

## MEDIDAS DE RIESGO PARA RIESGO OPERACIONAL CON UN MODELO DE PÉRDIDA AGREGADA POISSON-LINDLEY

Agustín Hernández-Bastida<sup>1</sup>  
bastida@ugr.es

Pilar Fernández Sánchez  
pilarfs@ugr.es

Universidad de Granada

fecha de recepción: 09/09/2010  
fecha de aceptación: 21/11/2010

### Resumen

En este trabajo se considera la determinación de medidas de riesgo en riesgo operacional, es decir, la determinación de cuantiles de alto orden. Se considera la aproximación basada en la distribución de la pérdida dentro de la aproximación avanzada. Se calculan, y se comparan entre sí, las medidas de riesgo a partir de la distribución de la pérdida agregada y a partir de la distribución predictiva considerando como funciones estructura para los perfiles de riesgo las distribuciones Triangular y Gamma.

**Palabras clave:** Modelo de pérdida agregada; Distribución de Poisson-Lindley; Distribución triangular; Distribución gamma.

### Abstract

This paper considers the determination of the risk measures in Operational Risk, i.e. the determination of a high level quantile. The Loss Distribution Approach in the Advanced Measurement Approach is adopted. The risk measures, obtained from the aggregate loss distribution and from the predictive distribution are determined and compared, using the Triangular and Gamma distributions as structure functions of the risk profiles.

**Keywords:** Aggregate Loss Model; Poisson-Lindley distribution; Triangular distribution; Gamma distribution.

---

<sup>1</sup> Departamento de Métodos Cuantitativos para la Economía de la Empresa, Facultad de Ciencias Económicas y Empresariales, Universidad de Granada, Campus de Cartuja, s/n, 18011-Granada.

## 1. Introducción

En una entidad financiera el riesgo operacional se refiere a sucesos que no pueden ser considerados como riesgos de mercado o de crédito. Se refiere a las pérdidas debidas a la inadecuación o a fallos en los procesos, al personal o a los sistemas internos y a las pérdidas causadas por acontecimientos externos a la entidad. Un excelente trabajo en lengua española sobre este tema es Carrillo y Suárez (2006) del cual se ha tomado el siguiente párrafo (ver op. cit. pág. 63):

“La importancia del riesgo operacional ha sido reconocida recientemente en el mundo financiero. A consecuencia de su relevancia, el Comité de supervisión bancaria de Basilea ha incluido en el documento “Convergencia internacional de medidas y normas de capital – Marco revisado”, conocido como Basilea II, un apartado específico sobre riesgo operacional. En Basilea II se describen diferentes metodologías para el cálculo de una cifra de capital regulatorio, requerido para hacer frente a posibles pérdidas por este tipo de riesgo, y que tienen un grado de complejidad creciente. En el enfoque básico (Basic Indicator Approach) el capital regulatorio se calcula como un porcentaje fijo de un indicador de exposición al riesgo operacional (los ingresos brutos) para la entidad en su conjunto. En el enfoque estándar (Standardised Approach) se aplican distintos porcentajes, fijados por el regulador, a este indicador desagregado para cada una de las líneas de negocio del banco. Finalmente para el enfoque avanzado (Advanced Measurement Approaches, AMA), Basilea II establece unas directrices generales, pero permite e incentiva que los propios bancos construyan su propio modelo de medición y gestión del riesgo operacional. Se espera que la metodología

## 1. Introduction

In a financial firm, “operational risk” refers to the risk of losses resulting from inadequate or failed internal processes, people and systems, or external events. An excellent paper on this question was published by Carrillo and Suarez (2006), who observed (p. 63).

“The importance of the operational risk has been recognised. Due to its importance, the Basel Committee on Banking Supervision has included in the document entitled “International Convergence of Capital Measurement and Capital Standard: a revised framework, known as Basel II (see BIS, 2005), a specific section on operational risk, and different methodologies are described to calculate the amount of regulatory capital needed to deal with possible losses derived from this risk. These methodologies are presented in increasing levels of complexity. In the Basic Indicator Approach, the regulatory capital is obtained as a fixed percentage of an indicator of the entire firm’s exposure to operational risk. In the Standardised Approach, diverse percentages are applied to the indicator, one for each of the bank’s lines of business. Finally, in the Advanced Measurement Approach (AMA), Basel II establishes general rules but allows and encourages banks to develop their own model for assessing exposure to operational risk. This advanced approach is believed to reflect in a more detailed form the specific risk profile of the bank”.

desarrollada dentro del enfoque avanzado refleje de manera más detallada el perfil de riesgo específico de cada entidad”.

Concretando, bajo las exigencias de Basilea II a los bancos se les solicita que evalúen el capital asignado para el riesgo operacional. Diversos aspectos de la modelización del riesgo operacional pueden verse en Chavez-Demoulin, Embrechts y Neslehová (2006) y en Cruz (2002).

Muchas entidades financieras han adoptado la Loss Distribution Approach (LDA) en la metodología AMA. En esta aproximación se deben estimar, en el horizonte de un año, la frecuencia y la severidad de las pérdidas operacionales para cada una de las celdas de riesgo (en la matriz de ocho líneas de negocio y siete tipos de sucesos). A partir de ello se estima el capital usando el cuantil de orden 0.999 de la distribución de la pérdida total. El modelo habitualmente usado para la pérdida anual en una celda de riesgo es un proceso compuesto  $S = X_1 + \dots + X_K$ , donde K es el número anual de sucesos modelado como una variable aleatoria con alguna distribución de probabilidad discreta y  $X_i; i = 1, 2, \dots$  son las severidades de los sucesos modeladas como variables aleatorias independientes con alguna distribución de probabilidad continua.

Con el objetivo de determinar el capital regulatorio para el riesgo operacional, se debe determinar la distribución de la pérdida total (en especial una medida de riesgo que es el cuantil de orden 0.999) para cada una de las celdas de riesgo (tipo de suceso / línea de negocio) de la entidad financiera.

Under the Basel II requirements, banks are required to quantify the capital charge for operational risk. Various aspects of the modelling of operational risk are considered in Chavez-Demoulin, Embrechts and Neslehová (2006) and Cruz (2002). Many banks have adopted the Loss Distribution Approach (LDA) in the AMA, where the frequency and severity of operational losses for each risk cell (in the matrix of eight business lines and seven event types) are estimated over a one year period. The capital is then estimated using the 0.999 quantile of the distribution for the total annual loss in the bank. The most commonly used model for the annual loss in a single risk cell is the compound process  $S = X_1 + \dots + X_K$ , where K is the annual number of events modelled as a random variable from a discrete distribution and  $X_i; i = 1, \dots, K$ , are the severities of the event, modelled as independent random variables from a continuous distribution.

For the purposes of the regulatory capital calculation of operational risk, the annual loss distribution (in particular its 0.999 quantile as a risk measure) should be quantified for each risk cell (event type / business line) in the bank.

La especificación de las distribuciones de probabilidad de la frecuencia y de la severidad es una tarea complicada debido entre otras cosas a la baja frecuencia de pérdidas de gran volumen. Habitualmente los datos internos (usualmente truncados) de las entidades financieras solo están disponibles para unos pocos años y contienen pocas (o ninguna) de las pérdidas poco frecuentes y de gran volumen. Los datos externos (pérdidas sufridas por otras entidades financieras) pueden estar disponibles a partir de bases de datos de terceras personas, pero son difíciles de usar debido a diversos factores. En definitiva es difícil estimar las distribuciones requeridas usando solamente datos.

También es claro que una estimación usando datos del pasado tendrá una capacidad limitada para predecir el futuro debido al constante cambio del entorno de una entidad financiera; si se introduce una nueva política en la entidad financiera para disminuir las pérdidas por riesgo operacional esto no puede ponerse de manifiesto en un modelo basado exclusivamente en los datos de pérdidas. Es muy importante incorporar un análisis del entorno en el modelo, pero claro esto puede ser muy subjetivo y además deberá combinarse con (y basarse en) los datos actuales de pérdidas.

La inferencia bayesiana es una técnica estadística muy apropiada para combinar opiniones de expertos con datos observados. En este artículo se adopta un punto de vista bayesiano para el problema que nos ocupa.

