá definir más adelante algunos tipos de procesos estocásticos particulares. Veremos a continuación algunas propiedades generales que pueden cumplir los procesos estocásticos.

Propiedad de Markov

Se dice que un proceso estocástico a tiempo discreto $\{X_n : n \geqslant 0\}$ cumple la propiedad de Markov si para cualesquiera estados o valores $x_0, x_1, \ldots, x_{n+1}$ se cumple la identidad

$$P(X_{n+1} = x_{n+1} \,|\, X_0 = x_0, \ldots, X_n = x_n) = P(X_{n+1} = x_{n+1} \,|\, X_n = x_n).$$

De esta forma la probabilidad del evento futuro $(X_{n+1} = x_{n+1})$ sólo depende del evento $(X_n = x_n)$, mientras que la información correspondiente al evento pasado $(X_0 = x_0, \ldots, X_{n-1} = x_{n-1})$ es irrelevante. Puede observarse que mediante esta propiedad se está especificando un tipo de dependencia entre las variables aleatorias del proceso. La propiedad de Markov puede extenderse sin dificultad al caso cuando el proceso es a tiempo continuo y daremos los detalles más adelante.

Procesos con incrementos independientes

Se dice que un proceso estocástico a tiempo continuo $\{X_t : t \geqslant 0\}$ tiene incrementos independientes si para cualesquiera tiempos $0 \leqslant t_1 < t_2 < \cdots < t_n$, las variables aleatorias

$$X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$$

son independientes. Esto quiere decir que los desplazamientos que tiene el proceso en estos intervalos disjuntos de tiempo son independientes unos de otros. Cuando el proceso es a tiempo discreto la definición es completamente análoga.

Procesos con incrementos estacionarios

Se dice que un proceso estocástico a tiempo continuo $\{X_t : t \geqslant 0\}$ tiene incrementos estacionarios si para cualesquiera tiempos $s < t$, y para cualquier $h > 0$, las variables aleatorias $X_{t+h} - X_{s+h}$ y $X_t - X_s$ tienen la misma distribución de probabilidad. Es decir, el incremento que tiene el proceso entre los tiempos s y t sólo depende de estos tiempos a través de la diferencia $t - s$, y no de los valores específicos de s y t. Cuando el proceso es a tiempo discreto la definición es nuevamente análoga.

Las propiedades generales de los procesos estocásticos que hemos mencionado serán identificadas en los varios modelos estocásticos que veremos más adelante.

6.2. Filtraciones y tiempos de paro

Definiremos en esta sección el concepto de filtración y en particular definiremos la filtración generada por un proceso estocástico. Con estos elementos definiremos el concepto de tiempo de paro. Estos conceptos serán usados principalmente en los resultados que mencionaremos sobre martingalas, las cuales constituyen un tipo de proceso estocástico importante.

Definición 6.2 *Dado un espacio de probabilidad $(\Omega, \mathscr{F}, P)$, a una familia de sub σ-álgebras $(\mathscr{F}_t)_{t\geqslant 0}$ de $\mathscr{F}$ se le llama filtración si para $0 \leqslant s \leqslant t$, se cumplen las contenciones*

$$\mathscr{F}_s \subseteq \mathscr{F}_t \subseteq \mathscr{F}.$$

Al espacio $(\Omega, \mathscr{F}, P, (\mathscr{F}_t)_{t\geqslant 0})$ se le llama espacio de probabilidad filtrado.

El concepto de filtración puede definirse de manera análoga también para tiempos discretos. En ese caso se escribe $(\mathscr{F}_n)_{n\geqslant 0}$ y naturalmente la condición de monotonía es $\mathscr{F}_n \subseteq \mathscr{F}_m \subseteq \mathscr{F}$ para $0 \leqslant n \leqslant m$. El concepto de adaptabilidad de un proceso estocástico respecto de una filtración que se enuncia a continuación también tiene su contraparte discreta.

Definición 6.3 *Se dice que un proceso a tiempo continuo $\{X_t : t \geqslant 0\}$ es adaptado a una filtración $(\mathscr{F}_t)_{t\geqslant 0}$ si para cada $t \geqslant 0$, la variable X_t es $\mathscr{F}_t$-medible.*

Todo proceso estocástico $\{X_t : t \geqslant 0\}$ determina una filtración natural dada por

$$\mathscr{F}_t = \sigma\{X_s : 0 \leqslant s \leqslant t\}, \quad t \geqslant 0.$$

No es difícil ver que todo proceso es adaptado a su filtración natural. En este caso a la σ-álgebra $\mathscr{F}_t$ se le interpreta como la historia del proceso al tiempo t, pues en ella se encuentran todos los posibles eventos o sucesos que el proceso haya tenido hasta ese momento.

Definición 6.4 *Un tiempo de paro respecto de una filtración* $(\mathscr{F}_t)_{t\geqslant 0}$ *es una función* $\tau : \Omega \to [0,\infty]$ *tal que* $(\tau \leqslant t) \in \mathscr{F}_t$ *para cada* $t \geqslant 0$.

Esto es, un tiempo de paro es una variable aleatoria no negativa, con posible valor infinito, tal que puede determinarse la ocurrencia o no ocurrencia del evento $(\tau \leqslant t)$ sólo con la información o historia del proceso hasta el tiempo t. Por ejemplo, y sin entrar en mayores detalles técnicos, el tiempo aleatorio que transcurre hasta la ocurrencia de un cierto evento es un tiempo de paro, este tiempo aleatorio puede ser, por ejemplo, el tiempo que transcurre hasta la llegada de la n-ésima reclamación a una compañía aseguradora. Pueden definirse también tiempos de paro para filtraciones a tiempo discreto de manera completamente análoga a la definición anterior. Finalmente mencionaremos que un tiempo de paro τ es finito si $P(\tau = \infty) = 0$.

En las siguientes secciones estudiaremos muy brevemente algunos tipos de procesos estocásticos que usaremos en la última parte del texto.

6.3. Caminatas aleatorias

Una caminata aleatoria simple sobre el conjunto de números enteros $\mathbb{Z}$ es un proceso estocástico a tiempo discreto $\{X_n : n = 0, 1, \ldots\}$ que evoluciona de la siguiente forma: iniciando en el estado 0, al tiempo 1 el proceso puede pasar al estado $+1$ con probabilidad p, o al estado -1 con probabilidad q, en donde $p + q = 1$. Se usa la misma regla para los siguientes tiempos, es decir, estando en el estado k, al siguiente instante el proceso pasa al estado $k + 1$ con probabilidad p, o al estado $k - 1$ con probabilidad q. El valor de X_n es el estado donde se encuentra el proceso al tiempo n y puede escribirse como

$$X_n = \xi_1 + \xi_2 + \cdots + \xi_n,$$

en donde $\xi_1, \xi_2, \ldots$ es una sucesión de variables aleatorias independientes con idéntica distribución

$$\begin{aligned} P(\xi = +1) &= p, \\ P(\xi = -1) &= q = 1 - p. \end{aligned}$$

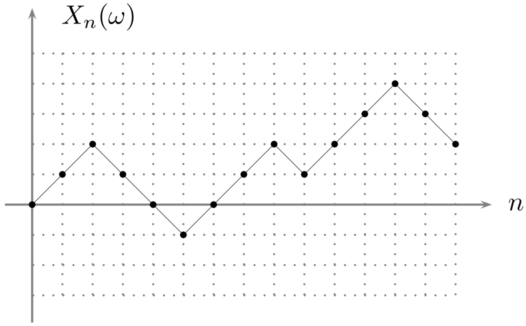


Figura 6.3

Una posible trayectoria de este proceso se muestra en la Figura 6.3 y es uno de los mejores ejemplos para entender el carácter aleatorio en la dinámica de un proceso estocástico a lo largo del tiempo. La independencia de las variables aleatorias $\xi_1, \xi_2, \ldots$ tiene como consecuencia que la caminata aleatoria simple cumpla la propiedad de Markov y que sea un proceso estocástico con incrementos independientes y estacionarios. En términos generales el objetivo es estudiar el comportamiento de la v.a. X_n al paso del tiempo.
Uno de los elementos fundamentales para estudiar a las caminatas aleatorias está dado por las probabilidades de transición. Para cualesquiera enteros i y j, las probabilidades de transición en un paso de la caminata aleatoria simple son:

$$P(X_{n+1} = j \,|\, X_n = i) = \begin{cases} p & \text{si } j = i+1, \\ q & \text{si } j = i-1, \\ 0 & \text{otro caso.} \end{cases}$$

Como estas probabilidades no dependen de n, se dice que son homogéneas en el tiempo, es decir, son las mismas para cualquier valor de n. Pueden considerarse también caminatas aleatorias que inician en cualquier valor entero o que permiten saltar de un estado en sí mismo al siguiente instante (es decir, no cambiar de estado), o bien, que los saltos no sean unitarios ni los tiempos en los que efectúan los saltos sean necesariamente tiempos enteros, e incluso pueden definirse caminatas en dimensiones mayores, por ejemplo, en $\mathbb{Z}^2$. Las caminatas aleatorias son procesos que pueden representar versiones discretas de procesos a tiempo continuo más complejos. En el siguiente capítulo estudiaremos algunas características de un proceso a tiempo discreto muy similar a una caminata aleatoria que representará el balance de una aseguradora. Para las caminatas aleatorias se pueden plantear diversos problemas matemáticos, por ejemplo, encontrar la distribución de la variable X_n para cualquier valor de n, o encontrar la probabilidad de que en algún momento la caminata regrese a su posición de origen, o que nunca tome valores negativos, o alguna otra condición interesante sobre sus trayectorias. En el texto de Norris [26] o Resnick [29] se puede encontrar la solución a algunos de estos problemas, así como un estudio más completo sobre caminatas aleatorias.

6.4. Cadenas de Markov

Este tipo de procesos es también de amplia aplicación y se cuenta con una teoría matemática bastante desarrollada.

Definición 6.5 *Una cadena de Markov es un proceso estocástico a tiempo discreto* $\{X_n : n = 0, 1, \ldots\}$ *con espacio de estados discreto y tal que satisface la propiedad de Markov, esto es, para cualquier entero* $n \geqslant 0$ *y cualesquiera estados* $x_0, x_1, \ldots, x_{n+1}$ *se cumple la identidad*

$$\begin{aligned} P(X_{n+1} = x_{n+1} \,|\, X_n = x_n, \ldots, X_1 = x_1, X_0 = x_0) \\ = P(X_{n+1} = x_{n+1} \,|\, X_n = x_n). \end{aligned}$$

A las probabilidades condicionales mencionadas en la definición anterior se les llama probabilidades de transición del tiempo n al tiempo $n+1$, y se les denota por $p_{x_n,x_{n+1}}(n, n+1)$. Adicionalmente, se dice que la cadena es estacionaria u homogénea en el tiempo si estas probabilidades no dependen explícitamente de los tiempos particulares n y $n+1$, sino únicamente de los estados involucrados. De esta forma, si de manera general se considera la probabilidad de transición del estado i al estado j de una unidad de tiempo cualquiera a la siguiente unidad de tiempo, la probabilidad de transición del primer estado al segundo se escribe como p_{ij}, o $p_{ij}(1)$, o también como $p_{ij}^{(1)}$, es decir, sin importar el valor del entero n:

$$p_{ij} = P(X_n = j \mid X_{n-1} = i).$$

A estas probabilidades se les llama probabilidades de transición en un paso. En general se considera como espacio de estados para una cadena de Markov el conjunto discreto $\{0, 1, \ldots\}$ o algún subconjunto finito de él. Haciendo variar los valores de los índices i y j en este espacio de estados se forma la matriz de probabilidades de transición en un paso:

$$P = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

En la entrada (i, j) de esta matriz aparece la probabilidad p_{ij}. Así, todas las entradas son números no negativos y la suma de ellos en cada renglón es 1. A tales matrices cuadradas se les llama matrices estocásticas. Recíprocamente, a partir de una matriz con estas características junto con una distribución inicial para X_0, se puede construir una cadena de Markov. Es por ello que a una matriz estocástica también se le puede llamar cadena de Markov. La cadena puede iniciar en un estado particular o en cualquiera de ellos de manera aleatoria por medio de una distribución de probabilidad inicial $(\pi_0, \pi_1, \ldots)$.

Probabilidades de transición en n pasos

Uno de los problemas que se pueden plantear para las cadenas de Markov es el de encontrar las probabilidades de transición de un estado a otro estado cualquiera en n pasos sucesivos. Estas probabilidades se definen como sigue

$$p_{ij}(n) = P(X_{m+n} = j \mid X_m = i).$$

Por la estacionariedad de las probabilidades de transición en un paso, el lado derecho de esta identidad no depende realmente del subíndice m. La solución al problema planteado se resuelve a través de la igualdad de Chapman-Kolmogorov, la cual establece que para cualesquiera estados i y j y cualesquiera tiempos $0 \leqslant r \leqslant n$,

$$p_{ij}(n) = \sum_k p_{ik}(r)\, p_{kj}(n-r).$$

De este modo la probabilidad de pasar del estado i al estado j en n pasos se descompone en la probabilidad de las distintas trayectorias que inician en i y terminan en j, pero que al tiempo intermedio r visitan el estado k, sumando estas probabilidades sobre todos los posibles estados intermedios k. Cuando $n = 0$ se define

$$p_{ij}(0) = \begin{cases} 1 & \text{si} \;\; i = j. \\ 0 & \text{si} \;\; i \neq j. \end{cases}$$

Por lo tanto, aplicando la identidad de Chapman-Kolmogorov $n-1$ veces tomando siempre como tiempo intermedio $r = 1$,

$$p_{ij}(n) = (P \cdot P \cdots P)_{ij} = (P^n)_{ij}.$$

Es decir, la probabilidad buscada $p_{ij}(n)$ es la entrada (i, j) de la matriz potencia P^n. Así, el problema de encontrar las probabilidades de transición en n pasos se reduce al problema de calcular las potencias de la matriz de probabilidades de transición en un paso. Las cadenas de Markov son modelos ampliamente conocidos y la propiedad de Markov hace que estos procesos sean muy atractivos de estudiar, pues dicha propiedad hace que ciertas probabilidades sean fáciles de calcular, como por ejemplo las probabilidades $p_{ij}(n)$ que acabamos de mencionar. Las cadenas de Markov con frecuencia aparecen dentro de otros modelos estocásticos de interés teórico y aplicado. Por ejemplo, el proceso de riesgo a tiempo discreto que presentaremos en el siguiente capítulo resulta ser una cadena de Markov. En los textos de Karlin y Taylor [20] o Norris [26] el lector puede encontrar una exposición muy detallada de las cadenas de Markov.

Distribuciones discretas tipo fase

En esta sección definiremos un tipo de distribución de probabilidad discreta que surge a través de una cadena de Markov. Para ello, la siguiente característica será de interés.

Definición 6.6 *Se dice que el estado i de una cadena de Markov es absorbente si*

$$p_{ii} = 1.$$

En caso contrario se dice que el estado es no absorbente.

Esto quiere decir que si en algún momento la cadena visita un estado absorbente, el proceso permanece en dicho estado el resto del tiempo. Consideremos entonces una cadena de Markov $\{X_n : n \geqslant 0\}$ con espacio de estados $\{0, 1, 2, \ldots, k\}$, en donde el estado 0 es absorbente y los estados $1, 2, \ldots, k$ son no absorbentes. Esta cadena tiene una matriz de probabilidades de transición en un paso compuesta por cuatro bloques de la siguiente forma:

$$P = \begin{pmatrix} 1 & \mathbf{0} \\ b^T & B \end{pmatrix},$$

en donde $\mathbf{0}$ es el vector renglón $(0, 0, \ldots, 0)$ de dimensión k, b es el vector rengón $(b_1, b_2, \ldots, b_k)$ con no todas sus entradas cero en donde el superíndice T indica transpuesta y B es una matriz cuadrada de $k \times k$ tal que la matriz completa es estocástica. Esta última condición se puede expresar mediante la ecuación

$$b^T + Be^T = e^T, \tag{6.1}$$

en donde e representa el vector renglón $(1, 1, \ldots, 1)$ de dimensión k. La idea es dejar que la cadena inicie en alguno de los estados no absorbentes $1, 2, \ldots, k$ o bien a través de una distribución inicial $\pi = (\pi_1, \pi_2, \ldots, \pi_k)$ sobre esos estados y observar el momento aleatorio τ (posiblemente infinito) en el que el proceso es atrapado en el estado absorbente 0. Esta variable aleatoria resulta ser un tiempo de paro y a la distribución de probabilidad de este tiempo de espera hasta la absorción se le llama distribución tipo fase. Más generalmente, tomaremos como distribución de probabilidad inicial al vector

$$\pi' = (\pi_0, \pi_1, \ldots, \pi_k),$$

en donde π_0 es la probabilidad de que la cadena inicie en el estado absorbente

0 y el subvector

$$\pi = (\pi_1, \pi_2, \ldots, \pi_k)$$

denota entonces el vector inicial cuyas entradas son las probabilidades iniciales de los estados no absorbentes. Cuando $\pi_0 = 0$ el vector π es efectivamente una distribución de probabilidad. Cuando $\pi_0 > 0$, se tiene que $P(\tau = 0) = \pi_0 > 0$. Puesto que, partiendo de un estado no absorbente, es posible hacer varias transiciones entre los estados no absorbentes antes de la absorción, en las fórmulas relacionadas con esta distribución tipo fase aparece el subvector π y no el vector completo π'.

Definición 6.7 *A la distribución de probabilidad de la variable aleatoria τ definida por*

$$\tau = \text{mín}\,\{n \geqslant 0 \,|\, X_n = 0\}$$

se le llama distribución tipo fase discreta con subdistribución inicial π y submatriz de transición B. Esto se escribe

$$\tau \sim PH(\pi, B).$$

La notación PH para la distribución del tiempo de absorción proviene de las primeras letras del término en inglés *phase*. En general no parece fácil encontrar la distribución de probabilidad del tiempo de absorción τ, sin embargo, como veremos más adelante, existen fórmulas matriciales sorprendentemente cortas para ciertas características de τ que facilitan el cálculo de probabilidades con este tipo de distribuciones. Antes de mencionar estas fórmulas veamos el ejemplo más sencillo de distribución discreta tipo fase.

Ejemplo 6.1 (Distribución geométrica) *Considere la cadena de Markov $\{X_n : n \geqslant 0\}$ con dos estados: un estado absorbente 0 y un estado no absorbente 1. La matriz de probabilidades de transición en un paso es*

$$P = \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}.$$

Suponga que la cadena empieza en el estado 1. No es difícil comprobar que *la distribución del tiempo de absorción τ es geo(p), esto es,*

$$P(\tau = n) = (1-p)^{n-1}\, p \quad n = 0, 1, 2, \ldots$$

Si se postula como distribución inicial a $\pi' = (\pi_0, \pi_1) = (1-\theta, \theta)$ sobre la totalidad de los estados $\{0, 1\}$ con $0 < \theta < 1$, entonces la distribución de τ es

$$P(\tau = n) = \begin{cases} 1-\theta & \textit{si} \;\; n = 0, \\ \theta(1-p)^{n-1}\, p & \textit{si} \;\; n \geqslant 1. \end{cases}$$

Concluimos esta sección mencionando algunas fórmulas generales sobre las distribuciones tipo fase discretas. Más adelante veremos este mismo tipo de distribuciones en su versión continua.

a) Usando la identidad (6.1) y el método de inducción, se puede demostrar que la matriz de probabilidades de transición en $n \geqslant 1$ pasos es

$$P^n = \begin{pmatrix} 1 & \mathbf{0} \\ e^T - B^n\, e^T & B^n \end{pmatrix}.$$

b) Partiendo de un estado no absorbente i, la probabilidad de estar en el estado no absorbente j después de $n-1$ pasos y en el siguiente paso llegar al estado absorbente es, por la ecuación de Chapman-Kolmogorov,

$$\pi_i\, (B^{n-1})_{ij}\, b_j^T.$$

Sumando sobre los valores de los índices i y j se obtiene la función de probabilidad del tiempo de absorción τ:

$$f(n) = \begin{cases} \pi_0 & \text{si} \;\; n = 0, \\ \pi\, B^{n-1}\, b^T & \text{si} \;\; n \geqslant 1. \end{cases}$$

c) La función de distribución del tiempo de absorción τ es

$$F(n) = P(\tau \leqslant n) = 1 - P(\tau > n),$$

en donde $P(\tau > n)$ es la probabilidad de que al tiempo n la cadena no se encuentre en el estado absorbente. La probabilidad de pasar del estado no absorbente i al estado no absorbente j en n pasos es $\pi_i\, (B^n)_{ij}$.

La suma sobre i y sobre j no absorbentes de estas probabilidades se puede escribir como $\pi\, B^n\, e^T$. Por lo tanto,

$$F(n) = \begin{cases} \pi_0 & \text{si} \;\; n = 0, \\ 1 - \pi\, B^n\, e^T & \text{si} \;\; n \geqslant 1. \end{cases}$$

6.5. Proceso de Poisson

En esta sección recordaremos la definición y algunos aspectos elementales de uno de los procesos estocásticos de mayor importancia: el proceso de Poisson.

Definición 6.8 *Un proceso estocástico de tiempo continuo* $\{N_t : t \geqslant 0\}$ *con espacio de estados el conjunto discreto* $\{0, 1, \ldots\}$ *es un proceso de Poisson de parámetro o intensidad* $\lambda > 0$ *si cumple las siguientes propiedades:*

a) $N_0 = 0$.

b) Tiene incrementos independientes.

c) La variable incremento $N_{t+s} - N_s$ *tiene distribución Poisson*(λt)*, para cualesquiera* $s \geqslant 0$, $t > 0$.

Observe que la tercera propiedad establece implícitamente que los incrementos del proceso de Poisson son estacionarios, pues la distribución de $N_{t+s} - N_s$ no depende de s. El proceso de Poisson se utiliza para modelar situaciones de conteo de ocurrencias de un evento particular en un intervalo de tiempo dado. Por ejemplo, N_t puede representar el número de llamadas recibidas en un número telefónico de emergencias, o el número de accidentes ocurridos en un cierto lugar o el número de clientes que buscan un servicio durante el intervalo de tiempo $[0, t]$. En particular, tomando $s = 0$ en la tercera propiedad en la definición, se tiene que la variable N_t tiene distribución Poisson(λt) para cualquier $t \geqslant 0$. Esto es

$$P(N_t = n) = e^{-\lambda t}\, \frac{(\lambda t)^n}{n!}, \qquad n = 0, 1, 2, \ldots$$

Por lo tanto,

$$\begin{aligned} E(N_t) &= \lambda t, \\ \text{y} \quad \mathrm{Var}(N_t) &= \lambda t. \end{aligned}$$

Así, para cada tiempo t desconocemos el valor que toma N_t, pero sabemos que N_t tiene distribución Poisson(λt). En nuestro caso, usaremos el proceso de Poisson para modelar el número de reclamaciones que llegan a una compañía aseguradora hasta un tiempo t cualquiera. Una posible trayectoria de un proceso de Poisson se muestra en la Figura 6.4.

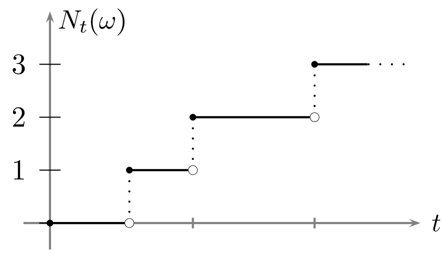


Figura 6.4

Tiempos de estancia exponenciales

No resulta evidente a partir de la definición que hemos dado del proceso de Poisson, pero la dinámica de este proceso es la siguiente: el proceso empieza en el estado cero y permanece en ese estado un tiempo aleatorio exponencial de parámetro λ, brinca después al estado uno y nuevamente permanece en ese nuevo estado otro tiempo exponencial con el mismo parámetro e independiente del tiempo de estancia anterior, después brinca al estado dos y así sucesivamente. Es por ello que una posible trayectoria de este proceso tiene las características que se muestran en la Figura 6.4: saltos unitarios hacia arriba y tiempos de estancia exponenciales. De este modo pueden definirse las variables $T_1, T_2, \ldots$ como los tiempos que transcurren entre un salto y el siguiente. A tales variables aleatorias se les llama también tiempos de interarribo. En el siguiente resultado que enunciamos sin demostración, se muestra la relación entre la propiedad de Markov del proceso de Poisson y los tiempos de estancia exponenciales.

Proposición 6.1 *La propiedad de Markov del proceso de Poisson implica que los tiempos de estancia $T_1, T_2, \ldots$ son independientes y cada uno de ellos tiene distribución exponencial. Recíprocamente, dada una sucesión de v.a.s independientes $T_1, T_2, \ldots$ con idéntica distribución exponencial, el proceso de conteo que puede construirse a partir de esta sucesión cumple la propiedad de Markov.*

Pérdida de memoria y sus consecuencias

Una de las propiedades que caracterizan de manera única a la distribución exponencial dentro del conjunto de distribuciones absolutamente continuas es que satisface la propiedad de pérdida de memoria, esto es, si T tiene distribución $\exp(\lambda)$, entonces para cualesquiera tiempos $s, t > 0$ se cumple la igualdad

$$P(T > t + s \,|\, T > s) = P(T > t).$$

En otras palabras, condicionada al evento $(T > s)$, la variable $T - s$ sigue teniendo distribución $\exp(\lambda)$. Esto significa que, para un valor de $s \geqslant 0$ fijo, todos los tiempos de interarribo a partir de s, incluyendo el primero, siguen teniendo distribución $\exp(\lambda)$, y por lo tanto el proceso de conteo de eventos a partir del tiempo s es nuevamente un proceso de Poisson. A esta propiedad se le conoce como la propiedad de renovación del proceso de Poisson. La situación se muestra gráficamente en la Figura 6.5.

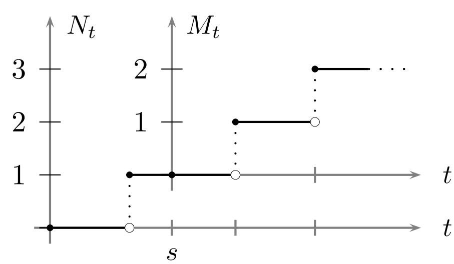


Figura 6.5

Por lo tanto, tenemos el siguiente resultado.

Proposición 6.2 *Si $\{N_t : t \geqslant 0\}$ es un proceso de Poisson de parámetro λ, entonces para cualquier $s \geqslant 0$, el proceso dado por $M_t = N_{t+s} - N_s$ es nuevamente un proceso de Poisson con el mismo parámetro.*

Observemos que la propiedad de renovación se cumple también en los momentos en los que el proceso de Poisson tiene saltos, es decir, si empezamos a estudiar al proceso a partir del momento en el que se observa un salto, ese nuevo proceso es también un proceso de Poisson con el mismo parámetro. Desde el punto de vista matemático esta propiedad será de mucha utilidad cuando estudiemos modelos en donde las reclamaciones que reciba una aseguradora sigan un proceso de Poisson.

Otras definiciones equivalentes del proceso de Poisson y algunos otros resultados sobre sus propiedades se pueden encontrar en [31].

6.6. Cadenas de Markov a tiempo continuo

Este tipo de procesos es una generalización del proceso de Poisson mencionado en la sección anterior. Ahora los saltos del proceso no son necesariamente unitarios hacia arriba sino que el proceso puede saltar a cualquier otro estado distinto del estado desde donde salta. Una trayectoria de este tipo de procesos se muestra en la Figura 6.6.
Vamos entonces a definir cadenas de Markov en donde el tiempo es continuo y las variables toman valores enteros. Consideremos un proceso a tiempo continuo $\{X_t : t \geqslant 0\}$ que inicia en un estado i_1 al tiempo cero. El proceso permanece en ese estado un tiempo aleatorio T_{i_1}, y después salta a un nuevo estado i_2 distinto del anterior. El sistema permanece ahora en el estado i_2 un tiempo aleatorio T_{i_2} al cabo del cual brinca a otro estado i_3 distinto del inmediato anterior, y así sucesivamente. Los tiempos aleatorios T son los tiempos en los que el proceso permanece constante en alguno de sus estados, y se llaman tiempos de estancia (*passage times*). Los momentos en donde el proceso tiene saltos son los tiempos $W_n = T_{i_1} + \cdots + T_{i_n}$, para $n \geqslant 1$. El

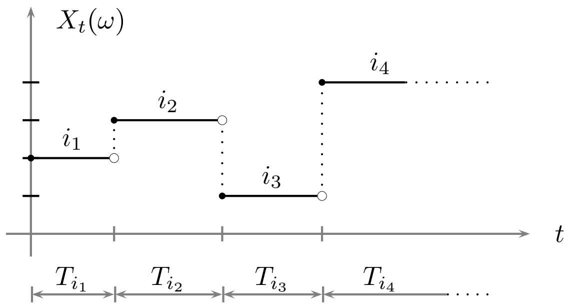


Figura 6.6

proceso puede entonces escribirse en la forma siguiente:

$$X_t = \begin{cases} i_1 & \text{si} \quad 0 \leqslant t < W_1, \\ i_2 & \text{si} \quad W_1 \leqslant t < W_2, \\ i_3 & \text{si} \quad W_2 \leqslant t < W_3, \\ \vdots & \end{cases}$$

A un proceso de estas características se llama proceso de saltos, y resulta ser una buena versión continua de las cadenas de Markov a tiempo discreto, pero es necesario requerir de algunas otras condiciones. Supondremos que

$$\lim_{n\to\infty} W_n = \infty \quad \text{c.s.}$$

y ello garantiza que para todo $t \geqslant 0$, el valor de X_t es finito, c.s. Por otro lado, sin pérdida de generalidad supondremos que el espacio de estados es el conjunto $S = \{0, 1, \ldots\}$ y que el tiempo de estancia asociado el estado i es la variable aleatoria T_i, la cual supondremos positiva con función de distribución $F_i(t)$. Como en el caso de cadenas a tiempo discreto, se denotará por p_{ij} a la probabilidad de que la cadena pase del estado i al estado j al efectuar un salto. Adicionalmente impondremos la condición $p_{ii} = 0$, y con ello se imposibilita que la cadena salte al mismo estado de partida. Las probabilidades de saltos deben entonces satisfacer las siguientes condiciones:

a) $p_{ij} \geqslant 0$.

b) $p_{ii} = 0$.

c) $\sum_j p_{ij} = 1$.

En forma de matriz, las probabilidades de saltos constituyen una matriz estocástica de la siguiente forma:

$$P = \begin{pmatrix} 0 & p_{01} & p_{02} & \cdots \\ p_{10} & 0 & p_{12} & \cdots \\ p_{20} & p_{21} & 0 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

Supondremos además que los tiempos de estancia $T_{i_1}, T_{i_2}, \ldots$ son independientes entre sí, y también son independientes del mecanismo mediante el cual se escoge el estado j al cual la cadena salta después de estar en cualquier otro estado i. Más aún, supondremos que cada variable T_i es finita con probabilidad 1, o bien, es infinita con probabilidad 1. En el primer caso se dice que el estado i es no absorbente, y en el segundo caso que es absorbente. El hecho de que $T_i = \infty$ se interpreta en el sentido de que el proceso deja de saltar y permanece en el estado i el resto del tiempo, es decir, el estado i es absorbente. Por lo tanto, sólo hay dos tipos de estados: absorbentes o no absorbentes. En otras palabras, con probabilidad uno el tiempo de estancia es finito o con probabilidad uno es infinito. Por otra parte, un resultado no trivial establece que un proceso de las características arriba especificadas satisface la propiedad de Markov si y sólo si, los tiempos de estancia en los estados no absorbentes tienen distribución exponencial. Éste es un resultado importante cuya demostración omitiremos y que simplifica drásticamente el modelo general planteado. Como deseamos estudiar procesos de saltos que cumplan la propiedad de Markov, pues tal propiedad ayuda a calcular probabilidades con cierta facilidad, tendremos que suponer entonces que el tiempo de estancia en un estado no absorbente i tiene distribución $\exp(\lambda_i)$, con $\lambda_i > 0$, es decir,

$$F_i(t) = 1 - e^{-\lambda_i t} \quad \text{para } t > 0.$$

Observe que puede considerarse que $\lambda_i = 0$ en el caso cuando $T_i = \infty$. La propiedad de Markov que consideraremos tiene la siguiente forma: para cualesquiera tiempos $0 \leqslant t_1 < t_2 < \cdots < t_n$,

$$p(X_{t_n} = x_{t_n} \,|\, X_{t_1} = x_{t_1}, \ldots, X_{t_{n-1}} = x_{t_{n-1}}) = p(X_{t_n} = x_{t_n} \,|\, X_{t_1} = x_{t_1}).$$

Supondremos nuevamente que estas probabilidades de transición son estacionarias en el tiempo, esto significa que para cada $s \geqslant 0$ y $t \geqslant 0$, se cumple la identidad

$$P(X_{t+s} = j \,|\, X_s = i) = P(X_t = j \,|\, X_0 = i),$$

es decir, no hay dependencia del valor de s. Esta probabilidad se escribe de manera breve mediante la expresión $p_{ij}(t)$, para i y j enteros no negativos, es decir,

$$p_{ij}(t) = P(X_{t+s} = j \,|\, X_s = i) = P(X_t = j \,|\, X_0 = i).$$

En particular para $t = 0$ se define $p_{ij}(0)$ como la función delta de Kronecker, es decir,

$$p_{ij}(0) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Haciendo variar los índices i y j en el espacio de estados se obtiene la matriz de probabilidades de transición al tiempo t, que denotaremos por P_t y en ocasiones se escribe también como $P(t)$:

$$P_t = (p_{ij}(t)) = \begin{pmatrix} p_{00}(t) & p_{01}(t) & \cdots \\ p_{10}(t) & p_{11}(t) & \cdots \\ \vdots & \vdots & \end{pmatrix}.$$

Puede demostrarse que cuando el espacio de estados es finito, esta matriz es siempre estocástica, es decir, los elementos de cada renglón suman 1.

