

Tiempos de pasada condicional para el Proceso de Riesgo de Crámer Lundberg

Roberto Vásquez Martínez¹, Kapioma Villarreal Haro², Christian Adrian Sánchez León³, Emiliano Iván Ortiz Arana⁴, Jorge Alfonso Cárdenas Treviño⁵, Ehyter M. Martin Gonzalez⁶

^{1,2,3,4,5,6}Universidad de Guanajuato

r.vasquezmartinez@ugto.mx¹

k.villarrealharo@ugto.mx²

ca.sanchezleon@ugto.mx³

ei.ortizarana@ugto.mx⁴

ja.cardenastrevino@ugto.mx⁵

ehyter.martin@ugto.mx⁶

Resumen

El proceso de Crámer Lundberg, propuesto por primera vez en (Lundberg, 1903), es un modelo clásico en Teoría de Riesgo. Particularmente, en el contexto de compañías aseguradoras, modela el capital de la compañía considerando ingresos constantes de los asegurados y reclamos que se presentan a lo largo del tiempo. En muchas ocasiones es de interés analizar los tiempos de llegada a cierto nivel de capital a que puede representar un nivel de riesgo para la estabilidad económica de la compañía antes de llegar a cierto nivel b que puede entenderse en este contexto como un nivel de rentabilidad. En este trabajo se estudia este proceso y esos tiempos de llegada. Se estudia de forma empírica y numérica la distribución de tiempos de llegada, bajo el supuesto de que la distribución de los saltos del proceso los rige una distribución en el dominio de atracción maximal Gumbel, lo que permitirá calcular probabilidades de interés en el contexto de su aplicación.

Palabras clave: Teoría de Extremos; Tiempos de Pasada; Crámer-Lundberg; Riesgo; Gumbel

Introducción

Un modelo estocástico describe un conjunto de posibles observaciones mediante sus pesos o probabilidades. La elección de si un modelo es estocástico o determinista la toma el observador, aunque muchas veces resulta claro cuando un fenómeno es aleatorio, es decir, condicionado a las probabilidades.

En el contexto de Teoría de Riesgo, el modelo clásico de riesgo o proceso de Crámer-Lundberg fue propuesto por primera vez en (Lundberg, 1903). Es un proceso sumamente útil debido a su simplicidad, razón por la cual ha sido fuertemente aplicado y estudiado a lo largo del tiempo. Actualmente existen generalizaciones de este modelo, que en realidad es simplemente un caso particular de una clase amplia de procesos, llamada Procesos de Lévy (Kyprianou, 2006, p. 2). Diversos artículos, por ejemplo (Hieber, 2013) y (Avram et al., 2014), han estudiado propiedades relacionadas con esta clase de procesos en el contexto del problema de salida fuera una barrera. Para ser más precisos, este problema consiste en estudiar la probabilidad de que un proceso estocástico que inicia dentro de una barrera, delimitada por un intervalo (a, b) , salga de dicha barrera y de qué manera dicha salida ocurre, esto es, si lo hace superando el nivel de rentabilidad b , o si lo hace cruzando por debajo del nivel crítico.

En Kyprianou, 2014 este problema es tratado y resuelto en términos de las llamadas funciones de escala (Kyprianou, 2014, pp 231-235). Dichas funciones están completamente determinadas por sus transformadas de Laplace, las cuales usualmente tienen expresiones simples. Sin embargo, como puede observarse en las fórmulas dadas en (Kyprianou, 2014), la distribución de los tiempos de salida de la barrera de interés no se conocen explícitamente; sólo se conoce su transformada de Laplace, la cual queda en términos de las funciones de escala, mismas que en general, no tienen expresión analítica cerrada. Esto último hace particularmente difícil utilizar las fórmulas en (Kyprianou, 2014) más que para propiedades teóricas de la distribución de los tiempos de salida de la barrera.

Debido a lo anterior, en este estudio aplicamos un método clásico de la Teoría de Valores Extremos para analizar datos simulados de los tiempos de salida del Proceso de Crámer-Lundberg y, con esto, hacer

inferencia sobre las propiedades de las distribuciones de dichos tiempos. Cabe mencionar que, pese a la elección particular de este proceso, dadas las propiedades de continuidad de los procesos de Lévy y funcionales asociados, es posible aproximar numéricamente cualquier proceso de Lévy sin componente gaussiana usando un proceso Poisson compuesto, como lo es el proceso de Crámer-Lundberg. Esto significa que la metodología utilizada puede replicarse posteriormente en cualquier otro proceso estocástico espectralmente negativo (Asmussen & Albrecher, 2010, p. 337).

En la primera sección, analizaremos la naturaleza matemática general de los procesos estocásticos y la abstracción de variables y cantidades de interés en estos modelos. En particular, en la segunda sección consideraremos el proceso de riesgo de Crámer-Lundberg, el contexto en el que se utiliza y cómo los conceptos que se derivan de este proceso proponen respuestas a las preguntas más naturales que surgen, por ejemplo, en fenómenos que involucran a compañías aseguradoras. Posteriormente, discutimos el marco teórico utilizado para aproximar de forma numérica la distribución de un variable de interés, haciendo énfasis en el método de excesos sobre un umbral introducido por (Pickands, 1975). Se presentan en las últimas secciones los resultados computacionales y las conclusiones que se derivan de ellos.

Procesos Estocásticos

Un proceso estocástico es una familia de variables aleatorias $(X_t)_{t \in T}$ indexadas por el parámetro t que se mueve a lo largo de un conjunto de índices T apropiado. En una situación común, t representa al tiempo de forma continua, por lo que $T = [0, \infty)$.

Un proceso estocástico que ha resultado importante es el [proceso Poisson Simple](#), el cual abarca las nociones de independencia y la utilidad del modelo Poisson (Taylor & Karlin, 1998, p. 270).

Definición. Un proceso de Poisson Simple con intensidad, o tasa, $\lambda > 0$ es un proceso estocástico N_t con valores en el conjunto de enteros no negativos para el cual se cumplen las siguientes condiciones,

1. Incrementos independientes: Para $t_0 = 0 < t_1 < t_2 < \dots < t_n$ las variables aleatorias,

$$N_{t_0} - N_{t_1}, N_{t_1} - N_{t_2}, \dots, N_{t_{n-1}} - N_{t_n},$$

son independientes.