En Shevchenko y Wüthrich (2006) se propone una forma para determinar el capital regulatorio que consiste en calcular la distribución a posteriori del número de reclamaciones y de las severidades; a partir

Estimating frequency and severity distributions is a challenging task, especially for low frequency high impact losses. Internal bank data (usually truncated) are usually available with respect to several years and contain few (or no) high impact, low frequency losses. External data (losses experienced by other banks) are available through third party databases, but these are difficult to use directly due to different volumes and other factors. Moreover, the data have a survival bias, because the data from all collapsed companies are not usually available, and so it is difficult to estimate distributions using only these data. Moreover, this estimation is backward looking and has a limited ability to predict the future, due to a constantly changing banking environment. For example, if a new policy is introduced by a bank, with the aim of decreasing operational risk losses, this would not be reflected by a model based only on the loss data.

It is very important to have a scenario analysis incorporated into the model. In fact, this is mandatory under regulatory requirements. In itself, the scenario analysis is very subjective and should be combined with (supported by) the actual loss data analysis. Bayesian inference is a statistical technique that is well suited to incorporating expert opinions into data analysis, and so a Bayesian analysis is adopted in this paper.

Shevchenko and Wüthrich (2006) proposed that the regulatory capital should be calculated by determining the posterior distributions for the number of claims and for the severities. A value for the parameters of the number of claims

de ella se simulan valores para el parámetro de la distribución del número de reclamaciones y para el parámetro de la distribución de las severidades y utilizando estos valores simulados para los parámetros se determina en la distribución de la pérdida agregada el cuantil de orden 0.999. Una metodología similar se propone en Lambrigger, Shevchenko y Wüthrich, (2007, 24). Tanto en esta metodología como cuando se usan estimadores puntuales de los parámetros (por ejemplo estimadores de máxima verosimilitud) lo que sucede, tal y como señala Shevchenko (2008), es que se ignora la aleatoriedad de los parámetros (ver trabajo indicado para detalles); por esta razón es por la que el autor citado propone la determinación de la distribución predictiva y a partir de ella calcular el cuantil de orden 0.999.

Esta es la aproximación que consideramos en este artículo, por tanto nuestro objetivo es el cálculo de la distribución predictiva y a continuación determinar el cuantil de orden 0.999.

Concretando, en este artículo se considera el modelo colectivo de riesgo (de ahora en adelante crm) con distribución primaria una distribución de Poisson-Lindley de parámetro  $\theta_1$  (de aquí en adelante el modelo PL) y con distribución secundaria una distribución Exponencial de parámetro  $\theta_2$  (de aquí en adelante el modelo E). Se desarrolla un análisis bayesiano del modelo colectivo de riesgo (de ahora en adelante crmPLE) suponiendo para el perfil de riesgo  $\theta_2$  una función estructura dada por una distribución Gamma; para el perfil de riesgo  $\theta_1$  una elección frecuente para la función estructura es la distribución Beta (ver por ejemplo, Hernández-Bastida, Fernández-Sánchez y Gómez-Déniz, 2010), en este artículo se considera una distribución Triangular.

and the severities was then simulated. From these values, the aggregate loss was simulated and its 0.999 quantile determined. A similar methodology was proposed in Lambrigger, Shevchenko and Wüthrich (2007). Both in these methodologies and when the capital is estimated by point estimators of the parameters (e.g. Maximum Likelihood Estimators) what happens is that, as Shevchenko (2008) pointed out, although the parameter estimates are uncertain, in practice this uncertainty is commonly ignored in the estimation of the operational risk charge (*op. cit.*). For this reason, in the present paper we propose to make use of the predictive function and the determination of the 0.999 quantile in the distribution. This is the line followed in this paper, and so our aim is to determine the predictive function and its 0.999 quantile.

In this paper, the Collective Risk Model (henceforth, the crm) in which the primary distribution is the Poisson-Lindley distribution, with parameter  $\theta_1$  (henceforth, the PL model) and the Exponential distribution, with parameter  $\theta_2$ , as the secondary one (henceforth, the E model). In the Bayesian analysis developed in the crm indicated (henceforth, the crmPLE), we consider for the risk profile  $\theta_2$  a structure function given by the Gamma distribution. For the risk profile  $\theta_1$ , a common choice for the structure function is the Beta distribution, see for example Hernández-Bastida, Fernández-Sánchez and Gómez-Déniz (2010). In the present paper, the Triangular distribution is considered, and the following aims are addressed:

Los objetivos de este artículo son:

- (i) Calcular cuantiles de alto orden en la distribución predictiva del crmPLE.
- (ii) Calcular cuantiles de alto orden en la distribución de la pérdida agregada con los parámetros estimados por la moda a posteriori.
- (iii) Poner de manifiesto las divergencias que se presentan entre (i) y (ii).

La estructura del artículo es la siguiente: en la sección 2 se desarrolla el modelo de pérdida agregada que se va a considerar, en la sección 3 se calculan las distribuciones marginal y a posteriori y en la sección 4 se calculan las medidas de riesgo y se hace un análisis comparativo.

**Notación**

En el resto del trabajo será muy cómodo utilizar las siguientes funciones. Sean  $\varepsilon_1, \varepsilon_2, \varepsilon_3$  números enteros positivos,  $a$  un número real entre 0 y  $c$  y  $d$  números reales y positivos. Escribimos,

$$J_2(\varepsilon_1, \varepsilon_2, \varepsilon_3 | s, a, c, d) = \int_0^a \frac{\theta_1^{1+\varepsilon_1} (1-\theta_1)^{\varepsilon_2}}{(s\theta_1 + d)^{c+\varepsilon_3}} d\theta_1; \quad (1)$$

$$J_3(\varepsilon_1, \varepsilon_2, \varepsilon_3 | s, a, c, d) = \int_a^1 \frac{\theta_1^{\varepsilon_1} (1-\theta_1)^{1+\varepsilon_2}}{(s\theta_1 + d)^{c+\varepsilon_3}} d\theta_1. \quad (2)$$

Es fácil ver que las dos integrales se pueden determinar con un cálculo directo pero tedioso. La primera de ellas se puede escribir de la siguiente forma,

$$J_2 = \sum_{i=0}^{\varepsilon_2} (-1)^i \binom{\varepsilon_2}{i} \int_0^a \frac{\theta_1^{1+\varepsilon_1+i}}{(s\theta_1 + d)^{c+\varepsilon_3}} d\theta_1$$

y haciendo el cambio  $u = s\theta_1 + d$  se obtiene,

$$J_2 = \sum_{i=0}^{\varepsilon_2} (-1)^i \binom{\varepsilon_2}{i} \frac{2}{s^{2+\varepsilon_1+i}} \sum_{j=0}^{1+\varepsilon_1+i} (-1)^j d^j \binom{1+\varepsilon_1+i}{j} \int_d^{d+as} u^{1+\varepsilon_1-\varepsilon_3-c+i-j} du. \quad (3)$$

- (i) To obtain the high order quantile for the predictive function in the crmPLE;
- (ii) To obtain the high order quantile for the aggregate loss with the parameter fixed in the bivariate posterior mode;
- (iii) To show the divergences between the quantiles obtained in (i) and (ii).

The rest of the paper is structured as follows. Section 2 develops the aggregate loss model to be considered. The marginal and posterior distributions are determined in Section 3. The risk measures are derived in Section 4, and a comparative analysis performed.

**Notation**

For the specification of several important magnitudes in this paper, the following functions are defined:

Let  $\varepsilon_1, \varepsilon_2, \varepsilon_3$  be any positive integers,  $a$  a real number between 0 and 1 and  $c, d$  positive and real numbers. Note,

It can be seen that both integrals can be obtained by direct but tedious calculus. The first one can be written in the following form, and its calculus obtained immediately,

making  $s\theta_1 + d = u$ , and hence,

La segunda integral puede escribirse como,

The second integral can be calculated from this expression,

$$J_3 = \sum_{i=0}^{1+\varepsilon_2} (-1)^i \binom{1+\varepsilon_2}{i} \int_a^1 \frac{\theta_1^{\varepsilon_1+i}}{(s\theta_1+d)^{c+\varepsilon_3}} d\theta_1$$

y haciendo el mismo cambio de antes  $u = s\theta_1 + d$  se obtiene,

making the same change,  $s\theta_1 + d = u$ , and hence,

$$J_3 = \sum_{i=0}^{1+\varepsilon_2} (-1)^i \binom{1+\varepsilon_2}{i} \frac{1}{s^{1+\varepsilon_1+i}} \sum_{j=0}^{\varepsilon_1+i} (-1)^j d^j \binom{\varepsilon_1+i}{j} \int_{as+d}^{s+d} u^{\varepsilon_1-\varepsilon_3-c+i-j} du. \quad (4)$$

también usaremos la siguiente función,

We also use the next function,

$$J_1(\varepsilon_1, \varepsilon_2, \varepsilon_3 | s, a, c, d) = \frac{2d^c \Gamma(c + \varepsilon_3)}{a \Gamma(c)} J_2 + \frac{2d^c \Gamma(c + \varepsilon_3)}{(1-a) \Gamma(c)} J_3. \quad (5)$$

## 2. El modelo de pérdida agregada

## 2. The aggregate loss model

El modelo colectivo de riesgo consiste en una distribución de frecuencias para el número de reclamaciones  $K$  y una secuencia de variables aleatorias independientes e idénticamente distribuidas que representan el tamaño de las severidades individuales  $X_i$ ; se supone que la frecuencia  $K$  y las severidades  $X_i$  son independientes, independencia que es condicionada a la distribución de los parámetros. Hay una amplia literatura sobre la modelización de procesos de riesgo, ver por ejemplo McNeil, Frey y Embrechts (2005).