Definición 6.9 *A un proceso de saltos con las características y postulados arriba señalados se le llama cadena de Markov a tiempo continuo.*

Observe que en un proceso de Markov a tiempo continuo las probabilidades de saltos p_{ij} y las probabilidades de transición $p_{ij}(t)$ representan aspectos distintos del proceso. Las primeras son probabilidades de cambio al estado j cuando el proceso se encuentra en el estado i y el proceso tiene un salto, mientras que las segundas son probabilidades de encontrar al proceso en el estado j, partiendo de i, al término de un intervalo de tiempo de longitud t. Observe además que un proceso de Markov a tiempo continuo queda completamente especificado por los siguientes tres elementos: una distribución de probabilidad inicial en el espacio de estados, el conjunto de los parámetros no negativos λ_i, y las probabilidades de saltos p_{ij}.

Ejemplo 6.2 (Proceso de Poisson) *El proceso de Poisson es una cadena de Markov a tiempo continuo que empieza en cero, es decir, la distribución de probabilidad inicial tiene el valor* 1 *en el estado cero. Los tiempos de estancia son exponenciales de parámetro* λ *y las probabilidades de saltos de un estado a otro son*

$$p_{ij} = \begin{cases} 1 & si \;\; j = i+1, \\ 0 & si \;\; j \neq i+1. \end{cases}$$

Las probabilidades de transición son

$$p_{ij}(t) = \begin{cases} e^{-\lambda t}\,\frac{(\lambda t)^{j-i}}{(j-i)!} & si\ j \geqslant i, \\ 0 & otro\ caso. \end{cases}$$

Ejemplo 6.3 (Cadena de dos estados) *Considere el proceso* $\{X_t : t \geqslant 0\}$ *con espacio de estados* $\{0,1\}$ *y definido por la siguiente dinámica: cuando el proceso entra al estado* 0 *permanece en él un tiempo* exp(λ) *y luego va al estado* 1*, entonces permanece en el estado* 1 *un tiempo* exp(μ) *y después regresa a* 0*, y así sucesivamente. Se postula además que los tiempos de estancia en cada estado son variables aleatorias independientes. Para este proceso pueden encontrarse explícitamente las probabilidades de transición* $p_{ij}(t)$*. Puede demostrarse que para cualquier* $t \geqslant 0$,

$$p_{00}(t) \;=\; \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}\, e^{-(\lambda+\mu)t}.$$

En consecuencia, por complemento o simetría,

$$\begin{aligned} p_{01}(t) &= \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}\, e^{-(\lambda+\mu)t}, \\ p_{11}(t) &= \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\, e^{-(\lambda+\mu)t}, \\ p_{10}(t) &= \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu}\, e^{-(\lambda+\mu)t}. \end{aligned}$$

En notación matricial,

$$\begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{pmatrix} = \frac{1}{\lambda+\mu}\begin{pmatrix} \mu & \lambda \\ \mu & \lambda \end{pmatrix} + \frac{1}{\lambda+\mu}\begin{pmatrix} \lambda & -\lambda \\ -\mu & \mu \end{pmatrix} e^{-(\lambda+\mu)t}.$$

Probabilidades de transición

Hemos mencionado antes que para una cadena de Markov a tiempo continuo las probabilidades de transición son los números $p_{ij}(t) = P(X_t = j \mid X_0 = i)$. El problema que puede plantearse es el de encontrar una expresión para las probabilidades de transición $p_{ij}(t)$ para cada par de estados i y j, y para cada tiempo $t \geqslant 0$. Este es un problema demasiado general y sólo en algunos pocos casos es posible encontrar explícitamente tales probabilidades. Los anteriores dos ejemplos son casos muy sencillos en donde es posible encontrar las probabilidades $p_{ij}(t)$.

El generador infinitesimal

A partir de los parámetros λ_i y p_{ij} de una cadena de Markov a tiempo continuo se pueden definir las cantidades g_{ij} de la siguiente forma:

$$g_{ij} = \begin{cases} -\lambda_i & \text{si } i = j, \\ \lambda_i\, p_{ij} & \text{si } i \neq j. \end{cases}$$

A estos números se les conoce con el nombre de parámetros infinitesimales del proceso. Haciendo variar los índices i y j, estos nuevos parámetros conforman una matriz G llamada el generador infinitesimal del proceso de Markov, es decir,

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 p_{01} & \lambda_0 p_{02} & \cdots \\ \lambda_1 p_{10} & -\lambda_1 & \lambda_1 p_{12} & \cdots \\ \lambda_2 p_{20} & \lambda_2 p_{21} & -\lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.$$

Puede demostrarse que el generador infinitesimal caracteriza de manera única a la cadena de Markov. Así, a esta misma matriz se le llama a veces cadena de Markov a tiempo continuo, y es el concepto equivalente a la matriz de probabilidades de transición en un paso para cadenas a tiempo discreto. Se trata de una matriz con las siguientes propiedades:

a) $g_{ij} \geqslant 0, \quad \text{si } i \neq j.$

b) $g_{ii} \leqslant 0.$

c) $\sum_j g_{ij} = 0.$

Ejemplo 6.4 *El generador infinitesimal para el proceso de Poisson de parámetro λ es*

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & \vdots & & \end{pmatrix}. \tag{6.2}$$

Ejemplo 6.5 *El generador infinitesimal para la cadena de Markov de dos estados del Ejemplo 6.3 es*

$$G = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}. \tag{6.3}$$

Ecuaciones de Kolmogorov

Puede demostrarse que las probabilidades de transición $p_{ij}(t)$ satisfacen el sistema de ecuaciones diferenciales dado por

$$p'_{ij}(t) = \sum_k g_{ik}\, p_{kj}(t), \quad t \geqslant 0.$$

Observe que se tiene una ecuación diferencial para cada par ordenado de estados (i,j). En términos de matrices la igualdad anterior se escribe

$$P'(t) = G\,P(t).$$

A este sistema de ecuaciones diferenciales se le conoce como las ecuaciones retrospectivas de Kolmogorov.

Ejemplo 6.6 (Proceso de Poisson) *El sistema de ecuaciones retrospectivas de Kolmogorov para el proceso de Poisson de parámetro λ está dado por*

$$\begin{aligned} p'_{ii}(t) &= -\lambda p_{ii}(t) \\ p'_{ij}(t) &= -\lambda p_{ij}(t) + \lambda p_{i+1,j}(t) \quad \textit{para } i < j. \end{aligned}$$

Y sabemos que la solución es $p_{ij}(t) = e^{-\lambda t}\,(\lambda t)^{j-i}/(j-i)!$ *para* $i \leqslant j$.

Ejemplo 6.7 (Cadena de dos estados) *El sistema de ecuaciones retrospectivas de Kolmogorov para la cadena de Markov de dos estados definida en el Ejemplo 6.3 está dado por*

$$\begin{aligned} p'_{00}(t) &= -\lambda\, p_{00}(t) + \lambda\, p_{10}(t), \\ p'_{01}(t) &= -\lambda\, p_{01}(t) + \lambda\, p_{11}(t), \\ p'_{10}(t) &= -\mu\, p_{10}(t) + \mu\, p_{00}(t), \\ p'_{11}(t) &= -\mu\, p_{11}(t) + \mu\, p_{01}(t). \end{aligned}$$

Pueden resolverse estas ecuaciones y encontrar las expresiones que fueron enunciadas anteriormente.

Al sistema de ecuaciones diferenciales dado por la igualdad $P'(t) = P(t)\,G$ se le llama sistema de ecuaciones prospectivas de Kolmogorov. La diferencia entre este sistema y el sistema retrospectivo mencionado antes es que el orden de los factores en el lado derecho es distinto. Más explícitamente, el sistema prospectivo es el siguiente

$$p'_{ij}(t) = \sum_k p_{ik}(t)\, g_{kj}.$$

En algunos casos los dos sistemas de ecuaciones son equivalentes y su solución produce las mismas probabilidades de transición $p_{ij}(t)$. En general, el sistema retrospectivo es el que siempre se satisface.
Los procesos de nacimiento y muerte son un tipo particular importante de cadena de Markov a tiempo continuo. Este tema y una exposición más detallada sobre este tipo de proceso estocástico puede ser encontrada, por ejemplo, en el texto de Hoel, Port y Stone [19], en Basu [3], en Karlin y Taylor [20] o en Rolski *et al.* [32]. Concluimos esta sección definiendo un tipo

de distribución de probabilidad que surge a partir de cadenas de Markov a tiempo continuo con espacio de estados finito y que permiten encontrar una fórmula para la probabilidad de ruina en un modelo de riesgo cuando las reclamaciones tienen este tipo de distribuciones. Estudiaremos esta fórmula en la última parte del texto.

Distribuciones continuas tipo fase

Definiremos ahora las distribuciones tipo fase, las cuales también se pueden definir a partir de ciertas cadenas de Markov a tiempo continuo. Sea $\{X_t : t \geqslant 0\}$ una cadena de Markov a tiempo continuo con espacio de estados $\{0, 1, 2, \ldots, k\}$ en donde el estado 0 es absorbente y los estados $1, 2, \ldots, k$ son no absorbentes. Suponga que el generador infinitesimal de esta cadena es de la forma:

$$G = \begin{pmatrix} 0 & \mathbf{0} \\ b^T & B \end{pmatrix},$$

en donde nuevamente $\mathbf{0}$ es el vector renglón $(0, 0, \ldots, 0)$ de dimensión k, b es el vector rengón $(b_1, b_2, \ldots, b_k)$ con no todas sus entradas cero y B es una matriz cuadrada de $k \times k$ tal que la matriz completa G es un generador infinitesimal. A la matriz B se le llama matriz de subintensidades. En particular,

$$b^T + Be^T = \mathbf{0}^T, \tag{6.4}$$

en donde e es el vector renglón $(1, 1, \ldots, 1)$ de dimensión k. Supondremos que la cadena inicia en cualquiera de sus estados, incluyendo el estado absorbente, a través de una distribución de probabilidad inicial

$$\pi' = (\pi_0, \pi_1, \ldots, \pi_k).$$

Y nuevamente denotaremos por π al subvector cuyas entradas corresponden a la distribución de probabilidad inicial para estados no absorbentes.

$$\pi = (\pi_1, \pi_2, \ldots, \pi_k).$$

Cuando $\pi_0 = 0$ el subvector π es efectivamente una distribución de probabilidad. Si la cadena inicia en alguno de sus estados no absorbentes, entonces visitará estados absorbentes un tiempo aleatorio exponencial en cada de uno de ellos y después posiblemente quedará atrapado en el estado absorbente 0. La idea es la misma que fue expresada en el caso de cadenas de Markov a tiempo discreto. La diferencia radica en que ahora el tiempo es continuo. A la distribución de probabilidad del tiempo de espera hasta la absorción se le llama nuevamente distribución tipo fase.

Definición 6.10 *A la distribución de probabilidad de la variable aleatoria* τ *definida por*

$$\tau = \text{ínf}\,\{t \geqslant 0 \,|\, X_t = 0\}$$

se le llama distribución continua tipo fase con subdistribución inicial π *y matriz de subintensidades* B*. Esto se escribe*

$$\tau \sim PH(\pi, B).$$

Hemos usado aquí la misma notación para las distribuciones tipo fase discretas y continuas. El contexto o especificación del modelo determinará sin ambigüedad si es de un tipo o del otro. Por otro lado, en algunos textos se considera que la totalidad de estados es $\{1, 2, \ldots, k+1\}$ en donde el estado $k+1$ es absorbente y el resto de los estados es no absorbente. Ambos modelos son equivalentes.

Ejemplo 6.8 (Distribución Erlang) *Sea* $\lambda > 0$ *fijo. Considere la cadena de Markov a tiempo continuo* $\{X_t : t \geqslant 0\}$ *con* $k+1$ *estados: un estado absorbente* 0 *y* k *estados no absorbentes* $1, 2, \ldots, k$*. Suponga que el generador infinitesimal es de la forma:*

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -\lambda & \lambda & 0 & \cdots & 0 \\ 0 & 0 & -\lambda & \lambda & \cdots & 0 \\ 0 & 0 & 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda & 0 & 0 & 0 & \cdots & -\lambda \end{pmatrix}.$$

Suponga además que $X_0 = 1$*. Así, la cadena inicia en el estado* 1 *y permanece en dicho estado un tiempo exponencial de parámetro* λ*, al término* *de este tiempo la cadena salta al estado* 1 *y permanece allí otro tiempo exponencial de parámetro* λ *e independiente del primero, y así sucesivamente hasta llegar al último estado transitorio* k*. Después de estar en este último estado un tiempo exponencial, la cadena finalmente salta al estado absorbente* 0 *y permanece allí el resto del tiempo. Es claro que el tiempo de absorción* τ *tiene distribución Erlang*(k,λ) *pues es la suma de* k *variables aleatorias independientes con distribución exponencial de parámetro* λ*. Por lo tanto la función de densidad de* τ *es*

$$f(t) = \frac{(\lambda t)^{k-1}}{(k-1)!}\,\lambda t\, e^{-\lambda t}, \quad t > 0.$$

En particular, cuando únicamente hay un estado transitorio, es decir, $k = 1$*, se obtiene la distribución exponencial.*

Del mismo modo que lo hicimos en el caso discreto, mencionaremos ahora algunas fórmulas que se conocen para las distribuciones tipo fase continuas. Las demostraciones de estas propiedades y algunas otras pueden encontrarse en Rolski *et al.* [32]. Recordemos que la función exponencial de una matriz cuadrada A de dimensión finita se define mediante la serie de potencias de Taylor:

$$e^A := \sum_{n=0}^{\infty} \frac{1}{n!} A^n.$$

a) Para cada $t \geqslant 0$,

$$P(\tau > t) = \pi_0 + \pi\,[\,\exp tB\,]\; e^T. \tag{6.5}$$

b) Si $\pi_0 = 0$ y la matriz B es no singular, entonces τ tiene función de densidad

$$f(t) = \pi\,[\,\exp tB\,]\; b^T, \quad t > 0.$$

c) Si la matriz B es no singular, entonces para cada $n \geqslant 1$,

$$E(\tau^n) = (-1)^n\, n!\, \pi\, (B^{-1})^n\, e^T.$$

Ejemplo 6.9 (Distribución Erlang) *Considere nuevamente la situación del Ejemplo 6.8 en donde el tiempo de absorción tiene distribución Erlang*(k,λ)*.*

El estado inicial es $X_0 = 1$ y por lo tanto $\pi_0 = 0$. Entonces la matriz de subintensidades B es no singular y tiene la forma

$$B = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & 0 \\ 0 & -\lambda & \lambda & \cdots & 0 \\ 0 & 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda \end{pmatrix}_{k\times k}.$$

Las fórmulas arriba mencionadas se reducen a las expresiones conocidas para la distribución Erlang(k,λ), es decir,

a) $P(\tau > t) = \sum_{n=1}^{k-1} e^{-\lambda t}\, \frac{(\lambda t)^n}{n!}, \quad t > 0.$

b) $f(t) = \frac{(\lambda t)^{k-1}}{(k-1)!}\, \lambda t\, e^{-\lambda t}, \quad t > 0.$

c) $E(\tau^n) = \frac{(n+k-1)!}{\lambda^n\,(k-1)!}, \quad n \geqslant 1.$

Estas expresiones se reducen aún más en el caso exponencial cuando $k = 1$.

En la última parte de este texto usaremos las distribuciones tipo fase para modelar los montos de las reclamaciones y aprovecharemos sus propiedades computacionales para encontrar una fórmula para la probabilidad de ruina en un modelo de riesgo a tiempo continuo. En particular haremos uso de la fórmula (6.5). Se puede encontrar mayor información sobre las distribuciones tipo fase, incluyendo las demostraciones de los resultados arriba enunciados, en el texto de Rolski *et al.* [32]. En el artículo de Bladt [5] puede también encontrarse una excelente exposición sobre el uso y aplicación de las distribuciones tipo fase en la teoría del riesgo.

6.7. Martingalas

Las martingalas son un tipo de proceso estocástico que aparece con frecuencia tanto en la teoría general de procesos como en las aplicaciones. Algunos resultados desarrollados para martingalas nos serán de utilidad para presentar una forma de resolver ciertos problemas de la teoría de la ruina. La siguiente es la definición de martingala, la cual puede también enunciarse para tiempos discretos.

Definición 6.11 *Un proceso a tiempo continuo* $\{X_t : t \geqslant 0\}$ *que es adaptado a una filtración* $(\mathscr{F}_t)_{t\geqslant 0}$ *y que es integrable, es decir, cada variable que conforma el proceso tiene esperanza finita, se llama una martingala si para* $0 \leqslant s \leqslant t$ *se cumple*

$$E(X_t \,|\, \mathscr{F}_s) = X_s \quad c.s. \tag{6.6}$$

La identidad (6.6) establece un tipo de dependencia probabilística particular entre las variables aleatorias que conforman una martingala. Este es entonces otro ejemplo de proceso estocástico. En el Apéndice se encuentra una revisión breve de la definición y algunas propiedades de la esperanza condicional como la que aparece en (6.6). Las martingalas son procesos que están relacionados con los juegos justos. Por ejemplo, si la variable X_t representa la fortuna de un jugador al tiempo t, y quien supondremos apuesta de manera continua, entonces la igualdad anterior se interpreta del siguiente modo: en promedio la fortuna del jugador al tiempo t dada toda la historia del juego hasta el tiempo s anterior a t es la fortuna del jugador al tiempo s, es decir, el juego es justo pues el jugador en promedio no pierde ni gana. Cuando en lugar de (6.6) se cumple $E(X_t \,|\, \mathscr{F}_s) \leqslant X_s$ se dice que el proceso es una supermartingala, se trata entonces de un juego desfavorable al jugador pues en promedio su fortuna disminuye. En caso de la desigualdad contraria el proceso es una submartingala, juego favorable al jugador. Cuando se toma esperanza en la ecuación (6.6) se obtiene

$$E(X_t) = E(X_s),$$

para $0 \leqslant s \leqslant t$. Esto quiere decir que todas las variables aleatorias que conforman una martingala tienen la misma esperanza. En particular, si la variable inicial X_0 es cero o su esperanza es cero, entonces $E(X_t) = 0$ para cualquier $t \geqslant 0$. Algunos ejemplos sencillos de martingalas aparecen en la sección de ejercicios. Finalmente enunciaremos un resultado de la teoría general de martingalas que será usado en la última parte del curso. Recuerde la notación $x \wedge y = \min\{x, y\}$.

Teorema 6.1 (Teorema de paro de martingalas). *Sea* $\{X_t : t \geqslant 0\}$ *una martingala y sea* τ *un tiempo de paro, ambos respecto de una filtración* $(\mathscr{F}_t)_{t\geqslant 0}$. *Entonces* $\{X_{t\wedge\tau} : t \geqslant 0\}$ *es también una martingala, es decir, para cualesquiera* $0 \leqslant s \leqslant t$,

$$E(X_{t\wedge\tau} \,|\, \mathscr{F}_s) = X_{s\wedge\tau} \quad c.s.$$

La demostración de este resultado puede encontrarse en [30] o [39]. En el siguiente capítulo usaremos la teoría de martingalas para estimar la probabilidad de ruina en el modelo clásico de Cramér-Lundberg.

Comentarios y referencias

Como preparación para los modelos dinámicos de riesgo que estudiaremos en los siguientes capítulos, hemos presentado aquí algunos conceptos elementales de los procesos estocásticos, así como algunos ejemplos particulares de este tipo de modelos matemáticos. Para mayor información sobre estos temas el lector puede consultar cualquiera de los textos sobre procesos estocásticos que aparecen en la bibliografía, por ejemplo, Basu [3], Hoel *et al.* [19], Karlin y Taylor [20], Resnick [29] y Stirzaker [37].

6.8. Ejercicios

Caminatas aleatorias

136. Para la caminata aleatoria simple $\{X_n : n \geqslant 0\}$ sobre $\mathbb{Z}$ demuestre que para $n \geqslant 1$:

   a) $E(X_n) = n(p - q)$.

   b) $\mathrm{Var}(X_n) = 4npq$.

   c) $E(e^{tX_n}) = (pe^t + qe^{-t})^n$.

Cadenas de Markov

137. Considere una cadena de Markov $\{X_n : n \geqslant 0\}$ con espacio de estados $\{0, 1, 2\}$ y con matriz de probabilidades de transición

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 2/3 & 0 & 1/3 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}.$$

Suponga $X_0 = 0$. Calcule:

a) $P(X_2 = 1)$.

b) $P(X_5 = 1 \,|\, X_3 = 2)$.

Ahora suponga que X_0 tiene distribución $(1/5, 1/2, 3/10)$. Calcule:

c) $P(X_2 = 1 \,|\, X_1 = 0)$.

d) $P(X_3 = 0)$.

138. Para la cadena de Markov $\{X_n : n \geqslant 0\}$ con espacio de estados $\{0, 1, 2, 3, 4\}$ con distribución inicial $\pi = (1/5, 1/5, 1/5, 1/5, 1/5)$ y matriz de probabilidades de transición

$$P = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 & 0 \\ 2/3 & 0 & 0 & 0 & 1/3 \\ 0 & 1/4 & 1/2 & 1/4 & 0 \\ 2/5 & 0 & 2/5 & 0 & 1/5 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix},$$

calcule:

a) $P(X_0 = 0, X_1 = 1, X_2 = 1, X_3 = 4, X_4 = 1)$.

b) $P(X_{101} = 3 \,|\, X_{100} = 3)$.

c) $P(X_0 = X_1)$.

d) $P(|X_1 - X_0| \leqslant 1)$.

e) $P(X_1 = 0, X_3 = 0)$.

139. Suponga que $\{X_n : n \geqslant 0\}$ es una cadena de Markov con espacio de estados $\{0, 1\}$ y con matriz de probabilidades de transición

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/10 & 9/10 \end{pmatrix}.$$

El estado 0 representa no contar con un seguro de vida y el estado 1 corresponde a tener algún seguro de vida. La dinámica de contratar o no contratar un seguro de vida para cada persona en periodos sucesivos de tiempo está dada por la cadena de Markov. Si inicialmente la mitad de la población está asegurada, calcule el porcentaje de la población que estará asegurada en los periodos 1, 2 y 3 por separado.

Proceso de Poisson

140. Suponga que las llegadas de reclamaciones a una compañía aseguradora siguen un proceso de Poisson de intensidad $\lambda = 5$, en donde la unidad de tiempo es un día, es decir, en promedio llegan 5 reclamaciones por día. Calcule la probabilidad de que:

    a) No se reciba ninguna reclamación en un día cualquiera.
    b) Se reciban más de 10 reclamaciones en un día cualquiera.
    c) Se reciba una sola reclamación en los siguientes tres días.

141. Los clientes ingresan a un establecimiento de acuerdo a un proceso de Poisson a razón de 10 clientes por hora en promedio. Calcule la probabilidad de que:

    a) En una hora cualquiera no llegue ningún cliente.
    b) No llegue ningún cliente en una jornada de 12 horas.
    c) Se presente exactamente un cliente en todas y cada una de las 12 horas en las que está abierto el establecimiento en un día.

142. *Superposición.* Demuestre que la suma de dos procesos de Poisson independientes es nuevamente un proceso de Poisson con parámetro la suma de los parámetros. Nota: la operación suma debe entenderse en el sentido de superponer los dos procesos puntuales.

143. *Thinning.* Sea $\{N_t : t \geqslant 0\}$ un proceso de Poisson de parámetro $\lambda > 0$. Considere una sucesión de variables aleatorias $Y_1, Y_2, \ldots$ independientes, con distribución común $\mathrm{Ber}(p)$ e independientes del proceso de Poisson. Defina el proceso $\{X_t : t \geqslant 0\}$ como aparece abajo, definiendo además a X_t como cero cuando N_t es cero.

$$X_t = \sum_{k=1}^{N_t} Y_k,$$

Así, el proceso $\{X_t : t \geqslant 0\}$ representa un subconteo del proceso de Poisson inicial.

a) Demuestre que $\{X_t : t \geqslant 0\}$ es un proceso de Poisson de parámetro λp.

b) Sea $t \geqslant 0$ fijo. Demuestre que las variables aleatorias X_t y $N_t - X_t$ son independientes. En general se cumple que los procesos estocásticos $\{X_t : t \geqslant 0\}$ y $\{N_t - X_t : t \geqslant 0\}$ son independientes.

144. Suponga que los accidentes automovilísticos en una cierta ciudad ocurren de acuerdo a un proceso de Poisson con un promedio de cinco accidentes por semana. Suponga además que la probabilidad de que en cada accidente haya personas lastimadas que requieran atención médica es 0.2 . Calcule la probabilidad de que:

a) En un día no se presente ningún accidente.

b) En una semana no se presente ningún accidente.

c) En un mes no se presente ningún accidente con personas lastimadas.

145. Sea $\{N_t : t \geqslant 0\}$ un proceso de Poisson con parámetro λ, y sea $a > 0$ una constante. Demuestre que $\{N_{at} : t \geqslant 0\}$ es un proceso de Poisson con parámetro λa.

146. Considere dos procesos de Poisson independientes de parámetros λ_1 y λ_2. Sea X el número de eventos que ocurren en el primer proceso

entre dos eventos sucesivos del segundo proceso. Demuestre que X tiene la siguiente distribución geométrica.

$$P(X = n) = \frac{\lambda_2}{\lambda_1 + \lambda_2} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^n, \quad n \geqslant 0.$$

Cadenas de Markov a tiempo continuo

147. Escriba el sistema retrospectivo de ecuaciones diferenciales de Kolmogorov para las probabilidades $p_{0n}(t) = p_n(t)$ del proceso de Poisson y compruebe que su solución es $p_n(t) = e^{-\lambda t}\,(\lambda t)^n/n!$

Martingalas

148. *Martingala del juego de apuestas.* Sea $\xi_1, \xi_2, \ldots$ una sucesión de variables aleatorias independientes idénticamente distribuidas y con esperanza finita. Para cada entero $n \geqslant 1$ defina

$$X_n = \xi_1 + \cdots + \xi_n.$$

Demuestre que el proceso a tiempo discreto $\{X_n : n \geqslant 1\}$ es una martingala si y sólo si $E(\xi) = 0$.

149. *Proceso de Poisson centrado.* Sea $\{N_t : t \geqslant 0\}$ un proceso de Poisson de parámetro o intensidad λ, junto con su filtración natural. Demuestre que el proceso centrado $\{N_t - \lambda t : t \geqslant 0\}$ es una martingala.

150. *Procesos con incrementos independientes.* Sea $\{X_t : t \geqslant 0\}$ un proceso estocástico a tiempo continuo tal que cada una de sus variables aleatorias tiene esperanza finita. Suponga que el proceso tiene incrementos independientes. Demuestre que el proceso centrado $\{X_t - E(X_t) : t \geqslant 0\}$ es una martingala.

151. *Martingala de de Moivre.* Sea $\xi_1, \xi_2, \ldots$ una sucesión de variables aleatorias independientes cada una de ellas con la misma distribución dada por

$$\begin{array}{rrcl} & P(\xi = +1) & = & p \\ \text{y} & P(\xi = -1) & = & q \;=\; 1 - p. \end{array}$$

Para cada entero $n \geqslant 1$ defina la variable $X_n = \xi_1 + \cdots + \xi_n$. Demuestre que el proceso estocástico a tiempo discreto $\{Y_n : n \geqslant 1\}$ dado por $Y_n = (q/p)^{X_n}$ es una martingala.

152. Demuestre que el proceso $\{X_t : t \geqslant 0\}$ es una submartingala si y sólo si $\{-X_t : t \geqslant 0\}$ es una supermartingala.

153. Demuestre que un proceso es una martingala si y sólo si es al mismo tiempo una submartingala y una supermartingala.

Capítulo 7

Teoría de la ruina: tiempo discreto

Este capítulo contiene una introducción elemental a uno de los problemas centrales de la teoría del riesgo: el problema de la ruina. Estudiaremos este problema en una versión discreta. Definiremos primero un proceso de riesgo a tiempo discreto y encontraremos una fórmula general recursiva para la probabilidad de ruina con horizonte infinito. Presentaremos también una fórmula para la probabilidad de ruina con horizonte finito, así como el concepto de coeficiente de ajuste para este modelo y su aplicación en la desigualdad de Lundberg. Estudiaremos también el problema de calcular la distribución de probabilidad de la severidad de ruina cuando ésta se presenta. Todos estos resultados serán extendidos en el siguiente capítulo, en donde se estudiará el modelo clásico de riesgo a tiempo continuo.

7.1. Un proceso de riesgo a tiempo discreto

Definiremos a continuación un proceso estocástico a tiempo discreto que modela de manera simplificada la evolución a lo largo del tiempo del capital de una compañía aseguradora respecto de una cartera de asegurados. Suponga que $u \in \{0, 1, 2 \ldots\}$ es el capital inicial de la aseguradora y que en cada unidad de tiempo la compañía aseguradora recibe una unidad monetaria por concepto de primas. Si $Y_1, Y_2, \ldots$ representan los montos de las reclamaciones en los periodos sucesivos, entonces el capital de la compañía aseguradora al tiempo $n \geqslant 1$ es la variable aleatoria C_n definida a continuación.

Definición 7.1 *El proceso de riesgo a tiempo discreto* $\{C_n : n \geqslant 0\}$ *está dado por*

$$C_n = u + n - \sum_{j=1}^{n} Y_j, \tag{7.1}$$

en donde $u \geqslant 0$ *es un entero y* $Y_1, Y_2, \ldots$ *son variables aleatorias independientes e idénticamente distribuidas con valores en el conjunto* $\{0, 1, \ldots\}$ *y tales que* $E(Y) < 1$.

Dada la hipótesis de independencia e idéntica distribución de las variables aleatorias $Y_1, Y_2, \ldots$, puede comprobarse que el proceso $\{C_n : n \geqslant 0\}$ tiene incrementos independientes y estacionarios. Observamos además que se puede escribir

$$C_n = u + \sum_{j=1}^{n} (1 - Y_j),$$

y por lo tanto $\{C_n : n \geqslant 0\}$ es una caminata aleatoria sobre $\mathbb{Z}$ en donde los saltos son de magnitud $1 - Y_j$ con valores en el conjunto $\{\ldots, -2, -1, 0, 1\}$. Por otro lado, hemos supuesto que las variables aleatorias $Y_1, Y_2, \ldots$ son tales que $E(Y) < 1$, a esta condición la llamaremos condición de ganancia neta y establece que por cada unidad de tiempo, la cantidad de dinero recibida por concepto de primas, en este caso una unidad monetaria, es mayor al valor promedio de las reclamaciones.

Condición de ganancia neta $\quad E(Y) < 1$

Debe observarse que no es realista la hipótesis de que las primas recibidas en cada unidad de tiempo sean unitarias. Se puede considerar un modelo más aplicable al admitir que el ingreso por primas en cada periodo es un entero cualquiera $c \geqslant 1$. En ese caso, mediante un cambio adecuado en la forma en la que se mide el tiempo, esa constante c puede considerarse

como una unidad monetaria. Sin embargo, tal cambio en el tiempo no se puede aplicar de manera tan inmediata pues afecta también los montos de las reclamaciones. Dado que nuestro objetivo en este capítulo es presentar de manera simple algunos problemas de la teoría de la ruina en un modelo discreto, mantendremos la hipótesis de primas unitarias pues las fórmulas y resultados adquieren expresiones más sencillas que en el caso general.
En la presente sección la variable Y representará a cualquiera de las variables Y_j que aparecen en la expresión (7.1), $F(y)$ será la función de distribución de Y y la correspondiente función de probabilidad se denotará por $f(y)$. En las fórmulas que encontraremos más adelante aparecerá con frecuencia el término $P(Y > y) = 1 - F(y)$ y para hacer las expresiones más cortas dicha probabilidad será denotada por $\overline{F}(y)$. En ocasiones a esta probabilidad se le llama función de supervivencia.

