



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

SIMULACIÓN DEL MODELO CRAMÉR-LUNDBERG

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

A C T U A R I O

P R E S E N T A:

ARRIGO COEN CORIA



**DIRECTOR DE TESIS:
M. EN C. GERARDO RUBIO HERNÁNDEZ
2009**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Simulación del modelo Cramér-Lundberg

Arrigo Coen Coria

2 de diciembre de 2009

Hoja de Datos del Jurado

1. Datos del alumno

Apellido paterno: Coen

Apellido materno: Coria

Nombre: Arrigo

Teléfono: 56 10 53 07

Universidad Nacional Autonoma de Mexico

Facultad de Ciencias

Carrera Actuarial

Número de cuenta 302735228

2. Datos del tutor

Grado M en C

Nombre(s) Gerardo

Apellido paterno Rubio

Apellido materno Hernandez

3. Datos del sinodal 1

Grado Dra

Nombre(s) María Asuncion Begoña

Apellido paterno Fernandez

Apellido materno Hernandez

4. Datos del sinodal 2

Grado Dr

Nombre(s) Luis Antonio

Apellido paterno Rincon

Apellido materno Hernandez

5. Datos del sinodal 3

Grado Dra

Nombre(s) Ruth Selene

Apellido paterno Fuentes

Apellido materno Garcia

6. Datos del sinodal 4

Grado Dr

Nombre(s) Ramses Humberto

Apellido paterno Mena

Apellido materno Chavez

7. Datos del trabajo escrito.

Título Simulación del modelo Cramer-Lundberg

Número de páginas 193 p

Año 2009

A Babbo y a Elí.

Agradecimientos

Probablemente uno espera que realizar una tesis sea un evento de suma importancia para el desarrollo académico y este caso no cayó en el complemento. Además, por ley, durante todos los momentos de su realización uno aprende algo nuevo. Por lo tanto quiero distribuir agradecimientos al conjunto de personas que me apoyaron durante este trabajo, sobre quienes he de decir que son mi familia, mas dada la condición de ser un gran número resaltaré a una muestra pequeña pero significativa.

Gracias Gerardo por toda la guía que me brindaste, la cual fue independiente de tu variable carga de trabajo. La medida en que me trataste hace que te considere como el amigo intersección director de tesis óptimo.

Gracias Laura por tu continuo apoyo el cual ha sido tu característica normal. Te agradezco cordialmente por ser esa combinación ideal exacta.

Estoy casi seguro de que para no simular cosas que no son propias de este trabajo solo me falta agregar... σ -álgebra.

Índice general

Introducción	1
1. El modelo de Cramér-Lundberg y la probabilidad de ruina	9
1.1. Introducción a conceptos	9
1.2. El modelo de Cramér-Lundberg	11
1.2.1. Resultados asintóticos	14
1.2.2. Probabilidad de ruina vista como caminata aleatoria	16
1.2.3. Condición de ganancia neta	19
1.3. Ecuaciones para la probabilidad de ruina	23
1.4. Fórmula de Pollaczek-Khinchin	32
1.5. Distribuciones de colas pesadas	34
1.5.1. Distribuciones subexponenciales	37
1.6. Cotas de Lundberg	39
1.7. Fórmula de Panjer	45
2. Simulación	51
2.1. Estimadores y su eficiencia	53
2.1.1. Método de Monte Carlo crudo y estimadores insesgados	53
2.1.2. Eficiencia de los estimadores	56
2.1.3. Error relativo acotado y eficiencia logarítmica	58
2.2. Eventos raros	61
2.2.1. Simulación de probabilidades de ruina vista como un evento raro	63
2.3. Métodos de reducción de varianza	64
2.3.1. Uso de covariables	65
2.3.2. Muestreo de importancia	70
2.3.3. Muestreo estratificado	77
2.3.4. Monte Carlo Condicionado	82
2.4. Algoritmos de simulación para la probabilidad de ruina en el modelo de C-L	84

2.4.1. Algoritmo I: Simulación a través de la fórmula de Pollaczek-Khinchine	84
2.4.2. Algoritmo II: Monte Carlo condicionado	85
2.4.3. Algoritmo III: Utilización de las estadísticas de orden	87
3. Implementación de algoritmos	95
3.1. Simulación a través de variables aleatorias uniformes	95
3.2. Resultados numéricos	97
3.2.1. Comparación de cotas y comportamiento asintótico	97
3.3. Simulación de algoritmos	100
3.3.1. Distribución de los sumandos: Pareto(1,2)	103
3.3.2. Distribución de los sumandos: Log-normal(-1.62,1.80)	110
3.3.3. Distribución de los sumandos: MPE(3)	117
Conclusiones	125
A. Tabla de distribuciones	129
B. Teoremas y resultados	131
B.1. Teoremas de probabilidad	131
B.2. Teoremas de análisis	133
C. Transformada de Laplace y convoluciones	135
C.1. Convoluciones geométricas	137
D. Distribuciones subexponenciales	141
D.1. Definición	141
D.2. Propiedades y resultados	141
E. Variación regular	149
E.1. Mezcla Pareto de distribuciones exponenciales	159
F. Verificación del método de las variables antitéticas al estimar el valor esperado de funciones monótonas	161
G. Programas utilizados	165
G.1. Lista de programas	165
G.2. Significado de parámetros	166
G.3. Programas	167
Bibliografía	183

Introducción

Una compañía de seguros representa un objeto de estudio muy interesante y complejo, el cual ha sido trabajado ampliamente por la teoría de riesgo. Existe una infinidad de factores a tomar en cuenta para las decisiones que la compañía debe tomar; por ejemplo, el tipo de siniestros que se está cubriendo, el monto que se deberá cobrar por concepto de primas, la existencia de un contrato de reaseguro, entre otros.

Todos estos factores son determinantes en la forma de operar de la compañía y las decisiones antes mencionadas representan distintos problemas que la misma debe enfrentar. Dada toda esta complejidad, nosotros solo nos enfocaremos en plantear un modelo que emule el comportamiento de la reserva y con él poder tener un indicador de la solvencia de la compañía aseguradora.

En este trabajo nos enfocaremos en el problema de que la reserva de la compañía se haga negativa, lo que denominaremos por **ruina**. La reserva es función de distintos factores, es decir

$$U(\text{tiempo}) = U(\text{tiempo} ; \text{Capital inicial, Monto de primas, Magnitud de siniestros, Frecuencia de siniestros, Número de asegurados, } \dots),$$

donde U representa a la reserva.

El planteamiento más básico (sin ser descabellado) sería plantear a la reserva en el tiempo $t \in \mathbb{R}_+$ como

$$U(t) = \text{Capital inicial} + \text{Entradas}(t) - \text{Salidas}(t),$$

donde $\text{Entradas}(t)$ y $\text{Salidas}(t)$ representan las entradas y las salidas al momento t respectivamente.

Como ya hemos mencionado existen muchísimas consideraciones, por lo cual, nosotros nos enfocaremos en una compañía aseguradora que cumpla con las siguientes hipótesis:

- Las entradas serán deterministas. Esta hipótesis la hacemos pensando que nuestra compañía cubre a suficientes asegurados como para que las entradas que recibe por concepto de primas se puedan considerar como deterministas; es decir, podemos escribir a las entradas como

$$\text{Entradas}(t) = ct,$$

donde c representa el monto de primas que se percibe por unidad de tiempo.

- Las salidas durante lapsos separados del mismo tamaño son independientes e idénticamente distribuidas. Esta suposición la hacemos no sólo por el hecho de que matemáticamente facilita los cálculos sino también para que se aplique al seguro de daños; es decir estamos pensando en seguros intemporales, los cuales no cambiarán sus características durante el tiempo. El modelo que estamos planteando no es útil para los seguros de vida ya que esta hipótesis no se cumple.¹

Dada la última hipótesis, las salidas de la compañía quedan forzosamente determinadas por un proceso Poisson, ya que ésta suposición encierra que la distribución de los siniestros tenga pérdida de memoria. Esto nos lleva a utilizar el modelo Cramér-Lundberg el cual plantea que la reserva de una compañía de seguros al tiempo t puede ser modelada por

$$U(t) := x + ct - S(t) \quad \text{para } t \in \mathbb{R}_+$$

donde x representa el capital inicial, c representa al monto obtenido por concepto de primas por unidad de tiempo y $S(t)$ es el monto a pagar por concepto de reclamos, definido por

$$S(t) := \sum_{k=1}^{N(t)} \xi_k,$$

donde $N(t)$ es un proceso Poisson y ξ_1, ξ_2, \dots son variables aleatorias que representan los montos de los siniestros ocurridos.

En este trabajo nos enfocaremos en siniestros cuya distribución sea de cola pesada, ya que consideramos fenómenos de escasa ocurrencia cuya aparición causa cambios drásticos en el entorno. Este tipo de fenómenos tiene la característica de que aún cuando su probabilidad es pequeña, su ocurrencia es lo

¹En los seguros de vida el tiempo cambia la distribución de los siniestros; por ejemplo, no es igual el seguro de vida para un niño que para un adulto.

suficientemente importante para tomarlos en cuenta.

El pedir que la reserva sea positiva es fundamental para el correcto funcionamiento de la compañía de seguros por lo que existen diferentes normativas, creadas por los organos regulatorios de las compañías de seguros que estipulan esta condición como obligatoria. Por esta razón plantearemos a la ruina como el evento de que la reserva se haga negativa

$$\text{Ruina} = \{U(t) < 0 \text{ para alguna } t > 0\}$$

y estamos interesados en cuál es la probabilidad de que ésta ocurra. En nuestro estudio planteamos a la probabilidad de ruina como función del capital inicial, lo cual es equivalente a denotar a la probabilidad de ruina por

$$\psi(x) := \mathbb{P}[U(t) < 0 \text{ para algún } t > 0 \mid U(0) = x].$$

Nos enfocaremos entonces en el estudio del comportamiento de $\psi(x)$, en especial, para que valores de x ésta se hace pequeña. Al observar el comportamiento de $\psi(x)$ para valores grandes de x surge el problema de que $\psi(x)$ es pequeña lo cual nos hará abordar la teoría de eventos raros.

Esta última ecuación es difícil de manejar por todos los factores considerados en el proceso $U(t)$. Es entonces necesario replantearla en términos más prácticos; es decir, buscamos formas de plantear a la probabilidad de ruina tales que nos faciliten su estudio. Este tipo de replanteamientos los haremos por medio de reescribir a la probabilidad de ruina a través de distintas ecuaciones entre las cuales destaca ponerla en términos de una ecuación integro-diferencial de la forma

$$\delta'(x) = \frac{\lambda}{c}\delta(x) - \frac{\lambda}{c} \int_0^x \delta(x-y)dF(y),$$

donde $\delta(x) = 1 - \psi(x)$. Esto sigue sin resolver el problema de encontrar una expresión analítica para la probabilidad de ruina, ya que sólo en el caso de que la distribución de los siniestros sea exponencial tendremos una solución explícita.

Otra forma de expresar a la probabilidad de ruina es a través de la fórmula de Pollaczek-Khinchin

$$\delta(x) = \frac{\rho}{1+\rho} \sum_{m=0}^{\infty} (1+\rho)^{-m} F_I^{*m}(x) \quad \text{para } x > 0$$

la cual tiene matemáticamente la virtud de plantearla en términos de una suma aleatoria ya que la ecuación anterior equivale a

$$\psi(x) = \mathbb{P}[S_M > x],$$

donde

$$S_M = \sum_{i=1}^M X_i \quad M \sim geo\left(\frac{\rho}{1+\rho}\right) \quad X \sim F_I(x) = \frac{1}{\mathbb{E}[\xi]} \int_0^x \bar{F}_\xi(x) dx.$$

Estas ecuaciones a pesar de que no expresan de manera analítica a la probabilidad de ruina, serán la base para poder aproximar su valor.

Dado que no tenemos expresiones analíticas para la probabilidad de ruina es conveniente analizar su comportamiento a través de diversas cotas. Entre ellas destaca la cota de Lundberg

$$\psi(x) \leq e^{-rx}.$$

La desventaja es que esta cota no existe cuando la distribución de los reclamos es de cola pesada. En cambio podemos utilizar tanto la cota superior como la cota inferior de Panjer en estos casos, las cuales se basan en discretizar la distribución de los siniestros. La desventaja de las cotas de Panjer es que para valores grandes del capital inicial requieren de un número inmenso de operaciones lo cual imposibilita su cálculo.

Al no lograr tener una expresión analítica con la cual calcular la probabilidad de ruina se requieren entonces métodos que nos aproximen su valor. Para lograr estas aproximaciones utilizaremos la simulación por el método de Monte Carlo como principal herramienta. Este método se basa en simular R variables aleatorias Z tales que $\mathbb{E}[Z] = \zeta$, donde ζ es el valor que estamos buscando y utilizar a

$$\hat{\zeta} = \sum_{i=1}^R Z_i$$

como estimador de ζ .

Surgen entonces las preguntas:

¿Que tan eficientes son nuestros estimadores?

Para responderla es importante observar el intervalo de confianza de nuestros estimadores. Por el teorema central del límite tenemos que para R grande el intervalo de confianza alrededor de $\hat{\zeta}$ al cual ζ pertenece con probabilidad $(1 - \alpha)$ está dado por

$$\left(\hat{\zeta} - z_{1-\alpha/2} \frac{\sigma_{\hat{\zeta}}}{\sqrt{R}}, \hat{\zeta} + z_{1-\alpha/2} \frac{\sigma_{\hat{\zeta}}}{\sqrt{R}} \right),$$

donde $z_{1-\alpha}$ es el cuantil $(1 - \alpha)$ de una distribución normal estándar y $\sigma_{\hat{\zeta}}$ representa la desviación estándar de nuestro estimador. Por lo tanto una cantidad que se puede tomar como medida de dispersión es

$$\frac{S/\sqrt{R}}{\hat{\zeta}},$$

donde S representa al estimador insesgado de la varianza. Esta ecuación para R grande converge al error relativo del estimador $\hat{\zeta}$, definido por

$$\varepsilon = \frac{\sqrt{\text{Var}(\hat{\zeta})}}{\mathbb{E}[\hat{\zeta}]}.$$

Entonces queremos lograr que el error relativo de nuestros estimadores sea pequeño para así lograr que nuestro intervalo de confianza se reduzca.

¿Cómo se puede mejorar la simulación, si se requiere que los estimadores cumplan con ciertos criterios de error?

Para responderla veamos qué problemas existen al simular la probabilidad de ruina por medio del método de Monte Carlo:

1. El método de Monte Carlo no se puede aplicar de manera directa para aproximar la probabilidad de ruina utilizando el proceso $U(t)$, ya que implicaría simular un proceso durante un tiempo infinito², lo cual es computacionalmente imposible; es decir, si uno tuviera un algoritmo que simulara el proceso de ruina, cada simulación podría ser detenida sólo si existiera la ruina pues en caso contrario nada garantizaría que al detener la simulación en un tiempo finito la ruina no ocurriera en un tiempo posterior.

²Esto pasa porque el evento ruina lo definimos para un horizonte infinito.

2. Supongamos que podemos simular de manera adecuada la probabilidad de ruina pero que el resultado de nuestras simulaciones solo son cero si en la simulación hubo ruina o uno en caso de no haber; es decir, el resultado de nuestras simulaciones son $Z \sim Ber(\psi(x))$. Como ya hemos mencionado queremos conocer el comportamiento de $\psi(x)$ cuando toma valores pequeños lo cual implica que estamos trabajando con probabilidades pequeñas, entonces es importante tener un intervalo de confianza similarmente pequeño. El tamaño de este intervalo tiene una relación directa con el error relativo, el cual está dado por

$$\varepsilon = \frac{\sqrt{\psi(x)(1 - \psi(x))/R}}{\psi(x)}.$$

Por como fue planteada la probabilidad de ruina tenemos que $\psi(x) \rightarrow 0$ cuando $x \rightarrow \infty$, lo cual implica

$$\varepsilon \approx \frac{1}{\sqrt{R\psi(x)}} \rightarrow \infty \quad \text{cuando } x \rightarrow \infty.$$

Éste es un grave problema ya que nosotros queremos estimadores cuyo intervalo de confianza sea pequeño.

Para resolver el primer problema utilizaremos la fórmula de Pollaczek-Khinchin la cual nos permite aproximar la probabilidad de ruina por medio de la simulación de una suma geométrica, lo que es computacionalmente sencillo.

Para resolver el segundo problema daremos una breve introducción a la simulación de eventos raros donde definiremos un evento raro como $A(x)$, el cual cumple que $\mathbb{P}[A(x)] \rightarrow 0$ cuando $x \rightarrow \infty$. Este tipo de teoría nos es necesaria ya que el comportamiento de la probabilidad de ruina cumple $\psi(x) \rightarrow 0$ cuando $x \rightarrow \infty$. Expondremos los diversos problemas que este tipo de simulación plantea así como las distintas técnicas para hacerles frente.

Existen dos caminos para reducir el error relativo de nuestras estimaciones, ya sea generando un mayor número de simulaciones o reduciendo la varianza de nuestro estimador. En el planteamiento de los eventos raros, es inapropiado el intentar hacer un mayor número de simulaciones para lograr reducir este tipo de error. Se requieren entonces métodos de reducción de varianza para lograr simular de manera adecuada a la probabilidad de ruina.

Estos métodos utilizan información a priori del valor estimado para así mejorar las simulaciones.

Existen diversos criterios para mostrar qué tan eficientes son los estimadores, nosotros nos enfocaremos en los que impliquen un buen comportamiento de los errores relativos. Por ejemplo, se dice que un estimador $\widehat{\zeta}(x)$ tiene error relativo acotado si

$$\limsup_{x \rightarrow \infty} \frac{Var \widehat{\zeta}(x)}{\zeta(x)^2} < \infty \quad (1)$$

el cual nos dice que el error relativo del estimador nunca pasa de cierta cantidad³, lo cual es una propiedad sumamente apreciada.

Es difícil obtener estimadores con error relativo acotado, ya que (en caso de existir) las hipótesis que requieren son muy restrictivas. Por lo tanto en este trabajo estudiaremos un algoritmo que cumple con un criterio menos restrictivo, eficiencia logarítmica, el cual pide que los estimadores cumplan

$$\limsup_{x \rightarrow \infty} \frac{Var \widehat{\zeta}(x)}{\zeta(x)^{2-\varepsilon}} = 0$$

Se puede observar que éste es un criterio menos estricto que el anterior (de hecho error relativo acotado implica eficiencia logarítmica).

Al final de este trabajo exhibiremos los resultados que hemos obtenido al comparar nuestras estimaciones con las diferentes cotas y aproximaciones que existen para la probabilidad de ruina.

³Hay que observar el parecido entre esta ecuación y el error relativo.

Capítulo 1

El modelo de Cramér-Lundberg y la probabilidad de ruina

1.1. Introducción a conceptos

En este capítulo daremos una breve introducción a la notación y terminología que se utilizará durante este trabajo al igual que se plantearán distintos modelos junto con sus interpretaciones e implicaciones. La motivación de nuestro estudio es el poder ponderar la solvencia de una compañía de seguros. Para poder estimar esta solvencia un factor importante es el comportamiento de la reserva de la compañía, ya que para empezar si ésta se hace cero o negativa es equivalente a que la compañía se arruine o que por lo menos entre en una etapa de ruina. Además existen normatividades las cuales exigen que las compañías de seguros al comenzar sus operaciones tengan cierto capital (capital inicial) para poder garantizar, con cierto grado de confianza, los pagos a los asegurados que sufran un siniestro. Por lo tanto tenemos la siguiente definición.

Definición 1.1.1. *Denotaremos por $\{U(t)\}_{t \geq 0}$ a la reserva de la compañía al tiempo t . Para $t = 0$ nombraremos a $x = U(0)$ como la reserva inicial de la compañía.*

Nos enfocaremos en estudiar el cambio que experimenta la reserva a lo largo del tiempo visto como un proceso aleatorio, que depende de la reserva con la que inicia. Dicho en otras palabras, modelaremos el proceso de la reserva como un proceso estocástico que es función del capital inicial de la compañía. Se estudiará entonces el capital inicial que la compañía debería tener para poder hacer que éste junto con las primas recibidas logren hacer, con cierto grado de confianza, frente a los reclamos efectuados.

El grado de confianza para garantizar el pago a los asegurados lo mediremos por medio de la probabilidad de ruina, lo cual nos otorga una noción de qué tan solvente es la compañía. Por ejemplo si conocemos cuál es la probabilidad de ruina de cierta compañía de seguros podríamos analizar si conviene o no invertir en ella (en el caso de ser inversionistas) o qué tipo de cambios hacer para disminuir esta probabilidad (si es que ya formamos parte de la compañía). Lo anterior nos lleva a la siguiente definición.

Definición 1.1.2. Denotaremos por $\psi(x)$ a la probabilidad de ruina (a horizonte infinito) cuando se tiene un capital inicial x , definida por

$$\psi(x) := \mathbb{P} \left[\inf_{t \geq 0} U(t) < 0 \right] = \mathbb{P} \left[\inf_{t \geq 0} U(t) < 0 \mid U(0) = x \right]. \quad (1.1)$$

Observación 1.1.1. La razón por la cual se le nombra como probabilidad de ruina a horizonte infinito es que existen otros estudios donde se considera la probabilidad de que la reserva se haga negativa durante un intervalo finito la cual se define como la probabilidad de ruina antes del tiempo $T \in \mathbb{R}_+$ o probabilidad de ruina en horizonte finito y se determina por

$$\psi(x, T) := \mathbb{P} \left[\inf_{0 \leq t \leq T} U(t) < 0 \right]. \quad (1.2)$$

Como este trabajo se centra en el estudio de la probabilidad de ruina a horizonte infinito, tendremos la convención de nombrar como probabilidad de ruina a la probabilidad de ruina en horizonte infinito.

Definición 1.1.3. Definimos el tiempo de ruina $\tau(x)$, como el primer momento cuando la reserva se vuelve negativa dado un capital inicial x ; es decir

$$\tau(x) := \inf \{ t > 0 \mid U(t) < 0 \}. \quad (1.3)$$

Entonces $\tau := \tau(x)$ es una v.a. la cual, dependiendo de las propiedades del modelo, podría tomar el valor ∞ con probabilidad positiva; es decir, que el rango de τ es $[0, \infty]$.

Utilizando estas definiciones, tenemos la siguiente expresión para la probabilidad de ruina

$$\psi(x) = \mathbb{P} [\tau < \infty] \quad (1.4)$$

Ésta última expresión nos permite exhibir a la probabilidad de ruina en términos de la variable aleatoria τ (que es el formato más usual para expresar probabilidades), la cual utilizaremos ampliamente durante nuestro estudio.

1.2. El modelo de Cramér-Lundberg

Estudiaremos un modelo básico para una compañía de seguros en donde el pago de las primas y el capital inicial serán deterministas mientras que el número de siniestros y el monto de cada uno será aleatorio. Estas suposiciones las podemos hacer dado que estamos pensando que tenemos una gran cantidad de asegurados, por lo cual, aunque una parte de éstos deje el contrato de seguro (ya sea por algún siniestro que le ocurra o por algún otro motivo) esta disminución en el pago de las primas no será significativa y por lo tanto se considerará constante. El capital inicial será nuestra variable independiente, para ver qué cambios existen en la probabilidad de ruina al ir cambiando este valor. Estudiaremos el siguiente modelo:

$$U(t) := x + ct - S(t) \quad \text{para } t \in \mathbb{R}_+ \quad (1.5)$$

En donde se tendrá que $U(t)$ es el capital de la empresa al tiempo $t \in \mathbb{R}_+$, x es el capital inicial, c es el monto obtenido por concepto de primas por unidad de tiempo y $S(t)$ es el monto que la compañía deberá pagar por concepto de siniestros. Para el caso particular del modelo Cramér-Lundberg se tendrá que $S(t)$ es un proceso de la forma

$$S(t) := \sum_{k=1}^{N(t)} \xi_k = \sum_{k=1}^{\infty} \xi_k 1_{\{[0,t]\}} T_k \quad t \geq 0 \quad (1.6)$$

que cumple con los siguientes supuestos:

Proceso de reclamaciones: El tamaño de las reclamaciones es determinado por $\{\xi_k\}_{k \in \mathbb{N}}$, las cuales son variables aleatorias independientes idénticamente distribuidas (v.a.i.i.d.) positivas con distribución F y que cumplen con $\mu := \mathbb{E}[\xi] < \infty$.

Tiempos de reclamación: El proceso $\{T_k\}_{k \in \mathbb{N}}$ corresponde al tiempo en el que ocurren las reclamaciones con relación al tiempo cero.

Arribos de reclamación: Serán el número de reclamaciones en el intervalo $[0,t]$ y se definen como $N(t) = \sup\{n \geq 1 \mid T_n \leq t\}$ para $t \geq 0$.

Tiempos inter-arribos: Se definen como $Y_1 := T_1$ y $Y_k := T_k - T_{k-1}$ para $k = 2, 3, \dots$ y son v.a.i.i.d. $\exp(\lambda)$ los cuales contarán el lapso transcurrido entre una y otra de las reclamaciones¹.

¹La motivación principal para utilizar v.a. exponenciales es por su propiedad de pérdida de memoria ya que se considera que la ocurrencia de un siniestro no incita ni evita a otros siniestros a ocurrir.

Condición de ganancia neta: Se denominará a $\rho = \frac{c}{\lambda\mu} - 1 > 0$ como la condición de ganancia neta. Se necesitará que ρ sea positiva para que la probabilidad de ruina sea distinta de uno (esto se demostrará más adelante en el Teorema 1.2.4) Una manera de interpretarla es como el cargo extra que se le realiza a la prima, por medio de $c = \lambda\mu(1 + \rho)$ (véase la observación 1.2.2).

Independencia entre procesos: Tendremos la independencia entre los procesos $\{\xi_k\}_{k \in \mathbb{N}}$ y $\{Y_k\}_{k \in \mathbb{N}}$.

Las hipótesis anteriores son la base del modelo con el cual trabajaremos, en ellas se encuentra el punto de partida para la teoría que desarrollaremos. Estas hipótesis han sido estudiadas durante más de un siglo, su primera presentación fue en la tesis doctoral de Filip Lundberg la cual fue terminada en el año de 1903 para después ser formalizada, al utilizar la teoría de los procesos estocásticos, por Harald Cramér en 1930. A la fecha sigue siendo motivo de numerosos estudios.

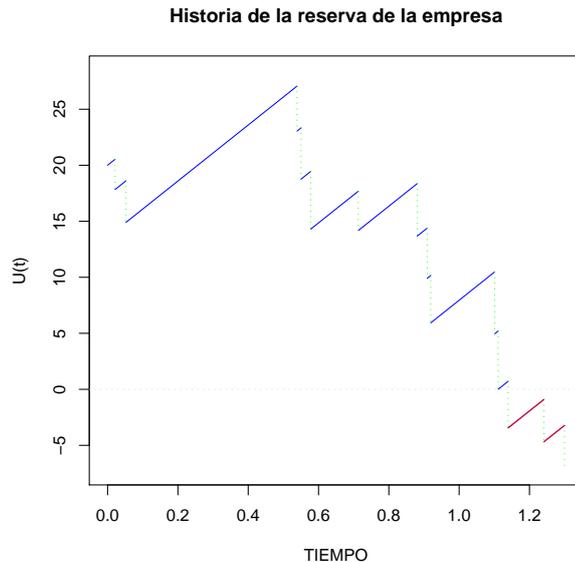


Figura 1.1: Posible trayectoria del proceso $\{U_t\}$.

En la figura 1.1 se muestra una posible trayectoria del proceso $\{U_t\}$, donde con un capital inicial de 20, en el tiempo $t \approx 1.139498$ ocurre la ruina.

Observación 1.2.1. Una consecuencia de las definiciones anteriores es que el proceso $\{N_t\}_{t \geq 0}$ es un proceso Poisson homogéneo con intensidad $\lambda > 0$ (la

demostración puede ser consultada en Rolski(1998)[32] página 156), lo cual implica que

$$\mathbb{P}[N(t) = k] = e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad k = 0, 1, 2, \dots \quad (1.7)$$

Con la notación anterior se obtiene el siguiente resultado.

Proposición 1.2.1.

$$\mathbb{E}[U(t)] = x + ct - \lambda\mu t \quad (1.8)$$

Demostración. Tenemos que $\mathbb{E}[U(t)] = x + ct - \mathbb{E}[S(t)]$

$$\begin{aligned} \mathbb{E}[S(t)] &= \sum_{n=0}^{\infty} \mathbb{E}[S(t) \mid N(t) = n] \mathbb{P}[N(t) = n] \\ &= \sum_{n=0}^{\infty} \mathbb{E} \left[\sum_{i=1}^{N(t)} \xi_i \mid N(t) = n \right] \mathbb{P}[N(t) = n] \\ &= \sum_{n=0}^{\infty} \mathbb{E} \left[\sum_{i=1}^n \xi_i \right] \mathbb{P}[N(t) = n] \\ &= \mu \sum_{n=0}^{\infty} n \mathbb{P}[N(t) = n] \\ &= \mu \mathbb{E}[N(t)] \end{aligned}$$

de esta última igualdad se sigue el resultado, ya que $\mathbb{E}[N(t)] = \lambda t$ por ser un proceso Poisson homogéneo. \square

Observación 1.2.2. Al utilizar la ley fuerte de los grandes números (ver el Apéndice B.1.7), el resultado anterior nos permite ver que

$$\frac{\mathbb{E}[U(t)]}{t} \xrightarrow{c.s} c - \lambda\mu, \quad (1.9)$$

lo cual nos sugiere que una condición deseada para que la compañía no tenga un alta probabilidad de ruina es $c - \lambda\mu > 0$ la cual es justamente la condición de ganancia neta.

Proposición 1.2.2.

$$\mathbb{E}[S(t)^2] = \mathbb{E}[N(t)] \mathbb{E}[\xi^2] + \mathbb{E}[N(t)(N(t) - 1)] \mathbb{E}[\xi]^2. \quad (1.10)$$

Demostración. Condicionando sobre el valor de $N(t)$ tenemos que

$$\begin{aligned}
\mathbb{E} [S(t)^2] &= \sum_{n=0}^{\infty} \mathbb{E} \left[\left(\sum_{i=0}^{N(t)} \xi_i \right)^2 \mid N(t) = n \right] \mathbb{P} [N(t) = n] \\
&= \sum_{n=0}^{\infty} \mathbb{E} \left[\left(\sum_{i=0}^n \xi_i \right)^2 \mid N(t) = n \right] \mathbb{P} [N(t) = n] \\
&= \sum_{n=0}^{\infty} \mathbb{E} \left[\left(\sum_{i=0}^n \xi_i \right)^2 \right] \mathbb{P} [N(t) = n] \\
&= \sum_{n=0}^{\infty} \mathbb{E} \left[\sum_{i=0}^n \xi_i^2 + 2 \sum_{1 \leq i < j \leq n} \xi_i \xi_j \right] \mathbb{P} [N(t) = n] \\
&= \sum_{n=0}^{\infty} n \mathbb{E} [\xi^2] \mathbb{P} [N(t) = n] + \sum_{n=2}^{\infty} n(n-1) \mathbb{E} [\xi]^2 \mathbb{P} [N(t) = n] \\
&= \mathbb{E} [N(t)] \mathbb{E} [\xi^2] + \mathbb{E} [N(t)(N(t) - 1)] \mathbb{E} [\xi]^2
\end{aligned}$$

□

Una relación que existe entre las variables $N(t)$ y T_n es

$$\{N(t) = n\} = \{T_n \leq t < T_{n+1}\}, \quad (1.11)$$

donde el lado izquierdo de esta igualdad expresa que al tiempo t se tienen n reclamaciones, mientras que el lado derecho expresa que el valor de t se encuentra entre la reclamación n y la $n + 1$ lo cual es equivalente. Esta expresión va a ser muy útil durante el resto de este capítulo.²

1.2.1. Resultados asintóticos

En esta sección presentaremos distintos resultados asintóticos de los procesos que hemos enunciado. También exhibiremos distintas convergencias, las cuales revelarán el comportamiento asintótico de los procesos con los que trabajaremos.

Teorema 1.2.1. *Para el proceso Poisson homogéneo $\{N(t)\}_{t \in \mathbb{R}}$ se tiene la siguiente convergencia en probabilidad*

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \lambda. \quad (1.12)$$

²Esta igualdad será de importancia ya que más adelante nos mostrará que los momentos en los que puede ocurrir la ruina son justamente los tiempos $\{T_k\}_{k \in \mathbb{N}}$

Para poder dar la demostración de este teorema, necesitaremos el siguiente resultado auxiliar.

Lema 1.2.1. *Sea $\{Z_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias tales que existe una v.a. Z para la cual $Z_n \xrightarrow{c.s.} Z$ cuando $n \rightarrow \infty$ y sea $\{M(t)\}_{t \geq 0}$ un proceso estocástico el cual toma valores en los enteros, tal que $M(t) \xrightarrow{c.s.} \infty$ cuando $t \rightarrow \infty$. Si M y Z_n están definidos en el mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ se tendrá que*

$$Z_{M(t)} \xrightarrow[t \rightarrow \infty]{c.s.} Z$$

Demostración. Sean

$$\Omega_1 = \{\omega \in \Omega \mid M(t, \omega) \rightarrow \infty\} \quad \Omega_2 = \{\omega \in \Omega \mid Z_n(\omega) \rightarrow Z(\omega)\}.$$

Por hipótesis tenemos que $\mathbb{P}[\Omega_1] = \mathbb{P}[\Omega_2] = 1$, se sigue que $\mathbb{P}[\Omega_1 \cap \Omega_2] = 1$, por lo tanto

$$\mathbb{P}[\{\omega \in \Omega \mid Z_{M(t, \omega)}(\omega) \rightarrow Z(\omega)\}] \geq \mathbb{P}[\Omega_1 \cap \Omega_2] = 1,$$

lo cual demuestra el lema. \square

Demostración del Teorema 1.2.1. Motivados por la ecuación (1.11) escribimos para $n \in \mathbb{N}$ la siguiente serie de desigualdades

$$\frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{T_{N(t)} + 1}{N(t) + 1} \frac{N(t) + 1}{N(t)}. \quad (1.13)$$

Utilizando la ley fuerte de los grandes números para las variables aleatorias Y_n tenemos $n^{-1}Y_n \xrightarrow{c.s.} \lambda^{-1}$. Entonces se tendrá que $N(t) \xrightarrow{c.s.} \infty$ cuando $t \rightarrow \infty$. Aplicando el Lema 1.2.1 con $Z_n = T_n/n$ y $M = N$ tenemos que

$$\frac{T_{N(t)}}{N(t)} \xrightarrow{c.s.} \frac{1}{\lambda}, \quad (1.14)$$

lo cual concluye la prueba al utilizar esta última ecuación en ambos lados de (1.13). \square

Observación 1.2.3. *Del Teorema 1.2.1 también se puede concluir que el proceso $N(t) \rightarrow \infty$ cuando $t \rightarrow \infty$, ya que en caso contrario el límite (1.12) tendería a cero.*

Teorema 1.2.2 (Comportamiento asintótico de $S(t)$). *El proceso $S(t) = \sum_{i=1}^{N(t)} \xi_i$ con $\mu := \mathbb{E}[\xi]$ y $\lambda := \mathbb{E}[Y]$ cumple que*

$$\lim_{t \rightarrow \infty} \frac{S(t)}{t} \stackrel{c.s.}{=} \lambda\mu \quad (1.15)$$

Demostración. Se sigue que

$$\frac{S(t)}{t} = \frac{S(t)}{N(t)} \frac{N(t)}{t}, \quad (1.16)$$

si definimos a Ω_1 y Ω_2 como

$$\Omega_1 = \{\omega \mid N(t)/t \rightarrow \lambda\} \quad \Omega_2 = \{\omega \mid S(t)/N(t) \rightarrow \mu\},$$

se obtendrá por el Teorema 1.2.1 que $\mathbb{P}[\Omega_1] = 1$.

Además como $N(t) \xrightarrow{c.s.} \infty$ y aplicando la ley fuerte de los grandes números a $S_n = X_1 + X_2 + \dots + X_n$ para $n \geq 1$ tenemos que

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} \stackrel{c.s.}{=} \mu. \quad (1.17)$$

Entonces aplicando el Lema 1.2.1 con $Z_n = S_n/n$ y $M = N$ tenemos que $\mathbb{P}[\Omega_2] = 1$, entonces $\mathbb{P}[\Omega_1 \cap \Omega_2] = 1$, por lo tanto al utilizar la ecuación (1.16) el resultado queda demostrado. \square

1.2.2. Probabilidad de ruina vista como caminata aleatoria

Por cómo se construyó el proceso $U(t)$, la ruina sólo podrá ocurrir en los tiempos $t = T_n$ para algún $n \geq 1$ (Tiempos de reclamación)³, ya que $U(t)$ incrementa linealmente cuando $t \in [T_n, T_{n+1})$ para $n \in \mathbb{N}$.

Por lo anterior se obtiene que el problema de la ruina se puede replantear como

$$\begin{aligned} \psi(x) &= \mathbb{P} \left[\inf_{t \geq 0} U(t) < 0 \right] \\ &= \mathbb{P} \left[\inf_{n \geq 1} U(T_n) < 0 \right] \\ &= \mathbb{P} \left[\inf_{n \geq 1} [x + cT_n - S(T_n)] < 0 \right] \\ &= \mathbb{P} \left[\inf_{n \geq 1} \left[x + cT_n - \sum_{i=1}^n \xi_i \right] < 0 \right]. \end{aligned}$$

Esto nos permite abordar el cálculo de la probabilidad de ruina mediante el estudio de las variables aleatorias

$$Z_n := \xi_n - cY_n, \quad R_0 = 0 \quad \text{y} \quad R_n := Z_1 + \dots + Z_n \quad n \geq 1. \quad (1.18)$$

³Esto pasa porque el resto de la ecuación (1.5) es determinista y creciente

Por lo tanto tenemos la siguiente expresión para la probabilidad de la ruina

$$\psi(x) = \mathbb{P} \left[\inf_{n \geq 1} (-R_n) < -x \right] = \mathbb{P} \left[\sup_{n \geq 1} R_n > x \right]. \quad (1.19)$$

Como las sucesiones de v.a. $\{Y_i\}$ y $\{\xi_i\}$ son i.i.d. (cada una con los miembros de su propia sucesión) y las dos son mutuamente independientes, entonces $\{R_n\}_{n \in \mathbb{N}}$ es una caminata aleatoria; es decir, el proceso $\{R_n\}_{n \in \mathbb{N}}$ cumple con que $R_n = R_{n-1} + Z_n$. Esto último nos permite expresar a $\psi(x)$ como la probabilidad de que la caminata aleatoria $\{R_n\}_{n \in \mathbb{N}}$ entre en el conjunto $\{r \in \mathbb{R} : r > x\}$ para alguna $n \in \mathbb{N}$. Más adelante en la Sección 1.3 retomaremos esta idea para poder expresar a la probabilidad de ruina en términos de una ecuación integro-diferencial.

Las siguientes definiciones tienen como finalidad demostrar por qué se necesita que $\rho > 0$ para que la probabilidad de ruina no sea idénticamente uno. Necesitaremos entonces definir el primer tiempo de entrada de la caminata aleatoria $\{R_n\}$ al intervalo $(0, \infty)$ como:

Definición 1.2.1. *El primer tiempo de entrada de la caminata aleatoria $\{R_n\}$ al intervalo $(0, \infty)$ se define como*

$$\nu^+ := \min\{n : R_n > 0\}, \quad (1.20)$$

sea $\nu^+ = \infty$ si $R_n \leq 0$ para todo $n \in \mathbb{N}$, al cual denotaremos por el primer salto positivo de $\{R_n\}$. De manera similar definimos el primer tiempo de entrada de la caminata aleatoria $\{R_n\}$ al intervalo $(-\infty, 0]$ como

$$\nu^- = \min\{n : R_n \leq 0\}, \quad (1.21)$$

sea $\nu^- = \infty$ si $R_n > 0$ para todo $n \in \mathbb{N}$, al cual denotaremos por el primer salto negativo de $\{R_n\}$.

A continuación se demostrará que el comportamiento de ν^+ y ν^- depende de como sea $\mathbb{E}[Z]$, ya que se tendrán tres casos dependiendo de si $\mathbb{E}[Z]$ es estrictamente negativa, cero o estrictamente positiva.

Observación 1.2.4. *Otra manera de expresar a ν^+ es*

$$\{\nu^+ = k\} = \{R_1 \leq 0, R_2 \leq 0, \dots, R_{k-1} \leq 0, R_k > 0\}. \quad (1.22)$$

De manera similar podemos expresar a ν^- como

$$\{\nu^- = k\} = \{R_1 > 0, R_2 > 0, \dots, R_{k-1} > 0, R_k \leq 0\}. \quad (1.23)$$

Definición 1.2.2. Definimos la sucesión $\{\nu_n^+ : n \in N\}$ por medio de la recursión

$$\nu_{n+1}^+ = \min\{j > \nu_n^+ : R_j > R_{\nu_n^+}\} \quad (1.24)$$

donde $\nu_0^+ = 0$. A ν_n^+ le llamaremos el n -ésimo salto ascendente. De manera similar definiremos la sucesión $\{\nu_n^- : n \in N\}$ por medio de la recursión

$$\nu_{n+1}^- = \min\{j > \nu_n^- : R_j \leq R_{\nu_n^-}\} \quad (1.25)$$

donde $\nu_0^- = 0$. A ν_n^- le llamaremos el n -ésimo salto descendiente.

Observación 1.2.5. De la definición anterior se sigue que $\nu_1^+ = \nu^+$ al igual que $\nu_1^- = \nu^-$.

Es importante notar que ambas sucesiones son crecientes. Más adelante en la demostración del Teorema 1.2.4 se exhibirá que estas dos sucesiones están bien definidas sobre los naturales, con lo cual se excluye a ∞ de su rango. Las siguientes dos proposiciones nos serán de gran utilidad.

Proposición 1.2.3. Las variables aleatorias $A_n = \{\nu_{n+1}^+ - \nu_n^+\}_{n \in N}$ son i.i.d.

Demostración. Utilizando la ecuación (1.22) se sigue que

$$\begin{aligned} \mathbb{P}[A_n = k] &= \mathbb{P}\left[Z_{\nu_n^+ + 1} \leq 0, Z_{\nu_n^+ + 1} + Z_{\nu_n^+ + 2} \leq 0, \dots, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + k} Z_i \leq 0, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + k + 1} Z_i > 0\right] \\ &= \mathbb{P}\left[Z_1 \leq 0, Z_1 + Z_2 \leq 0, \dots, \sum_{i=1}^{k-1} Z_i \leq 0, \sum_{i=1}^k Z_i > 0\right] \\ &= \mathbb{P}[\nu_1^+ - \nu_0^+ = k] = \mathbb{P}[A_0 = k] \end{aligned}$$

la segunda igualdad utiliza el hecho de que las variables Z_i son i.i.d., por lo tanto queda demostrada la igualdad en distribución.

$$\begin{aligned} \mathbb{P}[A_n = r_n \mid A_{n-1} = r_{n-1}, \dots, A_0 = r_0] &= \mathbb{P}\left[\left(\forall k_n < r_n, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + k_n} Z_i \leq 0, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + r_n + 1} Z_i > 0\right)\right] \\ &\quad \left(\forall k_{n-1} < r_{n-1}, \sum_{i=\nu_{n-1}^+ + 1}^{\nu_{n-1}^+ + r_{n-1}} Z_i \leq 0, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + k_n + 1} Z_i > 0\right), \dots, \\ &\quad \left(\forall k_0 < r_0, \sum_{i=\nu_0^+ + 1}^{\nu_0^+ + r_0} Z_i \leq 0, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + k_n + 1} Z_i > 0\right)\Big] \\ &= \mathbb{P}\left[\left(\forall k_n < r_n, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + k_n} Z_i \leq 0, \sum_{i=\nu_n^+ + 1}^{\nu_n^+ + r_n + 1} Z_i > 0\right)\right] = \mathbb{P}[A_n = r_n] \end{aligned}$$

Lo cual demuestra la independencia. \square

Proposición 1.2.4. *Las variables aleatorias $W_n = R_{\nu_{n+1}^+} - R_{\nu_n^+}$ son i.i.d.*

Demostración. Para demostrar que son idénticamente distribuidas, por la Proposición 1.2.3 se tiene que

$$\begin{aligned} \mathbb{P}[W_n < r] &= \mathbb{P}\left[R_{\nu_{n+1}^+} - R_{\nu_n^+} < r\right] = \mathbb{P}\left[\sum_{i=1}^{\nu_{n+1}^+} Z_i - \sum_{i=1}^{\nu_n^+} Z_i < r\right] \\ &= \mathbb{P}\left[\sum_{i=\nu_n^++1}^{\nu_{n+1}^+} Z_i < r\right] = \mathbb{P}\left[\sum_{i=1}^{\nu_{n+1}^+ - \nu_n^+} Z_i < r\right] \\ &= \mathbb{P}\left[\sum_{i=1}^{\nu_1^+ - \nu_0^+} Z_i < r\right] = \mathbb{P}[W_0 < r] \end{aligned}$$

lo cual no depende de n , entonces queda demostrada la igualdad en distribución. Para demostrar independencia se tendrá lo siguiente

$$\begin{aligned} &\mathbb{P}[W_0 < r_0, \dots, W_n < r_n] \\ &= \mathbb{P}\left[R_{\nu_1^+} < r_0, R_{\nu_2^+} - R_{\nu_1^+} < r_1, \dots, R_{\nu_{n+1}^+} - R_{\nu_n^+} < r_n\right] \\ &= \mathbb{P}\left[\sum_{i=1}^{\nu_1^+} Z_i < r_0, \sum_{i=\nu_1^++1}^{\nu_2^+} Z_i < r_1, \dots, \sum_{i=\nu_n^++1}^{\nu_{n+1}^+} Z_i < r_n\right] \\ &= \mathbb{P}[W_0 < r_0] \mathbb{P}[W_1 < r_1] \dots \mathbb{P}[W_n < r_n] \end{aligned}$$

Se utilizó el hecho de que $R_{\nu_n^+}$ está bien definido para toda $n \in N$ (lo cual será demostrado en Teorema 1.2.4) y el hecho de que los procesos $\{\nu_{n+1}^+ - \nu_n^+\}$ y $\{Z_i\}$ son formados por v.a.i.i.d.. Por lo tanto queda demostrada la independencia. \square

Los resultados que hemos expuesto nos serán de utilidad, ya que son la razón por la cual se pide la condición de ganancia neta. En la sección siguiente desarrollaremos esta teoría y después de esta sección se tomará a la condición de ganancia neta como un hecho cierto durante el resto de este trabajo.

1.2.3. Condición de ganancia neta

Una pregunta básica que se hacen las compañías de seguros es cómo poder escoger correctamente la cuota de las primas para con ella poder hacer frente

a las pérdidas que vayan surgiendo por los siniestros. En nuestro modelo pensamos que esta cuota es fija. Por el comportamiento de la función $S(t)$ parecería apropiado cobrar una gran cantidad c para poder hacer frente a los siniestros (esto por que en el Teorema 1.2.2 demostramos que crece sin cota). El problema surge al tomar un gran valor para c es que la compañía sería menos competitiva respecto al mercado y entonces el número de asegurados decrecería, dado que éstos irían a otra compañía la cual cobrara menores primas.

A continuación mostraremos que una condición necesaria para que la probabilidad de ruina sea distinta de uno es la condición de ganancia neta (ver la Sección 1.2), lo cual demostraremos a partir de los siguientes resultados. Antes de continuar, necesitaremos definir lo que es un tiempo de paro.

Definición 1.2.3. *Una variable aleatoria N , que toma valores en los naturales se le llamará tiempo de paro de la sucesión ξ_1, ξ_2, \dots , si la σ -álgebra generada por ξ_1, \dots, ξ_n cumple con*

$$\{N = n\} \in \sigma(\xi_1, \dots, \xi_n) \text{ para toda } n \in \mathbb{N}. \quad (1.26)$$

Teorema 1.2.3 (Teorema de Wald). *Sea $\{\xi_n\}_{n \in \mathbb{N}}$ v.a.i.i.d. con $\mathbb{E}[\xi_1] < \infty$ y sea α un tiempo de paro con respecto a estas variables aleatorias tal que $\mathbb{E}[\alpha] < \infty$, entonces*

$$\mathbb{E} \left[\sum_{i=1}^{\alpha} \xi_i \right] = \mathbb{E}[\xi_1] \mathbb{E}[\alpha] \quad (1.27)$$

La demostración puede ser consultada en el Apéndice B.1.5 y una mayor referencia de este resultado puede ser consultada en Rolski(1998)[32].

El corolario del siguiente teorema da la razón por la cual $\rho > 0$.