The Collective Risk Model (crm) is described by a frequency distribution for the number of claims  $K$  and a sequence of independent and identically distributed random variables representing the size of the single claims  $X_i$ . Frequency  $K$  and severities  $X_i$  are assumed to be independent. Note that the independence assumed here is conditional on the distribution parameters. There is an extensive bibliography on risk process modelling, see e.g. McNeil, Frey and Embrechts (2005).

La estimación de la distribución de la pérdida agregada a partir de las modelizaciones de la frecuencia y de la severidad es una técnica actuarial muy conocida que también se usa en modelos de solvencia en la industria del seguro, ver por ejemplo Sandström (2006) o Wüthrich (2006).

Estimation of the annual loss distribution by modelling the frequency and severity of losses is a well known actuarial technique, which is also used for model solvency requirements in the insurance industry, see e.g. Sandström (2006) and Wüthrich (2006).

La pérdida agregada  $S$  es la suma de las pérdidas individuales, es decir,  $S = X_1 + X_2 + \dots + X_K$ , para  $K > 0$ , y  $S = 0$ , para  $K = 0$ . Una modelización muy desarrollada en crm es considerar que la distribución primaria, es decir la distribución de  $K$  es una distribución de Poisson y la distribución secundaria, es decir la distribución de la severidad, es una distribución Exponencial. Resultados interesantes sobre una suma de variables Exponenciales con número de sumandos aleatorio pueden encontrarse en Kozubowski y Panoska (2005) y un amplio catálogo de formas aproximadas para el modelo compuesto de Poisson aparecen en Nadarajah y Kotz (2006a; 2006b). Por otra parte, es conocido que en muchos conjuntos de datos en los problemas que nos ocupan se presenta una varianza más grande que la media (fenómeno conocido como sobredispersión) y por esta razón se han considerado diferentes distribuciones alternativas para la variable aleatoria  $K$ ; en especial, las distribuciones mixturas de la Poisson, ver entre otros Grandell (1997) o Nikoloulopoulos y Karlis (2008).

En este artículo se realiza un análisis bayesiano en el crm con distribución primaria una distribución de Poisson-Lindley y con distribución secundaria una distribución Exponencial; un compendio de diferentes modelos bayesianos en diversas áreas de las matemáticas actuariales puede verse en Schmidt (1998).

Sea  $K$  la variable aleatoria "número de pérdidas", suponemos que sigue una distribución de Poisson-Lindley de parámetro  $\lambda > 0$ , con función de probabilidad,

The aggregate loss  $S$  is the sum of the individual claim sizes, i.e.  $S = X_1 + X_2 + \dots + X_K$ , for  $K > 0$ , and  $S = 0$ , for  $K = 0$ . A well developed model of the crm is to consider that the primary distribution, i.e., the  $K$  distribution, is the Poisson distribution, and the secondary one, the claim severity, is the Exponential distribution. Kozubowski and Panoska (2005) presented a set of interesting results for the sum of random variables with an Exponential distribution and a random number of summands. In addition, a comprehensive collection of approximate forms for the compound mixed Poisson distribution is to be found in Nadarajah and Kotz (2006a; 2006b). At the same time, there is known to exist a remarkable difficulty in handling so many data sets, in which the variance is larger than the average (overdispersion). For this reason, various alternative distributions have been considered for the random variable  $K$ , especially the mixed Poisson distributions, see for example Grandell (1997) and Nikoloulopoulos and Karlis (2008), among others.

In this paper, a Bayesian analysis in the crm is considered, in which the primary distribution is the Poisson Lindley distribution and the Exponential is the secondary distribution, denoted as crmPLE. A unifying survey of Bayesian models in different areas of actuarial mathematics is presented in Schmidt (1998).

Let  $K$  be the random variable "number of claims" and let us assume it has a Poisson-Lindley distribution with parameter  $\lambda > 0$ , and the probability function,

$$\Pr[K = k / \lambda] = \lambda^2 (\lambda + 2 + k) / (\lambda + 1)^{k+3}; k = 0, 1, \dots$$



Recordemos que la distribución de Poisson-Lindley es una distribución de probabilidad discreta propuesta por Sankaran (1971) que se obtiene como una mixtura de la distribución de Poisson con la distribución continua de Lindley (Lindley (1958); esta distribución ha recibido notable atención en diversos trabajos, ver por ejemplo Karlis y Xekalaki (2005), Ghitany y Al-Mutairi (2008) o Ghitany, Al-Mutairi y Nadarajah (2009). Con la reparametrización  $\theta_1 = \lambda/(\lambda+1)$  el modelo PL adopta la forma,

The Poisson-Lindley distribution is a discrete probability mass function, as proposed in Sankaran (1971) and obtained by mixing a Poisson distribution with the continuous Lindley distribution (Lindley, 1958). Studies addressing the use of this one-parameter discrete distribution include Karlis and Xekalaki (2005) and, recently, Ghitany and Al-Mutairi (2008) and Ghitany, Al-Mutairi and Nadarajah (2009). With a new parametrization,  $\theta_1 = \lambda/(\lambda+1)$ , the PL model follows,

$$f_{PL}[k/\theta_1] = \theta_1^2 (1-\theta_1)^k [2-\theta_1 + (1-\theta_1)k]; k=0, 1, 2, \dots \theta_1 \in (0,1).$$

La función generatriz de momentos viene dada por,

Its moment generating function is

$$M_1^{PL}(t; \theta_1) = [\theta_1^2 (2-\theta_1 - e^t + \theta_1 e^t)] / (1 - e^t + \theta_1 e^t)^2,$$

el valor de la esperanza es,

and its expected value is,

$$E_{PL}[K] = [(2-\theta_1)(1-\theta_1)] / \theta_1.$$

El momento de segundo orden es,

The second order moment is given by

$$E_{PL}[K^2] = [(1-\theta_1)(3\theta_1^2 - 8\theta_1 + 6)] / \theta_1^2$$

y la varianza es,

and its variance is

$$Var_{PL}[K] = [(1-\theta_1)(\theta_1^3 - 2\theta_1^2 + 2)] / \theta_1^2.$$

Es conocido que la distribución es sobredispersa, es decir,

Moreover, it is well known that the distribution is overdispersed, i.e.

$$Var_{PL}[K] / E_{PL}[K] = \theta_1^3 - 2\theta_1^2 + 2 / \theta_1 (2-\theta_1) > 1.$$

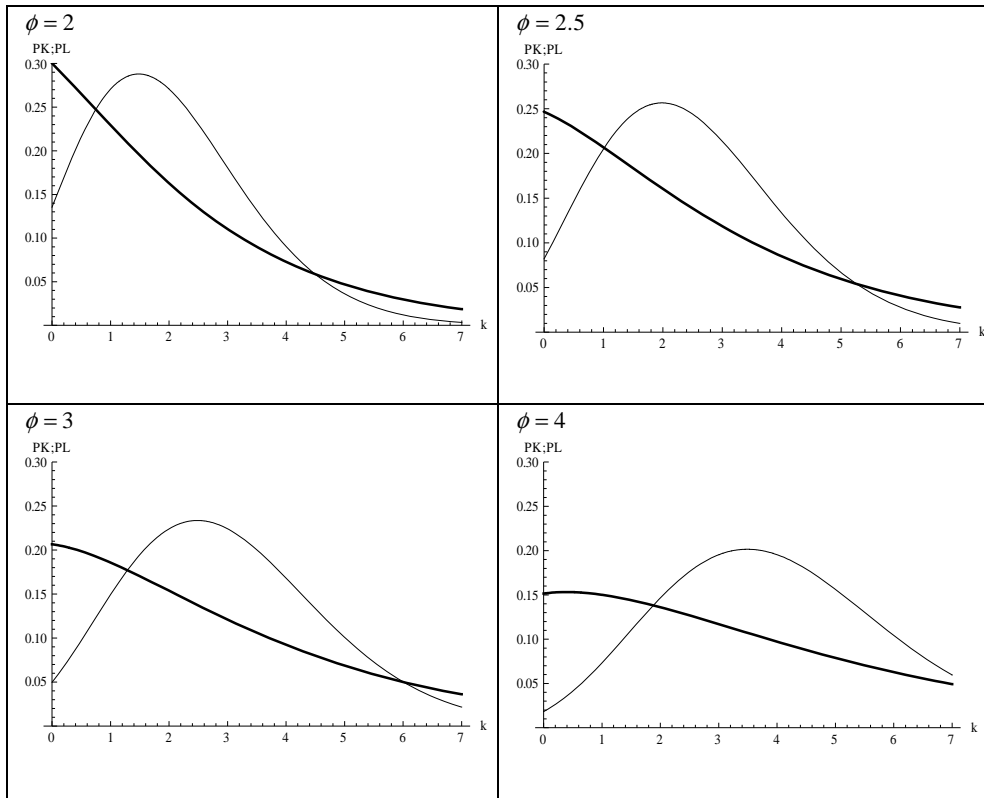
La siguiente figura muestra, para diversos valores de la media  $\phi$ , la diferencia entre las probabilidades de ambas distribuciones, usando la línea gruesa para