2. Estacionariedad. Para $s \geq 0$ y $t > 0$, la variable aleatoria $N_{t+s} - N_s$ tiene distribución Poisson con parámetro λt .
3. $X_0 = 0$.

Se puede demostrar que, en efecto, cualquier proceso de contar estacionario con incrementos independientes es precisamente el Proceso de Poisson Simple (Cont & Tankov, 2004, p. 67), y aquí radica la importancia de este proceso, pues las propiedades de independencia y estacionariedad son inherentes a muchos fenómenos. Este proceso puede modelar, por ejemplo, visitas de clientes a cierta tienda siguiendo una tasa constante de llegadas λ por unidad de tiempo. De esto último podemos intuir una utilidad de este proceso en el problema de modelación de capital de una compañía aseguradora: el proceso Poisson puede contar la cantidad aleatoria de reclamos en un intervalo de tiempo.

Si consideramos $(\tau_i)_{i \geq 1}$ una sucesión de variables aleatorias exponenciales independientes de parámetro λ , y consideremos $T_n = \sum_{i=1}^n \tau_i$ las sumas acumulativas de ellas, el proceso de Poisson simple de parámetro λ puede ser expresado de la forma

$$N_t = \sum_{n \geq 1} 1_{t \geq T_n},$$

lo cual puede ser de utilidad para ciertas aplicaciones y hace más evidente su cualidad de proceso de conteo.

Es también de interés calcular probabilidades de los tiempos de llegada en los cuales el proceso alcanza cierto nivel. Estos niveles se pueden entender como valores de control del proceso. Formalmente, la definición matemática de estos tiempos de llegada, que son también variables aleatorias, es la siguiente,

Definición. El tiempo de llegada del proceso X_t al nivel a , con $X_0 > a$, es la variable aleatoria definida como

$$\tau_a = \inf\{t \in T : X_t \leq a\}.$$

De forma análoga, si $X_0 < b$ entonces el tiempo de llegada al nivel b se define como

$$\tau_b = \inf\{t \in T : X_t \geq b\}.$$

En ese sentido, si X_t es el proceso estocástico que modela el capital de una compañía de seguros, entonces, estas variables aleatorias pueden ser interpretables y describir los primeros instantes en los que el proceso llega a un nivel inferior a , que puede entenderse como un nivel de alarma, o a un nivel b , uno de excedencia.

Claramente es de interés estudiar propiedades probabilísticas de estos tiempos de llegada, puntualmente, para describir probabilidades de llegada al nivel de alarma antes que al nivel de excedencia, situación de gran importancia para una compañía aseguradora.

Proceso de Crámer-Lundberg

Dentro de la gran familia de procesos estocásticos, existe un modelo clásico de riesgo propuesto por primera vez en (Lundberg, 1903) con la intención de estudiar el comportamiento del capital de una compañía bajo ciertas condiciones.

Definición. El proceso de Cramér-Lundberg es un proceso estocástico a tiempo continuo $X = (X_t)_{t \geq 0}$ definido como

$$X_t := x + ct - \sum_{i=1}^{N_t} Y_i$$

donde $x \geq 0$ denota el capital inicial; ct es entrada de capital por concepto de primas hasta el tiempo t con $c > 0$ una tasa constante de ingresos, conocida como *drift* del proceso, por unidad de tiempo; $(N_t)_{t \geq 0}$ es un proceso Poisson simple de intensidad $\lambda > 0$ que describe el número de reclamos hasta el tiempo t ; y $(Y_i)_{i \in \mathbb{N}}$ es una secuencia de variables aleatorias no negativas, independientes, idénticamente distribuidas entre sí e independientes de $(N_t)_{t \geq 0}$, que describe el tamaño de los reclamos. Como condiciones adicionales, supondremos que los reclamos tienen media finita y positiva $0 < \mu = E[Y_i] < \infty$, y que se satisface $c > \lambda\mu$ la *net profit condition*. Esta condición es realista debido a la suposición de que la tasa constante de ingresos es mayor a la media de los reclamos. De lo contrario la ruina, esto es, cuando $X_t \leq 0$, ocurriría con probabilidad uno, lo cual no es de interés para una compañía de seguros cuyo capital se modele con el proceso X .

El anterior es un modelo clásico de riesgo y podemos interpretar a la variable aleatoria X_t como el capital de la compañía al tiempo t . Notemos también que el modelo hace algunas suposiciones, como que sólo hay ingresos a la compañía provenientes de las primas y es un ingreso constante de capital (el término ct), y que no hay pérdidas o gastos, salvo las que corresponden a los reclamos (la suma aleatoria $\sum Y_i$).

El estudio de tiempos de llegada en este proceso resulta de suma importancia. Por ejemplo, el primer tiempo de pasada $\tau_0 = \inf\{t \in T : X_t \leq 0\}$ se conoce como tiempo de ruina del proceso estocástico en el contexto del modelo de Crámer-Lundberg. El problema clásico de ruina consiste en estudiar

$$\psi(u) = \mathbb{P}[\tau_0 < \infty | X_0 = u]$$

a la cual se le conoce como probabilidad de ruina del proceso estocástico. Es posible ver que esta función de probabilidad satisface que es decreciente y que $\lim_{u \rightarrow \infty} \psi(u) = 0$. En el contexto de compañías aseguradoras se busca determinar estrategias para que esta probabilidad sea mínima con base en algún contexto de optimalidad, para lo cual notamos que el estudio se reduce a la distribución de probabilidad de un tiempo de pasada.

Como parte de las estrategias que la compañía de seguros puede tomar para tratar de reducir su probabilidad de ruina, está el problema de salida bilateral, en el cual interesa justamente que el proceso estocástico rebase un valor fijo b antes de cruzar debajo de otro valor fijo a .

En este caso, una posible estrategia es optimizar la prima c de modo que con alta probabilidad, el cruce ocurra primero arriba de b , de modo que el exceso de capital por encima de b pueda ser utilizado en una reserva, contratar un reaseguro o realizar pago de dividendos.

Este problema, en el caso de procesos estocásticos más generales, también se encuentra por ejemplo en la industria, en lo referente a control de procesos de producción. En este caso, las franjas representan niveles bajo los cuales el proceso de producción se encuentra estable.

Método de Excesos sobre un umbral

En muchos contextos es de interés conocer los niveles en los que una variable aleatoria queda acotada superiormente con una probabilidad alta. En el caso de la variable aleatoria que representa los tiempos de llegada, es relevante conocer los valores máximos que puede alcanzar con casi toda certeza o probabilidad.