Teorema 1.2.4. *Con la notación antes mencionada se tendrán los siguientes límites*

1. Si $\mathbb{E}[Z] > 0$ entonces $\lim_{n \rightarrow \infty} R_n \stackrel{c.s.}{=} \infty$
2. Si $\mathbb{E}[Z] < 0$ entonces $\lim_{n \rightarrow \infty} R_n \stackrel{c.s.}{=} -\infty$
3. Si $\mathbb{E}[Z] = 0$ entonces $\limsup_{n \rightarrow \infty} R_n \stackrel{c.s.}{=} \infty$ y $\liminf_{n \rightarrow \infty} R_n \stackrel{c.s.}{=} -\infty$

Demostración. Por la ley fuerte de los grandes números tenemos que $R_n/n \xrightarrow{c.s.} \mathbb{E}[Z]$. Entonces $R_n \xrightarrow{c.s.} \infty$ si $\mathbb{E}[Z] > 0$ y $R_n \xrightarrow{c.s.} -\infty$ si $\mathbb{E}[Z] < 0$. Entonces 1 y 2 quedan demostrados. En el caso de 3 tenemos la hipótesis de que $\mathbb{E}[Z] = 0$

y definamos a $N = \min\{n : R_n = \max_{j \geq 0} R_j\}$ y $N = \infty$ en el caso de que $\{n : R_n = \max_{j \geq 0} R_j\} = \emptyset$. Entonces

$$\begin{aligned}
1 &\geq \mathbb{P}[N < \infty] = \sum_{n=0}^{\infty} \mathbb{P}[N = n] \\
&= \sum_{n=0}^{\infty} \mathbb{P}[\{R_j < R_n \quad \forall j = 0, 1, \dots, n-1\} \cap \{R_j \leq R_n \quad \forall j > n\}] \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left[\left\{\sum_{k=j+1}^n Z_k > 0 \quad \forall j = 0, 1, \dots, n-1\right\} \cap \left\{\sum_{k=n+1}^j Z_k \leq 0 \quad \forall j > n\right\}\right] \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left[\sum_{k=j+1}^n Z_k > 0 \quad \forall j = 0, 1, \dots, n-1\right] \mathbb{P}\left[\sum_{k=n+1}^j Z_k \leq 0 \quad \forall j > n\right] \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left[Z_n > 0, Z_n + Z_{n+1} > 0, \dots, \sum_{k=1}^n Z_k > 0\right] \mathbb{P}[R_j \leq 0 \quad \forall j \geq 0] \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left[Z_1 > 0, Z_1 + Z_2 > 0, \dots, \sum_{k=1}^n Z_k > 0\right] \mathbb{P}[\nu^+ = \infty],
\end{aligned}$$

donde en la última igualdad se utilizó el hecho de que $(Z_n, Z_{n-1}, \dots, Z_1) \stackrel{d}{=} (Z_1, Z_2, \dots, Z_n)$ para toda $n \geq 1$, entonces

$$1 \geq \sum_{n=0}^{\infty} \mathbb{P}[\nu^- > n] \mathbb{P}[\nu^+ = \infty] = \mathbb{E}[\nu^-] \mathbb{P}[\nu^+ = \infty]. \quad (1.28)$$

Esta última igualdad es válida porque al ser ν^- una v.a. no negativa se puede calcular su esperanza por medio de su función de distribución⁴. Se tienen entonces los siguientes casos:

1. Si $\mathbb{P}[\nu^+ = \infty] > 0$ entonces se tendrá al utilizar la ecuación (1.28) que $\mathbb{E}[\nu^-] = \sum_{n=0}^{\infty} \mathbb{P}[\nu^- > n] < \infty$ y usando la igualdad del teorema de Wald B.6 que $\mathbb{E}[R_{\nu^-}] = \mathbb{E}[\nu^-] \mathbb{E}[Z] = 0$.

Como $R_{\nu^-} \leq 0$ entonces $R_{\nu^-} \stackrel{c.s.}{=} 0$. Esto nos da una contradicción dado que $\mathbb{P}[R_{\nu^-} < 0] \geq \mathbb{P}[Z_1 < 0] > 0$ ya que supusimos $\mathbb{E}[Z] = 0$.

2. Si $\mathbb{P}[\nu^+ < \infty] = 1$, entonces $R_{\nu^+} > 0$, por definición de ν^+ . Ahora demostraremos que $\mathbb{P}[\nu_n^+ < \infty] = 1$ para toda $n \in \mathbb{N}$ utilizando inducción

⁴Si X es una v.a. no negativa entonces $\mathbb{E}[X] = \sum_{n=0}^{\infty} \mathbb{P}[X > n]$ en caso discreto y $\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}[X > x] dx$ en caso continuo.

sobre n . Supongamos que es válido para n entonces

$$\begin{aligned}
\mathbb{P}[\nu_{n+1}^+ = \infty] &= \mathbb{P}[\nu_{n+1}^+ = \infty \cap \nu_n^+ < \infty] \\
&= \sum_{i=1}^{\infty} \mathbb{P}[\nu_{n+1}^+ = \infty \cap \nu_n^+ = i] \\
&= \sum_{i=1}^{\infty} \mathbb{P}[\text{mín}\{j > \nu_n^+ \mid R_j > R_{\nu_n^+}\} = \infty \cap \nu_n^+ = i] \\
&= \sum_{i=1}^{\infty} \mathbb{P}[\text{mín}\{j > i \mid R_j > R_i\} = \infty \cap \nu_n^+ = i] \\
&\leq \sum_{i=1}^{\infty} \mathbb{P}[\text{mín}\{j > i \mid R_j > R_i\} = \infty] \\
&= \sum_{i=1}^{\infty} \mathbb{P}[R_j < R_i \quad \text{para toda } j > i] \\
&= \sum_{i=1}^{\infty} \mathbb{P}\left[\sum_{k=i+1}^j Z_k < 0 \quad \text{para toda } j > i\right] \\
&= \sum_{i=1}^{\infty} \mathbb{P}\left[\sum_{k=1}^{j-i-1} Z_k < 0 \quad \text{para toda } j > i\right] \\
&= \sum_{i=1}^{\infty} \mathbb{P}\left[\sum_{k=1}^j Z_k < 0 \quad \text{para toda } j > 0\right] \\
&= \sum_{i=1}^{\infty} \mathbb{P}[\nu^+ = \infty] = 0.
\end{aligned}$$

Entonces excluimos el caso de ∞ , por tanto tenemos que $R_{\nu_n^+}$ está bien definido para toda $n \in \mathbb{N}$.

Como resultado de lo último, podemos definir la sucesión

$$\{R_{\nu_{n+1}^+} - R_{\nu_n^+}\}_{n \in \mathbb{N}},$$

la cual es una sucesión de v.a.i.i.d. (esto fue demostrado en 1.2.4). Por como fué definido $R_{\nu_n^+}$ se tiene que $\mathbb{E}[R_{\nu_n^+}] > 0$ y utilizando la ley fuerte de los grandes números se sigue que

$$\frac{1}{n}R_{\nu_n^+} = \frac{1}{n} \sum_{k=0}^{n-1} (R_{\nu_{k+1}^+} - R_{\nu_k^+}) \xrightarrow{c.s} \mathbb{E}[R_{\nu_n^+}] > 0$$

Entonces $\lim_{n \rightarrow \infty} R_{\nu_n^+} = \infty$ lo cual implica que $\limsup_{n \rightarrow \infty} R_n = \infty$. La prueba para $\liminf_{n \rightarrow \infty} R_n = -\infty$ es análoga, considerando la caminata aleatoria $\{-R_n\}$ con $\mathbb{E}[-Z] = 0$. Con eso demostramos que existen dos subsucesiones una que tiende a ∞ y otra a $-\infty$, lo cual demuestra el último inciso. \square

Corolario 1.2.4.1. Si $\mathbb{E}[Z] \leq 0$ entonces la probabilidad de ruina $\psi(u) \equiv 1$ y si $\mathbb{E}[Z] > 0$ entonces $\psi(u) < 1$ para toda $u \in \mathbb{R}_+$.

Demostración. Para el caso cuando $\mathbb{E}[Z] \leq 0$ se tiene que $\sup R_n = \infty$ y utilizando la ecuación (1.19) se tendrá que $\psi(u) \equiv 1$.

Si $\mathbb{E}[Z] > 0$ bastará con probar que⁵

$$\psi(0) = \mathbb{P} \left[\sup_n R_n > 0 \right] < 1.$$

Si se tuviera que

$$\mathbb{P} \left[\sup_n R_n > 0 \right] = 1,$$

entonces R_n cruzaría el valor 0 al menos una vez casi seguramente. Como se tiene que $R_n \xrightarrow{c.s.} -\infty$ se tendrá con probabilidad uno que la sucesión cruzará en algún descenso otra vez el cero. Utilizando este argumento repetidas veces se tendrá que la sucesión cruzará una y otra vez el valor cero, lo cual contradice que $R_n \xrightarrow{c.s.} -\infty$. \square

Como $\mathbb{E}[Z] = \mathbb{E}[cY - X] = \frac{c}{\lambda} - \mu$ entonces el corolario anterior pide que $\frac{c}{\lambda} - \mu > 0$ para que la probabilidad de ruina no sea idénticamente uno, lo cual es justamente la condición de ganancia neta. De aquí en adelante tomaremos como cierta a la condición de ganancia neta.

1.3. Ecuaciones para la probabilidad de ruina

Como cada una de las sucesiones $\{T_i\}$ y $\{\xi_i\}$ consisten de v.a.i.i.d. y las dos sucesiones son independientes entre sí, la probabilidad de ruina $\psi(x)$ al utilizar la ecuación (1.19) se puede ver como la probabilidad de la cola del supremo de la caminata aleatoria $\{R_n\}$. Es claro entonces que esta probabilidad no será fácil de calcular dado que es una función de un proceso aleatorio muy sofisticado (ver figura 1.5).

En principio se debe evitar que la ruina tenga probabilidad 1, y la meta es entonces poder evitar que la probabilidad de que la caminata aleatoria

⁵Esto es suficiente porque para $x > 0$ se tiene que $\psi(x) \leq \psi(0)$

$\{R_n\}$ exceda el valor de x sea lo suficientemente baja (lo suficientemente baja según los requisitos de nuestra compañía).

El siguiente teorema, da una forma de calcular las probabilidades de ruina via una ecuación integral. Antes de dicho teorema necesitaremos las siguientes definiciones.

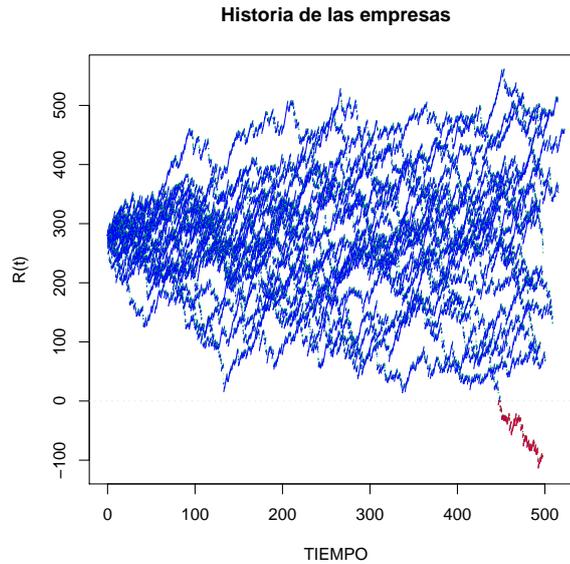


Figura 1.2: Este es un ejemplo del comportamiento caótico que presenta la probabilidad de ruina. En la figura se muestran distintas simulaciones del proceso de reserva y que solo una trayectoria incurrió en ruina durante las primeras 500 unidades de tiempo.

Definición 1.3.1 (Cola de la distribución). *Se define la cola de la distribución F como*

$$\bar{F}(x) := 1 - F(x), \quad x \geq 0. \quad (1.29)$$

Definición 1.3.2 (Probabilidad de supervivencia o de no ruina). *Para el caso particular de la probabilidad de ruina utilizaremos la notación*

$$\delta(x) = \bar{\psi}(x). \quad (1.30)$$

Las dos definiciones anteriores las hicimos tan solo para facilitar la notación. La definición siguiente tiene un significado más amplio.

Definición 1.3.3 (Cola integrada). Para una v.a. positiva ξ , con función de distribución F , se define la cola integrada de una distribución F como

$$F_I(x) := \frac{1}{\mu} \int_0^x \bar{F}(y) dy, \quad x \geq 0. \quad (1.31)$$

Observación 1.3.1. Hay que distinguir que F_I es una función de distribución, dado que, para una variable aleatoria positiva ξ , tenemos que $\mathbb{E}[\xi] = \int_0^\infty \bar{F}(y) dy$ y por tanto $F_I \rightarrow 1$ cuando $y \rightarrow \infty$, además de ser creciente y continua.

Los siguientes dos ejemplos serán de utilidad para la simulación de la probabilidad de ruina.

Ejemplo 1.3.1 (Cola integrada de la distribución Exponencial). Sea X una v.a. con distribución exponencial de parámetro $\alpha > 0$ ($X \sim \text{exp}(\alpha)$), entonces su función de distribución es

$$F(x) = 1 - e^{-\alpha x} \quad \text{para} \quad x \geq 0,$$

como su esperanza es $\mathbb{E}[X] = 1/\alpha$, se sigue que

$$\begin{aligned} F_I(x) &= \alpha \int_0^x e^{-\alpha y} dy \\ &= 1 - e^{-\alpha x}. \end{aligned}$$

Ésta última ecuación muestra que la distribución de la cola integrada de una exponencial vuelve a ser exponencial.

Ejemplo 1.3.2. [Cola integrada de la distribución Pareto] Sea X una v.a. con distribución Pareto de parámetros $a > 0$ y $b > 1$ ($X \sim \text{Pareto}(a, b)$), entonces su función de distribución es

$$F(x) = 1 - \left(\frac{a}{x}\right)^b \quad \text{para} \quad x \geq a,$$

como su esperanza es $\mathbb{E}[X] = ab/(b-1)$, se sigue que

$$\begin{aligned} F_I(x) &= \frac{b-1}{ab} \int_0^x 1 - \left[1 - \left(\frac{a}{y}\right)^b\right] 1_{\{x>a\}} dy \\ &= \frac{b-1}{ab} x 1_{\{x \leq a\}} + \frac{b-1}{ab} \left[a + \int_a^x \left(\frac{a}{y}\right)^b \right] 1_{\{x>a\}} dy \\ &= \frac{b-1}{ab} x 1_{\{x \leq a\}} + \left[1 - \frac{1}{b} \left(\frac{a}{x}\right)^{b-1}\right] 1_{\{x>a\}}. \end{aligned}$$

Ésta transformación hace que la distribución cambie radicalmente (ver figura 1.3 y 1.4); más adelante nos será muy útil.

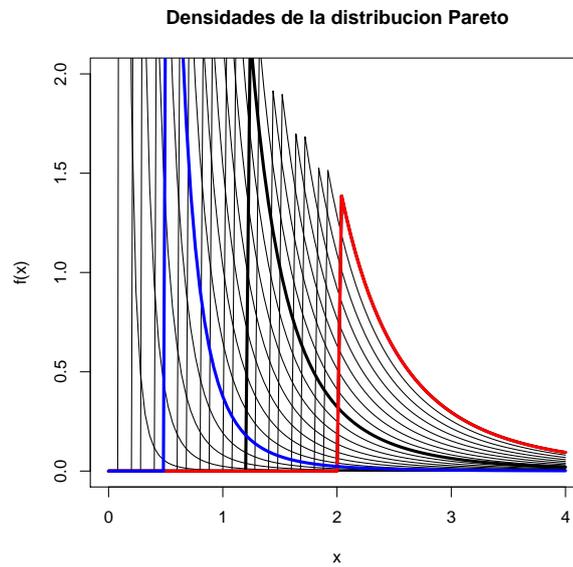


Figura 1.3: Esta figura muestra distintas gráficas de la densidad de Pareto al cambiar su parámetro de localización a .

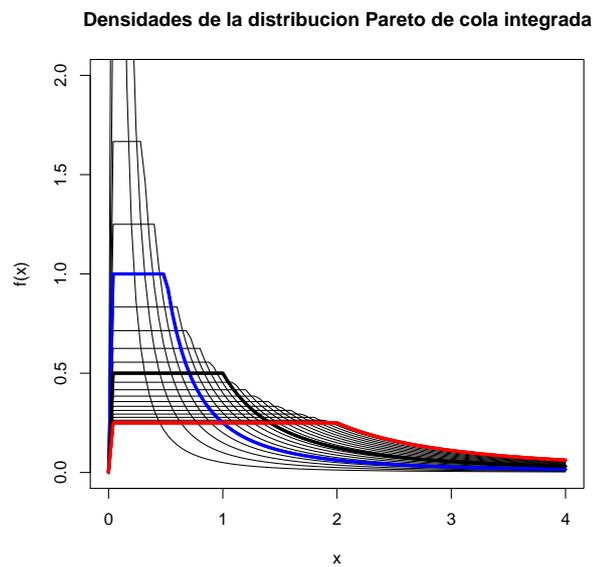


Figura 1.4: Esta figura muestra distintas gráficas de la densidad de una cola integrada de una distribución Pareto al cambiar su parámetro de localización a .

Generalmente no se tendrá una expresión analítica para la distribución de la cola. El siguiente teorema muestra una ecuación que cumple la probabilidad de ruina, la cual tampoco tendrá solución analítica en la mayor parte de los casos⁶, pero que más adelante nos llevará a expresar la probabilidad de ruina como una suma aleatoria que podremos simular y con esto obtener estimadores para la probabilidad de ruina.

Teorema 1.3.1. *Consideremos el modelo Cramér-Lundberg y supongamos que la condición de ganancia neta es cierta. Además, supongamos que la función de distribución de los reclamos F tiene densidad, entonces $\delta(x) \in C^1$ y cumple con*

$$\delta'(x) = \frac{\lambda}{c}\delta(x) - \frac{\lambda}{c} \int_0^x \delta(x-y)dF(y), \quad (1.32)$$

la cual es equivalente a

$$\delta(x) = \frac{\rho}{1+\rho} + \frac{1}{1+\rho} \int_0^x \delta(x-y)dF_I(y). \quad (1.33)$$

A estas ecuaciones se les denotará por ecuación diferencial de ruina y ecuación integral de ruina respectivamente.

Demostración. A $\delta(x)$ la podemos ver como la probabilidad de que nuestra compañía nunca tenga una reserva negativa, esto es

$$\begin{aligned} \delta(x) &= \mathbb{P}[S(t) - ct \leq x, \quad \text{para toda } t > 0] \\ &= \mathbb{P}\left[\sum_{k=1}^n \xi_k - cY_k \leq x, \quad \text{para toda } n \geq 1\right] \\ &= \mathbb{P}\left[\sum_{k=2}^n X_k - cY_k \leq x + cY_1 - \xi_1, \quad \text{para toda } n \geq 2, \quad \xi_1 - cY_1 \leq x\right] \\ &= \mathbb{P}[S'(t) - ct \leq x + cY_1 - \xi_1, \quad \text{para toda } t > 0, \quad \xi_1 - cY_1 \leq x]. \end{aligned}$$

La segunda igualdad es válida por ser v.a.i.. Para la tercera igualdad se define a $S'(t) = \sum_{k=2}^n X_k$ como una copia independiente (en distribución)⁷ del proceso S . Esto es válido porque

$$(\xi_2, \xi_3, \dots, \xi_n) \stackrel{d}{=} (\xi_1, \xi_2, \dots, \xi_{n-1}) \quad \text{para toda } n \geq 0 \quad (1.34)$$

⁶De hecho solo se tiene solución para el caso de que los reclamos tengan distribución exponencial.

⁷Lo cual quiere decir que su medida de probabilidad es igual que la del proceso original.

por ser v.a.i.i.d. Donde nuestro proceso original es definido sobre $(\xi_2, \xi_3, \dots, \xi_n)$ mientras que S' lo componen $(\xi_1, \xi_2, \dots, \xi_{n-1})$. Al tomar esperanza condicional sobre $\{Y_1, \xi_1\}$ se tiene que

$$\begin{aligned}
\delta(x) &= \mathbb{E} [\mathbb{P} [S'(t) - ct \leq x + cY_1 - \xi_1, \quad t > 0, \quad \xi_1 - cY_1 \leq x \mid Y_1, \xi_1]] \\
&= \iint_{\mathbb{R}^2} \mathbb{P} [S'(t) - ct \leq x + cY_1 - \xi_1, t > 0, \xi_1 - cY_1 \leq x \mid Y_1 = s, \xi_1 = z] dF(z) dF_y(s) \\
&= \iint_{\mathbb{R}^2} \mathbb{P} [S'(t) - ct \leq x + cs - z, t > 0, z - cs \leq x \mid Y_1 = s, \xi_1 = z] dF(z) \lambda e^{-\lambda s} ds \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{P} [S'(t) - ct \leq x + cs - z, \quad t > 0, \quad z - cs \leq x] dF(z) \lambda e^{-\lambda s} ds \\
&= \int_0^{\infty} \int_0^{x+cs} \mathbb{P} [S'(t) - ct \leq x + cs - z, \quad t > 0] dF(z) \lambda e^{-\lambda s} ds \\
&= \int_0^{\infty} \lambda e^{-\lambda s} \int_0^{x+cs} \delta(x + cs - z) dF(z) ds.
\end{aligned}$$

Utilizando la sustitución $w = x + cs$ en la última ecuación obtenemos

$$\delta(x) = \frac{\lambda}{c} e^{\frac{\lambda}{c}x} \int_x^{\infty} e^{-\frac{\lambda}{c}w} \int_0^w \delta(w - z) dF(z) dw. \quad (1.35)$$

Por hipótesis F tiene densidad, entonces la función

$$g(w) = \int_0^w \delta(w - z) dF(z),$$

es continua y nos permite reescribir a δ como

$$\delta(x) = \frac{\lambda}{c} e^{\frac{\lambda}{c}x} \int_0^{\infty} e^{-\frac{\lambda}{c}w} g(w) dw,$$

lo cual implica que $\delta(x)$ es diferenciable. Por lo tanto, derivando con respecto de x , ambos lados de la igualdad (1.35) se tendrá que

$$\delta'(x) = \frac{\lambda}{c} \delta(x) - \frac{\lambda}{c} \int_0^x \delta(x - z) dF(z). \quad (1.36)$$

Ahora integrando sobre $(0, t]$ la variable x obtenemos

$$\begin{aligned}
\delta(t) - \delta(0) - \frac{\lambda}{c} \int_0^t \delta(x) dx &= -\frac{\lambda}{c} \int_0^t \int_0^x \delta(x - z) dF(z) dx \\
&= -\frac{\lambda}{c} \left(\int_0^t \delta(0) F(x) dx + \int_0^t \int_0^x \delta'(x - z) F(z) dz dx \right) \\
&= -\frac{\lambda}{c} \left(\int_0^t \delta(0) F(x) dx + \int_0^t \int_z^t \delta'(x - z) dx F(z) dz \right) \\
&= -\frac{\lambda}{c} \left(\int_0^t \delta(0) F(x) dx + \int_0^t \int_z^t \delta'(x - z) dx F(z) dz \right)
\end{aligned}$$

en la primera igualdad utilizamos el hecho de que $F(0) = 0$ porque $X > 0$ c.s.. Lo anterior lo podemos reescribir como

$$\begin{aligned}\delta(t) - \delta(0) &= \frac{\lambda}{c} \int_0^t \delta(x) dx - \frac{\lambda}{c} \delta(0) \int_0^t F(x) dx - \frac{\lambda}{c} \int_0^t F(z) [\delta(t-z) - \delta(0)] dz \\ &= \frac{\lambda}{c} \int_0^t \delta(x) dx - \frac{\lambda}{c} \int_0^t F(z) \delta(t-z) dz \\ &= \frac{\lambda}{c} \int_0^t \bar{F}(z) \delta(t-z) dz.\end{aligned}$$

Se sigue que

$$\frac{c}{\lambda} (\delta(t) - \delta(0)) = \int_0^t \delta(t-z) \bar{F}(z) dz. \quad (1.37)$$

Antes de continuar necesitamos probar que $\lim_{t \rightarrow \infty} \delta(t) = 1$. Se tiene que

$$\begin{aligned}\delta(x) &= \mathbb{P}[S(t) - ct \leq x \quad \text{para toda } t > 0] \\ &= \mathbb{P}\left[\sup_{t>0} \{S(t) - ct\} \leq x\right]\end{aligned}$$

como $S(t) - ct \xrightarrow{c.s.} -\infty$ cuando $n \rightarrow \infty$ (utilizando la condición de ganancia neta), entonces $\sup_{t>0} \{S(t) - ct\} \stackrel{c.s.}{<} \infty$, lo cual demuestra que

$$\lim_{x \rightarrow \infty} \delta(x) = 1. \quad (1.38)$$

Utilizando la prueba M de Weierstrass (ver el Apéndice B.2.3) haciendo $f(z, t) = \delta(t-z) \bar{F}(z)$ y $M(z) = \bar{F}(z)$ junto con (1.38), tendremos que

$$\begin{aligned}\frac{c}{\lambda} (1 - \delta(0)) &= \lim_{t \rightarrow \infty} \int_0^\infty \delta(t-z) \bar{F}(z) dz \\ &= \int_0^\infty \left(\lim_{t \rightarrow \infty} \delta(t-z) \right) \bar{F}(z) dz \\ &= \int_0^\infty \bar{F}(z) dz \\ &= \mu.\end{aligned}$$

Entonces

$$\delta(0) = 1 - \frac{\lambda\mu}{c} = \frac{\rho}{1 + \rho}.$$

Sustituyendo en (1.37) se obtiene

$$\delta(t) = \frac{\rho}{1 + \rho} + \frac{1}{1 + \rho} \frac{1}{\mu} \int_0^t \delta(t-u) \bar{F}(u) du,$$

lo cual es equivalente a (1.33). □

Una propiedad muy utilizada del proceso $\{R_t\}$ es la de tener incrementos independientes y estacionarios. Considerando $\{R_t\}$ a partir del momento t , si no ha ocurrido la ruina, es como volver a empezar un nuevo proceso de reserva el cual tiene la misma distribución para sus reclamos que el original pero con reserva inicial $x = R(t)$. En particular si $R(t) = y$ y la ruina todavía no ha ocurrido para el tiempo t , entonces la probabilidad de ruina (condicionada a la falta de ruina al momento t) es simplemente $\psi(y)$. Además, si consideramos el proceso de reserva a partir del primer momento de reclamo T_1 es como iniciar un proceso de reserva con capital inicial $x + cT_1 - X_1$.

Existe una gran analogía entre la forma en que se demostró el teorema anterior y el análisis del primer paso de Markov. En el análisis del primer paso se utiliza la siguiente igualdad entre conjuntos

$$[X_\tau = j] = \bigcup_k [X_\tau = j, X_1 = k],$$

la cual se basa en descomponer el evento $[X_\tau = j]$ en la unión de los eventos ajenos, los cuales se caracterizan por que el “primer paso” es igual a k . En el caso de la demostración se utiliza este mismo tipo de análisis, pero sobre las intersecciones de la ruina con respecto a todos los posibles “primeros reclamos”, esto variando entre todos los posibles momentos en los que puede ocurrir y entre todos los montos posibles de éste.

El último teorema muestra que la probabilidad de ruina puede ser encontrada como la solución de una ecuación integro-diferencial o integral. En general no es fácil encontrar este tipo de soluciones de modo que utilizaremos métodos de simulación para encontrar estimadores de la probabilidad de ruina. Sin embargo, cuando las reclamaciones tienen distribución exponencial el sistema (1.32) tiene una solución analítica, como muestra el siguiente ejemplo.

Ejemplo 1.3.3. *[Solución explícita de la probabilidad de ruina cuando las reclamaciones son exponenciales] Consideremos el modelo Cramér-Lundberg en donde las reclamaciones tiene distribución $\exp(\alpha)$, lo cual implica que $\mu = 1/\alpha$. Entonces al sustituir en la ecuación diferencial de ruina (1.32) se tiene que*

$$\begin{aligned} \delta'(x) &= \frac{\lambda}{c} \delta(x) - \frac{\lambda}{c} \int_0^x \delta(x-y) \alpha e^{-\alpha y} dy \\ &= \frac{\lambda}{c} \left[\delta(x) - e^{-\alpha x} \int_0^x \delta(y) \alpha e^{\alpha y} dy \right] \end{aligned}$$

al derivar ambos lados de la ecuación anterior, se tendrá que

$$\begin{aligned}\delta''(x) &= \frac{\lambda}{c} \left[\delta'(x) - \delta(x) - \alpha e^{-\alpha x} \int_0^x \delta(y) \alpha e^{\alpha y} dy \right] \\ &= \left(\frac{\lambda}{c} - \alpha \right) \delta'(x).\end{aligned}$$

La solución de la última ecuación es $\delta(x) = a + b \exp \left\{ - \left(\alpha - \frac{\lambda}{c} \right) x \right\}$, donde a y b son constantes. Utilizando las condiciones $\psi(0) = \frac{\lambda}{\alpha c}$ y $\psi(\infty) = 0$, además del hecho de que $\delta = 1 - \psi$ se tiene entonces que $a = 1$ y $b = -\frac{\lambda}{\alpha c}$. Por lo tanto

$$\psi(x) = \frac{\lambda}{\alpha c} e^{-(\alpha - \frac{\lambda}{c})x}. \quad (1.39)$$

Entonces para el caso de que la distribución de los reclamos sea exponencial se tiene una expresión analítica para la probabilidad de ruina.

El siguiente corolario muestra otra forma de expresar las probabilidades de ruina.

Corolario 1.3.1.1. *La probabilidad de ruina $\psi(x)$ satisface la siguiente ecuación integral*

$$c\psi(x) = \lambda \left(\int_x^\infty \bar{F}_U(y) dy + \int_0^x \psi(x-y) \bar{F}_U(y) dy \right). \quad (1.40)$$

Demostración. Sustituyendo $\delta(u) = 1 - \psi(u)$ y $\delta(0) = 1 - \frac{\lambda\mu}{c}$ en la ecuación (1.37) y se tiene que

$$\begin{aligned}c\psi(u) &= \lambda\mu - \lambda \int_0^u [1 - \psi(u-x)] \bar{F}(x) dx \\ &= \lambda \left(\mu - \int_0^u \bar{F}(x) dx + \int_0^u \psi(u-x) \bar{F}(x) dx \right) \\ &= \lambda \left(\int_0^\infty \bar{F}(x) dx - \int_0^u \bar{F}(x) dx + \int_0^u \psi(u-x) \bar{F}(x) dx \right) \\ &= \lambda \left(\int_u^\infty \bar{F}(x) dx + \int_0^u \psi(u-x) \bar{F}(x) dx \right)\end{aligned}$$

□

1.4. Fórmula de Pollaczek-Khinchin

Un problema que surge en la simulación utilizando el método de Monte Carlo (ver Sección 2.1.1) para la probabilidad de ruina en horizonte infinito, es que a pesar de que podamos simular nuestro proceso $\{R_t\}$, de una manera adecuada, no podríamos simularlo para un tiempo infinito. No sabríamos si al simular nuestro proceso hasta cierta t y que éste no haya alcanzado la ruina esto implicaría que para algún otro tiempo futuro no la lograría alcanzar, esto por más grande que hiciéramos a t .

En esta sección se obtendrá la fórmula de Pollaczek-Khinchin, la cual facilitará la simulación de la probabilidad de ruina. La gran ventaja que esta fórmula aporta es poder simular una probabilidad de ruina a horizonte infinito por medio de simular una convolución geométrica, lo cual es computacionalmente sencillo. Por lo anterior es importante consultar el Apéndice C para una breve introducción al tema de las convoluciones, en especial la parte C.1 la cual habla de las convoluciones geométricas donde puede ser consultada la demostración del siguiente lema el cual utilizaremos para la demostración de la fórmula de Pollaczek-Khinchin.

Lema 1.4.1. *Sea H una función de distribución con función de densidad h y definimos las funciones G_1 y G_2 para $q \in (0, 1)$ como*

$$G_1(x) := (1 - q) \sum_{m=0}^{\infty} q^m H^{m*}(x) \quad G_2(x) := (1 - q) + q \int_0^x G_2(x - z) h(z) dz$$

para $x > 0$. Entonces sus transformadas de Laplace son iguales, lo cual implica que $G_1(x) = G_2(x)$.

Teorema 1.4.1. *[Fórmula de Pollaczek-Khinchin] Consideremos el modelo C-L con la condición de ganancia neta, además supongamos que ξ_i tiene densidad. y sean X_1, X_2, \dots v.a.i.i.d. con función de distribución F_I (ver la Definición 1.3.3), definamos δ^* por*

$$\delta^*(x) = \frac{\rho}{1 + \rho} \left[1 + \sum_{m=1}^{\infty} (1 + \rho)^{-m} \mathbb{P}[X_1 + X_2 + \dots + X_m \leq x] \right] \quad x > 0. \quad (1.41)$$

Entonces δ^* satisface (1.33), lo cual implica que $\delta^*(x) = \delta(x)$ para $x > 0$. Además, la función δ^* es la única función de distribución que la satisface.

Demostración. Para demostrar que δ^* satisface (1.33) es conveniente reescribir $q = (1 + \rho)^{-1}$ y $p = 1 - q = \rho/(1 + \rho)$ entonces al separar el primer

término de (1.41) tenemos

$$\begin{aligned}\delta(x) &= p + qp \left[F_I(x) + \sum_{m=2}^{\infty} q^{m-1} \int_0^x \mathbb{P}[y + X_2, \dots, X_m \leq x] dF_I(y) \right] \\ &= p + q \int_0^x p \left[1 + \sum_{m=1}^{\infty} q^m \mathbb{P}[X_2, \dots, X_m \leq x - y] dF_I(y) \right] \\ &= p + q \int_0^x \delta(x - y) dF_I(y)\end{aligned}$$

por lo tanto δ^* satisface (1.33). Para demostrar que es única utilizaremos el lema anterior el cual implica que $\delta^*(x) = \delta(x)$, con lo que concluye la prueba. \square

Observación 1.4.1. *Del teorema anterior podemos concluir que*

$$\delta(x) = \frac{\rho}{1 + \rho} \sum_{m=0}^{\infty} (1 + \rho)^{-m} F_I^{*m}(x) \quad \text{para } x > 0 \quad (1.42)$$

A ésta última fórmula la podemos identificar como la función de distribución de una convolución geométrica⁸ de parámetros $(\frac{\rho}{1+\rho}, F_I)$; es decir

$$\psi(x) = \mathbb{P}[S_M > x] \quad (1.43)$$

donde

$$S_M = \sum_{i=1}^M X_i \quad M \sim \text{geo} \left(\frac{\rho}{1 + \rho} \right) \quad X \sim F_I.$$

En este caso los sumandos tienen distribución F_I donde

$$F_I(x) = \frac{1}{\mu_\xi} \int_0^x \bar{F}(x) dx$$

Este es un hecho de gran importancia dado que permite la simulación de la probabilidad de ruina a horizonte infinito por medio de la simulación de convoluciones geométricas lo cual se usará en la Sección 2.4.1.

⁸Es recomendable consultar el Apéndice C.1.

1.5. Distribuciones de colas pesadas

Es importante cuando se está planteando un modelo estocástico que las distribuciones que se asignen al modelo reflejen las características más importantes⁹ que rigen al proceso real. En nuestro estudio quisieramos que la distribución que asignamos para los reclamos cumpla con este criterio. Por lo anterior haremos una división del conjunto de las distribuciones en dos subconjuntos; las distribuciones de colas pesadas y las de colas ligeras. Esta división se basa en la existencia de la función generadora de momentos. Una consecuencia de esto es que para las distribuciones de colas pesadas el decaimiento de la cola de su distribución es menor al decaimiento exponencial, se tendrá entonces que su convergencia a cero es sumamente lenta (lo cual se interpreta como que para valores grandes sigue existiendo una masa considerable bajo la curva de la densidad), algunos ejemplos son los de la figura 1.5.

Existen fenómenos en el mundo los cuales tienen baja ocurrencia (poca probabilidad) pero que su aparición cambia de manera radical a su entorno, por ejemplo los terremotos, incendios, plagas, pandemias entre otros.¹⁰ Desde el punto de vista de una aseguradora es sumamente importante tenerlos en cuenta ya que su aparición podría conllevar a la ruina. Este tipo de fenómenos, desde el punto de vista matemático, son variables aleatorias las cuales esperamos que tengan una gran masa bajo la curva de la cola de su distribución para valores lejanos del cero (ver Definición 1.3.1), ya que esto implica que para valores grandes de x las probabilidades de la forma $\mathbb{P}[X > x]$ sean considerables (relativamente altas).

Definición 1.5.1 (Distribución de cola ligera). *Sea X una v.a. positiva, se dice que X es de cola ligera (o tiene distribución de cola ligera) si existe un valor $s > 0$ para el cual $m_X(s) := \mathbb{E}[e^{sX}] < \infty$.*

Observación 1.5.1. *La definición de cola ligera permite que si X es una v.a. positiva de cola ligera entonces en una vecindad alrededor del origen exista la función generadora de momentos*

$$m_X(s) = \mathbb{E}[e^{sX}] \quad s \in (-s_0, s_0) \quad \text{para algún } s_0 > 0,$$

ya que por la desigualdad de Markov (ver Apéndice B.1.3) para $s \in (0, s_0)$

$$\mathbb{P}[X \geq x] = \mathbb{P}[e^{sX} \geq e^{sx}] \leq \frac{\mathbb{E}[e^{sX}]}{e^{sx}} = e^{-sx} m_X(s),$$

⁹Importantes con respecto a lo que estemos estudiando del fenómeno.

¹⁰Otro tipo de fenómenos son los que por sí solos no representan un cambio drástico en su entorno; por ejemplo los accidentes automovilísticos, los cuales (aún en el peor de los casos) un solo accidente no es causa suficiente para que una compañía se arruine. En este caso se necesitaría una racha de muchos accidentes para que la ruina ocurriera.

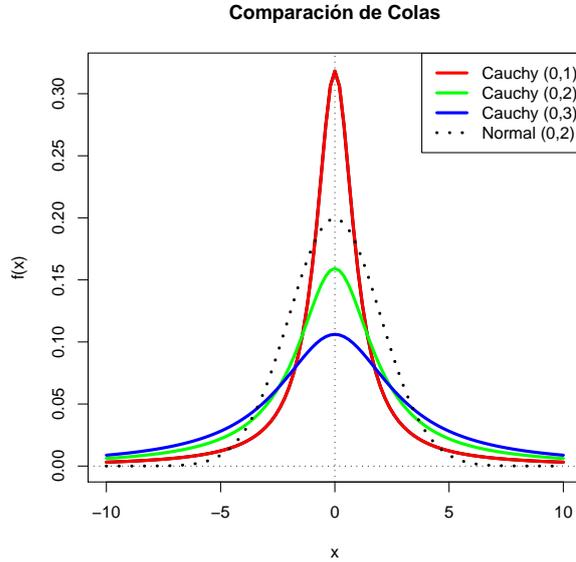


Figura 1.5: En esta figura se puede comparar el decaimiento de una distribución normal contra el de diversas distribuciones Cauchy

lo cual muestra que el decaimiento de $\mathbb{P}[X \geq x]$, para X de cola ligera, visto como una función de x es mayor al decaimiento exponencial.

Observación 1.5.2. Si existe un valor $s > 0$ para el cual $m_X(s) < \infty$ entonces también tendremos que $m_X(s') < \infty$ para $s' \leq s$ ya que $0 \leq e^{s'X} \leq e^{sX}$ implica (utilizando la monotonía de la esperanza) que $m_X(s') < \infty$.

Definición 1.5.2 (Distribución de cola pesada). Sea X una v.a. positiva, se dice que X es de cola pesada (o tiene distribución de cola pesada) si $m_X(s) = \infty$ para toda $s > 0$.

Algunos ejemplos de distribuciones de colas pesadas son la Lognormal, Pareto y la Weibull, esta última cuando su parámetro de forma es menor a 1 (ver figuras 1.6 y 1.7).

Observación 1.5.3. Una igualdad útil para obtener el comportamiento de la función generadora de momentos de una variable aleatoria positiva es

$$m_X(s) - 1 = s \int_0^{\infty} e^{sy} \bar{F}(y) dy, \quad (1.44)$$

la cuál utilizaremos frecuentemente.

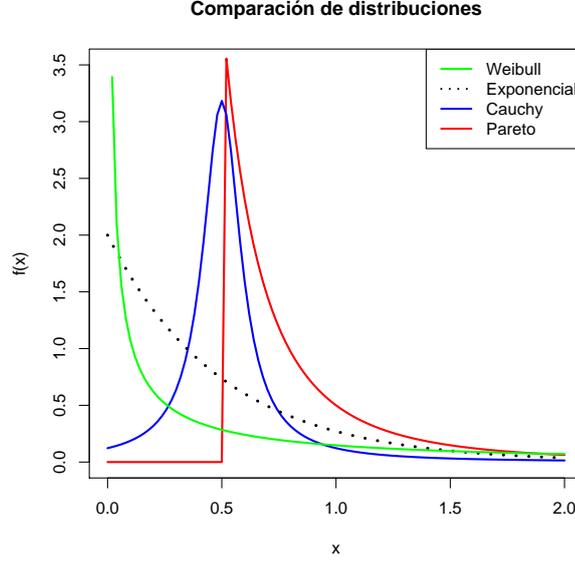


Figura 1.6: En esta figura se muestran distintas distribuciones de colas pesadas y la distribución exponencial, para así poder hacer una comparación. En esta parte de las distribuciones, no se distingue a simple vista cual es la diferencia.

El siguiente teorema nos dará otra forma de caracterizar a las distribuciones de colas pesadas lo cual facilitará su detección.

Teorema 1.5.1. *Si $\limsup_{x \rightarrow \infty} -\log \bar{F}(x)/x = 0$ entonces F es de cola pesada*

Demostración. Supongamos que $\limsup_{x \rightarrow \infty} -\log \bar{F}(x)/x = 0$, esto quiere decir que para toda $\varepsilon > 0$ existe una $x' > 0$ tal que $-\log \bar{F}(x) \leq \varepsilon x$ para toda $x > x'$. Entonces existe una constante $c = e^{\varepsilon x'} > 0$ tal que para toda $x \geq 0$

$$\bar{F}(x) \geq ce^{-\varepsilon x}, \quad (1.45)$$

de esta manera $e^{\varepsilon x} \bar{F}(x) \geq c$ implica que

$$\int_0^{\infty} e^{sx} \bar{F}(x) dx = \infty$$

Esto último para toda $s \geq \varepsilon$. Como $\varepsilon > 0$ es arbitraria se tendrá que esta última ecuación es válida para toda $s > 0$ (ver la observación 1.5.3), lo cual es la definición de ser de cola pesada. \square

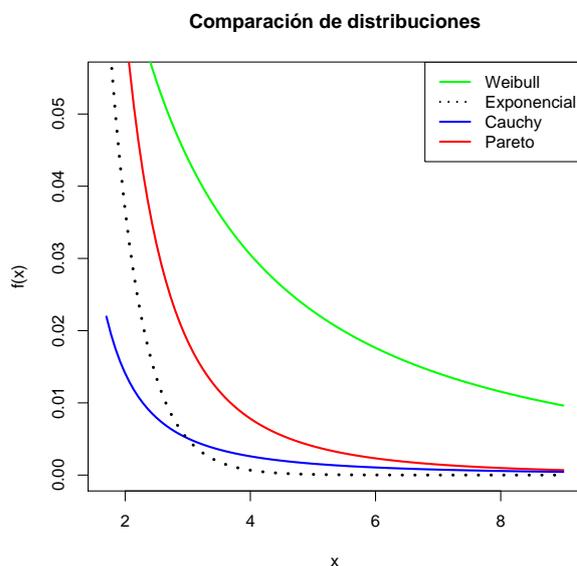


Figura 1.7: Esta es la continuación de las distribuciones de la figura anterior. En comparación con la otra figura, puede observarse la gran diferencia que existe de las colas de las distribuciones, en comparación con la cola de la distribución exponencial

1.5.1. Distribuciones subexponenciales

Si F es una distribución de cola pesada entonces se tiene

$$\lim_{x \rightarrow \infty} e^{sx} \bar{F}(x) = \infty$$

para toda $s > 0$. Esta es la razón por la cual se les llama distribuciones subexponenciales a un subconjunto de las distribuciones de colas pesadas. Este tipo de distribuciones tendrán la propiedad de que el decaimiento de su cola es lento (menor al decaimiento exponencial), esto por el hecho de que \bar{F} decae más lentamente que cualquier función exponencial (ya que su decaimiento es menor que el de $e^{-\varepsilon x}$ para toda $\varepsilon > 0$). Además se tiene que

$$\begin{aligned} m_F(\varepsilon) &= \int_0^{\infty} e^{\varepsilon x} dF(x) \\ &\geq \int_y^{\infty} e^{\varepsilon x} dF(x) \\ &\geq e^{\varepsilon y} \bar{F}(y) \xrightarrow{y \rightarrow \infty} \infty \end{aligned}$$

Lo cual muestra que no existe la función generadora de momentos en ninguna vecindad del origen.

Definición 1.5.3 (Distribución subexponencial). *Sea F una distribución en \mathbb{R}_+ , se dice que F es subexponencial si para toda $n \geq 2$*

$$\lim_{x \rightarrow \infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} = n. \quad (1.46)$$

Una manera más sencilla de comprobar que una distribución es subexponencial es por medio del siguiente lema, del cual su demostración se encuentra en el Apéndice D, en este apéndice también se encuentran diversos resultados que utilizaremos para algunas demostraciones por lo cual es recomendable consultarlo.

Lema 1.5.1 (Condición suficiente para ser subexponencial). *Sea F una distribución en \mathbb{R}_+ , si*

$$\limsup_{x \rightarrow \infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \leq 2 \quad (1.47)$$

entonces F es subexponencial

Terminaremos esta sección con un resultado que nos muestra el comportamiento asintótico de la probabilidad de ruina para el caso de que la cola integrada tenga distribución subexponencial.¹¹

Proposición 1.5.1. *[Relación asintótica cuando la distribución de los reclamos es subexponencial]Bajo el modelo C-L con $\mathbb{E}[X] < \infty$ y la condición de ganancia neta, suponiendo además que $\overline{F}_I(x)$ es subexponencial se tiene que la probabilidad de ruina cumple con*

$$\lim_{x \rightarrow \infty} \frac{\psi(x)}{\overline{F}_I(x)} = \rho^{-1} \quad (1.48)$$

Demostración. Por la fórmula de Pollaczek-Khinchin (1.42) tenemos que

$$\frac{\psi(x)}{\overline{F}_I(x)} = \frac{\rho}{1 + \rho} \sum_{n=1}^{\infty} (1 + \rho)^{-n} \frac{\overline{F_I^{*n}}(x)}{\overline{F}_I(x)}.$$

Como $\overline{F}_I(x)$ es subexponencial se tiene que

$$\lim_{x \rightarrow \infty} \frac{\overline{F_I^{*n}}(x)}{\overline{F}_I(x)} = n \quad \text{para } n \geq 1$$

¹¹Habrá que distinguir que la propiedad de ser subexponencial no corresponde directamente a la distribución del reclamo.

además de que por el teorema de la convergencia dominada (ver Apéndice B.2.2) y por el Lema D.2.3 (el cual puede ser consultado en el Apéndice D) es válido hacer el intercambio entre la integral y el límite. Por lo tanto tenemos que

$$\lim_{x \rightarrow \infty} \frac{\psi(x)}{\overline{F_I}(x)} = \frac{\rho}{1 + \rho} \sum_{n=1}^{\infty} (1 + \rho)^{-n} n = \rho^{-1}$$

con lo que se concluye la prueba. \square

Esta última proposición muestra que para valores grandes de x el cociente de la probabilidad de ruina y la cola integrada tienden a una constante, lo cual nos dice que la probabilidad de ruina vista como función del capital inicial tiene un decaimiento “muy lento” para valores grandes, bajo el supuesto de subexponencialidad.

1.6. Cotas de Lundberg

Siempre es útil tener ciertas cotas para las funciones que se están estudiando, ya que esto permite tener mayor información sobre ellas. Al simular surgen diversos problemas, entre los cuales se encuentra el de saber si nuestro resultado es factible o si hay algún factor que no ha sido incluido. Por lo tanto será útil tener alguna cota para las probabilidades de ruina.

Para el desarrollo de las cotas de Lundberg se necesitarán dos supuestos: la condición de ganancia neta y que la distribución de los reclamos sea de cola ligera.¹²

Definición 1.6.1 (Coeficiente de Lundberg). *Supongamos que la función generadora de Z_1 existe en alguna vecindad $(-s_0, s_0)$ del origen, con $s_0 > 0$. Si existe una única solución positiva r de la ecuación*

$$m_{Z_1}(s) = \mathbb{E} [e^{s(\xi_1 - cY_1)}] = 1 \quad (1.49)$$

entonces se llamará a r coeficiente de Lundberg.

Ejemplo 1.6.1. *Para el caso de que $\xi \sim \exp(\alpha)$ tenemos que*

$$m_{Z_1}(s) = \mathbb{E} [e^{s\xi_1}] \mathbb{E} [e^{-scY_1}] = \frac{\alpha}{\alpha - s} \frac{\lambda}{\lambda + sc} = 1,$$

lo cual implica que

$$r = \alpha - \frac{\lambda}{c}$$

¹²Es recomendable recordar la ecuación (1.18).

Lema 1.6.1. *Si existe r que cumpla con (1.49) entonces r es única.*

Demostración. Definimos la función $f(s) = m_{Z_1}(s)$, entonces f tiene derivadas de todos los ordenes en $(-s_0, s_0)$ ¹³. Además $f'(0) = \mathbb{E}[Z_1] < 0$ por la condición de ganancia neta y $f''(s) = \mathbb{E}[Z_1^2 e^{sZ_1}] > 0$ porque $Z_1 \neq 0$ casi seguramente.

La condición $f'(0) < 0$ junto con la continuidad de f implican que f decrece en alguna vecindad del cero. Por otro lado $f''(s) > 0$ implica que f es convexa. Estas dos propiedades de f hacen que si existe un $s^* \in (0, s_0)$ tal que $f'(s^*) = 0$ entonces f cambia su comportamiento de ser decreciente a creciente en s^* .