The following figure shows for several values of the mean  $\phi$  the difference between the probabilities of the two distributions, where the thick line refers to

referirnos a las probabilidades de la Poisson-Lindley. Se observa que en todas las situaciones, la distribución PL asigna una masa de probabilidad mayor al valor  $k = 0$  y tiene colas más pesadas.

the Poisson-Lindley probabilities. It is clear that in all situations the PL distribution assigns a higher probability to the value  $k = 0$ , and that it has heavy tails.

Figura 1/ Figure 1



Supongamos que la variable aleatoria  $X_i$ , “severidad de las pérdidas”, sigue una distribución Exponencial con parámetro  $\theta_2 > 0$ , el modelo E, y tiene como función de densidad  $f_E(x/\theta_2) = \theta_2 e^{-\theta_2 x}$ , para  $x > 0$ . Su función generatriz de momentos es  $M_2^E(t; \theta_2) = \theta_2 / (\theta_2 - t)$ ; y los valores de su media y de su varianza son,

Let  $X_i$  be the random variable “claim severity”, which is to be an Exponential distribution with parameter  $\theta_2 > 0$ , the E model, and density function,  $f_E(x/\theta_2) = \theta_2 e^{-\theta_2 x}$ , for  $x > 0$ . Its moment generating function is  $M_2^E(t; \theta_2) = \theta_2 / (\theta_2 - t)$ ; and the expected value and the variance are

respectivamente,  $E_E[X] = 1/\theta_2$  y  $Var_E[X] = 1/\theta_2^2$ .  $E_E[X] = 1/\theta_2$  and  $Var_E[X] = 1/\theta_2^2$ , respectively.

**Proposición 1** Se verifican las siguientes afirmaciones:

(i) La función de densidad de la variable aleatoria pérdida agregada total,  $S$ , crmPLE viene dada por,

**Proposition 1** These affirmations are verified:

(i) The probability density function of the total loss random variable,  $S$ , the crmPLE model is given by,

$$f_{PLE}(s/\theta_1, \theta_2) = \begin{cases} \theta_1^2 (1-\theta_1) \theta_2 [3-2\theta_1 + (1-\theta_1)^2 \theta_2 s] e^{-\theta_1 \theta_2 s}; & s > 0 \\ \theta_1^2 (2-\theta_1); & s = 0 \end{cases}$$

en la que aparece la discontinuidad habitual del crm en  $s=0$ .

(ii) La función generadora de momentos de  $S$ , que notamos  $M_3^{PLE}(t; \theta_1, \theta_2)$ , viene dada por,

with the usual discontinuity in the crm appearing at  $s=0$ .

(ii) The moment generating function of  $S$ , denoted as  $M_3^{PLE}(t; \theta_1, \theta_2)$ , is given by

$$M_3^{PLE}(t; \theta_1, \theta_2) = \theta_1^2 [(2-\theta_1)t^2 + (\theta_1 \theta_2 - 3\theta_2)t + \theta_2^2] / (\theta_1 \theta_2 - t)^2.$$

Se deduce que,

It follows that

$$\begin{aligned} E_{PLE}[S] &= [\theta_1^2 - 3\theta_1 + 2] / \theta_1 \theta_2; \\ E_{PLE}[S^2] &= [-2\theta_1^3 + 8\theta_1^2 - 9\theta_1 + 6] / \theta_1^2 \theta_2^2; \\ Var_{PLE}[S] &= [-\theta_1^4 + 4\theta_1^3 - 5\theta_1^2 + 3\theta_1 + 2] / \theta_1^2 \theta_2^2. \end{aligned}$$

**Demostración.**

La afirmación (i) se obtiene directamente sin más que observar que la convolución  $n$ -ésima de la Exponencial es  $\theta_2^n x^{n-1} e^{-\theta_2 x} / (n-1)!$ ; a partir de ello la función de densidad  $f_{PLE}(s/\theta_1, \theta_2)$  se obtiene de,

**Proof.**

Affirmation (i) is obtained directly noting that the convolution of  $n$  Exponential distributions is given by  $\theta_2^n x^{n-1} e^{-\theta_2 x} / (n-1)!$  Therefore, the density function  $f_{PLE}(s/\theta_1, \theta_2)$  is given by

$$\begin{aligned} f_{PLE}(s/\theta_1, \theta_2) &= \sum_{n=1}^{\infty} \frac{\theta_2^n s^{n-1} e^{-s\theta_2}}{(n-1)!} \theta_1^2 (1-\theta_1)^n [2-\theta_1 + (1-\theta_1)n] = \\ &= [2\theta_1^2 (1-\theta_1)\theta_2 - \theta_1^3 (1-\theta_1)\theta_2 + s\theta_1^2 (1-\theta_1)^3 \theta_2^2 + \theta_1^2 (1-\theta_1)^2 \theta_2] e^{-s\theta_2} \sum_{n=0}^{\infty} \frac{[s(1-\theta_1)\theta_2]^n}{n!} \end{aligned}$$

La afirmación (ii) se obtiene del siguiente resultado muy conocido, (ver Klugman, Panjer y Willmot, 2004),

$$M_3^{PLE} = M_1^{PL} (\log M_2^E).$$

La Figura 2 muestra la comparación entre la función de verosimilitud de crmPLE, notada  $f_{PLE}(s/\theta_1, \theta_2)$  y la verosimilitud de crmPE,  $f_{PE}(s/\phi, \theta_2)$ , para el mismo valor de la media. La línea gruesa corresponde a crmPLE.

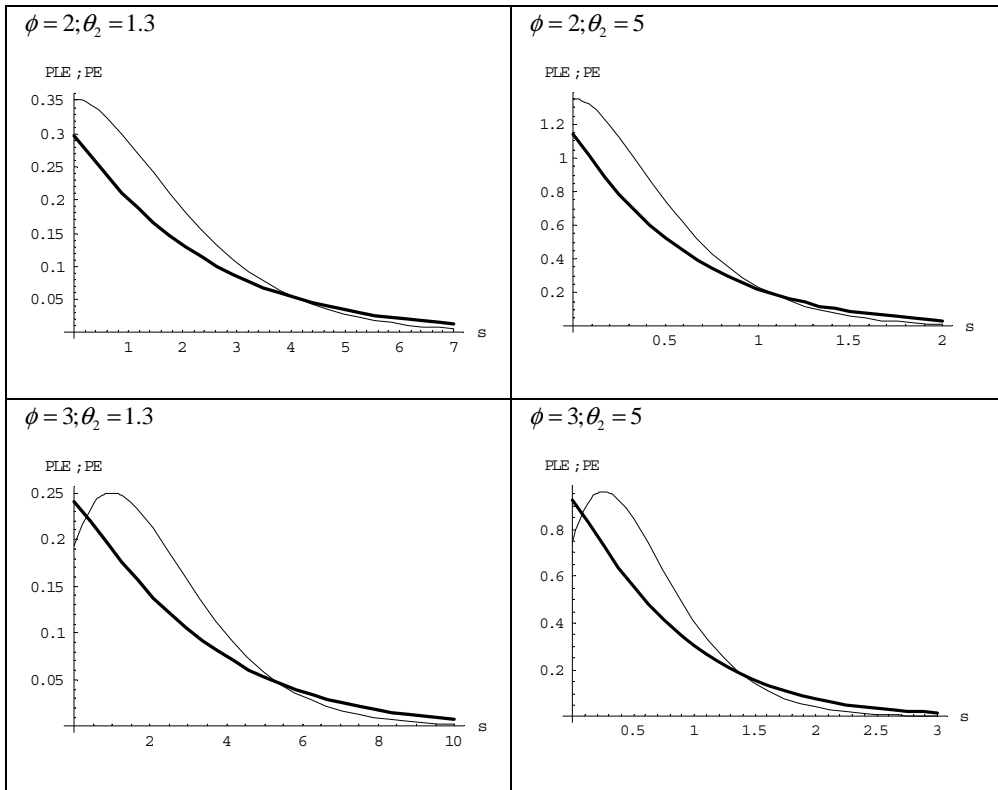
Affirmation (ii) is derived from the following well known result,

$$M_3^{PLE} = M_1^{PL} (\log M_2^E),$$

(see Klugman, Panjer and Willmot, 2004).