Sin embargo, a pesar de la simplicidad del proceso de Crámer-Lundberg, debido a que la distribución de los saltos se puede tomar de forma arbitraria pues sólo debe cumplir con tener soporte positivo, el problema de calcular estos valores extremos resulta retador.

En la Teoría de Valores Extremos existe una herramienta que permite caracterizar al máximo de ciertas distribuciones en una sólo familia de distribuciones de forma asintótica, como no podría ser de otra manera, a través de la distribución generalizada de valores extremos (DGVE). Este resultado puramente matemático es el denominado Teorema de Fisher-Tippet, nos dice que bajo ciertas condiciones, si el máximo normalizado de variables aleatorias con cierta distribución converge a una distribución no degenerada, entonces esta distribución límite es de la familia de valores extremos conformada por las funciones de distribución Fréchet, Gumbel y Weibull.

Este hecho matemático nos permite, al menos de forma aproximada, calcular la distribución del máximo de variables aleatorias con cierta distribución, digamos F , con alguna de las distribuciones Fréchet, Gumbel y Weibull que han sido ampliamente estudiadas. Aquí vemos cómo el problema en primera instancia es demasiado abierto pero puede ser aproximado con modelos ampliamente conocidos. Dependiendo de si la distribución del máximo de F es Gumbel, Weibull o Fréchet decimos que F pertenece al dominio de atracción maximal Gumbel, Weibull o Fréchet. En la práctica, nos valemos de resultados matemáticos para construir pruebas gráficas y estadísticas que nos permitan dilucidar en qué dominio de atracción maximal estamos, y de esa forma explotar todas sus propiedades.

Como dijimos al principio de esta sección, estamos interesados en niveles que concentran toda la probabilidad de una variable aleatoria. En el contexto de riesgo, podemos pensar en los valores extremos del tiempo de llegada τ_a condicionado al evento $\{\tau_a < \tau_b\}$. Más aún, dado estos niveles máximos, estamos interesados en el valor que podría esperarse de la variable aleatoria si supera ese valor extremo, esto exactamente se engloba en la siguiente definición.

A continuación definimos la distribución generalizada de Pareto, herramienta fundamental para el método de excesos sobre umbral.

Definición. Definimos a la siguiente función de distribución

$$P_{\xi, \alpha, \beta}(x) = \begin{cases} 1 - \left(1 + \xi \left(\frac{x-\beta}{\alpha}\right)\right)^{-\frac{1}{\xi}}, & \text{si } \xi \neq 0, \\ 1 - e^{-\frac{x-\beta}{\alpha}}, & \text{si } \xi = 0 \end{cases}$$

con $x \geq \beta$ si $\xi \geq 0$ y $0 \leq x \leq -\alpha/\xi + \beta$ si $\xi < 0$ como la distribución generalizada de Pareto (DGP).

Otra definición importante para el método de excesos sobre umbral es la siguiente.

Definición. Sea X una variable aleatoria no negativa con distribución F . Consideramos u tal que $\underline{F}(u) > 0$. A la función

$$e_F(u) := \mathbb{E}[X - u | X > u]$$

la denominamos la función de exceso promedio. En la práctica, y para las simulaciones realizadas, se utiliza un estimador numérico de esta función cuyo uso se justifica a través de teoremas de convergencia (Ghosh & Resnick, 2010).

Esta función tiene gran utilidad práctica por su interpretación y también nos permite construir un diagnóstico de en qué dominio de atracción maximal pudiésemos estar. Por ejemplo, existe un resultado que nos indica que, para una distribución F concentrada en la parte no negativa de la real con extremo derecho infinito, con media finita, si la función de exceso promedio tiene límite cuando $u \rightarrow \infty$, entonces este límite es infinito si y sólo si esta distribución es de cola pesada.

Este tipo de resultados nos permitirán evaluar si hay evidencia de estar en un dominio de atracción o en otro. Con esta información podemos utilizar la Teoría de Valores Extremos, con la finalidad de aproximar la distribución condicional del tiempo de llegada τ_a .

Será de particular interés el caso de la función de exceso promedio cuando la variable X tiene distribución exponencial. En este caso, es posible demostrar que

$$e_F(u) = \theta,$$

donde θ es la media de la distribución exponencial.

Existe teoría que caracteriza los dominios Fréchet y Weibull, que se podrían considerar como duales, pues el dominio Fréchet es el conjunto de distribución de variación regular, que son de cola pesada; mientras que el dominio Weibull las distribuciones tienen extremo derecho finito, pero tales que vía una transformación, se puede obtener una distribución de variación regular (Lindgren et al., 1983).

El problema del Dominio Gumbel.

Los dominios de atracción Weibull y Fréchet engloban familias de distribuciones con propiedades más específicas que el dominio Gumbel. Mientras que el dominio Fréchet es equivalente a las distribuciones de variación regular que son de cola pesada y el dominio Weibull solo incluye a distribuciones con extremo derecho finito, las cuales son de cola ligera, el dominio Gumbel puede incluir tanto distribuciones de cola pesada como de cola ligera con extremo derecho finito o infinito. Más aún, distribuciones de cola larga como la distribución log-normal, forman parte de este dominio. Puede pensarse al dominio Gumbel como el punto medio entre los dominios de atracción maximales, por lo que resulta el más problemático debido a que no tiene una caracterización tan precisa como los otros dos.

Esto motiva el estudio de este dominio de atracción pensando que la distribución de los reclamos en el proceso de Crámer-Lundberg forma parte de este dominio.

Distribución Pareto Generalizada.

La forma de aproximar los valores extremos se puede realizar a través de la distribución generalizada de Pareto (DGP) (Coles, 2011, p. 75) si la distribución subyacente que se quiere aproximar pertenece a algún dominio de atracción maximal. Este resultado se conoce como el Teorema de Pickands-Balkema-De Haan (Pickands, 1975). Se puede ver a partir de aquí la importancia de considerar dominios de atracción

maximales, pues con este resultado matemático podemos estimar los valores extremos que queremos, por ejemplo, del tiempo de llegada condicional, utilizando la distribución Generalizada de Pareto.