Entonces la solución r de la ecuación $f(s) = 1$ es única suponiendo que la función generadora de momentos de Z_1 existe en una vecindad del origen. \square

Observación 1.6.1. *Una condición suficiente para que esto ocurra es que exista un $s_1 \in (0, \infty]$ tal que $f(s) < \infty$ para $s < s_1$ y que $\lim_{s \rightarrow s_1} f(s) = \infty$. Esto implicaría, que la función generadora de momentos crece de manera continua al infinito y que en particular ésta asume el valor 1 para un valor s suficientemente grande.*

Observación 1.6.2. *El lema anterior muestra que la existencia de la función generadora de momentos de Z_1 en una vecindad del origen no es suficiente para asegurar la existencia del coeficiente de Lundberg.*

El siguiente teorema da una desigualdad, para el caso de reclamaciones de colas ligeras, la cual garantiza que la probabilidad de ruina sea muy pequeña dado un capital inicial x grande. Esta desigualdad también muestra que mientras menor sea el coeficiente de Lundberg mayor será la probabilidad de ruina.

Teorema 1.6.1 (Cota de Lundberg). *Suponiendo la condición de ganancia neta y la existencia del coeficiente de Lundberg r , entonces la siguiente desigualdad es válida para $x > 0$*

$$\psi(x) \leq e^{-rx} \tag{1.50}$$

Demostración. Para esta demostración definiremos la probabilidad de ruina hasta la n -ésima reclamación como

$$\psi_n(x) := \mathbb{P} \left[\max_{1 \leq k \leq n} S_k > x \right]$$

¹³Esta es una propiedad de las funciones generadoras de momentos

y utilizaremos el hecho de que esta probabilidad cumple con $\psi_n(x) \rightarrow \psi(x)$ cuando $n \rightarrow \infty$. Lo anterior por que

$$\alpha_n = \{ \max_{1 \leq k \leq n} S_k > x \}$$

forma una sucesión monótona decreciente de conjuntos que convergen a $\{ \max_{1 \leq k} S_k > x \}$. Por lo tanto bastará con probar que

$$\psi_n(x) \leq e^{-rx} \quad \text{para toda } n \geq 1, \quad x > 0, \quad (1.51)$$

usando inducción matemática sobre n . Además utilizaremos que

$$\psi_n(x) = \mathbb{P} \left[\max_{1 \leq k \leq n} S_k > x \right] = \mathbb{P} [S_k > x \quad \text{para algun } k \in \{1, \dots, n\}]. \quad (1.52)$$

Por la desigualdad de Markov (ver Apéndice B.1.3) se tiene que

$$\psi_1(x) \leq e^{-rx} m_{Z_1} = e^{-rx}$$

lo cual muestra que es válido para $n = 1$. Supongamos que es válido para $n = k \geq 1$ y utilizando la independencia de las Z_i para $i = 1, \dots$, se tiene que

$$\begin{aligned} \psi_{k+1}(x) &= \mathbb{P} \left[\max_{1 \leq n \leq k+1} S_n > x \right] \\ &= \mathbb{P} [Z_1 > x] + \mathbb{P} \left[\max_{2 \leq n \leq k+1} \{Z_1 + (S_n - Z_1)\} > x, Z_1 \leq x \right] \\ &= \int_x^\infty dF_{Z_1}(y) + \int_{-\infty}^x \mathbb{P} \left[\max_{1 \leq n \leq k+1} [y + S_n] > x \right] dF_{Z_1}(y) \\ &= p_1 + p_2 \end{aligned}$$

Considerando primero a p_2 y usando la hipótesis de inducción tenemos que

$$\begin{aligned} p_2 &= \int_{-\infty}^x \mathbb{P} \left[\max_{1 \leq n \leq k+1} S_n > x - y \right] dF_{Z_1}(y) = \int_{-\infty}^x \psi(x - y, k) dF_{Z_1}(y) \\ &\leq \int_{-\infty}^x e^{r(y-x)} dF_{Z_1}(y). \end{aligned}$$

Considerando a p_1 y utilizando la desigualdad de Markov tenemos que

$$p_1 = \mathbb{P} [X > x] \leq e^{rx} m_X(r) = e^{-rx} \int_0^\infty e^{ry} dF(y) \leq \int_x^\infty e^{r(y-x)} dF_{Z_1}(y).$$

Entonces por la definición del coeficiente de Lundberg

$$p_1 + p_2 \leq e^{rx} m_{Z_1}(r) = e^{-rx}$$

lo cual demuestra que la ecuación (1.51) es válida para $n = k + 1$, lo cual demuestra el teorema. \square

Observación 1.6.3. *Para el caso de que la distribución de los reclamos sea $\exp(\alpha)$ ya hemos obtenido tanto la probabilidad de ruina de forma analítica como un valor para el coeficiente de Lundberg (ver ejemplos 1.3.3 y 1.6.1). Al sustituir en la cota de Lundberg obtenemos*

$$\frac{\lambda}{\alpha c} e^{-(\alpha - \frac{\lambda}{c})x} \leq e^{-(\alpha - \frac{\lambda}{c})x},$$

lo cual es congruente ya que $\frac{\lambda}{\alpha c} < 1$ por la condición de ganancia neta.

En esta sección encontraremos dos funciones las cuales acotarán la función de distribución de las convoluciones geométricas, esto junto con el teorema de Pollaczek-Khinchin nos será de utilidad para acotar nuestra probabilidad de ruina. Si utilizamos (1.41) obtenemos

$$\begin{aligned} F_X(x) &= (1-p)\delta_0(x) + \sum_{k=1}^{\infty} (1-p)p^k F_U^{*k}(x) \\ &= (1-p)\delta_0(x) + \sum_{k=1}^{\infty} p(1-p)p^{k-1} F_U^{*k-1} * F_U(x) \\ &= (1-p)\delta_0(x) + \sum_{k=1}^{\infty} p(1-p)p^{k-1} F_U * F_U^{*k-1}(x) \\ &= (1-p)\delta_0(x) + pF_U * \sum_{k=1}^{\infty} (1-p)p^{k-1} F_U^{*k-1}(x) \end{aligned}$$

En este caso $\delta_0(x)$ significa la delta de Kronecker¹⁴

$$\delta_0(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$$

Por lo tanto tenemos que

$$F_X(x) = (1-p)\delta_0(x) + pF_U * F_X(x). \quad (1.53)$$

Supongamos que existe una $\gamma \geq 0$ tal que es una solución de

$$m_X(\gamma) = p^{-1}. \quad (1.54)$$

A esta $\gamma \geq 0$ le llamaremos coeficiente de ajuste.¹⁵ Si éste existe entonces

¹⁴Por como definimos a $\delta_0(x)$ la podemos sustituir por $\mathbb{1}_{\{x \in [0, \infty)\}}$

¹⁵El coeficiente de ajuste es distinto al coeficiente de Lundberg

también tendremos que $m_x(s) < \infty$ para $s \leq \gamma$ (ver la observación 1.5.2).

Ahora obtendremos las cotas de Lundberg para la cola de la distribución F_X , donde X se distribuye como una convolución geométrica de parámetros $\{p, F_U\}$. Utilizaremos la notación

$$x_0 = \sup\{x : F_U(x) < 1\}. \quad (1.55)$$

Teorema 1.6.2. *Si X es una convolución geométrica de parámetros $\{p, F_U\}$ y existe $\gamma \geq 0$ que cumple con (1.54) entonces*

$$a_- e^{-\gamma x} \leq \bar{F}_X(x) \leq a_+ e^{-\gamma x} \quad (1.56)$$

Donde

$$a_- = \inf_{x \in [0, x_0)} \frac{e^{\gamma x} \bar{F}_U(x)}{\int_x^\infty e^{\gamma y} d\bar{F}_U(y)} \quad a_+ = \sup_{x \in [0, x_0)} \frac{e^{\gamma x} \bar{F}_U(x)}{\int_x^\infty e^{\gamma y} d\bar{F}_U(y)} \quad (1.57)$$

Demostración. Notemos que a_+ no puede ser cero, ya que su numerador es positivo y su denominador es finito, entonces para $x \in [0, x_0)$ se tiene que

$$a_+ p \int_x^\infty e^{\gamma y} dF_U(y) \geq p e^{\gamma x} \bar{F}_U(x) \quad (1.58)$$

Utilizando que

$$1 = p \int_0^x e^{\gamma y} d\bar{F}_U(y) + p \int_x^\infty e^{\gamma y} d\bar{F}_U(y)$$

entonces (1.58) equivale a

$$a_+ \left(1 - p \int_0^x e^{\gamma y} d\bar{F}_U(y) \right) \geq p e^{\gamma x} \bar{F}_U(x) \quad (1.59)$$

lo cual no sólo es cierto para $x \in (x_0, \infty)$, sino también para $x \geq x_0$ (porque ambos lados de (1.59) son iguales a cero para $x \in [0, x_0)$). Se tiene que (1.59) es equivalente a

$$1 - p + p \left(F_U(x) + a_+ \int_0^x e^{\gamma(x-y)} dF_U(y) \right) \geq 1 - a_+ e^{-\gamma x}$$

Sea $F_0(x) = 1 - a_+ e^{-\gamma x} = (1 - a_+) \delta_0(x) + a_+ G(x)$ donde $G(x) = 1 - e^{-\gamma x}$ y definamos a $F_n = (1 - p) \delta_0 + p F_u * F_{n-1}$. Entonces para el caso de F_1 con

$x \geq 0$ tenemos lo siguiente

$$\begin{aligned}
F_1(x) &= 1 - p + pF_U * F_0(x) \\
&= 1 - p + p \left((1 - a_+)F_U + a_+ \int_0^x G(x-y)dF_U(y) \right) \\
&= 1 - p + p \left((1 - a_+)F_U + a_+ \int_0^x (1 - e^{-\gamma(x-y)})dF_U(y) \right) \\
&= 1 - p + p \left(F_U - a_+ \int_0^x e^{-\gamma(x-y)}dF_U(y) \right) \\
&\geq F_0(x)
\end{aligned}$$

Podemos ver que las convoluciones respetan desigualdades, esto es, si $F(x) \geq G(x)$ para toda $x \geq 0$ entonces

$$\int_0^x F(x-y)dF_U(y) \geq \int_0^x G(x-y)dF_U(y).$$

Entonces $F_U * F_1 \geq F_U * F_0$ y por lo tanto por inducción se tendrá que si $F_{k+1} \geq F_k$ entonces $F_U * F_{k+1} \geq F_U * F_k$ lo cual implica que

$$F_{k+2} = (1-p)\delta + pF_U * F_{k+1} \geq (1-p)\delta + pF_U * F_k = F_{k+1}.$$

De donde se sigue que $F_{n+1} \geq F_n$ para toda $n \geq 0$. La sucesión $\{F_n\}$ es creciente, además por como fue definida $\lim_{n \rightarrow \infty} F_n = F_X$. Por ser una sucesión creciente se tiene que $F_X \geq F_0$ lo cual implica que

$$\bar{F}(x) \leq a_+ e^{-\gamma x} \quad x \geq 0. \quad (1.60)$$

La demostración de la cota inferior es análoga. \square

Corolario 1.6.2.1. *Si existe $\gamma \geq 0$ tal que es una solución de $m_X(\gamma) = p^{-1}$ entonces la probabilidad de ruina $\psi(x)$ esta acotada por*

$$a_- e^{-\gamma x} \leq \psi(x) \leq a_+ e^{-\gamma x} \quad (1.61)$$

Donde

$$a_- = \inf_{x \in [0, x_0)} \frac{e^{\gamma x} \bar{F}_U(x)}{\int_x^\infty e^{\gamma y} d\bar{F}_U(y)} \quad a_+ = \sup_{x \in [0, x_0)} \frac{e^{\gamma x} \bar{F}_U(x)}{\int_x^\infty e^{\gamma y} d\bar{F}_U(y)} \quad (1.62)$$

Demostración. Utilizando el teorema anterior y la fórmula de Pollaczek-Khinchin el resultado es inmediato. \square

1.7. Fórmula de Panjer

Hasta ahora no tenemos, para el caso de distribuciones de colas pesadas, un criterio que establezca cotas para la probabilidad de ruina (además de saber que se encuentra en $[0,1]$) lo que más adelante cuando simulemos, nos será útil para comparar si nuestros algoritmos arrojan valores adecuados o no. Este tipo de cotas nos las dará el Teorema de Panjer 1.7.1. Pero antes de presentarlo necesitaremos definir cierta notación y demostrar algunos resultados.

Como ya hemos demostrado, la probabilidad de ruina puede ser modelada con el uso de una suma geométrica

$$S_M = \sum_{i=1}^M X_i.$$

Ahora supongamos que los sumandos tienen una distribución discreta que toma valores en $\mathbb{N} \cup \{0\}$ y definamos

$$S_M^* := \sum_{i=1}^M X_i^*, \quad (1.63)$$

como la suma aleatoria de variables aleatorias X_i^* donde el número de sumandos tiene distribución geométrica de parámetro p y los sumandos cumplen con $\mathbb{P}[X_i^* \in \mathbb{N} \cup \{0\}] = 1$. Hay que notar que la variable aleatoria S_M^* toma valores en $\mathbb{N} \cup \{0\}$.

Nuestra meta es demostrar una fórmula recursiva para el cálculo de la probabilidad de ruina. Para simplificar la notación definamos las siguientes probabilidades.

$$\begin{array}{ll} p_m := \mathbb{P}[M = m] & \text{para } m = 0, 1, \dots \\ f_r := \mathbb{P}[X^* = r] & \text{para } r = 0, 1, \dots \\ \delta_r^* := \mathbb{P}[S^* = r] & \text{para } r = 0, 1, \dots \\ f_r^{m*} := \mathbb{P}[X_1^* + \dots + X_m^* = r] & \text{para } r = 0, 1, \dots \end{array}$$

Por lo anterior tenemos que

$$p_0 = p$$

$$p_m = (1 - p)p_{m-1} \quad \text{para} \quad m = 1, 2, \dots$$

$$\delta_r^* = \sum_{m=1}^{\infty} \mathbb{P}[S^* = r \mid M = m] \mathbb{P}[M = m] \quad \text{para} \quad r \geq 1$$

$$= \sum_{m=1}^{\infty} f_r^{m*} p_m \quad \text{para} \quad r \geq 1$$

Utilizando la notación anterior y recordando que $f_r^{0*} = 1$ tenemos la siguiente proposición.

Proposición 1.7.1.

$$f_r^{m*} = \sum_{i=0}^r f_{r-i}^{*(m-1)} f_i \quad \text{para} \quad m = 1, 2, \dots \quad (1.64)$$

Demostración.

$$\begin{aligned} \sum_{i=0}^r f_{r-i}^{*(m-1)*} f_i &= \sum_{i=0}^r \mathbb{P}[X_2^* + \dots + X_m^* = r - i] \mathbb{P}[X_1^* = i] \\ &= \sum_{i=0}^r \mathbb{P}[X_1^* = i, X_2^* + \dots + X_m^* = r - i] \\ &= \sum_{i=0}^r \mathbb{P}\left[X_1^* = i, \sum_{j=1}^m X_j^* = r\right] \\ &= \sum_{i=0}^r \mathbb{P}\left[X_1^* = i \mid \sum_{j=1}^m X_j^* = r\right] f_r^{m*} \\ &= f_r^{m*} \sum_{i=0}^r \mathbb{P}\left[X_1^* = i \mid \sum_{j=1}^m X_j^* = r\right] \\ &= f_r^{m*} \end{aligned}$$

□

Ahora contamos con las herramientas necesarias para demostrar el siguiente teorema.

Teorema 1.7.1. [Teorema de Panjer] La probabilidad $\delta_r^* = \mathbb{P}[S^* = r]$ está dada por

$$\begin{aligned}\delta_r^* &= \frac{1-p}{1-(1-p)f_0} \sum_{i=1}^r f_i \delta_{r-i}^*, \quad \text{para } r \geq 1 \\ \delta_0^* &= \frac{p}{1-(1-p)f_0}.\end{aligned}\tag{1.65}$$

Demostración. Para $r = 0$ se tiene que

$$\begin{aligned}\delta_r^* &= \sum_{m=0}^{\infty} p_m f_0^{m*} \\ &= \sum_{m=0}^{\infty} p(1-p)^m (f_0)^m \\ &= p \frac{1}{1-(1-p)f_0},\end{aligned}$$

con lo cual queda demostrado para $r = 0$. Para $r \geq 1$ se sigue que

$$\begin{aligned}\delta_r^* &= \sum_{m=1}^{\infty} p_m f_r^{m*} \\ &= \sum_{m=1}^{\infty} (1-p)p_{m-1} \sum_{i=0}^r f_{r-i}^{(m-1)*} f_i \\ &= \sum_{i=0}^r (1-p)f_i \sum_{m=1}^{\infty} p_{m-1} f_{r-i}^{(m-1)*} \\ &= \sum_{i=0}^r (1-p)f_i \delta_{r-i}^* \\ &= (1-p)f_0 \delta_r^* + (1-p) \sum_{i=1}^r f_i \delta_{r-i}^*,\end{aligned}$$

entonces al despejar δ_r^* de la última ecuación se obtiene el resultado deseado. \square

La fórmula anterior tiene una interpretación muy sencilla; es decir, se puede observar que utiliza la idea de “fijar” un sumando con la cantidad i y al tener esto en cuenta el resto de los sumandos tienen (para que la suma valga r) que valer $r - i$, luego ir variando entre todos los posibles valores de i (los cuales son $0, \dots, r$). Una gran ventaja que tiene esta ecuación al ser recursiva

(además de ser computacionalmente sencilla) es que automáticamente esta tomando todas las combinaciones posibles para obtener un monto de tamaño r ya que en caso contrario se tendrían que hacer todas las combinaciones de sumandos tomados de $\{\mathbb{N} \cup \{0\}\}$ que sumen r .

La utilidad del teorema anterior es que nos permite acotar la probabilidad de ruina al “discretizar” nuestros sumandos X_i , ya que si definimos para $\kappa > 0$ las variables aleatorias

$$\begin{aligned}\bar{X}_i &= \inf\{n \in \mathbb{N} : X_i \leq n\kappa\} \\ \underline{X}_i &= \sup\{n \in \mathbb{N} : X_i \geq n\kappa\}\end{aligned}$$

donde utilizamos que $\inf \emptyset = 0$. Entonces podemos definir las sumas aleatorias enteras como $\bar{S} = \sum_{i=1}^M \bar{X}_i$ y $\underline{S} = \sum_{i=1}^M \underline{X}_i$. Se tendrá entonces que $\kappa\bar{S} \leq S \leq \kappa\underline{S}$, por lo tanto

$$\mathbb{P}[\bar{S} \leq x/\kappa] \leq \mathbb{P}[S \leq x] \leq \mathbb{P}[\underline{S} \leq x/\kappa] \quad (1.66)$$

Para nuestro estudio utilizaremos $\kappa = 1$ para facilitar cálculos.¹⁶

Por lo anterior se sigue que si utilizamos

$$\begin{aligned}\delta_r^* &= \frac{1}{1 + \theta - f_0} \sum_{i=1}^r f_i \delta_{r-i}^*, \quad \text{para } r \geq 1 \\ \delta_0^* &= \frac{\theta}{1 + \theta - f_0},\end{aligned}$$

podemos aproximar la probabilidad de ruina por medio de

$$\psi(x) \approx 1 - \sum_{r=0}^x \delta_r^* \quad (1.67)$$

y al sustituir f_r por

$$\begin{aligned}f_x^l &= F_I(\lfloor x \rfloor + 1) - F_I(\lfloor x \rfloor) \\ f_x^u &= F_I(\lfloor x \rfloor) - F_I(\lfloor x \rfloor - 1),\end{aligned}$$

obtendremos tanto una cota superior como una inferior para la probabilidad de ruina (donde $\lfloor x \rfloor$ representa la parte entera de x).

¹⁶Por como fué definido a \bar{S} y a \underline{S} se observa que mientras menor sea κ mejor será la aproximación, pero bastaría con cambiar los parámetros de nuestras distribuciones para superar este problema.

Al consultar Dickson(1995)[12] se obtiene que una buena aproximación a la probabilidad de ruina es $\psi_P(x)$ la cual se obtiene al utilizar

$$f_x^P = F_I(\lfloor x \rfloor + 1/2) - F_I(\lfloor x \rfloor - 1/2).$$

Estas últimas fórmulas nos permitirán comparar qué tan eficientes son los algoritmos y si es que son necesarios. El inconveniente que tienen, tanto el cálculo de la aproximación como el de las cotas, es que el número de operaciones visto como función del capital inicial crece de manera enorme, de hecho crece como la serie de Fibonacci.

Capítulo 2

Simulación

“**simular**. (Del lat. «simulare», deriv. de «símilis».)
«Aparentar. Fingir». Hacer parecer que existe u ocurre una cosa que no existe o no ocurre: ‘Simula una cojera. Simuló un accidente’. En los juegos de niños se expresan las simulaciones con el pretérito imperfecto de «ser»: ‘Esto era un castillo. Yo era el capitán’...”

María Moliner.-Diccionario de uso del español H-Z, 1973(pag. 1169).

El tema central del trabajo es la simulación de la probabilidad de ruina para el caso en que los reclamos tengan distribuciones de colas pesadas. Expondremos en este capítulo los diversos problemas que plantea este tipo de estimaciones y las diversas técnicas para hacerles frente. Por lo tanto, será necesario dar una pequeña introducción al tema de la simulación. Para este estudio definiremos como simulador a un programa de computadora (algoritmo) que represente a un fenómeno tomando en cuenta las variables que actúan sobre él (no necesariamente todas las variables, ya que por motivos de sencillez se puede solamente hacer uso de las que se consideren relevantes) y que arroje resultados similares al fenómeno. Este tipo de programas se realizan con la finalidad de deducir ciertas propiedades del fenómeno a través de los resultados que genere.

Por ejemplo en el caso de no tener una solución analítica de alguna ecuación se puede, por medio de un modelo de computadora, representar su comportamiento y con esto aproximar su solución; con base en esta aproximación podemos realizar cierta inferencia sobre la ecuación original. Un ejemplo de esto es el cálculo de integrales en varias dimensiones cuando no conocemos el comportamiento de la función a integrar, ya que puede ser un gran problema el hacer particiones de los intervalos de integración. Entonces se puede realizar una serie de simulaciones que arrojen una partición adecuada.

Una ventaja de la simulación es poder manipular las variables que rigen al fenómeno para así hacer inferencia sobre las reacciones del fenómeno real respecto a estos cambios, cosa que puede ser difícil o en ciertos casos imposible con el modelo real ya que existen ciertas condiciones que no podríamos cambiar (por ejemplo cuando queremos conocer como influye la luz solar en un cultivo). También la simulación es importante en el caso de que la información reportada acerca del fenómeno sea escasa (por ejemplo en el caso de un estudio sobre terremotos) es entonces útil un modelo con el cual se puedan obtener datos que representen una mayor cantidad de información.

Otra motivación para la simulación es la existencia de fenómenos que se pueden considerar como aleatorios o que por su contexto las condiciones que los rigen son lo suficientemente complejas para que no se puedan considerar como deterministas (por ejemplo la rapidez con que se disemina una pandemia, la probabilidad de que un huracán aparezca en los próximos tres años o el precio de ciertas acciones en el mercado bursátil). Es importante el estudio de estos fenómenos para tomar ciertas decisiones y con la simulación podemos deducir ciertas propiedades utilizando un modelo que incorpore este tipo de aleatoriedad.

En general, se simula para deducir el valor de cierta cantidad ζ desconocida relacionada con un modelo estocástico. Una simulación del sistema en cuestión produce los datos de salida Z , una variable aleatoria cuyo valor esperado es la cantidad de interés ζ . Una nueva simulación independiente de la anterior proporciona una nueva variable aleatoria con media ζ . Esto continúa hasta un total de R ejecuciones y R variables aleatorias independientes Z_1, \dots, Z_R , todas con la misma distribución y media ζ . El promedio de estos R valores, $\hat{\zeta} = \sum_{r=1}^R Z_r / R$ sirve entonces como estimador (como una aproximación) de ζ .

Existen diversos problemas en la simulación. Uno de ellos es decidir cuántas simulaciones hacer; es decir, el problema de hallar un valor adecuado para R (considerando que necesitamos tener cierta confianza en nuestro estimado) dado que mientras mayor fuera nuestra muestra mejor será la calidad de nuestro estimador. Este es un problema en el que nos enfocaremos ya que existirán casos en los que dado un algoritmo se necesitará una cantidad casi infinita para obtener la precisión deseada. Debido a la naturaleza aleatoria¹ de nuestro estimador $\hat{\zeta}$ quisiéramos que dado un valor $\varepsilon > 0$ saber qué tan probable es que nuestro estimador $\hat{\zeta}$ esté a una distancia menor que ε del valor ζ .

¹En realidad ningún programa puede arrojar números verdaderamente aleatorios, lo que arroja son números pseudoaleatorios con propiedades que permiten utilizarlos.

Definiremos como un estimador puntual a una expresión matemática en la cual se introducen los valores obtenidos de las simulaciones para producir un único número, que utilizamos como una aproximación del valor buscado. Por otro lado, definimos como intervalo de confianza a una expresión matemática en la cual se introducen los datos para así producir un intervalo numérico que cubrirá el valor verdadero con una cierta probabilidad de antemano especificada; es decir, un intervalo de confianza es un intervalo generado a través de los datos que nos permite tener una noción de la posición del valor que se esté estudiando, con la confianza de que el valor cae dentro de este intervalo con cierta probabilidad.

2.1. Estimadores y su eficiencia

Un estimador es un valor con el cual se aproxima una cantidad desconocida, este valor es el resultado de ejecutar un proceso al cual llamaremos algoritmo de simulación o simplemente algoritmo. Al estar estimando una cantidad desconocida surge la duda de qué tan bien se está estimando a la cantidad, o dado que necesitamos un cierto grado de confianza de nuestro estimador si es que él lo cumple o no. Por lo anterior será necesario tener ciertos criterios para nuestros estimadores. El siguiente método nos presenta una forma de obtener estimadores por medio de obtener el promedio de simulaciones del proceso que rige a nuestro valor.

2.1.1. Método de Monte Carlo crudo y estimadores insesgados

En diversos estudios surge el problema de calcular una cantidad desconocida $\zeta \in \mathbb{R}$. Supongamos que podemos hacer un modelo que emule el comportamiento que rige a ζ y que éste genere datos “cercaños” al valor desconocido. La idea que plantea el método de Monte Carlo crudo (o simplemente método de Monte Carlo) es tomar el promedio de estos valores como un estimador de ζ .

La idea anterior nos lleva a querer simular² repetidas veces (por ejemplo R veces) una v.a. Z que tenga una media ζ . El método de Monte Carlo crudo implementa esta “cercaña” al pedir que la variable aleatoria Z tenga esperanza ζ , a este tipo de estimadores cuya esperanza es el valor estimado

²Más adelante daremos el significado de simular a las variables aleatorias, por ahora debe de entenderse como realizar un proceso que asemeje los resultados de ellas.

se les llama estimadores insesgados³. Entonces el método de Monte Carlo plantea realizar R simulaciones de la v.a. Z y luego utiliza

$$\widehat{\zeta}_{MCC} := \widehat{\zeta} := \frac{1}{R} \sum_{r=1}^R Z_r, \quad (2.1)$$

como estimador de ζ , donde Z_r es el valor que arrojó la simulación r . La razón por la cual es válido este método es que utilizando la ley fuerte de los grandes números podemos asegurar que $\widehat{\zeta} \rightarrow \zeta$ con probabilidad 1 cuando $R \rightarrow \infty$, lo que nos dice que mientras mayor sea el valor de R más probable es que nuestro estimador se aproxime a ζ . Entonces se tiene el problema de encontrar a Z de tal forma que la velocidad de convergencia de nuestro estimador sea la adecuada, según las exigencias de nuestro problema.

Para fines prácticos lo anterior nos dice que si podemos “simular” a la variable aleatoria Z obteniendo los valores z_1, \dots, z_R entonces el promedio de estos valores nos será útil como una aproximación al valor de ζ . Uno entonces tiene el problema de entender que significa “simular” una variable aleatoria, ¿cómo se puede “simular” una función real medible?, además ¿cómo se puede simular un proceso “aleatorio” si sólo tengo a mi disposición el hacer procesos deterministas?. No existe una respuesta exacta⁴ para las preguntas anteriores, lo que sí se tiene es una interpretación de su significado; es decir, cuando uno habla de simular una variable aleatoria se refiere al hecho de representar un proceso aleatorio. Por aleatorio se debe de entender como un proceso el cual no se conocen los suficientes elementos para poder predecir su resultado. En este caso al querer representar un proceso aleatorio lo que se hace es crear una variable pseudo-aleatoria, la cual es en realidad el resultado de un proceso determinista que al tomar en cuenta ciertos criterios de lo que es y lo que no es aleatorio cumple con ser aleatorio.

Por ejemplo, cuando se quiere simular variables aleatorias con distribución $\mathcal{U}(0,1)$ una forma de hacerlo es tomar un número $x_0 \in (0,1)$ (el cual se denomina semilla) y realizar las operaciones siguientes

$$x_n = \alpha x_{n-1} \text{ modulo } (\beta), \quad \text{con } n \in \mathbb{N}$$

donde $\alpha, \beta \in \mathbb{N}$ fijos. Entonces los valores obtenidos x_1, x_2, \dots se utilizan como los resultados de la simulación.

Existen diversos problemas en los cuales se está interesado en estimar cierta cantidad desconocida ζ , la cual es la probabilidad de cierto evento

³En general se puede hacer uso de v.a. no insesgadas (sesgadas) para hacer inferencia, lo cual dependerá de la naturaleza de nuestro problema.

⁴No entremos en detalles de que significa “exacta”.

(por ejemplo la probabilidad de ruina); es decir $\zeta = \mathbb{P}[Z \in \mathcal{A}]$ donde Z es una variable aleatoria y \mathcal{A} es un subconjunto de su dominio. Se sigue que si uno puede simular a Z y calcular $1_{\{Z \in \mathcal{A}\}}$ se podrá utilizar el método de Monte Carlo al utilizar la ecuación (2.5) y obtener

$$\widehat{\zeta} = \frac{1}{R} \sum_{r=1}^R 1_{\{Z_r \in \mathcal{A}\}}, \quad (2.2)$$

donde se utiliza el hecho de que $\mathbb{E}[1_{\{Z \in \mathcal{A}\}}] = \mathbb{P}[Z \in \mathcal{A}]$. Lo anterior tiene una importante interpretación; al estar utilizando a $1_{\{Z \in \mathcal{A}\}}$ lo que se está realmente tomando en cuenta es la ocurrencia de la variable Z en el conjunto \mathcal{A} , es decir, se aporta un 1 en el caso de que $Z \in \mathcal{A}$ y cero en caso de que $Z \in \mathcal{A}^c$ lo cual es justamente la definición de una v.a. con distribución *Bernoulli* ($p = \mathbb{P}[Z \in \mathcal{A}]$). Se sigue entonces que la varianza de nuestro estimador es

$$\text{Var} \widehat{\zeta} = \zeta(1 - \zeta), \quad (2.3)$$

la cual será una ecuación importante que utilizaremos principalmente cuando desarrollemos la teoría de eventos raros.

En un planteamiento más general podría interesarnos simular una variable aleatoria W y utilizar el método de Monte Carlo. Puede suceder que la simulación directa de W no sea posible, en cambio tenemos acceso a simular una variable aleatoria Z la cual bajo cierta transformación H cumpla con $W = H(Z)$ y entonces tenemos que $\zeta = \mathbb{E}[H(Z)]$, donde la esperanza está siendo tomada bajo la distribución de Z y no de W .

Lo anterior implica que se puede replantear el problema a encontrar una variable aleatoria Z que cumpla con

$$\zeta = \mathbb{E}_F[H(Z)] = \int_{\Omega} H(z) dF(z), \quad (2.4)$$

donde F es la distribución de Z y Ω es su soporte. Por lo tanto un estimador insesgado para ζ es

$$\widehat{\zeta} := \frac{1}{R} \sum_{r=1}^R H(Z_r) \quad (2.5)$$

donde Z_1, \dots, Z_R son simulaciones independientes de la variable aleatoria Z .

2.1.2. Eficiencia de los estimadores

Un problema que surge al simular es comparar la eficiencia⁵ de los distintos algoritmos de simulación, por lo tanto se necesitará plantear criterios de comparación.⁶ Utilizando el teorema central del límite (ver Apéndice B.1.8) tenemos que para $\hat{\zeta}$ dado por (2.5) tiene aproximadamente la distribución $\mathcal{N}(\zeta, R^{-1}Var(Z))$ para R grande, además utilizando la varianza muestral que es un estimador insesgado de $Var(Z)$ definido por

$$S^2 = \frac{1}{R-1} \sum_{r=1}^R (Z_i - \hat{\zeta})^2, \quad (2.6)$$

se obtendrá por la ley fuerte de los grandes números que dicho estimador converge con probabilidad 1 a $Var(Z)$. Por el teorema de Slutsky (ver B.1.2) tenemos que $Var(Z)$ puede ser sustituida por S^2 y seguirá convergiendo a una v.a. con distribución normal. Entonces para $Var(Z) < \infty$ ⁷ y R grande tenemos que el intervalo de confianza alrededor de $\hat{\zeta}$ al cual ζ pertenece con probabilidad $(1 - \alpha)$ está dado por

$$\left(\hat{\zeta} - z_{1-\alpha/2} \frac{S}{\sqrt{R}}, \hat{\zeta} + z_{1-\alpha/2} \frac{S}{\sqrt{R}} \right), \quad (2.7)$$

donde $z_{1-\alpha}$ es el cuantil $(1 - \alpha)$ de una distribución normal estándar.

Un factor importante es el de ponderar qué tan bueno es el intervalo de confianza con respecto al estimador ya que dependiendo del valor estimado podremos dar mayor o menor libertad al tamaño del intervalo. Por ejemplo cuando uno está aproximando probabilidades un intervalo que tuviera una logitud de 2 no nos sería útil ya que las probabilidades se encuentran solamente en el intervalo $[0, 1]$ y un intervalo así estaría reportando valores fuera del dominio, en cambio si se estuviera estimando el tamaño de una población de un país, este intervalo podría pensarse que es bueno, ya que nos estaría reportando que nuestro estimador a lo más tiene un error de una persona.

⁵Esta eficiencia se refiere entre otras cosas a la velocidad de convergencia de nuestro estimador al valor estimado.

⁶La literatura escrita con respecto a criterios de comparación de algoritmos y estimadores es sumamente extensa por lo cual sólo se presentaran los resultados necesarios para nuestro estudio, para mayor referencia consultar Rubinstein(1981)[35] y Dimov(2008)[13].

⁷En el caso de simular eventos con varianza infinita surgen otro tipo de problemas, como por ejemplo que nuestras simulaciones nunca reportarán una varianza muestral infinita.

Por lo tanto una cantidad que se puede tomar como medida de dispersión es

$$\frac{S/\sqrt{R}}{\widehat{\zeta}}$$

la cual para R grande converge al error relativo del estimador $\widehat{\zeta}$, definido por

Definición 2.1.1 (Error relativo). *Se define el error relativo ε del estimador $\widehat{\zeta}$ como*

$$\varepsilon := \frac{\sqrt{\text{Var}(\widehat{\zeta})}}{\mathbb{E}[\widehat{\zeta}]} \quad (2.8)$$

Entonces el error relativo del estimador $\widehat{\zeta}$ viene dado por

$$\varepsilon = \frac{\sqrt{\text{Var}(\widehat{\zeta})}}{\mathbb{E}[\widehat{\zeta}]} = \frac{\sqrt{\text{Var}(Z)/R}}{\zeta} \quad (2.9)$$

De esta última ecuación se puede ver que para reducir el error relativo de nuestro estimador existen dos formas. La primera es por medio de aumentar el número de simulaciones hasta lograr el error relativo deseado, el problema que surge es que existen ocasiones en las que no es práctico, ya que exigen un número de simulaciones prácticamente infinito. La segunda forma de reducir el error relativo es a través de reducir la varianza de nuestros estimadores; para este fin existen los métodos de reducción de varianza los cuales utilizan información a priori que incorporan a la simulación para lograr reducir la varianza, los cuales expondremos en la Sección 2.3.

La razón por la cual el error relativo es importante es que nos brinda una noción de qué tanto está variando nuestro estimador con respecto al valor estimado. Por ejemplo, supongamos que se quiere estimar $\zeta = \mathbb{P}[Z \in \mathcal{A}]$ la cual es pequeña (casi cero), entonces se tendrá que al utilizar las ecuaciones (2.3) y (2.2) el error relativo del estimador de Monte Carlo está dado por

$$\varepsilon = \frac{\sqrt{\text{Var}(\widehat{\zeta})}}{\mathbb{E}[\widehat{\zeta}]} = \sqrt{\frac{1-\zeta}{R\zeta}} \approx \sqrt{\frac{1}{R\zeta}}, \quad (2.10)$$

lo cual representa un gran problema al momento de la simulación, ya que por hipotesis ζ dijimos que era pequeña, lo que implica que el error relativo es grande y por ende el intervalo de confianza. Más adelante en la Sección 2.2 se exhibirán los problemas que surgen al estar estimando probabilidades pequeñas por medio de plantearlas como eventos raros.

En la ecuación (2.9) se exhibe una de las razones por las cuales es importante hacer una reducción de la varianza. Se observa que al reducir la varianza se logrará que el error relativo sea menor y por ende que el intervalo de confianza se haga más pequeño y con esto lograr el grado de precisión que requiera nuestro estimador. Lo anterior sugiere que se buscará una reducción de la varianza de nuestro estimador que la disminuya a la mitad, a una décima parte o a una cantidad aún menor, para que sea efectivo este nuevo método de simulación, ya que en caso contrario se podría tan solo repetir unas cuantas veces más la simulación para lograr la precisión deseada.

Además, al observar la fórmula (2.10), se observa que para reducir el error estándar de un estimador en un factor de α se necesitarán alrededor de α^2 número de simulaciones. Esto significa que para un aumento relativamente pequeño en la precisión es necesario un gran aumento en el número de simulaciones.

2.1.3. Error relativo acotado y eficiencia logarítmica

El estimar la probabilidad de ruina de manera adecuada tiene diversos problemas, entre los que destaca que al trabajar con probabilidades de ruina pequeñas necesitaremos que el intervalo de confianza de nuestros estimadores sea también pequeño. Por ejemplo si estamos estimando por el método de Monte Carlo crudo una probabilidad de ruina del orden de 10^{-7} , necesitaríamos hacer 10^{14} simulaciones para lograr que nuestro valor estimado fuera del mismo orden que el valor que está estimando, lo cual se suma a que (por estar estimando por medio de promediar ceros y unos) la varianza de nuestro estimador va a tener graves problemas.⁸ Lo anterior plantea un gran problema al momento de simular ya que el método de Monte Carlo crudo se vuelve ineficiente no solo en su error relativo sino también en la estimación. En la Sección 2.2 presentaremos lo que son los eventos raros los cuales son un caso más general del problema anterior.

Considere un estimador $\widehat{\zeta}(x)$ para $\zeta(x)$, en este caso estamos planteando tanto a los estimadores como a los valores estimados como funciones de un parámetro x y supongamos que ζ decrece conforme $x \rightarrow \infty$. Utilizando la desigualdad de Jensen tenemos que

$$\mathbb{E} \left[\widehat{\zeta}(x)^2 \right] \geq \left(\mathbb{E} \left[\widehat{\zeta}(x) \right] \right)^2 = \zeta(x)^2.$$

⁸De hecho en el capítulo 3 mostraremos que al simular el método de Monte Carlo este presenta una varianza mínima y en algunos casos cero, ya que durante su ejecución solo fue simulada la constante cero, lo cual representa un pésimo estimador.

Quisiéramos entonces garantizar que el segundo momento de nuestro estimador $\widehat{\zeta}(x)$ convergiera proporcionalmente a $\zeta^2(x)$ cuando $x \rightarrow \infty$. Esto último para que nuestra varianza convergiera de manera similar a $\zeta^2(x)$ y con esto lograr que el error relativo de nuestras estimaciones se minimizara; es decir quisiéramos lograr que

$$\mathbb{E} \left[\widehat{\zeta}(x)^2 \right] \leq c\zeta(x)^2,$$

para toda x y donde c es una constante fija mayor o igual a 1. Lo anterior nos lleva a la siguiente definición.

Definición 2.1.2 (Error relativo acotado). *Se dice que un estimador $\widehat{\zeta}(x)$ de $\zeta(x)$ tal que $\zeta(x) \rightarrow 0$ cuando $x \rightarrow \infty$ tiene error relativo acotado si*

$$\limsup_{x \rightarrow \infty} \frac{\text{Var}\widehat{\zeta}(x)}{\zeta(x)^2} < \infty \quad (2.11)$$

Durante mucho tiempo se creía que lograr estimadores para la probabilidad de ruina con error relativo acotado era imposible, pero en fechas recientes se logró obtener el primer estimador de la probabilidad de ruina con esta propiedad, el cual fue publicado en el artículo de Asmussen, S. y D. P. Kroese (2006)[4]. Se decidió que para este trabajo no era apropiado el exhibir a este estimador ya que para su implementación se requieren demasiadas hipótesis extra y nosotros estamos planteando un escenario mucho más amplio.

Como un criterio más relajado (ya que tener un error relativo acotado es difícil) se tiene el siguiente

Definición 2.1.3 (Eficiencia logarítmica). *Se dice que nuestro estimador $\widehat{\zeta}(x)$ de $\zeta(x)$ tal que $\zeta(x) \rightarrow 0$ cuando $x \rightarrow \infty$ es logarítmicamente eficiente (o asintóticamente óptimo) si cumple con*

$$\limsup_{x \rightarrow \infty} \frac{\text{Var}\widehat{\zeta}(x)}{\zeta(x)^{2-\varepsilon}} = 0 \quad (2.12)$$

para toda $\varepsilon > 0$.

Por como fue definido es obvio que tener error relativo acotado implica tener eficiencia logarítmica. El siguiente teorema plantea una forma alternativa de definir la eficiencia logarítmica, además de ser el medio que utilizaremos para demostrar que un estimador es logarítmicamente eficiente a la par que explica razón del nombre de eficiencia logarítmica.

Teorema 2.1.1 (Criterio alternativo de eficiencia logarítmica). *Un estimador $\widehat{\zeta}(x)$ de $\zeta(x)$ tal que $\zeta(x) \rightarrow 0$ cuando $x \rightarrow \infty$ es logarítmicamente eficiente si solo si cumple con*

$$\liminf_{x \rightarrow \infty} \frac{|\ln \text{Var} \widehat{\zeta}(x)|}{|\ln \zeta(x)^2|} \geq 1. \quad (2.13)$$

Demostración. Supongamos que $\widehat{\zeta}(x)$ cumple con la ecuación (2.12). Entonces para toda $\delta \in (0, 1)$ se puede encontrar un valor x_0 tal que si $x > x_0$ $\ln \text{Var} \widehat{\zeta}(x) < 0$, $\ln \zeta(x) < 0$ y además que cumpla con

$$\frac{\text{Var} \widehat{\zeta}(x)}{\zeta(x)^{2-\varepsilon}} < \delta$$

lo que implica que

$$\ln \text{Var} \widehat{\zeta}(x) < \ln \delta + (2-\varepsilon) \ln \zeta(x) \quad \Leftrightarrow \quad \frac{\ln \text{Var} \widehat{\zeta}(x)}{\ln \zeta(x)^2} > \frac{\ln \delta}{2 \ln \zeta(x)} + (1 - \frac{\varepsilon}{2}).$$

Por lo tanto

$$\begin{aligned} \liminf_{x \rightarrow \infty} \frac{|\ln \text{Var} \widehat{\zeta}(x)|}{|\ln \zeta(x)^2|} &\geq \sup_{x \rightarrow \infty} \inf_{k > x} \left| \frac{\ln \delta}{\ln \zeta(x)} + 1 - \frac{\varepsilon}{2} \right| \\ &\geq 1 - \frac{\varepsilon}{2} \\ &\geq 1, \end{aligned}$$

con lo cual se demuestra una implicación. Para obtener la otra implicación, suponemos que se cumple la ecuación (2.13), entonces existe una x_0 tal que si $x > x_0$ $\ln \text{Var} \widehat{\zeta}(x) < 0$, $\ln \zeta(x) < 0$ y además cumple que para toda $\varepsilon > 0$ se tenga

$$\frac{|\ln \text{Var} \widehat{\zeta}(x)|}{|\ln \zeta(x)^2|} \geq 1 - \frac{\varepsilon}{2} + \frac{1}{x}.$$

Se sigue que

$$|\ln \text{Var} \widehat{\zeta}(x)| \geq |\ln \zeta(x)^{2-\varepsilon+\frac{2}{x}}|.$$

En la última ecuación cada cantidad encerrada en valor absoluto es negativa, por tanto

$$\ln \text{Var} \widehat{\zeta}(x) \leq \ln \zeta(x)^{2-\varepsilon+\frac{2}{x}} \quad \Leftrightarrow \quad \frac{\text{Var} \widehat{\zeta}(x)}{\zeta(x)^{2-\varepsilon}} \leq \zeta(x)^{\frac{2}{x}}$$

utilizando que $\zeta(x) \rightarrow 0$ cuando $x \rightarrow \infty$ se completa la prueba. \square

El teorema anterior nos será de gran utilidad ya que con él se demostrará si un estimador es logarítmicamente eficiente o no; la siguiente proposición muestra que el estimador de Monte Carlo no lo es.

Proposición 2.1.1. *El estimador de Monte Carlo crudo no es logarítmicamente eficiente*

Demostración. Considere (2.14) se tiene que

$$\sigma_Z^2 \leq \mathbb{E}[Z^2] = \mathbb{E}[Z] = \zeta$$

Entonces

$$\liminf_{x \rightarrow \infty} \frac{\ln \sigma_Z}{\ln \zeta} \leq \lim_{x \rightarrow \infty} \frac{\ln \mathbb{E}[Z^2]^{1/2}}{\ln \zeta} = \lim_{x \rightarrow \infty} \frac{1}{2} \frac{\ln \mathbb{E}[Z^2]}{\ln \zeta} = \frac{1}{2}.$$

Por lo tanto el estimador de M.C. crudo no es logarítmicamente eficiente. \square

La conclusión de la proposición anterior es que el estimador M.C. crudo no es logarítmicamente eficiente y por lo tanto es necesario encontrar mejores estimadores para el caso de eventos raros.

2.2. Eventos raros

Como ya hemos mencionado existen fenómenos que aun cuando su probabilidad es baja su ocurrencia genera grandes cambios en su entorno. Para modelarlos utilizaremos a los eventos raros, los cuales, como su nombre lo indica, son sucesiones de variables aleatorias las cuales tienen una convergencia casi segura a la constante cero. En esta sección exhibiremos los diversos problemas que surgen al simular este tipo de variables y como poder solucionarlos.

Una familia de eventos raros es un conjunto de eventos $\{A(x)\}$ donde $x \in (0, \infty)$ o $x \in \mathbb{N}$ tales que $\mathbb{P}[X \in A(x)] \rightarrow 0$ cuando $x \rightarrow \infty$. Se querrá entonces poder estimar el comportamiento de $\mathbb{P}[X \in A(x)]$ para valores grandes de x .

Se puede utilizar el método de Monte Carlo crudo para estimar a ζ por medio de (2.5) como

$$\hat{\zeta} = \frac{1}{R} \sum_{i=1}^R \mathbb{1}_{\{X_i \in A(x)\}}, \quad (2.14)$$

la ecuación anterior involucra generar una muestra aleatoria X_1, \dots, X_R tal que $X_i \sim \mathbb{P}$ para $i = 1, \dots, R$.

Un problema que surge al estimar eventos raros es el siguiente, supongamos que tenemos $\zeta = \mathbb{P}[A(x)]$ tal que $\zeta \rightarrow 0$ cuando $x \rightarrow \infty$, entonces podríamos estimar dicha probabilidad por medio de simulaciones de $Z = \mathbb{1}_{\{X \in A(x)\}}$. En nuestro estudio tenemos el ejemplo de $A = \{\tau(x) < \infty\}$. Con el método M.C. crudo obtenemos que la varianza es

$$\sigma_{\hat{\zeta}}^2 = \zeta(1 - \zeta)/R$$

la cual tiende a cero cuando $\zeta \rightarrow 0$, pero demostraremos que su velocidad de convergencia no es lo suficientemente rápida. El problema reside en el comportamiento del error relativo, ya que éste dictamina el intervalo de confianza, el cual para valores pequeños queremos que sea pequeño, porque en caso contrario caeríamos en el error de tener un estimador que podría variar demasiado. Por lo anterior es importante observar la velocidad de convergencia de nuestro estimador ya que según sea su velocidad será el intervalo de confianza que podamos dar a nuestras estimaciones; para ejemplificar, supongamos que $\zeta = 10^{-6}$, para poder estimar ζ adecuadamente con un error relativo, por ejemplo, de $\varepsilon = 0.01$ se necesitará, aproximadamente, una muestra de un tamaño tan grande como

$$R \approx \frac{1}{\varepsilon^2 \zeta} = 10^{10}.$$

Con lo anterior se muestra que estimar probabilidades pequeñas con el método de Monte Carlo crudo no es computacionalmente eficiente.

Otro ejemplo es que si estamos trabajando con valores del orden de 10^{-5} un intervalo de confianza de longitud 10^{-4} es demasiado grande con respecto a nuestros valores estimados (es tan grande que permite probabilidades negativas), aún cuando una longitud de 10^{-4} parecería buena.

Para la estimación de eventos raros el error relativo de nuestro estimador de Monte Carlo está dado por

$$\varepsilon = \frac{\sigma_{\hat{\zeta}}}{\mathbb{E}[\hat{\zeta}]} = \frac{\sqrt{\zeta(1 - \zeta)}}{R\zeta} \sim \frac{1}{\sqrt{R\zeta}} \rightarrow \infty. \quad (2.15)$$

Existen entonces dos formas de reducir el error relativo de nuestras estimaciones; la primera es aumentando el número de estimaciones, lo cual en la mayor parte de los casos es ineficiente ya que por estar trabajando con eventos raros se necesitarán al rededor de

$$R \approx \frac{1}{\varepsilon^2 \zeta}.$$

Esto implica que se necesita una cantidad enorme de simulaciones para lograr conseguir que el error relativo pequeño.