Figure 2 shows the comparison for the same mean value between the likelihood function in the crm.PLE, denoted as  $f_{PLE}(s/\theta_1, \theta_2)$ , and in the crm.PE,  $f_{PE}(s/\phi, \theta_2)$ . The thick line corresponds to the crm.PLE.

Figura 2/ Figure 2



Desde un punto de vista bayesiano los parámetros de interés en el problema han de ser estimados usando nuestro grado de conocimiento sobre ellos; aquí, consideraremos las siguientes distribuciones a priori para  $\theta_1$  y para  $\theta_2$ . Una elección natural de la distribución a priori para  $\theta_2$  es la distribución Gamma con parámetros  $c$  y  $d$ ,  $G(\theta_2, c, d) = (d^c / \Gamma(c)) \theta_2^{c-1} e^{-d\theta_2}$ ;  $\theta_2 \in \mathbb{R}^+$ ;  $c, d \in \mathbb{R}^+$ , en la que supondremos  $c > 1$ . La media y varianza a priori viene dada por  $E_G[\theta_2] = c/d$  y  $Var_G[\theta_2] = c/d^2$ , respectivamente y la moda es  $Mo_G[\theta_2] = (c-1)/d$ .

En este artículo usaremos la distribución Triangular como distribución a priori alternativa a la distribución Beta para el parámetro  $\theta_1$ . La distribución Triangular fue propuesta por Johnson (1997) como una aproximación sencilla (proxy) a la distribución Beta y analizó las diferencias entre ambas distribuciones. La función de densidad de la distribución Triangular  $T(\theta_1, a)$ , es la siguiente,

$$T(\theta_1, a) = \begin{cases} \frac{2\theta_1}{a} & 0 \leq \theta_1 \leq a \\ \frac{2(1-\theta_1)}{1-a} & a \leq \theta_1 \leq 1 \end{cases}$$

La distribución tiene una moda en  $a$ ,  $T(a, a) = 2$ , y puede mostrar asimetría positiva o negativa. La media es  $E_T(\theta_1) = (a+1)/3$ , es creciente en el valor de la moda y varía de 0.33 a 0.67. La varianza,  $V_T(\theta_1) = (1-a-a^2)/18$ , es igual a 0.056 cuando  $a = 0$ ; decrece si  $0 < a < 0.5$ , alcanzando un mínimo en  $a = 0.5$ , que es 0.042, y crece si  $0.5 < a < 1$  alcanzando de nuevo el valor 0.056 cuando  $a = 1$ .

Under a Bayesian point of view, the parameters of interest in this problem can be estimated using our state of knowledge about them. Let us consider the following prior distributions for  $\theta_1$  and for  $\theta_2$ .

The natural choice of a prior pdf for  $\theta_2$  is the Gamma distribution with parameters  $c$  and  $d$  (we assume  $c > 1$ ),  $G(\theta_2, c, d) = (d^c / \Gamma(c)) \theta_2^{c-1} e^{-d\theta_2}$ ;  $\theta_2 \in \mathbb{R}^+$ ;  $c, d \in \mathbb{R}^+$ , for  $c > 1$ . The prior mean and variance for  $\theta_2$  are given by  $E_G[\theta_2] = c/d$  and  $Var_G[\theta_2] = c/d^2$ , respectively, and the mode is  $Mo_G[\theta_2] = (c-1)/d$ .

In this paper we use a prior distribution, as an alternative to the Beta distribution for  $\theta_1$ , which is the Triangular distribution. Johnson (1997) proposed the Triangular distribution as a proxy for the Beta distribution, and analyzed the differences between the two distributions. The density function of the Triangular distribution,  $T(\theta_1, a)$ , is given by

It has a mode at  $a$ ;  $T(a, a) = 2$  and it can show positive or negative asymmetry. The mean,  $E_T(\theta_1) = (a+1)/3$ , increases with the value of the mode, ranging from 0.33 to 0.67. The variance,  $V_T(\theta_1) = (1-a-a^2)/18$ , is equal to 0.056 when  $a = 0$ ; it decreases if  $0 < a < 0.5$ , and reaches a minimum at  $a = 0.5$ , of 0.042, and increases if  $0.5 < a < 1$  again reaching the value 0.056 when  $a = 1$ .

Si  $\theta_1$  y  $\theta_2$  son independientes, entonces la distribución a priori bidimensional es  $TG(\theta_1, \theta_2) = T(\theta_1) \cdot G(\theta_2)$ . Un camino atractivo para modelar la dependencia, en los problemas que nos ocupan, es a través de la dependencia entre los perfiles de riesgo  $\theta_1$  y  $\theta_2$  ver Shevchenko y Wüthrich (2006, 21) o Hernández-Bastida, Fernández-Sánchez y Gómez-Déniz (2009). En este trabajo suponemos que los perfiles de riesgo son independientes.

### 3. La distribución a posteriori

En este apartado se calculan la distribución marginal para crmPLE, que notaremos  $m_{PLE}(s|T \cdot G)$ , la distribución a posteriori, que notaremos  $TG(\theta_1, \theta_2 | s)$ , y la moda de la distribución a posteriori. La distribución marginal se obtiene por cálculo directo usando (1), (2) y (5).

Para  $s > 0$ ,

$$\begin{aligned} m_{PLE}(s|T \cdot G) &= E[f_{PLE}(s|\theta_1, \theta_2)|T \cdot G] = \\ &= \frac{2d^c}{a\Gamma(c)} \int_{\theta_1=0}^a \int_{\theta_2=0}^\infty \left[ 3\theta_1^3(1-\theta_1)\theta_2^c - 2\theta_1^4(1-\theta_1)\theta_2^c + \theta_1^3(1-\theta_1)^3\theta_2^{c+1}s \right] e^{-(s\theta_1+d)\theta_2} d\theta_1 d\theta_2 + \\ &+ \frac{2d^c}{(1-a)\Gamma(c)} \int_{\theta_1=a}^1 \int_{\theta_2=0}^\infty \left[ 3\theta_1^2(1-\theta_1)^2\theta_2^c - 2\theta_1^3(1-\theta_1)^2\theta_2^c + \theta_1^2(1-\theta_1)^4\theta_2^{c+1}s \right] e^{-(s\theta_1+d)\theta_2} d\theta_1 d\theta_2 = \\ &= 3J_1(2,1,1|s) - 2J_1(3,1,1|s) + sJ_1(2,3,2|s). \end{aligned}$$

En el punto de discontinuidad,  $s=0$ , se obtiene la siguiente expresión más sencilla,

$$m_{PLE}(0|T \cdot G) = \frac{-3a^3 + 7a^2 + 7a + 7}{30}$$

Los momentos de la distribución marginal pueden ser útiles para la especificación de los hiperparámetros, en este trabajo se omiten porque no es el objetivo del mismo

If  $\theta_1$  and  $\theta_2$  are independent, then the joint prior distribution for the bivariate parameter  $(\theta_1, \theta_2)$  is  $TG(\theta_1, \theta_2) = T(\theta_1) \cdot G(\theta_2)$ . Observe that an attractive way to model the dependence is via that between the risk profiles  $\theta_1$  and  $\theta_2$  (in such analysis scenarios, the parameters are usually considered risk profiles), see Shevchenko and Wüthrich (2006) or Hernández-Bastida, Fernández-Sánchez and Gómez-Déniz (2009). Here, the risk profiles are assumed to be independent.

### 3. Posterior distribution

In this section, the marginal distributions for the crm.PLE model, denoted as  $m_{PLE}(s|T \cdot G)$ , the posterior distribution  $TG(\theta_1, \theta_2 | s)$  and the bivariate mode of the posterior distribution are determined. The marginal distribution is obtained directly, via several operations, using (1), (2) and (5).