La idea general de la aproximación numérica propuesta para los valores extremos de la distribución condicional de τ_a al evento $\{\tau_a < \tau_b\}$, es evaluar si esta distribución pertenece a un dominio de atracción maximal y luego utilizar la aproximación de estos cuantiles extremos que brinda la distribución Pareto Generalizada.

En el proceso de Crámer-Lundberg, si consideramos la distribución del tamaño de los reclamos en el dominio Gumbel, evaluaremos si la distribución del tiempo de llegada condicional pertenece a algún dominio de atracción para determinados valores de a, b . Además veremos si dicha distribución es, al menos, asintóticamente exponencial.

Detección del umbral.

La idea de la función de exceso promedio en un análisis exploratorio es la de justamente la detección del dominio de atracción. Para el caso de dominio de atracción Gumbel se tiene lo siguiente

1. Si F tiene extremo derecho infinito entonces,

$$\lim_{u \rightarrow \infty} \frac{e_F(u)}{u} = 0.$$

2. Si F tiene extremo derecho finito $\omega_F < \infty$ entonces

$$\lim_{u \rightarrow \infty} \frac{e_F(u)}{\omega_F - u} = 0.$$

Este resultado nos permite explorar, a través de un estimación empírica de la función de exceso promedio, la distribución de los tiempos de pasada, graficando las funciones e_F y $e_F(u)/u$ en la muestra de estadísticos ordenados de los tiempos de llegada condicional.

Ajuste de la DGP.

Para realizar el ajuste vía la DGP se utilizó la paquetería *evir* del lenguaje R (Pfaff, n.d.). Como el ajuste de la DGP funciona para valores extremos, se fija un umbral u de forma que el porcentaje de la muestra que excede a ese umbral es menor al 5% si la muestra lo permite. En nuestro caso, debido a que la muestra fue simulada, consideramos un tamaño de dicha muestra suficientemente grande para poder aplicar este criterio.

Con base en métodos gráficos, notamos que no existe evidencia en contra de la aplicabilidad del Teorema de Pickands-Balkema-De Haan, el cual nos dice que $\tau_a - u$ condicionado a $\{\tau_a < \tau_b\}$, se puede aproximar utilizando la distribución generalizada de Pareto. Para validar la bondad de ajuste realizamos gráficas PP y QQ, utilizando bandas de confianza construidas vía bootstrap (Efron, 1972) a través de la paquetería *qqplotr* (Almeida & Loy, 2021).

En general, los resultados numéricos parecieron indicar que la distribución del tiempo de pasada de interés pertenece también al dominio de atracción Gumbel, en cuyo caso su aproximación vía una DGP se reduce a aproximar vía una distribución exponencial. Para dar más fuerza a la validez de resultados, se usa también el estadístico de prueba

$$LR = 2 \left(\ell(X; \hat{\xi}_0, \hat{a}_0(u)) - \ell(X; \hat{\xi}_1, \hat{a}_1(u)) \right),$$

el cual tiene asintóticamente distribución χ^2 (ji-cuadrada) con el objetivo de contrastar los parámetros de escala y forma a y b de las aproximaciones con los de la DGP cuando $\xi = 0$, lo cual corresponde al caso en que la DGP es una distribución exponencial.

Con lo anterior en mente, se realiza también el ajuste de los datos a través del modelo exponencial para validar, al menos numéricamente, si existe evidencia de un buen ajuste asintótico a través de la distribución exponencial a la distribución de los tiempos de llegada condicional.

Simulaciones

El esquema general para simular los tiempos de tiempo de llegada τ_a condicionado a que $\{\tau_a < \tau_b\}$ fue simular el proceso de Crámer-Lundberg a través de la librería *ruin* del lenguaje R (Rudnytsky, 2018), y elegir las trayectorias condicionadas al evento de interés.

Se consideraron tres distribuciones diferentes para el tamaño de los saltos:

1. Exponencial con tasa $1/\theta$.
2. Lognormal con escala σ y localización μ .
3. Weibull con escala σ y forma κ .

En todas las simulaciones hechas se consideró $\lambda = 1$ en el proceso Poisson simple que describe los tiempos de llegada de los reclamos.

Como dijimos anteriormente, para no caer en un caso trivial, esto es, la ruina con probabilidad uno, consideramos el *drift* c de forma que se cumpla la *net profit condition*. En particular consideramos $c - \lambda\theta \in \{0.2, 0.45, 0.7\}$, donde θ es la media de la distribución de los saltos. Con esto se garantiza que se cumple la *net profit condition* y que el *drift* y la media de los saltos tienen aproximadamente el mismo valor.

Por otro lado, los valores de a y b los consideramos de forma que si el capital inicial es 0 entonces el nivel a sea crítico pero realista. Como estamos considerando un caso balanceado, pensamos al intervalo (a, b) simétrico alrededor de 0. Específicamente, dado que queremos estos niveles sean realistas los consideramos como sigue

$$(a, b) = (-n\lambda\theta, n\lambda\theta), \quad n \in \{2, 3, 4\}.$$

Posterior a las simulaciones, graficamos la función de exceso promedio empírica en cada una de ellas. Esto para determinar su tipo de cola, ya sea pesada o ligera, intentando hacer un diagnóstico con los resultados que se tienen para el dominio Gumbel.

En el caso de las simulaciones con la distribución Weibull, consideramos un parámetro de forma $\kappa < 1$, para no tener una media degenerada en 1.

En la siguiente tabla hacemos un resumen de los parámetros utilizados.

Tabla 1. Descripción de las distribuciones utilizadas en las simulaciones del Proceso de Cramer-Lundberg.

Exponencial	LogNormal	Weibull
$1/\theta = 0.1$	$\sigma = 1, \theta$ = 2	$\sigma = 1, \kappa$ = 0.35
$1/\theta = 0.2$	$\sigma = 1, \theta$ = 5	$\sigma = 1, \kappa$ = 0.4
$1/\theta = 0.3$	$\sigma = 1, \theta$ = 10	$\sigma = 1, \kappa$ = 0.45

Para cada una de estas distribuciones se prueban los tres diferentes valores del *drift* c y los tres intervalos (a, b) considerados. La implementación de estas 81 simulaciones se encuentra en el notebook Simulaciones Tiempos de Llegada Condicional al que se debe pedir acceso. En cada simulación se considera un valor máximo de iteraciones de 10000 y en todas se alcanza a obtener una muestra de tamaño 2000. En general, en los tres casos el comportamiento es el esperado y similar entre los tres casos y las múltiples simulaciones realizadas.