La otra forma de reducir el error relativo es por medio de métodos de reducción de varianza los cuales expondremos en la Sección 2.3. Durante el transcurso de los distintos métodos de reducción de varianza daremos ejemplos de cómo solucionar los problemas que surgen al simular eventos raros.

2.2.1. Simulación de probabilidades de ruina vista como un evento raro

Para nuestro estudio queremos encontrar un método eficiente de estimar la probabilidad de ruina $\psi(x)$ bajo el modelo Cramér-Lundberg, donde contamos con un capital inicial x , el cual suponemos grande. En este caso supondremos también que la distribución de los reclamos, F_x , es de cola pesada.

Los algoritmos que expondremos se basan en que podemos representar a la probabilidad de ruina $\psi(x) = \zeta = \mathbb{E}[Z]$ para alguna v.a. Z que puede ser simulada, luego entonces simularemos R réplicas independientes Z_1, Z_2, \dots, Z_R de Z y estimaremos a $\psi(x)$ por medio de $\hat{\zeta} = (Z_1 + \dots + Z_R)/R$ y utilizaremos el estimador de la varianza para lograr determinar los intervalos de confianza de nuestro estimador.

Ya hemos mencionado (ver la Sección 2.1.2) que una medida de la eficiencia de nuestro estimador es el error relativo (2.8)⁹. Por lo tanto surgen los siguientes problemas

1. El problema de la ruina para horizonte infinito implica simular un proceso durante un tiempo infinito, lo cual es computacionalmente imposible; es decir, si uno tuviera un algoritmo que simulara el proceso de ruina cada simulación podría ser detenida sólo si existiera la ruina pues en caso contrario nada garantizaría que al detener la simulación en un tiempo finito la ruina no ocurriera en un tiempo posterior.
2. Como se espera tener un capital inicial grande entonces la probabilidad de ruina será pequeña, por lo que se puede ver que es un evento raro. Suponiendo que logramos superar el primer problema y que podemos generar $Z = 1_{\{\tau(x) < \infty\}}$ entonces podríamos simular la probabilidad de ruina por medio del método de Monte Carlo crudo. Este procedimiento

⁹Supondremos que el tiempo que toma una corrida de cualquier algoritmo es la misma ya que siendo más estrictos sería importante notar esta diferencia para poder ponderar la eficiencia de cada algoritmo

sería generar variables aleatorias Bernoulli con parámetro $\psi(x)$, entonces se tendrá que el error relativo es

$$\varepsilon = \frac{\sigma_Z}{\psi(x)} = \frac{\sqrt{\psi(x)(1-\psi(x))/R}}{\psi(x)} \approx \frac{1}{\sqrt{R\psi(x)}} \rightarrow \infty \quad \text{cuando } x \rightarrow \infty \quad (2.16)$$

lo cual muestra que el método de Monte Carlo crudo no es computacionalmente factible para el problema de la ruina.

El primer problema se puede resolver por medio de la formula de Pollaczek-Khinchine ya que permite simular la probabilidad de la ruina por medio de una convolución geométrica. En el Capítulo 3 se expondrán a través de los resultados obtenidos, cuáles son las mejoras computacionales al aplicar los métodos de reducción de varianza para el caso en que se tengan reclamos con distribución de colas pesadas, tomando en cuenta los problemas que implica este tipo de simulación. En la Sección 2.4 se expondrán diversos algoritmos para poder afrontar el segundo de los problemas.

2.3. Métodos de reducción de varianza

Los métodos de reducción de varianza se basan en la información a priori que se tenga sobre el problema. Esta información se implementa en la simulación para hacerla más eficaz, ya sea por ejemplo conocer el comportamiento que presenta el fenómeno cuando se condiciona sobre algún otro evento o cuales eventos tienen mayor peso para la estimación, entre otros. Además se pueden implementar varios métodos de reducción de varianza a la vez según sean las condiciones del problema. Estos métodos harán que el algoritmo a iterar sea mayor, pero (si se implementan de manera adecuada) reducirán en gran medida el tiempo computacional que se requiere para la precisión que el problema exija.

Entre los métodos que utilizaremos se encuentra muestreo por importancia el cual mejora la simulación al cambiar la función de distribución $F(z)$ (con la que se generan los datos) por alguna otra. El método de covariables crea variables con algún tipo de correlación entre sí para mejorar la eficiencia (ya sea para variables antitéticas o variables de control). El beneficio del muestreo estratificado viene de partir el dominio de F que denotamos como Ω en subregiones $\Omega_1, \dots, \Omega_m$ llamadas estratos e ir tomando muestras de cada región y con esto eliminar en la mayor medida posible la variabilidad del estimador. La mejora que se obtiene en el caso del muestreo por importancia se debe a que se asigna mayor peso a eventos que sean más importantes que

otros al momento de simular. En Monte Carlo condicionado se utiliza que las esperanzas condicionadas con respecto a una nueva variable aleatoria W son conocidas, luego entonces se simula a W para obtener un estimador de ζ .

Cabe mencionar que no siempre un método de reducción de varianza es eficiente, ya que al aplicarlo aunque se disminuya el número de simulaciones no forzosamente se está reduciendo el tiempo de cómputo dado que al implementar uno de estos métodos el algoritmo de la simulación puede crecer, de manera que cada simulación tarde mayor tiempo en ser realizada y por ende no exista una mejora computacional.

2.3.1. Uso de covariables

Una forma de reducir la varianza es utilizando covariables, esto es, por medio del uso de variables extras para mejorar la simulación. Cualquier variable que tenga alguna relación con la variable original (la variable de estudio) tiene la posibilidad de ayudar a reducir la varianza de nuestro estimador.

Este tipo de variables pueden ser útiles si son fáciles de generar y se tiene conocimiento de algunas de sus propiedades (o estas propiedades son fáciles de aproximar). Los métodos de variables antitéticas y variables de control utilizan este tipo de variables para obtener estimadores con menor varianza.

VARIABLES ANTITÉTICAS

Para este método se generan M réplicas de vectores aleatorios, los cuales están formados por parejas de v.a. no necesariamente independientes. De hecho para mayor eficiencia del método se buscará que estas parejas sean tan negativamente correlacionadas como se pueda. En otras palabras se toma $R = 2M$ y se generan M parejas de vectores aleatorios i.i.d.

$$(Z_1, Z_2), (Z_3, Z_4), \dots, (Z_{M-1}, Z_M)$$

tales que la distribución marginal de Z_r es la misma (como en el método de Monte Carlo crudo) para toda r (sea par o impar) pero se tendrá que Z_{2j-1} y Z_{2j} pueden ser dependientes. El estimador es entonces

$$\hat{\zeta}_{Anti} = \frac{Z_1 + \dots + Z_R}{R}, \quad (2.17)$$

con varianza

$$\begin{aligned}
 \sigma_{Anti}^2 &= Var\left(\frac{Z_1 + \dots + Z_R}{R}\right) \\
 &= \frac{M}{M^2} Var\left(\frac{Z_1 + Z_2}{2}\right) \\
 &= \frac{1}{M} Var\left(\frac{Z_1 + Z_2}{2}\right) \\
 &= \frac{1}{4M} (\sigma_z^2 + \sigma_z^2 + 2\sigma_z^2\rho) \\
 &= \frac{1}{2R} \sigma_z^2 (1 + \rho)
 \end{aligned}$$

donde $\sigma_z^2 = Var(Z_i)$ y $\rho = Cor(Z_{2i-1}, Z_{2i})$ para $i = 1, 2, \dots, M$. Entonces ρ debería ser negativa para obtener reducción en la varianza y preferiblemente tan cercana a -1 como se pueda para que este método sea más eficiente.

Ejemplo 2.3.1. *Supongamos que queremos lograr que X_1 y X_2 estén correlacionados en forma negativa, suponiendo que X_1 es una función de n v.a.i.i.d. con distribución $\mathcal{U}(0, 1)$; es decir, supongamos que*

$$X_1 = h(U_1, \dots, U_n)$$

donde U_1, \dots, U_m son v.a.i.i.d. con $U_i \sim \mathcal{U}(0, 1)$ para $i = 1, \dots, m$. Ahora, si $U \sim \mathcal{U}(0, 1)$, entonces la distribución de $1 - U$ también es $\mathcal{U}(0, 1)$. Por lo tanto, la variable aleatoria

$$X_2 = h(1 - U_1, \dots, 1 - U_n)$$

tiene la misma distribución que X_1 . Además como $1 - U_i$ está correlacionada de manera negativa con U_i para $i = 1, \dots, n$, supondríamos que X_2 está correlacionada de manera negativa con X_1 . Este resultado se puede demostrar en el caso particular de que h sea una función monótona (ver Apéndice F). Por lo tanto si h es monótona, después de generar U_1, \dots, U_n para calcular a X_1 en vez de generar un nuevo conjunto independiente de m números aleatorios, podemos utilizar $1 - U_1, \dots, 1 - U_n$ para calcular a X_2 . En este caso se obtuvo un estimador con menor varianza a la vez que se simuló la mitad de las variables aleatorias.

Variabes de control

Supongamos que Z es una v.a. y queremos estimar por medio del método de Monte Carlo a $\mathbb{E}[Z]$. Supongamos también que W es una v.a. con esperanza conocida $\omega = \mathbb{E}[W]$. La idea de este método es tomar una v.a. W

que tenga una gran correlación (ya sea positiva o negativa) con Z , generar $(Z_1, W_2), \dots, (Z_R, W_R)$ en vez de generar solamente Z_1, \dots, Z_R y con esto hacer un estimador $\widehat{\zeta}_{VC}$ con menor varianza que $\widehat{\zeta}_{MCC}$ obtenido por el método de Monte Carlo crudo para $\mathbb{E}[Z]$. El siguiente teorema muestra cómo obtener este estimador.

Teorema 2.3.1. Sean Z y W v.a. con varianzas σ_Z, σ_W respectivamente y con covarianza σ_{ZW} . Supongamos que se conoce a $\omega = \mathbb{E}[W]$ y se desconoce $\zeta = \mathbb{E}[Z]$, entonces

1. La función

$$\pi(\alpha) = Z - \alpha(W - \omega) \quad (2.18)$$

es un estimador insesgado de ζ para toda $\alpha \in (-\infty, \infty)$, con

$$\text{Var}\pi(\alpha) = \sigma_Z^2 - 2\alpha\sigma_{ZW} + \alpha^2\sigma_W^2 \quad (2.19)$$

2. $\alpha^* = \sigma_{ZW}^2/\sigma_W^2$ minimiza (2.19) con

$$\text{Var}\pi(\alpha^*) = \sigma_Z^2(1 - \rho_{ZW}^2) \leq \text{Var}\pi(0) = \sigma_Z^2$$

donde

$$\rho_{ZW} = \text{cor}(Z, W) = \frac{\sigma_{ZW}}{\sqrt{\sigma_Z^2\sigma_W^2}}$$

Demostración. Para la primera parte tenemos que

$$\begin{aligned} \text{Var}\pi(\alpha) &= \text{Var}Z + \text{Var}(\alpha(W - \omega)) + 2\text{Cov}(Z, \alpha(W - \omega)) \\ &= \sigma_Z^2 + \alpha^2\sigma_W^2 - 2\alpha\sigma_{ZW} \end{aligned}$$

Lo cual demuestra el primer punto. Para la segunda demostración se tiene que (2.19) es convexa (porque su segunda derivada es positiva) vista como función de α , y al derivar e igualar a cero se obtiene que esta alcanza su máximo en

$$\alpha^* = \frac{\sigma_{ZW}^2}{\sigma_W^2} \quad (2.20)$$

Al substituir tenemos que

$$\text{Var}\pi(\alpha^*) = \sigma_Z^2 - \frac{\sigma_{ZW}^2}{\sigma_W^2} = \sigma_Z^2(1 - \rho_{ZW}^2). \quad (2.21)$$

Además se tiene que $\text{Var}\pi(0) = \sigma_Z^2$ y al usar el hecho de que $|\rho_{ZW}| \leq 1$ se completa la prueba. \square

Ejemplo 2.3.2. Supongamos que estamos interesados en calcular

$$\zeta = e - 1 = \int_0^1 e^x dx = \mathbb{E}[e^U],$$

donde $U \sim \mathcal{U}(0, 1)$. Y utilizaremos a la variable de control U , se tiene que

$$\begin{aligned} \text{Var}(e^U) &= \int_0^1 e^{2x} dx - (e - 1)^2 \\ &= \frac{e^2 - 1}{2} - (e - 1)^2. \\ &\approx 0.242035607 \end{aligned}$$

$$\begin{aligned} \text{Cov}(e^U, U) &= \mathbb{E}[Ue^U] - \mathbb{E}[U] \mathbb{E}[e^U] \\ &= \int_0^1 xe^x dx - \frac{e - 1}{2} \\ &= 1 - \frac{e - 1}{2} \\ &= \frac{3 - e}{2}. \end{aligned}$$

Utilizando la ecuación (2.21) se obtiene

$$\begin{aligned} \text{Var}\left(e^U + \alpha^* \left(U - \frac{1}{2}\right)\right) &= \text{Var}(e^U) - 12\text{Cov}(e^U, U)^2 \\ &\approx 0.003940222922. \end{aligned}$$

Por lo tanto se obtiene una reducción de la varianza aproximadamente del 98.37%.

Observación 2.3.1. En la formula (2.21) hay que notar que mientras mayor sea el valor de $|\rho_{ZW}|$ se tendrá una mayor reducción en la varianza.

Para el método de las variables de control se utilizará como estimador de $\zeta = \mathbb{E}[Z]$ a

$$\hat{\zeta}_{VC} := \hat{\zeta} - \alpha^*(\hat{w} - \omega) \quad (2.22)$$

El cual por el teorema anterior es un estimador insesgado y utilizando una variable aleatoria W altamente correlacionada con Z se obtendrá una reducción de varianza.

Observación 2.3.2. Las fórmulas (2.20) y (2.21) se pueden fácilmente generalizar para el caso de múltiples variables de control. Sea $W = (W_1, \dots, W_m)^T$

un vector aleatorio columna de tamaño m de variables de control con vector de medias $\mu_W = \mathbb{E}[W] = (\mu_1, \dots, \mu_m)^T$, con $\mu_i = \mathbb{E}[W_i]$ para $i = 1, \dots, m$ entonces la versión vectorial de (2.18) puede ser escrita como

$$\pi(\alpha) = Z - \alpha^T(W - \mu_W) \quad (2.23)$$

donde α es un vector m -dimensional. Se tendrá entonces que el valor de α^* que minimiza $\text{Var}\pi(\alpha)$ es dado por

$$\alpha^* = \Sigma_W^{-1} \sigma_{ZW} \quad (2.24)$$

donde Σ_W denota la matriz de covarianzas (de $m \times m$) de W y σ_{ZW} denota el vector (de tamaño m) cuya componente i es la covarianza entre Z y W_i para $i = 1, \dots, m$. Por lo tanto la correspondiente varianza de $\pi(\alpha^*)$ será

$$\text{Var}\pi(\alpha^*) = (1 - R_{ZW}^2) \text{Var}(Z) \quad (2.25)$$

donde

$$R_{ZW}^2 = (\sigma_{ZW})^T \Sigma_W^{-1} \sigma_{ZW} / \text{Var}(Z)$$

la cual es la raíz cuadrada del coeficiente de correlación de Z y W . De nuevo se tendrá que mientras mayor sea $|R_{ZW}|$ será mejor la reducción de la varianza.

Por lo general, las cantidades $\text{Var}(Z)$, $\text{Var}(W)$ y $\text{Cov}(Z, W)$ no se conocen de antemano y deben de estimarse a partir de una simulación piloto. Si se realizan R ejecuciones de simulación y se obtienen los datos de salida Z_i , W_i para $i = 1, \dots, R$, entonces se puede estimar el valor óptimo de α^* por medio de remplazar σ_{ZW}^2 y σ_W^2 por sus valores empíricos

$$\hat{\alpha} := \frac{s_{ZW}^2}{s_W^2}$$

donde

$$s_Z^2 := s^2 := \frac{1}{R-1} \sum_{r=1}^R (Z_r - \bar{Z})^2, \quad s_W^2 := \frac{1}{R-1} \sum_{r=1}^R (W_r - \bar{W})^2,$$

$$s_{ZW}^2 := \frac{1}{R-1} \sum_{r=1}^R (Z_r - \bar{Z})(W_r - \bar{W})$$

y entonces se puede utilizar el estimador

$$\hat{\zeta}_{VC} := \hat{\zeta} - \hat{\alpha}(\hat{w} - \omega) \quad (2.26)$$

que asintóticamente tiene las mismas propiedades que $\hat{\zeta} - \alpha^*(\hat{w} - \omega)$

Observación 2.3.3. Otra forma de hacer los cálculos es observar la similitud que se tiene con la regresión lineal simple, ya que si consideramos el modelo

$$Z = a + bW + e$$

donde e es una variable aleatoria con media 0 y varianza σ^2 , entonces \hat{a} y \hat{b} , los estimadores por mínimos cuadrados de a y b con base en los datos Z_i, W_i para $i = 1, \dots, R$ son

$$\hat{b} = \frac{\sum_{r=1}^R (Z_r - \bar{Z})(W_r - \bar{W})}{\sum_{r=1}^R (W_r - \bar{W})^2}$$

$$\hat{a} = \bar{Z} - \hat{b}\bar{W}$$

por lo tanto, $\alpha^* = \hat{\alpha}$. Además, como

$$\bar{Z} - \hat{\alpha}(\bar{W} - \omega) = \bar{Z} - \hat{b}(\bar{W} - \omega) = \hat{a} + \hat{b}\omega$$

se tiene que la estimación de la variable de control es la evaluación de la recta de regresión estimada en el valor $W = \omega$. Por lo tanto se puede utilizar algún programa de cómputo que contenga la estimación de los valores de la regresión lineal para facilitar los cálculos.

2.3.2. Muestreo de importancia

Supongamos que queremos estimar $\zeta = \mathbb{E}[H(Z)]$ para una v.a. Z . Entonces puede ser que algunas observaciones de Z tengan mayor importancia que otros para determinar el valor de ζ y quisiéramos que estos valores fueran muestreados con mayor frecuencia.¹⁰ Entonces una manera de estimar ζ más eficientemente sería produciendo estos valores de manera más frecuente.

Supongamos que estamos interesados en estimar

$$\zeta = \mathbb{E}_f[H(Z)] = \int H(z)f(z)dz$$

Sea g otra densidad de probabilidad tal que el soporte de g está contenido en el de Hf . Esto es, si $g(z) = 0$ entonces $H(z)f(z) = 0$. Usando la densidad g podemos representar a ζ como

$$\zeta = \int H(z)\frac{f(z)}{g(z)}g(z)dz = \mathbb{E}_g\left[H(Z)\frac{f(Z)}{g(Z)}\right] \quad (2.27)$$

¹⁰Un ejemplo es tomar ζ como la probabilidad de ocurrencia de un evento raro, entonces dada su poca probabilidad necesitaríamos una inmensa cantidad de simulaciones para que ocurriera en alguna de ellas el evento raro, por tanto quisiéramos que su aparición tuviera mayor probabilidad. Este cambio de probabilidad luego se tendría que ajustar, lo cual es la idea básica del muestreo por importancia.

Entonces si Z_1, \dots, Z_R son una muestra aleatoria con densidad g se tendrá que

$$\widehat{\zeta}_{MI} := \frac{1}{R} \sum_{r=1}^R H(Z_r) \frac{f(Z_r)}{g(Z_r)} \quad (2.28)$$

es un estimador insesgado de ζ . A este se le llama estimador del muestreo por importancia.

La forma en que se está utilizando la densidad g para dar más importancia a ciertos valores es que si Z se distribuye de acuerdo con g por lo general $f(Z)$ será pequeño con respecto de $g(Z)$ ¹¹, así cuando Z se simule de acuerdo con g , el cociente de verosimilitud $f(Z)/g(Z)$ será pequeño con respecto de 1. Pero se tiene que

$$\mathbb{E}_g \left[\frac{f(Z)}{g(Z)} \right] = \int \frac{f(z)}{g(z)} g(z) dz = 1$$

Entonces como $f(Z)/g(Z)$ por lo general es menor que 1 pero su esperanza es igual a 1, se implica que su varianza es grande. Entonces para lograr una reducción de varianza se necesitará elegir a g para que los valores z para los cuales $f(z)/g(z)$ es grande sean los valores para los cuales $H(z)$ sea pequeña de modo $H(z)f(z)/g(z)$ casi siempre sea pequeño.

Ejemplo 2.3.3. Sea N una variable aleatoria tal que $\mathbb{P}[N = n] = \pi(1-\pi)^{n-1}$ para $n \in \mathbb{N}$.¹² Y queremos estimar

$$\zeta = \mathbb{P}[N \leq m] = \sum_{n=1}^m \pi(1-\pi)^{n-1} = 1 - (1-\pi)^m,$$

donde $\pi = \pi(x)$ y $m = m(x)$, son funciones del parámetro x de tal forma que $\zeta = \zeta(x) \rightarrow 0$ cuando $x \rightarrow \infty$. Lo cual implica que $(1-\pi)^m \rightarrow 1$; es decir, que $m\pi \rightarrow 0$. Entonces se sigue que $\zeta \sim m\pi$.

Para poder simular a ζ utilizaremos simulaciones con distribución $N \sim \text{geo}(\tilde{\pi})$. Definimos la variable aleatoria

$$Z := 1_{\{N \leq m\}} \frac{\pi(1-\pi)^{N-1}}{\tilde{\pi}(1-\tilde{\pi})^{N-1}}.$$

¹¹Esto por el hecho de que en una muestra saldrán más frecuentemente los valores con altas probabilidades

¹²Hay que notar que no es una v.a. $\text{geo}(\pi)$ como hemos utilizado normalmente, ya que esta es una v.a. geométrica que cuenta los ensayos totales hasta el primer éxito.

Entonces

$$\begin{aligned}\mathbb{E}[Z^2] &= \frac{\pi^2}{\tilde{\pi}^2} \mathbb{E} \left[\left(\frac{1-\pi}{1-\tilde{\pi}} \right)^{2(N-1)} ; N \leq m \right] = \frac{\pi^2}{\tilde{\pi}^2} \sum_{n=1}^m \frac{(1-\pi)^{2(n-1)}}{(1-\tilde{\pi})^{n-1}} \\ &= \frac{\pi^2 (1-\pi)^{2m} / (1-\tilde{\pi})^m - 1}{\tilde{\pi}^2 (1-\pi)^2 / (1-\tilde{\pi}) - 1}\end{aligned}$$

Como queremos estimar $\mathbb{P}[N \leq m]$ una decisión que parece ser buena es el hacer que la media de nuestra suma sea m ; es decir, hacer $\mathbb{E}[N] = m$ lo que implica que $\tilde{\pi} = 1/m$. Utilizando que $\pi \sim \zeta/m$ y la ecuación anterior, obtenemos

$$\mathbb{E}[Z^2] \sim \frac{\zeta^2/m^2}{1/m} \frac{1/e^{-1} - 1}{1/(1-1/m) - 1} \sim \zeta^2/m \frac{1/e^{-1} - 1}{1/m} = \zeta^2(e-1),$$

por lo tanto el estimador tiene error relativo acotado.

Existen ocasiones en las cuales la siguiente transformación de la función de densidad es útil.

Definición 2.3.1. Sea f una función de densidad con función generadora de momentos m , a la transformación definida por

$$f_t(x) = \frac{e^{tx} f(x)}{m(t)} \quad (2.29)$$

se le llamará el cambio de medida exponencial de f , para $t \in \mathbb{R}$

Ejemplo 2.3.4. Sean X_1, \dots, X_n v.a.i. con funciones de densidad de probabilidad f_i para $i = 1, \dots, n$. Se quiere aproximar la probabilidad de que su suma sea al menos tan grande como a , donde a es un valor mucho mayor que la media de la suma. Se quiere entonces aproximar

$$\zeta = \mathbb{P}[S \geq a]$$

donde $S = \sum_{i=1}^n X_i$ (pensando que $a > \sum_{i=1}^n \mathbb{E}[X_i]$). Se tiene que

$$\zeta = \mathbb{E}_f [1_{\{S \geq a\}}]$$

donde f es la función de densidad conjunta de $\{f_1, \dots, f_n\}$. Supongamos que simulamos $X_{i,r}$ de acuerdo con el cambio de medida exponencial de $f_{i,t}$ para $i = 1, \dots, n$, $r = 1, \dots, R$ para una $t > 0$ fija. Entonces utilizamos las variables aleatorias

$$Z_r = 1_{\{S \geq a\}} \prod_{i=1}^n \frac{f_i(X_{i,r})}{f_{i,t}(X_{i,r})} = 1_{\{S \geq a\}} \prod_{i=1}^n m_i(t) e^{-tX_{i,r}}$$

para aproximar a ζ con $r = 1, \dots, R$. Entonces

$$Z_r = \mathbb{1}_{\{S \geq a\}} m(t) e^{-tS}$$

donde $m(t) = \prod_{i=1}^n m_i(t) e^{-tX_{i,r}}$ la cual es la función generadora de momentos de S . Como $t > 0$ se tendrá que

$$\mathbb{1}_{\{S \geq a\}} e^{-tS} \leq e^{-at}$$

por lo tanto

$$Z_r \leq m(t) e^{-at}$$

lo cual es una cota para el estimador. Para lograr que esta cota se minimice vemos que la función

$$\pi(t) = m(t) e^{-at}.$$

Al derivarla e igualarla a cero tenemos que

$$\pi'(t) = e^{-at} [m'(t) - m(t)] = 0 \quad \Leftrightarrow \quad \frac{d}{dt} \ln m(t) = a,$$

usando la condición inicial de que $m(0) = 1$ se tiene que

$$m(t) = e^{-at^*}$$

donde t^* es la t que minimiza a $m(t) e^{-at}$, lo cual se tendrá que al derivar y evaluar en 0

$$\mathbb{E}_{f_{t^*}} [S] = \mathbb{E}_{f_{t^*}} \left[\sum_{i=1}^n X_i \right] = a$$

donde f_{t^*} quiere decir que el valor esperado se debe considerar bajo la hipótesis de que la distribución de X_i es f_{i,t^*} para $i = 1, \dots, n$. Al hacer esto se ha logrado un estimador el cual cada una de sus simulaciones estará acotada entre 0 y $m(t^*) e^{-at^*}$, lo cual reducirá la varianza del estimador. Además hay que notar que estas nuevas variables justamente tienen media a , lo cual mejora en gran medida las estimaciones que obtenemos.

Observación 2.3.4. El método del muestreo por importancia permite estimar varias cantidades distintas en una única simulación. Por ejemplo, supongamos que

$$\zeta_1 = \mathbb{E} [h(Y)] \quad \zeta_2 = \mathbb{E} [h(W)]$$

donde Y y W son vectores aleatorios con funciones de densidad f y g respectivamente. Si se simula Y , podemos utilizar $h(Y)$ y $h(Y)g(W)/h(W)$ como estimadores de ζ_1 y ζ_2 respectivamente.

Muestreo por importancia condicionado

El muestreo de importancia es útil al estimar una esperanza condicional cuando se está condicionando sobre un evento raro. Supongamos que Z es un vector aleatorio con función de densidad f y que queremos estimar

$$\zeta = \mathbb{E}[H(Z) \mid Z \in \mathcal{A}] \quad (2.30)$$

donde H es una función con valores reales y $\mathbb{P}[Z \in \mathcal{A}]$ es una probabilidad pequeña desconocida. Entonces tenemos que

$$\zeta = \frac{\int_{z \in \mathcal{A}} H(z) f(z) dz}{\mathbb{P}[Z \in \mathcal{A}]} = \frac{\mathbb{E}[H(Z) \mathbb{1}_{\{Z \in \mathcal{A}\}}]}{\mathbb{E}[\mathbb{1}_{\{Z \in \mathcal{A}\}}]}. \quad (2.31)$$

Por lo tanto necesitamos simular

$$\mathbb{E}[N] = \mathbb{E}[H(Z) \mathbb{1}_{\{Z \in \mathcal{A}\}}] \quad \mathbb{E}[D] = \mathbb{E}[\mathbb{1}_{\{Z \in \mathcal{A}\}}] \quad (2.32)$$

Entonces en lugar de simular Z de acuerdo con la densidad f , lo cual haría bastante improbable que Z estuviese en \mathcal{A} , la simulamos de acuerdo con alguna otra densidad g que haga este evento más probable. Si simulamos R v.a. Z_1, \dots, Z_R con densidad g entonces podemos estimar a $\mathbb{E}[N]$ mediante $\frac{1}{R} \sum_{r=1}^R N_r$ y $\mathbb{E}[D]$ mediante $\frac{1}{R} \sum_{r=1}^R D_r$ donde

$$N_r = \frac{H_r(Z_r) f(Z_r) \mathbb{1}_{\{Z_r \in \mathcal{A}\}}}{g(Z_r)}$$

$$D_r = \frac{f(Z_r) \mathbb{1}_{\{Z_r \in \mathcal{A}\}}}{g(Z_r)}.$$

Entonces tenemos el siguiente estimador de ζ

$$\hat{\zeta}_{MIC} = \frac{\sum_{r=1}^R H_r(Z_r) f(Z_r) \mathbb{1}_{\{Z_r \in \mathcal{A}\}} / g(Z_r)}{\sum_{r=1}^R f(Z_r) \mathbb{1}_{\{Z_r \in \mathcal{A}\}} / g(Z_r)} \quad (2.33)$$

Ejemplo 2.3.5. Supongamos que estamos interesados en simular a $S = \sum_{i=1}^4 X_i$ donde $X_i \sim \exp(1/(i+2))$ para $i = 1, 2, 3, 4$, para poder estimar

$$\zeta = \mathbb{E}[S \mid S > 62].$$

En este caso utilizaremos el muestreo por importancia junto con el cambio de medida dado por la transformación exponencial.

Por ejemplo, si tomamos el valor de $t = 0.14$ entonces $\mathbb{E}_t[S] = 62.42999$. Lo anterior nos permite generar k conjuntos de variables aleatorias exponenciales tales que $X_i \sim \exp(1/(i+2) - 0.14)$ para $i = 1, 2, 3, 4$. Si definimos a

S_j como la suma del j -ésimo conjunto, para $j = 1, \dots, k$. De esta manera al utilizar el muestreo por importancia condicionado podemos estimar

$$\alpha = \mathbb{E} [S \mathbb{1}_{\{S > 62\}}]$$

y

$$\beta = \mathbb{E} [\mathbb{1}_{\{S > 62\}}]$$

utilizando

$$\hat{\alpha} = \frac{C}{k} \sum_{j=1}^k S_j \mathbb{1}_{\{S > 62\}} e^{-0.14S_j}$$

y

$$\hat{\beta} = \frac{C}{k} \sum_{j=1}^k \mathbb{1}_{\{S > 62\}} e^{-0.14S_j}$$

donde definimos

$$C = \prod_{i=1}^4 \frac{1}{1 - 0.14(i + 2)}.$$

Por lo tanto tenemos el estimador

$$\hat{\zeta} = \frac{\sum_{j=1}^k S_j \mathbb{1}_{\{S > 62\}} e^{-0.14S_j}}{\sum_{j=1}^k \mathbb{1}_{\{S > 62\}} e^{-0.14S_j}}$$

Método de la minimización de la varianza

Como la elección de la densidad de importancia g es un factor decisivo para la reducción de varianza del estimador $\hat{\zeta}_{MI}$ dado por (2.28), entonces será importante considerar el problema de minimizar la varianza de $\hat{\zeta}_{MI}$ con respecto de g , esto es

$$\min_g \text{Var}_g \left(H(Z) \frac{f(Z)}{g(Z)} \right). \quad (2.34)$$

Como se tiene en general que la varianza es mayor o igual a cero y alcanza el valor de cero si y solo si la variable aleatoria es constante, tendremos que para una constante c

$$H(z) \frac{f(z)}{g(z)} = c \quad \Leftrightarrow \quad g(z) = \frac{H(z)f(z)}{c}. \quad (2.35)$$

Entonces al tener que ser g una densidad¹³, esto determinará el valor de c . Se tendrá que el valor de g^* que minimiza la varianza está dado por

$$g^*(z) = \frac{|H(z)|f(z)}{\int |H(z)|f(z)dz} \quad (2.36)$$

donde alcanza una varianza igual a cero. La densidad g^* es llamada la densidad óptima del muestreo por importancia

Observación 2.3.5. *El problema de tomar la densidad óptima surge con que la integral $\int |H(z)|f(z)dz$ puede no ser calculable o dado el caso de que $H(z) > 0$ ser exactamente el valor ζ .*

Es importante que aunque (2.28) es un estimador insesgado para cualquier función de densidad g que domine a Hf no todas las funciones g son computacionalmente posibles. Una forma de realizar una buena elección para g es haciendo que el estimador (2.28) tenga varianza finita, lo cual es equivalente a

$$\mathbb{E}_g \left[H^2(Z) \frac{f^2(Z)}{g^2(Z)} \right] = \mathbb{E}_f \left[H^2(Z) \frac{f(Z)}{g(Z)} \right] < \infty \quad (2.37)$$

Esto sugiere tomar a g con una cola más ligera que la de f para que así el cociente f/g sea acotado.

Ejemplo 2.3.6. *Para mostrar cómo es el método de la minimización de la varianza utilizaremos el uso de variables antitéticas. Supongamos que queremos estimar una integral de la forma $\zeta = \int_0^1 g(x)dx$, donde g es una función monótona, entonces al utilizar las variables $Z_1 = g(U)$ y $Z_2 = g(U)$, para $U \sim \mathcal{U}(0, 1)$ el Corolario F.0.1.1 nos garantiza que su covarianza es menor o igual a cero.*

En el caso de que g sea un polinomio se tendrán los siguientes casos:

- Si $g(x) = x$, entonces $Z_1 + Z_2 = 1$ lo cual implica que se tiene varianza cero
- Si $g(x) = x^2$, entonces

$$\begin{aligned} \text{Cov}(Z_1, Z_2) &= \frac{\mathbb{E}[U^2(1-U)^2] - (\mathbb{E}[U^2])^2}{\text{Var}(U^2)} \\ &= \frac{1/5 + 1/3 - 1/2 - 1/3^2}{1/5 - 1/3^2} \\ &= -\frac{7}{8} \end{aligned}$$

¹³Por lo tanto g tiene que ser positiva e integrar uno.

- Si $g(x) = x^n$, entonces

$$\begin{aligned} |Cov(Z_1, Z_2)| &\leq \frac{\mathbb{E}[U^n(1-U)^n] - (\mathbb{E}[U^n])^2}{Var(U^n)} \\ &\leq \frac{1/4^n + 1/(n+1)^2}{1/(2n+1) - 1/(n+1)^2} \\ &\sim \frac{2}{n}, \end{aligned}$$

donde se utilizó que $x^n(1-x)^n$ tiene un mínimo en $1/4^n$ y que en el caso de que n sea impar la covarianza será negativa. Entonces se obtiene que la varianza se minimiza mientras mayor sea el valor de n .

2.3.3. Muestreo estratificado

En el muestreo estratificado la región de integración Ω de

$$\zeta = \mathbb{E}_f[H(Z)] = \int_{\Omega} H(z)f(z)dz$$

es partida en m subregiones Ω_i , $i = 1, \dots, m$ que cumplan que $\Omega = \cup_{i=1}^m \Omega_i$ con $\Omega_i \cap \Omega_j = \emptyset$ para $i \neq j$, donde \emptyset es el conjunto vacío. A estas subregiones Ω_i se les llama estratos.

La idea que utiliza el muestreo estratificado es similar a la del muestreo por importancia. Las muestras son obtenidas sobre las subregiones Ω_i y la reducción de la varianza se obtiene al favorecer a los subconjuntos Ω_i que sean relativamente más importantes.

Para esto utilizamos que ζ se puede ver como

$$\zeta = \sum_{i=1}^m \zeta_i \quad (2.38)$$

donde

$$\zeta_i = \int_{\Omega_i} H(z)f(z)dz. \quad (2.39)$$

Entonces podemos reescribir a (2.39) como

$$\begin{aligned} \zeta_i &= \int_{\Omega_i} H(z)f(z)dz \frac{p_i}{p_i} \\ &= p_i \int_{\Omega_i} H(z) \frac{f(z)}{p_i} dz \\ &= p_i \mathbb{E}_{f_i}[H_i(Z)] \end{aligned} \quad (2.40)$$

donde

$$\begin{aligned} p_i &= \int_{\Omega_i} f(z) d(z) \\ f_i(z) &= \begin{cases} f(z)/p_i & \text{si } z \in \Omega_i \\ 0 & \text{e.o.c.} \end{cases} \\ H_i(z) &= \begin{cases} H(z) & \text{si } z \in \Omega_i \\ 0 & \text{e.o.c.} \end{cases} \end{aligned}$$

Entonces podemos representar a $\zeta = \mathbb{E}_f [H(Z)]$ como

$$\zeta = \sum_{i=1}^m p_i \mathbb{E}_{f_i} [H_i(Z)]. \quad (2.41)$$

Asumiendo que podemos obtener a p_i , $i = 1, \dots, m$ de manera analítica, el estimador del muestreo estratificado de ζ basado en una muestra aleatoria $\{Z_{ij}\}$ para $i = 1, \dots, m$ y $j = 1, \dots, R_i$ es

$$\widehat{\zeta}_{ME} = \sum_{i=1}^m \frac{p_i}{R_i} \sum_{j=1}^{R_i} H_i(Z_{ij}) \quad (2.42)$$

donde R_i es el tamaño de la muestra asignado a la región Ω_i tal que

$$\sum_{i=1}^m R_i = R$$

y para $i = 1, \dots, m$, $j = 1, \dots, R_i$ donde Z_{ij} es la j -ésima observación de la subregión Ω_i . Además se tiene que

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^m \frac{p_i}{R_i} \sum_{j=1}^{R_i} H_i(Z_{ij}) \right) &= \sum_{i=1}^m \frac{p_i^2}{R_i^2} \sum_{j=1}^{R_i} \text{Var}(H_i(Z_{ij})) \\ &= \sum_{i=1}^m \frac{p_i^2}{R_i} \text{Var}(H_i(Z_{i1})) \\ &= \sum_{i=1}^m \frac{p_i^2}{R_i} \text{Var}_{f_i}(H(Z)) \end{aligned}$$

entonces podemos reescribir a la varianza de nuestro estimador como

$$\text{Var}(\widehat{\zeta}_{ME}) := \sum_{i=1}^m \frac{p_i^2 \sigma_i^2}{R_i} \quad (2.43)$$

donde

$$\sigma_i^2 := \text{Var}_{f_i}(H(Z)) = \int_{\Omega_i} [H(z) - \zeta_i]^2 f_i(z) dz. \quad (2.44)$$

Observación 2.3.6. *Se tiene que $p_1 + \dots + p_m = 1$. Esto permite interpretar al muestreo estratificado como una ponderación de cada uno de los estratos por medio de las $\{p_i\}$.*

Existen dos puntos importantes en el muestreo estratificado. Por una parte es importante la partición de nuestro espacio Ω y por otro dado que se tiene una vez elegida esta partición de qué forma asignar a cada Ω_i el número de muestras R_i , todo esto para reducir la varianza de nuestro estimador $\widehat{\zeta}_{ME}$. El siguiente teorema resuelve el segundo problema para el caso donde se conocen a $\{p_i\sigma_i\}$ para $i = 1, \dots, m$.

Teorema 2.3.2. *Sea $\{\Omega_1, \dots, \Omega_m\}$ una partición de Ω ; es decir, $\Omega = \cup_{i=1}^m \Omega_i$ y $\Omega_i \cap \Omega_j = \emptyset$ para $i \neq j$. Entonces el problema de minimización*

$$\min_{(R_1, \dots, R_m)} \left\{ \text{Var}(\widehat{\zeta}_{ME}) = \sum_{i=1}^m \frac{p_i^2 \sigma_i^2}{R_i} \right\} \quad (2.45)$$

sujeto a

$$\sum_{i=1}^m R_i = R$$

tiene como solución a

$$R_i^* = R \frac{p_i \sigma_i}{\sum_{k=1}^m p_k \sigma_k} \quad (2.46)$$

Además la varianza del estimador $\widehat{\zeta}_{ME}^*$ que utiliza a R_1^*, \dots, R_m^* está dada por

$$\text{Var}(\widehat{\zeta}_{ME}^*) = \frac{1}{R} \left[\sum_{i=1}^m p_i \sigma_i \right]^2 \quad (2.47)$$

Demostración. Al utilizar el método de los polinomios de Lagrange, utilizando a la variable λ como multiplicador de Lagrange se obtiene que

$$\frac{p_i \sigma_i}{R_i} = \frac{\lambda}{R} \quad \text{para toda } i = 1, \dots, m.$$

Como se tiene que es válido para toda i , se tendrá que para toda j con $j = 1, \dots, m$

$$\frac{p_i \sigma_i}{R_i} = \frac{p_j \sigma_j}{R_j} \quad \Leftrightarrow \quad p_j \sigma_j R_i = p_i \sigma_i R_j$$

Al sumar sobre todas las j se tiene que

$$\left(\sum_{j=1}^m p_j \sigma_j \right) R_i = p_i \sigma_i \left(\sum_{j=1}^m R_j \right)$$

Lo cual implica que

$$R_i = \frac{p_i \sigma_i R}{\sum_{j=1}^m p_j \sigma_j}. \quad (2.48)$$

Sustituyendo (2.48) en (2.43) se obtiene el valor mínimo de la varianza. \square

El Teorema 2.3.2 muestra la varianza mínima del estimador $\widehat{\zeta}_{ME}$ al tomar muestras de tamaño R_i^* las cuales son proporcionales a $p_i \sigma_i$. Una dificultad que surge es que aunque las probabilidades p_i se pueden suponer conocidas las desviaciones estándar $\{\sigma_i\}$ difícilmente se conocen.¹⁴

Se tendrá que el muestreo estratificado puede utilizar cualquier plan de muestreo que particione a Ω en conjuntos disjuntos y que tome cualquier número de muestras por estrato. Lo que uno quisiera es tener alguna partición y un número de muestras por estrato tal que $Var(\widehat{\zeta}_{ME}) \leq Var(\widetilde{\zeta}_{MCC})$ donde $\widetilde{\zeta}_{MCC}$ es el estimador del método de Monte Carlo crudo de una muestra de tamaño R .

El siguiente teorema describe una selección de los R_i tal que garantiza una reducción de la varianza con respecto del estimador de Monte Carlo pero sin la necesidad de conocer a las varianzas $\{\sigma_i^2\}$.

Teorema 2.3.3. *Para una R fija se tendrá que la selección R_1, \dots, R_m dada por*

$$R_i = R p_i \quad (2.49)$$

garantiza que

$$Var(\widehat{\zeta}_{ME}) \leq Var(\widehat{\zeta}_{MCC})$$

Demostración. Tenemos por (2.43) que para $R_i = R p_i$

$$Var(\widehat{\zeta}_{ME}) = \frac{1}{R} \sum_{i=1}^m p_i \sigma_i^2 \quad (2.50)$$

Usando la desigualdad de Cauchy-Schwarz se tiene

$$\begin{aligned} \zeta^2 &= \left(\sum_{i=1}^m \zeta_i \right)^2 = \left(\sum_{i=1}^m \frac{\zeta_i}{\sqrt{p_i}} \sqrt{p_i} \right)^2 \\ &\leq \sum_{i=1}^m \frac{\zeta_i^2}{p_i} \sum_{j=1}^m p_j = \sum_{i=1}^m \frac{\zeta_i^2}{p_i} \end{aligned} \quad (2.51)$$

¹⁴Este problema se puede resanar por medio de una simulación piloto para estimar a las $\{\sigma_i^2\}$ y así no tener este inconveniente.

Pero se tiene de (2.50) y de (2.39) que

$$R \cdot \text{Var}(\widehat{\zeta}_{ME}) = \sum_{i=1}^m p_i \sigma_i^2 = \int_{\Omega} H^2(z) f(z) dz - \sum_{i=1}^m \frac{\zeta_i}{p_i}. \quad (2.52)$$

Entonces combinando (2.52) y (2.51) se tiene que

$$\sum_{i=1}^m p_i \sigma_i^2 \leq \int_{\Omega} H^2(z) f(z) dz - \zeta^2 = R \cdot \text{Var}(\widehat{\zeta}_{MCC}) \quad (2.53)$$

lo cual completa la prueba. \square

El teorema anterior muestra que para toda partición $\Omega_1, \dots, \Omega_r$ de Ω esta inducirá una reducción de varianza cuando se utilice (2.49), suponiendo que se conocen p_1, \dots, p_r .

Supongamos ahora que estamos interesados en estimar la integral de una función sobre el intervalo $[0, 1]$

$$\zeta = \mathbb{E}[h(U)] = \int_0^1 h(x) dx$$

donde $U(0, 1)$. Podemos entonces partir nuestro intervalo en n estratos utilizando a la variable aleatoria S definida como

$$S = i \quad \text{cuando} \quad \frac{i-1}{n} \leq U < \frac{i}{n} \quad \text{para} \quad i = 1, \dots, n.$$

Se sigue que

$$\begin{aligned} \zeta &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[h(U) \mid S = i] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[h(U_i^*)] \end{aligned}$$

donde $U_i^* \sim \mathcal{U}((i-1)/n, i/n)$. Lo anterior para en lugar de tener que simular U_1, \dots, U_n y luego utilizar $\sum_{i=1}^n h(U_i)/n$ como estimador es conveniente el estimador

$$\widehat{\zeta} = \frac{1}{n} \sum_{i=1}^n h\left(\frac{U_i + i - 1}{n}\right). \quad (2.54)$$

el cual tiene una menor varianza.

Ejemplo 2.3.7. Supongamos que queremos estimar π , entonces podemos utilizar la fórmula del círculo para obtener que

$$\frac{\pi}{4} = \int_0^1 \sqrt{1-x^2} dx = \mathbb{E} \left[\sqrt{1-U^2} \right].$$

Se sigue que utilizando la ecuación (2.54) podemos plantear el estimador

$$\hat{\pi}_1 = \frac{4}{n} \sum_{i=1}^n \sqrt{1 - (U_i + i - 1)^2/n^2}.$$

Como ya hemos mencionado se pueden utilizar distintos métodos de reducción de varianza en un mismo algoritmo; por ejemplo, podemos mejorar nuestro estimador utilizando variables antitéticas y hacer el estimador

$$\hat{\pi}_2 = \frac{2}{n} \sum_{i=1}^n \left(\sqrt{1 - (U_i + i - 1)^2/n^2} + \sqrt{1 - (i - U_i)^2/n^2} \right).$$

La siguiente tabla muestra cuáles fueron los valores obtenidos al simular¹⁵ cada uno de los estimadores anteriores, al ir variando el valor de R .

	$R = 5$	$R = 10$	$R = 100$	$R = 1000$	$R = 10000$	$R = 100000$
$\hat{\pi}_1$	3.108618	3.111939	3.140088	3.141641	3.141596	3.141593
$\hat{\pi}_2$	3.141958	3.111702	3.141142	3.141643	3.141590	3.141593

Los valores anteriores tuvieron los siguientes errores absolutos

	$R = 5$	$R = 10$	$R = 100$	$R = 1000$	$R = 10000$	$R = 100000$
$\varepsilon_{\hat{\pi}_1}$	3.297514e-02	2.965392e-02	1.504272e-03	4.826670e-05	3.509558e-06	4.857466e-08
$\varepsilon_{\hat{\pi}_2}$	3.655597e-04	2.989027e-02	4.506198e-04	5.016205e-05	2.989698e-06	6.912356e-09

Las tablas anteriores muestran una tendencia del segundo estimador a ser mejor que el primero. Esta mejora se obtuvo por utilizar más de un método de reducción de varianza.

2.3.4. Monte Carlo Condicionado

Sea

$$\zeta = \int_{\Omega} H(z)f(z)dz = \mathbb{E} [H(Z)]$$

donde Z es una v.a. con densidad f . Supongamos que la v.a. $W \sim g(w)$ es tal que la esperanza condicional $\mathbb{E} [H(z) | W = w]$ puede ser obtenida analíticamente para todos los valores del espacio muestral de W . Entonces se tendrá que

$$\zeta = \mathbb{E} [H(z)] = \mathbb{E} [\mathbb{E} [H(z) | W]]$$

¹⁵El programa utilizado puede ser consultado en el Apéndice G.

Por lo tanto $\mathbb{E}[H(z) | W]$ es un estimador insesgado de ζ . Además como se tiene que

$$\text{Var}H(z) = \mathbb{E}[\text{Var}H(z) | W] + \text{Var}\mathbb{E}[H(z) | W]$$

Entonces se tendrá que

$$\text{Var}\mathbb{E}[H(z) | W] \leq \text{Var}H(z)$$

Lo cual nos da una reducción de varianza. Por lo tanto el algoritmo de Monte Carlo condicional será el siguiente

- Generar una muestra aleatoria W_1, \dots, W_R de $g(w)$
- Calcular $\mathbb{E}[H(z) | W_i]$ para $i = 1, \dots, R$
- Estimar $\zeta = \mathbb{E}[H(z)]$ por medio de

$$\hat{\zeta}_C = \frac{1}{R} \sum_{i=1}^R \mathbb{E}[H(z) | W_i] \quad (2.55)$$

El problema generalmente con este método surge para encontrar a la v.a. W tal que se conozca analíticamente $\mathbb{E}[X | W = w]$ y que además W sea fácil de generar. Un punto importante para encontrar a dicha v.a. W es que $\mathbb{E}[\text{Var}H(z) | W]$ debe de ser grande en comparación a $\text{Var}\mathbb{E}[H(z) | W]$ para que el estimador (2.55) tenga una menor varianza.