Notación	$\overline{F}(y) := 1 - F(y)$

Dado que la variable Y es discreta con valores en el conjunto $\{0, 1, \ldots\}$, su esperanza puede entonces escribirse de la siguiente forma:

$$E(Y) = \sum_{y=0}^{\infty} \overline{F}(y).$$

Por otro lado, dadas las características que hemos solicitado para la definición del proceso $\{C_n : n \geqslant 0\}$, éste resulta ser una cadena de Markov con espacio de estados o valores dado por el conjunto discreto $\mathbb{Z}$. Y hemos considerado valores enteros negativos, pues alguna reclamación puede ser demasiado grande y llevar al proceso a estados críticos para la aseguradora. Justamente esta situación es nuestro objeto de interés y es el contenido de la siguiente definición.

Definición 7.2 *Se dice que la compañía aseguradora se encuentra en ruina al tiempo* $n \geqslant 1$ *si*

$$C_n \leqslant 0,$$

y se define el tiempo de ruina τ *como el primer momento en que la ruina se presenta, es decir,*

$$\tau = \text{mín}\,\{n \geqslant 1 : C_n \leqslant 0\}. \tag{7.2}$$

En la expresión (7.2) se debe entender que cuando el conjunto indicado es vacío se define $\tau = \infty$, y equivale a la situación cuando la ruina nunca se presenta. El problema de la ruina consiste en encontrar la probabilidad de que la ruina ocurra en algún conjunto de tiempos específico. Por ejemplo, la probabilidad ruina con horizonte infinito es $P(\tau < \infty)$ y se denota usualmente por $\psi(u)$. Con esta notación se hace énfasis en que tal probabilidad depende, entre otros parámetros del modelo, particularmente del capital inicial u. Tenemos entonces que

$$\begin{aligned} \psi(u) &= P(\tau < \infty \,|\, C_0 = u) \\ &= P(\tau \in \{1, 2, \ldots\} \,|\, C_0 = u). \end{aligned}$$

Observe que técnicamente no hay ruina al tiempo cero, aun cuando se considere al capital inicial u igual a cero, pues de acuerdo con la Definición 7.2, la ruina sólo puede ocurrir en los tiempos $n \geqslant 1$. Intuitivamente es claro que la función $u \mapsto \psi(u)$ es decreciente, es decir, a mayor capital inicial menor probabilidad de ruina. Esta propiedad puede escribirse de la siguiente forma: para cualquier $u \geqslant 0$,

$$\psi(u + 1) \leqslant \psi(u).$$

En la Figura 7.1 se muestra una posible trayectoria del proceso $\{C_n : n \geqslant 0\}$, en donde para fines de visualización se han unido los valores del proceso mediante líneas punteadas indicando además los incrementos unitarios por concepto de primas en cada intervalo. Se muestra además un posible momento τ en donde se presenta la ruina. La propiedad de Markov del proceso de riesgo nos permitirá encontrar una fórmula recursiva para la probabilidad de ruina en este modelo discreto. En las expresiones que escribiremos

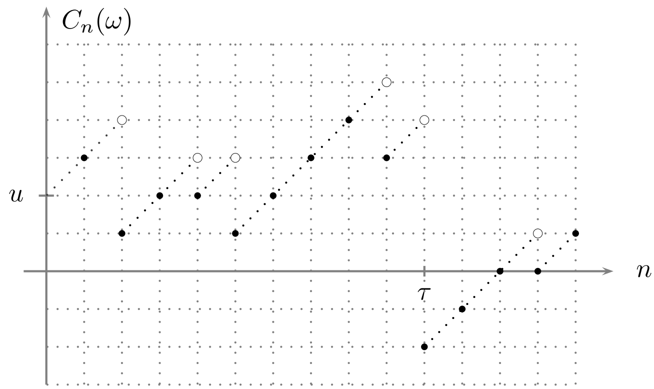


Figura 7.1

a continuación las sumas $\sum_a^b$ se definen como cero cuando los índices a y b son tales que $a > b$. Usaremos además el hecho intuitivamente claro de que la probabilidad de ruina es cero cuando el capital inicial es infinito. Más adelante daremos la demostración de esta propiedad y la comprobaremos de manera sencilla en el caso cuando podamos aplicar la desigualdad de Lundberg. Así, tenemos que

$$\lim_{u \to \infty} \psi(u) = 0.$$

7.2. Probabilidad de ruina con horizonte infinito

El siguiente resultado provee de un mecanismo recursivo para calcular la probabilidad de ruina con horizonte infinito en el proceso de riesgo a tiempo discreto.

Proposición 7.1 *Para el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$ con valor inicial $u \geqslant 0$,*

1. $\psi(u) = \psi(0) + \sum_{y=0}^{u-1} \psi(u-y)\,\overline{F}(y) - \sum_{y=0}^{u-1} \overline{F}(y), \quad u \geqslant 1.$
2. $\psi(0) = E(Y).$

Demostración. Para cualquier capital inicial $w \geqslant 0$ y condicionando sobre el valor de Y_1 tenemos los siguientes cálculos, los cuales explicaremos en el siguiente párrafo.

$$\begin{aligned}
\psi(w) &= \sum_{y=0}^{\infty} P(\tau < \infty \,|\, Y_1 = y)\, P(Y_1 = y) \\
&= \sum_{y=0}^{w} P(\tau < \infty \,|\, Y_1 = y)\, f(y) + \sum_{y=w+1}^{\infty} P(\tau < \infty \,|\, Y_1 = y)\, f(y) \\
&= \sum_{y=0}^{w} \psi(w+1-y)\, f(y) + \sum_{y=w+1}^{\infty} f(y) \\
&= \sum_{y=1}^{w+1} \psi(y)\, f(w+1-y) + \overline{F}(w). \qquad (7.3)
\end{aligned}$$

Observe que en la segunda igualdad se han separado dos casos: la primera suma se refiere al caso cuando el monto reclamado Y_1 no produce ruina y la segunda suma cuando se presenta la ruina. En el segundo caso la probabilidad condicional indicada es 1. En el primer caso la probabilidad condicional se reduce a la probabilidad $\psi(w+1-y)$, pues siendo la primera reclamación de magnitud y y no habiendo ruina en esta primera reclamación, la probabilidad condicional original se reduce a la probabilidad de que no se presente la ruina desde el momento de la primera reclamación en adelante pero ahora con capital inicial $w+1-y$. Por la propiedad de incrementos independientes, esta probabilidad es $\psi(w+1-y)$. Despejando el último término de la suma (7.3) y escribiendo u en lugar de w se obtiene

$$\psi(u+1)f(0) = \psi(u) - \sum_{y=1}^{u} \psi(y)\, f(u+1-y) - \overline{F}(u). \qquad (7.4)$$

Sumando ahora las ecuaciones de (7.3) para valores de w de cero a cualquier $u \geqslant 0$, se tiene que

$$\begin{aligned}
\sum_{w=0}^{u} \psi(w) &= \sum_{w=0}^{u} \sum_{y=1}^{w+1} \psi(y)\, f(w+1-y) + \sum_{w=0}^{u} \overline{F}(w) \\
&= \sum_{y=1}^{u+1} \psi(y) \sum_{w=y-1}^{u} f(w+1-y) + \sum_{w=0}^{u} \overline{F}(w) \\
&= \sum_{y=1}^{u+1} \psi(y) F(u+1-y) + \sum_{w=0}^{u} \overline{F}(w) \\
&= \psi(u+1)\, f(0) + \sum_{y=1}^{u} \psi(y) F(u+1-y) + \sum_{w=0}^{u} \overline{F}(w)
\end{aligned}$$

Para obtener la última identidad se separa el último sumando de la primera suma y se observa que $F(0) = f(0)$. Despejando este término se obtiene

$$\psi(u+1)\, f(0) = \psi(0) + \sum_{y=1}^{u} \psi(y)(1 - F(u+1-y)) - \sum_{y=0}^{u} \overline{F}(y). \qquad (7.5)$$

Igualando el lado derecho de esta ecuación con el lado derecho de la ecuación (7.4) se obtiene

$$\begin{aligned}
\psi(u) &= \psi(0) + \sum_{y=1}^{u} \psi(y)(1 - F(u+1-y) + f(u+1-y)) - \sum_{y=0}^{u-1} \overline{F}(y) \\
&= \psi(0) + \sum_{y=1}^{u} \psi(y)\, \overline{F}(u-y) - \sum_{y=0}^{u-1} \overline{F}(y) \\
&= \psi(0) + \sum_{y=0}^{u-1} \psi(u-y)\, \overline{F}(y) - \sum_{y=0}^{u-1} \overline{F}(y). \qquad (7.6)
\end{aligned}$$

Esto demuestra la primera parte de la proposición. Resta entonces demostrar que $\psi(0) = E(Y)$. Para ello tomaremos el límite cuando $u \to \infty$ en la ecuación (7.6). Tenemos entonces que

$$0 = \lim_{u\to\infty} \psi(u) = \psi(0) + \lim_{u\to\infty} \sum_{y=0}^{u-1} \psi(u-y)\,\overline{F}(y) - \sum_{y=0}^{\infty} \overline{F}(y).$$

Observamos que la segunda suma es $E(Y)$, de modo que es suficiente demostrar que el límite de la primera suma es cero. Como $E(Y) < \infty$, para cualquier $\epsilon > 0$ puede encontrarse un valor de n natural tal que

$$\sum_{y=n+1}^{\infty} \overline{F}(y) < \epsilon.$$

Entonces

$$\begin{aligned} \sum_{y=0}^{u-1} \psi(u-y)\,\overline{F}(y) &\leqslant \sum_{y=0}^{\infty} \psi(u-y)\,\overline{F}(y) \\ &\leqslant \sum_{y=0}^{n} \psi(u-y) + \sum_{y=n+1}^{\infty} \overline{F}(y). \end{aligned}$$

Por lo tanto al tomar el límite cuando $u \to \infty$ obtenemos que

$$\lim_{u\to\infty} \sum_{y=0}^{u-1} \psi(u-y)\,\overline{F}(y) \leqslant \epsilon.$$

Siendo ϵ arbitrario, el límite es efectivamente cero. ■

En analogía con la notación $\overline{F}(y) = 1 - F(y)$, la expresión $\overline{\psi}(u)$ denotará la probabilidad $1-\psi(u)$, esto representa la probabilidad de que la ruina nunca se presente.

Notación	$\overline{\psi}(u) := 1 - \psi(u)$

Usando esta nomenclatura, la ecuación recursiva para $\psi(u)$ también puede ser escrita como sigue

$$\overline{\psi}(u) = \overline{\psi}(0) + \sum_{y=0}^{u-1} \overline{\psi}(u-y)\,\overline{F}(y), \quad u \geqslant 1. \tag{7.7}$$

Por otro lado, recordando nuevamente que Y es una variable aleatoria con valores enteros en el conjunto $\{0, 1, \ldots\}$ y con esperanza finita, se puede escribir

$$E(Y) = \sum_{y=0}^{\infty} \overline{F}(y).$$

De modo que uniendo los dos resultados de la Proposición 7.1 y después de un cambio de variable se encuentra que la fórmula recursiva para $\psi(u)$ también puede ser escrita de la forma siguiente:

$$\psi(u) = \sum_{y=0}^{u-1} \psi(u-y)\overline{F}(y) + \sum_{y=u}^{\infty} \overline{F}(y), \quad u \geqslant 1. \tag{7.8}$$

En el siguiente capítulo encontraremos una fórmula similar a (7.8) para la probabilidad de ruina considerando un modelo de riesgo a tiempo continuo.

Ejemplo 7.1 *Suponga que las reclamaciones Y tienen distribución dada por la tabla que aparece abajo. Usando la fórmula recursiva de la Proposición 7.1, encontraremos $\psi(u)$ para los primeros valores de u.*

y	*0*	*1*	*2*
$f(y)$	*0.5*	*0.2*	*0.3*

Primeramente tenemos que $\psi(0) = E(Y) = 0.8 < 1$. Para $u = 1$ la fórmula recursiva establece que $\psi(1) = \psi(1)\overline{F}(0) + \overline{F}(1)$. Sustituyendo las probabilidades correspondientes se obtiene $\psi(1) = 0.6$. Para $u = 2$ se tiene que $\psi(2) = \psi(1)\overline{F}(1) + \psi(2)\overline{F}(0)$, de donde se obtiene $\psi(2) = 0.36$. Análogamente se obtienen las probabilidades de ruina que se muestran en las siguientes tablas.

u	$\psi(u)$
0	*0.8*
1	*0.6*
2	*0.36*
3	*0.216*
4	*0.1296*
5	*0.07776*

u	$\psi(u)$
6	*0.046656*
7	*0.0279936*
8	*0.01679616*
9	*0.010077696*
10	*0.006046618*
11	*0.0036279708*

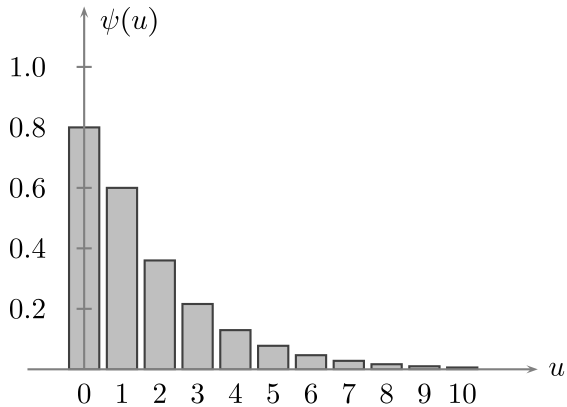


Figura 7.2

En la Figura 7.2 se muestra una gráfica de barras de los valores encontrados de $\psi(u)$. Como era de esperarse, se observa un comportamiento decreciente de la probabilidad de ruina conforme el capital inicial se incrementa.

Ejemplo 7.2 (Problema de la ruina del jugador) *Considere el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$ en donde las reclamaciones tienen la siguiente distribución de probabilidad:*

$$\begin{aligned} P(Y=0) &= p, \\ P(Y=2) &= 1-p, \end{aligned}$$

en donde $1/2 < p < 1$. De esta forma al término de cada periodo el proceso se incrementa en una unidad cuando $Y = 0$, o se reduce en una unidad cuando $Y = 2$. La condición que hemos mencionado para el valor de p garantiza que se cumple la condición de ganancia neta $E(Y) = 2(1-p) < 1$. La variable C_n representa entonces el capital de un jugador A que apuesta una unidad monetaria en cada unidad de tiempo, su capital inicial es u y el jugador contrario B puede considerarse que tiene capital infinito. Encontraremos la probabilidad de que el jugador A (i.e. la compañía aseguradora) eventualmente se arruine. Por los resultados de la Proposición 7.1,

$$\psi(0) = E(Y) = 2(1-p) < 1.$$

Por otro lado, la función $\overline{F}(y)$ es:

$$\begin{aligned}\overline{F}(0) &= 1-p,\\ \overline{F}(1) &= 1-p,\\ \overline{F}(y) &= 0 \quad \text{para } y \geqslant 2,\end{aligned}$$

y entonces, para $u=1$ se tiene que

$$\psi(1) = \psi(0) + \psi(1)\overline{F}(0) - \overline{F}(0).$$

De donde se obtiene

$$\psi(1) = \frac{1-p}{p}.$$

Para $u=2$ se obtiene la ecuación

$$\psi(2) = \psi(0) + \psi(2)\overline{F}(0) + \psi(1)\overline{F}(1) - \overline{F}(0) - \overline{F}(1),$$

que produce la solución

$$\psi(2) = \left(\frac{1-p}{p}\right)^2.$$

Usando inducción sobre u, se puede demostrar que para cualquier capital inicial $u \geqslant 1$,

$$\psi(u) = \left(\frac{1-p}{p}\right)^u.$$

Esta es la probabilidad de ruina del jugador A (compañía aseguradora) cuando su capital inicial es u y es una versión particular de la solución al problema de la ruina del jugador, véase por ejemplo [31]. Se observa que esta probabilidad converge a cero cuando u tiende a infinito.

En general, dada una distribución de probabilidad particular para las reclamaciones, la fórmula recursiva para la probabilidad de ruina de la Proposición 7.1 no produce expresiones compactas. El siguiente es uno de los pocos ejemplos en los que la probabilidad de ruina tiene una fórmula muy corta.

Ejemplo 7.3 (Reclamaciones geométricas) *Considere el modelo de riesgo a tiempo discreto en donde las reclamaciones tienen distribución geo(p), es decir, la función de probabilidad es*

$$f(y) = (1-p)^y\, p, \quad y = 0, 1, 2, \ldots$$

A partir de las fórmulas que aparecen en la Proposición 7.1 pueden encontrarse los valores de $\psi(u)$ de manera sucesiva para $u = 0, 1, 2, \ldots$. Más formalmente, usando inducción sobre el parámetro u puede demostrarse que para cualquier capital inicial $u \geqslant 0$ y para cualquier $p > 1/2$,

$$\psi(u) = \left(\frac{1-p}{p}\right)^{u+1}.$$

Concluimos esta sección mostrando el comportamiento límite de la probabilidad de ruina cuando el capital inicial crece a infinito. La demostración de este resultado se basa en el análisis presentado antes para justificar la hipótesis de ganancia neta en la sección introductoria sobre los principios para el cálculo de primas.

Proposición 7.2 *Para el modelo de riesgo a tiempo discreto y bajo la condición de ganancia neta,*

$$\lim_{u\to\infty} \psi(u) = 0.$$

Demostración. Por la ley fuerte de los grandes números y la condición de ganancia neta tenemos que

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} C_n &= \lim_{n\to\infty} \frac{1}{n}(u + n - \sum_{j=1}^{n} Y_j) \\
&= 1 - \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} Y_j \\
&= 1 - E(Y) > 0.
\end{aligned}$$

Así, este comportamiento límite implica forzosamente la v.a. C_n diverge a infinito casi seguramente cuando $n \to \infty$. En consecuencia, la variable $\inf_{n\geqslant 0} C_n$ está acotada por abajo casi seguramente. Por lo tanto, tomando un capital inicial u suficientemente grande, la cota inferior de $\inf_{n\geqslant 0} C_n$ puede hacerse igual a cero, es decir,

$$\inf_{n\geqslant 0} C_n \geqslant 0.$$

Esto quiere decir que $\psi(u) = 0$ cuando $u \to \infty$. ■

7.3. Probabilidad de ruina con horizonte finito

La probabilidad de ruina con horizonte finito $n \geqslant 1$ se define como

$$\begin{aligned} \psi(u,n) &= P(\tau \leqslant n \,|\, C_0 = u) \\ &= P(\tau \in \{1,2,\ldots,n\} \,|\, C_0 = u), \end{aligned}$$

y corresponde a la probabilidad de que la ruina se presente en alguno de los tiempos: $1,2,\ldots,n$. Puesto que siempre estaremos en el caso $C_0 = u$, se omitirá esta condición y la probabilidad de ruina con horizonte finito se puede escribir simplemente como $P(\tau \leqslant n)$. Observe entonces que $\psi(u,n)$ es la función de distribución del tiempo de ruina τ evaluada en el valor entero n. Naturalmente estamos interesados en conocer esta función. A partir de observar la contención de los eventos correspondientes se puede verificar que

$$\psi(u,1) \leqslant \psi(u,2) \leqslant \cdots \leqslant \psi(u,n) \leqslant \psi(u).$$

En particular cuando el capital tiende a infinito esta probabilidad de ruina también se anula, es decir, para cualquier $n \geqslant 1$ fijo,

$$\lim_{u\to\infty} \psi(u,n) \leqslant \lim_{u\to\infty} \psi(u) = 0.$$

Condicionando sobre el monto de la primera reclamación tal como se hizo en el caso de horizonte infinito, mostramos a continuación una forma recursiva de encontrar $\psi(u,n)$.

Proposición 7.3 *Para el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$ con valor inicial $u \geqslant 0$, la probabilidad de ruina con horizonte finito $\psi(u,n)$ puede calcularse de la siguiente forma*

1. $\psi(u,1) = \overline{F}(u)$.
2. $\psi(u,n) = \overline{F}(u) + \sum_{y=0}^{u} \psi(u+1-y, n-1)\, f(y), \quad n \geqslant 2.$

Demostración.
Para $n = 1$,

$$\psi(u,1) = P(\tau = 1) = P(u + 1 - Y_1 \leqslant 0) = P(Y_1 \geqslant u + 1) = \overline{F}(u).$$

Para $n \geqslant 2$, condicionando sobre el valor de la primera reclamación,

$$\begin{aligned}
\psi(u,n) &= P(\tau \leqslant n) \\
&= \sum_{y=0}^{\infty} P(\tau \leqslant n \,|\, Y_1 = y)\, P(Y_1 = y) \\
&= \sum_{y=0}^{u} \psi(u+1-y, n-1)\, f(y) + \sum_{y=u+1}^{\infty} f(y) \\
&= \sum_{y=0}^{u} \psi(u+1-y, n-1)\, f(y) + \overline{F}(u).
\end{aligned}$$

■

Ejemplo 7.4 *Consideremos nuevamente el caso cuando los montos Y tienen distribución dada por la tabla que aparece abajo. Usando la fórmula recursiva de la Proposición 7.3 encontraremos $\psi(u,n)$, cuando $u = 0$ y $n = 1, 2, 3, 4, 5$.*

y	0	1	2
$f(y)$	0.5	0.2	0.3

Para $n = 1$ tenemos que $\psi(0,1) = \overline{F}(0) = 0.5$. Para $n = 2$, la fórmula recursiva lleva a la ecuación $\psi(0,2) = \psi(0,1) + \psi(1,1)f(0)$, usando el hecho de que $\psi(1,1) = 0.3$ se obtiene $\psi(0,2) = 0.65$. Para $n = 3$, se tiene que $\psi(0,3) = \psi(0,1) + \psi(1,2)f(0)$, en donde $\psi(1,2)$ se calcula usando la misma fórmula recursiva. Al hacer los cálculos se obtiene $\psi(0,3) = 0.68$. Análogamente y utilizando repetidamente la fórmula recursiva se obtienen las probabilidades de ruina restantes que se muestran en la siguiente tabla.

n	1	2	3	4	5
$\psi(0,n)$	0.5	0.65	0.68	0.7085	0.7232

El comportamiento creciente de $n \mapsto \psi(0,n)$ *se muestra en la Figura 7.3 y tales probabilidades son siempre menores o iguales a* $\psi(0) = 0.8$*, es decir,*

$$\psi(0,1) \leqslant \psi(0,2) \leqslant \cdots \leqslant \psi(0,5) \leqslant \cdots \leqslant \psi(0) = 0.8\,.$$

El valor $\psi(0) = 0.8$ *fue calculado en el Ejemplo 7.1.*

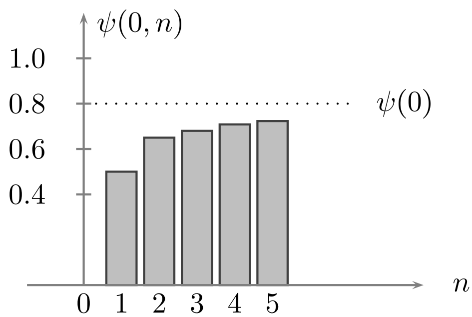


Figura 7.3

7.4. Coeficiente de ajuste

Para el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$, vamos a definir ahora un número llamado coeficiente de ajuste o exponente de Lundberg. La definición aparece a continuación y por ahora se presenta sin motivación ni justificación alguna. Las razones por las cuales tal número se define de esa manera serán evidentes en la siguiente sección en donde encontraremos una cota superior para la probabilidad de ruina en el modelo de riesgo mencionado y en donde aparece de manera natural el coeficiente de ajuste.

Definición 7.3 (Coeficiente de ajuste) *Suponga que la función generadora de momentos de las reclamaciones Y en el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$ existe y que $E(Y) < 1$. El coeficiente de ajuste para este proceso de riesgo se define como la única solución positiva r de la ecuación*

$$E[e^{r(Y-1)}] = 1. \tag{7.9}$$

A tal solución positiva se le denota por R.

Explícitamente el coeficiente de ajuste es aquel número positivo r tal que

$$\sum_{y=0}^{\infty} e^{r(y-1)} f(y) = 1.$$

Para justificar que esta ecuación tiene efectivamente una única solución positiva se define la función

$$\theta(r) = E[e^{r(Y-1)}]$$

y se comprueba que $\theta(r)$ cumple las siguientes propiedades:

a) $\theta(r) > 0 \quad$ para $r \geqslant 0$.

b) $\theta(0) = 1$.

c) $\theta'(r) = E[(Y-1)e^{r(Y-1)}]$.

d) $\theta''(r) = E[(Y-1)^2 e^{r(Y-1)}]$.

e) $\lim\limits_{r\to\infty} \theta(r) = \infty$, cuando $F(1) < 1$.

Así, el comportamiento de $\theta(r)$ alrededor de cero es el siguiente: toma el valor 1 en $r = 0$, la derivada por la derecha en $r = 0$ es $\theta'(0) = E(Y-1)$, que resulta ser negativa por la hipótesis de ganancia neta $E(Y) < 1$. La segunda derivada en cero es $\theta''(0) = E[(Y-1)^2] > 0$. Finalmente, si $F(1) < 1$ entonces $\theta(r)$ crece a infinito cuando $r \to \infty$. Esta afirmación se verifica más abajo. Todas estas observaciones acerca de la función $\theta(r)$ demuestran

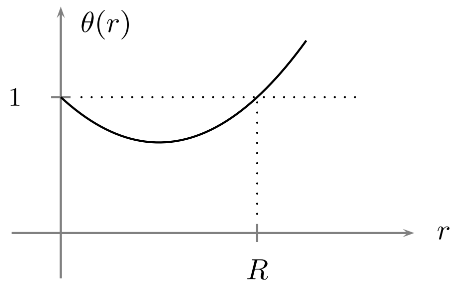


Figura 7.4

que existe un único valor r tal que $\theta(r) = 1$ y ese valor es el coeficiente de ajuste R. Una gráfica aproximada de $\theta(r)$ se muestra en la Figura 7.4.

Veamos más detenidamente el comportamiento límite de la función $\theta(r)$. Supongamos que la distribución de las reclamaciones es tal que $F(1) < 1$. Entonces para cualquier $r > 0$,

$$\theta(r) = \sum_{y=0}^{\infty} e^{r(y-1)} f(y) > \sum_{y=2}^{\infty} e^{r(y-1)} f(y) > e^{r} \sum_{y=2}^{\infty} f(y) = e^{r}(1 - F(1)).$$

Así, cuando $F(1) < 1$ el último factor es positivo y la función $\theta(r)$ crece sin límite cuando $r \to \infty$.

Ejemplo 7.5 (Reclamaciones Bernoulli) *Consideremos el caso de reclamaciones con distribución Ber(p), con $0 \leqslant p < 1$. En este caso, $F(1) = 1$ y por lo tanto no se satisface la condición $F(1) < 1$. La función $\theta(r)$ es*

$$\theta(r) = (1 - p)\, e^{-r} + p.$$

Es claro que esta función decrece exponencialmente cuando $r \to \infty$ y por lo tanto no cruza su valor inicial $\theta(0) = 1$. Por lo tanto no existe el coeficiente de ajuste en este caso. Para este tipo de reclamaciones las probabilidades de ruina son: $\psi(0) = p$ y $\psi(u) = 0$ para $u \geqslant 1$.

Observamos que de todas las distribuciones para las reclamaciones Y tales que $E(Y) < 1$, la única distribución de probabilidad que no satisface las

condición $F(1) < 1$ es la distribución Ber(p) de parámetro $0 \leqslant p < 1$. En conclusión, exceptuando esta distribución, se puede garantizar la existencia del coeficiente de ajuste para cualquier distribución de probabilidad sobre el conjunto $\{0, 1, 2, \ldots\}$ con función generadora de momentos finita y con esperanza menor a 1.
En general no es sencillo resolver la ecuación (7.9) para encontrar el coeficiente de ajuste. En una situación particular, a menudo es necesario usar algún procedimiento numérico como el método de Newton-Raphson para aproximar el valor de este coeficiente. En el apéndice el lector puede encontrar una breve descripción del método de Newton-Raphson. El siguiente ejemplo es atípico, pues se puede encontrar el coeficiente de ajuste con facilidad.

Ejemplo 7.6 (Problema de la ruina del jugador, continuación) *Considere nuevamente la distribución de probabilidad para las reclamaciones como en el Ejemplo 7.2 del problema de la ruina del jugador:*

$$\begin{aligned} P(Y = 0) &= p, \\ P(Y = 2) &= 1 - p, \end{aligned}$$

en donde $1/2 < p < 1$. El coeficiente de ajuste en este caso se calcula como la solución positiva r de la ecuación

$$E(e^{r(Y-1}) = pe^{-r} + (1-p)e^{r} = 1.$$

Esto equivale a resolver una ecuación cuadrática, cuyas soluciones son

$$r = \left\{ \begin{array}{l} 0, \\ \ln\left(\frac{p}{1-p}\right). \end{array} \right.$$

Así, el coeficiente de ajuste es $R = \ln\left(\frac{p}{1-p}\right)$.

7.5. Desigualdad de Lundberg

Demostraremos a continuación que es posible encontrar una cota superior para las probabilidades de ruina en el modelo de riesgo a tiempo discreto. Tal cota superior está dada en términos del coeficiente de ajuste.

Proposición 7.4 (Desigualdad de Lundberg) *Para el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$ con capital inicial $u \geqslant 0$ y cuando el coeficiente de ajuste R existe, las probabilidades de ruina con horizonte finito e infinito satisfacen las siguientes desigualdades:*

1. $\psi(u,n) \leqslant e^{-Ru}, \quad n \geqslant 1.$
2. $\psi(u) \leqslant e^{-Ru}.$

Demostración. Primeramente observamos que el segundo resultado es consecuencia del primero, pues

$$\psi(u) = \lim_{n\to\infty} \psi(u,n) \leqslant \lim_{n\to\infty} e^{-Ru} = e^{-Ru}.$$

Para demostrar el primer resultado usaremos el método de inducción sobre el parámetro n. Para $n = 1$ tenemos que, por la Proposición 7.3 y la existencia del coeficiente de ajuste R,

$$\begin{aligned} \psi(u,1) &= \overline{F}(u) \\ &= \sum_{y=u+1}^{\infty} f(y) \\ &\leqslant \sum_{y=u+1}^{\infty} e^{-R(u+1-y)} f(y) \\ &\leqslant \sum_{y=0}^{\infty} e^{-R(u+1-y)} f(y) \\ &= e^{-Ru} \sum_{y=0}^{\infty} e^{R(y-1)} f(y) \\ &= e^{-Ru}. \end{aligned}$$

En la penúltima igualdad la suma indicada vale 1, pues ésa es justamente la definición de coeficiente de ajuste. Supongamos ahora que $\psi(u,n) \leqslant e^{-Ru}$ para algún valor entero fijo $n \geqslant 1$. Entonces, nuevamente por la fórmula general de la Proposición 7.3,

$$\begin{aligned}
\psi(u,n+1) &= \psi(u,1)+\sum_{y=0}^{u}\psi(u+1-y,n)\,f(y)\\
&\leqslant \sum_{y=u+1}^{\infty} f(y)+\sum_{y=0}^{u} e^{-R(u+1-y)}\,f(y)\\
&\leqslant \sum_{y=u+1}^{\infty} e^{-R(u+1-y)}\,f(y)+\sum_{y=0}^{u} e^{-R(u+1-y)}\,f(y)\\
&= \sum_{y=0}^{\infty} e^{-R(u+1-y)}\,f(y)\\
&= e^{-Ru}\sum_{y=0}^{\infty} e^{R(y-1)}\,f(y)\\
&= e^{-Ru}.
\end{aligned}$$

■

En el caso cuando el coeficiente de ajuste existe, es una consecuencia inmediata de la desigualdad de Lundberg que las probabilidades de ruina se anulan cuando el capital inicial tiende a infinito, es decir,

$$\lim_{u\to\infty}\psi(u)=\lim_{u\to\infty}\psi(u,n)=0.$$

Con ayuda del siguiente resultado daremos una demostración alternativa de la desigualdad de Lundberg, esta vez usando la teoría de martingalas.