Resultados reclamos Exponenciales

En cada una de estas simulaciones el ajuste gráfico fue bastante similar en el conjunto de 27 simulaciones. Mostramos en la Figura 1 el resultado del análisis exploratorio hecho a través de la función de exceso promedio.

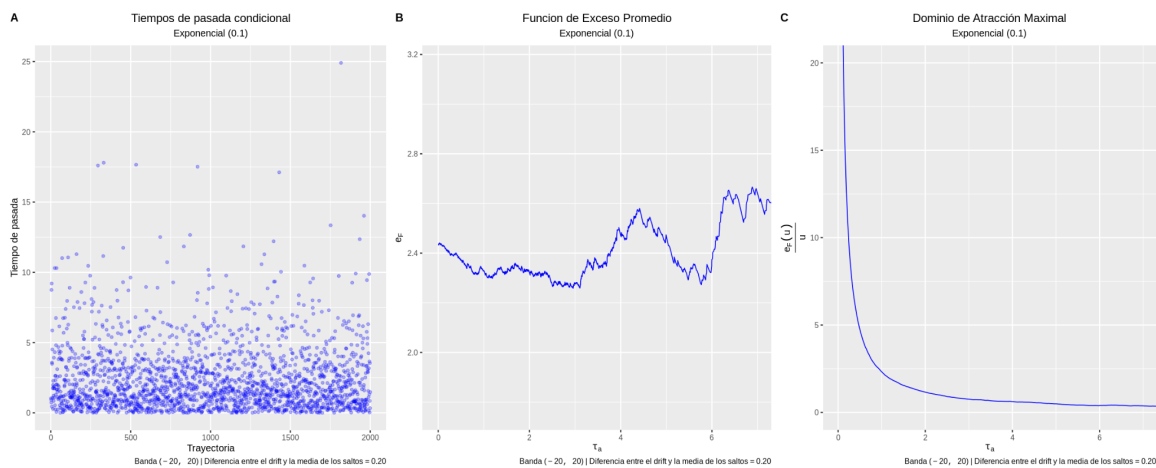


Figura 1. Análisis Exploratorio Distribución Exponencial con media 10 y drift 10.2 .

Se ve de hecho que hay evidencia a favor del dominio Gumbel, más aún, que la distribución subyacente es de cola ligera. El ajuste que se presenta con la DGP y el modelo exponencial se muestra en la Figura 2.

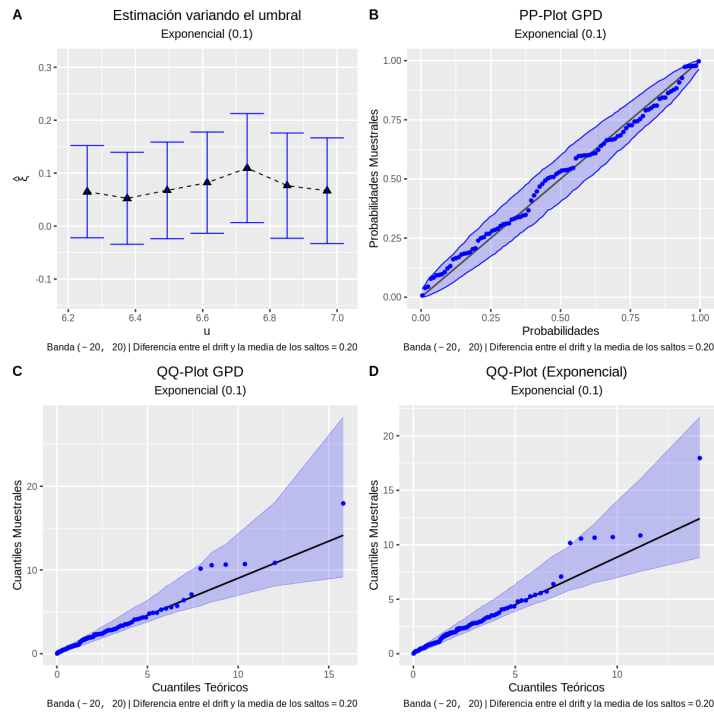


Figura 2. Bondad de ajuste del modelo DGP y Exponencial.

El ajuste es bueno utilizando el modelo DGP y Exponencial. En este caso, al realizar una prueba de razón de verosimilitudes entre el modelo DGP y Exponencial de forma global obtenemos un estadístico de valor de 0.54, mientras que el valor crítico de esta prueba con nivel de significancia 0.95 es 3.84. Luego, no hay evidencia en contra del dominio Gumbel.

Además, debido a que el modelo DGP con localización y forma 0 es el modelo exponencial, comparamos el inverso de la escala del modelo DGP con la tasa del modelo exponencial para cada simulación. El resultado correspondiente se presenta en la Figura 3. Debido a que se observa un ajuste similar en ambos modelos esperamos que la gráfica de estos puntos quede cercana a la recta identidad, evidencia a favor de que la distribución de los excesos de los tiempos de llegada τ_a condicionado a que $\{\tau_a < \tau_b\}$ es al menos asintóticamente exponencial en el sentido del Teorema de Pickands-Balkema-De Haan..

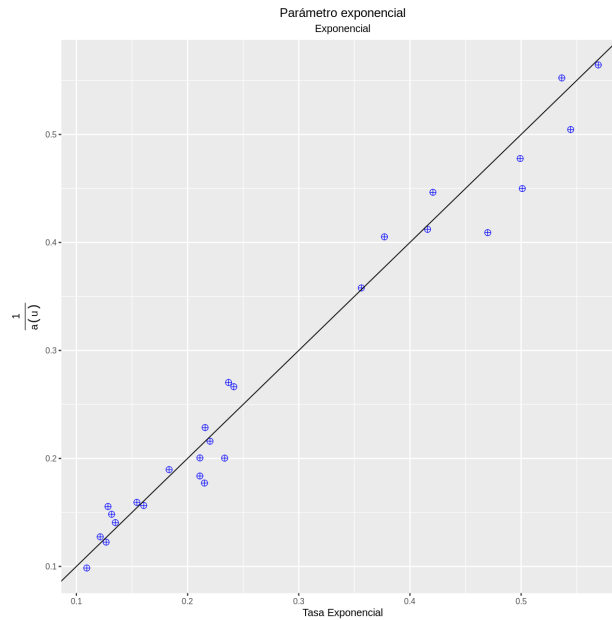


Figura 3. Tasas exponenciales estimadas contra el inverso de la escala DGP

Resultados reclamos Lognormales

Obtuvimos resultados similares considerando la distribución de los saltos lognormal. Primero, parece que hay evidencia de que los datos tienen una distribución con derecho extremo infinito, tal como ilustra el gráfico de dispersión de los datos de pasada. Ahora, dado que e_F parece tender a un límite finito, tenemos evidencia para decir que no es una distribución de cola pesada. Ya que no hay evidencia en contra de una distribución de cola ligera, y como $e_F(u)/u$ tiende a cero, bajo las suposiciones que hemos hecho la evidencia apunta en favor de que los datos pertenezcan al dominio de atracción Gumbel.