Ejemplo 2.3.8. *Supongamos que se quiere estimar*

$$\zeta = \mathbb{P}[S \leq z] = \mathbb{E}[1_{\{S_N \leq z\}}]$$

para

$$S_N = \sum_{i=1}^N Z_i$$

donde N es una v.a. que toma valores en los naturales y Z_i son v.a.i.i.d. con $Z_i \sim F(z)$. Definamos a F^n como la función de distribución de la variable S_n dado que $N = n$. Tenemos que

$$F^n(z) = \mathbb{P}\left[\sum_{i=1}^n Z_i \leq z\right] = \mathbb{E}\left[F\left(z - \sum_{i=2}^n Z_i\right)\right].$$

Entonces

$$\zeta = \mathbb{E}_R[\mathbb{E}[1_{\{S_R \leq z\}} | R]] = \mathbb{E}_R\left[\mathbb{E}\left[F\left(z - \sum_{i=2}^n Z_i\right)\right]\right]$$

Por lo tanto podemos tomar como un estimador condicionado de ζ a

$$\widehat{\zeta}_C = \frac{1}{R} \sum_{i=1}^R F \left(z - \sum_{j=2}^{N_i} Z_{ij} \right)$$

Utilizando que $\sum_{i=2}^1 Z_i = 0$. Se tendrá que simular primero a N_i el cual nos dirá cuantas simulaciones de Z tenemos que hacer, donde i varía desde 1 hasta R . En total se tendrán que hacer a lo más $N_1 + \dots + N_R - R$ simulaciones.

2.4. Algoritmos de simulación para la probabilidad de ruina en el modelo de C-L

En esta sección se hará el desarrollo de algoritmos para calcular la probabilidad de ruina para el modelo Cramér-Lundberg en el caso de que los reclamos tengan distribución subexponencial. Estos algoritmos serán aplicados en el siguiente capítulo para así poder comparar su efectividad.

2.4.1. Algoritmo I: Simulación a través de la fórmula de Pollaczek-Khinchine

Utilizando la ecuación (1.42) tenemos que

$$\psi(x) = 1 - \frac{\rho}{1 + \rho} \sum_{n=0}^{\infty} (1 + \rho)^{-n} F_I^{*n}(x) \quad \text{para } x > 0 \quad (2.56)$$

Lo cual nos dice que $1 - \psi(x)$ es una convolución geométrica, por lo tanto

$$\psi(x) = \mathbb{P}[S_M > x]$$

donde $S_M = X_1 + X_2 + \dots + X_M$, M tiene distribución geométrica con parámetro $\rho/(1+\rho)$ a la vez que es independiente de X_1, X_2, \dots y que X_1, X_2, \dots son v.a.i.i.d. con distribución F_I (ver Sección 1.3). Lo anterior nos permite aplicar el método de M.C. crudo para $\psi(x) = z = \mathbb{E}[Z]$ donde $Z = 1_{\{S_M > x\}}$. Por lo tanto tenemos el siguiente algoritmo

Definición 2.4.1 (Algoritmo I).

1. Generar $M_i \sim \text{geo}(\rho/(1 + \rho))$
2. Generar $X_1^i, \dots, X_{M_i}^i$ con distribución $F_I(x)$ y hacer $S_{M_i} = X_1^i + \dots + X_{M_i}^i$

3. Si $S_{M_i} > x$ hacer $Z_i = 1$ en otro caso $Z_i = 0$
4. Repetir pasos 1 a 3 R veces (para $i = 1, \dots, R$)
5. Estimar $\mathbb{E}[Z]$ utilizando $\hat{\zeta} = \frac{1}{R} \sum_{i=1}^R Z_i$

Este tipo de algoritmo ya hemos probado que tiene el problema de que no es logarítmicamente eficiente (ver Proposición 2.1.1). Por lo tanto presentaremos algoritmos que si cumplan con (2.13).

2.4.2. Algoritmo II: Monte Carlo condicionado

El siguiente algoritmo utiliza las ideas expuestas en la Sección 2.3.4, en especial la del Ejemplo 2.3.8. Además utilizaremos el hecho de que al quitar variables a simular reducimos la varianza que estas aportaban. Se demostrará también que para el caso de reclamos con distribución subexponencial (ver Sección 1.5.1) este algoritmo no es logarítmicamente eficiente.

El argumento base de este algoritmo es que la probabilidad de ruina cumple con

$$\begin{aligned} \psi(x) &= \mathbb{P}[X_1 + \dots + X_M > x] \\ &= \mathbb{E}[\mathbb{P}[X_1 + \dots + X_M > x \mid X_2, \dots, X_M]] \\ &= \mathbb{E}[\bar{F}_I(x - X_2 - \dots - X_M)] \end{aligned}$$

Por lo tanto es suficiente con generar X_1, \dots, X_{M-1} y hacer $Z = \bar{F}_I(W)$ donde $W = x - X_1 - \dots - X_{M-1}$. Este algoritmo se basa en analizar los $M - 1$ primeros reclamos, ya que conociendo el monto de los $M - 1$ primeros reclamos se tendrá que la probabilidad de ruina (si no ha ocurrido hasta este momento) es simplemente la probabilidad de que el nuevo reclamo sobrepase el capital restante despues de pagar los $M - 1$ primeros reclamos. Por lo tanto tenemos el siguiente algoritmo

Definición 2.4.2 (Algoritmo II).

1. Generar $M_i \sim \text{geo}(\rho/(1 + \rho))$
2. Generar $X_1^i, \dots, X_{M_i-1}^i$ con distribución $F_I(x)$ y hacer $W_i = x - X_1^i - \dots - X_{M_i-1}^i$
3. Sea $Z_i = 1$ si $W_i < 0$ y $Z_i = \bar{F}_I(W_i)$ en otro caso
4. Repetir pasos 1 a 3 R veces (para $i = 1, \dots, R$)

5. Estimar $\mathbb{E}[Z]$ utilizando $\hat{\zeta} = \frac{1}{R} \sum_{i=1}^R Z_i$

Este algoritmo da como resultado a $\hat{\zeta}$ el cual es un estimador insesgado de $\psi(x)$. Aun cuando al aplicar M.C. condicionado siempre se tiene una reducción de varianza y por ende el Algoritmo II tiene menor varianza que el Algoritmo I, se tendrá que para el caso de distribuciones subexponenciales este nuevo algoritmo no será asintóticamente mejor. Utilizaremos el siguiente resultado, el cual es un caso particular de un resultado más general que se puede consultar en Embrechts and Veraverbeke (1982)[15].

Proposición 2.4.1. *Supongamos que la distribución de los reclamos F es subexponencial. Entonces para el Algoritmo II se tiene que*

$$\lim_{x \rightarrow \infty} \frac{\ln \sigma_Z}{\ln \zeta} = \frac{1}{2} \quad (2.57)$$

Demostración.

$$\begin{aligned} \mathbb{E}[Z^2] &= \mathbb{E}\left[\overline{F}_I^2(x - X_1 - \dots - X_{M-1})\right] \\ &\geq \mathbb{E}\left[\overline{F}_I^2(x); M \leq 1\right] + \mathbb{E}\left[\overline{F}_I^2(x - X_1); M \geq 2\right] \\ &= \left(1 - \left(\frac{1}{1+\rho}\right)^2\right) \overline{F}_I^2(x) + \mathbb{E}\left[\overline{F}_I^2(x - X_1); M \geq 2\right] \\ &\geq \left(1 - \left(\frac{1}{1+\rho}\right)^2\right) \overline{F}_I^2(x) + \mathbb{E}\left[\overline{F}_I^2(x - X_1); X_1 > x, M \geq 2\right] \\ &= \left(1 - \left(\frac{1}{1+\rho}\right)^2\right) \overline{F}_I^2(x) + \left(\frac{1}{1+\rho}\right)^2 \overline{F}_I(x). \end{aligned}$$

Donde la última igualdad utiliza el hecho de que si $X_1 > x$ entonces $\overline{F}_I(x - X_1) = 1$. De lo anterior tenemos

$$\mathbb{E}[Z^2] \geq \left(1 - \left(\frac{1}{1+\rho}\right)^2\right) \overline{F}_I^2(x) + \left(\frac{1}{1+\rho}\right)^2 \overline{F}_I(x). \quad (2.58)$$

Por otro lado se tiene por la Proposición 1.5.1 que $\overline{F}_I(x) \sim \rho\psi(x)$, de donde

$$\mathbb{E}[Z]^2 = \psi(x)^2 \sim \frac{1}{\rho^2} \overline{F}_I^2(x) \quad (2.59)$$

Entonces al utilizar (2.58) y (2.59) tenemos que σ_Z^2 es a lo menos del mismo orden que $\bar{F}_I(x)$ para valores grandes de x .¹⁶ Como $\bar{F}_I(x) \sim \rho\psi(x)$ se obtiene entonces que $\ln \sigma_Z$ no puede converger a $-\infty$ más rápido que $\ln \psi(x)/2$, lo cual quiere decir que

$$\ln \sigma_Z \geq \frac{\ln \psi(x)}{2} \quad \Rightarrow \quad \frac{\ln \sigma_Z}{\ln \psi(x)} \leq \frac{1}{2}$$

por lo tanto $1/2$ es una cota superior para el límite del lado izquierdo de (2.57).

Para obtener una cota inferior utilizamos la Proposición 2.1.1 y que la varianza de M.C. crudo es mayor (y por tanto también su desviación estándar) a la de M.C. condicionado, lo cual nos dice que

$$\ln \sigma_{Z_{M.C.C}} \geq \ln \sigma_Z \quad \Rightarrow \quad \frac{\ln \sigma_{Z_{M.C.C}}}{\ln \zeta} \leq \frac{\ln \sigma_Z}{\ln \zeta}.$$

Por lo tanto como el lado izquierdo de la ecuación (2.57) cuando se utiliza M.C. crudo es igual a $1/2$.

$$\lim_{x \rightarrow \infty} \frac{\ln \sigma_Z}{\ln \zeta} \geq \frac{1}{2}.$$

Entonces se tendrá que $1/2$ también será la cota inferior, con lo cual se completa la prueba. \square

La proposición anterior muestra que este algoritmo también carece de ser logarítmicamente eficiente. Este problema es causado porque al estar trabajando con distribuciones subexponenciales la probabilidad de que una exceda un valor alto es (asintóticamente) igual a la de que la suma lo exceda. Por esta razón quisiéramos un algoritmo el cual descartara al reclamo más alto para que así este algoritmo arrojará un mejor estimador.

2.4.3. Algoritmo III: Utilización de las estadísticas de orden

Este algoritmo se basa en que cuando se trabaja con distribuciones subexponenciales el hecho de que la suma de las variables exceda cierto valor es a causa de que una variable tomó un valor excesivamente grande y no porqué cada uno de los sumandos contribuyera significativamente. Los siguientes lemas definirán al Algoritmo III.

¹⁶Esto porque $\bar{F}_I(x) \rightarrow 0$ cuando $x \rightarrow \infty$, lo cual quiere decir que para valores grandes de x tenemos que $\bar{F}_I^2(x) + \bar{F}_I(x) \sim \bar{F}_I(x)$

Lema 2.4.1. Sean X_1, \dots, X_m v.a.i.i.d. no negativas con función de distribución F y densidad f , denotando por $X_{(1)}, \dots, X_{(m)}$ a sus estadísticas de orden. Entonces

$$\mathbb{P} [X_{(m)} > x \mid X_{(1)}, \dots, X_{(m-1)}] = \frac{\bar{F}(X_{(m-1)} \vee x)}{\bar{F}(X_{(m-1)})}. \quad (2.60)$$

Demostración. Dado que X_1, \dots, X_m son i.i.d. y absolutamente continuas (que tienen densidad), se tiene el resultado de que $\{X_{(n)}\}$ forma una cadena de Markov (este resultado puede ser consultado en Arnold(1992)[1]), por lo tanto

$$\mathbb{P} [X_{(m)} > x \mid X_{(1)}, \dots, X_{(m-1)}] = \mathbb{P} [X_{(m)} > x \mid X_{(m-1)}] \quad (2.61)$$

con

$$\mathbb{P} [X_{(m)} > x \mid X_{(m-1)} = y] = \begin{cases} 1 & \text{si } x < y \\ \int_x^\infty f_{X_{(m)}|X_{(m-1)}}(u \mid y) du & \text{si } x \geq y \end{cases}$$

donde

$$\int_x^\infty f_{X_{(m)}|X_{(m-1)}}(u \mid y) du = \frac{\bar{F}(x)}{\bar{F}(y)}.$$

Entonces

$$\mathbb{P} [X_{(m)} > x \mid X_{(1)}, \dots, X_{(m-1)}] = \frac{\bar{F}(X_{(m-1)} \vee x)}{\bar{F}(X_{(m-1)})}.$$

□

Lema 2.4.2. Sean $S_m = X_1 + \dots + X_m$ y $S_{(k)} = X_{(1)} + \dots + X_{(k)}$ para $k = 1, \dots, m$. Con la notación del lema anterior se tiene que

$$\mathbb{P} [S_m > x] = \mathbb{E} \left[\frac{\bar{F}((x - S_{(m-1)}) \vee X_{(m-1)})}{\bar{F}(X_{(m-1)})} \right] \quad (2.62)$$

Demostración. Al condicionar con respecto a $X_{(1)}, \dots, X_{(m-1)}$ se tiene que

$$\begin{aligned} \mathbb{P} [S_m > x] &= \mathbb{E} [\mathbb{P} [S_m > x \mid X_{(1)}, \dots, X_{(m-1)}]] \\ &= \mathbb{E} [\mathbb{P} [X_{(m)} + S_{(m-1)} > x \mid X_{(1)}, \dots, X_{(m-1)}]] \\ &= \mathbb{E} [\mathbb{P} [X_{(m)} > x - S_{(m-1)} \mid X_{(1)}, \dots, X_{(m-1)}]], \end{aligned}$$

esto último es válido ya que $S_{(m-1)}$ es medible con respecto a $X_{(1)}, \dots, X_{(m-1)}$. Aplicando el Lema 2.4.1 se completa la demostración. □

El lema anterior nos plantea una ecuación con respecto a una m fija, pero como

$$\begin{aligned} \mathbb{P}[S_M > x \mid M = m] &= \mathbb{P}[S_m > x \mid M = m] \\ &= \mathbb{P}[X_1 + \dots + X_m > x \mid M = m] \\ &= \mathbb{P}[X_1 + \dots + X_m > x] \\ &= \mathbb{P}[S_m > x] \\ &= \mathbb{E} \left[\frac{\bar{F}((x - S_{(m-1)}) \vee X_{(m-1)})}{\bar{F}(X_{(m-1)})} \right], \end{aligned}$$

lo cual nos lleva a que

$$\mathbb{P}[S_M > x] = \mathbb{E} \left[\frac{\bar{F}((x - S_{(M-1)}) \vee X_{(M-1)})}{\bar{F}(X_{(M-1)})} \right]. \quad (2.63)$$

Por lo tanto podemos planter el siguiente algoritmo el cual da un estimador insesgado de la probabilidad de ruina.

Definición 2.4.3 (Algoritmo III).

1. Generar $M_i \sim \text{geo}(\rho/(1 + \rho))$
2. Generar $X_1^i, \dots, X_{M_i}^i$ con distribución $F_I(x)$ y hacer $W_i = x - X_{(1)}^i - \dots - X_{(M_i-1)}^i$, $K_i = X_{(M_i-1)}^i$
3. Sea $Z_i = \frac{\bar{F}_I(W_i \vee K_i)}{\bar{F}_I(K_i)}$
4. Repetir pasos 1 a 3 R veces (para $i = 1, \dots, R$)
5. Estimar $\mathbb{E}[Z]$ utilizando $\hat{\zeta} = \frac{1}{R} \sum_{i=1}^R Z_i$

El Teorema 2.4.1 es muy importante ya que muestra porqué para el caso de distribuciones de variación regular (ver Apéndice E) el Algoritmo III es logarítmicamente eficiente.

Teorema 2.4.1 (Eficiencia logarítmica del Algoritmo III). *Supongamos que $\bar{F}_I(x) = \frac{L(x)}{x^\alpha}$ donde $\alpha > 1$ y L es de variación lenta. Entonces el Algoritmo III satisface*

$$\liminf_{x \rightarrow \infty} \frac{\ln \sigma_Z}{\ln \psi(x)} \geq 1. \quad (2.64)$$

Para poder demostrarlo necesitaremos auxiliarnos de los siguientes lemas.

Lema 2.4.3. *Para el Algoritmo III se tiene que*

$$\sigma_Z^2 \leq \mathbb{E} \left[M^2 \left(\frac{1}{2} \bar{F}_I^2 \left(\frac{x}{2} \right) + \bar{F}_I^2 \left(\frac{x}{M} \right) \left| \ln \bar{F}_I \left(\frac{x}{2} \right) \right| \right) \right] \quad (2.65)$$

Demostración. Suponiendo que X tiene densidad f_I se tiene que la densidad de X_{M-1} dado $M = m$ es

$$f_{X_{(M-1)}|M=m}(x) = m(m-1) \bar{F}_I^{m-2}(x) \bar{F}_I(x) f_I(x) \quad (2.66)$$

Entonces

$$\begin{aligned} \mathbb{E} [Z^2 | M] &= \mathbb{E} \left[\left(\frac{\bar{F}_I((x - S_{(M-1)}) \vee X_{(M-1)})}{\bar{F}_I(X_{(M-1)})} \right)^2 \mid M \right] \\ &= \mathbb{E} \left[\left(\frac{\bar{F}_I(x - S_{(M-1)})}{\bar{F}_I(X_{(M-1)})} \right)^2 ; X_{(M-1)} \leq \frac{x}{M} \mid M \right] \\ &\quad + \mathbb{E} \left[\left(\frac{\bar{F}_I((x - S_{(M-1)}) \vee X_{(M-1)})}{\bar{F}_I(X_{(M-1)})} \right)^2 ; \frac{x}{M} < X_{(M-1)} \leq \frac{x}{2} \mid M \right] \\ &\quad + \mathbb{E} \left[1 ; X_{(M-1)} > \frac{x}{2} \mid M \right] \end{aligned} \quad (2.67)$$

El primer sumando de (2.67) se puede acotar notando que si $X_{(M-1)} \leq \frac{x}{M}$ entonces $x - S_{(M-1)} \geq \frac{x}{M}$ lo cual implica que $\bar{F}_I(x - S_{(M-1)}) \leq \bar{F}_I(\frac{x}{M})$, por lo tanto

$$\begin{aligned} \mathbb{E} \left[\left(\frac{\bar{F}_I(x - S_{(M-1)})}{\bar{F}_I(X_{(M-1)})} \right)^2 ; X_{(M-1)} \leq \frac{x}{M} \mid M \right] &\leq \bar{F}_I^2 \left(\frac{x}{M} \right) \int_0^{\frac{x}{M}} \frac{f_{X_{(M-1)}|M}(y)}{\bar{F}_I^2(y)} dy \\ &= \bar{F}_I^2 \left(\frac{x}{M} \right) \int_0^{\frac{x}{M}} \frac{M(M-1) \bar{F}_I^{M-2}(y) f_I(y)}{\bar{F}_I(y)} dy \\ &\leq M(M-1) \bar{F}_I^2 \left(\frac{x}{M} \right) \int_0^{\frac{x}{M}} \frac{f_I(y)}{\bar{F}_I(y)} dy \\ &= -M(M-1) \bar{F}_I^2 \left(\frac{x}{M} \right) \ln \left(\bar{F}_I \left(\frac{x}{M} \right) \right) \end{aligned}$$

Para acotar el segundo sumando de (2.67) utilizamos el hecho de que cuando $\frac{x}{M} < X_{(M-1)} \leq \frac{x}{2}$ entonces $((x - S_{(M-1)}) \vee X_{(M-1)}) \geq X_{(M-1)} > \frac{x}{M}$ y por tanto $\bar{F}_I((x - S_{(M-1)}) \vee X_{(M-1)}) \leq \bar{F}_I(\frac{x}{M})$. Entonces

$$\begin{aligned} & \mathbb{E} \left[\left(\frac{\bar{F}_I((x - S_{(M-1)}) \vee X_{(M-1)})}{\bar{F}_I(X_{(M-1)})} \right)^2 ; \frac{x}{M} < X_{(M-1)} \leq \frac{x}{2} \mid M \right] \\ & \leq \bar{F}_I\left(\frac{x}{M}\right) \int_{\frac{x}{M}}^{\frac{x}{2}} \frac{f_{X_{(M-1)}|M}(y)}{\bar{F}_I^2(y)} dy \\ & = \bar{F}_I\left(\frac{x}{M}\right) \int_{\frac{x}{M}}^{\frac{x}{2}} \frac{M(M-1)F_I^{M-2}(y)f_I(y)}{\bar{F}_I(y)} dy \\ & \leq M(M-1)\bar{F}_I\left(\frac{x}{M}\right) \int_{\frac{x}{M}}^{\frac{x}{2}} \frac{f_I(y)}{\bar{F}_I(y)} dy \\ & = -M(M-1)\bar{F}_I\left(\frac{x}{M}\right) \left[\ln(\bar{F}_I\left(\frac{x}{2}\right)) - \ln(\bar{F}_I\left(\frac{x}{M}\right)) \right] \end{aligned}$$

Por último para acotar el tercer sumando de (2.67) tenemos que

$$\begin{aligned} \mathbb{E} \left[1; X_{(M-1)} > \frac{x}{2} \mid M \right] &= \int_{\frac{x}{2}}^{\infty} f_{X_{(M-1)}|M}(y) dy \\ &= M(M-1) \int_{\frac{x}{2}}^{\infty} F_I^{M-2}(y)\bar{F}_I(y)f_I(y) dy \\ &\leq M(M-1) \int_{\frac{x}{2}}^{\infty} \bar{F}_I(y)f_I(y) dy \\ &= M(M-1) \frac{1}{2} \bar{F}_I^2\left(\frac{x}{2}\right). \end{aligned}$$

Juntando las tres desigualdades se obtiene

$$\begin{aligned} \mathbb{E} [Z^2 \mid M] &\leq M(M-1) \left(\frac{1}{2} \bar{F}_I^2\left(\frac{x}{2}\right) - \bar{F}_I^2\left(\frac{x}{M}\right) \ln\left(\bar{F}_I\left(\frac{x}{2}\right)\right) \right) \\ &\leq M^2 \left(\frac{1}{2} \bar{F}_I^2\left(\frac{x}{2}\right) + \bar{F}_I^2\left(\frac{x}{M}\right) \ln\left|\bar{F}_I\left(\frac{x}{2}\right)\right| \right) \end{aligned} \quad (2.68)$$

Por lo tanto, utilizando la esperanza condicional, podemos concluir que

$$\begin{aligned} \sigma_Z^2 &= \mathbb{E} [\mathbb{E} [Z^2 \mid M]] - \mathbb{E} [Z]^2 \\ &\leq \mathbb{E} \left[M^2 \left(\frac{1}{2} \bar{F}_I^2\left(\frac{x}{2}\right) + \bar{F}_I^2\left(\frac{x}{M}\right) \ln\left|\bar{F}_I\left(\frac{x}{2}\right)\right| \right) \right]. \end{aligned}$$

□

Lema 2.4.4. Si $\bar{F}_I(x) = \frac{L(x)}{x^\alpha}$, donde L es una función de variación lenta, entonces para cualquier $\varepsilon > 0$ existen constantes $C_-(\varepsilon)$ y $C_+(\varepsilon)$ tales que para toda $x > 0$ y para toda $d > 0$

$$C_-(\varepsilon)d^{\alpha-\varepsilon}x^{-\alpha-\varepsilon} \leq \bar{F}_I\left(\frac{x}{d}\right) \leq C_+(\varepsilon)d^{\alpha-\varepsilon}x^{-\alpha+\varepsilon}. \quad (2.69)$$

Demostración. Se tiene por $\bar{F}_I(x)x^{\alpha-\varepsilon} = x^{-\varepsilon}L(x)$ que $\lim_{x \rightarrow 0} x^{-\varepsilon}L(x) = 0$ y que L es una función continua. Como L es de variación lenta entonces $x^{-\varepsilon}L(x) \in RV_{-\varepsilon}$ y al utilizar el Corolario E.0.3.3 tenemos que $\lim_{x \rightarrow \infty} x^{-\varepsilon}L(x) = 0$. Por lo tanto existe una constante $C_+(\varepsilon)$ tal que $L(x) \leq C_+(\varepsilon)x^\varepsilon$ para toda $x > 0$ y entonces

$$\bar{F}_I\left(\frac{x}{d}\right) = \frac{L(x/d)}{(x/d)^\alpha} \leq \frac{C_+(\varepsilon)(x/d)^\varepsilon}{(x/d)^\alpha}.$$

La prueba para la cota inferior es similar, utilizando que existe una constante $C_-(\varepsilon)$ tal que $L(x) \geq C_-(\varepsilon)x^\varepsilon$. \square

Lema 2.4.5. Si $\bar{F}_I(x) = \frac{L(x)}{x^\alpha}$, donde L es de variación lenta entonces para toda $\varepsilon > 0$ existen constantes $D_1(\varepsilon)$ y $D_2(\varepsilon)$ tales que

$$\mathbb{E}[Z^2] \leq (D_1(\varepsilon) + D_2(\varepsilon)|\ln x|)x^{2\varepsilon-2\alpha} \quad (2.70)$$

Demostración. Del Lema 2.4.3, observando la ecuación (2.68) tenemos

$$\mathbb{E}[Z^2] \leq \mathbb{E}\left[M^2\left(\frac{1}{2}\bar{F}_I^2\left(\frac{x}{2}\right) + \bar{F}_I^2\left(\frac{x}{M}\right)\ln\left|\bar{F}_I\left(\frac{x}{2}\right)\right|\right)\right]$$

Y utilizando el Lema 2.4.4

$$\begin{aligned} &\leq \mathbb{E}[M^2] \frac{1}{2}C_+^2(\varepsilon)2^{2\alpha-2\varepsilon}x^{-2\alpha+2\varepsilon} \\ &\quad + \mathbb{E}[M^{2\alpha-2\varepsilon+2}] C_+^2(\varepsilon) |\ln(C_-(\varepsilon)2^{\alpha-\varepsilon}x^{-\alpha-\varepsilon})| x^{-2\alpha+2\varepsilon} \\ &= (D_1(\varepsilon) + D_2(\varepsilon)|\ln x|)x^{2\varepsilon-2\alpha} \end{aligned}$$

donde

$$\begin{aligned} D_1(\varepsilon) &= \mathbb{E}[M^2] \frac{1}{2}C_+^2(\varepsilon)2^{2\alpha-2\varepsilon} + \mathbb{E}[M^{2\alpha-2\varepsilon+2}] C_+^2(\varepsilon) |\ln(C_-(\varepsilon)2^{\alpha-\varepsilon}x^{-\alpha-\varepsilon})| \\ D_2(\varepsilon) &= \mathbb{E}[M^{2\alpha-2\varepsilon+2}] C_+^2(\varepsilon)(\alpha + \varepsilon) \end{aligned}$$

\square

Gracias a los lemas anteriores ahora podemos demostrar el Teorema 2.4.1, el cual demuestra la eficiencia logarítmica para el Algoritmo III.

Demostración del Teorema 2.4.1.

$$\begin{aligned}\ln \sigma_Z &\leq \ln \sqrt{(D_1(\varepsilon) + D_2(\varepsilon)|\ln x|)x^{2\varepsilon-2\alpha}} \\ &= \frac{1}{2} \ln(D_1(\varepsilon) + D_2(\varepsilon)|\ln x|) + (\varepsilon - \alpha) \ln x\end{aligned}$$

Por lo tanto

$$\lim_{x \rightarrow \infty} \frac{\ln \sigma_Z}{\ln \psi(x)} \geq \lim_{x \rightarrow \infty} \frac{\frac{1}{2} \ln(D_1(\varepsilon) + D_2(\varepsilon)|\ln x|) + (\varepsilon - \alpha) \ln x}{\ln \psi(x)}$$

Usando la Proposición 1.5.1 obtenemos

$$\begin{aligned}&= \lim_{x \rightarrow \infty} \frac{\frac{1}{2} \ln(D_1(\varepsilon) + D_2(\varepsilon)|\ln x|) + (\varepsilon - \alpha) \ln x}{\ln(\bar{F}_I(x)/\rho)} \cdot \frac{\ln(\bar{F}_I(x)/\rho)}{\ln \psi(x)} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{1}{2} \ln(D_1(\varepsilon) + D_2(\varepsilon)|\ln x|) + (\varepsilon - \alpha) \ln x}{-\ln \rho + \ln L(x) - \alpha \ln x} \cdot \lim_{x \rightarrow \infty} \frac{\ln(\bar{F}_I(x)/\rho)}{\ln \psi(x)} \\ &= \frac{\varepsilon - \alpha}{-\alpha} \cdot 1 \\ &= 1 - \frac{\varepsilon}{\alpha}\end{aligned}$$

Como es válido para toda $\varepsilon > 0$ podemos hacer $\varepsilon \rightarrow 0$, con lo cual queda demostrado el teorema. \square

En conclusión hemos logrado plantear tres algoritmos para calcular la probabilidad de ruina de los cuales uno tiene eficiencia logarítmica. En el siguiente capítulo exhibiremos los resultados que obtuvimos al simularlos.

Capítulo 3

Implementación de algoritmos

En este capítulo exhibiremos los distintos programas para simular las probabilidades de ruina utilizando el programa **R** como plataforma, el cual es un software libre que tiene una orientación principal hacia temas estadísticos. Antes necesitaremos demostrar algunos resultados básicos para el estudio de la simulación.

3.1. Simulación a través de variables aleatorias uniformes

Es bien sabido que las computadoras no pueden generar verdaderas variables aleatorias ya que lo que las computadoras ejecutan son programas pre-determinados de forma concreta, por lo tanto lo que en realidad generan son variables pseudo-aleatorias. Lograr generar estos algoritmos de variables pseudo aleatorias no es sencillo si uno piensa en que para cada variable aleatoria se necesitaría un algoritmo. El siguiente lema demuestra cómo con variables aleatoria $\mathcal{U}(0, 1)$ y conociendo la inversa de la función de distribución (a evaluar) se puede obtener las variables aleatorias deseadas.

Lema 3.1.1. *Sea X una v.a. con función de distribución F y suponga que F^{-1} es su inversa, entonces*

$$F^{-1}(U) \sim F$$

donde U es una v.a. $\mathcal{U}(0, 1)$.

Demostración.

$$\mathbb{P} [F^{-1}(U) < x] = \mathbb{P} [U < F(x)] = F(x)$$

□

Este lema nos permitirá en el caso de conocer la función inversa de las funciones de distribución obtener variables aleatorias a partir de simular una serie de variables uniformes y luego evaluarlas en las funciones inversas. Aunque en la mayor parte de los casos no se tenga a dichas inversas podremos aproximarlas de manera numérica (ver el Apéndice G). El siguiente ejemplo es un caso en el cual se puede obtener la función inversa de manera analítica.

Ejemplo 3.1.1. *En el Ejemplo 1.3.2 obtuvimos que si X es una v.a. cuya función de distribución F_I es la cola integrada de una función Pareto(a, b) entonces*

$$F_I(x) = \frac{b-1}{ab} x 1_{\{x \leq a\}} + \left[1 - \frac{1}{b} \left(\frac{a}{x} \right)^{b-1} \right] 1_{\{x > a\}} \quad \text{para } x \geq 0.$$

Entonces sea $u = F_I(x)$ y tenemos dos casos

- Si $0 \leq x \leq a$ entonces $0 \leq u \leq \frac{b-1}{b}$ y

$$u = \frac{b-1}{ab} x \quad \Leftrightarrow \quad x = \frac{ab}{b-1} u \quad \Rightarrow \quad F_I^{-1}(u) = \frac{ab}{b-1} u.$$

- Si $x > a$ entonces $\frac{b-1}{b} < u \leq 1$ y

$$u = 1 - \frac{1}{b} \left(\frac{a}{x} \right)^{b-1} \quad \Leftrightarrow \quad x = \frac{a}{\sqrt[b-1]{b(1-u)}} \quad \Rightarrow \quad F_I^{-1}(u) = \frac{a}{\sqrt[b-1]{b(1-u)}}.$$

Por lo tanto

$$F_I^{-1}(u) = \begin{cases} \frac{ab}{b-1} u & \text{si } 0 \leq u \leq \frac{b-1}{b} \\ \frac{a}{\sqrt[b-1]{b(1-u)}} & \frac{b-1}{b} < u \leq 1 \end{cases}$$

Antes de continuar nos es apropiado demostrar la siguiente proposición la cual nos permite la simulación de la probabilidad de ruina sin tener que preocuparnos por estar simulando variables aleatorias con varianza infinita.

Proposición 3.1.1. *Sea X una v.a. que tiene un soporte acotado; es decir, existe un valor $a \in \mathbb{R}$ tal que $\mathbb{P}[|X| \leq a/2] = 1$, entonces $\text{Var}(X) < \infty$.*

Demostración.

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 dF(x) \\ &= \int_{-a/2}^{a/2} (x - \mathbb{E}[X])^2 dF(x) \\ &\leq \int_{-a/2}^{a/2} a^2 dF(x) \\ &= a^2, \end{aligned}$$

donde utilizamos el hecho de que $-a/2 \leq \mathbb{E}[X] \leq a/2$. \square

La razón por la cual es pertinente la proposición anterior es por que al calcular las estimadores de la probabilidad de ruina, en el fondo lo que estamos haciendo es simular variables aleatorias las cuales tienen soporte en $[0, 1]$, por lo tanto no estamos simulando v.a. con varianza infinita. Más aún, esto nos permite que aunque cuando estamos simulando v.a. con varianza infinita (por ejemplo el caso de Pareto(1,2)) los estimadores que brindan los algoritmos continúen teniendo varianza finita, y con esto poder reportar una mejora al aplicar los métodos de reducción de varianza.

3.2. Resultados numéricos

En esta sección mostraremos los resultados obtenidos al simular la probabilidad de ruina con distintas distribuciones y la comparación que tenemos con las cotas y aproximaciones. También mostraremos los distintos problemas prácticos que surgen así como las formas de solucionarlos o de aminorarlos.

Compararemos tanto a los algoritmos entre sí como con las distintas aproximaciones que tenemos. La siguiente sección abordará el tema de cuándo usar qué aproximación así como los problemas prácticos que surgen al utilizarlas.

3.2.1. Comparación de cotas y comportamiento asintótico

Antes de continuar nos es pertinente hacer algunos comentarios sobre los resultados que exhibiremos. Dado que ya tenemos los algoritmos para estimar la probabilidad de ruina ahora necesitamos un punto de comparación, ya que sin él no sabríamos si los resultados obtenidos al simular son eficientes. Para estos puntos de partida utilizaremos tanto las cotas de Panjer como el comportamiento asintótico.

Las cotas de Panjer nos brindan un intervalo (más adelante exhibiremos que el intervalo que otorgan es muy preciso) pero el problema que tienen es que para valores grandes del capital inicial, el tiempo de ejecución es demasiado grande, lo cual las hace inoperables. En cambio, el comportamiento asintótico tiene el problema opuesto, ya que para valores pequeños del capital inicial su aproximación es burda (porque su convergencia es demasiado lenta) pero tiene la ventaja de que su cálculo sólo requiere de una operación.

En la figura 3.1 se puede observar que el comportamiento asintótico se va aproximando a la probabilidad de ruina por abajo. Esto último es el resultado

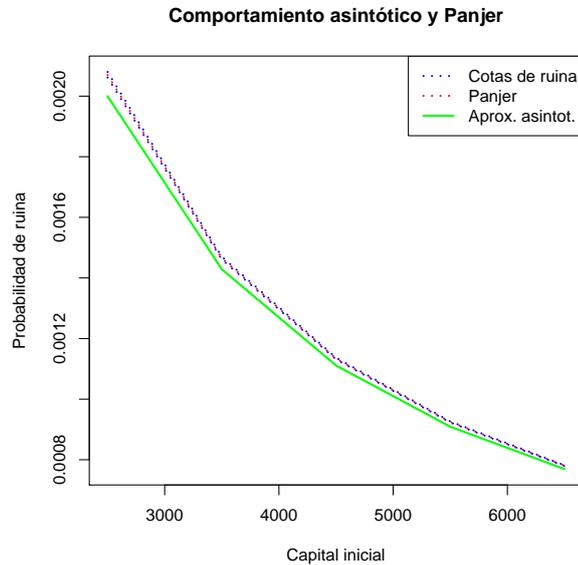


Figura 3.1: Comportamiento asintótico Pareto(1,2)

de que al estar trabajando con distribuciones subexponenciales uno puede pensar que el hecho de que la suma exceda a un valor x es a causa de un solo sumando, lo cual es justamente lo que hace la aproximación al sólo calcular la cola de la distribución de un sumando en el punto x (ver la proposición 1.5.1). Además muestra lo mucho que tarda en lograr aproximarse a las cotas. La gráfica anterior fue basada en la siguiente tabla de resultados la cual muestra la comparación entre el comportamiento asintótico y la aproximación de Panjer, esta tabla cuenta con las siguientes columnas:

x: Representa al capital inicial.

C. Asintótico: Es el valor que tomó la aproximación asintótica para el valor x , dada por $\psi(x) \sim \rho^{-1} \bar{F}_I(x)$.

Panjer: Estimación de Panjer.

E. absoluto: Error absoluto entre la aproximación asintótica y la estimación de Panjer.

E. relativo: Error relativo entre la aproximación asintótica y la estimación de Panjer.

Cota inf.: Cota inferior de Panjer.

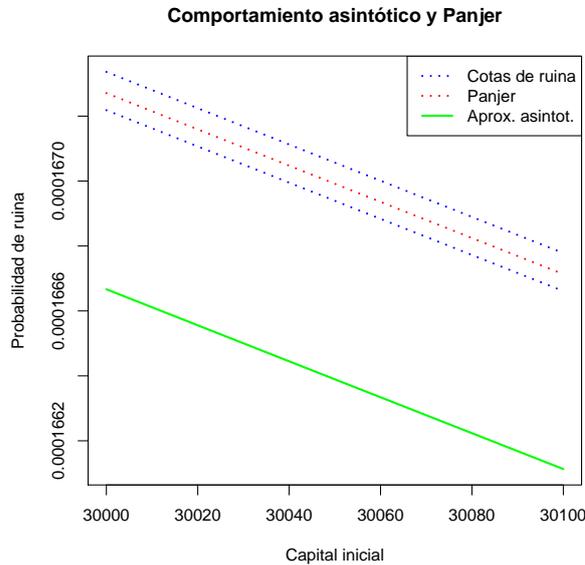


Figura 3.2: Comportamiento asintótico Pareto(1,2)

Cota sup.: Cota superior de Panjer.

<i>Parámetros</i>	
<i>Distribución</i>	ρ
Pareto (1,2)	0.10

<i>Tabla de resultados</i>						
x	<i>C. Asintótico</i>	<i>Panjer</i>	<i>E. absoluto</i>	<i>E. relativo</i>	<i>Cota inf.</i>	<i>Cota sup.</i>
3000	1.666666e-3	1715770e-3	4.91041e-5	0.02861930	0.001710142	0.001722844
5500	0.909090e-3	0924504e-3	1.54135e-5	0.01667218	0.000922883	0.000926530
8000	0.625000e-3	0632538e-3	7.53849e-6	0.01191783	0.000631782	0.000633481
10500	0.476190e-3	0480676e-3	4.48643e-6	0.00933357	0.000480241	0.000481220
13000	0.384615e-3	0387599e-3	2.98447e-6	0.00769988	0.000387316	0.000387952

Esto muestra que aún cuando el error relativo de la aproximación asintótica va decreciendo con respecto a las cotas de Panjer, no lo hace de manera rápida ya que se puede ver tanto en las figuras 3.1 y 3.2 que nunca logra caer dentro de las cotas para los valores calculados.

La figura 3.2 nos muestra que aún para valores mayores no se logra caer dentro del intervalo que dan las cotas. De hecho no fue posible mostrar para que valores ya caía el comportamiento asintótico dentro de las cotas ya que el tiempo que tarda las cotas en ser calculadas es muy grande. Lo que si nos

muestra la tabla de resultados es que su error relativo sigue decayendo (aún cuando su decaimiento no es lo suficientemente veloz). La gráfica 3.2 utiliza los siguientes datos

<i>Parámetros</i>	
<i>Distribución</i>	ρ
Pareto (1,2)	0.10

<i>Tabla de resultados</i>						
x	<i>C. Asintótico</i>	<i>Panjer</i>	<i>E. absoluto</i>	<i>E. relativo</i>	<i>Cota inf.</i>	<i>Cota sup.</i>
30000	1.666667e-4	1.672708e-4	6.041336e-7	3.611710e-3	1.672183e-4	1.673362e-4
30050	1.663894e-4	1.669916e-4	6.022132e-7	3.606249e-3	1.669392e-4	1.670567e-4
30100	1.661130e-4	1.667133e-4	6.003021e-7	3.600806e-3	1.666611e-4	1.667782e-4

Lo anterior nos permitirá comparar los resultados de nuestros algoritmos ya que en caso de que el valor inicial sea relativamente pequeño utilizaremos las cotas de ruina y para valores mayores utilizaremos el comportamiento asintótico.

3.3. Simulación de algoritmos

En las siguientes secciones se mostrarán los resultados obtenidos por los distintos programas. Dado que tenemos muchas variables que podemos cambiar (por ejemplo: el capital inicial, la distribución de los sumandos, los parámetros de la distribución, el número de simulaciones, entre otras) utilizaremos solo tres distribuciones, para cada una de ellas se mantendrán los mismos parámetros, tanto de los sumandos como el parámetro de la geométrica y mostraremos los cambios que tienen las simulaciones con respecto al capital inicial y el número de simulaciones.

Se exhibirán los resultados obtenidos para las distribuciones Pareto(1,2), Log-normal(-1.62,1.80) y MPE(3) (para consultar la definición de esta última distribución consultar E.1). Estas distribuciones son representativas por dos razones; la primera es porque aún cuando las tres pertenecen a las distribuciones de variación regular su comportamiento es muy distinto entre sí y hacen buena representación de comportamientos probables, la segunda razón es porque cada una exhibe distintos problemas para su simulación esto en consecuencia de la existencia o no de expresiones analíticas para las funciones que se les tiene que calcular.

Para cada distribución se exhibirá el comportamiento que se obtuvo con respecto a tres criterios.

- Utilizando el programa `sim_3geom_con_Panjer` el cual para cada capital inicial simula tres sumas geométricas, una para cada algoritmo y exhibe las cotas de Panjer. Éste es el programa más lento de ejecutar, ya que para cada punto que simula realiza la ejecución completa de los algoritmos, además de calcular tanto las cotas como la aproximación de Panjer. La razón por la cual este programa opera así es para mostrar el comportamiento de cada uno de los algoritmos por separado, ya que a diferencia de los programas cada algoritmo tiene su propia suma geométrica. Lo anterior restringe mucho el número de simulaciones que puede realizar.
- Utilizando el programa `sim_1geom_con_Panjer` el cual durante toda su ejecución solo simula una suma geométrica y luego utiliza el criterio de cada algoritmo para obtener las probabilidades y exhibe las cotas de Panjer. Este programa tiene el tiempo de ejecución medio ya que al solo simular una suma geométrica lo hace más veloz que el anterior pero el hecho de calcular tanto las cotas como la aproximación de Panjer lo hace más lento que el programa siguiente. La razón por la cual fue diseñado así es para mostrar el comportamiento de los algoritmos entre sí con respecto a la misma simulación y a la vez comparar estos resultados con Panjer.
- Utilizando el programa `sim_1geom_con_asintota` el cual durante su ejecución solo simula una suma geométrica y luego utiliza el criterio de cada algoritmo para obtener las probabilidades y exhibe el comportamiento asintótico. Este es el programa más rápido al sólo tener que calcular una suma geométrica y no tener que calcular las cotas y aproximaciones de Panjer. La razón por la cual fue diseñado así es para mostrar el comportamiento de los algoritmos al estar estimando probabilidades pequeñas, ya que permite el tener grandes montos para el capital inicial. La desventaja que tiene es que la aproximación asintótica no es muy buena dado el tipo de convergencia que tiene.

Estos programas fueron diseñados de esa manera para poder exhibir distintos aspectos de nuestras estimaciones y aproximaciones. Por ejemplo, el programa `sim_3geom_con_Panjer` es el más lento de ejecución ya que simula repetidas veces sumas geométricas, además de que calcula las cotas de Panjer¹, en cambio el programa `sim_1geom_con_asintota` es sumamente veloz (en

¹En este caso al estar simulando una suma geométrica por cada algoritmo luego hacer esto por cada capital inicial se tiene que se está haciendo una cantidad enorme de simulaciones y con esto poder exhibir las ventajas de cada algoritmo sin tener cierta tendencia por consecuencia de tener solo una simulación de la distribución geométrica.

comparación) ya que no incurre en ninguno de estos cálculos, mientras que el programa `sim_1geom_con_Panjer` cae en un término medio.

Para mostrar los resultados obtenidos se han creado tanto gráficas como tablas que muestran estos resultados. Cada gráfica representa la ejecución de un solo programa y para cada punto de cada gráfica se ha creado una tabla que exhibe distintas características de este punto. Las tablas cuentan con las siguientes columnas y valores

x: Representa al capital inicial. Con respecto de la gráfica representa los valores que se obtuvieron para este punto.

Alg.: Esta columna es para identificar a qué algoritmo pertenecen los resultados, siendo I para el Algoritmo I, II para el Algoritmo II y III para el Algoritmo III.

Aprox: Esta columna representa los valores obtenidos por los distintos algoritmos.

E. absoluto: Esta columna representa el error absoluto que tiene nuestro estimador con respecto a la aproximación de Panjer, para cada uno de los algoritmos.

E. relativo: Esta columna representa el error relativo que tiene nuestro estimador con respecto a la aproximación de Panjer, para cada uno de los algoritmos.

\widehat{s} : Esta columna representa a la varianza muestral que tiene nuestro estimador con respecto a la aproximación de Panjer, para cada uno de los algoritmos. El cual es un estimador insesgado de la varianza de nuestro estimador.

$\log \widehat{s} / \log \psi_{\mathbf{P}}$: Esta columna representa a el cociente que existe entre el logaritmo de la varianza muestral y el logaritmo de la aproximación de Panjer, para cada uno de los algoritmos. La cual se puede tomar como una estimación de cómo se comporta su eficiencia logarítmica.

Intervalo de confianza; Extremos de intervalo: Muestra los extremos del intervalo de confianza dado por $\widehat{\psi}(x) \pm 1.96 \sqrt{\widehat{s}/R}$, donde $\widehat{\psi}(x)$ representa a la aproximación obtenida y R es el número de simulaciones. Se toma el valor de 1.96 por ser el cuantíl 0.95 de la distribución normal estándar.

Intervalo de confianza; Longitud: Muestra la longitud del intervalo de confianza.

Aproximación con Panjer; up: Es el valor que toma la cota superior de Panjer.

Aproximación con Panjer; ψ_P : Es el valor que toma la aproximación de Panjer.

Aproximación con Panjer; lower: Es el valor que toma la cota inferior de Panjer.

Aproximación con Panjer; Longitud: Es la diferencia entre la cota superior y la cota inferior de Panjer.

Aproximación asintótica: Es el valor que tomó la aproximación asintótica para el valor x , dada por $\psi(x) \sim \rho^{-1} \bar{F}_I(x)$.

Además al inicio de cada grupo de tablas aparecerá una pequeña tabla la cual indica cuales son los parámetros que se utilizaron para la distribución así como el número de simulaciones y el valor de ρ .

3.3.1. Distribución de los sumandos: Pareto(1,2)

La gran ventaja que otorga la distribución Pareto es que la inversa de su cola integrada tiene una representación analítica (ver Ejemplo 3.1.1) lo cual facilita su simulación. Ésta fue la distribución más rápida de simular, ya que aún cuando se utilizaron valores grandes para el número de simulaciones (en algunos casos hasta 500 veces el número de simulaciones con otras distribuciones) su tiempo de ejecución fue corto en comparación de las otras distribuciones.

Resultados de `sim_3geom_con_Panjer` utilizando Pareto(1,2)

Para esta primera ejecución tomamos a $R = 2000$ y consideramos capitales iniciales a partir de $x = 1000$. Se observa que la probabilidad de ruina tiene un ligero decaimiento conforme se aumenta el capital inicial lo cual es una tendencia que esperábamos. También es importante notar que con estos valores la probabilidad de ruina ya se encuentra alrededor del valor 0.0013. En la figura 3.3 se puede ver una muy clara ventaja de la estimación del Algoritmo III en comparación de los otros algoritmos.

En la tabla de resultados se puede ver el excelente comportamiento del Algoritmo III al tener un error relativo sumamente pequeño al igual que su desviación estándar. Otro factor que exhibe la tabla es cómo se comporta

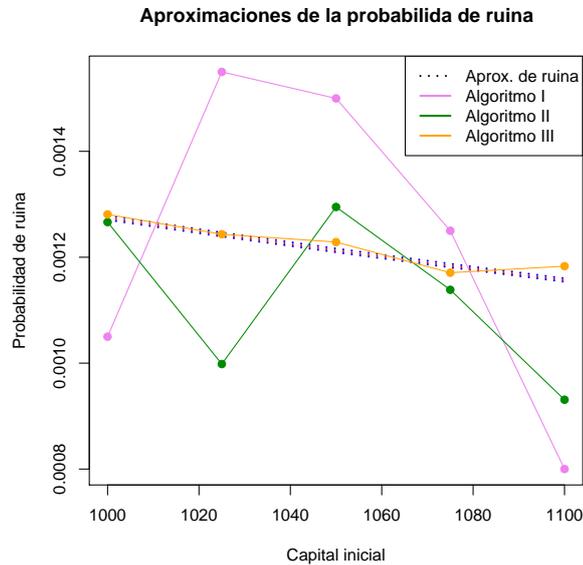


Figura 3.3: `sim_3geom_con_Panjer` utilizando Pareto(1,2)

el límite que determina la eficiencia logarítmica ya que se observa como el Algoritmo III tiene una tendencia a aproximarse al valor de 1.