Proposición 7.5 *Sea $\{C_n : n \geqslant 0\}$ el proceso de riesgo a tiempo discreto. Suponga que el coeficiente de ajuste R existe. Entonces el proceso abajo especificado es una martingala a tiempo discreto.*

$$\{e^{-RC_n} : n \geqslant 0\}.$$

Demostración. Sea $(\mathscr{F}_n)_{n\geqslant 0}$ la filtración natural del proceso $\{e^{-RC_n} : n \geqslant 0\}$. La existencia del coeficiente de ajuste garantiza que cada variable de este proceso tiene esperanza finita. En efecto, no es difícil comprobar que

$$E(e^{-RC_n}) = e^{-Ru}.$$

Por otro lado, para cualquier $n \geqslant 0$,

$$\begin{aligned} E(e^{-RC_{n+1}} \,|\, \mathscr{F}_n) &= E(e^{-R(C_n+1-Y_{n+1})} \,|\, \mathscr{F}_n) \\ &= e^{-RC_n}\, E(e^{R(Y-1)}) \\ &= e^{-RC_n}. \end{aligned}$$

En la segunda igualdad hemos usado la hipótesis de independencia de las reclamaciones y para la última igualdad se usó la definición de coeficiente de ajuste. ■

Demostración. *(Desigualdad de Lundberg, segunda demostración).* Sea τ el tiempo de paro definido como el momento de la ruina en el proceso de riesgo a tiempo discreto. Como $\{C_n : n \geqslant 0\}$ es una martingala, también lo es $\{C_{n\wedge\tau} : n \geqslant 0\}$. Recordemos que $x \wedge y = \text{mín}\{x, y\}$. Ambas martingalas comienzan en el valor e^{-Ru}. Entonces para cualquier $n \geqslant 1$,

$$\begin{aligned} e^{-Ru} &= e^{-RC_0} \\ &= E(e^{-RC_{n\wedge\tau}}) \\ &= E(e^{-RC_{n\wedge\tau}} \,|\, \tau \leqslant n)\, P(\tau \leqslant n) \\ &\qquad + E(e^{-RC_{n\wedge\tau}} \,|\, \tau > n)\, P(\tau > n) \\ &\geqslant E(e^{-RC_{n\wedge\tau}} \,|\, \tau \leqslant n)\, P(\tau \leqslant n) \\ &= E(e^{-RC_{\tau}} \,|\, \tau \leqslant n)\, P(\tau \leqslant n). \end{aligned}$$

Haciendo $n \to \infty$ monótonamente puede demostrarse que el evento $(\tau \leqslant n)$ converge crecientemente al evento $(\tau < \infty)$ y que la esperanza condicional indicada converge a $E(e^{-RC_\tau} \,|\, \tau < \infty)$. Por lo tanto,

$$\begin{aligned} e^{-Ru} &\geqslant E(e^{-RC_\tau} \,|\, \tau < \infty)\, P(\tau < \infty) \\ &\geqslant E(\,1 \,|\, \tau < \infty)\, P(\tau < \infty) \\ &= P(\tau < \infty) \\ &= \psi(u). \end{aligned}$$

Las probabilidades de ruina con horizonte finito también tienen la misma cota superior, pues para cualquier $n \geqslant 1$,

$$\psi(u,n) \leqslant \psi(u) \leqslant e^{-Ru}.$$

■

Ejemplo 7.7 (Problema de la ruina del jugador, continuación) *En el Ejemplo 7.2 hemos calculado la probabilidad exacta de ruina para el modelo de riesgo cuando las reclamaciones son tales que* $P(Y=0)=p$ *y* $P(Y=2)=1-p$*, con* $2(1-p)<1$*. Esta probabilidad es, para* $u \geqslant 1$*,*

$$\psi(u) = \left(\frac{1-p}{p}\right)^{u}.$$

Por otro lado, en el Ejemplo 7.6 hemos encontrado que el coeficiente de ajuste en este caso es

$$R = \ln\left(\frac{p}{1-p}\right).$$

Podemos entonces comparar $\psi(u)$ *con la cota de Lundberg* e^{-Ru}*. Después de algunos cálculos sencillos puede comprobarse que estas cantidades coinciden, es decir,*

$$\psi(u) = e^{-Ru}, \quad u \geqslant 1.$$

Esto demuestra que sin ninguna otra condición adicional, la cota superior de Lundberg es óptima.

7.6. Severidad de la ruina

Consideremos nuevamente el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$ con capital inicial $u \geqslant 0$. Sea τ el tiempo de ruina en este modelo y defina la probabilidad

$$\varphi(u,z) = P(\tau < \infty, -C_\tau \leqslant z),$$

para $u = 0,1,2,\ldots$ y $z = 0,1,2,\ldots$ Observe que para hacer la escritura más corta, hemos omitido escribir esta expresión como una probabilidad condicional cuando $C_0 = u$. Esta función representa la probabilidad conjunta de que la ruina ocurra en un tiempo finito y que el déficit de la aseguradora al momento de la ruina sea menor o igual al valor z. A la variable aleatoria $-C_\tau$ se le llama severidad de la ruina o déficit al momento de la ruina. Se muestra gráficamente su significado en la Figura 7.5.

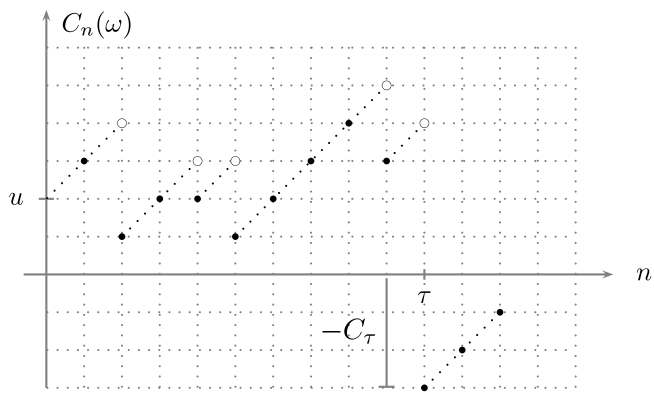


Figura 7.5

Demostraremos a continuación una fórmula recursiva para la probabilidad $\varphi(u, z)$. Para ello usaremos el hecho de que dicha probabilidad se anula cuando el capital inicial es infinito, es decir,

$$\lim_{u\to\infty} \varphi(u, z) \leqslant \lim_{u\to\infty} \varphi(u) = 0.$$

El método de la demostración es muy similar al presentado para la probabilidad de ruina con horizonte infinito $\psi(u)$. En la sección de ejercicios aparece una fórmula recursiva para la severidad de la ruina con horizonte finito.

Proposición 7.6 *Para el proceso de riesgo a tiempo discreto* $\{C_n : n \geqslant 0\}$ *con capital inicial* $u \geqslant 0$ *y tiempo de ruina* τ, *la función*

$$\varphi(u,z) = P(\tau < \infty, -C_\tau \leqslant z),$$

para $z = 0, 1, \ldots$, *satisface las siguientes identidades:*

1. $\varphi(u,z) = \varphi(0,z) + \sum_{y=0}^{u-1} \varphi(u-y,z)\, \overline{F}(y) - \sum_{y=0}^{z} [F(u+y) - F(y)].$

2. $\varphi(0,z) = \sum_{y=0}^{z} \overline{F}(y).$

Demostración. Condicionando sobre el monto de la primera reclamación,

$$\begin{aligned} \varphi(u,z) &= \sum_{y=0}^{\infty} P(\tau < \infty, C_\tau > -z \,|\, Y_1 = y)\, f(y) \\ &= \sum_{y=0}^{u} \varphi(u+1-y,z)\, f(y) + \sum_{y=u+1}^{u+z+1} f(y). \end{aligned}$$

Cuando el monto y toma un valor entre cero y u, no hay ruina, pues al tiempo uno el capital de la aseguradora se ha incrementado a $u+1$, por la propiedad de incrementos independientes del proceso, la probabilidad condicional se reduce a $\varphi(u+1-y,z)$. Si el monto y está entre $u+1$ y $u+z+1$, entonces hay ruina y se cumple la condición $-C_\tau \leqslant z$, por lo tanto la probabilidad condicional es idénticamente uno. Finalmente, si y es mayor a $u+z+1$, hay ruina y la ruina es severa en el sentido de que no se cumple la condición $-C_\tau \leqslant z$, en este caso la probabilidad condicional es cero y es por ello que no aparece en la última igualdad. La última ecuación puede escribirse de la siguiente forma:

$$\varphi(u,z) = \sum_{y=1}^{u+1} \varphi(y,z)\, f(u+1-y) + F(u+z+1) - F(u). \tag{7.10}$$

Escribiendo k en lugar de u y sumando estas expresiones para valores k de 0 a u se obtiene

$$\begin{aligned}\sum_{k=0}^{u}\varphi(k,z) &= \sum_{k=0}^{u}\sum_{y=1}^{k+1}\varphi(y,z)\,f(k+1-y)+\sum_{k=0}^{u}[F(k+z+1)-F(k)]\\ &= \sum_{y=1}^{u+1}\sum_{k=y-1}^{u}\varphi(y,z)\,f(k+1-y)+\sum_{k=0}^{u}[F(k+z+1)-F(k)]\\ &= \sum_{y=1}^{u+1}\varphi(y,z)\,F(u+1-y)+\sum_{k=0}^{u}[F(k+z+1)-F(k)]\end{aligned}$$

Despejando el término $\varphi(u+1,z)$,

$$\begin{aligned}\varphi(u+1,z)F(0) &= \sum_{k=0}^{u}\varphi(k,z)-\sum_{y=1}^{u}\varphi(y,z)\,F(u+1-y)\\ &\quad -\sum_{k=0}^{u}[F(k+z+1)-F(k)]\\ &= \varphi(0,z)+\sum_{y=1}^{u}\varphi(y,z)\,(1-F(u+1-y))\\ &\quad -\sum_{k=0}^{u}[F(k+z+1)-F(k)]. \qquad (7.11)\end{aligned}$$

Por otro lado, de la ecuación (7.10) se obtiene

$$\varphi(u+1,z)F(0)=\varphi(u,z)-\sum_{y=1}^{u}\varphi(y,z)\,f(u+1-y)-[F(u+z+1)-F(u)]. \qquad (7.12)$$

Igualando el lado derecho de las ecuaciones (7.11) y (7.12),

$$\varphi(u,z)=\varphi(0,z)+\sum_{y=1}^{u}\varphi(y,z)\,\overline{F}(u-y)-\sum_{k=0}^{u-1}[F(k+z+1)-F(k)].$$

La segunda suma puede expresarse de la siguiente forma:

$$
\begin{aligned}
\sum_{k=0}^{u-1} [F(k+z+1) - F(k)] &= \sum_{k=0}^{u-1} \sum_{y=0}^{z} [F(k+y+1) - F(k+y)] \\
&= \sum_{y=0}^{z} \sum_{k=0}^{u-1} [F(k+y+1) - F(k+y)] \\
&= \sum_{y=0}^{z} [F(u+y) - F(y)].
\end{aligned}
$$

Por lo tanto,

$$
\varphi(u,z) = \varphi(0,z) + \sum_{y=1}^{u} \varphi(y,z)\,\overline{F}(u-y) - \sum_{y=0}^{z} [F(u+y) - F(y)]. \tag{7.13}
$$

Haciendo un cambio de variable en la primera suma se obtiene la primera fórmula de la proposición. Para obtener la segunda fórmula hacemos u tender a infinito en la expresión recién encontrada (7.13). Así, tenemos que

$$
\begin{aligned}
0 &= \lim_{u\to\infty} \varphi(u,z) \\
&= \varphi(0,z) + \lim_{u\to\infty} \sum_{y=1}^{u} \varphi(y,z)\,\overline{F}(u-y) - \sum_{y=0}^{z} (1 - F(y)).
\end{aligned}
$$

Resta demostrar que el último límite indicado es cero. Tenemos que

$$
\begin{aligned}
\sum_{y=1}^{u} \varphi(y,z)\,\overline{F}(u-y) &= \sum_{v=0}^{u-1} \varphi(u-v,z)\,\overline{F}(v) \\
&\leqslant \sum_{v=0}^{\infty} \varphi(u-v,z)\,\overline{F}(v) \\
&\leqslant \sum_{v=0}^{n} \varphi(u-v,z) + \sum_{v=n+1}^{\infty} \overline{F}(v),
\end{aligned}
$$

en donde n es cualquier entero tal que $\sum_{v=n+1}^{\infty} \overline{F}(v) < \epsilon$, para cualquier $\epsilon > 0$. Ello es posible pues $E(Y) = \sum_{v=0}^{\infty} \overline{F}(v) < \infty$. Así, al hacer u tender a infinito la primera suma se anula y la segunda suma, como hemos visto, puede hacerse tan pequeña como se desee. ■

Es interesante observar que cuando $z \to \infty$, es decir, cuando no hay cota para la severidad de la ruina, la probabilidad $\varphi(u,z)$ converge a la probabilidad de ruina $\psi(u)$ y las ecuaciones de la Proposición 7.6 se reducen a las estudiadas antes en la Proposición 7.1.

Por otro lado, incorporando la expresión de $\varphi(0,z)$ en la fórmula recursiva para $\varphi(u,z)$ y haciendo algunas simplificaciones, puede encontrarse la siguiente expresión alternativa

$$\varphi(u,z) = \sum_{y=0}^{u-1} \varphi(u-y,z)\, \overline{F}(y) + \sum_{y=0}^{z} \overline{F}(u+y).$$

Comparando esta expresión con la fórmula para $\psi(u)$ dada por la ecuación (7.8) de la página 171 podemos escribir la diferencia entre $\psi(u)$ y $\varphi(u,z)$ de la siguiente forma:

$$\psi(u) - \varphi(u,z) = \sum_{y=0}^{u-1} [\psi(u-y) - \varphi(u-y,z)]\, \overline{F}(y) + \sum_{y=u+z+1}^{\infty} \overline{F}(y).$$

Pero observemos que

$$\begin{aligned} \psi(u) - \varphi(u,z) &= P(\tau < \infty) - P(\tau < \infty, -C_\tau \leqslant z) \\ &= P(\tau < \infty, -C_\tau > z). \end{aligned}$$

Por lo tanto hemos demostrado que la función

$$\varphi_1(u,z) = P(\tau < \infty, -C_\tau > z)$$

satisface la ecuación recursiva

$$\varphi_1(u,z) = \sum_{y=0}^{u-1} \varphi_1(u-y,z)\, \overline{F}(y) + \sum_{y=u+z+1}^{\infty} \overline{F}(y).$$

Ejemplo 7.8 (Reclamaciones geométricas) *Considere el modelo de riesgo a tiempo discreto en donde las reclamaciones tienen distribución geo(p), es decir, la correspondiente función de probabilidad es*

$$f(y) = (1-p)^y\, p, \quad y = 0, 1, 2, \ldots$$

A partir de las fórmulas que aparecen en la Proposición 7.6 y usando inducción sobre el parámetro u puede demostrarse que

$$\varphi(u,z) = [\,1-(1-p)^{z+1}\,]\left(\frac{1-p}{p}\right)^{u+1}.$$

Al hacer z tender a infinito obtenemos las probabilidades de ruina con horizonte infinito cuando las reclamaciones son geométricas. Estas probabilidades aparecen en el Ejemplo 7.3.

$$\lim_{z\to\infty} \varphi(u,z) = \left(\frac{1-p}{p}\right)^{u+1} = \psi(u).$$

En el siguiente capítulo estudiaremos un modelo de riesgo a tiempo continuo muy similar al que hemos estudiado hasta ahora. Algunos de los conceptos, notación, técnicas y definiciones que hemos utilizado serán extendidas al caso cuando el tiempo se mide de manera continua.

Comentarios y referencias

En el presente capítulo hemos estudiado el problema de la ruina en el proceso de riesgo a tiempo discreto no trivial más sencillo posible. El modelo que hemos estudiado puede plantearse con características más generales, pero al considerar el modelo elemental estudiado, nuestro objetivo ha sido mostrar el tipo de problemas matemáticos que pueden plantearse y en algunos casos mostrar las respuestas que se obtienen en casos particulares. En la respuesta general a estos problemas hemos encontrado varias fórmulas recursivas que invitan a su implementación en computadora. En algunos casos las probabilidades involucradas pueden tomar valores muy cercanos a cero y ello podría provocar problemas numéricos. Debe entonces tenerse cuidado al programar estas fórmulas en una computadora. El material expuesto en este capítulo está basado en el capítulo introductorio a la teoría de la ruina del texto de Dickson [13]. Otras referencias en el tema son [7] y [23]. El lector interesado en el estudio más profundo sobre los modelos de riesgo a tiempo discreto puede consultar artículos de investigación panorámicos como [24].

7.7. Ejercicios

Proceso de riesgo a tiempo discreto

154. Encuentre la matriz de probabilidades de transición en un paso de la cadena de Markov dada por el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$.

155. El total de montos por reclamaciones durante cada periodo unitario en el proceso de riesgo a tiempo discreto se ha modelado mediante una variable aleatoria discreta Y con valores en el conjunto $\{0, 1, \ldots\}$ y se ha supuesto que la esperanza de esta variable es finita. Demuestre que
$$E(Y) = \sum_{y=0}^{\infty} \overline{F}(y).$$

156. Para el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$, use la ley fuerte de los grandes números para demostrar que, en el sentido casi seguro,
$$\lim_{n\to\infty} C_n = \begin{cases} \infty & \text{si } E(Y) < 1, \\ -\infty & \text{si } E(Y) > 1. \end{cases}$$

Probabilidad de ruina con horizonte infinito

157. Calcule la probabilidad de ruina $\psi(u)$ para $u = 0, 1, 2, 3, 4, 5$ en el modelo de riesgo a tiempo discreto cuando las reclamaciones tienen la distribución de probabilidad dada por la siguiente tabla.

y	0	1	2
$f(y)$	3/4	1/8	1/8

158. Suponga que las reclamaciones en el modelo de riesgo a tiempo discreto tienen la siguiente función de probabilidad: para algún entero $k \geqslant 2$ fijo,
$$\begin{aligned} f(0) &= 1 - p, \\ f(k) &= p. \end{aligned}$$

Demuestre que:

a) $\psi(u) = \dfrac{p}{1-p}\left(k - u + \sum_{y=1}^{u-1} \psi(u-y)\right)$ para $1 \leqslant u < k$.

b) $\psi(u) = \dfrac{p}{1-p}\sum_{y=1}^{k-1} \psi(u-y)$ para $u \geqslant k$.

159. A partir de la fórmula recursiva para la probabilidad de ruina $\psi(u)$ en el proceso de riesgo a tiempo discreto que aparece en el enunciado de la Proposición 7.1, demuestre que:

a) $\psi(u+1) \leqslant \psi(u)$.

b) $\lim\limits_{u\to\infty} \psi(u) = 0$.

160. Hemos demostrado que la función $u \mapsto \psi(u)$ satisface la ecuación recursiva dada en la Proposición 7.1. Demuestre que esta ecuación recursiva tiene una única solución. Sugerencia: suponga que $\psi_1(u)$ y $\psi_2(u)$ son dos funciones que cumplen las propiedades que se enuncian en la Proposición 7.1. Use inducción sobre u para demostrar que $\psi_1(u) - \psi_2(u) = 0$, $u \geqslant 0$.

Probabilidad de ruina con horizonte finito

161. Calcule la probabilidad de ruina con horizonte finito $\psi(u, n)$ para $u = 1$ y $n = 1, 2, 3, 4, 5$ en el modelo de riesgo a tiempo discreto, cuando las reclamaciones tienen la distribución de probabilidad dada por la siguiente tabla.

y	0	1	2
$f(y)$	4/10	5/10	1/10

Coeficiente de ajuste

162. Calcule el coeficiente de ajuste para el modelo de riesgo a tiempo discreto cuando las reclamaciones tienen la distribución de probabilidad dada por la siguiente tabla.

y	0	1	2
$f(y)$	1/3	1/2	1/6

163. Considere el modelo de riesgo a tiempo discreto en donde las reclamaciones tienen distribución geo(p), es decir, la correspondiente función de probabilidad es

$$f(y) = (1-p)^y p, \quad y = 0, 1, 2, \ldots$$

Recuerde que el valor esperado de esta distribución es $E(Y) = (1-p)/p$. Para que se cumpla la condición de ganancia neta $E(Y) < 1$ se necesita suponer $p > 1/2$. Demuestre que el coeficiente de ajuste es

$$R = \ln\left(\frac{p}{1-p}\right).$$

164. Suponga que las reclamaciones tienen la siguiente distribución de probabilidad

$$\begin{aligned} P(Y=0) &= 1-p, \\ P(Y=3) &= p. \end{aligned}$$

Suponga además que $p < 1/3$, con esta restricción se cumple la condición de ganancia neta $E(Y) < 1$ en el modelo de riesgo a tiempo discreto. Demuestre que el coeficiente de ajuste es

$$R = \ln\left(\frac{-p + \sqrt{p^2 + 4p(1-p)}}{2p}\right).$$

165. Sea $k \geqslant 2$ un entero fijo. Considere el modelo de riesgo a tiempo discreto en donde las reclamaciones tienen función de probabilidad

$$\begin{aligned} f(0) &= 1-p, \\ f(k) &= p. \end{aligned}$$

Suponga que $p < 1/k$. Con esta hipótesis se cumple la condición de ganancia neta $E(Y) < 1$. Demuestre que el coeficiente de ajuste es $R = \ln x$, en donde x es la única solución mayor a uno de la ecuación

$$px^{k-1} + px^{k-2} + \cdots + p - 1 = 0.$$

166. Para el proceso de riesgo a tiempo discreto $\{C_n : n \geqslant 0\}$ con capital inicial $u \geqslant 0$ y suponiendo que el coeficiente de ajuste R existe, demuestre que para cualquier $n \geqslant 0$,

$$E(e^{-RC_n}) = e^{-Ru}.$$

Desigualdad de Lundberg

167. Considere los datos del modelo que aparece en el Ejemplo 7.1, en donde se calcularon explícitamente las probabilidades $\psi(u)$ para $u = 0, 1, \ldots, 11$. Demuestre que el coeficiente de ajuste es

$$R = \ln(10/6) = 0.5108256238\,.$$

Para los valores de u indicados, verifique que efectivamente se cumple la desigualdad de Lundberg $\psi(u) \leqslant e^{-Ru}$. El uso de una hoja de cálculo podría ser útil para realizar estas verificaciones.

Severidad de la ruina

168. *Propiedades.* A partir de la definición de la función $\varphi(u, z)$ demuestre que:

a) $u \mapsto \varphi(u, z)$ es decreciente.

b) $z \mapsto \varphi(u, z)$ es creciente.

c) $\varphi(u,z) \leqslant \psi(u)$.

d) $\lim\limits_{z\to\infty} \varphi(u,z) = \psi(u)$.

169. *Severidad con horizonte finito.* Para cada entero $u = 0, 1, 2, \ldots$ y $z = 0, 1, \ldots$ se define la severidad de la ruina con horizonte finito como la función

$$\varphi(u,z;n) = P(\tau \leqslant n, -C_\tau \leqslant z).$$

Demuestre que:

a) $\varphi(u,z;1) = F(u+z+1) - F(u)$.

b) $\varphi(u,z;n) = \varphi(u,z;1) + \sum_{y=0}^{u} \varphi(u+1-y, z; n-1)\, f(y), \quad n \geqslant 2.$

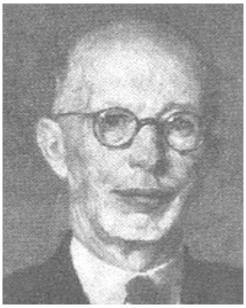
Capítulo 8

Teoría de la ruina: tiempo continuo

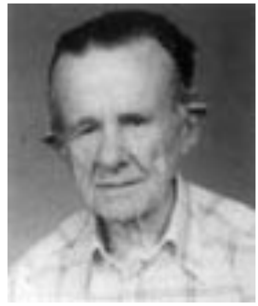
En este capítulo se presenta una versión a tiempo continuo del proceso de riesgo conocida como el modelo clásico de Cramér-Lundberg. Encontraremos que la probabilidad de ruina para este modelo satisface una ecuación integral. Estudiaremos además algunos otros resultados relacionados al cálculo y estimación de la probabilidad de ruina.

8.1. Modelo clásico de Cramér-Lundberg

El modelo de Cramér-Lundberg es una versión a tiempo continuo del modelo a tiempo discreto que estudiamos en el capítulo anterior y tiene sus orígenes en la tesis doctoral de Filip Lundberg defendida en el año de 1903. En este trabajo, Lundberg analiza el reaseguro de riesgos colectivos y presenta el proceso de Poisson compuesto. Lundberg utilizó términos un tanto distintos a los actuales pues en aquellos años aún no se había formalizado la teoría de los procesos estocásticos como la entendemos hoy en día. En 1930 Harald Cramér retoma las ideas originales de Lundberg y las pone en el contexto de los procesos estocásticos, en ese entonces de reciente creación. El modelo ha sido estudiado en extenso, y varias formas de generalizarlo se han propuesto y analizado.



Ernest Filip Oskar Lundberg
(Suecia, 1876–1965)



Carl Harald Cramér
(Suecia, 1893–1985)

Definición 8.1 *El modelo clásico de Cramér-Lundberg es el proceso estocástico a tiempo continuo* $\{C_t : t \geqslant 0\}$ *dado por*

$$C_t = u + c\,t - \sum_{j=1}^{N_t} Y_j, \tag{8.1}$$

en donde u *y* c *son constantes positivas,* $Y_1, Y_2, \ldots$ *es una sucesión de v.a.i.i.d. positivas e independientes del proceso de Poisson* $\{N_t : t \geqslant 0\}$ *de parámetro* λ.

La constante u representa el capital inicial de la compañía aseguradora, ct corresponde a la entrada por primas hasta el tiempo t, Y_j es el monto de la j-ésima reclamación, y el proceso de Poisson $\{N_t : t \geqslant 0\}$ modela la forma en la que las reclamaciones son recibidas. Observe que para un proceso de reclamaciones con distribución Poisson compuesta como en la ecuación (8.1), la esperanza es justamente de la forma ct, y el principio del valor esperado para el cálculo de primas lleva a que el proceso de primas sea lineal como el sugerido en el modelo. La variable C_t representa el balance más sencillo de ingresos menos egresos a tiempo continuo de una compañía aseguradora. Al proceso $\{C_t : t \geqslant 0\}$ se le llama nuevamente proceso de riesgo (*risk* *process*), o proceso de superávit (*surplus process*), y tiene trayectorias como se muestra en la Figura 8.1.

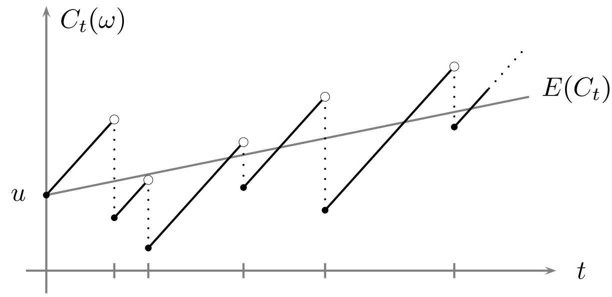


Figura 8.1

Estas trayectorias comienzan siempre en el capital inicial u. Los intervalos en donde estas trayectorias son continuas y crecientes corresponden a periodos en donde no hay reclamaciones. El crecimiento es de la forma ct. Las discontinuidades son siempre saltos hacia abajo, y aparecen en el momento en que se efectúa una reclamación, la cual supondremos que se paga de manera inmediata. El tamaño de un salto es el tamaño de la reclamación dada por la variable Y. Hemos supuesto que los montos $Y_1, Y_2, \ldots$ son variables aleatorias independientes, positivas e idénticamente distribuidas, y en algunas ocasiones supondremos que tienen función generadora de momentos $M_Y(r)$. Los momentos, cuando existan, se denotarán nuevamente por

$$\mu_n = E(Y^n), \quad n \geqslant 1.$$

En particular μ denotará el primer momento μ_1. No es difícil comprobar que

$$\begin{aligned} E(C_t) &= u + (c - \lambda\mu)t, \\ \mathrm{Var}(C_t) &= \lambda\mu_2\, t. \end{aligned}$$

Usando la independencia de las reclamaciones en el proceso de riesgo y la estacionariedad de los incrementos en el proceso de Poisson, puede demostrarse que el proceso de riesgo $\{C_t : t \geqslant 0\}$ tiene incrementos independientes y estacionarios.

La condición de ganancia neta

Sean $T_0, T_1, T_2, \ldots$ los tiempos aleatorios (tiempos de paro) en donde la aseguradora recibe las reclamaciones. Supondremos $T_0 = 0$. Para cada entero $k \geqslant 1$ defina la variable aleatoria $X_k = c\,(T_k - T_{k-1}) - Y_k$, que puede ser interpretada como el balance de la compañía aseguradora entre dos siniestros sucesivos. La esperanza de esta variable es

$$\begin{aligned} E(X_k) &= c\,E(T_k - T_{k-1}) - E(Y_k) \\ &= c\left(\frac{1}{\lambda}\right) - \mu. \end{aligned}$$

Si denotamos por $C_{[k]}$ al proceso de riesgo al momento de la k-ésima reclamación, entonces tenemos que

$$C_{[k]} = u + \sum_{j=1}^{k} X_k.$$

De modo que por la ley de los grandes números,

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{k} C_{[k]} &= \lim_{k\to\infty} \frac{1}{k}\left(u + \sum_{j=1}^{k} X_k\right) \\ &= E(X_k) \\ &= c\left(\frac{1}{\lambda}\right) - \mu. \end{aligned}$$

De la misma forma que se argumentó en el capítulo sobre los principios generales para el cálculo de primas, se puede demostrar que la ruina ocurre casi seguramente si y sólo si, $E(X_k) \leqslant 0$. Como deseamos que esta situación no ocurra supondremos que $E(X_k) > 0$, es decir, tenemos la hipótesis:

Condición de ganancia neta $\quad c > \lambda\mu$

Esta condición ya la habíamos mencionado antes en otros modelos. Ahora la interpretamos para el proceso de riesgo a tiempo continuo de la siguiente forma: la entrada por primas por unidad de tiempo, c, es mayor que el total de reclamaciones promedio por unidad de tiempo, $\lambda\mu$.

Ruina

La trayectoria promedio de este proceso de riesgo es la línea recta que inicia en $u > 0$ y tiene pendiente $c - \lambda\mu$, la cual es positiva por la condición o hipótesis de ganancia neta. Véase la Figura 8.1. Por razones naturales y legales es importante que C_t permanezca por arriba de cierto nivel mínimo. Supongamos que tal nivel mínimo es a, con $0 < a < u$. Ajustando el capital inicial u, esto es, suponiendo un nuevo capital inicial de magnitud $u - a$, se puede suponer, sin pérdida de generalidad, que este nivel mínimo es cero, y así lo haremos en nuestro análisis. De esta forma cuando $C_t < 0$ para algún $t > 0$ se dice que hay ruina.

Definición 8.2 *Se dice que el proceso de riesgo se encuentra en ruina al tiempo $t > 0$ si*

$$C_t < 0,$$

y se define el tiempo de ruina τ como el primer momento en que la ruina se presenta, es decir,

$$\tau = \text{ínf}\,\{t > 0 : C_t < 0\}. \tag{8.2}$$

Nuevamente definiremos $\tau = \infty$ cuando el conjunto indicado en la expresión (8.2) es vacío y corresponde a la situación cuando la ruina nunca se presenta. Observe que hemos definido ruina en el modelo a tiempo continuo cuando ocurre el evento $(C_t < 0)$, a diferencia del evento $(C_n \leqslant 0)$ para el modelo discreto estudiado en el capítulo anterior. La ruina casi nunca sucede en la práctica, es solamente un término técnico que produce alguna toma de decisión. Por ejemplo, si el capital de una compañía aseguradora asignado a una cartera decrece en forma significativa, automáticamente la aseguradora puede tomar ciertas medidas para subsanar esta situación y no se trata de un evento insalvable. Por otro lado, es natural suponer que la compañía aseguradora posea varios portafolios de modo que ruina en uno de ellos no significa necesariamente bancarrota que el término ruina podría sugerir.

En las siguientes secciones nos interesará calcular o estimar probabilidades de ruina en el modelo de Cramér-Lundberg. De manera análoga al modelo de riesgo a tiempo discreto presentado en el capítulo anterior, denotaremos a la probabilidad de ruina con horizonte infinito en el modelo de Cramér-Lundberg como $\psi(u)$, es decir,

$$\psi(u) = P(\tau < \infty \,|\, C_0 = u).$$

Y no habrá ambigüedad en su definición pues el modelo o contexto en el que se estudia determinará el caso correspondiente. Nuevamente escribiremos a esta probabilidad como función del capital inicial u, aunque en realidad depende de todos los parámetros del modelo. Y observamos además, sin proveer una demostración, la monotonía decreciente de esta función, es decir, si $u_1 \leqslant u_2$, entonces

$$\psi(u_1) \geqslant \psi(u_2).$$

Para deducir una ecuación para la probabilidad de ruina $\psi(u)$ tomaremos como preliminarme cierto el hecho de que cuando el capital inicial es infinito la probabilidad de ruina es cero, es decir,

$$\psi(\infty) = \lim_{u\to\infty} \psi(u) = 0.$$

Más adelante daremos una demostración de este resultado y corroboraremos su validez usando la desigualdad de Lundberg en el caso cuando el coeficiente de ajuste existe. Como en el capítulo anterior, usaremos la siguiente nomenclatura.

Notación	$\overline{\psi}(u) := 1 - \psi(u)$

8.2. Probabilidad de ruina con horizonte infinito

Presentaremos a continuación tres resultados generales sobre la probabilidad de ruina con horizonte infinito. A diferencia del primer capítulo, y para hacer la notación más apegada a la literatura existente en el tema, recordemos que estamos denotando por $F(y)$ a la función de distribución de una reclamación Y cualquiera. La función de densidad será $f(y)$, cuando ésta exista.