For  $s > 0$

For the point of discontinuity,  $s=0$ , the following, simpler, expression is obtained:

For the specification of the hyperparameters, the raw order moments of the marginal distribution can be useful. However, they have been omitted here

pero de las expresiones (3) y (4) se deduce que el cálculo no es difícil aunque tedioso.

La Tabla siguiente recoge la masa de probabilidad asignada al punto de discontinuidad  $s=0$  por la distribución marginal, para valores de la moda  $a$ , desde 0 hasta 1 con saltos de 0.1.

because this is not the aim of our paper; however, from expressions (3) and (4) it can be shown that the calculus is not difficult, though tedious.

The next table shows the mass of probability assigned by the marginal distribution to the point of discontinuity  $s=0$  for the values of the mode of  $T(\theta_1)$ ,  $a$ , from 0 to 1 with steps of 0.1.

**Tabla 1/Table 1**

$a$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$m_{PLE}(0 T \cdot G)$	0.233	0.258	0.288	0.322	0.358	0.396	0.436	0.477	0.518	0.559	0.6

En la tabla anterior se observa que en todos los casos la masa de probabilidad en  $s=0$  es notable, y que, como era de esperar del comportamiento de la distribución PL, es monótona no decreciente en el valor de  $a$ .

La tabla siguiente muestra los valores de la función de densidad marginal para  $s > 0$ . Se han seleccionado valores grandes de la moda  $a$  porque ilustran mejor la alta probabilidad del valor  $k=0$  en conjuntos de datos habituales en el problema que nos ocupa. Para los hiperparámetros  $c$  y  $d$  se han seleccionado pares de valores que recogen una amplia variedad de valores de la media y la varianza de la distribución  $G(\theta_2)$ . En los análisis numéricos del resto de este artículo siempre consideraremos algunos de los casos aquí recogidos.

Observe that in all the cases of the table below, the mass of marginal probability in  $s=0$  is notable and, as expected from the behaviour of the PL distribution, it is monotonously non-decreasing in the value of the mode  $a$ .

The following table shows the values for the marginal density function for  $s > 0$ . High values of the mode of  $T(\theta_1)$ ,  $a$ , were selected because they better illustrate the high probability of the value  $k=0$  from the usual data in the problem in which we are interested. For the hyper-parameters  $c$  and  $d$  sets of pairs of values were selected which indicate a wide variety of the values of the mean and variance of the  $G(\theta_2)$  distribution. In the rest of the numerical analysis in this paper, we consider some of the cases indicated above.

**Tabla 2/Table 2**

$(c, d)$	$\alpha = 0.4$				$\alpha = 0.6$			
	(1.5,2)	(4,6)	(9,7)	(25,2)	(1.5,2)	(4,6)	(9,7)	(25,2)
S								
0.01	0.32610	0.1249	0.24033	2.23632	0.34546	0.13244	0.25475	2.35181
0.51	0.22084	0.10714	0.18521	0.33404	0.21851	0.11054	0.18732	0.25636
1.01	0.15846	0.09281	0.14567	0.09381	0.14851	0.09337	0.14113	0.06479
1.51	0.11842	0.08106	0.11654	0.03692	0.10623	0.07967	0.10858	0.02482
2.51	0.07204	0.06315	0.07773	0.00996	0.06564	0.05955	0.06775	0.00664
3.51	0.04747	0.05034	0.05018	0.00399	0.03815	0.05034	0.04476	0.00266
4.51	0.03309	0.04087	0.03910	0.00198	0.02574	0.04581	0.03095	0.00132
6.51	0.01808	0.02814	0.02208	0.00069	0.01347	0.02364	0.01645	0.00046
8.51	0.01101	0.02027	0.01353	0.00038	0.00799	0.01640	0.00971	0.00021
10.51	0.00723	0.01512	0.00883	0.00017	0.00515	0.01188	0.00618	0.00011
12.51	0.00501	0.01159	0.00605	0.00010	0.00352	0.00890	0.00417	0.00007
14.51	0.00362	0.00909	0.00431	0.00007	0.00252	0.00685	0.00294	0.00004

Ya se ha comentado con anterioridad que una de las aproximaciones al problema considerado es obtener las medidas de riesgo sustituyendo en  $f_{PLE}(s/\theta_1, \theta_2)$  los parámetros por un estimador adecuado y entonces determinar el cuantil de orden 0.999 en la función de distribución de la pérdida agregada, con esta idea a continuación se determina la moda a posteriori como el estimador más conveniente para el parámetro bidimensional  $(\theta_1, \theta_2)$ . La moda a posteriori se puede determinar por cálculo directo con las tres ecuaciones siguientes que se obtienen de las derivadas parciales del logaritmo de la distribución a posteriori,

As commented above, in one of the approaches to the problem in question, the risk measures were obtained by substituting the parameters in  $f_{PLE}(s/\theta_1, \theta_2)$  by a convenient estimator and then obtaining the 0.999 quantile in the aggregate loss distribution. With this purpose, the posterior mode, which is the most convenient estimator of the bivariate parameter  $(\theta_1, \theta_2)$ , is now determined. The bivariate mode is calculated directly using the following three equations derived from the partial derivatives of the logarithm of  $TG(\theta_1, \theta_2 | s)$ .

$$(I) \quad \frac{3 - 4\theta_1}{\theta_1 - \theta_1^2} - \frac{2 + 2s(1 - \theta_1)\theta_2}{3 - 2\theta_1 + s(1 - \theta_1)^2 \theta_2} - s\theta_2 = 0;$$

$$(II) \quad \frac{2 - 4\theta_1}{\theta_1 - \theta_1^2} - \frac{2 + 2s(1 - \theta_1)\theta_2}{3 - 2\theta_1 + s(1 - \theta_1)^2 \theta_2} - s\theta_2 = 0;$$



$$(III) \quad \frac{c}{\theta_2} + \frac{s(1-\theta_1)^2}{3-2\theta_1+s(1-\theta_1)^2\theta_2} - s\theta_1 - d = 0.$$

Del sistema de ecuaciones formado por (I) y (III) se obtienen las soluciones que verifican  $\theta_1 \in (0, a)$  y  $\theta_2 > 0$ . Del sistema formado por (II) y (III) se obtienen las soluciones que verifican  $\theta_1 \in (a, 1)$  and  $\theta_2 > 0$ . Entonces se evalúa la distribución a posteriori en las soluciones anteriores y en  $\theta_1 = a$ , y el valor más grande nos indica el valor de la moda que notamos  $(\hat{\theta}_1, \hat{\theta}_2)$ .

En el punto de discontinuidad  $s=0$ , de las derivadas parciales del logaritmo de  $TG(\theta_1, \theta_2 | 0) = \theta_1^2(2-\theta_1)T(\theta_1)G(\theta_2)$  se obtiene directamente que,  $\hat{\theta}_2 = (c-1)/d$ ; la obtención de  $\hat{\theta}_1$  se hace con los siguientes pasos:

Para  $\theta_1 < a$ , se determina una raíz de la ecuación  $3\theta_1^3 - 4\theta_1^2 - 1 = 0$ , pero para  $\theta_1 \in (0, 1)$ , el polinomio de tercer grado nunca se anula y es siempre negativo, por tanto la distribución a posteriori es monótona decreciente y el máximo se alcanza en el extremo inferior del intervalo, es decir, en  $\theta_1 = 0$ .

Para  $\theta_1 > a$ , hay que buscar una raíz de la ecuación  $3\theta_1^3 - 7\theta_1^2 + 4\theta_1 - 1 = 0$ , pero ocurre lo mismo que en el caso anterior y el máximo se alcanza en  $\theta_1 = a$ .

Finalmente comparamos los valores  $\theta_1^2(2-\theta_1)T(\theta_1)$  en  $\theta_1 = 0$  y en  $\theta_1 = a$ , y de la comparación se obtiene que  $\hat{\theta}_1 = a$ . En definitiva para  $s = 0$  se obtiene que la

From (I) and (III), the solutions verifying  $\theta_1 \in (0, a)$  and  $\theta_2 > 0$  are selected. From (II) and (III), the solutions verifying  $\theta_1 \in (a, 1)$  and  $\theta_2 > 0$  are selected. Then,  $TG(\theta_1, \theta_2 | s)$  is evaluated in all the previous solutions and in  $\theta_1 = a$ , and the highest value indicates the mode, denoted as  $(\hat{\theta}_1, \hat{\theta}_2)$ .