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
Pareto (1,2)	0.40	20000

Tabla de resultados con $x = 1000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.001050000	2.242137e-04	0.175962370	0.03238750	0.5145932
II	0.001266411	7.802701e-06	0.006123542	0.02998880	0.5261376
III	0.001281032	6.818551e-06	0.005351183	0.00520254	0.7889381
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	1.270636e-03
I	0.0006011322	0.001498868	0.0008977356	ψ_P	1.274214e-03
II	0.0008507874	0.001682034	0.0008312470	lower	1.278579e-03
III	0.0012089287	0.001353136	0.0001442070	Logitud	1.270636e-03

Tabla de resultados con $x = 1025$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.001550000	3.073767e-04	0.2473611117	0.039340499	0.4835941
II	0.000998324	2.442994e-04	0.1965997247	0.025480355	0.5485137
III	0.001243447	8.239481e-07	0.0006630715	0.005177391	0.7867020
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	1.239222e-03
I	0.0010047685	0.002095232	0.0010904630	ψ_P	1.242623e-03
II	0.0006451842	0.001351464	0.0007062794	lower	1.246773e-03
III	0.0011716923	0.001315202	0.0001435100	Logitud	7.550742e-06

Tabla de resultados con $x = 1050$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.001500000	2.874403e-04	0.23705245	0.038701743	0.4835941
II	0.001294931	8.237147e-05	0.06793189	0.030798621	0.5182836
III	0.001228704	1.614404e-05	0.01331401	0.003696325	0.8340131
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	1.209322e-03
I	0.0009636212	0.002036379	0.0010727576	ψ_P	1.212560e-03
II	0.0008680841	0.001721778	0.0008536942	lower	1.216509e-03
III	0.0011774753	0.001279932	0.0001024569	Logitud	7.187049e-06

Tabla de resultados con $x = 1075$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.001250000	6.608506e-05	0.05581910	0.035334118	0.4960590
II	0.001138551	4.536430e-05	0.03831720	0.028265542	0.5291808
III	0.001170807	1.310752e-05	0.01107133	0.003357998	0.8452991
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	1.180829e-03
I	0.0007602941	0.001739706	9.794118e-04	ψ_P	1.183915e-03
II	0.0007468102	0.001530291	7.834809e-04	lower	1.187678e-03
III	0.0011242679	0.001217347	9.307896e-05	Logitud	6.848984e-06

Tabla de resultados con $x = 1100$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000800000	3.565909e-04	0.30831201	0.028273662	0.5273111
II	0.000930940	2.256508e-04	0.19509992	0.024482760	0.5486000
III	0.001183101	2.650982e-05	0.02292066	0.003937349	0.8188435
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	1.153646e-03
I	0.0004081470	0.001191853	0.0007837059	ψ_P	1.156591e-03
II	0.0005916263	0.001270254	0.0006786275	lower	1.160180e-03
III	0.0011285318	0.001237670	0.0001091378	Logitud	6.534197e-06

Tanto en la figura 3.3 como en los resultados se puede observar una marcada mejora respecto del Algoritmo III con respecto a los otros. Por ejemplo, para un capital inicial de 1100 el Algoritmo III muestra una mejora tanto

en términos del error relativo como en la desviación estándar al reducir en una cifra estos valores. Esta mejora también se ve representada en términos de la longitud del intervalo de confianza. Para estos resultados utilizamos 20000 simulaciones ($R = 20000$) lo cual nos fue posible gracias al hecho de que tenemos la inversa de la cola integrada.

Resultados de `sim_1geom_con_Panjer` utilizando Pareto(1,2)

En este caso se pudo hacer $R = 100000$ gracias a que sólo se está simulando una vez la suma geométrica. Además se aumentó el capital inicial y con esto se obtuvo que la probabilidad de ruina fuera aún menor. Los resultados fueron los siguientes

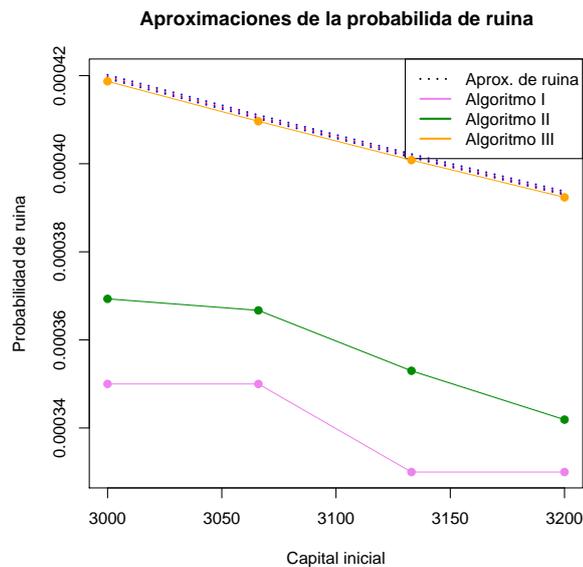


Figura 3.4: `sim_1geom_con_Panjer` utilizando Pareto(1,2)

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
Pareto (1,2)	0.40	100000

Tabla de resultados con $x = 3000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0003500000	6.970452e-05	0.166079975	0.018705106	0.5117000
II	0.0003693229	5.038158e-05	0.120040603	0.015807797	0.5333428
III	0.0004187084	9.961055e-07	0.002373350	0.001399836	0.8450918
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	4.193196e-04
I	0.0002340646	0.0004659354	2.318709e-04	ψ_P	4.197045e-04
II	0.0002713452	0.0004673007	1.959555e-04	lower	4.201713e-04
III	0.0004100321	0.0004273847	1.735255e-05	Logitud	8.517260e-07

Tabla de resultados con $x = 3066$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0003500000	6.061260e-05	0.147615056	0.018705106	0.5102629
II	0.0003667074	4.390516e-05	0.106925983	0.015807825	0.5318446
III	0.0004096345	9.780976e-07	0.002382045	0.001367429	0.8457222
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	4.102442e-04
I	0.0002340646	0.0004659354	2.318709e-04	ψ_P	4.106126e-04
II	0.0002687295	0.0004646854	1.959558e-04	lower	4.110593e-04
III	0.0004011591	0.0004181099	1.695082e-05	Logitud	8.151121e-07

Tabla de resultados con $x = 3133$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0003300000	7.177702e-05	0.178648894	0.018162995	0.5126045
II	0.0003529841	4.879295e-05	0.121442847	0.015403404	0.5336794
III	0.0004008168	9.601724e-07	0.002389814	0.001336037	0.8463394
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	4.014244e-04
I	0.0002174246	0.0004425754	2.251508e-04	ψ_P	4.017770e-04
II	0.0002575128	0.0004484554	1.909426e-04	lower	4.022047e-04
III	0.0003925360	0.0004090977	1.656169e-05	Logitud	7.803055e-07

Tabla de resultados con $x = 3200$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0003300000	6.331354e-05	0.160974723	0.018162995	0.5112126
II	0.0003418957	5.141788e-05	0.130730003	0.015162582	0.5342400
III	0.0003923709	9.426238e-07	0.002396622	0.001306062	0.8469354
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	3.929756e-04
I	0.0002174246	0.0004425754	2.251508e-04	ψ_P	3.933135e-04
II	0.0002479170	0.0004358743	1.879573e-04	lower	3.937233e-04
III	0.0003842759	0.0004004660	1.619012e-05	Logitud	7.476808e-07

Ya que este programa sólo utiliza una suma geométrica se puede observar la eficiencia que tiene el Algoritmo III, ya que las operaciones que realiza hacen que con la misma simulación obtenga resultados mucho más precisos.

Por ejemplo en el caso de un capital inicial de 3200 se obtuvo una disminución del error relativo de dos cifras, así como la reducción en una cifra de la desviación estándar. También se observa que con esta simulación se tiene una probabilidad de ruina sub-valorada, ya que todas las aproximaciones nos dieron inferiores a la cota inferior.

Resultados de `sim_1geom_con_asintota` utilizando Pareto(1,2)

En este caso se tomó tanto el mismo número de simulaciones y capital inicial, esto para hacer una comparación de los resultados obtenidos utilizando Panjer en comparación con los del comportamiento asintótico.

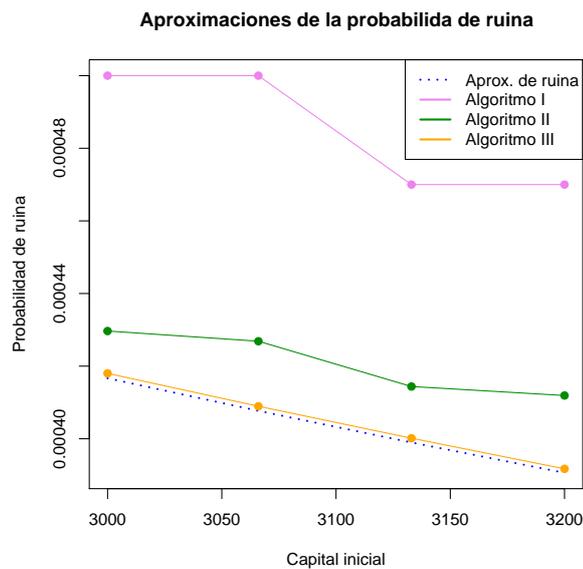


Figura 3.5: `sim_1geom_con_asintota` utilizando Pareto(1,2)

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
Pareto (1,2)	0.40	500000

Tabla de resultados con $x = 3000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0005000000	8.333333e-05	0.200000000	0.022355201	0.4883190
II	0.0004296714	1.300469e-05	0.031211252	0.017602370	0.5190293
III	0.0004180131	1.346438e-06	0.003231452	0.001400351	0.8442558
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.0004166667	
I	0.0003614410	0.0006385590	2.771179e-04		
II	0.0003205707	0.0005387720	2.182012e-04		
III	0.0004093336	0.0004266926	1.735893e-05		

Tabla de resultados con $x = 3066$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0005000000	9.230267e-05	0.226400000	0.022355201	0.4869575
II	0.0004268655	1.916821e-05	0.047015790	0.017602291	0.5175828
III	0.0004089533	1.255939e-06	0.003080566	0.001367393	0.8449534
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.0004076973	
I	0.0003614410	0.0006385590	2.771179e-04		
II	0.0003177654	0.0005359657	2.182003e-04		
III	0.0004004781	0.0004174285	1.695038e-05		

Tabla de resultados con $x = 3133$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0004700000	7.102139e-05	0.178008000	0.021674496	0.4895635
II	0.0004143880	1.540934e-05	0.038621963	0.017316195	0.5182470
III	0.0004001494	1.170736e-06	0.002934332	0.001335498	0.8456352
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.0003989786	
I	0.0003356601	0.0006043399	2.686798e-04		
II	0.0003070611	0.0005217148	2.146538e-04		
III	0.0003918718	0.0004084269	1.655501e-05		

Tabla de resultados con $x = 3200$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0004700000	7.937500e-05	0.203200000	0.021674496	0.4882435
II	0.0004119131	2.128810e-05	0.054497546	0.017316212	0.5168495
III	0.0003917166	1.091649e-06	0.002794621	0.001305068	0.8462922
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.000390625	
I	0.0003356601	0.0006043399	2.686798e-04		
II	0.0003045861	0.0005192401	2.146540e-04		
III	0.0003836278	0.0003998055	1.617779e-05		

En la simulación anterior se tenía una marcada tendencia a subvaluar la probabilidad de ruina, en cambio, en este caso ocurre lo contrario ya que se tiene una tendencia a supervalorar la probabilidad de ruina. Lo anterior sólo

es el resultado de la simulación ya que en el caso anterior se tuvieron sumas altas en comparación a las este caso.

3.3.2. Distribución de los sumandos: Log-normal(-1.62,1.80)

Con la distribución Log-normal se tienen diversos problemas ya que toma mucho tiempo tanto su simulación como el cálculo con los diversos algoritmos, estos problemas surgen por no tener expresiones analíticas ni de su cola integrada y mucho menos de la inversa de esta; lo anterior se vió reflejado en un menor número de simulaciones, ya que por un lado el simularla tardaba mucho tiempo (por tener que ir aproximando su inversa) y por otro al momento de aplicar los algoritmos II y III era necesario el cálculo de su distribución lo cual también consumía mucho tiempo (en comparación con las otras distribuciones). A la par se tiene que el tiempo necesario para calcular las cotas de Panjer aumenta considerablemente ya que al discretizar el proceso es necesario el cálculo de la distribución.

Dados estos problemas es obvio que el implementar métodos de reducción de varianza es algo necesario, ya que sin ellos se necesitaría una cantidad de tiempo inmensa para poder lograr una precisión buena.

Resultados de `sim_3geom_con_Panjer` utilizando Log-normal(-1.62,1.80)

Ya que este programa es el más tardado sólo se simuló $R = 200$ por cada punto. En la figura 3.6 se puede observar que con este número de simulaciones el Algoritmo I resulta ineficiente ya que solo logró reportar durante todas las simulaciones que ejecutó una ruina (haciendo cálculos realizó 1000 simulaciones y sólo una de ellas resultó en ruina, lo que significaría que se necesitarían más o menos 5000 simulaciones para que cada punto reportara una ruina).

En la figura 3.6 se puede ver una marcada mejoría que resulta de utilizar el Algoritmo III, ya que éste logra oscilar alrededor de la cotas, a diferencia de los otros algoritmos los cuales al no lograr ruinas, subvaloran su estimación.

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
Log-normal (-1.62,1.80)	0.10	200

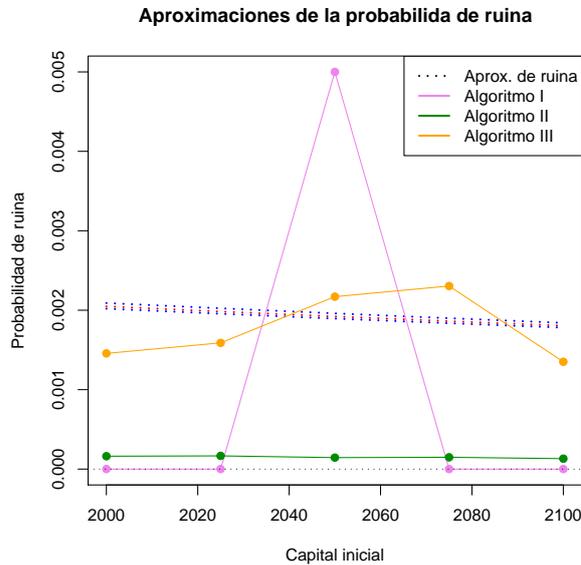


Figura 3.6: sim_3geom_con_Panjer utilizando Log-normal(-1.62,1.80)

Tabla de resultados con $x = 2000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0000000000	0.0020482015	1.0000000	0.0000000000	Inf
II	0.0001620148	0.0018861867	0.9208990	0.0001726125	1.3995721
III	0.0014561931	0.0005920084	0.2890382	0.0024416780	0.9716154
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	0.0020193001
I	0.0000000000		0.000000e+00	ψ_P	0.0020482015
II	0.0001380919		4.784573e-05	lower	0.0020905688
III	0.0011177938		6.767986e-04	Logitud	0.0000712687

Tabla de resultados con $x = 2025$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0000000000	0.0019833209	1.0000000	0.0000000000	Inf
II	0.0001661567	0.0018171642	0.9162230	0.0001571382	1.4074256
III	0.0015886542	0.0003946666	0.1989928	0.0029708411	0.9350677
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	1.955546e-03
I	0.0000000000		0.000000e+00	ψ_P	1.983321e-03
II	0.0001443785		4.355649e-05	lower	2.024002e-03
III	0.0011769167		8.234751e-04	Logitud	6.845604e-05

Tabla de resultados con $x = 2050$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0050000000	0.003078770	1.6024994	0.0707106781	0.4235408
II	0.0001442311	0.001776999	0.9249277	0.0001010871	1.4707974
III	0.0021704200	0.000249190	0.1297034	0.0044704156	0.8649810
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	
I	-0.0048000000	0.0148000000	0.0196000000	ψ_P	1.894567e-03
II	0.0001302212	0.0001582411	0.0000280199	lower	1.921230e-03
III	0.0015508520	0.0027899880	0.0012391360	Logitud	1.960255e-03
					6.568704e-05

Tabla de resultados con $x = 2075$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0000000000	0.0018617682	1.0000000	0.0000000000	Inf
II	0.0001485078	0.0017132604	0.9202329	0.0001073617	1.4538617
III	0.0023045018	0.0004427336	0.2378028	0.0115250145	0.7100021
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	
I	0.0000000000	0.0000000000	0.000000e+00	ψ_P	1.836185e-03
II	0.0001336282	0.0001633874	2.975915e-05	lower	1.861768e-03
III	0.0007072163	0.0039017873	3.194571e-03	Logitud	1.899189e-03
					6.300378e-05

Tabla de resultados con $x = 2100$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0000000000	0.0018047870	1.0000000	0.000000e+00	Inf
II	0.0001311202	0.0016736667	0.9273486	5.891957e-05	1.54169
III	0.0013507576	0.0004540294	0.2515695	4.336350e-03	0.86124
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	
I	0.0000000000	0.0000000000	0.000000e+00	ψ_P	0.0017802433
II	0.0001229544	0.0001392861	1.633167e-05	lower	0.0018047870
III	0.0007497702	0.0019517450	1.201975e-03	Logitud	0.0018406667
					0.0000604234

Aún en el caso cuando el capital inicial fue de $x = 2050$ se tiene que el error relativo del Algoritmo III fue menor respecto a los otros algoritmos. En toda esta simulación se obtuvo que el Algoritmo III tenía la mayor varianza, lo cual no es significativo ya que los otros algoritmos subvaluaron la probabilidad de ruina. También se puede observar que durante toda la simulación ningún algoritmo logró caer dentro de las cotas.

Resultados de `sim_1geom_con_Panjer` utilizando Log-normal(-1.62,1.80)

En esta ocasión se obtuvieron subvaloraciones de la probabilidad de ruina con los tres algoritmos, lo cual es por el hecho de haber utilizado la misma suma geométrica para todas las estimaciones. Una ventaja de sólo realizar una suma geométrica es que se pudo aumentar de $R = 200$ hasta $R = 500$ y

que aún así tener un tiempo de simulación menor con respecto a la simulación anterior.

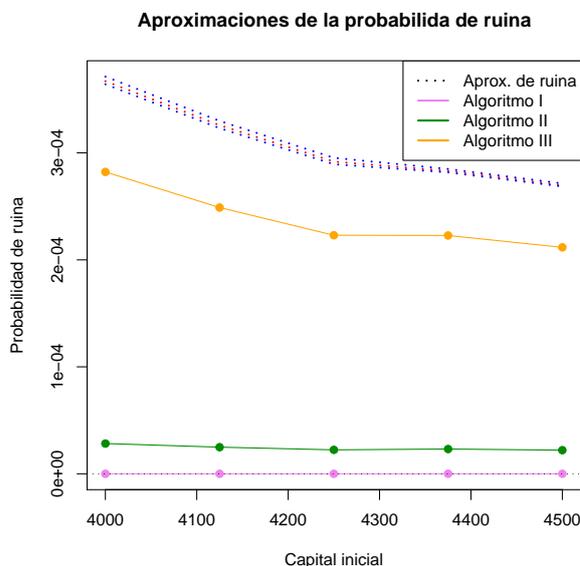


Figura 3.7: `sim_1geom_con_Panjer` utilizando $\text{Log-normal}(-1.62, 1.80)$

Se puede observar en la figura 3.7 una marcada mímica que hace el Algoritmo III con respecto de la probabilidad de ruina, lo cual se coteja al observar los errores relativos obtenidos. Cabe destacar cómo se va comportando el cociente $\log \hat{s} / \log \psi_{\mathbf{P}}$ el cual tiene una marcada tendencia a ser constante durante todos los capitales iniciales, lo cual se interpreta como la velocidad que tiene su convergencia con respecto a su eficiencia logarítmica.

Parámetros		
Distribución	ρ	R
Log-normal (-1.62, 1.80)	0.10	500

Tabla de resultados con $x = 4000$					
Alg.	Aprox.	$E. \text{ absoluto}$	$E. \text{ relativo}$	\hat{s}	$\log \hat{s} / \log \psi_{\mathbf{P}}$
I	0.000000e+00	3.670415e-04	1.0000000	0.000000e+00	Inf
II	2.827923e-05	3.387622e-04	0.9229536	1.211132e-05	1.4312666
III	2.821948e-04	8.484668e-05	0.2311638	6.516262e-04	0.9274344
Intervalo de confianza				Aproximación con Panjer	
Alg.	Extremos de intervalo		Longitud	up	$3.641331e-04$
I	0.000000e+00	0.000000e+00	0.000000e+00	$\psi_{\mathbf{P}}$	3.670415e-04
II	2.721763e-05	2.934084e-05	2.123207e-06	$lower$	3.712766e-04
III	2.250772e-04	3.393123e-04	1.142351e-04	$Logitud$	7.143550e-06

Tabla de resultados con $x = 4125$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	3.258419e-04	1.0000000	0.000000e+00	Inf
II	2.489427e-05	3.009476e-04	0.9236001	1.077836e-05	1.424565
III	2.489079e-04	7.693401e-05	0.2361084	5.986135e-04	0.924250
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	3.230807e-04
II	2.394951e-05	2.583904e-05	1.889530e-06	lower	3.298469e-04
III	1.964371e-04	3.013787e-04	1.049416e-04	Logitud	6.766199e-06

Tabla de resultados con $x = 4250$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	2.918257e-04	1.0000000	0.000000e+00	Inf
II	2.238034e-05	2.694454e-04	0.9233092	9.675362e-06	1.4185313
III	2.229947e-04	6.883103e-05	0.2358635	5.411855e-04	0.9241211
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	2.893792e-04
II	2.153226e-05	2.322842e-05	1.696166e-06	lower	2.953655e-04
III	1.755577e-04	2.704317e-04	9.487399e-05	Logitud	5.986355e-06

Tabla de resultados con $x = 4375$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	2.830926e-04	1.0000000	0.000000e+00	Inf
II	2.314284e-05	2.599498e-04	0.9182499	9.279487e-06	1.418369
III	2.227820e-04	6.031058e-05	0.2130419	4.993550e-04	0.930531
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	2.817694e-04
II	2.232946e-05	2.395623e-05	1.626766e-06	lower	2.849389e-04
III	1.790117e-04	2.665524e-04	8.754079e-05	Logitud	3.169574e-06

Tabla de resultados con $x = 4500$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	2.697760e-04	1.0000000	0.000000e+00	Inf
II	2.209458e-05	2.476814e-04	0.9181003	8.699403e-06	1.417908
III	2.116822e-04	5.809374e-05	0.2153407	4.619037e-04	0.934562
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	2.685874e-04
II	2.133205e-05	2.285712e-05	1.525073e-06	lower	2.714341e-04
III	1.711946e-04	2.521699e-04	8.097528e-05	Logitud	2.846657e-06

Resultados de `sim_1geom_con_asintota` utilizando `Log-normal(-1.62,1.80)`

Vuelven a repetirse dos factores; por una parte se tiene que la probabilidad de ruina es sobrevaluada por la estimación (a excepción de un solo punto de la gráfica) y por otra se tiene que el Algoritmo III tiene un comportamiento mímico con respecto de la estimación asintótica. En esta ocasión también se aumentó el valor de R , haciéndolo llegar a $R = 1000$, el cual es un número grande en comparación a las otras simulaciones al utilizar la distribución Log-normal pero relativamente pequeño en comparación del que se utilizó con las otras distribuciones.

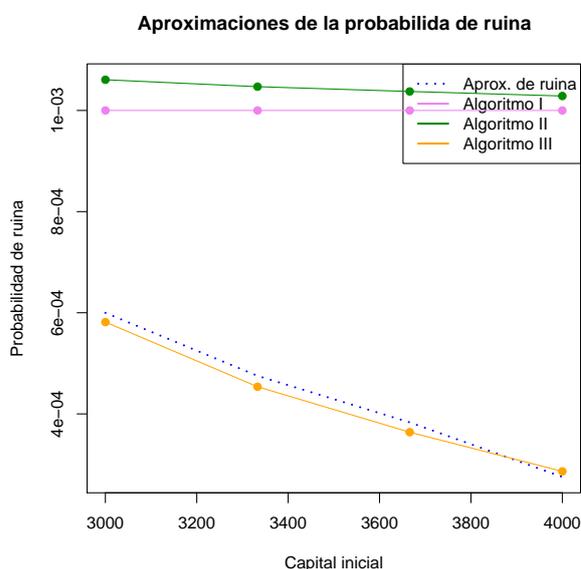


Figura 3.8: `sim_1geom_con_asintota` utilizando `Log-normal(-1.62,1.80)`

En esta simulación se puede observar cómo es manifestada la eficiencia logarítmica del Algoritmo III, al tener en el cociente de los logaritmos valores que se aproximan a uno. También se observa que el error relativo tiene una disminución considerable (en un caso de hasta dos cifras). Este tipo de ventajas también se ven reflejadas en el intervalo de confianza ya que en ocasiones los otros algoritmos llegan a caer en valores negativos.

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
Log-normal (-1.62,1.80)	0.10	1000

Tabla de resultados con $x = 3000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0010000000	4.001429e-04	0.66706381	0.031622777	0.4655563
II	0.0010607209	4.608638e-04	0.76828936	0.031620906	0.4655642
III	0.0005815652	1.829184e-05	0.03049367	0.001268287	0.8990768
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.0005998571	
I	-0.0009600000	0.0029600000	0.0039200000		
II	-0.0008991632	0.0030206049	0.0039197681		
III	0.0005029560	0.0006601745	0.0001572185		

Tabla de resultados con $x = 3333$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0010000000	5.246952e-04	1.10391321	0.0316227766	0.4513956
II	0.0010470094	5.717046e-04	1.20281682	0.0316213006	0.4514017
III	0.0004537453	2.155950e-05	0.04535931	0.0009705364	0.9066996
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.0004753048	
I	-0.0009600000	0.0029600000	0.0039200000		
II	-0.0009128992	0.0030069179	0.0039198170		
III	0.0003935908	0.0005138998	0.0001203089		

Tabla de resultados con $x = 3666$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0010000000	6.166957e-04	1.60889312	0.0316227766	0.4390514
II	0.001037568	6.542640e-04	1.70690494	0.0316215923	0.4390562
III	0.000363879	1.942534e-05	0.05067864	0.0007643726	0.9122596
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.0003833043	
I	-0.0009600000	0.0029600000	0.0039200000		
II	-0.0009223582	0.0029974950	0.0039198532		
III	0.0003165027	0.0004112553	0.0000947526		

Tabla de resultados con $x = 4000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.0010000000	7.245499e-04	2.63042129	0.0316227766	0.4213534
II	0.0010286108	7.531607e-04	2.73429063	0.0316218735	0.4213569
III	0.0002864003	1.095020e-05	0.03975385	0.0006155093	0.9019113
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	0.0002754501	
I	-0.0009600000	0.0029600000	3.920000e-03		
II	-0.0009313332	0.002988555	3.919888e-03		
III	0.0002482507	0.000324550	7.629932e-05		

3.3.3. Distribución de los sumandos: MPE(3)

La simulación de la distribución MPE (ver Apéndice E.1) tiene una gran ventaja, ya que aún cuando no se tiene una expresión analítica para la inversa de su cola integrada, sí se tiene una expresión cuasi-analítica para su cola integrada.² Esto por que al quedar en términos de distribuciones gamma truncadas es fácil su calculo ya que existen programas que lo hacen.

Resultados de `sim_3geom_con_Panjer` utilizando MPE(3)

Se observa en la figura 3.9 que el Algoritmo III tiene un comportamiento oscilatorio alrededor de la probabilidad de ruina. En cambio, los otros algoritmos se comportan como constantes; en este caso es importante no hacer énfasis en la varianza de nuestros estimadores ya que al tener este comportamiento tendrán una varianza pequeña pero que su estimación es sumamente ineficiente.

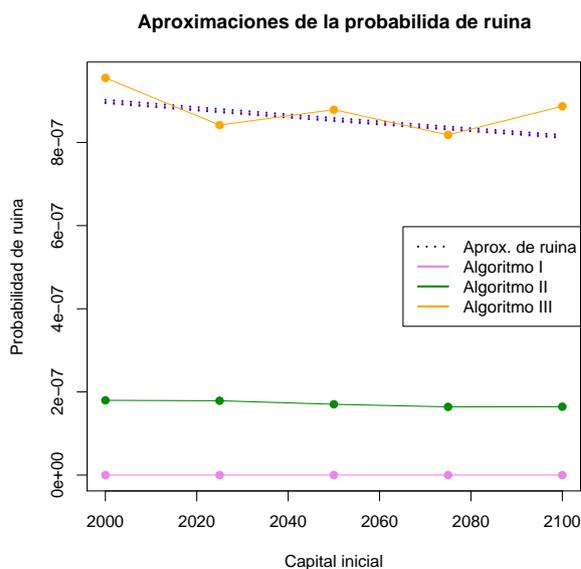


Figura 3.9: `sim_3geom_con_Panjer` utilizando MPE(3)

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
MPE(3)	0.25	1000

²De hecho sí se tiene una expresión analítica en el caso de que su parámetro r sea entero ya que se puede obtener al hacer r -veces la fórmula de integración por partes y lograr que la función gamma truncada sea analítica.

Tabla de resultados con $x = 2000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	8.979143e-07	1.00000000	0.000000e+00	Inf
II	1.797977e-07	7.181167e-07	0.79976078	8.854693e-08	1.166380
III	9.552896e-07	5.737526e-08	0.06389837	2.724555e-06	0.920278
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	8.946830e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	8.979143e-07
II	1.743095e-07	1.852859e-07	1.097639e-08	lower	9.028897e-07
III	7.864199e-07	1.124159e-06	3.377393e-07	Logitud	8.206702e-09

Tabla de resultados con $x = 2025$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	8.757699e-07	1.00000000	0.000000e+00	Inf
II	1.787631e-07	6.970068e-07	0.79587895	8.352186e-08	1.1684806
III	8.418644e-07	3.390547e-08	0.03871504	1.658480e-06	0.9542195
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	8.726576e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	8.757699e-07
II	1.735863e-07	1.839398e-07	1.035348e-08	lower	8.805615e-07
III	7.390708e-07	9.446581e-07	2.055873e-07	Logitud	7.903954e-09

Tabla de resultados con $x = 2050$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	8.544347e-07	1.00000000	0.000000e+00	Inf
II	1.702534e-07	6.841813e-07	0.80074146	8.491219e-08	1.1652366
III	8.787661e-07	2.433142e-08	0.02847662	3.837355e-06	0.8924986
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	8.514356e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	8.544347e-07
II	1.649905e-07	1.755163e-07	1.052582e-08	lower	8.590515e-07
III	6.409244e-07	1.116608e-06	4.756834e-07	Logitud	7.615916e-09

Tabla de resultados con $x = 2075$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	8.338698e-07	1.00000000	0.000000e+00	Inf
II	1.640491e-07	6.698207e-07	0.80326770	8.437775e-08	1.1636595
III	8.182340e-07	1.563583e-08	0.01875092	1.630266e-06	0.9521031
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	8.309785e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	8.338698e-07
II	1.588193e-07	1.692789e-07	1.045957e-08	lower	8.383202e-07
III	7.171890e-07	9.192790e-07	2.020899e-07	Logitud	7.341707e-09

<i>Tabla de resultados con $x = 2100$</i>					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	8.140385e-07	1.0000000	0.000000e+00	Inf
II	1.643619e-07	6.496766e-07	0.7980907	7.913356e-08	1.1662383
III	8.870068e-07	7.296821e-08	0.0896373	3.048519e-06	0.9058285
<i>Intervalo de confianza</i>				<i>Aproximación con Panjer</i>	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	<i>up</i>	
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	8.140385e-07
II	1.594572e-07	1.692667e-07	9.809498e-09	<i>lower</i>	8.183304e-07
III	6.980576e-07	1.075956e-06	3.778984e-07	<i>Logitud</i>	7.080505e-09

Para esta simulación es importante hacer énfasis en el error relativo, el cual es marcadamente menor para el Algoritmo III. También en estos casos se tiene que las cotas de Panjer son sumamente cercanas (con respecto a lo lejano que caen nuestras estimaciones) y que nuestros estimadores en ninguna ocasión logran caer dentro de ellas. Lo anterior es resultado de que para estos capitales iniciales las estimaciones de Panjer son eficientes ya que su ejecución es rápida. Cabe mencionar que en esta simulación tardaba el doble el cálculo de las cotas que el de los algoritmos.

Resultados de `sim_1geom_con_Panjer` utilizando MPE(3)

En esta simulación se pudo aumentar el capital inicial sólo un poco ya que aún cuando el sólo calcular una suma geométrica reduce el tiempo de simulación, el tener que ejecutar las cotas de Panjer lo aumenta de manera considerable. También se aumentó de manera similar el número de simulaciones al hacer $R = 2000$.

Otra vez es importante el resaltar el error relativo ya que ocurre otra vez que la probabilidad de ruina es subvaluada por el hecho de no obtener ruina con ninguna de las simulaciones. Lo anterior se ve reflejado en el comportamiento casi constante de los primeros dos algoritmos.

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
MPE(3)	0.25	2000

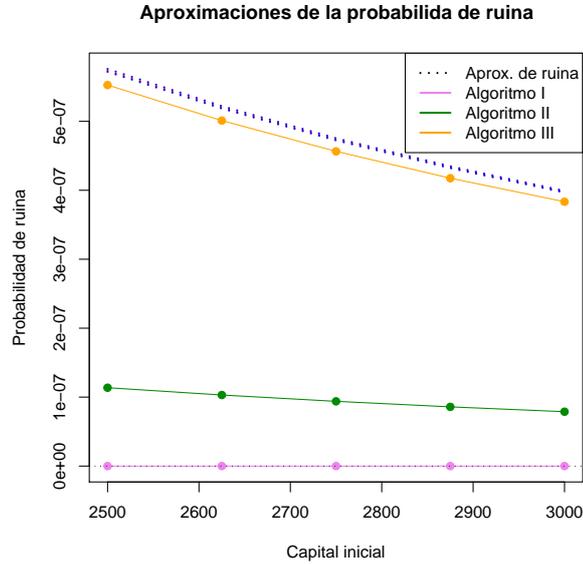


Figura 3.10: sim_1geom_con_Panjer utilizando MPE(3)

Tabla de resultados con $x = 2500$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	5.734940e-07	1.00000000	0.000000e+00	Inf
II	1.136442e-07	4.598497e-07	0.80183882	5.755157e-08	1.1599738
III	5.526033e-07	2.089073e-08	0.03642711	1.119913e-06	0.9534316
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	5.718461e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	5.734940e-07
II	1.111219e-07	1.161666e-07	5.044618e-09	lower	5.760262e-07
III	5.035209e-07	6.016856e-07	9.816472e-08	Logitud	4.180143e-09

Tabla de resultados con $x = 2625$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	5.199744e-07	1.00000000	0.000000e+00	Inf
II	1.030597e-07	4.169147e-07	0.80179855	5.219086e-08	1.1588772
III	5.010836e-07	1.889085e-08	0.03633035	1.015228e-06	0.9537586
Intervalo de confianza				Aproximación con Panjer	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	5.185520e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	5.199744e-07
II	1.007723e-07	1.053470e-07	4.574731e-09	lower	5.221594e-07
III	4.565892e-07	5.455779e-07	8.898869e-08	Logitud	3.607422e-09

<i>Tabla de resultados con $x = 2750$</i>					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	4.736116e-07	1.00000000	0.000000e+00	Inf
II	9.388784e-08	3.797238e-07	0.80176196	4.754573e-08	1.1578462
III	4.564467e-07	1.716490e-08	0.03624257	9.245641e-07	0.9540658
<i>Intervalo de confianza</i>				<i>Aproximación con Panjer</i>	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	4.723753e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	4.736116e-07
II	9.180405e-08	9.597162e-08	4.167567e-09	lower	4.755100e-07
III	4.159259e-07	4.969675e-07	8.104162e-08	Logitud	3.134727e-09

<i>Tabla de resultados con $x = 2875$</i>					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	4.331842e-07	1.00000000	0.000000e+00	Inf
II	8.588804e-08	3.472961e-07	0.80172858	4.349424e-08	1.1568740
III	4.175191e-07	1.566505e-08	0.03616256	8.455241e-07	0.9543551
<i>Intervalo de confianza</i>				<i>Aproximación con Panjer</i>	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	4.321029e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	4.331842e-07
II	8.398182e-08	8.779426e-08	3.812439e-09	lower	4.348440e-07
III	3.804624e-07	4.545758e-07	7.411346e-08	Logitud	2.741140e-09

<i>Tabla de resultados con $x = 3000$</i>					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	3.977205e-07	1.00000000	0.000000e+00	Inf
II	7.886878e-08	3.188518e-07	0.80169800	3.99394e-08	1.1559548
III	3.833671e-07	1.435348e-08	0.03608935	7.76203e-07	0.9546284
<i>Intervalo de confianza</i>				<i>Aproximación con Panjer</i>	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	up	3.967694e-07
I	0.000000e+00	0.000000e+00	0.000000e+00	ψ_P	3.977205e-07
II	7.711836e-08	8.061920e-08	3.500843e-09	lower	3.991802e-07
III	3.493485e-07	4.173856e-07	6.803719e-08	Logitud	2.410781e-09

Resultados de `sim_1geom_con_asintota` utilizando MPE(3)

En esta última simulación se aumentó de manera considerable el capital inicial; por ejemplo en las ejecuciones anteriores sólo se llegó a un máximo de $x = 3000$ en cambio para ésta se logró llegar hasta $x = 5000$. También en este caso se aumentó el valor de R ya que en la ejecución anterior era de $R = 2000$ y en ésta llegó a ser de $R = 10000$.

Se puede observar que el Algoritmo III casi está sobre la estimación asintótica. En cambio para los otros algoritmos el aumento tanto del valor x como de R significó que su comportamiento siguiera siendo el mismo.

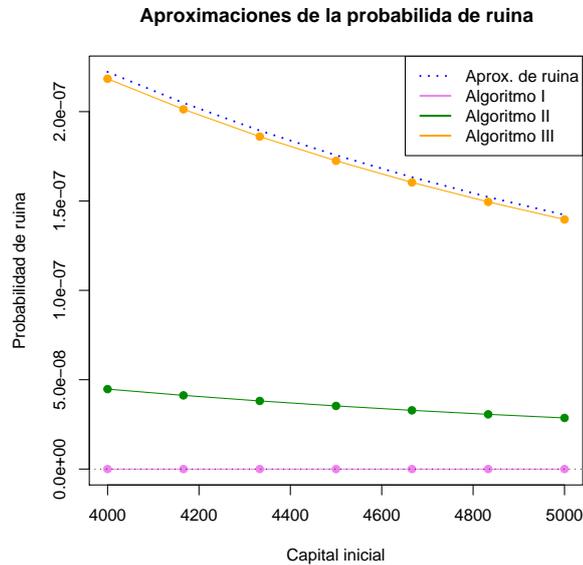


Figura 3.11: `sim_1geom_con_asintota` utilizando MPE(3)

Hay que notar que en este caso la probabilidad de ruina es sumamente pequeña y que esto no llega a afectar a la estimación que nos da el Algoritmo III. Este es un factor importante ya que en el momento de la simulación el hecho de que la probabilidad de ruina sea pequeña a efecto de aumentar el capital inicial hace que sea infactible el cálculo de las cotas de Panjer; es decir, cuando se está calculando probabilidades de ruina para grandes capitales iniciales es enorme el tiempo que demanda el cálculo de las cotas de Panjer, en cambio la simulación no presenta este problema, ya que sin importar el capital inicial el tiempo de ejecución permanece igual.

<i>Parámetros</i>		
<i>Distribución</i>	ρ	R
MPE(3)	0.25	10000

Tabla de resultados con $x = 4000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	2.222222e-07	1.00000000	0.000000e+00	Inf
II	4.473805e-08	1.774842e-07	0.79867880	2.214014e-08	1.1505449
III	2.183845e-07	3.837719e-09	0.01726973	4.881859e-07	0.9486267
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	2.222222e-07	
I	0.000000e+00	0.000000e+00	0.000000e+00		
II	4.430410e-08	4.517199e-08	8.678935e-10		
III	2.088161e-07	2.279529e-07	1.913689e-08		

Tabla de resultados con $x = 4166$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	2.048656e-07	1.00000000	0.000000e+00	Inf
II	4.123970e-08	1.636259e-07	0.79869870	2.040880e-08	1.1497566
III	2.012930e-07	3.572505e-09	0.01743829	4.498839e-07	0.9489228
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	2.048656e-07	
I	0.000000e+00	0.000000e+00	0.000000e+00		
II	4.083969e-08	4.163971e-08	8.000249e-10		
III	1.924753e-07	2.101108e-07	1.763545e-08		

Tabla de resultados con $x = 4333$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	1.893782e-07	1.00000000	0.000000e+00	Inf
II	3.811859e-08	1.512597e-07	0.79871717	1.886416e-08	1.1490022
III	1.860462e-07	3.332064e-09	0.01759476	4.157260e-07	0.9492051
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	1.893782e-07	
I	0.000000e+00	0.000000e+00	0.000000e+00		
II	3.774885e-08	3.848833e-08	7.394750e-10		
III	1.778980e-07	1.941944e-07	1.629646e-08		

Tabla de resultados con $x = 4500$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	1.755830e-07	1.00000000	0.000000e+00	Inf
II	3.533884e-08	1.402442e-07	0.79873427	1.748847e-08	1.1482833
III	1.724682e-07	3.114763e-09	0.01773955	3.853158e-07	0.9494733
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	1.75583e-07	
I	0.000000e+00	0.000000e+00	0.000000e+00		
II	3.499607e-08	3.568161e-08	6.855479e-10		
III	1.649160e-07	1.800204e-07	1.510438e-08		

Tabla de resultados con $x = 4666$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	1.633120e-07	1.00000000	0.000000e+00	Inf
II	3.286653e-08	1.304454e-07	0.79875004	1.626493e-08	1.147601
III	1.603931e-07	2.918899e-09	0.01787315	3.582783e-07	0.949727
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	1.633120e-07	
I	0.000000e+00	0.000000e+00	0.000000e+00		
II	3.254773e-08	3.318532e-08	6.375853e-10		
III	1.533708e-07	1.674153e-07	1.404451e-08		

Tabla de resultados con $x = 4833$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	1.522208e-07	1.00000000	0.000000e+00	Inf
II	3.063217e-08	1.215886e-07	0.79876481	1.515917e-08	1.1469446
III	1.494811e-07	2.739706e-09	0.01799824	3.338513e-07	0.9499702
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	1.522208e-07	
I	0.000000e+00	0.000000e+00	0.000000e+00		
II	3.033505e-08	3.092929e-08	5.942393e-10		
III	1.429376e-07	1.560245e-07	1.308697e-08		
Tabla de resultados con $x = 5000$					
<i>Alg.</i>	<i>Aprox.</i>	<i>E. absoluto</i>	<i>E. relativo</i>	\hat{s}	$\log \hat{s} / \log \psi_P$
I	0.000000e+00	1.422222e-07	1.00000000	0.000000e+00	Inf
II	2.861816e-08	1.136041e-07	0.79877858	1.416245e-08	1.1463158
III	1.396459e-07	2.576347e-09	0.01811494	3.118398e-07	0.9502026
Intervalo de confianza				Aproximación asintótica	
<i>Alg.</i>	<i>Extremos de intervalo</i>		<i>Longitud</i>	1.422222e-07	
I	0.000000e+00	0.000000e+00	0.000000e+00		
II	2.834057e-08	2.889574e-08	5.551680e-10		
III	1.335338e-07	1.457579e-07	1.222412e-08		

Se observa que el comportamiento del error relativo es excelente para el Algoritmo III por que aún cuando disminuimos de manera considerable a la probabilidad de ruina su error relativo permaneció sin presentar un aumento ad hoc.

Conclusiones

- El comportamiento de la probabilidad de ruina representa un problema sumamente complejo aún en el caso de su planteamiento más sencillo.
- A pesar de no lograr expresiones analíticas para la probabilidad de ruina, el replantearla en términos de otras ecuaciones no sólo nos permite ampliar el conocimiento que se tiene de su comportamiento sino además permite el poder obtener estimadores de ella.
- En general, el pedir ciertas propiedades de los estimadores obtenidos es muy importante, ya que el no observar si los estimadores cumplen con ciertos criterios puede ocasionar estimadores sumamente defectuosos.
- El utilizar los métodos de reducción de varianza para la estimación de eventos raros es de suma importancia, ya que optimizan altamente el proceso de simulación.
- Los algoritmos reportan una gran ventaja en comparación a las cotas y a la estimación de Panjer para valores grandes del capital inicial por el hecho de que su tiempo de ejecución no depende del valor del capital inicial sino del número de iteraciones; es decir, para valores grandes del capital inicial los valores obtenidos por Panjer requieren de inmensos lapsos para ser calculados, en cambio los algoritmos permanecen sin cambios en el tiempo requerido.
- Algunos de los programas que fueron creados para la simulación permitían conocer cual era el tiempo utilizado en las distintos calculos que realizaban, lo cual exhibió que la mitad del tiempo era destinado a la simulación y la otra a la obtención del estimador. Por lo tanto para lograr programas aún más veloces se puede utilizar mejores algoritmos de simulación para las variables aleatorias; por ejemplo, estas mejoras pueden hacerse por medio de implementar métodos de reducción de varianza.

- Un problema que puede surgir al momento de estar simulando eventos raros es el periodo de los números pseudoaleatorios ya que al necesitar una cantidad grande de simulaciones esta puede ser lo suficientemente grande como para superar este periodo, lo cual implicaría que el programa repitiera valores ya simulados; es decir, que a partir de cierto valor simulado los valores siguientes fueran iguales (tanto en valor como en posición) que los ya simulados. Lo anterior quitaría completamente lo “aleatorio” a nuestros números pseudoaleatorios y hace que nuestras simulaciones sean sumamente ineficientes.
- Otra ventaja de los algoritmos sobre la estimación por Panjer es el hecho de ser reutilizables; nos referimos a que con una sola corrida de la simulación además de poder ser utilizada para los tres algoritmos puede ser utilizada para distintos valores del capital inicial, lo cual no ocurre para la estimación por Panjer ya que por cada capital inicial requiere una ejecución del programa y que tampoco son reutilizables los valores obtenidos aún para el mismo valor del capital inicial para sus otras estimaciones. Por ejemplo, los valores obtenidos por la cota superior para el capital inicial x , no pueden ser reutilizados para la estimación de esta misma cota en $x + \kappa$ y tampoco pueden ser utilizados para la estimación de la cota inferior en el valor x , a diferencia de los algoritmos que sí ofrecen esta libertad.
- El Algoritmo I tiene una tendencia a subestimar la probabilidad de ruina ya que requiere un número de iteraciones extremadamente alto (de magnitud inversa) para probabilidades de ruina pequeñas. Este algoritmo presenta muy pocos valores (en algunas ocasiones ninguno) que aporten peso para la estimación a causa de sólo manejar ceros y unos como estimadores.
- El Algoritmo II sigue manteniendo una alta subestimación de la probabilidad de ruina aún para valores no tan altos del capital inicial. La mejora que presenta con respecto al Algoritmo I es apenas apreciables para valores altos del capital inicial.
- Para el cálculo de las probabilidades de ruina, el Algoritmo III tiene un comportamiento increíblemente eficiente ya que las restricciones que pide no son altas con respecto a los resultados que otorga. Esto por que aún cuando ya hemos mencionado que existen algoritmos que poseen error relativo acotado, éstos requieren de un gran número de supuestos para su ejecución, a diferencia del Algoritmo III.

- La aproximación asintótica tiene una convergencia que se aproxima por abajo al valor de la ruina y solo es útil cuando las cotas de Panjer son inoperables. Además su convergencia es sumamente lenta lo cual es exhibido por el error relativo que presenta en comparación a la cota inferior.
- La simulación presenta una herramienta sumamente útil para la estimación de la probabilidad de ruina. Los métodos deterministas requieren de lapsos enormes (en algunos casos lapsos imposibles) para su ejecución cuando la probabilidad de ruina es pequeña mientras que los métodos que incorporan factores aleatorios llegan a ser mucho más rápidos, cuando son creados tomando en cuenta los factores que rigen al fenómeno de estudio.