Proposición 8.1 *Suponga que la función de distribución $F(y)$ de una reclamación en el modelo de Cramér-Lundberg es continua. Entonces*

1. $\displaystyle \frac{d}{du}\,\overline{\psi}(u) = \frac{\lambda}{c}\left(\overline{\psi}(u) - \int_0^u \overline{\psi}(u-y)\,dF(y)\right).$

2. $\displaystyle \psi(0) = \frac{\lambda\mu}{c}.$

3. $\displaystyle \psi(u) = \frac{\lambda}{c}\left(\int_u^\infty \overline{F}(y)\,dy + \int_0^u \psi(u-y)\,\overline{F}(y)\,dy\right).$

Demostración. Usaremos análisis del primer paso condicionando sobre el monto de la primera reclamación Y_1 y el momento T_1 en el que esta reclamación ocurre. Usaremos además el hecho de que T_1 tiene distribución $\exp(\lambda)$.

$$
\begin{aligned}
\overline{\psi}(u) &= P(\text{ “No ruina en } (0,\infty)\text{”} \,|\, C_0 = u) \\
&= \int_0^\infty \int_0^\infty P(\text{ “No ruina en } (0,\infty)\text{”} \,|\, Y_1 = y, T_1 = t)\,dF(y)\,f_{T_1}(t)\,dt \\
&= \int_0^\infty \int_0^{u+ct} P(\text{ “No ruina en } (0,\infty)\text{”} \,|\, Y_1 = y, T_1 = t)\,dF(y)\,f_{T_1}(t)\,dt \\
&= \int_0^\infty \lambda e^{-\lambda t} \int_0^{u+ct} P(\text{ “No ruina en } (t,\infty)\text{”} \,|\, Y_1 = y, T_1 = t)\,dF(y)\,dt \\
&= \int_0^\infty \lambda e^{-\lambda t} \int_0^{u+ct} \overline{\psi}(u+ct-y)\,dF(y)\,dt.
\end{aligned}
$$

En el análisis anterior hemos separado dos casos para el monto de la primera reclamación:

$$(0, u+ct) \cup (u+ct, \infty).$$

Cuando el monto de esta primera reclamación excede el capital $u+ct$, hay ruina y por lo tanto la probabilidad que no ruina es cero. Cuando la reclamación es menor o igual a $u+ct$, no hay ruina y por lo tanto la probabilidad de no ruina en $(0,\infty)$ se reduce a la probabilidad de no ruina desde el tiempo t en adelante pero con capital inicial $u+ct-y$, esto es $\overline{\psi}(u+ct-y)$. Haciendo el cambio de variable $s(t)=u+ct$ en la última ecuación se obtiene

$$\overline{\psi}(u)=\frac{\lambda}{c}\,e^{\lambda u/c}\int_u^\infty e^{-\lambda s/c}\int_0^s \overline{\psi}(s-y)\,dF(y)\,ds.$$

A partir de esta fórmula se puede verificar que la función $u\mapsto\overline{\psi}(u)$ es diferenciable. Así, derivando esta expresión se encuentra el resultado del primer inciso. Demostraremos a continuación el segundo resultado. Integrando la ecuación diferencial del primer inciso entre 0 y u se obtiene

$$\begin{aligned}
\overline{\psi}(u)-\overline{\psi}(0) &= \frac{\lambda}{c}\left(\int_0^u \overline{\psi}(x)dx-\int_0^u\int_0^x \overline{\psi}(x-y)\,dF(y)\,dx\right)\\
&= \frac{\lambda}{c}\left(\int_0^u \overline{\psi}(x)dx-\int_0^u\int_y^u \overline{\psi}(x-y)\,dx\,dF(y)\right)\\
&= \frac{\lambda}{c}\left(\int_0^u \overline{\psi}(x)dx-\int_0^u\int_0^{u-y} \overline{\psi}(x)\,dx\,dF(y)\right)\\
&= \frac{\lambda}{c}\left(\int_0^u \overline{\psi}(x)dx-\int_0^u\int_0^{u-x} \overline{\psi}(x)\,dF(y)\,dx\right)\\
&= \frac{\lambda}{c}\int_0^u \overline{\psi}(x)\,\overline{F}(u-x)\,dx.
\end{aligned}$$

Haciendo el cambio de variable $y=u-x$ se obtiene

$$\begin{aligned}
\overline{\psi}(u)-\overline{\psi}(0) &= \frac{\lambda}{c}\int_0^u \overline{\psi}(u-y)\,\overline{F}(y)\,dy \qquad (8.3)\\
&= \frac{\lambda}{c}\int_0^\infty \overline{\psi}(u-y)\,\overline{F}(y)\,1_{[0,u]}(y)\,dy.
\end{aligned}$$

El siguiente paso es hacer u tender a infinito. En tal caso, $\overline{\psi}(u)$ tiende a uno. Además el integrando que aparece en el lado derecho es una función monótona creciente en u y cuyo límite es la función integrable $\overline{F}(x)$. Entonces por el teorema de convergencia monótona se obtiene

$$1-\overline{\psi}(0)=\frac{\lambda}{c}\int_0^\infty \overline{F}(y)\,dy=\frac{\lambda\mu}{c}.$$

Por lo tanto,

$$\psi(0) = 1 - \overline{\psi}(0) = \frac{\lambda\mu}{c}. \tag{8.4}$$

De esta forma se obtiene el segundo resultado. Finalmente de (8.3) y (8.4) se sigue que

$$\begin{aligned}\psi(u) &= \frac{\lambda}{c}\left(\mu - \int_0^u \overline{\psi}(u-y)\overline{F}(y)\,dy\right)\\ &= \frac{\lambda}{c}\left(\int_u^\infty \overline{F}(y)dy + \int_0^u \psi(u-y)\overline{F}(y)\,dy\right).\end{aligned}$$

■

Observe que la última expresión corresponde a una ecuación íntegro diferencial para la probabilidad de ruina. En general no es fácil resolver este tipo de ecuaciones, sin embargo, cuando las reclamaciones tienen distribución exponencial la ecuación es soluble como se muestra en un ejemplo más adelante. Para resolver la ecuación observamos primero que mediante un cambio de variable en la integral de la primera fórmula de la Proposición 8.1 se obtiene

$$\frac{d}{du}\,\overline{\psi}(u) = \frac{\lambda}{c}\left(\overline{\psi}(u) - \int_0^u \overline{\psi}(y)\,f(u-y)\,dy\right). \tag{8.5}$$

En el texto de Rolski *et al.* [32] se pueden encontrar los detalles de la demostración de la diferenciabilidad de la función $u \mapsto \overline{\psi}(u)$, por simplicidad hemos supuesto tal propiedad en nuestros argumentos. Por otro lado, en el texto de Embrechts *et al* [14] puede encontrarse una argumentación más formal sobre la propiedad de renovación del proceso de riesgo en los tiempos aleatorios en los que ocurre una reclamación. Esta propiedad fue utilizada en la demostración anterior y nos ayudó a encontrar una ecuación íntegro diferencial para la probabilidad de ruina.

Ejemplo 8.1 (Reclamaciones exponenciales) *Encontraremos la probabilidad de ruina cuando las reclamaciones son exponenciales. Éste es uno de los pocos modelos para los cuales tal probabilidad puede encontrarse de manera explícita. Consideremos entonces el modelo de Cramér-Lundberg en donde las reclamaciones tienen distribución exp*(α) *y cuya esperanza es*

$\mu = 1/\alpha$. *Por la ecuación (8.5),*

$$\overline{\psi}\,'(u) = \frac{\lambda}{c}\left(\overline{\psi}(u) - e^{-\alpha u}\int_0^u \overline{\psi}(y)\,\alpha e^{\alpha y}\,dy\right).$$

Derivando esta expresión y sustituyéndola nuevamente en la ecuación encontrada se obtiene

$$\overline{\psi}\,''(u) = \left(\frac{\lambda}{c} - \alpha\right)\overline{\psi}\,'(u),$$

cuya solución es $\overline{\psi}(u) = a + be^{-(\alpha-\lambda/c)u}$, *en donde a y b son constantes. Usando las condiciones $\psi(0) = \lambda/(\alpha c)$ y $\psi(\infty) = 0$ se encuentra que $a = 1$ y $b = -\lambda/(\alpha c)$. Por lo tanto,*

$$\psi(u) = \frac{\lambda}{\alpha c}\,e^{-(\alpha-\lambda/c)u}.$$

La gráfica de esta función se encuentra en la Figura 8.2. Observe que debido a la condición de ganancia neta, el exponente $-(\alpha-\lambda/c)$ es negativo, y por lo tanto la probabilidad de ruina decae a cero exponencialmente cuando el capital inicial u crece a infinito.

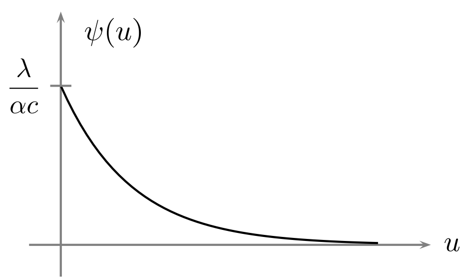


Figura 8.2

Demostraremos a continuación que la probabilidad de ruina se anula cuando el capital inicial es infinito. La técnica de la demostración es similar a la presentada en el modelo de riesgo a tiempo discreto.

Proposición 8.2 *Para el modelo de de Cramér-Lundberg y bajo la condición de ganancia neta,*

$$\lim_{u\to\infty} \psi(u) = 0.$$

Demostración. Por la ley fuerte de los grandes números y la condición de ganancia neta tenemos que

$$\begin{aligned}
\lim_{t\to\infty} \frac{1}{t} C_t &= \lim_{t\to\infty} \frac{1}{t}(u + ct - \sum_{j=1}^{N_t} Y_j) \\
&= c - \lim_{t\to\infty} \frac{1}{t} \sum_{j=1}^{N_t} Y_j \\
&= c - (\lim_{t\to\infty} \frac{N_t}{t})(\lim_{t\to\infty} \frac{1}{N_t} \sum_{j=1}^{N_t} Y_j) \\
&= c - \lambda\mu > 0.
\end{aligned}$$

Para el primer límite hemos usado uno de los resultados que establece la velocidad de crecimiento del proceso de Poisson, esto es, que N_t/t converge casi seguramente al parámetro λ cuando $t \to \infty$. Para el segundo límite hemos usado una extensión del teorema central del límite. De esta forma concluimos nuevamente que la v.a. C_t diverge a infinito casi seguramente cuando $t \to \infty$. En consecuencia, la variable $\inf_{t\geqslant 0} C_t$ está acotada por abajo casi seguramente. Por lo tanto, tomando un capital inicial u suficientemente grande, la cota inferior de $\inf_{t\geqslant 0} C_t$ puede hacerse igual a cero, es decir,

$$\inf_{t\geqslant 0} C_t \geqslant 0.$$

Esto quiere decir que $\psi(u) = 0$ cuando $u \to \infty$. ■

8.3. Probabilidad de ruina con horizonte finito

Dado un valor $x > 0$ fijo, la probabilidad de ruina en el intervalo $(0, x]$ o también llamada probabilidad de ruina con horizonte finito es

$$\psi(u, x) = P(\tau \leqslant x \,|\, C_0 = u),$$

y corresponde a la función de distribución del tiempo de ruina. Es claro que cuando $0 < x_1 \leqslant x_2$, se cumplen las desigualdades

$$\psi(u, x_1) \leqslant \psi(u, x_2) \leqslant \psi(u).$$

Además, cuando el horizonte x crece a infinito, en el límite la probabilidad con horizonte finito converge a la probabilidad con horizonte infinito, es decir,

$$\psi(u) = \lim_{x \to \infty} \psi(u, x).$$

Mientras que cuando el capital inicial es infinito y como consecuencia del resultado $\lim_{u\to\infty} \psi(u) = 0$, la probabilidad de ruina con horizonte finito también es cero, es decir, para cualquier $x > 0$,

$$\lim_{u \to \infty} \psi(u, x) \leqslant \lim_{u \to \infty} \psi(u) = 0.$$

Nuevamente, usaremos la siguiente nomenclatura.

Notación	$\overline{\psi}(u, x) := 1 - \psi(u, x)$

Encontraremos a continuación una ecuación integral para la probabilidad de ruina con horizonte finito. El análisis es más elaborado que en el caso con horizonte infinito. A la fórmula explícita para $\overline{\psi}(0, x)$ que aparece en la siguiente proposición y a la ecuación integral para $\overline{\psi}(u, x)$ se les conoce como fórmulas de Seal.

Proposición 8.3 (Fórmulas de Seal) *Considere el proceso de riesgo de Cramér-Lundberg* $C_t = u+ct-S(t)$, *en donde* $S(t) = \sum_{j=1}^{N_t} Y_j$. *Suponga que las reclamaciones tienen distribución absolutamente continua con función de densidad* $f(y)$ *y defina la función*

$$\tilde{f}_{S(t)}(y) = e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^n}{n!} f^{*n}(y).$$

Entonces

1. $F_{S(t)}(x) = e^{-\lambda t} + \int_0^x \tilde{f}_{S(t)}(y)\, dy, \quad x \geqslant 0.$

2. $\overline{\psi}(0,x) = \dfrac{1}{cx} \int_0^{cx} F_{S(x)}(y)\, dy.$

3. $\dfrac{\partial}{\partial u}\overline{\psi}(u,x) = \dfrac{\lambda}{c}\left(\overline{\psi}(u,x) - \int_0^u \overline{\psi}(u-y,x)\, dF(y) + \dfrac{1}{\lambda}\dfrac{\partial}{\partial x}\overline{\psi}(u,x)\right).$

4. $\overline{\psi}(u,x) = F_{S(x)}(u+cx) - c\int_0^x \overline{\psi}(0,x-y)\, \tilde{f}_{S(y)}(u+cy)\, dy.$

Demostración. La primera identidad se obtiene condicionando sobre el número de reclamaciones: para cualquier $x \geqslant 0$,

$$\begin{aligned}
F_{S(t)}(x) &= P(S(t) \leqslant x) \\
&= \sum_{n=0}^{\infty} P(S(t) \leqslant x \mid N_t = n)\, P(N_t = n) \\
&= e^{-\lambda t} + \sum_{n=1}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \int_0^x f^{*n}(y)\, dy.
\end{aligned}$$

Para la segunda identidad, suponiendo que el capital inicial u es cero, tenemos que

$$\begin{aligned}\overline{\psi}(0,x) &= P(\tau > x) \\ &= P(\bigcap_{t\leqslant x}(C_t \geqslant 0)\,) \\ &= P(\bigcap_{t\leqslant x}(S(t) \leqslant ct)\,) \\ &= \int_0^\infty P(\bigcap_{t\leqslant x}(S(t) \leqslant ct)\,|\,S(x)=y)\,f_{S(x)}(y)\,dy\end{aligned}$$

Puede demostrarse que

$$P(\bigcap_{t\leqslant x}(S(t) \leqslant ct)\,|\,S(x)=y) = \frac{1}{cx}\,(cx-y)_+.$$

Recordemos la notación

$$x_+ = \begin{cases} x & \text{si } x \geqslant 0, \\ 0 & \text{si } x < 0. \end{cases}$$

Por lo tanto,

$$\begin{aligned}\overline{\psi}(0,x) &= \frac{1}{cx}\int_0^\infty (cx-y)_+\,f_{S(x)}(y)\,dy \\ &= \frac{1}{cx}\,E(cx-S(x))_+ \\ &= \frac{1}{cx}\int_0^\infty y\,dF_{cx-S(x)}(y)\,dy \\ &= \frac{1}{cx}\int_0^{cx} y\,dF_{cx-S(x)}(y)\,dy.\end{aligned}$$

Aplicando integración por partes,

$$\begin{aligned}\overline{\psi}(0,x) &= \frac{1}{cx}\left(y\,F_{cx-S(x)}(y)\Big|_0^{cx} - \int_0^{cx} F_{cx-S(x)}(y)\,dy\right) \\ &= \frac{1}{cx}\left(cx - \int_0^{cx} P(S(x)>v)\,dv\right) \\ &= \frac{1}{cx}\int_0^{cx} F_{S(x)}(v)\,dv.\end{aligned}$$

Esto demuestra la segunda identidad. Para la tercera identidad usaremos análisis del primer paso condicionando sobre el monto de la primera reclamación Y_1 y el momento T_1 en el que esta reclamación ocurre. Usaremos además el hecho de que T_1 tiene distribución $\exp(\lambda)$.

$$\begin{aligned}
\overline{\psi}(u,x) &= P(\text{ “No ruina en } (0,x]\text{” } \,|\, C_0 = u) \\
&= \int_0^\infty \int_0^\infty P(\text{ “No ruina en } (0,x]\text{” } \,|\, Y_1 = y, T_1 = t)\, dF(y)\, f_{T_1}(t)\, dt \\
&= \int_0^x \int_0^\infty P(\text{ “No ruina en } (0,x]\text{” } \,|\, Y_1 = y, T_1 = t)\, dF(y)\, f_{T_1}(t)\, dt \\
&\quad + \int_x^\infty \int_0^\infty P(\text{ “No ruina en } (0,x]\text{” } \,|\, Y_1 = y, T_1 = t)\, dF(y)\, f_{T_1}(t)\, dt.
\end{aligned}$$

Observe que hemos separado dos casos: uno cuando la primera reclamación ocurre al tiempo t dentro del intervalo $(0,x]$ y el otro cuando ocurre después de x. En el primer caso la probabilidad del evento de interés es distinta de cero cuando la reclamación es menor o igual a $u+ct$. En el segundo caso la probabilidad del evento es uno. Por lo tanto,

$$\begin{aligned}
\overline{\psi}(u,x) &= \int_0^x \int_0^{u+ct} P(\text{ “No ruina en } (0,x]\text{” } \,|\, Y_1 = y, T_1 = t)\, dF(y)\, f_{T_1}(t)\, dt \\
&\quad + \int_x^\infty \int_0^\infty dF(y)\, f_{T_1}(t)\, dt \\
&= \int_0^x \lambda e^{-\lambda t} \int_0^{u+ct} \overline{\psi}(u+ct-y, x-t)\, dF(y)\, dt + P(T_1 > x).
\end{aligned}$$

Haciendo el cambio de variable $s(t) = u + ct$ se obtiene

$$\overline{\psi}(u,x) = e^{\lambda u/c} \int_u^{u+cx} \lambda e^{-\lambda s/c} \int_0^s \overline{\psi}(s-y, x-(s-u)/c)\, dF(y)\, \frac{1}{c}\, ds + e^{-\lambda x}.$$

Derivando esta expresión respecto de u y respecto de x puede verificarse el cumplimiento de la ecuación integro diferencial

$$\frac{\partial}{\partial u}\,\overline{\psi}(u,x) = \frac{\lambda}{c}\left(\overline{\psi}(u,x) - \int_0^u \overline{\psi}(u-y,x)\, dF(y) + \frac{1}{\lambda}\frac{\partial}{\partial x}\,\overline{\psi}(u,x)\right). \quad (8.6)$$

Para encontrar la cuarta identidad reescribiremos esta ecuación en términos de la transformada de Laplace. Recordemos que la expresión $L_{\overline{\psi}}(s,x)$ denotará la transformada de Laplace de la función $u \mapsto \overline{\psi}(u,x)$, es decir,

$$L_{\overline{\psi}}(s,x) = \int_0^\infty e^{-su}\,\overline{\psi}(u,x)\,du.$$

Así, calculando la transformada de Laplace término a término de la ecuación diferencial (8.6) se obtiene

$$s\,L_{\overline{\psi}}(s,x) - \overline{\psi}(0,x) = \frac{\lambda}{c}\left(L_{\overline{\psi}}(s,x) - L_{\overline{\psi}}(s,x)\,L_f(s) + \frac{1}{\lambda}\frac{\partial}{\partial x}\,L_{\overline{\psi}}(s,x)\right).$$

O bien,

$$\frac{\partial}{\partial x}\,L_{\overline{\psi}}(s,x) = L_{\overline{\psi}}(s,x)\;(cs + \lambda(L_f(s)-1)) - c\overline{\psi}(0,x).$$

De esta forma la derivada respecto de la variable u en la ecuación diferencial parcial (8.6) ha sido absorbida por la transformada de Laplace y hemos obtenido una ecuación diferencial ordinaria en la variable x. La solución general de esta ecuación diferencial es

$$L_{\overline{\psi}}(s,x) = \left(a - \int_0^x \overline{\psi}(0,y)\,c\,e^{-(cs+\lambda(L_f(s)-1))y}\,dy\right)\,e^{(cs+\lambda(L_f(s)-1))x}, \quad (8.7)$$

en donde a es una constante. Evaluando en $x = 0$ se obtiene que esta constante es

$$a = L_{\overline{\psi}}(s,0) = \int_0^\infty e^{-su}\,\overline{\psi}(u,0)\,du = \int_0^\infty e^{-su}\,du = \frac{1}{s}.$$

Así, la ecuación (8.7) adquiere la forma

$$L_{\overline{\psi}}(s,x) = \frac{1}{s}\,e^{(cs+\lambda(L_f(s)-1))x} - \int_0^x \overline{\psi}(0,y)\,c\,e^{(cs+\lambda(L_f(s)-1))(x-y)}\,dy. \quad (8.8)$$

Demostraremos que los dos términos del lado derecho son también una transformada de Laplace. Después de algunos cálculos pueden comprobarse los siguientes resultados:

a) $L[F_{S(x)}(u)](s) = \dfrac{1}{s}\, e^{\lambda\,(L_f(s)-1)x}.$

b) $L[F_{S(x)}(u+cx)](s) = \dfrac{1}{s}\, e^{(cs+\lambda\,(L_f(s)-1))x}.$

c) $L[\tilde{f}_{S(x)}(u+cx)](s) = e^{(cs+\lambda\,(L_f(s)-1))x}.$

Así, el inciso (b) muestra que el primer sumando de (8.8) es la transformada de Laplace de $u \mapsto F_{S(x)}(u+cx)$, y por el inciso (c) el segundo sumando es

$$\begin{aligned}
&\int_0^x \overline{\psi}(0,y)\, c \left(L[\tilde{f}_{S(x-y)}(u+c(x-y))](s) \right) dy \\
&= \int_0^x \overline{\psi}(0,y)\, c \left(\int_0^\infty e^{-sv}\, \tilde{f}_{S(x-y)}(u+c(x-y))\, du \right) dy \\
&= \int_0^\infty e^{-sv}\, c \int_0^x \overline{\psi}(0,x-z)\ \tilde{f}_{S(z)}(u+cz)\, dz\, du.
\end{aligned}$$

Por lo tanto, cada término de la ecuación (8.8) es una transformada de Laplace. Por la propiedad de unicidad, las funciones originales deben coincidir, y así es como se obtiene la última ecuación de la proposición. ■

Observe que la función $x \mapsto \overline{\psi}(u,x)$ es monótona decreciente y tiene como límite la probabilidad $\overline{\psi}(u)$ cuando $x \to \infty$. Así, suponiendo válido el intercambio de este límite con la derivada respecto de x, la ecuación diferencial encontrada para $\overline{\psi}(u,x)$ se reduce a la ecuación para $\overline{\psi}(u)$.

8.4. Severidad de la ruina

Consideremos nuevamente el proceso de riesgo $\{C_t : t \geqslant 0\}$ con capital inicial u. En esta sección nos convendrá escribir la variable C_t como $C(t)$. Supongamos que estamos en la situación cuando el tiempo de ruina τ es finito. En este caso podemos considerar las siguientes variables aleatorias

$$\begin{aligned}
X &:= C(\tau-), \\
Z &:= -C(\tau).
\end{aligned}$$

Estas cantidades representan, respectivamente, el valor del proceso de riesgo un instante antes de la ruina y el valor del proceso justo en el momento de la ruina. Las variables X y Z se muestran gráficamente en la Figura 8.3.

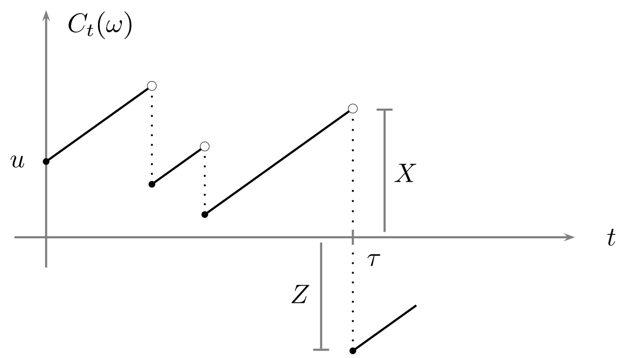


Figura 8.3

Como antes, a la variable Z le llamaremos severidad de la ruina. El problema que se plantea ahora es el de encontrar la probabilidad conjunta

$$\varphi(u, x, z) = P(\tau < \infty, X > x, Z > z),$$

para cualesquiera $x \geqslant 0$, $z \geqslant 0$. Es claro que esta probabilidad es una versión más elaborada que la probabilidad de ruina $\psi(u) = P(\tau < \infty)$ y que se cumplen las siguientes relaciones:

$$\begin{aligned}
\varphi(u, x, z) &\leqslant \varphi(u, 0, z), \\
\varphi(u, x, z) &\leqslant \varphi(u, x, 0), \\
\varphi(u, x, z) &\leqslant \psi(u), \\
\varphi(u, 0, 0) &= \psi(u).
\end{aligned}$$

Como es de esperarse, no se cuenta con una fórmula explícita para $\varphi(u, x, z)$, pero encontraremos una ecuación integral para esta probabilidad. La demostración de la ecuación integral para $\varphi(u, x, z)$ sigue la misma técnica que la que se presentó para el caso de la probabilidad de ruina, es decir, condicionaremos sobre el monto y el momento de la primera reclamación. Observemos que la probabilidad $\varphi(u, x, z)$ se anula cuando el capital inicial crece a infinito, es decir

$$\lim_{u \to \infty} \varphi(u, x, z) \leqslant \lim_{u \to \infty} \psi(u) = 0.$$

Proposición 8.4 *Suponga que la función de distribución $F(y)$ de las reclamaciones en el modelo de Cramér-Lundberg es continua. Entonces, para $x \geqslant 0$, y $z \geqslant 0$,*

1. $$\frac{d}{du}\varphi(u,x,z) = \frac{\lambda}{c}\left(\varphi(u,x,z) - \int_0^u \varphi(u-y,x,z)\,dF(y) - 1_{(u>x)}\,\overline{F}(u+z)\right).$$

2. $$\varphi(0,x,z) = \frac{\lambda}{c}\int_{x+z}^{\infty}\overline{F}(y)\,dy.$$

3. $$\varphi(u,x,z) = \frac{\lambda}{c}\left(\int_{u+z}^{\infty}\overline{F}(y)\,dy + \int_0^u \varphi(u-y,x,z)\overline{F}(y)\,dy + 1_{(u<x)}\int_{x+z}^{u+z}\overline{F}(y)\,dy\right).$$

Demostración. Usaremos nuevamente análisis del primer paso condicionando sobre el monto y momento de la primera reclamación. Dado un valor t para el momento de la primera reclamación, separaremos tres conjuntos para el monto de dicha reclamación:

$$(0,u+ct)\cup(u+ct,u+ct+z)\cup(u+ct+z,\infty).$$

En el primer caso no hay ruina y la probabilidad buscada se reduce a la probabilidad del mismo evento pero ahora con capital inicial $u+ct-y$. En el segundo caso hay ruina pero no es severa en el sentido de que no se cumple que $Z>z$ y por lo tanto la probabilidad $\varphi(u,x,z)$ es cero. En el tercer caso hay ruina y se cumple la condición $Z>z$, así es que la probabilidad buscada se reduce a $1_{(u+ct>x)}$. Así, tenemos que

$$\begin{aligned}
\varphi(u,x,z) &= \int_0^\infty \int_0^\infty P(\tau<\infty, X>x, Z>z \,|\, T_1=t, Y_1=y)\, dF(y)\, dF_{T_1}(t) \\
&= \int_0^\infty \lambda e^{-\lambda t} \int_0^{u+ct} \varphi(u+ct-y,x,z)\, dF(y)\, dt \\
&\quad + \int_0^\infty \lambda e^{-\lambda t} \int_{u+ct+z}^\infty 1_{(u+ct>x)}\, dF(y)\, dt.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\varphi(u,x,z) &= \int_0^\infty \lambda e^{-\lambda t} \int_0^{u+ct} \varphi(u+ct-y,x,z)\, dF(y)\, dt \\
&\quad + \int_0^\infty \lambda e^{-\lambda t}\, 1_{(u+ct>x)}\, \overline{F}(u+ct+z)\, dt.
\end{aligned}$$

Efectuando el cambio de variable $s(t) = u + ct$ se obtiene

$$\begin{aligned}
\varphi(u,x,y) &= \frac{1}{c}\left(\int_u^\infty \lambda e^{-\lambda(s-u)/c} \int_0^s \varphi(s-y,x,z)\, dF(y)\, ds \right.\\
&\quad \left. + \int_u^\infty \lambda e^{-\lambda(s-u)/c}\, 1_{(s>x)}\, \overline{F}(s+z)\, ds \right).
\end{aligned}$$

Derivando esta expresión respecto de u se encuentra el resultado del primer inciso. Escribiendo w por u en la ecuación diferencial encontrada e integrando entre 0 y u se obtiene

$$\begin{aligned}
\varphi(u,x,y) - \varphi(0,x,y) &= \frac{\lambda}{c}\left(\int_0^u \varphi(w,x,z)\, dw \right.\\
&\quad - \int_0^u \int_0^w \varphi(w-y,x,z)\, dF(y)\, dw \\
&\quad \left. - \int_0^u 1_{(w>x)} \overline{F}(w+z)\, dw \right),
\end{aligned}$$

en donde la integral doble puede escribirse de la siguiente forma

$$\begin{aligned}
\int_0^u \int_0^w \varphi(w-y,x,z)\,dF(y)\,dw &= \int_0^u \int_y^u \varphi(w-y,x,z)\,dw\,dF(y) \\
&= \int_0^u \int_0^{u-y} \varphi(v,x,z)\,dv\,dF(y) \\
&= \int_0^u \int_0^{u-v} \varphi(v,x,z)\,dF(y)\,dv \\
&= \int_0^u \varphi(v,x,z)F(u-v)\,dv.
\end{aligned}$$

Observe además que

$$\int_0^u 1_{(w>x)}\overline{F}(w+z)\,dw = 1_{(u\geqslant x)} \int_x^u \overline{F}(w+z)\,dw.$$

Por lo tanto,

$$\begin{aligned}
\varphi(u,x,z)-\varphi(0,x,z) &= \frac{\lambda}{c}\Bigg(\int_0^u \varphi(v,x,z)(1-F(u-v))\,dv \\
&\qquad -1_{(u\geqslant x)} \int_x^u \overline{F}(w+z)\,dw\Bigg) \\
&= \frac{\lambda}{c}\Bigg(\int_0^u \varphi(v,x,z)\,\overline{F}(u-v)\,dv \\
&\qquad -1_{(u\geqslant x)} \int_{x+z}^{u+z} \overline{F}(y)\,dy\Bigg) \\
&= \frac{\lambda}{c}\Bigg(\int_0^u \varphi(u-y,x,z)\overline{F}(y)\,dy \\
&\qquad -1_{(u\geqslant x)} \int_{x+z}^{u+z} \overline{F}(y)\,dy\Bigg). \qquad (8.9)
\end{aligned}$$

Haciendo u tender a infinito, $\varphi(u,x,z)\to 0$ y no es difícil demostrar que la primera integral se anula. Por lo tanto se obtiene la identidad

$$\varphi(0,x,z) = \frac{\lambda}{c}\int_{x+z}^{\infty} \overline{F}(y)\,dy. \qquad (8.10)$$

Este el segundo inciso del enunciado. Sustituyendo (8.10) en la ecuación (8.9) y simplificando se obtiene el tercer inciso del enunciado. ■

Observe que tomando $x = 0$ y $z = 0$, la ecuación integral para $\varphi(u, x, z)$ se reduce a la ecuación integral para la probabilidad de ruina con horizonte infinito $\psi(u)$, es decir,

$$\psi(u) = \varphi(u, 0, 0) = \frac{\lambda}{c} \left(\int_u^\infty \overline{F}(y)\, dy + \int_0^u \varphi(u - y, 0, 0)\, \overline{F}(y)\, dy \right).$$

Esta es la tercera ecuación de la Proposición 8.1. Además suponiendo cada caso por separado, $x = 0$ o $y = 0$, se obtienen fórmulas integrales para las probabilidades marginales de X y de Z, considerando siempre tiempo de ruina finito. Es decir, para la variable Z tenemos que

$$\varphi(u, 0, z) = \frac{\lambda}{c} \left(\int_{u+z}^\infty \overline{F}(y)\, dy + \int_0^u \varphi(u - y, 0, z)\, \overline{F}(y)\, dy \right).$$

Para la variable X,

$$\varphi(u, x, 0) = \frac{\lambda}{c} \left(\int_u^\infty \overline{F}(y)\, dy + \int_0^u \varphi(u - y, x, 0)\, \overline{F}(y)\, dy + 1_{(u<x)} \int_x^u \overline{F}(y)\, dy \right).$$

8.5. El coeficiente de ajuste

Como hemos visto en el modelo de riesgo a tiempo discreto en el capítulo anterior, este coeficiente es un número que aparece en el problema de calcular o estimar probabilidades de ruina. En esta sección vamos a definir y estudiar este coeficiente para el caso del modelo de riesgo a tiempo continuo. Existen varias maneras de definirlo, una de ellas, un tanto artificial pero que después justificaremos, es la siguiente: se define primero la función

$$\theta(r) = \lambda(M_Y(r) - 1) - cr,$$

en donde $M_Y(r)$ es la función generadora de Y. Naturalmente esta función está bien definida para valores de r en donde $M_Y(r)$ existe. En tal caso, $\theta(r)$ es diferenciable y se tiene que sus primeras dos derivadas son:

$$\begin{aligned} \theta'(r) &= \lambda M_Y'(r) - c, \\ \theta''(r) &= \lambda M_Y''(r) = \lambda E(Y^2 e^{rY}) > 0. \end{aligned}$$

Por lo tanto, $\theta(r)$ es una función estrictamente convexa tal que $\theta(0) = 0$, y por la condición de ganancia neta, $\theta'(0) = \lambda\mu - c < 0$. Entonces es posible

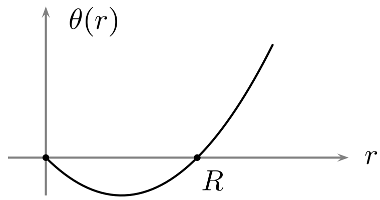


Figura 8.4

que exista un valor $r > 0$ tal que $\theta(r) = 0$. Una gráfica de la función $\theta(r)$ presentando esta situación se muestra en la Figura 8.4.