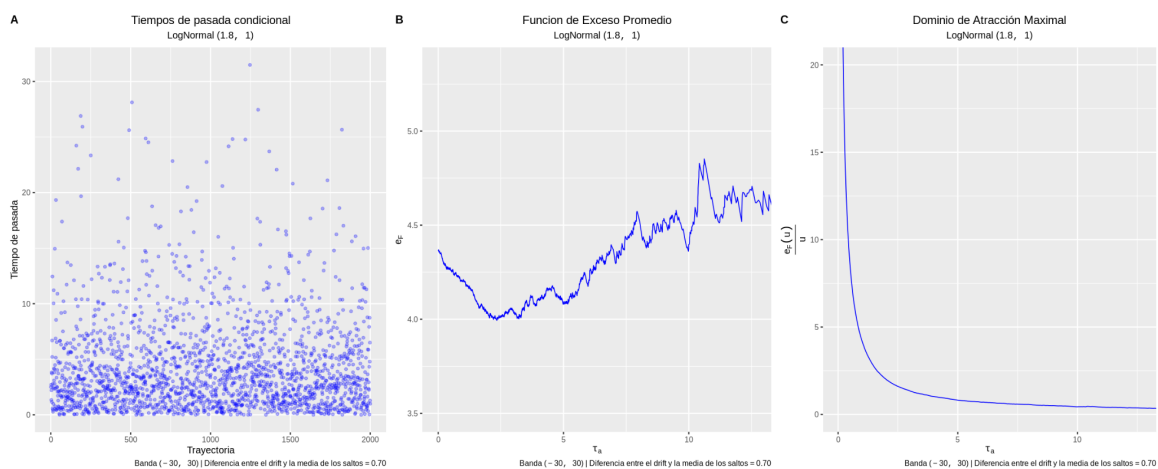


Figura 4. Análisis Exploratorio Distribución Lognormal con media 10

Al realizar el ajuste usando el modelo DGP, notamos que ambos, tanto el modelo DGP como el exponencial, reflejan en las gráficas de bondad de ajuste que son buenas estimaciones. Además, en la gráfica de estimación variando el umbral, no se muestra gran discrepancia entre los valores mostrados.

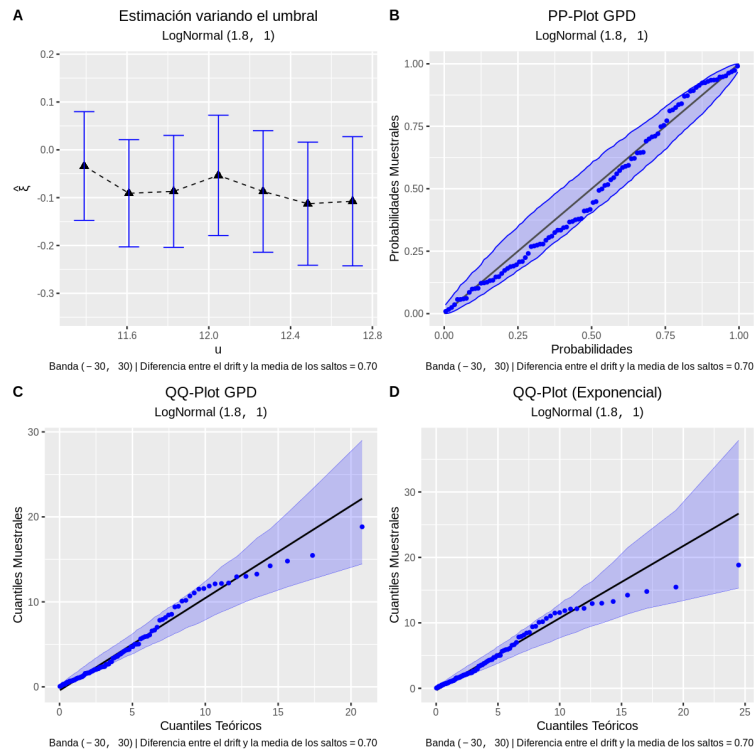


Figura 5. Bondad de ajuste del modelo DGP y Exponencial

Para dar más solidez a lo observado, una prueba de hipótesis indica que no hay evidencia para rechazar el dominio exponencial. De manera similar al caso anterior, y graficando las tasas exponenciales estimadas contra los inversos de las escalas de la DGP estimadas, parece que el ajuste se encuentra razonablemente cerca de la recta identidad, lo que da más fuerza a la hipótesis que habíamos planteado inicialmente.

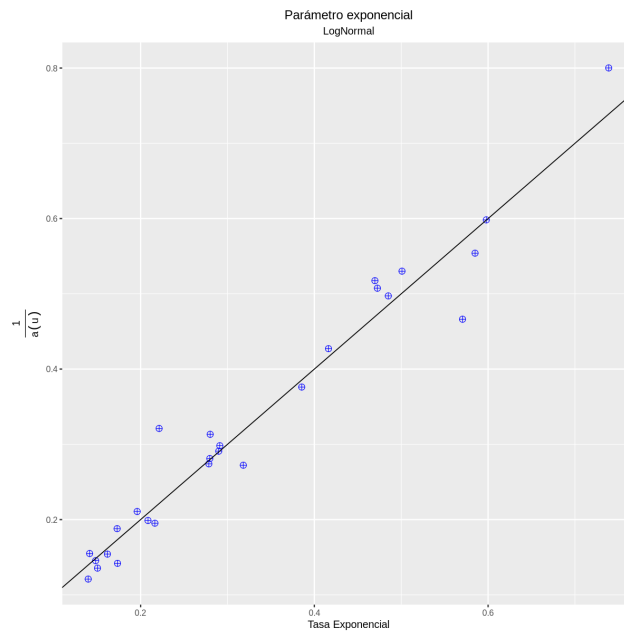


Figura 6. Tasas exponenciales estimadas contra el inverso de la escala DGP.