Apéndice A

Tabla de distribuciones

$X \sim$	$f(x)$	$F(x)$	$\mathbb{E}[X]$	$Var(X)$
<i>Ber</i> (p)	$p^x(1-p)^{1-x}$ $p \in (0, 1)$ $x \in \{0, 1\}$	-	p	$p(1-p)$
<i>Poisson</i> (λ)	$e^{-\lambda} \frac{\lambda^x}{x!}$ $\lambda > 0$ $x = 0, 1, \dots$	-	λ	λ
<i>Exp</i> (λ)	$\lambda e^{-\lambda x}$ $\lambda > 0$ $x \geq 0$	$1 - e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
<i>Geo</i> (p)	$p(1-p)^x$ $p \in (0, 1)$ $x = 0, 1, \dots$	-	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
<i>Log-normal</i> (μ, σ^2)	$\frac{\exp(-(\ln x - \mu)^2 / 2\sigma^2)}{x\sqrt{2\pi\sigma^2}}$ $\mu \in \mathbb{R}, \sigma^2 > 0$ $x \geq 0$	-	$\exp\left(\mu + \frac{\sigma^2}{2}\right)$	$\exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2)$
<i>Pareto</i> (a, b)	$ba^b x^{-(b+1)}$ $a > 0, b > 1$ $x \geq a$	$1 - \left(\frac{a}{x}\right)^b$	$\frac{ab}{b-1}$ Si $b > 2$	$\frac{a^2 b(b^2 - 3b + 3)}{(b-1)(b-2)}$
<i>MPE</i> (r)	$r \binom{r-1}{r} x^{-(r+1)}$ $\times \gamma\left(r+1, \frac{r}{r-1}x\right)$ $r > 1, x \geq 0$	-	1	-

Apéndice B

Teoremas y resultados

B.1. Teoremas de probabilidad

Teorema B.1.1 (Desigualdad de Jensen). *Sea X una variable aleatoria tal que $\mathbb{E}[X] < \infty$ y h una función convexa, entonces*

$$h(\mathbb{E}[X]) \leq \mathbb{E}[h(X)] \quad (\text{B.1})$$

Teorema B.1.2 (Teorema de Slutsky). *Sean $\{X_n\}$ y $\{Y_n\}$ sucesiones de variables aleatorias tales que $X_n \xrightarrow{d} X$ y $Y_n \xrightarrow{d} c \in \mathbb{R}$ entonces*

$$X_n Y_n \xrightarrow{d} cX, \quad X_n + Y_n \xrightarrow{d} X + c \quad (\text{B.2})$$

Teorema B.1.3 (Desigualdad de Markov). *Sea X una v.a. positiva con esperanza finita. Para cualquier $\varepsilon > 0$*

$$\mathbb{P}[X \geq \varepsilon] \leq \frac{\mathbb{E}[X]}{\varepsilon}.$$

Teorema B.1.4 (Desigualdad de Chebyshev). *Sea X una v.a. y f, g funciones monótonas no decrecientes positivas, entonces*

$$\mathbb{E}[f(X)g(X)] \geq \mathbb{E}[f(X)] \mathbb{E}[g(X)].$$

Teorema B.1.5 (Teorema de convergencia monótona). *Sea $0 \leq X_1 \leq X_2 \leq \dots$, una sucesión de v.a. convergente casi seguramente a una variable aleatoria X , entonces*

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

Una mayor referencia de este resultado puede ser consultada en Rolski(1998)[32]

Teorema B.1.6 (Ley débil de los grandes números). Sean X_1, X_2, \dots una sucesión de v.a.i.i.d. con media μ . Entonces

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mu \quad \text{cuando } n \rightarrow \infty. \quad (\text{B.3})$$

Teorema B.1.7 (Ley fuerte de los grandes números). Sean X_1, X_2, \dots una sucesión de v.a.i.i.d. con media μ . Entonces

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{c.s.} \mu \quad \text{cuando } n \rightarrow \infty. \quad (\text{B.4})$$

Teorema B.1.8 (Teorema central del límite). Sean X_1, X_2, \dots una sucesión de v.a.i.i.d. con $\mathbb{E}[X_i] = \mu$ y $\text{Var}(X_i) = \sigma^2 < \infty$. Entonces

$$\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} \mathcal{N}(0, 1) \quad \text{cuando } n \rightarrow \infty. \quad (\text{B.5})$$

Teorema B.1.9 (Teorema de Wald). Sea $\{X_n\}_{n \in \mathbb{N}}$ v.a.i.i.d. con $\mathbb{E}[|X_1|] < \infty$ y sea α un tiempo de paro con respecto a estas variables aleatorias tal que $\mathbb{E}[\alpha] < \infty$, entonces

$$\mathbb{E} \left[\sum_{i=1}^{\alpha} X_i \right] = \mathbb{E}[X_1] \mathbb{E}[\alpha]. \quad (\text{B.6})$$

Demostración. Tenemos que

$$\mathbb{E} \left[\sum_{i=1}^{\alpha} X_i \right] = \mathbb{E} \left[\sum_{i=1}^{\infty} X_i \mathbb{1}_{\{i, \infty\}} \alpha \right]$$

Además se tiene que

$$\mathbb{E} \left[\sum_{i=1}^{\infty} |X_i| \mathbb{1}_{\{i, \infty\}} \alpha \right] = \mathbb{E} \left[\sum_{i=1}^{\infty} |X_i| \mathbb{1}_{\{i, \infty\}} \alpha \right]$$

Como todos los términos de la suma son positivos se tiene que por el teorema de convergencia monótona

$$\sum_{i=1}^{\infty} \mathbb{E}[X_i] \mathbb{1}_{\{i, \infty\}} \alpha = \mathbb{E}[X] \mathbb{E}[\alpha] < \infty$$

se tiene entonces que y utilizando que

$$\mathbb{E} \left[\sum_{i=1}^{\infty} X_i \mathbb{1}_{\{[i, \infty)\}} \alpha \right] = \sum_{i=1}^{\infty} \mathbb{E} [X_i \mathbb{1}_{\{[i, \infty)\}} \alpha]$$

Como la función indicadora es sobre el conjunto $[i \leq \alpha] = [\alpha < i]^c = [\alpha \leq i - 1]^c$ lo cual solo depende de las v.a. X_1, \dots, X_{i-1} tenemos entonces que X_i y $\mathbb{1}_{\{[0, i)\}} \alpha$ son independientes, lo cual implica que

$$\begin{aligned} \sum_{i=1}^{\infty} \mathbb{E} [X_i \mathbb{1}_{\{[i, \infty)\}} \alpha] &= \sum_{i=1}^{\infty} \mathbb{E} [X_i] \mathbb{E} [\mathbb{1}_{\{[i, \infty)\}} \alpha] = \mathbb{E} [X_1] \sum_{i=1}^{\infty} \mathbb{P} [\alpha \geq i] \\ &= \mathbb{E} [X_1] \sum_{i=0}^{\infty} \mathbb{P} [\alpha > i] = \mathbb{E} [X_1] \mathbb{E} [\alpha] \end{aligned}$$

Esta última igualdad por que α es una v.a. no negativa □

B.2. Teoremas de análisis

Teorema B.2.1 (Teorema de convergencia acotada). *Sea (f_n) una sucesión acotada de funciones integrables con respecto a una función g monótonamente creciente de $J = [a, b]$ a \mathbb{R} . Supongamos que existe $B > 0$ tal que $\|f_n(x)\| \leq B$ para toda $n \in \mathbb{N}$, $x \in J$. Si la función $f(x) = \lim(f_n(x))$, $x \in J$, existe y es integrable con respecto a g en J , entonces,*

$$\int_a^b f dg = \lim \int_a^b f_n dg. \quad (\text{B.7})$$

Teorema B.2.2 (Teorema de convergencia dominada). *Sea (f_n) una sucesión de funciones de valor real, tal que $f(x) = \lim(f_n(x))$ para toda $x \geq a$, y que f y $f_n, n \in \mathbb{N}$, son Riemann integrables sobre $[a, c]$ para toda $c > a$. Además existe una función M que tiene una integral sobre $x \geq a$ y que*

$$|f_n(x)| \leq M(x) \quad x \geq a, \quad n \in \mathbb{N}$$

Entonces, f tiene una integral sobre $x \geq a$ y

$$\int_a^{\infty} f = \lim \int_a^{\infty} f_n. \quad (\text{B.8})$$

Teorema B.2.3 (Prueba M de Weierstrass). *Supongamos que f es Riemann integrable sobre $[a, c]$ para toda $c \geq a$ y toda $t \in [\alpha, \beta]$. Suponga que existe una función M definida para $x \geq a$ tal que*

$$|f(x, t)| \leq M(x) \quad \text{para } x \geq a, \quad t \in [\alpha, \beta] \quad (\text{B.9})$$

y tal que la integral infinita $\int_a^\infty M(x)dx$ exista. Entonces, para cada $t \in [\alpha, \beta]$, la integral

$$F(t) = \int_a^\infty f(x, t)dx \quad (\text{B.10})$$

es (absolutamente) convergente y la convergencia es uniforme en $[\alpha, \beta]$.

Teorema B.2.4. Si f es continua por pedazos sobre $[0, \infty)$ con transformada de Laplace $\hat{f}(s)$ y además existen constantes $M > 0$ y $T > 0$ tales que $|f(t)| \leq Me^{ct}$ para toda $t > T$ entonces $\hat{f}(s)$ existe para $s > c$.

Apéndice C

Transformada de Laplace y convoluciones

Definición C.0.1 (Transformada de Laplace). *Sea f una función definida sobre $t \geq 0$ entonces la integral*

$$\hat{f}(s) := \int_0^{\infty} e^{-st} f(t) dt \quad (\text{C.1})$$

se llamará la transformada de Laplace de f , si es que la integral existe.

Daremos por hecho que la transformada de Laplace es única y que caracteriza a las funciones.

Observación C.0.1. *Si X tiene función de densidad $f(x)$ entonces la transformada de Laplace $\hat{f}(-s)$ será la función generadora de momentos y tendremos la notación*

$$m_X(s) = \mathbb{E}[e^{sx}] = \hat{f}(-s). \quad (\text{C.2})$$

En caso de ser obvio sobre qué función se está aplicando la transformada se utilizará la notación $m(s) = m_X(s)$

Definición C.0.2 (Convoluciones). *Sea F una función de distribución sobre R_+ y g una función definida sobre R_+ y tal que g es localmente acotada (acotada en intervalos finitos) entonces definimos la convolución de F y g como la función*

$$F * g(t) := \int_0^t g(t-x) dF(x), \quad \text{para } t \geq 0. \quad (\text{C.3})$$

Además por notación utilizaremos

$$\begin{aligned} F^{0*}(x) &:= \mathbb{1}_{\{x \in [0, \infty)\}} \\ F^{1*}(x) &:= F(x) \end{aligned}$$

y para $n \geq 1$

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

Proposición C.0.1. [Propiedades de las convoluciones] Sea F una función de distribución sobre \mathbb{R}_+ y g una función definida sobre \mathbb{R}_+ y tal que g es localmente acotada, entonces:

i. $F * g(t) \geq 0$

ii. $F * g(t)$ es localmente acotada, además

$$\sup_{0 \leq s \leq t} |F * g(s)| \leq \left(\sup_{0 \leq s \leq t} |g(s)| \right) F(t).$$

iii. Si X_1 se distribuye F_1 y X_2 se distribuye F_2 cuya densidad es f_2 entonces $\mathbb{P}[X_1 + X_2 \leq z] = F_1 * F_2(z)$.

iv. Si f_1 y f_2 son funciones de densidad entonces $\widehat{f_1 * f_2} = \hat{f}_1 \times \hat{f}_2$.

v. Si F_1 y F_2 son funciones de distribución entonces $F_1 * F_2(t) = F_2 * F_1(t)$.

Demostración.

i. Dado que la integral de dos funciones positivas es positiva se tiene que la convolución es positiva

ii. Sea $\|g\| := \sup_{0 \leq s \leq t} |g(s)|$ lo cual es finito para cada t por hipótesis. Entonces para cualquier $s \leq t$

$$\begin{aligned} |F * g(s)| &= \left| \int_0^s g(s-x)F(dx) \right| \\ &\leq \int_0^s |g(s-x)| |F(dx)| \\ &\leq \|g\| \int_0^s F(dx) \\ &\leq \|g\| F(t) < \infty. \end{aligned}$$

iii. Para $t \geq 0$ se tiene que

$$\begin{aligned} \mathbb{P}[X_1 + X_2 \leq t] &= \mathbb{P}[(X_1, X_2) \in \{(x, y) \in \mathbb{R}_+^2 : x + y \leq t\}] \\ &= \iint_{\{(x, y) \in \mathbb{R}_+^2 : x + y \leq t\}} F_1(dx) F_2(dy) \\ &= \int_0^t \left[\int_0^{t-x} F_2(dy) \right] F_1(dx) \\ &= \int_0^t F_2(t-x) F_1(dx). \end{aligned}$$

iv.

$$\begin{aligned}\hat{f}_1(s) \cdot \hat{f}_2(s) &= \left(\int_0^\infty e^{-s\tau} f_1(\tau) d\tau \right) \cdot \left(\int_0^\infty e^{-s\beta} f_2(\beta) d\beta \right) \\ &= \int_0^\infty \int_0^\infty e^{-s(\tau+\beta)} f_1(\tau) f_2(\beta) d\beta d\tau \\ &= \int_0^\infty f_1(\tau) \int_0^\infty e^{-s(\tau+\beta)} f_2(\beta) d\beta d\tau\end{aligned}$$

haciendo $t = \tau + \beta$ tenemos

$$\hat{f}_1(s) \cdot \hat{f}_2(s) = \int_0^\infty e^{-st} \int_\tau^\infty f_1(\tau) f_2(t - \tau) dt d\tau$$

con lo cual queda demostrado.

v. Esta es una consecuencia inmediata de iii.

□

C.1. Convoluciones geométricas

La fórmula de Pollaczek-Khinchin se basa en la transformada de Laplace de una convolución geométrica, la cual es una suma de v.a.i.i.d. X_i donde el número de sumandos se distribuye de forma geométrica, también independiente de los sumandos. Supongamos que M es una v.a. geométrica, entonces tenemos que

$$P_m = \mathbb{P}[M = m] = pq^m \quad m = 0, 1, 2, \dots \quad \text{para } p = (1 - q) \in (0, 1).$$

Entonces la suma aleatoria¹

$$S_M = \sum_{i=1}^M X_i$$

tiene distribución geométrica compuesta (o convolución geométrica) de parámetros (p, F) donde F es la función de distribución de X_i . Dado que M y la sucesión de v.a.i.i.d. X_i son independientes entre sí, tenemos que

¹Utilizaremos que $\sum_{i=1}^0 X_i = 0$

$$\begin{aligned}
\mathbb{P}[S_M \leq x] &= \mathbb{P}[M = 0] + \mathbb{P}\left[\sum_{i=1}^M X_i \leq x\right] \\
&= \mathbb{P}[M = 0] + \sum_{m=1}^{\infty} \mathbb{P}\left[\sum_{i=1}^M X_i \leq x \mid M = m\right] \mathbb{P}[M = m] \\
&= \mathbb{P}[M = 0] + \sum_{m=1}^{\infty} \mathbb{P}\left[\sum_{i=1}^m X_i \leq x\right] \mathbb{P}[M = m] \\
&= p + \sum_{i=1}^{\infty} F^{m*}(x) pq^m \\
&= \sum_{i=0}^{\infty} F^{m*}(x) pq^m.
\end{aligned}$$

Lema C.1.1. Sea H una función de distribución con función de densidad h y definimos las funciones G_1 y G_2 para $q \in (0, 1)$ como

$$G_1(x) := (1 - q) \sum_{m=0}^{\infty} q^m H^{m*}(x) \quad G_2(x) := (1 - q) + q \int_0^x G_2(x - z) h(z) dz$$

para $x > 0$. Entonces sus transformadas de Laplace son iguales, lo cual implica que $G_1(x) = G_2(x)$.

Demostración. Para G_1 tenemos

$$\begin{aligned}
\hat{G}_1(s) &= \int_0^{\infty} e^{-sx} G_1(x) dx \\
&= \int_0^{\infty} e^{-sx} (1 - q) \sum_{m=0}^{\infty} q^m H^{m*}(x) dx \\
&= (1 - q) \sum_{m=0}^{\infty} q^m \int_0^{\infty} e^{-sx} H^{m*}(x) dx \\
&= \frac{1 - q}{1 - q\hat{h}(s)},
\end{aligned}$$

donde para el intercambio de la integral y la suma utilizamos el Teorema de Convergencia Acotada (ver Apéndice B.2.1) utilizando que $\sum q^n H^{m*}$ converge uniformemente en un intervalo alrededor del cero y de que $q^n H^{m*}$ es continua.

Se tiene que la transformada de Laplace para G_2 es

$$\hat{G}_2(s) = (1 - q) + q\hat{g}_2(s)\hat{h}(s).$$

En esta igualdad utilizamos el hecho de que la transformada de Laplace de una convolución es el producto de las transformadas de Laplace (ver proposición C.0.1), con lo cual se concluye la demostración. \square

Apéndice D

Distribuciones subexponenciales

D.1. Definición

Definición D.1.1 (Distribución subexponencial). *Sea F una distribución en \mathbb{R}_+ , se dice que F es subexponencial si para toda $n \geq 2$*

$$\lim_{x \rightarrow \infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} = n. \quad (\text{D.1})$$

D.2. Propiedades y resultados

Lema D.2.1 (Condición suficiente para ser subexponencial). *Sea F una distribución en \mathbb{R}_+ , si*

$$\limsup_{x \rightarrow \infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \leq 2 \quad (\text{D.2})$$

entonces F es subexponencial

Demostración. La demostración se hará por inducción sobre n . Se tiene que si F es función de distribución de X_1 y X_2 se sigue que

$$F^{*2}(x) = \mathbb{P}[X_1 + X_2 \leq x] \leq \mathbb{P}[X_1 \leq x] \mathbb{P}[X_2 \leq x] = F^2(x),$$

lo cual implica que $\overline{F^{*2}}(x) \geq \overline{F^2}(x)$ para toda $x \geq 0$. Por lo tanto

$$\overline{F^{*2}}(x) \geq 1 - F^2(x) = (1 - F(x))(1 + F(x))$$

entonces

$$\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \geq 1 + F(x),$$

se sigue que

$$2 \geq \limsup_{x \rightarrow \infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \geq \lim_{x \rightarrow \infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \geq 1 + F(x) = 2$$

por lo tanto la ecuación (D.1) es válida para $n = 2$.

Supongamos que es válida para $n \geq 2$, como

$$\begin{aligned} F(x) - F^{*(n+1)}(x) &= \int_0^x dF(t) - \int_0^x F^{*n}(x-t)dF(t) \\ &= \int_0^x \overline{F^{*n}}(x-t)dF(t), \end{aligned} \quad (\text{D.3})$$

se sigue que para $x \geq y > 0$

$$\begin{aligned} \frac{\overline{F^{*(n+1)}}(x)}{\overline{F}(x)} &= \frac{\overline{F}(x) + F(x) - F^{*(n+1)}(x)}{\overline{F}(x)} \\ &= 1 + \frac{F(x) - F^{*(n+1)}(x)}{\overline{F}(x)} \\ &= 1 + \int_0^x \frac{\overline{F^{*n}}(x-t)}{\overline{F}(x)} dF(t) \\ &= 1 + \left(\int_0^{x-y} + \int_{x-y}^x \right) \left(\frac{\overline{F^{*n}}(x-t)}{\overline{F}(x-t)} \frac{\overline{F}(x-t)}{\overline{F}(x)} \right) dF(t) \\ &= 1 + I_1(x) + I_2(x). \end{aligned} \quad (\text{D.4})$$

Por hipótesis podemos hacer $(\overline{F^{*n}}(x-t)/\overline{F}(x-t)) - n$ arbitrariamente pequeño para $0 \leq t \leq x - y$ y y suficientemente grande se tiene que

$$\begin{aligned} I_1(x) &= \int_0^{x-y} \frac{\overline{F^{*n}}(x-t)}{\overline{F}(x-t)} \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\ &= \int_0^{x-y} \left(\frac{\overline{F^{*n}}(x-t)}{\overline{F}(x-t)} - n + n \right) \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\ &= \int_0^{x-y} (o(1) + n) \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\ &= (n + o(1)) \int_0^{x-y} \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t). \end{aligned}$$

Al utilizar (D.3) tenemos que esta última integral la podemos expresar como

$$\begin{aligned}
\int_0^{x-y} \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) &= \int_0^x \frac{\overline{F}(x-t)}{\overline{F}(x)} - \int_{x-y}^x \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\
&= \frac{F(x) - F^{*2}(x)}{\overline{F}(x)} - \int_{x-y}^x \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\
&= \frac{F(x) - F^{*2}(x)}{\overline{F}(x)} - J(x, y) \\
&= (1 + o(1)) - J(x, y)
\end{aligned}$$

donde en esta última igualdad utilizamos el hecho de que F cumple con (D.2). Tenemos que

$$\begin{aligned}
J(x, y) &= \int_{x-y}^x \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\
&= \int_{x-y}^x \frac{dF(t)}{\overline{F}(x)} dF(t) - \int_{x-y}^x \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\
&= \frac{F(x) - F(x-y)}{\overline{F}(x)} - \int_{x-y}^x \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\
&\leq \frac{F(x) - F(x-y)}{\overline{F}(x)}
\end{aligned}$$

esta última desigualdad se obtiene porque la integral se hace sobre una función positiva. Si utilizamos el Lema D.6 (que se expone más adelante) tenemos que

$$\lim_{x \rightarrow \infty} \frac{F(x) - F(x-y)}{\overline{F}(x)} = \lim_{x \rightarrow \infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} - \lim_{x \rightarrow \infty} \frac{\overline{F}(x)}{\overline{F}(x)} = 1 - 1 = 0.$$

Por lo tanto $J(x, y) \rightarrow 0$ cuando $x \rightarrow \infty$ lo que nos da $\lim_{x \rightarrow \infty} I_1(x) = n$. Por último, como $\overline{F^{*n}}(x-t)/\overline{F}(x-t)$ es acotado (por hipótesis de inducción) para $x-y \leq t \leq x$ tenemos

$$\begin{aligned}
I_2(x) &= \int_{x-y}^x \frac{\overline{F^{*n}}(x-t)}{\overline{F}(x-t)} \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\
&\leq K \int_{x-y}^x \frac{\overline{F}(x-t)}{\overline{F}(x)} dF(t) \\
&= KJ(x, y)
\end{aligned}$$

como $J(x, y) \rightarrow 0$ cuando $x \rightarrow \infty$ se tendrá que $\lim_{x \rightarrow \infty} I_2(x) = 0$, lo cual completa la prueba. \square

Observación D.2.1. Como consecuencia de (D.1) se tiene que si F es subexponencial y X_1, \dots, X_n son v.a.i.i.d. con distribución F entonces cuando $x \rightarrow \infty$

$$\mathbb{P}[X_1 + \dots + X_n > x] \sim \mathbb{P}[\text{máx}\{X_1, \dots, X_n\} > x]$$

esto porque la probabilidad del máximo podemos reescribirla como

$$\begin{aligned} \mathbb{P}[\text{máx}\{X_1, \dots, X_n\} > x] &= 1 - F^n(x) \\ &= (1 - F(x))(1 + F(x) + \dots + F^{n-1}(x)) \end{aligned}$$

y por lo tanto

$$\begin{aligned} 1 &= \lim_{x \rightarrow \infty} \frac{1 - F^{*n}(x)}{n(1 - F(x))} = \lim_{x \rightarrow \infty} \frac{1 - F^{*n}(x)}{(1 + F(x) + \dots + F^{n-1}(x))(1 - F(x))} \\ &= \lim_{x \rightarrow \infty} \frac{1 - F^{*n}(x)}{1 - F^n(x)}. \end{aligned}$$

Esto quiere decir que en el caso de distribuciones subexponenciales el hecho de que se tenga un gran monto a pagar no dependerá de la suma de montos sino de que uno de los sumandos causa que el monto sea grande.

Si utilizamos la siguiente identidad

$$\begin{aligned} \overline{F^{*2}}(x) &= 1 - \int_0^x F(x-y)dF(y) = 1 - \int_0^x [1 - \overline{F}(x-y)]dF(y) \\ &= \overline{F}(x) + \int_0^x \overline{F}(x-y)dF(y) \end{aligned}$$

tenemos que

$$\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} = 1 + \int_0^x \frac{\overline{F}(x-y)}{\overline{F}(x)} dF(y) \quad (\text{D.5})$$

Esta última la utilizaremos en la demostración de los siguientes teoremas, los cuales establecen la conexión entre las distribuciones subexponenciales y las de colas pesadas, además de que serán útiles cuando demos demos algunas propiedades asintóticas de nuestros estimadores.

Lema D.2.2. Si F es subexponencial entonces para toda $x' > 0$ que esté contenida en un intervalo compacto de $(0, \infty)$ se tendrá convergencia uniforme, es decir

$$\lim_{x \rightarrow \infty} \frac{\overline{F}(x - x')}{\overline{F}(x)} = 1 \quad (\text{D.6})$$

Demostración. Para $x' \leq x$ se tendrá utilizando (D.5) que

$$\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} = 1 + \int_0^{x'} \frac{\overline{F}(x-y)}{\overline{F}(x)} dF(y) + \int_{x'}^x \frac{\overline{F}(x-y)}{\overline{F}(x)} dF(y). \quad (\text{D.7})$$

Para acotar la primera integral utilizamos en hecho de que $\overline{F}(x-y)$ es una función creciente con respecto a y y entonces tomará su mínimo en el extremo inferior de la integral

$$\int_0^{x'} \frac{\overline{F}(x-y)}{\overline{F}(x)} dF(y) \geq \frac{\overline{F}(x-0)}{\overline{F}(x)} [F(x') - F(0)] = F(x').$$

Utilizando un argumento análogo al anterior para la segunda integral obtenemos que

$$\int_{x'}^x \frac{\overline{F}(x-y)}{\overline{F}(x)} dF(y) \geq \frac{\overline{F}(x-x')}{\overline{F}(x)} [F(x) - F(x')].$$

Como consecuencia de las ultimas dos desigualdades y utilizando (D.7) tenemos

$$\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \geq 1 + F(x') + \frac{\overline{F}(x-x')}{\overline{F}(x)} (F(x) - F(x')).$$

Al tomar un x lo suficientemente grande para que $F(x) - F(x') \neq 0$ podemos reescribir la ultima desigualdad como

$$\frac{\overline{F}(x-x')}{\overline{F}(x)} \leq \left(\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} - 1 - F(x') \right) (F(x) - F(x'))^{-1},$$

como $\overline{F}(x-y)$ es una función creciente con respecto a y se sigue que

$$1 \leq \frac{\overline{F}(x-x')}{\overline{F}(x)} \leq \left(\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} - 1 - F(x') \right) (F(x) - F(x'))^{-1}.$$

Al ser F subexponencial el lado derecho de la desigualdad tiende a 1 cuando $x \rightarrow \infty$, con lo cual se obtiene el límite deseado. \square

Lema D.2.3. *Si F es subexponencial, dado $\varepsilon > 0$ existe una constante K tal que para toda $n \geq 2$*

$$\frac{\overline{F^{n*}}(x)}{\overline{F}(x)} \leq K(1 + \varepsilon)^n \quad \text{para} \quad x \geq 0 \quad (\text{D.8})$$

Demostración. Sea $\alpha_n = \sup_{x \geq 0} \overline{F^{n*}}(x)/\overline{F}(x)$. Utilizando (D.4) y (D.3) se tiene para $T < \infty$

$$\begin{aligned} \alpha_{n+1} &\leq 1 + \sup_{0 < x \leq T} \int_0^x \frac{\overline{F^{n*}}(x-y)}{\overline{F}(x)} dF(y) + \sup_{x > T} \int_0^x \frac{\overline{F^{n*}}(x-y)}{\overline{F}(x-y)} \frac{\overline{F}(x-y)}{\overline{F}(x)} dF(y) \\ &\leq 1 + A_T + \alpha_n \sup_{x \geq T} \frac{F(x) - F^{*2}(x)}{\overline{F}(x)} \end{aligned}$$

donde para acotar definimos $A_T := (\overline{F}(T))^{-1} < \infty$. Como F es subexponencial tenemos que dada $\varepsilon > 0$ podemos escoger a T de tal forma que

$$\begin{aligned} \sup_{x \geq T} \frac{F(x) - F^{*2}(x)}{\overline{F}(x)} &\leq \sup_{x \geq T} \frac{1 - F^{*2}(x)}{\overline{F}(x)} + \sup_{x \geq T} \frac{F(x) - 1}{\overline{F}(x)} \\ &= \sup_{x \geq T} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} - 1 \\ &\leq 2 + \varepsilon - 1 \\ &= 1 + \varepsilon \end{aligned}$$

esto último porque el límite de $\overline{F^{*2}}(x)/\overline{F}(x)$ es 2. Por lo tanto tenemos que

$$\alpha_{n+1} \leq 1 + A_T + \alpha_n(1 + \varepsilon).$$

Entonces utilizando que $\alpha_1 = 1$ tenemos que

$$\begin{aligned} \alpha_n &\leq (1 + A_T) \sum_{i=0}^{n-1} (1 + \varepsilon)^i \\ &= (1 + A_T) \frac{(1 + \varepsilon)^n - 1}{(1 + \varepsilon) - 1} \\ &\leq (1 + A_T) \varepsilon^{-1} (1 + \varepsilon)^n \end{aligned}$$

lo cual implica (D.8). □

El siguiente teorema muestra que las distribuciones subexponenciales son un subconjunto de las distribuciones de colas pesadas.

Teorema D.2.1. *Si F es una distribución subexponencial entonces F es de cola pesada*

Demostración. Sea F una distribución subexponencial. Utilizando el Teorema 1.5.1 tenemos que es suficiente demostrar que $\limsup_{x \rightarrow \infty} -\log \overline{F}(x)/x = 0$. Tomando logaritmos en (D.6) tenemos que

$$\lim_{x \rightarrow \infty} (\log \overline{F}(x-y) - \log \overline{F}(x)) = 0$$

para toda $y \geq 0$. Entonces para toda $\varepsilon > 0$ existe $x_0 > 0$ tal que para toda $x \geq x_0$ tenemos que $-\log \bar{F}(x) + \log \bar{F}(x-1) < \varepsilon^1$ y por inducción tenemos que

$$-\log \bar{F}(x) \leq -\log \bar{F}(x-1) + \varepsilon \leq -\log \bar{F}(x-2) + 2\varepsilon \leq \dots \leq -\log \bar{F}(x-n) + n\varepsilon$$

donde n es tal que $x_0 \leq x - n < x_0 + 1$, entonces

$$-\log \bar{F}(x) \leq \sup_{x_0 \leq x' \leq x_0+1} -\log \bar{F}(x') + (x - x_0)\varepsilon, \quad \text{para toda } x \geq x_0.$$

Al dividir ambos lados de esta última desigualdad entre x y como ε es arbitraria tenemos el límite deseado. \square

¹Esto pasa porque $-\log \bar{F}(x)$ es una función creciente.

Apéndice E

Variación regular

La teoría de las funciones de variación regular es una herramienta esencial cuando se trabaja con funciones de colas pesadas. En este trabajo se dará un tratamiento somero a esta teoría, la cual es muy vasta y puede ser consultada para un primer acercamiento en Mikosch(2009)[22] y Resnick(1987)[25] o para un estudio más profundo en Bingham(1987)[9] y Feller(1971)[16]. Una forma de interpretar a las funciones de variación regular es pensando que su comportamiento asintótico es el de una función potencia, lo cual será demostrado por medio de los distintos teoremas que expondremos.

Definición E.0.1 (Variación regular). *Una función medible $U : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ se dice que tiene variación regular (que es de variación regular) al ∞ con índice $\rho \in \mathbb{R}$, lo cual denotaremos por $U \in RV_\rho$, si para $x > 0$*

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho. \quad (\text{E.1})$$

Si $\rho = 0$ llamaremos a U de variación lenta. Se denotará normalmente a las funciones de variación lenta como $L(x)$.

Algunos ejemplos de variación regular son x^ρ , $x^\rho \ln(1+x)$, $(x \ln(1+x))^\rho$ y $x^\rho \ln(\ln(e+x))$.

Definición E.0.2. *Una v.a. positiva X y su distribución se dice que son de variación regular con índice $\rho \geq 0$ si la cola de su distribución se puede representar como*

$$\mathbb{P}[X > x] = L(x)x^{-\rho} \quad \text{para } x > 0 \quad (\text{E.2})$$

para $L(x) \in VR_0$.

Observación E.0.2. Si $U \in RV_\rho$, entonces $U(x)/x^\rho \in RV_0$ y haciendo $L(x) = U(x)/x^\rho$ se puede observar que siempre es posible representar una función ρ -variable utilizando una función de variación lenta como $x^\rho L(x)$, lo cual permite que la definición anterior implique que la cola de la distribución es de variación regular.

Ejemplos de distribuciones de variación regular (con índice positivo) utilizadas para representar las distribuciones de los reclamos son la Pareto, Burr, α -estable, log-gamma y PME. Aunque es posible representar distribuciones con colas más pesadas que las de variación regular, las funciones de variación regular resultan ser adecuadas por su sencillez matemática.

La demostración del siguiente teorema puede ser consultada en el libro de Bingham(1987)[9] en el Teorema 1.2.1.

Teorema E.0.2 (Convergencia uniforme de las funciones de variación regular). Si $U \in VR_\rho$ (si $\rho > 0$ también se necesitará que U sea acotada en intervalos de la forma $(0, x]$, para $x > 0$), entonces para $0 < a \leq b < \infty$

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho \quad (\text{E.3})$$

converge uniformemente en x sobre

1. $[a, b]$ si $\rho = 0$
2. $(0, b]$ si $\rho > 0$
3. $[a, \infty)$ si $\rho < 0$.

Los siguientes resultados exhiben diversas propiedades de las funciones de variación regular. Para el caso de la integración las funciones de variación regular tienen un comportamiento parecido a las funciones x^ρ . Asumiremos que todas las funciones serán localmente acotadas y también que son integrables en intervalos que contengan al 0 para propósitos de observar su comportamiento en ∞ .

Teorema E.0.3 (Teorema de Karamata). Supongamos que $\rho \geq -1$ y $U \in RV_\rho$. Entonces $\int_0^x U(t)dt \in RV_{\rho+1}$ y

$$\lim_{x \rightarrow \infty} \frac{xU(x)}{\int_0^x U(t)dt} = \rho + 1. \quad (\text{E.4})$$

Si $\rho < -1$ (o si $\rho = -1$ y $\int_x^\infty U(t)dt < \infty$) entonces $U \in RV_\rho$ implica que $\int_x^\infty U(t)dt$ es finita, $\int_x^\infty U(t)dt \in RV_{\rho+1}$ y

$$\lim_{x \rightarrow \infty} \frac{xU(x)}{\int_x^\infty U(t)dt} = -\rho - 1. \quad (\text{E.5})$$

Demostración. Si $\rho > -1$ demostraremos que $\int_0^\infty U(t)dt = \infty$, para así al quitar una vecindad finita a esta integral se siga teniendo el mismo lím sup. Para $U \in RV_\rho$ tenemos

$$\lim_{s \rightarrow \infty} \frac{U(2s)}{U(s)} = 2^\rho > 2^{-1}$$

porque $\rho > -1$. Entonces existe s_0 tal que para $s > s_0$ se tiene que $U(2s) > 2^{-1}U(s)$. Lo cual nos permite que para n que cumpla con $2^n > s_0$ tengamos que

$$\int_{2^{n+1}}^{2^{n+2}} U(s)ds = 2 \int_{2^n}^{2^{n+1}} U(2s)ds > \int_{2^n}^{2^{n+1}} U(s)ds,$$

lo cual muestra que al incrementar el índice que tiene los límites de integración también incrementa la integral. Haciendo $n_0 = \inf\{n : 2^n > s_0\}$ obtenemos

$$\int_{s_0}^\infty U(s)ds \geq \sum_{n: 2^n > s_0} \int_{2^{n+1}}^{2^{n+2}} U(s)ds > \sum_{n \geq n_0} \int_{2^{n_0+1}}^{2^{n_0+2}} U(s)ds = \infty.$$

Por lo tanto para $\rho > -1$, $x > 0$ y para cualquier $N < \infty$ (utilizando el supuesto de que U es integrable sobre intervalos finitos alrededor del origen) se tiene que

$$\int_0^t U(sx)ds \sim \int_N^t U(sx)ds \quad \text{cuando } t \rightarrow \infty.$$

Con lo cual hemos demostrado que la integral diverge. Por otro lado, por ser U de variación regular tenemos que para una x fija y dada $\varepsilon > 0$ existe N tal que para $s > N$

$$(1 - \varepsilon)x^\rho U(s) \leq U(sx) \leq (1 + \varepsilon)x^\rho U(s),$$

y entonces

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{\int_0^{tx} U(s)ds}{\int_0^t U(s)ds} &= \limsup_{t \rightarrow \infty} \frac{x \int_0^t U(sx)ds}{\int_0^t U(s)ds} \\ &= \limsup_{t \rightarrow \infty} \frac{x \int_N^t U(sx)ds}{\int_N^t U(s)ds} \\ &\leq \limsup_{t \rightarrow \infty} x^{\rho+1} (1 + \varepsilon) \frac{\int_N^t U(s)ds}{\int_N^t U(s)ds} \\ &= x^{\rho+1} (1 + \varepsilon). \end{aligned}$$

Un argumento análogo se utiliza para probar que

$$\liminf_{t \rightarrow \infty} \frac{\int_0^{tx} U(s) ds}{\int_0^t U(s) ds} \geq x^{\rho+1}(1 - \varepsilon).$$

Por lo tanto hemos demostrado que

$$\int_0^x U(s) ds \in RV_{\rho+1}$$

cuando $\rho > -1$.

En el caso de que $\rho = -1$ si $\int_0^\infty U(s) ds < \infty$ en cuyo caso $\int_0^x U(s) ds \in RV_{-1+1} = RV_0$ (ya que se puede aplicar la definición y obtener el cociente de los límites). Si ocurriera que $\rho = -1$ y $\int_0^\infty U(s) ds = \infty$ en el cual se puede repetir el argumento anterior. Por lo tanto hemos demostrado para $\rho \geq -1$ que $\int_0^x U(s) ds \in RV_{\rho+1}$.

Ahora probaremos la ecuación (E.4) cuando $U \in RV_\rho$, $\rho \geq -1$. Definimos la función

$$b(x) := xU(x) / \int_0^x U(t) dt, \quad (\text{E.6})$$

entonces necesitaremos probar que $b(x) \rightarrow \rho + 1$. Al integrar $b(x)/x$ de 1 a x obtenemos que

$$\int_1^x \frac{b(s)}{s} ds = \int_1^x \frac{U(s)}{\int_0^s U(t) dt} ds = \ln \int_0^x U(t) dt + c \quad (\text{E.7})$$

por lo cual

$$\int_0^x U(t) dt = c \exp \left\{ \int_1^x t^{-1} b(t) dt \right\} \quad (\text{E.8})$$

y sustituyendo en la ecuación (E.6) se tiene que

$$U(x) = cx^{-1} b(x) \exp \left\{ \int_1^x t^{-1} b(t) dt \right\}. \quad (\text{E.9})$$

Por otro lado se tiene

$$\begin{aligned} \liminf_{x \rightarrow \infty} \frac{1}{b(x)} &= \liminf_{x \rightarrow \infty} \frac{\int_0^x U(t) dt}{xU(x)} \\ &= \liminf_{x \rightarrow \infty} \int_0^1 \frac{U(sx)}{U(x)} ds, \end{aligned}$$

donde utilizamos el cambio de variable $s = x^{-1}t$. Haciendo uso del Lema de Fatou se tiene que

$$\begin{aligned} &\geq \int_0^1 \liminf_{x \rightarrow \infty} \frac{U(sx)}{U(x)} ds \\ &= \int_0^1 s^\rho ds \\ &= \frac{1}{\rho + 1}, \end{aligned}$$

y podemos concluir que

$$\limsup_{x \rightarrow \infty} b(x) \leq \rho + 1. \quad (\text{E.10})$$

Para el caso $\rho = -1$ obtenemos que $b(x) \rightarrow 0$ cuando $x \rightarrow \infty$. Supongamos que $\rho > -1$, entonces tenemos las siguientes propiedades de $b(x)$:

1. $b(x)$ es acotada en intervalos semi-infinitos de ∞^1 (por E.10)
2. b es de variación lenta, porque $xU(x) \in RV_{\rho+1}$ y $\int_0^x U(s)ds \in RV_{\rho+1}$ lo cual conduce a

$$\lim_{t \rightarrow \infty} \frac{b(tx)}{b(t)} = \lim_{t \rightarrow \infty} \frac{txU(tx)}{tU(t)} \cdot \lim_{t \rightarrow \infty} \frac{\int_0^{tx} U(s)ds}{\int_0^t U(s)ds} = x^{\rho+1} \cdot x^{-\rho-1} = 1$$

3. Se tiene que

$$b(xt) - b(x) \rightarrow 0 \quad \text{cuando} \quad x \rightarrow \infty$$

porque las funciones de variación lenta convergen uniformemente en t para intervalos de la forma $[a, b]$ y haciendo uso del siguiente límite al notar que su denominador es acotado para valores grandes de x

$$\lim_{x \rightarrow \infty} \frac{b(xt) - b(x)}{b(x)} = 0.$$

Utilizando el punto 3 y el teorema de convergencia dominada tenemos que

$$\lim_{x \rightarrow \infty} \int_1^x t^{-1}(b(xt) - b(x))dt = 0,$$

¹Por intervalos semi-infinitos nos referimos a intervalos de la forma $[a, \infty)$

donde el lado izquierdo de la última ecuación se puede reescribir como

$$\lim_{x \rightarrow \infty} \left\{ \int_1^s t^{-1} b(xt) dt - b(x) \ln s \right\} = 0. \quad (\text{E.11})$$

Por (E.9) se tiene que

$$c \exp \left\{ \int_1^x t^{-1} b(t) dt \right\} = \int_0^x U(s) ds \in RV_{\rho+1}$$

lo cual por ser de variación regular y usando (E.7) implica que

$$\begin{aligned} (\rho + 1) \ln s &= \lim_{x \rightarrow \infty} \ln \left(\frac{\int_0^{xs} U(t) dt}{\int_0^x U(t) dt} \right) \\ &= \lim_{x \rightarrow \infty} \left\{ \ln \int_0^{xs} U(t) dt - \ln \int_0^x U(t) dt \right\} \\ &= \lim_{x \rightarrow \infty} \left\{ \int_1^{xs} t^{-1} b(t) dt - c - \int_1^x t^{-1} b(t) dt + c \right\} \\ &= \lim_{x \rightarrow \infty} \int_x^{xs} t^{-1} b(t) dt \\ &= \lim_{x \rightarrow \infty} \int_1^s t^{-1} b(xt) dt. \end{aligned}$$

Utilizando esta última ecuación junto con (E.11) se tiene que

$$\lim_{x \rightarrow \infty} \{(\rho + 1) \ln s - b(x) \ln s\} = 0$$

lo cual implica que

$$\lim_{x \rightarrow \infty} b(x) = \rho + 1$$

con lo cual termina la demostración. \square

El siguiente corolario es muy importante ya que permite una identificación de las funciones de variación regular.

Corolario E.0.3.1.

i) *La función L es de variación acotada si solo si L puede ser representada como*

$$L(x) = c(x) \exp \left\{ \int_1^x t^{-1} \varepsilon(t) dt \right\} \quad \text{para } x > 0 \quad (\text{E.12})$$

donde $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, $\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ y

$$\lim_{x \rightarrow \infty} c(x) = c \in (0, \infty) \quad (\text{E.13})$$

$$\lim_{x \rightarrow \infty} \varepsilon(x) = 0 \quad (\text{E.14})$$

ii) Una función $U : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ es de variación regular con índice ρ si y solo si U tiene la representación

$$U(x) = c(x) \exp \left\{ \int_1^x t^{-1} \rho(t) dt \right\} \quad \text{para } x > 0 \quad (\text{E.15})$$

donde $c(\cdot)$ satisface (E.13) y $\lim_{t \rightarrow \infty} \rho(t) = \rho$.

Demostración.

i. Si L se puede representar por medio de (E.12) entonces demostraremos que es de variación lenta ya que para $x > 1$

$$\lim_{t \rightarrow \infty} L(tx)/L(t) = \lim_{t \rightarrow \infty} (c(tx)/c(t)) \exp \left\{ \int_t^{tx} s^{-1} \varepsilon(x) dx \right\},$$

dada ε , existe t_0 tal que

$$-\varepsilon < \varepsilon(t) < \varepsilon \quad \text{para } t \geq t_0.$$

Se sigue que

$$-\varepsilon \ln x = -\varepsilon \int_t^{tx} s^{-1} ds \leq \int_t^{tx} s^{-1} \varepsilon(s) ds \leq \varepsilon \int_t^{tx} s^{-1} ds = \varepsilon \ln x$$

lo cual implica que $\lim_{t \rightarrow \infty} \int_t^{tx} s^{-1} \varepsilon(s) ds = 0$ y con esto se logra probar que $\lim_{t \rightarrow \infty} L(tx)/L(t) = 1$.

Ahora demostraremos la otra implicación; supongamos que $L \in RV_0$. Definimos (al igual que en (E.6))

$$b(x) := xL(x) / \int_0^x L(s) ds$$

y utilizando el teorema de Karamata E.0.3 tenemos que $b(x) \rightarrow 1$ cuando $x \rightarrow \infty$. Observando que

$$L(x) = x^{-1} b(x) \int_0^x L(s) ds$$

tenemos que al hacer $\varepsilon(x) = b(x) - 1$ entonces $\varepsilon(x) \rightarrow 0$ cuando $x \rightarrow \infty$ lo cual implica

$$\begin{aligned} \int_1^x t^{-1} \varepsilon(t) dt &= \int_1^x \left(L(t) / \int_0^t L(s) ds \right) dt - \ln x \\ &= \int_1^x d \left(\ln \int_0^t L(s) ds \right) dt - \ln x \\ &= \ln \left(x^{-1} \int_0^x L(s) ds / \int_0^1 L(s) ds \right). \end{aligned}$$

Por lo tanto

$$\begin{aligned} \exp \left\{ \int_1^x t^{-1} \varepsilon(t) dt \right\} &= x^{-1} \int_0^x L(s) ds / \int_0^1 L(s) ds \\ &= L(x) / \left(b(x) \int_0^1 L(s) ds \right). \end{aligned}$$

Se obtiene la representación deseada utilizando

$$c(x) = b(x) \int_0^1 L(s) ds.$$

ii. Usando que si $U \in RV_\rho$ entonces $U(x) = x^\rho L(x)$ para alguna $L \in RV_0$ se tendrá que

$$\begin{aligned} U(x) &= x^\rho L(x) \\ &= \exp \left\{ \int_1^x \rho t^{-1} dt \right\} c(x) \exp \left\{ \int_1^x t^{-1} \varepsilon(t) dt \right\} \\ &= c(x) \exp \left\{ \int_1^x t^{-1} (\varepsilon(t) + \rho) dt \right\} \end{aligned}$$

Se obtiene la representación (E.15) utilizando

$$\rho(t) = \varepsilon(t) + \rho.$$

□

Corolario E.0.3.2. *Si U satisface*

$$\lim_{x \rightarrow \infty} \frac{xU(x)}{\int_0^x U(t) dt} = \lambda \in (0, \infty), \quad (\text{E.16})$$

entonces $U \in RV_{\lambda-1}$.

Si $\int_x^\infty U(t)dt < \infty$ y

$$\lim_{x \rightarrow \infty} \frac{xU(x)}{\int_x^\infty U(t)dt} = \lambda \in (0, \infty), \quad (\text{E.17})$$

entonces $U \in RV_{-\lambda-1}$.

Demostración. Supongamos que U satisface (E.16) queremos probar que $U \in RV_{\lambda-1}$. Sea

$$b(x) = xU(x) / \int_0^x U(t)dt,$$

entonces $b(x) \rightarrow \lambda$ cuando $x \rightarrow \infty$. De (E.9) tenemos que

$$\begin{aligned} U(x) &= cx^{-1}b(x) \exp \left\{ \int_1^x t^{-1}b(t)dt \right\} \\ &= cb(x) \exp \left\{ \int_1^x t^{-1}(b(t) - 1)dt \right\}, \end{aligned}$$

como $b(t) - 1 \rightarrow \lambda - 1$, U satisface ser $(\lambda - 1)$ -variable utilizando el Corolario E.0.3.1, lo cual concluye la prueba. \square

El Teorema E.0.3 muestra que para propósitos de integración el comportamiento de las funciones de variación regular se asemeja a que fueran extraídas del interior de la integral, ya que al considerar

$$\int_0^x U(t)dt = \int_0^x t^\rho L(t)dt;$$

ahora sacando $L(t)$ de la integral como si fuera un factor $L(x)$ obtenemos

$$\begin{aligned} &\sim L(x) \int_0^x t^\rho dt = L(x)x^{\rho+1}/(\rho+1) \\ &= xx^\rho L(x)/(\rho+1) = xU(x)/(\rho+1), \end{aligned}$$

lo cual es equivalente a (E.4).

Para concluir esta sección demostraremos el siguiente corolario el cual nos será útil para el planteamiento de los algoritmos de simulación. Este nos habla sobre el comportamiento límite de las funciones de variación regular, el cual vuelve a ser igual que el de las funciones del tipo x^ρ .

Corolario E.0.3.3. Si $U \in VR_\rho$ para algún $\rho \neq 0$ entonces

$$\lim_{x \rightarrow \infty} U(x) = \begin{cases} \infty & \text{si } \rho > 0 \\ 0 & \text{si } \rho < 0 \end{cases} \quad (\text{E.18})$$

Demostración. Usando el corolario anterior tenemos que

$$U(x) = c(x) \exp \left\{ \int_1^x t^{-1} \rho(t) dt \right\} \quad \text{para } x > 0$$

y utilizando que $\rho(t) \rightarrow \rho$ cuando $t \rightarrow \infty$.