Definición 8.3 *A la posible solución $r > 0$ de la siguiente ecuación se le llama coeficiente de ajuste o exponente de Lundberg.*

$$\theta(r) = \lambda(M_Y(r) - 1) - cr = 0.$$

A tal solución positiva se le denota por R.

Observe que la existencia del coeficiente de ajuste depende de la distribución de las reclamaciones y los parámetros del modelo. Aquellas distribuciones para las cuales el coeficiente de ajuste existe se les llama distribuciones con cola ligera, y la razón de ello es que la función de densidad decae a cero exponencialmente rápido, asignando probabilidades pequeñas a reclamaciones grandes. Más adelante daremos condiciones para la existencia del coeficiente de ajuste y formalizaremos la interpretación de cola ligera para la distribución de las reclamaciones. Veremos a continuación algunos ejemplos. Demostraremos en particular que en el caso de reclamaciones exponenciales, el coeficiente de ajuste existe y es fácil calcularlo.

Ejemplo 8.2 (Reclamaciones exponenciales) *Suponga que las reclamaciones siguen una distribución exp(α), es decir, la función generadora de* *momentos es* $M_Y(r) = \alpha/(\alpha - r)$*, para* $r < \alpha$*. Entonces*

$$\begin{aligned} \theta(r) &= \lambda\,(M_Y(r) - 1) - cr \\ &= \lambda\,(\frac{\alpha}{\alpha - r} - 1) - cr \\ &= \lambda(\frac{r}{\alpha - r}) - cr \\ &= (\frac{\lambda}{\alpha - r} - c)\,r. \end{aligned}$$

De modo que $\theta(r)$ *es cero cuando* $r = 0$*, o bien cuando* $\lambda/(\alpha - r) - c = 0$*. Despejando* r *de la segunda condición y escribiendo ahora* R *como la variable se obtiene* $R = \alpha - \lambda/c$*. Más aún, por lo encontrado antes en el caso de reclamaciones exponenciales, la probabilidad de ruina puede ahora escribirse de la forma siguiente:*

$$\psi(u) = \frac{\lambda}{\alpha c}\, e^{-(\alpha - \lambda/c)u} = \frac{\lambda}{\alpha c}\, e^{-Ru} < e^{-Ru},$$

en donde la desigualdad es consecuencia de la condición de ganancia neta. Este tipo de cota superior para la probabilidad de ruina, llamada desigualdad de Lundberg, será demostrada más adelante para cualquier distribución de las reclamaciones para la cual el coeficiente de ajuste exista.

Ejemplo 8.3 (Reclamaciones gama) *Suponga que las reclamaciones siguen una distribución gama*(γ, α) *con* $\gamma = 2$*. La función generadora de momentos es* $M_Y(r) = (\alpha/(\alpha - r))^\gamma$*, para* $r < \alpha$*. Por lo tanto la función* $\theta(r)$ *correspondiente es*

$$\theta(r) = \lambda\left[\left(\frac{\alpha}{\alpha - r}\right)^\gamma - 1\right] - cr.$$

La condición $\theta(r) = 0$ *produce la ecuación cuadrática*

$$cr^2 + r(\lambda - 2\alpha c) + (c\alpha^2 - 2\alpha\lambda) = 0,$$

cuyas raíces son

$$r = \frac{2\alpha c - \lambda \pm \sqrt{\lambda^2 + 4\alpha c\lambda}}{2c}.$$

El caso con raíz cuadrada positiva es inválido pues resulta $r > \alpha$, en efecto, usando la condición de ganancia neta, $c > \lambda(2/\alpha)$, se obtiene

$$\begin{aligned}(2\alpha c - \lambda + \sqrt{\lambda^2 + 4\alpha c\lambda})/2c &\geqslant (2\alpha c - \lambda + \sqrt{\lambda^2 + 8\lambda^2})/2c \\ &= \alpha + \lambda/c \\ &> \alpha.\end{aligned}$$

Por lo tanto la raíz con el signo negativo es el coeficiente de ajuste.

Ejemplo 8.4 *Aplicaremos el método de Newton-Raphson para encontrar de manera aproximada el coeficiente de ajuste cuando las reclamaciones siguen una distribución gama(γ, α) con $\gamma = 3$. Véase el Apéndice para una explicación breve sobre este método numérico. La condición $\theta(r) = 0$ produce la siguiente ecuación que puede reducirse a una ecuación cúbica en r,*

$$g(r) = \lambda(\alpha^3 - (\alpha - r)^3) - cr(\alpha - r)^3 = 0.$$

La raíz buscada r es tal que por restricciones de la función generadora de momentos debe satisfacer $0 < r < \alpha$. Tomaremos $\alpha = 3$, $\lambda = 1$, y $c = 2$. Se han escogido estos valores de los parámetros por simplicidad pero al mismo tiempo cuidando que se verifique la condición de ganancia neta $c > \lambda(3/\alpha)$. El esquema general de Newton-Raphson es

$$r_{n+1} = r_n - \frac{g(r_n)}{g'(r_n)}.$$

Más adelante demostraremos que si el coeficiente de ajuste existe, entonces éste se encuentra siempre dentro del intervalo $(0, 2(c - \lambda\mu)/(\lambda\mu_2))$. Véase el enunciado de la Proposición 8.9. Con base en este resultado, tomaremos como condición inicial $r_0 = 2(c - \lambda\mu)/(\lambda\mu_2) = 3/2$. Usando estos datos, la iteración de Newton-Raphson arroja la siguiente sucesión de valores:

n	r_n
0	*1.5*
1	*0.8333333*
2	*0.8405056*
3	*0.8404738*
4	*0.8404738*

Por lo tanto, el coeficiente de ajuste se encuentra muy cercano al valor $R = 0.8404738$.

El siguiente resultado proporciona una forma equivalente de definir el coeficiente de ajuste y permite comprobar su existencia a través de la determinación de la finitud de una integral.

Proposición 8.5 *Sea $F(y)$ la función de distribución de las reclamaciones en el modelo de Cramér-Lundberg. Sea $\overline{F}(y) = 1 - F(y)$. Entonces la ecuación $\theta(r) = \lambda(M_Y(r) - 1) - cr = 0$ tiene una solución $r > 0$ si y sólo si se cumple la identidad*

$$\int_0^\infty e^{ry}\,\overline{F}(y)\,dy = \frac{c}{\lambda}.$$

Demostración. Tenemos que

$$\begin{aligned} 0 &= \theta(r) \\ &= \lambda(M_Y(r) - 1) - cr \\ &= \lambda(\int_0^\infty e^{ry}\,dF(y) - \int_0^\infty dF(y)) - cr \\ &= \lambda\int_0^\infty (e^{ry} - 1)\,dF(y) - cr \\ &= -\lambda\int_0^\infty (e^{ry} - 1)\,d\overline{F}(y) - cr. \end{aligned}$$

Para la última igualdad hemos usado el hecho de que $dF(x) = -d\overline{F}(x)$. Integrando ahora por partes tenemos que

$$0 = -\lambda\left((e^{ry} - 1)\,\overline{F}(y)\Big|_0^\infty - \int_0^\infty re^{ry}\,\overline{F}(y)\,dy \right) - cr.$$

Usando la hipótesis de existencia de la función generadora de momentos de la distribución $F(y)$, puede demostrarse que las evaluaciones indicadas en el primer término del lado derecho se anulan. Por lo tanto,

$$0 = \lambda r \int_0^\infty e^{ry}\,\overline{F}(y)\,dy - cr.$$

Despejando la integral se obtiene el resultado enunciado. ■

Este resultado permite además dar una interpretación de aquellas distribuciones de probabilidad para las cuales el coeficiente de ajuste existe: para que la integral de la proposición anterior sea finita, la cola de la distribución, es decir, $\overline{F}(y)$, debe decaer a cero lo suficientemente rápido para anular el comportamiento creciente de la expresión e^{ry} dentro de la integral. Esto ha originado el término de cola ligera para aquellas distribuciones para las cuales exista el coeficiente de ajuste.

Definición 8.4 (Distribución con cola ligera/pesada) *A una distribución de probabilidad con soporte en el intervalo* $(0,\infty)$ *y para la cual*

a) existe el coeficiente de ajuste se le llama distribución con cola ligera.

b) no existe el coeficiente de ajuste se le llama distribución con cola pesada.

Ejemplo 8.5 *Usaremos el criterio de la Proposición 8.5 para demostrar que para la distribución Burr*(c,α) *no existe el coeficiente de ajuste. En este caso la cola de la distribución está dada por*

$$\overline{F}(y) = \left(\frac{1}{1+y^c}\right)^\alpha.$$

Entonces para cualquier valor de $r > 0$,

$$\int_0^\infty e^{ry}\,\overline{F}(y)\,dy = \int_0^\infty e^{ry}\left(\frac{1}{1+y^c}\right)^\alpha dy \approx \int_0^\infty e^{ry}\,y^{-c\alpha}\,dy = \infty.$$

Por lo tanto la distribución Burr es una distribución con cola pesada. Según el criterio de la Proposición 8.5, para que exista el coeficiente de ajuste, la *cola de la distribución debe decaer a cero lo suficientemente rápido para anular el comportamiento creciente del término* e^{ry} *dentro de la integral. En este ejemplo la cola de la distribución Burr decae a cero en la forma* $y^{-c\alpha}$ *que es insuficiente para hacer que la integral sea finita.*

Una distribución con cola ligera asigna probabilidades muy pequeñas a los valores grandes de la variable aleatoria. Esto puede no ser muy conveniente para modelar algunos riesgos, pues de esa manera se está subestimando la posibilidad de registrar grandes montos en las reclamaciones. A continuación mencionamos algunos ejemplos de distribuciones con cola ligera y pesada.

Ejemplo 8.6 (Distribuciones con cola ligera) *Las siguientes distribuciones tienen como soporte el intervalo* $(0, \infty)$ *y tienen función generadora de momentos.*

a) Distribución exp(α)*, parámetro* $\alpha > 0$*.*

$$\overline{F}(y) = e^{-\alpha y}, \quad y > 0.$$

b) Distribución gama(γ, α)*, parámetros* $\gamma > 0$*,* $\alpha > 0$*.*

$$f(y) = \frac{(\alpha y)^{\gamma-1}}{\Gamma(\gamma)}\,\alpha\, e^{-\alpha y}, \quad y > 0.$$

c) Distribución Weibull(r, α)*, parámetros* $\alpha > 0$*,* $r \geqslant 1$*.*

$$\overline{F}(y) = e^{-(\alpha y)^r}, \quad y > 0.$$

d) Distribución normal truncada (valor absoluto de $N(0,1)$ *).*

$$\overline{F}(y) = 2(1 - \Phi(y)), \quad y > 0.$$

Ejemplo 8.7 (Distribuciones con cola pesada) *Las siguientes distribuciones tienen como soporte el intervalo* $(0, \infty)$ *y no poseen función generadora de momentos.*

a) Distribución lognormal(μ, σ^2)*, parámetros* $\mu \in \mathbb{R}$*,* $\sigma^2 > 0$*.*

$$\overline{F}(y) = 1 - \Phi\left(\frac{\ln y - \mu}{\sigma}\right), \quad y > 0.$$

b) Distribución Pareto(a, b)*, parámetros* $a > 0$*,* $b > 0$*.*

$$\overline{F}(y) = \left(\frac{b}{b + y}\right)^a, \quad y > 0.$$

c) *Distribución Burr*(c,α), *parámetros* $c>0$, $\alpha>0$.

$$\overline{F}(y) = \left(\frac{1}{1+y^c}\right)^{\alpha}, \quad y>0.$$

d) *Distribución Weibull*(r,α), *parámetros* $\alpha>0$, $0<r<1$.

$$\overline{F}(y) = e^{-(\alpha y)^r}, \quad y>0.$$

e) *Distribución loggama*(γ,α), *parámetros* $\gamma>0$, $\alpha>0$.

$$f(y) = \frac{(\alpha \ln y)^{\gamma-1}}{\Gamma(\gamma)}\,\alpha\, y^{-(\alpha+1)}, \quad y>0.$$

El siguiente resultado establece condiciones suficientes bajo las cuales el coeficiente de ajuste existe.

Proposición 8.6 *Suponga que la función generadora de momentos* $M_y(r)$ *de las reclamaciones* Y *en el modelo de Cramér Lundberg es tal que:*

a) $M_Y(r)<\infty$ *para* $r<r^*$ *para algún* $r^*\in(0,\infty]$.

b) $\lim\limits_{r\nearrow r^*} M_Y(r)=\infty$.

Entonces existe el coeficiente de ajuste, es decir, existe un único valor $r>0$ *tal que*

$$\theta(r) = \lambda(M_Y(r)-1)-cr = 0.$$

Demostración. Dadas las propiedades que se observaron que posee la función $\theta(r)$ alrededor de $r=0$, es suficiente demostrar que

$$\lim_{r\nearrow r^*} \theta(r) = \infty.$$

Se tienen dos casos. Supongamos primero que $r^*<\infty$. Entonces la afirmación se obtiene de la segunda hipótesis, pues como $\lambda>0$,

$$\lim_{r\nearrow r^*} \theta(r) = \lim_{r\nearrow r^*} \lambda(M_Y(r)-1)-cr = \infty.$$

Supongamos ahora que $r^* = \infty$. Sea y^* tal que $F_Y(y^*) < 1$. Entonces

$$\begin{aligned} M_Y(r) &= \int_0^\infty e^{ry}\, dF(y) \\ &= \int_0^{y^*} e^{ry}\, dF(y) + \int_{y^*}^\infty e^{ry}\, dF(y) \\ &\geqslant \int_{y^*}^\infty e^{ry}\, dF(y) \\ &= e^{ry^*}\, \overline{F}(y^*). \end{aligned}$$

Por lo tanto,

$$\begin{aligned} \lim_{r\to\infty} \theta(r) &= \lim_{r\to\infty} \lambda(M_Y(r) - 1) - cr \\ &\geqslant \lim_{r\to\infty} \lambda(e^{ry^*}\, \overline{F}(y^*) - 1) - cr \\ &= \infty. \end{aligned}$$

■

No es necesario suponer la existencia del coeficiente de ajuste para observar el comportamiento de cola ligera en una distribución. Demostraremos a continuación que la existencia de la función generadora de momentos garantiza tal comportamiento.

Proposición 8.7 *Sea Y una variable aleatoria positiva con función de distribución $F(y)$ y con función generadora de momentos tal que $M(r) < \infty$ para algún $r > 0$. Entonces para cualquier $y > 0$,*

$$\overline{F}(y) \leqslant M(r)\, e^{-ry}.$$

Demostración. Por la desigualdad de Markov, véase el Apéndice,

$$\begin{aligned}\overline{F}(y) &= P(Y > y) \\ &= P(e^{rY} > e^{ry}) \\ &\leqslant \frac{E(e^{rY})}{e^{ry}} \\ &= M(r)\, e^{-ry}.\end{aligned}$$

■

8.6. Desigualdad de Lundberg

Vamos a demostrar ahora que para aquellas distribuciones de probabilidad que se supongan como distribuciones de las reclamaciones en el modelo de Cramér-Lundberg y para las cuales el coeficiente de ajuste R existe, se cumple la desigualdad

$$\psi(u) \leqslant e^{-Ru}.$$

Existen varias formas de demostrar este resultado. El método que presentaremos hace uso de la teoría de martingalas.

Proposición 8.8 *Sea $\{C_t : t \geqslant 0\}$ el proceso de riesgo de Cramér-Lundberg en donde la distribución de las reclamaciones tienen a $M_Y(t)$ como función generadora de momentos. Sea $\theta(r) = \lambda(M_Y(r) - 1) - cr$. Entonces el siguiente proceso es una martingala:*

$$\{\, e^{-rC_t - \theta(r)t} : t \geqslant 0 \,\}.$$

Demostración. La adaptabilidad del proceso es evidente pues implícitamente estamos usando la filtración natural. Acerca de la integrabilidad tenemos los siguientes cálculos

$$\begin{aligned}
E(e^{-rC_t-\theta(r)t}) &= e^{-\theta(r)t}E(e^{-r(u+ct-\sum_{j=1}^{N_t}Y_j)}) \\
&= e^{-\theta(r)t-r(u+ct)}E(e^{r\sum_{j=1}^{N_t}Y_j}) \\
&= e^{-\theta(r)t-r(u+ct)}e^{\lambda t(M_Y(r)-1)} \\
&< \infty.
\end{aligned}$$

Finalmente tenemos la propiedad de martingala. Para $0 \leqslant s < t$ y por la propiedad de incrementos independientes y estacionarios del proceso de riesgo,

$$\begin{aligned}
E(e^{-rC_t-\theta(r)t} \,|\, \mathscr{F}_s) &= e^{-\theta(r)t}E(e^{-rC_t} \,|\, \mathscr{F}_s) \\
&= e^{-\theta(r)t}E(e^{-r(C_t-C_s)-rC_s} \,|\, \mathscr{F}_s) \\
&= e^{-\theta(r)t-rC_s}E(e^{-r(C_t-C_s)}) \\
&= e^{-\theta(r)t-rC_s}E(e^{-r(c(t-s)-\sum_{j=N_s+1}^{N_t}Y_j)}) \\
&= e^{-\theta(r)t-rC_s-rc(t-s)}E(e^{r\sum_{j=1}^{N_{t-s}}Y_j}) \\
&= e^{-\theta(r)t-rC_s-rc(t-s)}e^{\lambda(t-s)(M_Y(r)-1)} \\
&= e^{-rC_s-\theta(r)s}.
\end{aligned}$$

■

En particular, si el coeficiente de ajuste existe, es decir, si $\theta(R) = 0$, entonces el proceso $\{e^{-RC_t} : t \geqslant 0\}$ es una martingala. Este es el resultado clave para demostrar la cota superior para la probabilidad de ruina.

Teorema 8.1 (Desigualdad de Lundberg) *Suponga que el coeficiente de ajuste R existe para la distribución de las reclamaciones en el modelo de riesgo de Cramér-Lundberg. Entonces*

$$\psi(u) \leqslant e^{-Ru}.$$

Demostración. Sea τ el tiempo de paro correspondiente al tiempo de ruina. Como el proceso $\{e^{-RC_t} : t \geqslant 0\}$ es una martingala, se tiene que el proceso detenido $\{e^{-RC_{t\wedge\tau}} : t \geqslant 0\}$ también es una martingala. Ambos procesos inician en el valor e^{-Ru}. Por lo tanto,

$$
\begin{aligned}
e^{-Ru} &= e^{-RC_0} \\
&= E(e^{-RC_{t\wedge\tau}}) \\
&= E(e^{-RC_{t\wedge\tau}} \,|\, \tau \leqslant t)\, P(\tau \leqslant t) \\
&\quad +E(e^{-RC_{t\wedge\tau}} \,|\, \tau > t)\, P(\tau > t) \\
&\geqslant E(e^{-RC_{t\wedge\tau}} \,|\, \tau \leqslant t)\, P(\tau \leqslant t) \\
&= E(e^{-RC_{\tau}} \,|\, \tau \leqslant t)\, P(\tau \leqslant t) \\
&= \left(\int e^{-RC_\tau} 1_{(\tau \leqslant t)}\, dP \right) P(\tau \leqslant t).
\end{aligned}
$$

Haciendo $t \to \infty$ monótonamente, el evento $(\tau \leqslant t)$ converge crecientemente al evento $(\tau < \infty)$. Usando el teorema de convergencia monótona puede demostrarse que

$$
\begin{aligned}
e^{-Ru} &\geqslant E(e^{-RC_\tau} \,|\, \tau < \infty)\, P(\tau < \infty) \\
&\geqslant E(\,1 \,|\, \tau < \infty)\, P(\tau < \infty) \\
&= P(\tau < \infty) \\
&= \psi(u).
\end{aligned}
$$

■

Con ayuda de la desigualdad de Lundberg confirmamos que la probabilidad de ruina se anula cuando el capital inicial crece a infinito, al menos en situaciones cuando tal desigualdad es válida.
Como hemos visto, el coeficiente de ajuste no siempre existe, y aun cuando conozcamos su existencia no siempre es fácil calcularlo. El siguiente resultado proporciona algunas cotas para el valor de este coeficiente, suponiendo su existencia.

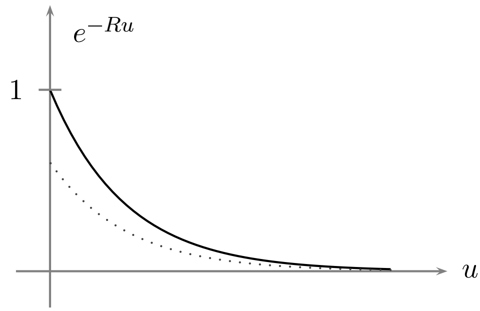


Figura 8.5: Cota superior para la probabilidad de ruina.

Proposición 8.9 *Si el coeficiente de ajuste R existe para la distribución de probabilidad de las reclamaciones en el modelo de Cramér-Lundberg, entonces*

$$\frac{1}{M}\ln\left(\frac{c}{\lambda\mu}\right) < R < \frac{2(c-\lambda\mu)}{\lambda\mu_2},$$

en donde la primera desigualdad es válida bajo la hipótesis adicional de que $Y \leqslant M$ c.s., para alguna constante $M > 0$.

Demostración. Demostraremos primero la cota superior. Consideremos nuevamente la función $\theta(r) = \lambda(M_Y(r) - 1) - cr$, para $r \geqslant 0$. Sabemos que $\theta(0) = 0$. Derivando esta función dos veces se obtiene

$$\begin{aligned} \theta'(r) &= \lambda M_Y'(r) - c, \\ \theta''(r) &= \lambda E(Y^2 e^{rY}) \\ &> \lambda E(Y^2) \\ &= \lambda\mu_2. \end{aligned}$$

En donde $\theta'(0) = \lambda\mu - c$. Por el teorema fundamental del cálculo,

$$\begin{aligned} \theta'(r) &= \theta'(0) + \int_0^r \theta''(s)ds \\ &> (\lambda\mu - c) + \lambda\mu_2 r. \end{aligned}$$

Por lo tanto,

$$\begin{aligned}\theta(r) &= \theta(0) + \int_0^t \theta'(s)ds \\ &> (\lambda\mu - c)r + \frac{1}{2}\lambda\mu_2 r^2.\end{aligned}$$

Evaluando la última desigualdad en R se obtiene

$$\begin{aligned}0 &> (\lambda\mu - c)R + \frac{1}{2}\lambda\mu_2 R^2 \\ &= [\,(\lambda\mu - c) + \frac{1}{2}\lambda\mu_2 R\,]\,R.\end{aligned}$$

Como $R > 0$, la expresión en el paréntesis debe ser negativa. Despejando de allí R se obtiene la cota superior anunciada. Demostraremos ahora la cota inferior. Supongamos que $Y \leqslant M$ c.s. y defina la función

$$h(x) = \frac{x}{M}(e^{RM} - 1) - (e^{Rx} - 1).$$

Entonces $h''(x) = -R^2 e^{Rx} < 0$. Por lo tanto $h(x)$ es una función cóncava con $h(0) = h(M) = 0$. Esto quiere decir que $h(x) > 0$ para $0 < x < M$, es decir,

$$\frac{x}{M}\,(e^{RM} - 1) - (e^{Rx} - 1) > 0,$$

o bien,

$$(e^{Rx} - 1) \leqslant \frac{x}{M}\,(e^{RM} - 1). \tag{8.11}$$

Sea $g(x) = xe^x - e^x + 1$. Entonces $g'(x) = xe^x > 0$. Por lo tanto $g(x)$ es una función creciente para $x > 0$, es decir, $g(x) > g(0) = 0$, para $x > 0$. Por lo tanto, $xe^x - e^x + 1 > 0$, para $x > 0$. En particular, evaluando en $x = RM$ se obtiene $RMe^{RM} - e^{RM} + 1 > 0$. Entonces

$$\frac{e^{RM} - 1}{RM} < e^{RM}. \tag{8.12}$$

Por otro lado, usando (8.11),

$$\begin{aligned} M_Y(R) - 1 &= \int_0^M (e^{Rx} - 1)\, dF(x) \\ &\leqslant \int_0^M \frac{x}{M}\,(e^{RM} - 1)\, dF(x) \\ &= \frac{1}{M}(e^{RM} - 1) \int_0^M x\, dF(x) \\ &= \frac{\mu}{M}(e^{RM} - 1). \end{aligned} \tag{8.13}$$

Así, usando (8.13) y luego (8.12),

$$\begin{aligned} 0 &= \lambda(M_Y(R) - 1) - cR \\ &\leqslant \frac{\lambda\mu}{M}(e^{RM} - 1) - cR \\ &< \lambda\mu R e^{RM} - cR \\ &= (\lambda\mu e^{RM} - c)R. \end{aligned}$$

Por lo tanto $\lambda\mu e^{RM} - c > 0$. Despejando finalmente R de esta desigualdad se obtiene la cota inferior buscada. ■

Ejemplo 8.8 (Reclamaciones exponenciales) *Consideremos nuevamente el caso cuando las reclamaciones son exponenciales de parámetro α. Sabemos que en este caso el coeficiente de ajuste existe y está dado por*

$$R = \alpha - \frac{\lambda}{c}.$$

Usando la condición de ganancia neta puede verificarse que efectivamente se cumple la cota superior para R demostrada en la Proposición 8.9, es decir,

$$R = \alpha - \frac{\lambda}{c} < \frac{2(c - \lambda\mu)}{\lambda\mu_2}.$$

Es interesante notar que la cota superior para el coeficiente de ajuste que aparece en la Proposición 8.9 no involucra hipótesis adicionales, de modo que cuando el coeficiente de ajuste R existe, éste se encuentra siempre dentro del intervalo $(0, 2(c - \lambda\mu)/\lambda\mu_2)$. Este resultado fue usado en el Ejemplo 8.4. Observe además que cuando las reclamaciones están acotadas superiormente por una constante positiva M, podemos usar la cota inferior para R para encontrar una cota superior para la probabilidad de ruina sin conocer el valor exacto del coeficiente de ajuste, pues

$$\begin{aligned} \psi(u) &\leqslant e^{-Ru} \\ &< e^{-\frac{u}{M}\ln(c/\lambda\mu)} \\ &= e^{\ln(\lambda\mu/c)^{u/M}} \\ &= \left(\frac{\lambda\mu}{c}\right)^{u/M}. \end{aligned}$$

Esta cota superior converge a 1 cuando $M \to \infty$ y por lo tanto se vuelve no informativa. Cuando hacemos $u \to \infty$, se corrobora que la probabilidad de ruina se anula cuando el capital inicial es infinito.

8.7. Aproximación de De Vylder

Consideremos nuevamente el modelo de riesgo de Cramér-Lundberg en donde las reclamaciones tienen distribución arbitraria. De Vylder propone aproximar la probabilidad de ruina de este modelo mediante la probabilidad de ruina de un modelo con reclamaciones exponenciales. Se define entonces el proceso

$$\tilde{C}_t = u + \tilde{c}\, t - \sum_{j=1}^{\tilde{N}_t} \tilde{Y}_j, \tag{8.14}$$

en donde $\tilde{c}$ es una nueva tasa de ingreso por primas, $\{\tilde{N}_t : t \geqslant 0\}$ es un proceso de Poisson de parámetro $\tilde{\lambda}$, y $\tilde{Y}_j$ son variables aleatorias con distribución $\exp(\tilde{\alpha})$. Para lograr que los modelos tengan algún parecido se igualan los tres primeros momentos de los dos procesos y se encuentran los valores de los parámetros $\tilde{c}$, $\tilde{\lambda}$ y $\tilde{\alpha}$ en términos de los parámetros del riesgo original. Así, la probabilidad de ruina en el modelo original se aproxima por la probabilidad de ruina en el modelo con reclamaciones exponenciales.

Proposición 8.10 (Aproximación de De Vylder) *Suponga que las reclamaciones en el modelo de riesgo de Cramér-Lundberg tienen tercer momento finito. Entonces la probabilidad de ruina en este modelo tiene como aproximación la fórmula*

$$\psi(u) \approx \frac{\tilde{\lambda}}{\tilde{c}\,\tilde{\alpha}}\, e^{-(\tilde{\alpha}-\tilde{\lambda}/\tilde{c})u},$$

en donde

$$\tilde{\alpha} = 3\,\frac{\mu_2}{\mu_3}, \qquad \tilde{\lambda} = \frac{9}{2}\,\frac{\mu_2^3}{\mu_3^2}\,\lambda, \qquad y \quad \tilde{c} = c - \lambda\mu + \frac{3}{2}\,\frac{\mu_2^2}{\mu_3}\,\lambda.$$

Demostración. El método consiste en igualar los tres primeros momentos de las variables C_t y $\tilde{C}_t$. Primeramente veamos la igualación de las esperanzas. La condición $E(C_t) = E(\tilde{C}_t)$ produce la ecuación

$$u + ct - \lambda\mu t = u + \tilde{c}t - \tilde{\lambda}\,(\frac{1}{\tilde{\alpha}})\,t.$$

De donde se obtiene

$$\tilde{c} = c - \lambda\mu + \frac{\tilde{\lambda}}{\tilde{\alpha}}. \tag{8.15}$$

El siguiente paso es igualar las varianzas. De los resultados generales del primer capítulo, recordemos que la varianza de un riesgo S que sigue un modelo colectivo Poisson(λ) está dada por $\mathrm{Var}(S) = \lambda\mu_2$. Por lo tanto, de la condición $\mathrm{Var}(C_t) = \mathrm{Var}(\tilde{C}_t)$ se obtiene

$$\lambda\mu_2 = 2\,\frac{\tilde{\lambda}}{\tilde{\alpha}^2}. \tag{8.16}$$

Hemos usado el hecho de que si X tiene distribución $\exp(\alpha)$, entonces $E(X^2) = 2/\alpha^2$. Finalmente, recordemos que el tercer momento central de un riesgo S que sigue un modelo colectivo Poisson(λ) está dado por $E(S - E(S))^3 = \lambda\mu_3$. Por lo tanto, la tercera condición $E(C_t - E(C_t))^3 = E(\tilde{C}_t - E(\tilde{C}_t))^3$ produce la ecuación

$$\lambda\mu_3 = 6\,\frac{\tilde{\lambda}}{\tilde{\alpha}^3}. \tag{8.17}$$

Hemos usado la fórmula $E(X^3) = 6/\alpha^3$, válida cuando X tiene distribución $\exp(\alpha)$. Despejando $\tilde{\lambda}$ de (8.16) y (8.17) e igualando,

$$\frac{1}{2}\lambda\mu_2\tilde{\alpha}^2 = \frac{1}{6}\lambda\mu_3\tilde{\alpha}^3.$$

Por lo tanto,

$$\tilde{\alpha} = 3\,\frac{\mu_2}{\mu_3}. \tag{8.18}$$

Sustituyendo en (8.16),

$$\tilde{\lambda} = \frac{9}{2}\,\frac{\mu_2^3}{\mu_3^2}\,\lambda. \tag{8.19}$$

Sustituyendo (8.18) y (8.19) en (8.15),

$$\tilde{c} = c - \lambda\mu + \frac{3}{2}\,\frac{\mu_2^2}{\mu_3}\,\lambda. \tag{8.20}$$

De esta forma hemos encontrado los parámetros $\tilde{c}$, $\tilde{\lambda}$ y $\tilde{\alpha}$ en términos de los parámetros del riesgo original c, λ y los momentos de Y. ■

No se espera que la aproximación de De Vylder sea muy precisa pero es muy interesante la idea de poder utilizar lo conocido para aproximar la solución desconocida de un problema general. El lector puede identificar el método de momentos para la estimación de parámetros en la aproximación de De Vylder. No es difícil comprobar que la aproximación de De Vylder es exacta cuando las reclamaciones son exponenciales.