Resultados reclamos Weibull

La lógica de análisis para este caso sigue el mismo patrón que en los dos casos anteriores. Nuevamente, parece haber evidencia para decir que la distribución de la que provienen los datos tiene extremo derecho infinito. Dado que e_F parece en efecto estabilizarse con el tiempo, no hay evidencia en contra de que la distribución sea de cola ligera. Y el hecho de que $e_F(u)/u$ tienda a cero, nos lleva a suponer que es posible que los datos pertenezcan al dominio Gumbel.

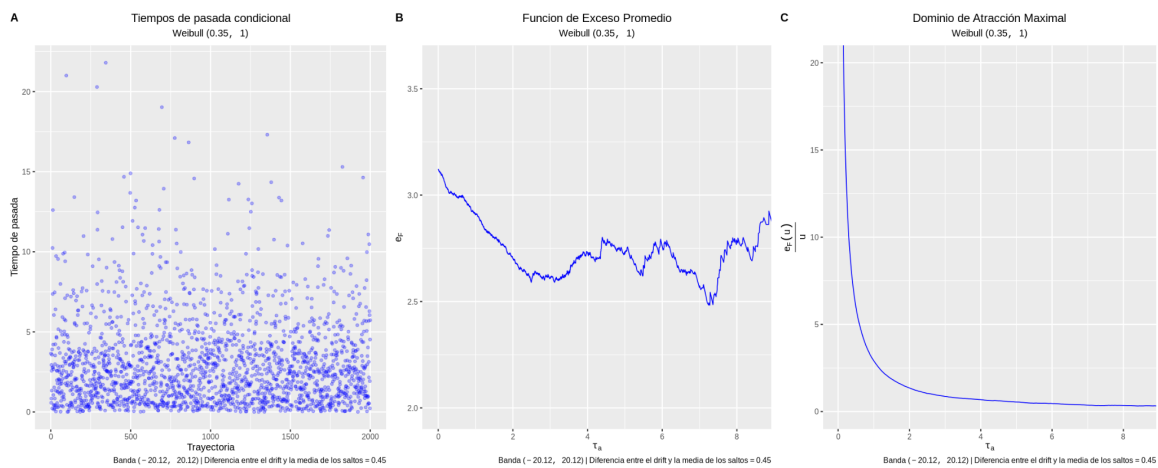


Figura 7. Análisis Exploratorio Distribución Weibull con media 5.03.

De hecho, para este caso, visualmente las gráficas de bondad de ajuste se comportan de manera particularmente buena: en la gráfica PP del ajuste con la DGP los puntos no sólo caen en la región de confianza graficada alrededor de la recta identidad, sino que yacen particularmente cercanos a la recta identidad; en el caso de ambas gráficas QQ, no sólo se mantienen cercanos los puntos a la recta identidad y dentro de la región de confianza, si no que no se desvían tanto en el extremo derecho. De los tres comentados, es este caso en el que las gráficas de bondad de ajuste tanto para la DGP como la exponencial se comportan mejor.

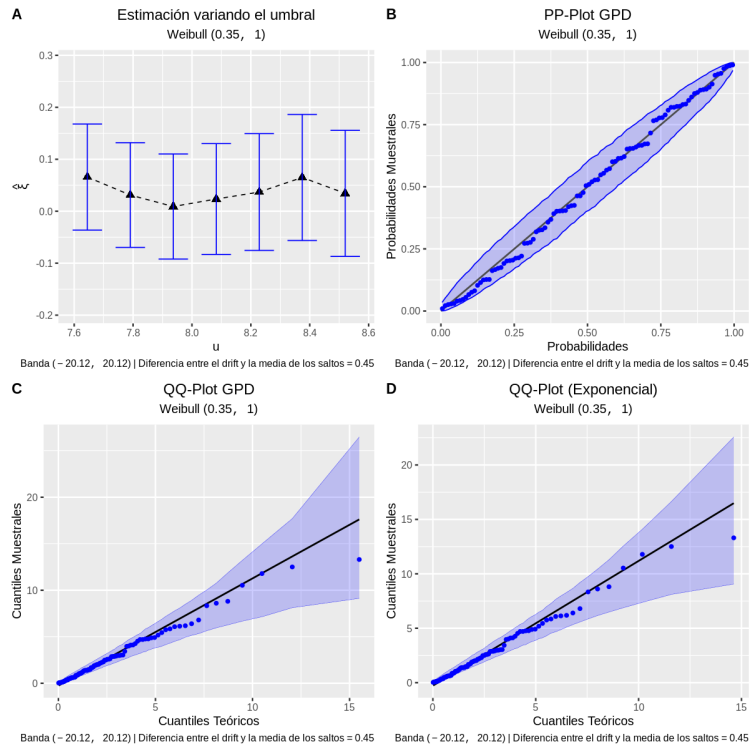


Figura 8. Bondad de ajuste del modelo DGP y Exponencial.

Finalmente, en la Figura 9, observamos una cercanía razonable entre la recta identidad y la mayoría de los puntos correspondientes a las simulaciones. Lo cual nuevamente es buen indicador sobre el planteamiento que hicimos de la distribución de τ_a condicionado a que $\{\tau_a < \tau_b\}$.

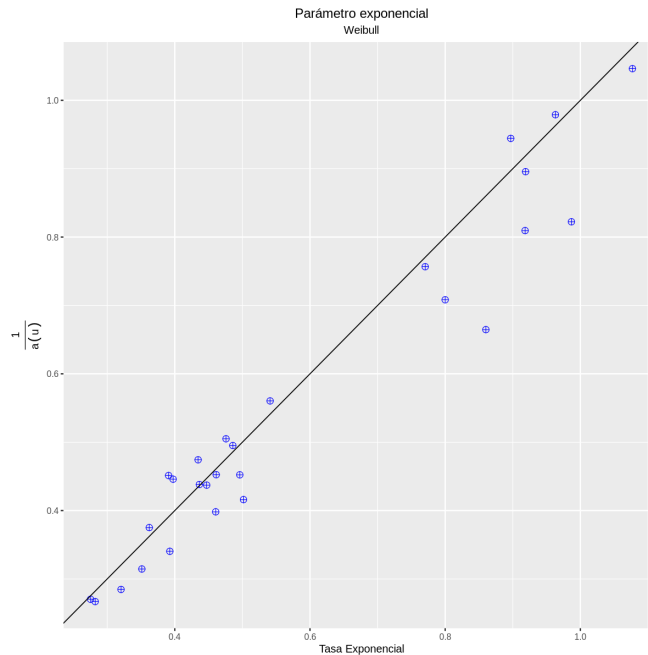


Figura 9. Tasas exponenciales estimadas contra el inverso de la escala DGP.