Caso 1: $\rho > 0$

Si tomamos $\varepsilon = \frac{\rho}{2}$ entonces existe un t_0 tal que para $t > t_0$, $\rho(t) > \varepsilon$.
Lo anterior implica que

$$\begin{aligned} \lim_{x \rightarrow \infty} \int_1^x t^{-1} \rho(t) dt &= \lim_{x \rightarrow \infty} \left\{ \int_1^{t_0} t^{-1} \rho(t) dt + \int_{t_0}^x t^{-1} \rho(t) dt \right\} \\ &> \int_1^{t_0} t^{-1} \rho(t) dt + \varepsilon \lim_{x \rightarrow \infty} \int_{t_0}^x t^{-1} dt \\ &= \infty \end{aligned}$$

Por lo tanto

$$\lim_{x \rightarrow \infty} U(x) = \infty.$$

Caso 2: $\rho < 0$

Si tomamos $\varepsilon = \frac{\rho}{2}$ entonces existe un t_0 tal que para $t > t_0$, $\rho(t) < \varepsilon$.
Entonces

$$\begin{aligned} \lim_{x \rightarrow \infty} \int_1^x t^{-1} \rho(t) dt &= \lim_{x \rightarrow \infty} \left\{ \int_1^{t_0} t^{-1} \rho(t) dt + \int_{t_0}^x t^{-1} \rho(t) dt \right\} \\ &< \int_1^{t_0} t^{-1} \rho(t) dt + \varepsilon \lim_{x \rightarrow \infty} \int_{t_0}^x t^{-1} dt \\ &= -\infty \end{aligned}$$

Por lo tanto

$$\lim_{x \rightarrow \infty} U(x) = 0.$$

□

Observación E.0.3. *Un ejemplo trivial de que en el caso $\rho = 0$ no se tiene conocimiento del límite de $U(t)$ es cuando se toman las funciones constantes, ya que para estas su límite son ellas mismas y por tanto no se puede fijar un criterio general para el límite.*

E.1. Mezcla Pareto de distribuciones exponenciales

La Mezcla Pareto de distribuciones exponenciales se genera por medio de hacer una transformación a una distribución Pareto. Sea $X \sim \text{Pareto}(\frac{r-1}{r}, r)$ entonces X tiene densidad

$$f_X(x) = r \left(\frac{r-1}{r} \right)^r x^{-(r+1)} \quad \text{para} \quad x \geq \frac{r-1}{r}, \quad (\text{E.19})$$

la cual tiene media 1. Entonces la densidad de una Mezcla Pareto de distribuciones exponenciales es definida por

$$f_{MPE}(x) = \int_{\frac{r-1}{r}}^{\infty} f_X(x) \frac{1}{y} e^{-\frac{x}{y}} dy \quad \text{para} \quad x \geq 0. \quad (\text{E.20})$$

Al resolver la integral se obtiene que

$$f_{MPE}(x) = r \left(\frac{r-1}{r} \right)^r x^{-(r+1)} \gamma \left(r+1, \frac{r}{r-1}x \right) \quad \text{para} \quad x \geq 0, \quad (\text{E.21})$$

donde $\gamma(a, u) = \int_0^u t^{a-1} e^{-t} dt$ es una función gamma incompleta. La cual sigue teniendo media 1.

Al integrar esta última ecuación de 0 a x se obtiene que su distribución es

$$F_{MPE}(x) = \left(1 - e^{-\frac{r}{r-1}x} \right) - \frac{1}{y^r} \left(\frac{r-1}{r} \right)^r \gamma \left(r+1, \frac{r}{r-1}y \right) \quad \text{para} \quad x \geq 0. \quad (\text{E.22})$$

Además la distribución de su cola integrada es

$$F_I(x) = \frac{r-1}{r} (1 - e^{-\frac{r}{r-1}x}) + \left(\frac{r-1}{r} \right)^r \frac{1}{r-1} x^{1-r} \gamma \left(r+1, \frac{r}{r-1}x \right) + \frac{1}{r} \gamma \left(2, \frac{r}{r-1}x \right). \quad (\text{E.23})$$

Apéndice F

Verificación del método de las variables antitéticas al estimar el valor esperado de funciones monótonas

El siguiente teorema junto con su corolario muestra cómo en el caso de variables aleatorias que sean funciones monótonas de $U(0, 1)$ se tendrá una reducción de varianza a la par de no necesitar generar un nuevo conjunto de valores para la siguiente simulación.

Teorema F.0.1. *Si X_1, \dots, X_n son independientes, entonces para cualesquiera funciones crecientes f y g de n variables*

$$\mathbb{E}[f(\mathbf{X})g(\mathbf{X})] \geq \mathbb{E}[f(\mathbf{X})]\mathbb{E}[g(\mathbf{X})] \quad (\text{F.1})$$

donde $\mathbf{X} = (X_1, \dots, X_n)$.

Demostración. La demostración es por inducción sobre n . Para $n = 1$ tenemos que f y g son funciones de una sola variable. Entonces, para cualesquiera x y y ,

$$[f(x) - f(y)][g(x) - g(y)] \geq 0.$$

Por lo tanto, para cualesquiera variables aleatorias X y Y ,

$$[f(X) - f(Y)][g(X) - g(Y)] \geq 0.$$

Lo cual implica que

$$\mathbb{E}[f(X) - f(Y)][g(X) - g(Y)] \geq 0.$$

que es equivalente a

$$\mathbb{E}[f(X)g(X)] + \mathbb{E}[f(Y)g(Y)] \leq \mathbb{E}[f(X)g(Y)] + \mathbb{E}[f(Y)g(X)]$$

Si utilizamos la hipótesis de que X y Y son v.a.i.i.d. tenemos que

$$\begin{aligned}\mathbb{E}[f(X)g(X)] &= \mathbb{E}[f(Y)g(Y)] \\ \mathbb{E}[f(X)g(Y)] &= \mathbb{E}[f(Y)g(X)] = \mathbb{E}[f(X)] \mathbb{E}[g(X)]\end{aligned}$$

lo cual muestra el resultado para $n = 1$. Supongamos que la ecuación (F.1) se cumple para $n - 1$ variables, y ahora supongamos que X_1, \dots, X_n son independientes y que f y g son funciones crecientes. Entonces

$$\begin{aligned}\mathbb{E}[f(\mathbf{X})g(\mathbf{X}) \mid X_n = x_n] &= \mathbb{E}[f(X_1, \dots, X_{n-1}, x_n)g(X_1, \dots, X_{n-1}, x_n) \mid X_n = x_n] \\ &= \mathbb{E}[f(X_1, \dots, X_{n-1}, x_n)g(X_1, \dots, X_{n-1}, x_n)] \\ &\geq \mathbb{E}[f(X_1, \dots, X_{n-1}, x_n)] \mathbb{E}[g(X_1, \dots, X_{n-1}, x_n)] \\ &= \mathbb{E}[f(\mathbf{X}) \mid X_n = x_n] \mathbb{E}[g(\mathbf{X}) \mid X_n = x_n]\end{aligned}$$

donde en la desigualdad utilizamos el hecho de la hipótesis de inducción. Por lo tanto

$$\mathbb{E}[f(\mathbf{X})g(\mathbf{X}) \mid X_n] \geq \mathbb{E}[f(\mathbf{X}) \mid X_n] \mathbb{E}[g(\mathbf{X}) \mid X_n]$$

al obtener esperanzas de ambos lados de esta desigualdad

$$\begin{aligned}\mathbb{E}[f(\mathbf{X})g(\mathbf{X})] &\geq \mathbb{E}[\mathbb{E}[f(\mathbf{X}) \mid X_n] \mathbb{E}[g(\mathbf{X}) \mid X_n]] \\ &\geq \mathbb{E}[\mathbb{E}[f(\mathbf{X}) \mid X_n]] \mathbb{E}[\mathbb{E}[g(\mathbf{X}) \mid X_n]] \\ &\geq \mathbb{E}[f(\mathbf{X})] \mathbb{E}[g(\mathbf{X})].\end{aligned}$$

En la penúltima desigualdad se utilizó el hecho de que $\mathbb{E}[f(\mathbf{X}) \mid X_n]$ y $\mathbb{E}[g(\mathbf{X}) \mid X_n]$ son ambas funciones crecientes de X_n y el resultado para $n = 1$ \square

Corolario F.0.1.1. Si $h(x_1, \dots, x_n)$ es una función monótona de cada uno de sus argumentos, entonces para un conjunto U_1, \dots, U_n de números aleatorios independientes se tiene que

$$\text{Cov}[h(U_1, \dots, U_n), h(1 - U_1, \dots, 1 - U_n)] \leq 0$$

Demostración. Sin pérdida de generalidad supongamos que h es creciente en los primeros r argumentos y decreciente en los últimos $n - r$. Por lo tanto, si

$$\begin{aligned}f(x_1, \dots, x_n) &= h(x_1, \dots, x_r, 1 - x_{r+1}, \dots, 1 - x_n) \\ g(x_1, \dots, x_n) &= -h(1 - x_1, \dots, 1 - x_r, x_{r+1}, \dots, x_n)\end{aligned}$$

entonces f y g son funciones crecientes. Utilizando el teorema anterior

$$\text{Cov}[f(U_1, \dots, U_n), g(U_1, \dots, U_n)] \geq 0$$

lo cual equivale a

$$\text{Cov}[h(U_1, \dots, U_r, 1 - U_{r+1}, \dots, 1 - U_n), h(1 - U_1, \dots, 1 - U_r, U_{r+1}, \dots, U_n)] \leq 0$$

. Como se tiene que $(h(U_1, \dots, U_n), h(1 - U_1, \dots, 1 - U_n))$ tiene la misma distribución conjunta que el vector aleatorio $(h(U_1, \dots, U_r, 1 - U_{r+1}, \dots, 1 - U_n), h(1 - U_1, \dots, 1 - U_r, U_{r+1}, \dots, U_n))$, se concluye la prueba. \square

Apéndice G

Programas utilizados

G.1. Lista de programas

Los siguientes fueron todos los programas utilizados durante la simulación

- phiexp
- media
- denpareto
- denparetovarias
- gamma_in
- pF I
- tope
- encuentra
- simula
- rsim
- Algoritmos I, II y III
- Algoritmos rapidos Alg.I, II y III
- sumgeom
- Algoritmos
- f.disc-lo,up,pa
- noruin.aprox
- ruin.aprox
- noruin.aprox.(lo/up)
- ruin.aprox.(lo/up)
- ruina_aprox_asintot
- sim_3geom_con_Panjer
- sim_1geom_con_Panjer
- sim_1geom_con_asintota
- imprime_de_totales1

- `imprime_de_totales2`
- `imprime_de_totales3`
- `dibuja_algoritmos1`
- `dibuja_algoritmos2`
- `dibuja_algoritmos3`

Librerías auxiliares

```
library(grDevices)
library(MASS)
library(stats)
library(evd)
library(lmomco)
```

G.2. Significado de parámetros

x: Capital inicial.

df: Dictamina que distribución se está utilizando

df=1 Distribución exponencial.

df=2 Distribución Log-normal.

df=3 Distribución Pareto.

df=4 Distribución Weibull.

df=5 Distribución MPE.

par1 y par2: Son los parámetros que utilizará la distribución

- Exponencial ($\alpha=\text{par1}$).
- Log-normal ($\mu=\text{par1}, \sigma^2=\text{par2}$).
- Pareto ($a=\text{par1}, b=\text{par2}$).
- Weibull ($r=\text{par1}, \lambda=\text{par2}$).
- MPE ($r=\text{par1}$).

R: Número de veces que se ejecuta el algoritmo.

u: Resultado de una simulación de una variable $\mathcal{U}(0, 1)$.

mu: Esperanza de la distribución.

eps: Error en precisión de la inversa.

n: Número de simulaciones necesarias.

rho: Parámetro de la distribución geométrica ($\text{geom}(\text{rho})$, ver Sección 1.4).

Z: Vector de resultados de las simulaciones.

c: Parámetro c del modelo (ver Sección 1.2).

lambda: Media de Proceso Poisson (ver Sección 1.2).

x.inicial, lar, veces y valores: Parámetros que dictaminan a qué capitales iniciales calcular la probabilidad de ruina; $x.inicial$ = Primer valor a calcular, lar = Distancia entre el primer valor y el último, $veces$ = Cuantos puntos intermedios tomar, $valores$ = Vector que contiene los capitales iniciales.

A, B, D, E, F, G y H Resultados obtenidos por

- A = Simulación de Algoritmo I.
- B = Simulación de Algoritmo II.
- D = Simulación de Algoritmo III.
- E = Aproximación de Panjer.
- F = Cota inferior de Panjer.
- G = Cota superior de Panjer.
- H = Aproximación asintótica.

G.3. Programas

Programas auxiliares

P. media: Realiza el cálculo de la esperanza.

```
media<-function(par1,par2,df){ #par1,par2=parámetros de la distribución
  if(df==1){return(par1)}#Exponencial
  if(df==2){return(exp(par1+par2^2/2))}#Log-normal
  if(df==3){#Pareto
    if(par1<0||par2<1){
      return("parámetros incorrectos para Pareto")}
    return(par1*par2/(par2-1))
  }
}
```

```

if(df==4){#Weibull
  if(par1<0|| par2<0){
    return("Parámetros incorrectos para Weibull")}
  return(par2*gamma(1+1/par1))
}
if(df==5){#MPE (Mezcla Pareto de distribuciones exponenciales)
  if(par1<1){return("parámetros incorrectos para MPE")}
  return(1)
}
}

```

P. gamma_in: Realiza el cálculo de la función gama inversa.

```

gamma_in<-function(x,par1){
  return(pgamma(x*par1/(par1-1),par1+1,1)*gamma(par1))
}

```

P. tope: Regresa el primer intervalo donde se puede encontrar la función inversa; esta búsqueda la realiza por medio de pasos de tamaño mu.

```

tope<-function(u,par1,par2,df,mu){
  #Encuentra el valor i para el cual F_I((i-1)*mu)<=u<F_I(i*mu)
  if(u>1 || u<0){return("u no se encuentra en (0,1)")}
  i=1
  while(u>pF_I(mu*i,par1,par2,df,mu)){i=i+1}
  return(i)
}

```

P. encuentra: Encuentra el valor de la inversa con un error de eps.

```

encuentra<-function(u,a,b,eps,par1,par2,df,mu){
  ### Valor aproximado de la función inversa en el punto u ###
  if(b-a<eps){return((a+b)/2)}
  if(b-a>=eps){
    if(u>=pF_I((a+b)/2,par1,par2,df,mu)){
      return(encuentra(u,(a+b)/2,b,eps,par1,par2,df,mu))}
    if(u<pF_I((a+b)/2,par1,par2,df,mu)){
      return(encuentra(u,a,(a+b)/2,eps,par1,par2,df,mu))}
  }
}

```

P. phiexp: Realiza el cálculo de probabilidad ruina en el caso exponencial.

```

phiexp<-function(x,c,lambda,par1){
  if(lambda*par1>=c){
    return("Condicion de ganancia neta no se satisface")}
  return(lambda*par1*exp(-(1/par1-lambda/c)*x)/c)
}

```

P. estim_pi_1: Realiza una aproximación de π utilizando muestreo estratificado.

```
estim_pi_1<-function(n){
  est=rep(0,n)
  for(i in 1:n){est[i]=sqrt(1-(runif(1)+i-1)^2/(n^2))}
  return(4*sum(est)/n)
}
```

P. estim_pi_2: Realiza una aproximación de π utilizando muestreo estratificado y variables antitéticas.

```
estim_pi_2<-function(n){
  est=rep(0,n)
  for(i in 1:n){est[i]=sqrt(1-(runif(1)+i-1)^2/(n^2))+sqrt(1-(i-runif(1))^2/(n^2))}
  return(2*sum(est)/n)
}
```

Programas de distribuciones

P. denpareto: Realiza el cálculo de la densidad Pareto para un real.

```
denpareto<-function(x,par1,par2){
  if(par1<=0){
    return("Parametros incorrectos para distribucion Pareto, par1<=0")}
  if(par2<=1){
    return("Parametros incorrectos para distribucion Pareto, par2<=1")}
  }
  if(x<=par1){return(0)}
  if(x>par1){return((par1^(par2))*par2/x^(par2+1))}
}
```

P. denparetovarias: Realiza el cálculo de la densidad Pareto para cada entrada de un vector.

```
denparetovarias<-function(x,par1,par2){
  y=0
  for(i in 1:length(x)){y[i]=denpareto(x[i],par1,par2)}
  return(y)
}
```

P. pF_I: Realiza el cálculo de la distribución de la cola integrada.

```
pF_I<-function(x,par1,par2,df,mu){
  if(x<=0){return(0)}
  if(df==1){return(pexp(x,1/par1))}#Exponencial
```

```

if(df==2){#Log-normal
  f=as.function(alist(x=,a=par1,b=par2,plnorm(x,meanlog=a,sdlog=b,lower.tail=F)))
  return(integrate(f,0,x)$value/mu)
}
if(df==3){#Pareto
  if(par1<=0 || par2<=1){return("parámetros incorrectos para distribucion Pareto")}
  if(x<par1){return(x/mu)}
  if(x>=par1){return(1-(((par1/x)^(par2-1))/par2))}
}
if(df==4){#Weibull
  f=as.function(alist(x=,a=par1,b=par2,pweibull(x,shape=a,scale=b,lower.tail=F)))
  return(integrate(f,0,x)$value/mu)
}
if(df==5){#MPE
  if(par1<1){print("Parámetro incorrecto de MPE")}
  z=par1/(par1-1)
  return((1-exp(-z*x))/z+(z^(-par1))*(x^(1-par1))*
    gamma_in(x*z,(par1+1))/(1-par1)+gamma_in(x*z,2)/par1)
}
}

```

Programas de simulaciones

P. reserva: Realiza una simulación de reserva donde ocurren n siniestros

```

reserva <- function(c,x0,mu,lamda,n,reclamos){
  t=matrix(0,ncol=reclamos,nrow=n)      ###cuenta tiempos
  truina=rep(0,n)                        ###momentos de ruina
  R=matrix(0,ncol=reclamos,nrow=n)     ###reserva despues de reclamo
  R[,1]=x0
  t[,1]=0
  lamda=1/lamda
  for(j in 1:n){
    for(i in 2:reclamos){
      sig=rexp(1,mu)
      t[j,i]=t[j,i-1]+sig
      R[j,i]=R[j,i-1]+c*sig-rexp(1,lamda)
    }
    if(min(R[j,])<0){truina[j]=1 }
  }
  return(list(R=R,t=t,truina=truina))
}

```

P. simula: Simula una v.a. con distribución F_I .

```

simula<-function(eps,par1,par2,df,mu){
  ##simula con distribución F_I

```

```

if(df==1){ #Exponenciales
  return(rexp(1,1/par1))
}
u=runif(1)      #simula una uniforme (0,1)
if(df==3){     #Pareto
  if(par1<0 || par2<1){
    return("parámetros incorrectos para distribucion Pareto")}
  if(u<=(par2-1)/par2){return(par1*par2*u/(par2-1))}
  if(u>(par2-1)/par2){return(par1*(par2*(1-u))^(1/(1-par2)))}
}
if(df!=3 && df!=1){ #para que no simule pareto ni expo
  i=tope(u,par1,par2,df,mu)
  #tope da cuantas medias se necesitan para pasarse del valor u
  a=mu*(i-1)      #extremo inferior del intervalo
  b=mu*(i)        #extremo superior del intervalo
  return(encuentra(u,a,b,eps,par1,par2,df,mu)) #encuentra la  $F^{-1}(u)$ 
}
}

```

P. rsim: `rsim<-function(n,eps,par1,par2,df,mu){`
 ##simula n veces F_I
 if(n<0){return("numero de simulaciones incorrectas")}
 if(n==0){return(0)}
 a=rep(0,n)
 for(i in 1:n){a[i]=simula(eps,par1,par2,df,mu)}
 return(a)
}

P. sumgeom: `sumgeom<-function(R,eps,rho,par1,par2,df,mu){`
 ### Simula una suma geométrica
 M=rgeom(R,(rho/(1+rho)))
 X=matrix(0,nrow=R,ncol=max(M))
 for(i in 1:R){X[i,1:M[i]]=rsim(M[i],eps,par1,par2,df,mu)}
 return(list(X=X,M=M))
}

Programas de algoritmos (versiones totales)

P. AlgoritmoI: Realiza el Algoritmo I conservando todos los valores que fué obteniendo.

```

AlgoritmoI<-function(x,R,eps,rho,par1,par2,df,mu){
  Z=rep(0,R)
  X=rep(0,R)
  M=rgeom(R,(rho/(1+rho)))
  for(j in 1:R){
    X[j]=sum(rsim(M[j],eps,par1,par2,df,mu))
  }
}

```

```

        if(X[j]>x){Z[j]=1}
      }
      return(matrix(c(M,X,Z),nrow=3,dimnames= list(c("M","X","Z"),1:R),
        byrow=TRUE))
    }

```

P. AlgoritmoII Realiza el Algoritmo II conservando todos los valores que fué obteniendo.

```

AlgoritmoII<-function(x,R,eps,rho,par1,par2,df,mu){
  Z=rep(0,R)
  X=rep(0,R)
  W=rep(0,R)
  M=rgeom(R,(rho/(1+rho)))
  for(j in 1:R){
    if(M[j]>1){X[j]=sum(rsim((M[j]-1),eps,par1,par2,df,mu))}
    W[j]=x-X[j]
    if(W[j]<=0){Z[j]=1}
    if(W[j]>0 && M[j]>0){Z[j]=1-pF_I(W[j],par1,par2,df,mu)}
  }
  return(matrix(c(M,X,W,Z),nrow=4,dimnames = list(c("M","X","W","Z"),1:R),
    byrow=TRUE))
}

```

P. AlgoritmoIII Realiza el Algoritmo III conservando todos los valores que fué obteniendo.

```

AlgoritmoIII<-function(x,R,eps,rho,par1,par2,df,mu){
  Z=rep(0,R)
  X=rep(0,R)
  W=rep(0,R)
  K=rep(0,R) #con esta variable se lleva un control de las X's
  estord=0 #los vectores que va haciendo las estadisticas de orden
  M=rgeom(R,(rho/(1+rho)))
  for(j in 1:R){
    estord=0
    if(M[j]>1){
      estord=rev(sort(rsim(M[j],eps,par1,par2,df,mu)))
      #acomoda los elementos de X en orden decreciente
      X[j]=sum(estord[-1])
      K[j]=estord[2]
      W[j]=x-X[j]
      if(W[j]<=0){Z[j]=1}
      if(W[j]>0){Z[j]=(1-pF_I(max(W[j],K[j]),
        par1,par2,df,mu))/(1-pF_I(K[j],par1,par2,df,mu))}
    }
    if(M[j]==1){
      W[j]=x
    }
  }
}

```

```

        Z[j]=1-pF_I(x,par1,par2,df,mu)
      }
      if(M[j]==0){
        W[j]=x
        Z[j]=0
      }
    }
  return(matrix(c(M,X,W,K,Z),nrow=5,
    dimnames = list(c("M","X","W","K","Z"),1:R),byrow=TRUE))
}

```

Programas de algoritmos (versiones rápidas)

P. Alg.I Realiza el Algoritmo I y solo regresa los valores que obtuvo en cada simulación.

```

Alg.I<-function(x,R,eps,rho,par1,par2,df,mu){
  Z=rep(0,R)
  M=rgeom(R,(rho/(1+rho)))
  for(j in 1:R){if(sum(rsim(M[j],eps,par1,par2,df,mu))>x){Z[j]=1}}
  return(Z)
}

```

P. Alg.II Realiza el Algoritmo II y solo regresa los valores que obtuvo en cada simulación.

```

Alg.II<-function(x,R,eps,rho,par1,par2,df,mu){
  Z=rep(0,R)
  M=rgeom(R,(rho/(1+rho)))
  for(j in 1:R){
    X=0
    if(M[j]>1){X=sum(rsim((M[j]-1),eps,par1,par2,df,mu))}
    W=x-X
    if(W<=0){Z[j]=1}
    if(W>0 && M[j]>0){Z[j]=1-pF_I(W,par1,par2,df,mu)}
  }
  return(Z)
}

```

P. Alg.III Realiza el Algoritmo III y solo regresa los valores que obtuvo en cada simulación.

```

Alg.III<-function(x,R,eps,rho,par1,par2,df,mu){
  Z=rep(0,R)
  M=rgeom(R,(rho/(1+rho)))
  for(j in 1:R){
    if(M[j]==1){Z[j]=1-pF_I(x,par1,par2,df,mu)}
  }
}

```

```

    if(M[j]>1){
      estord=rev(sort(rsim(M[j],eps,par1,par2,df,mu)))
      #acomoda los elementos de X en orden decreciente
      Z[j]=(1-pF_I(max((x-sum(estord[-1])),estord[2]),par1,par2,df,mu))/
        (1-pF_I(estord[2],par1,par2,df,mu))
    }
  }
  return(Z)
}

```

P. Algoritmos Realiza todos los Algoritmos y solo regresa los valores que obtuvo para cada simulación.

```

Algoritmos<-function(x,sgeomX,sgeomM,R,eps,rho,par1,par2,df,mu){
  AI=as.numeric(rowSums(sgeomX)>x)
  AII=rep(0,R)
  for(i in 1:R){
    if(sgeomM[i]>0){AII[i]=1-pF_I(x-sum(sgeomX[i,-1]),par1,par2,df,mu)}}
  AIII=rep(0,R)
  for(i in 1:R){
    if(sum(sgeomX[i,])==0){AIII[i]=0}
    if(sgeomX[i,1]>0 && sgeomX[i,2]==0){AIII[i]=1-pF_I(x,par1,par2,df,mu)}
    if(sgeomX[i,1]>0 && sgeomX[i,2]>0){
      estord=rev(sort(sgeomX[i,]))
      #acomoda los elementos de X en orden decreciente
      if((x-sum(estord[-1]))<=0){AIII[i]=1}
      if((x-sum(estord[-1]))>0){
        AIII[i]=(1-pF_I(max((x-sum(estord[-1])),estord[2]),
          par1,par2,df,mu))/(1-pF_I(estord[2],par1,par2,df,mu))}
    }
  }
  return(matrix(c(AI,AII,AIII),ncol=R,byrow=T))
}

```

Programas de discretización de Panjer y aproximación asintótica

P. f.disclo Realiza la discretización para la cota inferior

```

f.disclo<-function(x,par1,par2,df,mu){ #Para cota inferior
  return(pF_I(floor(x)+1,par1,par2,df,mu)-pF_I(floor(x),par1,par2,df,mu))
}

```

P. f.discup Realiza la discretización para la cota superior.

```

f.discup<-function(x,par1,par2,df,mu){ #Para cota superior
  return(pF_I(floor(x),par1,par2,df,mu)-pF_I(floor(x)-1,par1,par2,df,mu))
}

```

P. f.discpa Realiza la discretización para la aproximación de Panjer.

```
f.discpa<-function(x,par1,par2,df,mu){#Para Panjer
  return(pF_I(floor(x)+.5,par1,par2,df,mu)-pF_I(floor(x)-.5,par1,par2,df,mu))
}
```

P. noruin.aprox Realiza la aproximación de la probabilidad de no-ruina para Panjer.

```
noruin.aprox<-function(x,rho,par1,par2,df,mu){
  if(floor(x)<x){return("x no es entero (survival)")}
  a=1/(1+rho-f.discpa(0,par1,par2,df,mu))
  fi=rep(0,x)
  fi[1]=rho/(1+rho-f.discpa(0,par1,par2,df,mu))
  g=0
  for(i in 1:x){g[i]=f.discpa(i,par1,par2,df,mu)} #g[x]:=g(x)
  if(x>=2){
    for(i in 2:x){
      #fi[x+1]:=fi(x), hay que llenarlo como (fi1,fi0)->(fi2,fi1,fi0)
      fi[i]=a*g[1:(i-1)]%*%rev(fi[1:(i-1)])
    }
  }
  return(fi)
}
```

P. ruin.aprox Realiza la aproximación de Panjer.

```
ruin.aprox<-function(x,rho,par1,par2,df,mu){
  return(1-sum(noruin.aprox(x,rho,par1,par2,df,mu)))
}
```

P. noruin.aprox.lo Realiza la aproximación de la probabilidad de no-ruina para la cota inferior de Panjer.

```
noruin.aprox.lo<-function(x,rho,par1,par2,df,mu){
  if(floor(x)<x){return("x no es entero (survival)")}
  a=1/(1+rho-f.disclo(0,par1,par2,df,mu))
  fi=rep(0,x)
  fi[1]=rho/(1+rho-f.disclo(0,par1,par2,df,mu))
  g=0
  for(i in 1:x){g[i]=f.disclo(i,par1,par2,df,mu)} #g[x]:=g(x)
  if(x>=2){
    for(i in 2:x){
      #fi[x+1]:=fi(x), hay que llenarlo como (fi1,fi0)->(fi2,fi1,fi0)
      fi[i]=a*g[1:(i-1)]%*%rev(fi[1:(i-1)])
    }
  }
  return(fi)
}
```

P. ruin.aprox.lo Realiza la cota inferior de Panjer.

```
ruin.aprox.lo<-function(x,rho,par1,par2,df,mu){
  return(1-sum(noruin.aprox.lo(x,rho,par1,par2,df,mu)))
}
```

P. noruin.aprox.up Realiza la aproximación de la probabilidad de no-ruina para la cota superior de Panjer.

```
noruin.aprox.up<-function(x,rho,par1,par2,df,mu){
  if(floor(x)<x){return("x no es entero (survival)")}
  a=1/(1+rho-f.discup(0,par1,par2,df,mu))
  fi=rep(0,x)
  fi[1]=rho/(1+rho-f.discup(0,par1,par2,df,mu))
  g=0
  for(i in 1:x){g[i]=f.discup(i,par1,par2,df,mu)} #g[x]:=g(x)
  if(x>=2){
    for(i in 2:x){
      #fi[x+1]:=fi(x), hay que llenarlo como (fi1,fi0)->(fi2,fi1,fi0)
      fi[i]=a*g[1:(i-1)]%*%rev(fi[1:(i-1)])
    }
  }
  return(fi)
}
```

P. ruin.aprox.up Realiza la cota superior de Panjer.

```
ruin.aprox.up<-function(x,rho,par1,par2,df,mu){
  return(1-sum(noruin.aprox.up(x,rho,par1,par2,df,mu)))
}
```

P. ruina_aprox_asintot Realiza la aproximación asintótica de la probabilidad de la ruina.

```
ruina_aprox_asintot<-function(x,rho,par1,par2,df,mu){
  return((1-pF_I(x,par1,par2,df,mu))/rho)
}
```

Programas de resultados finales

P. sim_3geom_con_Panjer Realiza los tres algoritmos utilizando en cada paso una suma aleatoria para cada uno de ellos y exhibe las cotas de Panjer.

```

sim_3geom_con_Panjer<-function(R,x.inicial,par1,par2,df,mu,rho,lar,veces){
  valores=floor(seq(x.inicial,x.inicial+lar,lar/(veces-1)))
  A=B=D=matrix(0,nrow=veces,ncol=R,dimnames=list(valores,1:R))
  E=F=G=rep(0,veces)
  for(j in 1:veces){
    plot(x=j,y=0,main="Alg. I",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    A[j,]=Alg.I(valores[j],R,eps,rho,par1,par2,df,mu)
    points(x=j,y=0,main="Alg. II",ylim=c(-1,1),xlim=c(0,veces),
          xlab="No. de simulacion",ylab="")
    B[j,]=Alg.II(valores[j],R,eps,rho,par1,par2,df,mu)
    plot(x=j,y=0,main="Alg. III",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    D[j,]=Alg.III(valores[j],R,eps,rho,par1,par2,df,mu)
    plot(x=j,y=0,main="Panjer aprox",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    E[j]=ruin.aprox(valores[j],rho,par1,par2,df,mu)
    plot(x=j,y=0,main="Panjer up",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    F[j]=ruin.aprox.lo(valores[j],rho,par1,par2,df,mu)
    plot(x=j,y=0,main="Panjer low",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    G[j]=ruin.aprox.up(valores[j],rho,par1,par2,df,mu)
  }
  return(list(A=A,B=B,D=D,E=E,F=F,G=G,valores=valores))
}

```

P. sim_1geom_con_Panjer Realiza los tres algoritmos utilizando solo una suma geométrica durante toda su ejecución y exhibe las cotas de Panjer.

```

sim_1geom_con_Panjer<-function(R,x.inicial,par1,par2,df,mu,rho,lar,veces){
  plot(x=0,y=0,main="Simulando geométrica",xlab="",ylab="")
  sgeomtotal=sumgeom(R,eps,rho,par1,par2,df,mu)
  sgeomX=sgeomtotal$X
  sgeomM=sgeomtotal$M
  valores=floor(seq(x.inicial,x.inicial+lar,lar/(veces-1)))
  A=B=D=matrix(0,nrow=veces,ncol=R,dimnames=list(valores,1:R))
  E=F=G=rep(0,veces)
  for(j in 1:veces){
    plot(x=j,y=0,main="Algoritmos",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    algos=Algoritmos(valores[j],sgeomX,sgeomM,R,eps,rho,par1,par2,df,mu)
    A[j,]=algos[1,]
    B[j,]=algos[2,]
    D[j,]=algos[3,]
    plot(x=j,y=0,main="Panjer aprox",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    E[j]=ruin.aprox(valores[j],rho,par1,par2,df,mu)
    plot(x=j,y=0,main="Panjer up",ylim=c(-1,1),xlim=c(0,veces),

```

```

        xlab="No. de simulacion",ylab="")
F[j]=ruin.aprox.lo(valores[j],rho,par1,par2,df,mu)
plot(x=j,y=0,main="Panjer low",ylim=c(-1,1),xlim=c(0,veces),
     xlab="No. de simulacion",ylab="")
G[j]=ruin.aprox.up(valores[j],rho,par1,par2,df,mu)
}
return(list(A=A,B=B,D=D,E=E,F=F,G=G, valores=valores))
}

```

P. sim_1geom_con_asintota Realiza los tres algoritmos utilizando solo una suma geométrica durante toda su ejecución y exhibe la aproximación asintótica.

```

sim_1geom_con_asintota<-function(R,x.inicial,par1,par2,df,mu,rho,lar,veces){
  plot(x=0,y=0,main="Simulando geométrica",xlab="",ylab="")
  geomtotal=sumgeom(R,eps,rho,par1,par2,df,mu)
  sgeomX=sgeomtotal$X
  sgeomM=sgeomtotal$M
  valores=floor(seq(x.inicial,x.inicial+lar,lar/(veces-1)))
  A=B=D=matrix(0,nrow=veces,ncol=R,dimnames=list(valores,1:R))
  H=rep(0,veces)
  for(j in 1:veces){
    plot(x=j,y=0,main="Algoritmos",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    algos=Algoritmos(valores[j],sgeomX,sgeomM,R,eps,rho,par1,par2,df,mu)
    A[j,]=algos[1,]
    B[j,]=algos[2,]
    D[j,]=algos[3,]
    plot(x=j,y=0,main="Asintota",ylim=c(-1,1),xlim=c(0,veces),
         xlab="No. de simulacion",ylab="")
    H[j]=ruina_aprox_asintot(valores[j],rho,par1,par2,df,mu)
  }
  return(list(A=A,B=B,D=D,H=H, valores=valores))
}

```

Programas de tablas

P. imprime_de_totales1 Realiza una tabla con los resultados obtenidos para los programas `sim_3geom_con_Panjer` y `sim_1geom_con_Panjer`.

```

imprime_de_totales1<-function(A,B,D,E,F,G, valores,veces,
  R,eps,rho,par1,par2,df){
  for(i in 1:veces){
    tabla<-matrix(ncol=8,nrow=3,dimnames=list(c("Algoritmo I",
      "Algoritmo II","Algoritmo III"),
      c("  Aprox. ","  E. absoluto","  E. relativo","S.D.",
      "(Aprox-1.96*S.D./R^.5)",",Aprox+1.96*S.D./R^.5)","Long.inter",

```

```

    "log(S.D.)/log(fi.aprox)"))
tabla[1,]<-c(mean(A[i,]),abs(mean(A[i,])-E[i]),
  abs(mean(A[i,])-E[i])/E[i],sd(A[i,]),mean(A[i,])-1.96*sd(A[i,])/sqrt(R),
  mean(A[i,])+1.96*sd(A[i,])/sqrt(R),3.92*sd(A[i,])/sqrt(R),
  log(sd(A[i,]))/log(E[i]))
tabla[2,]<-c(mean(B[i,]),abs(mean(B[i,])-E[i]),
  abs(mean(B[i,])-E[i])/E[i],sd(B[i,]),mean(B[i,])-1.96*sd(B[i,])/sqrt(R),
  mean(B[i,])+1.96*sd(B[i,])/sqrt(R),3.92*sd(B[i,])/sqrt(R),
  log(sd(B[i,]))/log(E[i]))
tabla[3,]<-c(mean(D[i,]),abs(mean(D[i,])-E[i]), abs(mean(D[i,])-E[i])/E[i],
  sd(D[i,]),mean(D[i,])-1.96*sd(D[i,])/sqrt(R),
  mean(D[i,])+1.96*sd(D[i,])/sqrt(R),3.92*sd(D[i,])/sqrt(R),
  log(sd(D[i,]))/log(E[i]))
parametros<-c(valores[i],R,rho*100,par1,par2,df)
names(parametros)<-c("x","R","rho*100","par1","par2","df")
aproximacion.panjer=c(F[i],E[i],G[i],G[i]-F[i])
names(aproximacion.panjer)<-c("up","Panjer","lower"," Logitud de intervalo")
print("Parametros")
print(parametros)
print("Tabla de resultados")
print(tabla)
print("Aproximacion con Panjer")
print(aproximacion.panjer)
print("*****")
}
}

```

P. imprime_de_totales2 Realiza una tabla con los resultados obtenidos para el programa `sim_lgeom_con_asintota`.

```

imprime_de_totales2<-function(A,B,D,H,valores,veces,R,eps,rho,par1,par2,df){
  for(i in 1:veces){
    tabla<-matrix(ncol=8,nrow=3,dimnames=list(c("Algoritmo I",
      "Algoritmo II","Algoritmo III"),c("  Aprox.,"  E. absoluto",
      "  E. relativo","S.D. ","(Aprox-1.96*S.D./R^.5",
      ",Aprox+1.96*S.D./R^.5)","Long.inter","  log(S.D.)/log(fi.aprox)"))
    tabla[1,]<-c(mean(A[i,]),abs(mean(A[i,])-H[i]),
      abs(mean(A[i,])-H[i])/H[i],sd(A[i,]),mean(A[i,])-1.96*sd(A[i,])/sqrt(R),
      mean(A[i,])+1.96*sd(A[i,])/sqrt(R),3.92*sd(A[i,])/sqrt(R),
      log(sd(A[i,]))/log(H[i]))
    tabla[2,]<-c(mean(B[i,]),abs(mean(B[i,])-H[i]),
      abs(mean(B[i,])-H[i])/H[i],sd(B[i,]), mean(B[i,])-1.96*sd(B[i,])/sqrt(R),
      mean(B[i,])+1.96*sd(B[i,])/sqrt(R),3.92*sd(B[i,])/sqrt(R),
      log(sd(B[i,]))/log(H[i]))
    tabla[3,]<-c(mean(D[i,]),abs(mean(D[i,])-H[i]),
      abs(mean(D[i,])-H[i])/H[i],sd(D[i,]),mean(D[i,])-1.96*sd(D[i,])/sqrt(R),
      mean(D[i,])+1.96*sd(D[i,])/sqrt(R),3.92*sd(D[i,])/sqrt(R),
      log(sd(D[i,]))/log(H[i]))
    parametros<-c(valores[i],R,rho*100,par1,par2,df)
  }
}

```

```

names(parametros)<-c("x","R","rho*100","par1","par2","df")
aproximacion.asintotica=H[i]
print("Parametros")
print(parametros)
print("Tabla de resultados")
print(tabla)
print("Aproximacion asintotica")
print(aproximacion.asintotica)
print("*****")
}
}

```

P. imprime_de_totales3 Realiza una tabla con los resultados obtenidos para comparar los resultados de Panjer contra los de la aproximación asintótica.

```

imprime_de_totales3<-function(E,F,G,H,valores,veces,eps,rho,par1,par2,df){
  print("*****")
  parametros<-c(rho*100,par1,par2,df)
  names(parametros)<-c("rho*100","par1","par2","df")
  print("Parametros")
  print(parametros)
  tabla<-matrix(ncol=6,nrow=veces,dimnames=list(valores,
    c("Asintotico","Panjer","E. absoluto","E. relativo",
    "Cota inferior","Cota superior")))
  for(i in 1:veces){
    tabla[i,]<-c(H[i],E[i],abs(H[i]-E[i]),abs(H[i]-E[i])/E[i],F[i],
    G[i])}
  print("*****")
  print("Tabla de resultados, con respecto a capitla inicial")
  print(tabla)
  print("*****")
}
}

```

Programas para graficar

P. dibuja_algoritmos1 Grafica los resultados obtenidos para los programas `sim_3geom_con_Panjer` y `sim_1geom_con_Panjer`.

```

dibuja_algoritmos1<-function(A,B,D,E,F,G,valores){
  A.m=rowMeans(A)
  B.m=rowMeans(B)
  D.m=rowMeans(D)
  plot(y=E,x=valores,type="l",col="red",

```

```

ylim=c(min(A.m,B.m,D.m,E,F,G),max(A.m,B.m,D.m,E,F,G)),lwd=2,lty= 3,
xlab="Capital inicial",ylab="Probabilidad de ruina",
main="Aproximaciones de la probabilidad de ruina")
legend(par('usr')[2], par('usr')[4], xjust=1,
c("Aprox. de ruina","Algoritmo I","Algoritmo II","Algoritmo III"),
lwd=c(2,2,2,2),lty=c(3,1,1,1),
col=c("black","violet","green4","orange"))
lines(y=F,x=valores,col="blue",lwd=2,lty= 3)
lines(y=G,x=valores,col="blue",lwd=2,lty= 3)
points(y=A.m,x=valores,type="l",col="violet")
points(y=B.m,x=valores,type="l",col="green4")
points(y=D.m,x=valores,type="l",col="orange")
points(y=A.m,x=valores,pch=19,col="violet")
points(y=B.m,x=valores,pch=19,col="green4")
points(y=D.m,x=valores,pch=19,col="orange")
abline(h=0,lty=3)
abline(v=0,lty=3)
}

```

P. dibuja_algoritmos2 Grafica los resultados obtenidos para el programa `sim_1geom_con_asintota`.

```

dibuja_algoritmos2<-function(A,B,D,E,valores){
  A.m=rowMeans(A)
  B.m=rowMeans(B)
  D.m=rowMeans(D)
  plot(y=E,x=valores,type="l",col="blue",
        ylim=c(min(A.m,B.m,D.m,E),max(A.m,B.m,D.m,E)),lwd=2,lty= 3,
        xlab="Capital inicial",ylab="Probabilidad de ruina",
        main="Aproximaciones de la probabilidad de ruina")
  legend(par('usr')[2], par('usr')[4], xjust=1,
        c("Aprox. de ruina","Algoritmo I","Algoritmo II","Algoritmo III"),
        lwd=c(2,2,2,2),lty=c(3,1,1,1),
        col=c("blue","violet","green4","orange"))
  points(y=A.m,x=valores,type="l",col="violet")
  points(y=B.m,x=valores,type="l",col="green4")
  points(y=D.m,x=valores,type="l",col="orange")
  points(y=A.m,x=valores,pch=19,col="violet")
  points(y=B.m,x=valores,pch=19,col="green4")
  points(y=D.m,x=valores,pch=19,col="orange")
  abline(h=0,lty=3)
  abline(v=0,lty=3)
}

```

P. dibuja_algoritmos3 Grafica los resultados obtenidos para comparar los resultados de Panjer contra los de la aproximación asintótica.

```

dibuja_algoritmos3<-function(E,F,G,H,valores){

```

```

plot(y=E,x=valores,type="l",col="red",
     ylim=c(min(E,F,G,H),max(E,F,G,H)),lwd=2,lty= 3,
     xlab="Capital inicial",ylab="Probabilidad de ruina",
     main="Comportamiento asintótico y Panjer")
legend(par('usr')[2], par('usr')[4], xjust=1,
       c("Cotas de ruina","Panjer","Aprox. asintot."),
       lwd=c(2,2,2),lty=c(3,3,1),
       col=c("blue","red","green"))
lines(y=F,x=valores,col="blue",lwd=2,lty= 3)
lines(y=G,x=valores,col="blue",lwd=2,lty= 3)
lines(y=H,x=valores,col="green",lwd=2,lty= 1)
abline(h=0,lty=3)
abline(v=0,lty=3)
}

```

P. dibujatodo Grafica varios procesos de reserva.

```

dibujatodo<-function(t,R,c){
  n=dim(R)[1]
  m=dim(R)[2]
  plot(NULL,xlim=c(0,max(t)), ylim=c(min(R),1+max(R[,-1]+
  c(R[,-1]-R[,2:m]),0)),
  main="Historia de las empresas", xlab="TIEMPO", ylab="R(t)");
  abline(h=0, col = "lightgray", lty=3)
  for(k in 1:n){
    for(j in 1:(m-1)){
      segments(t[k,j],R[k,j],t[k,j+1],R[k,j]+c*(t[k,j+1]-t[k,j]),
              col="blue")
      segments(t[k,j+1],R[k,j]+c*(t[k,j+1]-t[k,j]),
              t[k,j+1],R[k,j+1],col="green",lty=3)
      if(min(R[k,j])<0)segments(t[k,j],R[k,j],t[k,j]+
      (min(R[k,j]+c*(t[k,j+1]-t[k,j]),0)-R[k,j])/c,
      min(R[k,j]+c*(t[k,j+1]-t[k,j]),0),col="red")
    }
  }
}

```

Bibliografía

- [1] Arnold, B. C., Balakrishnan, N., and Naqaraja, H. N., *A First Course in Order Statistics*. Wiley, New York, 1992.
- [2] Asmussen, S. (2000). *Ruin Probabilities*. Singapore: World Scientific Publishing.
- [3] Asmussen, S. (2003). *Applied Probability and Queues*. New York: Springer-Verlag.
- [4] Asmussen, S. & D. P. Kroese (2006). *Improved algorithms for rare event simulation with heavy tails*. *Advances in Applied Probability* 38, 545–558.
- [5] Asmussen, S. & K. Binswanger (1997). *Simulation of ruin probabilities for subexponential claims*. *ASTIN Bulletin* 27, 297–318.
- [6] Asmussen, S., K. Binswanger, & B. Højgaard (2000). *Rare events simulation for heavy-tailed distributions*. *Bernoulli* 6, 303–322.
- [7] Asmussen, S. & P. W. Glynn (2007). *Stochastic Simulation: Algorithms and Analysis*. New York: Springer-Verlag.
- [8] Bartle, R. G. (1995). *The Elements of Integration and Lebesgue Measure*. Wiley Classics Edition. John Wiley & Sons. New York.
- [9] Bingham, N.H., Goldie, C.M. & Teugels, J.L. (1987) *Regular Variation*. Cambridge University Press, Cambridge.
- [10] Crawley, M. J. (2007). *The R book*. Imperial College London at Silwood Park. John Wiley & Sons. UK.
- [11] Crawley, M.J. (2005). *Statistics: An Introduction using R*. Imperial College London at Silwood Park. John Wiley & Sons. UK.

- [12] Dickson, C. M. (1995). *A review of Panjer's recursion formula and its applications*. British Actuarial Journal .
- [13] Dimov, I.T. (2008). *Monte Carlo methods for applied scientists*. World Scientific Publishing Co. Pte. Ltd.
- [14] Embrechts P., Kluppelberg C. & Mikosch T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Heidelberg.
- [15] Embrechts,P., & Veraverbeke,N.(1982). *Estimates for the probability of ruin with special emphasis on the possibility of large claims*. Insurance: Mathematics and Economics I (1982), 55-72.
- [16] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications II*. Wiley, New York.
- [17] Fishman, G. (1996). *Monte Carlo. Concepts, Algorithms and Applications*. Springer-Verlag.
- [18] Gentle, J.E. (2003). *Random Number Generation and Monte Carlo Methods*. 2nd Edition. Springer-Verlag.
- [19] Gut, A. (2005), *Probability: A graduate course*. Springer-Verlag.
- [20] Kaas, R., Goovaerts, M. Dhaene, J. & Denuit, M. (2008). *Modern Actuarial Risk Theory: Using R*. Springer-Verlag Berlin Heidelberg.. Kluwer.
- [21] Karr, A. F. (1993). *Probability*. Springer-Verlag.
- [22] Mikosch, T. (2009). *Non-Life Insurance Mathematics An Introduction with the Poisson Process*. 2nd Edition. Springer-Verlag Berlin Heidelberg 2009.
- [23] Panjer, H. H. (2006). *Operational Risk: Modeling Analytics*. Wiley series in probability and statistics.
- [24] Panjer, H. & Willmot, G. (1992). *Insurance Risk Models*. Chicago: Society of Actuarie
- [25] Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
- [26] Resnick, S. I. (1992). *Adventures in Stochastic Processes*. Boston: Birkhäuser.

- [27] Resnick, S. I. (2006). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer-Verlag.
- [28] Rincón, L. (2007). *Curso intermedio de probabilidad*. Departamento de Matemáticas. Facultad de Ciencias UNAM.
- [29] Rincón, L. (2007). *Introducción a la teoría del riesgo*. Departamento de Matemáticas. Facultad de Ciencias UNAM.
- [30] Ripley, B. (1987). *Stochastic Simulation*. Wiley, New York.
- [31] Rojas, N. L. (2008). *Risk Probabilities: Asymptotics and Simulation*. PhD Dissertation.
- [32] Rolski, T., Schmidli, H., Schmidt, V. & Teugels, J. (1998). *Stochastic processes for insurance and finance*. Wiley Series in Probability and Statistics.
- [33] Ross, S. (1976). *A First Course in Probability*. Prentice Hall.
- [34] Ross, S. (1999). *Simulación*. 2a Edición. Prentice Hall.
- [35] Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo Method*. Wiley. Wiley Series in Probability and Statistics.
- [36] Rubinstein, R. Y. & Melamed, B. (1998). *Modern Simulation and Modeling*. John Wiley & Sons. New York.
- [37] Thompson, J. R. (2000). *Simulation. A Modeler's Approach*. Wiley.