8.8. Fórmula de Pollaczek-Khinchin

La fórmula de Pollaczek-Khinchin es una expresión general que permite escribir a la probabilidad de ruina en términos de una serie infinita de convoluciones. La demostración que presentaremos hace uso de la transformada de Laplace de la cual se recuerda su definición y algunas propiedades en el Apéndice. Para obtener la fórmula de Pollaczek-Khinchin se necesita conocer primero la transformada de Laplace de una distribución geométrica compuesta. Sea entonces X una variable aleatoria con distribución geométrica compuesta, es decir, una variable aleatoria de la forma

$$X = \sum_{j=1}^{N} U_j, \tag{8.21}$$

en donde N tiene distribución geo$(1-p)$, es decir, para $n = 0, 1, 2, \ldots$

$$P(N = n) = p^n\,(1-p),$$

y las variables $U_1, U_2, \ldots$ son independientes, no negativas, idénticamente distribuidas, y con función de distribución $H(x)$ definida de la siguiente forma particular

$$H(x) = \frac{1}{\mu} \int_0^x \overline{F}(y)\,dy,$$

en donde $F(y)$ es una función de distribución de una variable aleatoria no negativa con media finita μ. Puede comprobarse que $H(x)$ es efectivamente una función de distribución. Siguiendo la misma notación que en el modelo de Cramér-Lundberg, $F(y)$ representará la función de distribución del monto de una reclamación.

Proposición 8.11 *Sea $G(x)$ la función de distribución de la v.a. X con distribución geométrica compuesta dada por (8.21) junto con la notación e hipótesis arriba enunciados. Para valores de s suficientemente pequeños, la transformada de Laplace de $\overline{G}(x) = 1 - G(x)$ es*

$$\int_0^\infty e^{-sx}\,\overline{G}(x)\,dx = \frac{p}{s}\left(1 - (1-p)\,\frac{L_{\overline{F}}(s)}{\mu - p\,L_{\overline{F}}(s)}\right).$$

Demostración. Primeramente tenemos que

$$G(x) = \sum_{n=0}^{\infty} H^{*n}(x)(1-p)p^n.$$

Entonces

$$\begin{aligned}
\overline{G}(x) &= 1-(1-p)\sum_{n=0}^{\infty} p^n\, H^{*n}(x) \\
&= p-(1-p)\sum_{n=1}^{\infty} p^n\, H^{*n}(x).
\end{aligned}$$

La transformada de Laplace de esta función es

$$\begin{aligned}
\int_0^{\infty} e^{-sx}\,\overline{G}(x)\,dx &= \frac{p}{s}-(1-p)\int_0^{\infty} e^{-sx}\,(\sum_{n=1}^{\infty} p^n\, H^{*n}(x)\,)\,dx \\
&= \frac{p}{s}-(1-p)\sum_{n=1}^{\infty} p^n \int_0^{\infty} e^{-sx}\, H^{*n}(x)\,dx \\
&= \frac{p}{s}-(1-p)\sum_{n=1}^{\infty} p^n\, s^{n-1}\, L_H^n(s) \\
&= \frac{p}{s}-\frac{(1-p)}{s}\sum_{n=1}^{\infty} (psL_H(s))^n \\
&= \frac{p}{s}-(1-p)\frac{p\,L_H(s)}{1-ps\,L_H(s)}.
\end{aligned}$$

En la última igualdad hemos supuesto que s es suficientemente pequeño de tal forma que $psL_H(s)<1$. A partir de la definición de $H(x)$ puede demostrarse que

$$L_H(s)=\frac{1}{\mu s}L_{\overline{F}}(s).$$

Por lo tanto,

$$\int_0^{\infty} e^{-sx}\,\overline{G}(x)\,dx = \frac{p}{s}-\frac{(1-p)}{s}\,\frac{p\,L_{\overline{F}}(s)}{\mu-p\,L_{\overline{F}}(s)}.$$

■

Con la notación e hipótesis utilizadas en el modelo de Cramér-Lundberg, tenemos el siguiente resultado importante: la probabilidad de ruina está dada por la cola de la función de distribución de la variable aleatoria geométrica compuesta dada por (8.21), es decir, $\psi(u)=\overline{G}(x)$.

Teorema 8.2 (Fórmula de Pollaczek-Khinchin) *Considere el modelo de riesgo de Cramér-Lundberg en donde las reclamaciones tienen función de distribución $F(y)$ y media finita μ. Entonces la probabilidad de ruina está dada por*

$$\psi(u) = (1-p)\sum_{n=1}^{\infty} p^n\,\overline{H^{*n}}(u),$$

en donde $p = \dfrac{\lambda\mu}{c}$ *y* $H(u) = \dfrac{1}{\mu}\displaystyle\int_0^u \overline{F}(y)\,dy.$

Demostración. Sabemos que la probabilidad de ruina satisface la ecuación integral

$$\begin{aligned}
\psi(u) &= \frac{\lambda}{c}\left(\int_u^{\infty}\overline{F}(y)\,dy + \int_0^u \psi(u-y)\overline{F}(y)\,dy\right)\\
&= \frac{\lambda}{c}\left(\mu - \int_0^u \overline{F}(y)\,dy + (\psi * \overline{F})(u)\right).
\end{aligned}$$

Tomando la transformada de Laplace de esta función tenemos que

$$L_\psi(s) = \frac{\lambda}{c}\left(\frac{\mu}{s} - \frac{1}{s}L_{\overline{F}}(s) + L_\psi(s)L_{\overline{F}}(s)\right),$$

de donde se obtiene

$$\begin{aligned}
L_\psi(s) &= \frac{\frac{\lambda}{cs}(\mu - L_{\overline{F}}(s))}{1-\frac{\lambda}{c}L_{\overline{F}}(s)}\\
&= \frac{\frac{\lambda}{cs}(\mu - (p+(1-p))L_{\overline{F}}(s))}{1-\frac{\lambda}{c}L_{\overline{F}}(s)}.
\end{aligned}$$

Definiendo $p = \lambda\mu/c$, se puede escribir

$$\begin{aligned}
L_\psi(s) &= \frac{\frac{p}{s}[\mu - (p+(1-p))L_{\overline{F}}(s)]}{\mu - p\,L_{\overline{F}}(s)}\\
&= \frac{p}{s}\left(1-(1-p)\,\frac{L_{\overline{F}}(s)}{\mu - p\,L_{\overline{F}}(s)}\right).
\end{aligned}$$

De esta forma hemos encontrado que la transformada de Laplace de la probabilidad de ruina coincide con la transformada de Laplace de la función $\overline{G}(x) = 1 - G(x)$, en donde $G(x)$ es la función de distribución geométrica compuesta dada por (8.21) con las características descritas en la Proposición 8.11. Por la unicidad de la transformada de Laplace tenemos que ambas funciones deben ser iguales, es decir,

$$\begin{aligned}
\psi(u) &= \overline{G}(u) \\
&= p - (1-p) \sum_{n=1}^{\infty} p^n \, H^{*n}(u) \\
&= p - (1-p) \sum_{n=1}^{\infty} p^n \, (1 - \overline{H^{*n}}(u)) \\
&= (1-p) \sum_{n=1}^{\infty} p^n \, \overline{H^{*n}}(u).
\end{aligned}$$

■

En general no es inmediato calcular las convoluciones de la función de distribución $H(u)$ ni tampoco llevar a cabo la suma infinita que aparece en la fórmula de Pollaczek-Khinchin, pero sin duda es una fórmula bastante atractiva desde el punto de vista matemático pues, como hemos visto, esta expresión corresponde a la cola de la distribución de la variable aleatoria geométrica compuesta dada por (8.21). Pueden llevarse a cabo simulaciones de esta variable aleatoria, calcular la magnitud de las colas y conocer de manera aproximada la probabilidad de ruina cuando las reclamaciones tienen distribución continua arbitraria. A continuación reduciremos la fórmula de Pollaczek-Khinchin a la solución conocida para $\psi(u)$ en el caso cuando las reclamaciones son exponenciales.

Ejemplo 8.9 (Reclamaciones exponenciales) *Suponga que las reclamaciones tienen distribución exponencial de parámetro α. Entonces puede comprobarse fácilmente que $H(u)$ es nuevamente la función de distribución exp(α). Por lo tanto, $H^{*n}(u)$ es la función de distribución gama(n, α) y tiene la expresión*

$$H^{*n}(u) = \sum_{k=n}^{\infty} e^{-\alpha u} \frac{(\alpha u)^k}{k!}.$$

Entonces

$$\overline{H^{*n}}(u) = \sum_{k=0}^{n-1} e^{-\alpha u} \frac{(\alpha u)^k}{k!}.$$

Por lo tanto,

$$\begin{aligned}
\sum_{n=1}^{\infty} p^n \, \overline{H^{*n}}(u) &= \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} p^n \, e^{-\alpha u} \frac{(\alpha u)^k}{k!} \\
&= \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} p^n \, e^{-\alpha u} \frac{(\alpha u)^k}{k!} \\
&= \frac{p}{1-p} \sum_{k=0}^{\infty} e^{-\alpha u} \frac{(\alpha p u)^k}{k!} \\
&= \frac{p}{1-p} \, e^{-\alpha(1-p)u}.
\end{aligned}$$

De donde se confirma nuevamente que la probabilidad de ruina en el modelo de Cramér-Lundberg cuando las reclamaciones son exponenciales de parámetro α es

$$\begin{aligned}
\psi(u) &= (1-p) \sum_{n=1}^{\infty} p^n \, \overline{H^{*n}}(u) \\
&= p \, e^{-\alpha(1-p)u} \\
&= \frac{\lambda}{\alpha c} \, e^{-(\alpha - \frac{\lambda}{c})u}.
\end{aligned}$$

En la siguiente sección veremos otro caso particular para la distribución de las reclamaciones en el modelo de Cramér-Lundberg en donde la fórmula de Pollaczek-Khinchin permite encontrar una forma de expresar la probabilidad de ruina.

8.9. Probabilidad de ruina con reclamaciones tipo fase

Como último resultado presentamos una expresión para la probabilidad de ruina en el modelo de Cramér-Lundberg cuando las reclamaciones son continuas tipo fase. Para un recordatorio sobre este tipo de distribuciones continuas y algunas fórmulas elementales, véase la sección sobre el tema en el capítulo de procesos estocásticos. Para llegar a nuestro objetivo con mayor rapidez, utilizaremos los siguientes dos lemas técnicos cuyas demostraciones pueden encontrarse en el texto de Rolski *et al.* [32].

Lema 8.1 *Sea $F(y)$ la función de distribución continua PH(π, B) con B no singular, $\pi_0 = 0$ y con media μ finita. Entonces la función de distribución*

$$H(x) = \frac{1}{\mu} \int_0^x \overline{F}(y)\, dy$$

es nuevamente PH$(-\frac{1}{\mu}\,\pi\, B^{-1}, B)$.

Lema 8.2 *Sean $U_1, U_2, \ldots$ v.a.s independientes con idéntica distribución continua PH(π, B), en donde B es no singular y $\pi_0 = 0$. Sea N otra v.a. independiente de las anteriores y con distribución geo(p). Entonces la variable aleatoria*

$$X = \sum_{j=1}^{N} U_j$$

correspondiente a un modelo geométrico compuesto tiene distribución

$$\mathit{PH}(p\pi, B + pb^T\pi).$$

Observe que el vector inicial $p\pi$ de la nueva distribución tipo fase no es de probabilidad pues la suma de sus entradas es p. La entrada cero de este vector es $1-p$ y corresponde a la probabilidad de que la variable geométrica tome el valor cero, es decir, la probabilidad de que no haya sumandos.

Teorema 8.3 (Probabilidad de ruina con reclamaciones PH) *Suponga que las reclamaciones en el modelo de Cramér-Lundberg tienen distribución continua PH(π, B) con $\pi_0 = 0$, en donde la matriz de subintensidades B es no singular y se cumple la condición de ganancia neta $p = \frac{\lambda\mu}{c} < 1$. Entonces para cualquier capital inicial $u \geqslant 0$,*

$$\psi(u) = p(-\frac{1}{\mu}\,\pi\, B^{-1}) \left[\exp u(B + pb^T(-\frac{1}{\mu}\,\pi\, B^{-1})) \right] e^T. \tag{8.22}$$

Demostración. La fórmula de Pollaczek-Khinchin establece que

$$\begin{aligned}\psi(u) &= (1-p)\sum_{n=1}^{\infty} p^n\,\overline{H^{*n}}(u)\\ &= \sum_{n=1}^{\infty} p^n\,(1-p)\,P(X_1+\cdots+X_n>u),\end{aligned}$$

en donde $X_1, X_2, \ldots$ son v.a.s independientes con idéntica función de distribución

$$H(u)=\frac{1}{\mu}\int_0^u \overline{F}(y)\,dy,$$

con $F(y)$ la función de distribución $\mathrm{PH}(\pi, B)$. La idea es identificar los sumandos como probabilidades condicionales de la siguiente forma:

$$\begin{aligned}\psi(u) &= \sum_{n=1}^{\infty} P(N=n)\,P(\sum_{k=1}^{N} X_k > u \,|\, N=n)\\ &= P(\sum_{k=1}^{N} X_k > u),\end{aligned}$$

en donde N es una v.a. con distribución $\mathrm{geo}(1-p)$, es decir,

$$P(N=n)=p^n(1-p), \quad n=0,1,\ldots$$

Así, puesto que conocemos la fórmula (6.5) de la página 153 para la cola de una distribución tipo fase y los Lemas 8.1 y 8.2 nos proporcionan los parámetros de dicha distribución, podemos deducir que $H(x)$ es

$$\mathrm{PH}(-\frac{1}{\mu}\,\pi\,B^{-1}, B),$$

y por lo tanto la variable aleatoria geométrica es

$$\mathrm{PH}(p(-\frac{1}{\mu}\,\pi\,B^{-1}), B+pb^T(-\frac{1}{\mu}\,\pi\,B^{-1})).$$

Substituyendo estos parámetros en el fórmula (6.5) se obtiene el resultado enunciado:

$$\psi(u)=p(-\frac{1}{\mu}\,\pi\,B^{-1})\left[\exp\, u(B+pb^T(-\frac{1}{\mu}\,\pi\,B^{-1}))\right] e^T.$$

■

Ejemplo 8.10 (Reclamaciones exponenciales) *Consideremos nuevamente el modelo de Cramér-Lundberg cuando las reclamaciones son exponenciales de parámetro α. Esta vez consideraremos a la distribución exponencial como una distribución tipo fase continua. El correspondientes generador infinitesimal de la cadena de Markov a tiempo continuo que produce el tiempo de espera hasta la absorción exponencial de parámetro α es*

$$G = \begin{pmatrix} 0 & 0 \\ \alpha & -\alpha \end{pmatrix}.$$

La matriz $B = -\alpha$ y el vector $b = \alpha$ son unidimensionales, $\mu = 1/\alpha$ y $p = \lambda/(\alpha c)$. Substituyendo estos términos en la fórmula (8.22) se obtiene nuevamente la expresión

$$\psi(u) = \frac{\lambda}{\alpha c}\, e^{-(\alpha - \frac{\lambda}{c})u}.$$

Comentarios y referencias

En este capítulo final hemos estudiado algunos aspectos del modelo clásico de riesgo a tiempo continuo llamado modelo de Cramér-Lundberg. Hemos encontrado expresiones para algunas probabilidades de ruina en este modelo en términos de ecuaciones diferenciales, integrales o ambas, y también en algunos casos en términos de sumas infinitas. Muchos otros problemas matemáticos pueden estudiarse para este modelo o para sus generalizaciones. El lector interesado en profundizar en el tema puede consultar textos más avanzados como Asmussen [2] o Rolski *et al.* [32]. En el texto de Schmidli [35] puede encontrarse una exposición sobre la interesante relación entre algunos problemas de riesgo y la teoría del control estocástico.

8.10. Ejercicios

Modelo de Cramér-Lundberg

170. Considere el proceso de Cramér-Lundberg $\{C_t : t \geqslant 0\}$ con la notación e hipótesis usuales. Demuestre que:

a) $E(C_t) = u + (c - \lambda\mu)t$.
b) $\mathrm{Var}(C_t) = \lambda t \mu_2$.
c) $M_{C_t}(r) = \exp\left[\, r(u + ct) + \lambda t(M_Y(-r) - 1)\,\right]$.

Probabilidades de ruina

171. Suponga que las reclamaciones en el modelo de Cramér-Lundberg siguen una distribución exponencial de parámetro $\alpha = 1$. Suponga además que $\lambda = 1/2$ y $c = 2$. Observe que se cumple la condición de ganancia neta $c > \lambda\mu$. ¿Cuál debe ser el capital inicial u para que la probabilidad de ruina sea menor o igual a 0.01?

Coeficiente de ajuste

172. Demuestre que si las reclamaciones son variables aleatorias acotadas, entonces el coeficiente de ajuste existe.

173. Se ha demostrado que cuando las reclamaciones tienen distribución $\exp(\alpha)$ el coeficiente de ajuste es $R = \alpha - \lambda/c$. Usando la condición de ganancia neta compruebe que este valor de R efectivamente cumple la desigualdad

$$0 < R < \frac{2(c - \lambda\mu)}{\lambda\mu_2}.$$

Desigualdad de Lundberg

174. Demuestre que efectivamente el evento $(\tau \leqslant t)$ converge monótonamente al evento $(\tau < \infty)$ cuando t tiende a infinito monótonamente. Este resultado fue usado en la demostración de la desigualdad de Lundberg.

Aproximación de De Vylder

175. Demuestre que la aproximación de De Vylder coincide con la probabilidad de ruina en el caso cuando las reclamaciones tienen distribución $\exp(\alpha)$.

Apéndice: Formulario y resultados varios

Fórmula de De Pril en R

```
####################################
#  F\'ormula de De Pril en R v1.1  #
####################################
I <- 5 # Montos de reclamaciones
J <- 3 # \'Indice m\'aximo para tasas de mortalidad
R <- 20 # Valor m\'aximo para r en g(r)
#
n <- array(1:15, dim=c(5,3))
n[1,1]<-1
n[2,1]<-3
n[3,1]<-5
n[4,1]<-2
n[5,1]<-2
n[1,2]<-3
n[2,2]<-5
n[3,2]<-3
n[4,2]<-2
n[5,2]<-3
n[1,3]<-1
n[2,3]<-4
n[3,3]<-4
n[4,3]<-6
n[5,3]<-4
#
q <- array(1:3, dim=c(3))
```

```
q[1]<-0.03
q[2]<-0.04
q[3]<-0.05
#..............................
# Funci\'on h(i,k)
#..............................
h <- function(i,k) {
aux <- 0
for (j in 1:J) {
aux <- aux+n[i,j]*(q[j]/(1-q[j]))^k
}
aux <- i*((-1)^(k-1))*aux
return(aux)
}
#..............................
# C\'alculo de la densidad de S
#..............................
gc <- array(1:R, dim=c(R))
gc0 <- g(0)
#
g <- function(r) {
if (r==0) {
aux <- 1
for (i in 1:I) {
    for (j in 1:J) {
    aux <- aux*((1-q[j])^n[i,j])
    }
}
return(aux)
}
else
{
aux <- 0
for (i in 1:min(r,I)) {
    for (k in 1:floor(r/i)) {
        if (r-i*k==0) { aux <- aux + gc0*h(i,k) }
        else {aux <- aux + gc[r-i*k]*h(i,k)}
        }
}
aux <- aux/r
gc[r] <- aux
```

```
return(aux)
}
}
#.............................
# Asignaci\'on en el arreglo "gc" y graficaci\'on.
#.............................
for (i in 1:R) {
gc[i] <- g(i)
}
# Nota: Se omite en la gr\'afica el valor de la densidad en cero "gc0".
barplot(gc,main="Funci?n de densidad de S",xlab="r", ylab="g(r)")
#################################################################
# Fin de c\'odigo
#################################################################
```

Fórmula de Panjer en R (Caso Poisson)

```
####################################
#  F\'ormula de Panjer en R v1.0   #
#  [Caso Poisson]                  #
####################################
#
R <- 20 # Valor m\'aximo para r en g(r)
#
#.............................
# c\'alculo de p_k=P(N=k) (Caso Poisson)
#.............................
a <- 0
b <- 3.5  #lambda
p0 <- 2.7172^{-b}
p <- array(1:R, dim=c(R))
p[1] <- (a+b)*p0
for (k in 2:R) {
p[k] <- (a+b/k)*p[k-1]
}
#.............................
# c\'alculo de f_r=P(Y=r), r>=1
#.............................
#
```

```
f <- array(1:R, dim=c(R))
f[1] <- 0.1
f[2] <- 0.1
f[3] <- 0.2
f[4] <- 0.3
f[5] <- 0.3
for (i in 5:R) { f[i] <- 0 }
#...............................
# C\'alculo de la densidad de S
#...............................
g0 <- p0
g <- array(1:R, dim=c(R))
g[1] <- (a+b)*f[1]*g0
for (r in 2: R) {
aux <- 0
for (i in 1:{r-1}) {
    aux <- aux + (a+b*i/r)*f[i]*g[r-i]
               }
aux <- aux + (a+b)*f[r]*g0
g[r] <- aux
}
#..............................
# Graficaci\'on
#..............................
# Nota: Se omite en la gr\'afica el valor de la densidad en cero "g0".
barplot(g,main="Funcin de densidad de S",xlab="r", ylab="g(r)")
#
#################################################################
# Fin de c\'odigo
#################################################################
```

Alfabeto griego

$\mathrm{A}\ \alpha$	alpha	$\mathrm{I}\ \iota$	iota	$\mathrm{P}\ \rho, \varrho$	rho
$\mathrm{B}\ \beta$	beta	$\mathrm{K}\ \kappa$	kappa	$\Sigma\ \sigma, \varsigma$	sigma
$\Gamma\ \gamma$	gamma	$\Lambda\ \lambda$	lambda	$T\ \tau$	tau
$\Delta\ \delta$	delta	$M\ \mu$	mu	$\Upsilon\ \upsilon$	upsilon
$\mathrm{E}\ \epsilon, \varepsilon$	epsilon	$\mathrm{N}\ \nu$	nu	$\Phi\ \phi, \varphi$	phi
$\mathrm{Z}\ \zeta$	zeta	$\Xi\ \xi$	xi	$X\ \chi$	chi
$H\ \eta$	eta	$\mathrm{O}\ \mathrm{o}$	omikron	$\Psi\ \psi$	psi
$\Theta\ \theta, \vartheta$	theta	$\Pi\ \pi$	pi	$\Omega\ \omega$	omega

Función indicadora

La función indicadora de un conjunto $A \subseteq \Omega$ es la función $1_A : \Omega \to \{0, 1\}$ definida por

$$1_A(\omega) = \begin{cases} 1 & \text{si } \ \omega \in A, \\ 0 & \text{si } \ \omega \notin A. \end{cases}$$

De este modo la función 1_A toma el valor uno dentro del conjunto A y cero fuera de él, y cumple las siguientes propiedades:

a) $1_{A\cup B} = \text{máx}\{1_A, 1_B\} = 1_A + 1_B - 1_A\, 1_B$.

b) $1_{A\cap B} = \text{mín}\{1_A, 1_B\} = 1_A\, 1_B$.

c) $1_{A^c} = 1 - 1_A$.

d) $1_{A-B} = 1_A - 1_A\, 1_B$.

e) $1_{A\triangle B} = |1_A - 1_B| = 1_A + 1_B - 2\, 1_A\, 1_B = (1_A - 1_B)^2$.

f) $A \subseteq B \Rightarrow 1_A \leqslant 1_B$.

Función generadora de probabilidad

La función generadora de probabilidad de una variable aleatoria X es la función dada por

$$G(t) = E(t^X) = \sum_x t^x\, P(X = x),$$

definida para valores reales de t tal que la suma indicada sea absolutamente convergente. Esta función se utiliza principalmente para variables aleatorias discretas y sin pérdida de generalidad se postula que la variable toma los valores $0, 1, \ldots$ y por lo tanto esta función existe por lo menos para valores de t en el conjunto $[-1, 1]$. Para hacer énfasis en la variable aleatoria particular X para la que se utiliza se escribe a veces como $G_X(t)$. Esta función cumple las siguientes propiedades:

a) La función generadora de probabilidad adquiere su nombre a partir del siguiente resultado: para cada entero $n \geqslant 0$,

$$\lim_{t \to 0} \frac{d^n}{dt^n} G_X(t) = P(X = n).$$

b) La función generadora de probabilidad determina de manera única a la distribución de probabilidad de la variable aleatoria, es decir, si X y Y tienen la misma distribución de probabilidad, entonces claramente tienen la misma función generadora de probabilidad. Inversamente, si X y Y son tales que sus funciones generadoras de probabilidad $G_X(t)$ y $G_y(t)$ coinciden en un intervalo no trivial alrededor del cero, entonces X y Y tienen la misma distribución de probabilidad.

c) Si el n-ésimo momento factorial de X existe, entonces

$$\lim_{t \nearrow 1} \frac{d^n}{dt^n} G_X(t) = E[X(X-1)\cdots(X-n+1)].$$

d) Si X y Y son independientes, entonces

$$G_{X+Y}(t) = G_X(t)\, G_Y(t).$$

Para un estudio más detallado de la función generadora de probabilidad, el lector puede consultar el libro de Gut [17].

Función gama
Para valores de α donde la integral es convergente se define

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx,$$

y se cumplen las siguientes propiedades:

a) $\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$.

b) $\Gamma(\alpha+1) = \alpha!$ cuando α es entero positivo.

c) $\Gamma(2) = \Gamma(1) = 1$.

d) $\Gamma(1/2) = \sqrt{\pi}$.

Pueden consultars algunas otras propiedades de la función gama en [36].

Funciones de utilidad
Una función de utilidad es una función $v(x)$ que representa el valor o utilidad que un individuo o institución asocia a cada cantidad de un bien. Matemáticamente a una función de utilidad se le define a través de las siguientes dos propiedades:

a) Es estrictamente creciente.

b) Es cóncava.

Suponiendo diferenciabilidad, estas propiedades se escriben como $v'(x) > 0$ y $v''(x) \leqslant 0$, respectivamente. La primera propiedad representa el hecho evidente de que una cantidad mayor de dinero siempre tiene un valor o utilidad mayor. Para interpretar la segunda propiedad consideremos la situación en la que tanto una persona pobre como una persona rica incrementan ambos su capital por una misma cantidad. Entonces este incremento representa para una persona con poco dinero un gran incremento en su utilidad, mientras que el mismo incremento representa un menor incremento en la utilidad para una persona con mayor cantidad de dinero. En otras palabras, cuando hay un incremento de capital fijo, el valor o utilidad crece más rápido cuando uno tiene poco dinero y más lento cuando uno tiene mucho dinero. Dos funciones de utilidad $u(x)$ y $v(x)$ son equivalentes si existen constantes a y b, con $a > 0$, tales que $v(x) = au(x) + b$. La razón que subyace en esta definición de equivalencia radica en el hecho de que si un persona con capital x, función de utilidad $u(x)$, y usando el principio de la utilidad esperada, prefiere la inversión I_1 sobre la inversión I_2 si

$$E[u(x + I_1)] > E[u(x + I_2)],$$

y tal decisión no cambia si en lugar de usar la función de utilidad $u(x)$ utiliza ahora $v(x) = au(x) + b$, pues la desigualdad anterior es equivalente a

$$E[au(x + I_1) + b] > E[au(x + I_2) + b].$$

Del conjunto de funciones de utilidad equivalentes a una función de utilidad dada $u(x)$, uno puede escoger aquella función de utilidad $v(x)$ tal que $v(0) = 0$ y $v(1) = 1$. Tal función de utilidad $v(x)$ está dada por

$$v(x) = \frac{u(x) - u(0)}{u(1) - u(0)},$$

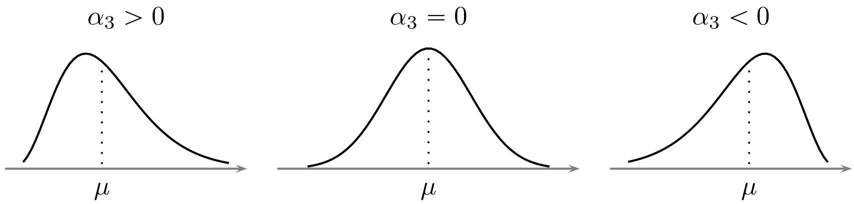
suponiendo que los valores $x = 0$ y $x = 1$ pertenecen al dominio de definición de la función $u(x)$. Es por ello que a una función de utilidad $u(x)$ se le puede pedir la condición $u(0) = 0$, sin perder generalidad ni provocar cambios en la toma de decisiones bajo el criterio de utilidad esperada. El lector puede encontrar exposiciones breves sobre la teoría de utilidad y los seguros en el primer capítulo de los textos [7] y [22].

Coeficiente de asimetría de Fisher

Para una variable aleatoria X con tercer momento finito se define el coeficiente de asimetría de Fisher α_3 como el número

$$\alpha_3 = \frac{E[(X - E(X))^3]}{[\mathrm{Var}(X)]^{3/2}}.$$

Esta cantidad es una medida de la asimetría de la distribución alrededor de su media. Cuando $\alpha_3 = 0$ la distribución es completamente simétrica alrededor de su media, como es el caso, por ejemplo, de la distribución normal. Cuando $\alpha_3 > 0$ se dice que la distribución es asimétrica positiva o que tiene mayor sesgo hacia la derecha (es decir, la cola a la derecha de la media es más larga que la de la izquierda, o bien, que hay valores más separados de la media a la derecha). Cuando $\alpha_3 < 0$, se dice que la distribución es asimétrica negativa o que tiene mayor sesgo a la izquierda (es decir, la cola a la izquierda de la media es más larga que la de la derecha, o bien, que hay valores más separados de la media a la izquierda).



Desigualdad de Jensen
Sea u una función convexa y sea X una variable aleatoria tal que tanto X como $u(X)$ tienen esperanza finita. Entonces

$$u(E(X)) \leqslant E(u(X)).$$

En el caso cuando u es cóncava, el resultado es

$$u(E(X)) \geqslant E(u(X)).$$

Una demostración de la desigualdad de Jensen puede encontrarse por ejemplo en [17].

Desigualdad de Markov
Sea X una variable aleaotoria positiva con esperanza finita. Entonces para cualquier $\epsilon > 0$,

$$P(X \geqslant \epsilon) \leqslant \frac{E(X)}{\epsilon}.$$

La demostración de este resultado es un buen ejercicio para el lector.

Esperanza condicional
Sean $(\Omega, \mathscr{F}, P)$ un espacio de probabilidad, X una variable aleatoria con esperanza finita y $\mathscr{G}$ una sub-σ-álgebra de $\mathscr{F}$. La esperanza condicional de X dado $\mathscr{G}$ es una variable aleatoria denotada por $E(X \,|\, \mathscr{G})$ que cumple las siguientes tres propiedades:

1. Es $\mathscr{G}$-medible.
2. Tiene esperanza finita.
3. Para cualquier evento G en $\mathscr{G}$, $\quad E[\, E(\, X \,|\, \mathscr{G}\,)\, 1_G\,] = E[\, X\, 1_G\,]$.

Puede demostrarse que esta variable aleatoria existe y es única casi seguramente, esto significa que si existe otra variable aleatoria con las tres propiedades anteriores, entonces con probabilidad uno coincide con $E(X\,|\,\mathscr{G})$. Cuando $\mathscr{G} = \sigma(Y)$ para alguna variable aleatoria Y, se escribe $E(X\,|\,Y)$ en lugar de $E(X\,|\,\sigma(Y))$. En particular, el término $P(A\,|\,Y)$ significa $E(1_A\,|\,Y)$. Se enuncian a continuación algunas propiedades de esta esperanza.

a) $E(X\,|\,\{\emptyset,\Omega\}\,) = E(X)$.

b) $E(1_A\,|\,\{\emptyset,\Omega\}\,) = P(A)$.

c) $E(1_A\,|\,\{\emptyset,B,B^c,\Omega\}\;) = P(A|B)1_B + P(A|B^c)1_{B^c}$.

d) $E(E(X\,|\,\mathscr{G})) = E(X)$. En particular, $E(P(A\,|\,Y)) = E(1_A) = P(A)$.

e) Si X es $\mathscr{G}$-medible, entonces $E(X\,|\,\mathscr{G}) = X$. En particular, si c es una constante, entonces $E(c\,|\,\mathscr{G}) = c$.

f) $E(aX + Y\,|\,\mathscr{G}) = aE(X\,|\,\mathscr{G}) + E(Y\,|\,\mathscr{G})$.

g) Si $X \geqslant 0$, entonces $E(X\,|\,\mathscr{G}) \geqslant 0$.

h) *Teorema de convergencia monótona.*
Si $0 \leqslant X_n \nearrow X$, entonces $E(X_n\,|\,\mathscr{G}) \nearrow E(X\,|\,\mathscr{G})$ c.s.

i) *Teorema de convergencia dominada.*
Si $|X_n| \leqslant Y$, $E|Y| < \infty$ y $X_n \to X$ c.s., entonces $E(X_n\,|\,\mathscr{G}) \to E(X\,|\,\mathscr{G})$ c.s.

j) *Desigualdad de Jensen.*
Si φ es convexa, entonces $\varphi(E(X\,|\,\mathscr{G})) \leqslant E(\varphi(X)\,|\,\mathscr{G})$.

k) Si $\mathscr{H}$ es una sub σ-álgebra de $\mathscr{G}$, entonces $E(E(X\,|\,\mathscr{G})\,|\,\mathscr{H}) = E(X\,|\,\mathscr{H})$.

l) Si Z es $\mathscr{G}$-medible y acotada, entonces $E(Z\,X\,|\,\mathscr{G}) = Z\,E(X\,|\,\mathscr{G})$.

m) Si X es independiente de $\mathscr{G}$, entonces $E(X\,|\,\mathscr{G}) = E(X)$.