Conclusiones y Trabajo Futuro

- El análisis numérico presenta evidencia de que la distribución del tiempo de pasada τ_a condicionado a $\{\tau_a < \tau_b\}$, pertenece al mismo dominio de atracción que los saltos considerados para el proceso de Crámer-Lundberg, esto es, el dominio de atracción Gumbel.
- Lo anterior implica que podemos aproximar la distribución de interés utilizando una DGP, en particular una distribución exponencial la cual corresponde con el caso en que el parámetro $\xi = 0$.
- En el análisis, el ancho de banda es también un factor relevante a considerar. En general, bajo el mismo esquema de simulación, una amplitud del ancho de entre barreras más o menos estrecha puede modificar ligeramente la bondad de ajuste.
- El método utilizado para el análisis es simple, por lo que puede ser replicado para procesos estocásticos más generales.

Como trabajo futuro, nos interesa la construcción de la demostración formal que refleje los resultados de la aproximación observada numéricamente, indicando condiciones precisas para que esta se cumpla. Además de esto, se desea estudiar el comportamiento de los casos en que la distribución de los saltos pertenece a los otros dos dominios de atracción restantes (Fréchet y Weibull), así como extender los resultados a procesos de Lévy más generales que puedan ser de utilidad no sólo en el contexto financiero.

Referencias

- Albrecher, H., Ivanovs, J., & Zhou, X. (2016). Exit identities for Lévy processes observed at Poisson arrival times. *Bernoulli*, 22(3), 1364–1382.
- Almeida, A., & Loy, A. (2021, April 23). An Introduction to qqplotr. Retrieved July 22, 2022, from <https://cran.r-project.org/web/packages/qqplotr/vignettes/introduction.html>
- Asmussen, S., & Albrecher, H. (2010). *Ruin Probabilities*. World Scientific.
- Avram, F., Biard, R., Dutang, C., Loisel, S., & Rabehasaina, L. (2014). A survey of some recent results on Risk Theory. *ESAIM: PROCEEDINGS*, 44(0), 322 - 337. ESAIM: Proceedings and Surveys. <https://doi.org/10.1051/proc/201444020>
- Biffis, E., & Kyprianou, A.E. (2009). A note on scale functions and the time value of ruin for Lévy insurance risk processes. *Insurance: Mathematics and Economics*, 46, 85–91.
- Biffis, E., & Morales, M. (2010). On a generalization of the EDPF to path-dependent penalties. *Insurance: Mathematics and Economics*, 46, 92-97.
- Coles, S. (2011). *An Introduction to Statistical Modeling of Extreme Values*. Springer.
- Cont, R., & Tankov, P. (2004). *Financial Modeling with Jump Processes*. Taylor & Francis.
- Dickson, D. C. M., & Hipp, C. (2001). On the time to ruin for Erlang(2) risk processes. *Insurance: Mathematics and Economics*, 29(3).
- Efron, B. (1972, Enero). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1-26.
- Furrer, H. (1998). Risk processes perturbed by an α -stable Lévy motion. *Scandinavian Actuarial Journal*, 1.
- Ghosh, S., & Resnick, S. (2010). A discussion on mean excess plots. *Stochastic Processes and their Applications*, 120(8), 1492–1517. Elsevier Science. 10.1016/j.spa.2010.04.002
- Griffin, P. S., & Maller, R.A. (2011). Stability of the exit time for Lévy processes. *Advances in Applied Probability*, (43), 712 - 734.
- Grzywny, T. (2019). First exit times from a bounded interval for Lévy processes.
- Heffernan, J.E., & Stephenson, A. C. (2018). Package 'ismev'. Retrieved July 23, 2022, from <https://cran.r-project.org/web/packages/ismev/ismev.pdf>
- Hieber, P. (2013). Hieber, P. (2013). First-exit times and their applications in default risk management. Technische Universität München Department of Mathematics. <https://mediatum.ub.tum.de/doc/1165519/>
- Kassambara, A. (2020). Package 'ggpubr'. Retrieved July 23, 2022, from <https://cran.r-project.org/web/packages/ggpubr/ggpubr.pdf>
- Kolkovska, E. T., & Martín-González, E. M. (2016). Gerber-Shiu functionals for classical risk processes perturbed by an α -stable motion. *Insurance: Mathematics and Economics*, 66, 22-28.
- Kyprianou, A. (2006). *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer.
- Kyprianou, A. E. (2014). *Fluctuations of Lévy Processes with Applications: Introductory Lectures*. Springer Berlin Heidelberg.
- Lindgren, G., Rootzen, H., & Leadbetter, M. R. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer New York.
- Lundberg, F. (1903). Approximerad framställning af sannolikhetsfunktionen: Återförsäkring af kollektiv risker. *Almqvist & Wiksell*.
- Pfaff, B. (n.d.). Package 'evir'. Retrieved July 22, 2022, from <https://cran.r-project.org/web/packages/evir/evir.pdf>
- Pickands, J. (1975). Statistical Inference using Extreme Order Statistics. *The Annals of Statistics*, 3(1), 119-131.
- Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1), 119-131. ISSN 0090-5364
- Rudnytsky, I. (2018, July 30). Package 'ruin'. Retrieved July 22, 2022, from <https://cran.r-project.org/web/packages/ruin/ruin.pdf>
- Taylor, H. M., & Karlin, S. (1998). *An Introduction to Stochastic Modeling* (H. E. Taylor, Ed.). Elsevier Science.
- Xiong, S., & Yang, W.-S. (2011). Ruin probability in the Cramér–Lundberg model with risky investments. *Stochastic Processes and their Applications*, 121(5), 1125 - 1137. Elsevier Science. <https://doi.org/10.1016/j.spa.2011.01.008>.