For  $s=0$ , from the partial derivatives of the logarithm of  $TG(\theta_1, \theta_2 | 0) = \theta_1^2(2-\theta_1)T(\theta_1)G(\theta_2)$  it is obtained, directly, that  $\hat{\theta}_2 = (c-1)/d$ . To obtain  $\hat{\theta}_1$  the following steps are taken:

For  $\theta_1 < a$ , determine a root in  $3\theta_1^3 - 4\theta_1^2 - 1 = 0$ , but for  $\theta_1 \in (0, 1)$ , the expression is never null and it is strictly negative, so the posterior distribution is monotonously non-increasing and the maximum is obtained at the lower extreme of the interval, i.e. in  $\theta_1 = 0$ .

For  $\theta_1 > a$ , search for a root in  $3\theta_1^3 - 7\theta_1^2 + 4\theta_1 - 1 = 0$ . The procedure is the same as in the previous case and so the maximum is achieved at  $\theta_1 = a$ .

Compare the values of  $\theta_1^2(2-\theta_1)T(\theta_1)$  for  $\theta_1 = 0$  and  $\theta_1 = a$ . The comparison leads to  $\hat{\theta}_1 = a$ . Finally, for  $s=0$ , it is obtained that  $(\hat{\theta}_1, \hat{\theta}_2) = (a, (c-1)/d)$ , which are, respectively, the mode of  $T(\theta_1)$  and the mode of  $G(\theta_2)$ .

moda de la distribución a posteriori es  $(\hat{\theta}_1, \hat{\theta}_2) = (a, (c-1)/d)$ , que son la moda de  $T(\theta_1)$  y de  $G(\theta_2)$ , respectivamente.

En la siguiente tabla se muestran valores de la moda bidimensional a posteriori para  $s = 0$ ,  $s = 1.51$  y  $s = 2.01$ , y para dos ternas de hiperparámetros.

The following table shows the values of the posterior mode, for  $s = 0$ ,  $s = 1.51$ ;  $s = 2.01$  for two pairs of hyper-parameters.

**Tabla 3/Table 3**

	$a$	$c$	$d$	$\hat{\theta}_1$	$\hat{\theta}_2$
$s = 0$	0.4	3.5	2	0.4	1.25
$s = 0$	0.6	9	7	0.6	1.14
$s = 1.51$	0.4	3.5	2	0.4	1.44
$s = 2.01$	0.6	9	7	0.6	1.12

#### 4. Determinación de las medidas de riesgo

En este apartado se determinan las medidas de riesgo para el problema de riesgo operacional considerando las dos aproximaciones propuestas en la literatura y antes indicadas.

La primera de ellas es determinar un estimador apropiado para el parámetro bidimensional  $(\theta_1, \theta_2)$ , en nuestro caso la moda a posteriori  $(\hat{\theta}_1, \hat{\theta}_2)$  y en la distribución de la pérdida agregada de  $f(s|\hat{\theta}_1, \hat{\theta}_2)$  determinar el cuantil de orden 0.999. La segunda de ellas es considerar la distribución predictiva, que entendemos es la distribución apropiada al problema que nos ocupa, y en ella determinar el cuantil 0.999. Como se podrá observar las diferencias entre las dos medidas de riesgo son muy significativas.

#### 4. Determining the risk measures

In this section, our aim is to determine the risk measures in the operational risk problem, considering both the approaches proposed in the literature and those described above.

Firstly, we determine the posterior mode  $(\hat{\theta}_1, \hat{\theta}_2)$  as an appropriate estimator of the bivariate parameter  $(\theta_1, \theta_2)$ , and then the 0.999 quantile in the loss aggregate distribution function (LADF) of  $f(s|\hat{\theta}_1, \hat{\theta}_2)$ . The second step is to consider the predictive distribution function (PDF), which, in our opinion, is the appropriate distribution in this problem, and then to determine the 0.999 quantile in this distribution. It can be observed that the differences between the two distributions are considerable.

La siguiente tabla recoge diversos valores de la función de distribución de la pérdida agregada, destacando los valores altos de la probabilidad que nos indican el valor del cuantil de orden 0.999. En la tabla se consideran dos ternas de hiperparámetros y se supone que la observación disponible,  $s_1$  es 0. Recordemos la discontinuidad del crm.PLE en el valor 0, lo que produce un comportamiento anómalo del modelo para ese valor.

The following table shows several values of the loss aggregate distribution function; note the high values of the probability indicating the value of the 0.999 order quantile. Two different sets of hyper-parameters are considered and it is assumed that the available observation,  $s_1$  is 0. The discontinuity of the crm.PLE at the value 0 produces an anomalous behaviour of the model for this value.

**Tabla 4/ Table 4**

$a = 0.4 ; c = 3.5 ; d = 2 ; s_1 = 0 ; (\hat{\theta}_1, \hat{\theta}_2) = (0.4; 1.25)$											
S	4.5	5	5.5	6	6.5	7	8	9	10	11	11.5
LADF	0.870	0.895	0.914	0.931	0.944	0.955	0.970	0.981	0.988	0.992	0.999
$a = 0.6 ; c = 9 ; d = 7 ; s_1 = 0 ; (\hat{\theta}_1, \hat{\theta}_2) = (0.6; 1.14)$											
S	3	3.5	4	4.5	5	5.5	6	7	8	9	19
LADF	0.919	0.941	0.956	0.968	0.977	0.983	0.987	0.993	0.996	0.998	0.999

La siguiente tabla muestra diversos valores de la función de distribución de la pérdida agregada (que notamos LADF), señalando los valores del cuantil de orden 0.999 para el conjunto de hiperparámetros indicados y suponiendo que la observación disponible,  $s_1$  no es 0; de la discontinuidad antes indicada del crm.PLE se deduce que los valores de esta tabla no son comparables con los de la tabla anterior.

The next table shows several values for the loss aggregate distribution function, with the values of the 0.999 quantile for the pair of hyper-parameters indicated, and it is assumed that the available observation,  $s_1$  is not 0. From the discontinuity of the crm.PLE at the value 0, as indicated above, it follows that the values in this table cannot be compared with those of the previous table.

**Tabla 5 / Table 5**

$a = 0.4 ; c = 3.5 ; d = 2 ; s_1 = 1.51 ; (\hat{\theta}_1, \hat{\theta}_2) = (0.4; 1.44)$											
S	4.5	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5
LADF	0.902	0.923	0.940	0.953	0.963	0.971	0.978	0.983	0.986	0.989	0.999
$a = 0.6 ; c = 9 ; d = 7 ; s_1 = 2.01 ; (\hat{\theta}_1, \hat{\theta}_2) = (0.6; 1.12)$											
S	4	5	6	7	8	9	10	11	12	13	14
LADF	0.909	0.945	0.967	0.980	0.988	0.992	0.995	0.997	0.998	0.999	0.999

A continuación consideramos la distribución predictiva del crmPLE, que notamos  $p_{PLE}^{(T-G)}(s|s_1)$ , la cual es, en nuestra opinión, la distribución adecuada en el problema que nos ocupa para tratar correctamente el carácter aleatorio de los parámetros  $\theta_1$  y  $\theta_2$ .

Para la determinación de los cuantiles se precisa la función de distribución predictiva, que notamos  $P_{PLE}^{(T-G)}(s|s_1)$ , y que viene dada en todos los casos por un punto de masa de probabilidad en cero más la correspondiente integral, es decir,

$$P_{PLE}^{(T-G)}(s|s_1) = p_{PLE}^{(T-G)}(0|s_1) + \int_0^s p_{PLE}^{(T-G)}(s|s_1) ds$$

los cuantiles se obtienen por evaluaciones sucesivas de la función anterior.

La distribución predictiva en el crmPLE, se deriva a partir de la expresión que la define,

$$p_{PLE}^{(T-G)}(s|s_1) = E_{TG(\theta_1, \theta_2|s_1)} [f_{PLE}(s|\theta_1, \theta_2)] = \frac{E_{TG(\theta_1, \theta_2)} [f_{PLE}(s|\theta_1, \theta_2) \cdot f_{PLE}(s_1|\theta_1, \theta_2)]}{m_{PLE(s_1|T-G)}}$$

A continuación, realizando el producto  $f_{PLE}(s|\theta_1, \theta_2) \cdot f_{PLE}(s_1|\theta_1, \theta_2)$ , para los cuatro casos posibles de  $(s, s_1)$ , que son  $ss_1 \neq 0$ ;  $s=0, s_1 \neq 0$ ,  $s \neq 0, s_1=0$  y  $s=s_1=0$ , con algunos cálculos tediosos pero sencillos, y usando (1), (2) y (5), se obtiene en todos los casos un cociente de combinaciones lineales de las funciones  $J_1$

Para  $s_1 > 0$ , es decir, suponiendo que la observación disponible  $s_1$  no es nula, la función de densidad predictiva viene dada por,

Let us now consider the predictive distribution in the crmPLE, denoted as  $p_{PLE}^{(T-G)}(s|s_1)$ , which is, in our opinion, the appropriate distribution for the problem considered, in order to correctly deal with the random character of the parameters  $\theta_1$  and  $\theta_2$ .