Un estudio más detallado de la esperanza condicional puede ser encontrado en libros dedicados a probabilidad como [21] o [40].

Integral de Riemann-Stieltjes
La integral de Riemann-Stieltjes generaliza a la integral usual de Riemann. Se trata de la integral de una función $h(x)$ respecto de otra función $F(x)$. Su definición es análoga al caso de la integral de Riemann: si $a = x_0 < x_1 < \cdots < x_n = b$ es una partición del intervalo $[a, b]$, entonces definimos de manera informal

$$\int_a^b h(x)\, dF(x) := \lim_{\Delta x \to 0} \sum_{i=1}^{n} h(x_i)(F(x_i) - F(x_{i-1})),$$

en donde Δx es el máximo de las distancias $x_i - x_{i-1}$, y las funciones $h(x)$ y $F(x)$ deben cumplir ciertas condiciones para que la integral tenga sentido y esté bien definida. En particular, a la función integradora $F(x)$ se le pide que sea continua por la derecha, monótona no decreciente y tal que $F(b) - F(a) < M$, para algún número $M > 0$. Observe que $F(x)$ debe cumplir propiedades semejantes a las de una función de distribución, justamente usaremos a las funciones de distribución como funciones integradoras. En particular, cuando $F(x) = x$ sobre el intervalo $[a, b]$ se tiene que

$$\int_a^b h(x)\, dF(x) = \int_a^b h(x)\, dx,$$

suponiendo la existencia de tal integral. Igualmente, bajo la hipótesis de existencia de todas las integrales que aparecen a continuación, la integral de Riemann-Stieltjes cumple las siguientes propiedades:

a) Es lineal en el integrando, es decir,

$$\int_a^b (\alpha h_1(x) + h_2(x))\, dF(x) = \alpha \int_a^b h_1(x)\, dF(x) + \int_a^b h_2(x)\, dF(x).$$

b) Es también lineal en el integrador, es decir,

$$\int_a^b h(x)\, d(\alpha F_1(x) + F_2(x)) = \alpha \int_a^b h(x)\, dF_1(x) + \int_a^b h(x)\, dF_2(x).$$

c) Cuando $h(x)$ tiene primera derivada continua se cumple la siguiente fórmula de integración por partes:

$$\int_a^b h(x)\, dF(x) = h(b)F(b) - h(a)F(a) - \int_a^b F(x)h'(x)\, dx.$$

d) De especial interés en la teoría de la probabilidad es el caso cuando $F(x)$ es diferenciable. Bajo tal hipótesis se tiene la igualdad

$$\int_a^b h(x)\, dF(x) = \int_a^b h(x)F'(x)\, dx.$$

En particular, tomando $h(x) = x$ y si X es una v.a. absolutamente continua con esperanza finita, con función de distribución $F(x)$ y función de densidad $f(x)$, entonces

$$E(X) = \int_{-\infty}^{\infty} x\, dF(x) = \int_{-\infty}^{\infty} x f(x)\, dx.$$

e) Otro caso interesante para la teoría de la probabilidad ocurre cuando $h(x)$ es continua y $F(x)$ es constante excepto en los puntos $x_0, x_1, \ldots$, en donde la función tiene saltos positivos de tamaño $p(x_0), p(x_1), \ldots$ respectivamente. En este caso y suponiendo convergencia,

$$\int_a^b h(x)\, dF(x) = \sum_{i=0}^{\infty} h(x_i)\, p(x_i).$$

Esto significa que integrar respecto de la función de distribución de una variable aleatoria discreta se reduce a efectuar una suma. Nuevamente tomando $h(x) = x$ y si X es una v.a. discreta con esperanza finita y con función de distribución como se mencionó antes, entonces

$$E(X) = \int_{-\infty}^{\infty} x\, dF(x) = \sum_{i=0}^{\infty} x_i p(x_i).$$

En el caso cuando X es una v.a. mixta con esperanza finita, en el cálculo de la esperanza se separa la parte continua de la parte discreta de la siguiente forma

$$E(X) = \int_{-\infty}^{\infty} x\, dF(x) = \int_{-\infty}^{\infty} x f(x)\, dx + \sum_{i=0}^{\infty} x_i\, p(x_i).$$

Un tratamiento más riguroso y completo de la integral de Riemann-Stieltjes puede encontrarse en el texto de probabilidad de Harris [18], o en libros de análisis matemático como el de Apostol [1].

Variables aleatorias mixtas

Una variable aleatoria mixta X es aquella que no es ni continua ni es discreta. Su función de distribución puede escribirse de la siguiente forma

$$F(x) = \int_{-\infty}^{x} f(u)\,du + \sum_{x_i \leqslant x} p(x_i),$$

en donde $f(u)$ es una función no negativa y $p(x_i) = P(X = x_i) > 0$ para ciertos valores $x_0, x_1, \ldots$ Si g es una función tal que $g(X)$ es una variable aleatoria integrable, entonces su esperanza se calcula de la siguiente forma

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)\,dF(x) = \int_{-\infty}^{\infty} g(x) f(x)\,dx + \sum_{i=0}^{\infty} g(x_i) p(x_i).$$

En el libro de Gut [17] puede encontrarse un tratamiento completo sobre variables aleatorias mixtas.

Varianza condicional

Sea X una variable aleatoria con segundo momento finito, y sea $\mathscr{G}$ una sub-σ-álgebra de $\mathscr{F}$. La varianza condicional de X dado $\mathscr{G}$ se define como la variable aleatoria dada por

$$\mathrm{Var}(X \,|\, \mathscr{G}) = E[\ (X - E(X|\mathscr{G}))^2 \ |\mathscr{G}\,].$$

Nuevamente, cuando la sub-σ-álgebra $\mathscr{G}$ es $\sigma(Y)$ para alguna variable aleatoria Y, entonces $\mathrm{Var}(X \,|\, \mathscr{G})$ se escribe $\mathrm{Var}(X \,|\, Y)$, y puede tomarse como definición la igualdad

$$\mathrm{Var}(X \,|\, Y) = E[\ (X - E(X \,|\, Y))^2 \ | Y\,].$$

Se enuncian a continuación algunas propiedades de esta variable aleatoria.

a) $\mathrm{Var}(X \,|\, \{\emptyset, \Omega\}) = \mathrm{Var}(X)$.

b) $\mathrm{Var}(1_A \,|\, \{\emptyset, \Omega\}) = P(A)(1 - P(A))$.

c) $\mathrm{Var}(X \,|\, \mathscr{G}) = E(X^2 \,|\, \mathscr{G}) - E^2(X \,|\, \mathscr{G})$.

d) $\mathrm{Var}(X) = E[\mathrm{Var}(X \,|\, \mathscr{G})] + \mathrm{Var}[E(X \,|\, \mathscr{G})]$.

Siendo la varianza condicional parte del concepto general de esperanza condicional, en los textos de Karr [21] y Williams [40] se puede encontrar mayor información sobre estas variables aleatorias.

Ley de los grandes números

Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con media μ. Entonces

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu.$$

Cuando la convergencia es en probabilidad este resultado se conoce como la ley débil. Y cuando la convergencia es casi segura se llama ley fuerte. Las demostraciones de estos resultados pueden encontrarse en [17].

Teorema central del límite

Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con media μ y varianza σ^2. Entonces

$$\frac{(\frac{1}{n}\sum_{i=1}^{n} X_i) - \mu}{\sqrt{\sigma^2/n}} \xrightarrow{d} \mathrm{N}(0,1).$$

La demostración de este importante resultado así como de los siguientes dos teoremas de convergencia pueden encontrarse en [17].

Teorema de convergencia dominada

Sea $X_1 \leqslant X_2 \leqslant \cdots$ una sucesión de variables aleatorias para la cual existe otra variable aleatoria Y con esperanza finita y tal que $|X_n| \leqslant Y$, para $n \geqslant 1$. Si X_n converge casi seguramente a una variable X, entonces tanto X como X_n tienen esperanza finita y

$$\lim_{n\to\infty} E(X_n) = E(X).$$

Teorema de convergencia monótona

Sea $0 \leqslant X_1 \leqslant X_2 \leqslant \cdots$ una sucesión monótona de variables aleatorias convergente casi seguramente a una variable X. Entonces

$$\lim_{n\to\infty} E(X_n) = E(X).$$

Fórmulas recursivas para calcular convoluciones
Sean $X_1, X_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas. Nos interesa encontrar la distribución de $S_n = X_1 + \cdots + X_n$. Cuando las variables $\{X_i\}$ tienen función de distribución F, a la función de distribución de S_n se le llama la n-ésima convolución de F, y se le denota por F^{*n}, es decir, $F^{*n}(x) = P(S_n \leqslant x)$. Cuando las variables $\{X_i\}$ tienen función de probabilidad o de densidad f, a la función de probabilidad o de densidad de S_n se le llama la n-ésima convolución de f, y se le denota por f^{*n}, es decir, en el caso discreto, $f^{*n}(x) = P(S_n = x)$.

1. Cuando las variables $\{X_i\}$ son discretas con valores en el conjunto $\{0, 1, 2, \ldots\}$, se cumplen las siguientes fórmulas recursivas.

   a) $P(S_n = x) = \displaystyle\sum_{j=0}^{x} P(S_{n-1} = x - j)\, P(X_n = j)$.

   b) $P(S_n \leqslant x) = \displaystyle\sum_{j=0}^{x} P(S_{n-1} \leqslant x - j)\, P(X_n = j)$.

2. Cuando las variables $\{X_i\}$ son continuas con soporte en el intervalo $(0, \infty)$, con función de distribución F, y con función de densidad f, se cumplen las siguientes fórmulas recursivas.

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

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

Transformada de Laplace
Para una función $\psi(u) : [0, \infty) \to \mathbb{R}$, la transformada de Laplace es

$$s \mapsto L_\psi(s) = \int_0^\infty e^{-su} \psi(u)\, du,$$

para valores reales de s donde tal integral exista. Hemos usado la notación para la probabilidad de ruina $\psi(u)$ como función objeto en la transformada de Lapace pues a este tipo de funciones se le aplicará con mayor frecuencia esta transformación. Cuando sea conveniente denotaremos la transformada de Laplace también como $L[\psi(u)](s)$. Revisaremos a continuación algunas propiedades que cumple esta transformación.

a) La transformada de una constante c es $L_c(s) = \dfrac{c}{s}$.

b) Linealidad: si α_1 y α_2 son constantes, entonces

$$L_{\alpha_1\psi_1+\alpha_2\psi_2}(s) = \alpha_1\, L_{\psi_1}(s) + \alpha_2\, L_{\psi_2}(s).$$

c) La transformada de Laplace de la derivada de una función es

$$L_{\psi'}(s) = s\, L_\psi - \psi(0).$$

d) La transformada de Laplace de la integral $H(u) = \displaystyle\int_0^u \psi(x)\, dx$ es

$$L_H(s) = \frac{1}{s}\, L_\psi(s).$$

e) La transformada de Laplace de la convolución de dos funciones es

$$L_{\psi_1 * \psi_2}(s) = L_{\psi_1}(s)\, L_{\psi_2}(s),$$

en donde $(\psi_1 * \psi_2)(u) = \displaystyle\int_0^u \psi_1(u-x)\psi_2(x)dx$. Observe sin embargo la siguiente situación particular: si $G(x)$ es la convolución de dos funciones de distribución continuas $F_1(x)$ y $F_2(x)$ de variables aleatorias no negativas, es decir,

$$G(x) = \int_0^x F_1(x-y)dF_2(y) = \int_0^x F_1(x-y)f_2(y)dy = (F_1 * f_2)(x),$$

en donde $f_2(y)$ es la función de densidad de $F_2(y)$, entonces

$$L_G(s) = L_{F_1}(s)L_{f_2}(s) = sL_{F_1}(s)L_{F_2}(s).$$

f) Existencia: se dice que una función $\psi(u)$ definida para $u \geqslant 0$ es de orden exponencial cuando u tiende a infinito si existe una constante u_0 tal que para todo $u \geqslant u_0$,

$$|\psi(u)| \leqslant m\, e^{cu},$$

en donde m y c son constantes. Puede demostrarse que cuando esta condición se cumple la transformada de Laplace $L_\psi(s)$ existe para $s \geqslant c$.

g) Unicidad: sean $\psi_1(u)$ y $\psi_2(u)$ dos funciones continuas definidas para $u \geqslant 0$, ambas de orden exponencial, si existe una constante s_0 tal que $L_{\psi_1}(s) = L_{\psi_2}(s)$ para todo $s > s_0$, entonces $\psi_1(u) = \psi_2(u)$ para todo $u \geqslant 0$.

Un tratamiento más completo sobre la transformada de Laplace y sus aplicaciones puede encontrarse en [33].

Método de Newton-Raphson

Sea $g(x)$ una función diferenciable que tiene una raíz cerca de $x = x_0$. Véase la Figura 6. El método de Newton-Raphson permite conocer el valor de esta raíz mediante aproximaciones sucesivas.

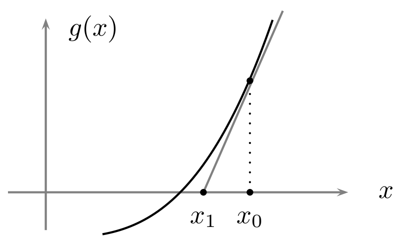


Figura 6

La primera aproximación de la raíz es el valor inicial x_0. La ecuación de la recta tangente que pasa por el punto $(x_0, g(x_0))$ es

$$y - g(x_0) = g'(x_0)(x - x_0).$$

Esta línea recta cruza el eje horizontal cuando el valor de x es igual a

$$x_1 = x_0 - \frac{g(x_0)}{g'(x_0)}.$$

Se toma a x_1 como nueva raíz aproximada y se repite el procedimiento. De este modo se encuentra una sucesión de valores $x_0, x_1, x_2, \ldots$, que bajo ciertas condiciones converge a la raíz de la función $g(x)$. La fórmula general es entonces

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}.$$

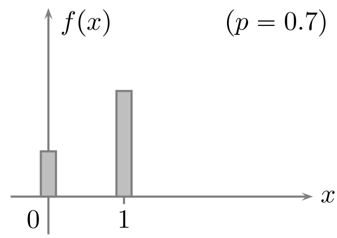
En el presente texto se ha mencionado este mecanismo para encontrar de manera aproximada el coeficiente de ajuste en los modelos de riesgo estudiados. El lector puede consultar, por ejemplo, el texto de Burden y Faires [10] para profundizar sobre el método de Newton-Raphson.

Distribuciones de probabilidad

En esta sección se presentan en orden alfabético algunas distribuciones de probabilidad utilizadas en el texto. La función generadora de probabilidad se denota por $G(t)$, y la función generadora de la momentos por $M(t)$.

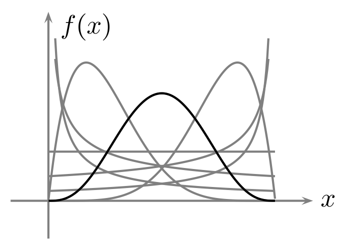
Distribución Bernoulli

$X \sim \text{Ber}(p)$ con $p \in (0,1)$.
$f(x) = p^x(1-p)^{1-x}$ para $x = 0, 1$.
$E(X) = p$.
$\text{Var}(X) = p(1-p)$.
$G(t) = 1 - p + pt$.
$M(t) = (1-p) + pe^t$.



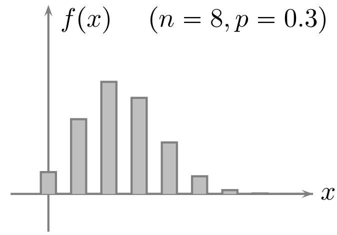
Distribución beta

$X \sim \text{beta}(a, b)$ con $a > 0, b > 0$.
$f(x) = x^{a-1}(1-x)^{b-1}/B(a,b)$ para $x \in (0,1)$.
en donde $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$,
con $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}\,dt$.
$E(X) = a/(a+b)$.
$\text{Var}(X) = ab/[(a+b+1)(a+b)^2]$.
Cuando $a = b = 1$ se obtiene la dist. unif$(0,1)$.



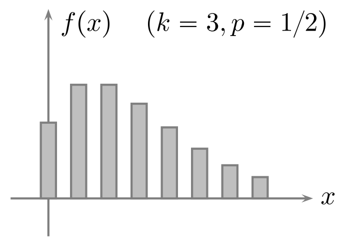
Distribución binomial

$X \sim \text{bin}(n, p)$ con $n \in \{1, 2, \ldots\}$ y $p \in (0, 1)$.
$f(x) = \binom{n}{k} p^x (1-p)^{n-x}$ para $x = 0, 1, \ldots, n$.
$E(X) = np$.
$\text{Var}(X) = np(1-p)$.
$G(t) = (1 - p + pt)^n$.
$M(t) = [(1-p) + pe^t]^n$.



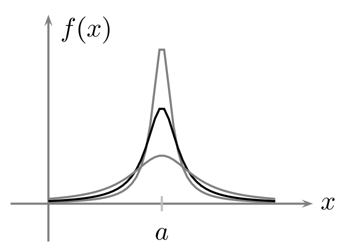
Distribución binomial negativa

$X \sim \text{bin neg}(k, p)$ con $p \in (0, 1)$ y $k \in \{1, 2, \ldots\}$.
$f(x) = \binom{k + x - 1}{x} p^k (1-p)^x$ para $x = 0, 1, \ldots$
$E(X) = k(1-p)/p$.
$\text{Var}(X) = k(1-p)/p^2$.
$G(t) = [p/(1 - (1-p)t)]^k$.
$M(t) = [p/(1-(1-p)e^t)]^k$ para $t < -\ln(1-p)$.
Cuando $r = 1$ se obtiene la distribución geo(p).

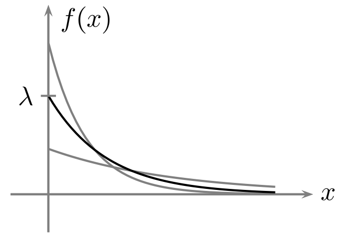


Distribución Cauchy

$X \sim \text{Cauchy}(a, b)$ con $a \in \mathbb{R}$ y $b > 0$.
$f(x) = \dfrac{1}{b\pi[1 + ((x-a)/b)^2]}$, para $x \in \mathbb{R}$.
$F(x) = 1/2 + (1/\pi)\arctan((x-a)/b)$.
La esperanza y varianza no existen.
La función generadora de momentos no existe.
Cuando $a = 0$ y $b = 1$ se obtiene la distribución Cauchy estándar, Cauchy$(0, 1)$.

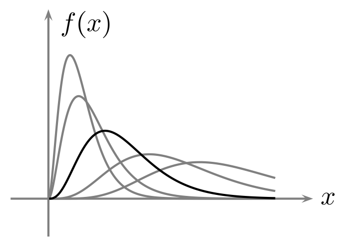


Distribución exponencial



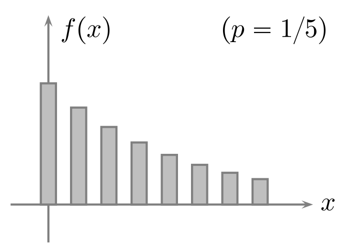
$X \sim \exp(\lambda)$ con $\lambda > 0$.
$f(x) = \lambda e^{-\lambda x}$ para $x > 0$.
$F(x) = 1 - e^{-\lambda x}$ para $x > 0$.
$E(X) = 1/\lambda$.
$\text{Var}(X) = 1/\lambda^2$.
$M(t) = \lambda/(\lambda - t)$ para $t < \lambda$.

Distribución gama



$X \sim \text{gama}(\gamma, \alpha)$ con $\gamma > 0$ y $\alpha > 0$.
$f(x) = \dfrac{(\alpha x)^{\gamma-1}}{\Gamma(\gamma)} \alpha e^{-\alpha x}$ para $x > 0$.
$F(x) = 1 - e^{-\alpha x} \sum_{j=0}^{\gamma-1} (\alpha x)^j / j!$ para $x > 0$ y γ entero.
$E(X) = \gamma/\alpha$.
$\text{Var}(X) = \gamma/\alpha^2$.
$E(X^n) = \Gamma(\gamma + n)/(\Gamma(\gamma)\alpha^n)$.
$M(t) = [\alpha/(\alpha - t)]^\gamma$ para $t < \alpha$.
Cuando $\gamma = 1$ se obtiene la distribución $\exp(\alpha)$. Cuando γ es entero se conoce también con el nombre de distribución Erlang.

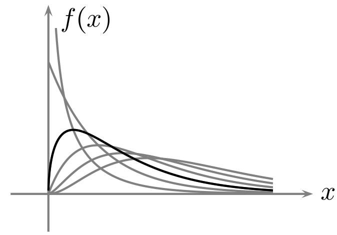
Distribución geométrica



$X \sim \text{geo}(p)$ con $p \in (0, 1)$.
$f(x) = p(1 - p)^x$ para $x = 0, 1, \ldots$
$E(X) = (1 - p)/p$.
$\text{Var}(X) = (1 - p)/p^2$.
$G(t) = p/(1 - (1 - p)t)$.
$M(t) = p/(1-(1-p)e^t)$ para $t < -\ln(1-p)$.

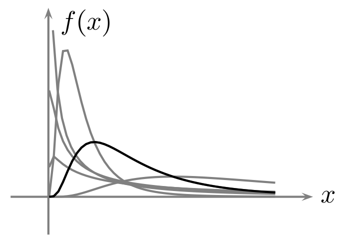
Distribución ji-cuadrada

$X \sim \chi^2(n)$ con $n > 0$.
$f(x) = \dfrac{(1/2)^{n/2}}{\Gamma(n/2)} x^{n/2-1} e^{-x/2}$ para $x > 0$.
$E(X) = n$.
$\text{Var}(X) = 2n$.
$M(t) = (1-2t)^{-n/2}$ para $t < 1/2$.



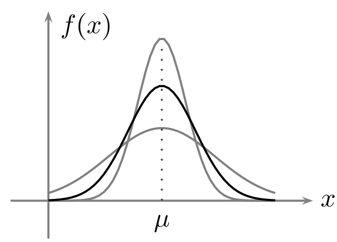
Distribución log normal

$X \sim \text{lognormal}(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ y $\sigma^2 > 0$.
$f(x) = \dfrac{1}{x\sqrt{2\pi\sigma^2}} \exp[-(\ln x - \mu)^2/2\sigma^2]$, $x > 0$.
$F(x) = \Phi((\ln x - \mu)/\sigma)$.
$E(X) = \exp(\mu + \sigma^2/2)$.
$E(X^n) = \exp(n\mu + n^2\sigma^2/2)$.
$\text{Var}(X) = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2)$.
Si $X \sim \text{N}(\mu, \sigma^2)$, entonces $e^X \sim \text{lognormal}(\mu, \sigma^2)$.



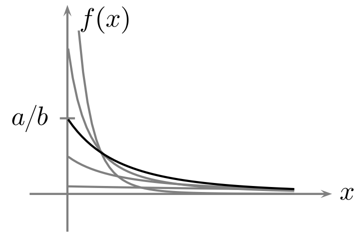
Distribución normal

$X \sim \text{N}(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ y $\sigma^2 > 0$.
$f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2}$.
$E(X) = \mu$.
$\text{Var}(X) = \sigma^2$.
$M(t) = \exp(\mu t + \sigma^2 t^2/2)$.
$\phi(t) = \exp(i\mu t - \sigma^2 t^2/2)$.
Cuando $\mu = 0$ y $\sigma^2 = 1$ se obtiene la distribución normal estándar, $\text{N}(0,1)$.



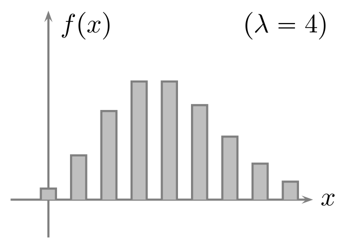
Distribución Pareto

$X \sim \text{Pareto}(a, b)$ con $a, b > 0$.
$f(x) = \dfrac{ab^a}{(b+x)^{a+1}}$ para $x > 0$.
$F(x) = 1 - [b/(b+x)]^a$ para $x > 0$.
$E(X) = b/(a-1)$ para $a > 1$.
$\text{Var}(X) = ab^2/[(a-1)^2(a-2)]$, $a > 2$.



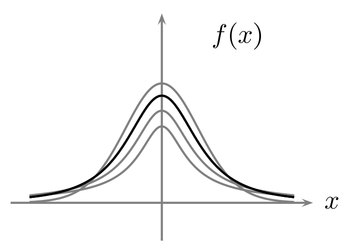
Distribución Poisson

$X \sim \text{Poisson}(\lambda)$ con $\lambda > 0$.
$f(x) = e^{-\lambda}\dfrac{\lambda^x}{x!}$ para $x = 0, 1, \ldots$
$E(X) = \lambda$.
$\text{Var}(X) = \lambda$.
$G(t) = e^{-\lambda(1-t)}$.
$M(t) = \exp[\lambda(e^t - 1)]$.

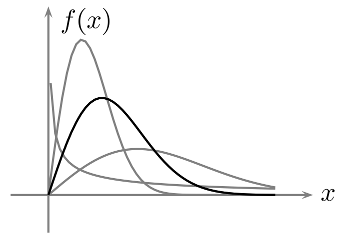


Distribución t

$X \sim \text{t}(n)$ con $n > 0$.
$f(x) = \dfrac{\Gamma(n+1/2)}{\sqrt{n\pi}\,\Gamma(n/2)}\,(1 + \dfrac{x^2}{n})^{-n-1/2}$.
$E(X) = 0$.
$\text{Var}(X) = n/(n-2)$ para $n > 2$.
$M(t)$ no existe para $t \neq 0$.
$\phi(t) = \exp(|t|)$.

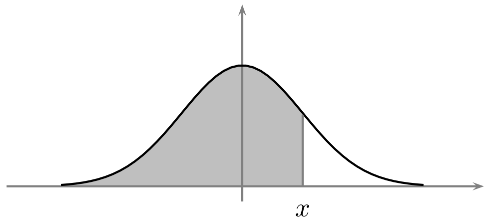


Distribución Weibull



$X \sim \text{Weibull}(r, \alpha)$ con $r, \alpha > 0$.
$f(x) = e^{-(\alpha x)^r} r\alpha^r x^{r-1}$ para $x > 0$.
$F(x) = 1 - e^{-(\alpha x)^r}$ para $x > 0$.
$E(X) = \Gamma(1 + 1/r)/\alpha$.
$\text{Var}(X) = [\Gamma(1 + 2/r) - \Gamma^2(1 + 1/r)]/\alpha^2$.

Tabla de la distribución normal estándar



$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8399
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Bibliografía

[1] Apostol T. M. (1974) *Mathematical analysis.* Addison–Wesley.

[2] Asmussen S. (2000) *Ruin probabilities.* World Scientific.

[3] Basu A. K. (2003) *Introduction to stochastic processes.* Alpha Science International Ltd.

[4] Beard R. E., Pentikäinen T., Pesonen E. (1984) *Risk theory.* Tercera edición. Chapman and Hall.

[5] Bladt M. (2005) *A review of phase-type distributions and their use in risk theory.* ASTIN Bulletin, Vol. 35, Núm. 1, pp. 145–161.

[6] Boland P. (2007) *Statistical and probabilistic methods in actuarial science.* Chapman & Hall / CRC.

[7] Bowers N. L. Jr., Gerber H. U., Hickman J. C., Jones D. A., Nesbitt C. J. (1997) *Actuarial mathematics.* The Society of Actuaries.

[8] Bühlmann H. (1970) *Mathematical methods in risk theory.* Springer–Verlag.

[9] Bühlmann H., Gisler A. (2005) *A course in credibility theory and its applications.* Springer.

[10] Burden J. D., Faires J. D. (2010) *Numerical analysis.* Brooks Cole.

[11] Daykin C. D., Pentikäinen T., Pesonen M. (1994) *Practical risk theory for actuaries.* Chapman and Hall.

[12] De Pril N. (1986) *On the exact computation of the aggregate claims distribution in the individual life model.* ASTIN Bulletin, Vol. 16, Núm. 2, pp. 109–112.

[13] Dickson D. C. M. (2005) *Insurance risk and ruin.* Cambridge University Press.

[14] Embrechts P., Kluppelberg C., Mikosch T. (1999) *Modelling extremal events for insurance and finance.* Springer.

[15] Gerber H. U. (1979) *An introduction to mathematical risk theory.* Monograph No. 8. S. S. Huebner Foundation for Insurance Education. Wharton School. University of Pennsylvania.

[16] Gerber H. U. (1988) *Mathematical fun with the compound binomial process.* ASTIN Bulletin, Vol. 18, Núm 2, pp. 161-168.

[17] Gut A. (2010) *Probability: a graduate course.* Springer.

[18] Harris B. (1966) *Theory of probability.* Addison–Wesley.

[19] Hoel P. G., Port S. C., Stone C. J. (1972) *Introduction to stochastic processes.* Hougthon Mifflin Company.

[20] Karlin S., Taylor H. M. (1975) *A first course in stochastic processes.* Academic Press.

[21] Karr A. F. (1993) *Probability.* Springer–Verlag.

[22] Kass R., Goovaerts M., Dhaene J., Denuit M. (2001) *Modern actuarial risk theory.* Kluwer Academic Publishers.

[23] Klugman S. A., Panjer H. H., Willmot G. E. (2008) *Loss models: from data to decisions.* Wiley.

[24] Li S., Lu Y., Garrido J. (2009) *A review of discrete-time risk models.* Rev. R. Acad. Cien. serie A. Mat. Vol. 103, Núm. 2, pp. 321-337.

[25] Melnikov A. (2003) *Risk analysis in finance and insurance.* Chapman & Hall/CRC.

[26] Norris J. (1997) *Markov chains.* Cambridge University Press.

[27] Panjer H. H. (1981) *Recursive evaluation of a family of compound distributions.* ASTIN Bulletin, Vol. 12, Núm 1, pp. 22–26.

[28] Panjer H. H. (editor) (1986) *Actuarial mathematics.* Proceedings of Symposia in Applied Mathematics, Vol. 35, AMS.

[29] Resnick S. (1992) *Adventures in stochastic processes.* Birkhäuser.

[30] Revuz D., Yor M. (1991) *Continuous martingales and Brownian motion.* Springer-Verlag.

[31] Rincón L. (2012) *Introducción a los procesos estocásticos.* Facultad de Ciencias, UNAM.

[32] Rolski T., Schmidli H., Teugels J. (1999) *Stochastic processes for insurance and finance.* John Wiley & Sons.

[33] Schiff J. L. (1999) *The Laplace transform: theory and applications.* Springer.

[34] Schmidli H. *Lecture notes on risk theory.* Manuscrito.

[35] Schmidli H. (2008) *Stochastic control in insurance.* Springer.

[36] Spiegel M. R. (1968) *Mathematical handbook of formulas and tables.* Schaum, McGraw-Hill.

[37] Stirzaker D. (2005) *Stochastic processes and models.* Oxford University Press.

[38] Teugels J. L., Sundt B. (editores) (2004) *Encyclopedia of actuarial science.* Wiley.

[39] Tudor C. (1994) *Procesos estocásticos.* Serie Textos, Sociedad Matemática Mexicana.

[40] Williams D. (1991) *Probability with martingales.* Cambridge University Press.

Índice analítico

El presente texto contiene material básico para un curso introductorio a ciertos temas de la teoría del riesgo aplicada a seguros. Está basado en el curso semestral de Teoría del Riesgo que el autor ha impartido a estudiantes de último semestre de la carrera de Actuaría en la Facultad de Ciencias de la UNAM. La intención es que este libro sea útil para los numerosos alumnos de las distintas escuelas de actuaría y matemáticas aplicadas de países de habla hispana y que también contribuya a apoyar el trabajo docente de sus profesores en esta importante área de las matemáticas aplicadas.

978-607-02-3773-7

9 786070 237737