In order to determine the quantiles, we need the predictive distribution function, which is given in all cases by a point of mass of probability at 0 plus the corresponding integral, i.e.

The quantiles are obtained by successively evaluating each previous function.

The predictive distribution,  $p_{PLE}^{(T-G)}(s|s_1)$ , is derived from its definition:

Next, performing the product  $f_{PLE}(s|\theta_1, \theta_2) \cdot f_{PLE}(s_1|\theta_1, \theta_2)$ , for the four possible cases of  $(s, s_1)$ , which are  $ss_1 \neq 0$ ;  $s=0; s_1 \neq 0$ ,  $s \neq 0, s_1=0$  and  $s=s_1=0$ . Following tedious but straightforward calculus, and using (1), (2), and (5), we obtain, in all cases, a division between the linear combinations of the  $J_1$  functions.

For  $s_1 > 0$ , i.e., when it is assumed that the available observation  $s_1$  is not null, the predictive density function is given by,

$$p_{PLE}^{(T,G)}(s | s_1) = \frac{1}{m_{PLE}(s_1 | T \cdot G)} \{9J_1(4, 2, 2 | s + s_1) - 12J_1(5, 2, 2 | s + s_1) + 4J_1(6, 2, 2 | s + s_1) + 3(s + s_1)J_1(4, 4, 3 | s + s_1) - 2(s + s_1)J_1(5, 4, 3 | s + s_1) + s \cdot s_1 J_1(4, 6, 4 | s + s_1)\}.$$

Para  $s = 0$ , se obtiene el punto de masa de probabilidad en cero, que aprovechando la notable simplificación de  $f_{PLE}(0 | \theta_1, \theta_2)$ , viene dado por el siguiente número,

For  $s = 0$ , the mass point of probability at 0 is obtained. Taking into account that there is a notable simplification in  $f_{PLE}(0 | \theta_1, \theta_2)$ , this mass of probability is given by the following value,

$$p_{PLE}^{(T,G)}(0 | s_1) = \frac{1}{m_{PLE}(s_1 | T \cdot G)} \{6J_1(4, 1, 1 | s_1) - 7J_1(5, 1, 1 | s_1) + 2J_1(6, 1, 1 | s_1) + 2s_1 J_1(4, 3, 2 | s_1) - s_1 J_1(5, 3, 3 | s_1)\}.$$

En las siguientes Tablas 6 y 7 se recogen valores de la función de distribución predictiva destacando los valores de probabilidad altos.

Tables 6 and 7 show the values for the predictive distribution function, revealing the high values of the probability.

En la Tabla 6 se considera una terna de hiperparámetros y cuatro casos de observación disponible  $s_1$  no nula.

For Table 6, a set of hyper-parameters and four cases of non-null available observations,  $s_1$ , are considered.

**Tabla 6/Table 6**

$a = 0.4 ; c = 3.5 ; d = 2 ; s_1 = 1.51$											
S	4.5	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5
PDF	0.851	0.868	0.883	0.895	0.905	0.914	0.921	0.928	0.933	0.938	0.942
$a = 0.4 ; c = 3.5 ; d = 2 ; s_1 = 2.01$											
PDF	0.829	0.849	0.865	0.879	0.891	0.901	0.910	0.917	0.924	0.929	0.934
$a = 0.4 ; c = 3.5 ; d = 2 ; s_1 = 3.01$											
PDF	0.788	0.811	0.830	0.847	0.862	0.875	0.886	0.896	0.904	0.912	0.918
$a = 0.4 ; c = 3.5 ; d = 2 ; s_1 = 4.01$											
PDF	0.748	0.773	0.796	0.815	0.832	0.847	0.860	0.872	0.883	0.892	0.900

De la lectura de la Tabla 6 por columnas se observa que las probabilidades son decrecientes en la observación disponible  $s_1$ , para un valor de  $s$  fijo, lo que implica

On observing the columns of the Table 6 it can be seen that for a fixed  $s$  the probabilities are non-increasing in the available observation  $s_1$ , and so the

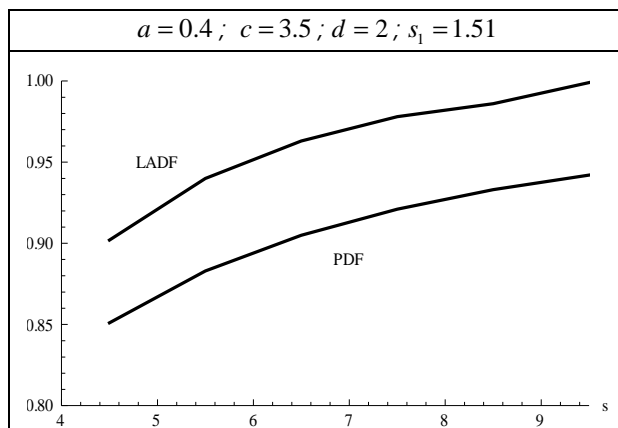
que los cuantiles son crecientes en  $s_1$ , es decir, la esperable conclusión de que a mayor pérdida observada  $s_1$  mayor será la medida de riesgo.

Si se compara el primer caso de la Tabla 6 con el primer caso de la Tabla 5 se observa claramente que los cuantiles, por tanto las medidas de riesgo, obtenidas con la distribución predictiva son apreciablemente más altos que los obtenidos con la función de distribución de la pérdida agregada. La siguiente Figura 3 ilustra este hecho, corroborando las conclusiones obtenidas con otras modelizaciones en Shevchenko and Wüthrich (2006) o en Hernández-Bastida, Fernández-Sánchez y Gómez-Déniz (2009), es decir, la notable diferencia en las medidas de riesgo obtenidas en cada una de las dos aproximaciones.

quantiles are non-decreasing in  $s_1$ , i.e., as expected, the greater the observed loss in  $s_1$ , the larger the risk measure.

Comparing the first cases in Tables 5 and 6, it is clear that the quantiles, as risk measures, obtained from the predictive distribution are considerably higher than the quantiles obtained from the loss aggregate distribution function. Figure 3 illustrates this situation, coinciding with the conclusions obtained in other models such as Shevchenko and Wüthrich (2006) and Hernández-Bastida, Fernández-Sánchez and Gómez-Déniz (2009), highlighting the remarkable difference in the risk measures obtained in both approximations.

**Figura 3/ Figure 3**



En la Tabla 7 se considera otra terna de hiperparámetros con moda de  $\theta_1$  mayor y media y varianza de  $\theta_2$  más pequeñas que en la tabla 6, y también cuatro casos de observación disponible,  $s_1$  no nula.

Table 7 contains a set of hyper-parameters with a greater mode of  $\theta_1$  and a mean and variance of  $\theta_2$  smaller than in Table 6, with four cases of non-null available observations,  $s_1$ .

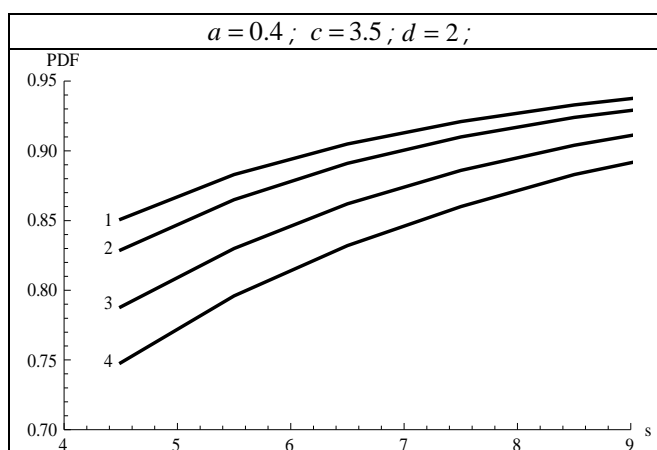
**Tabla 7/Table 7**

$a = 0.6 ; c = 9 ; d = 7 ; s_1 = 1.51$											
S	4	5	6	7	8	9	10	11	12	13	14
PDF	0.849	0.885	0.911	0.929	0.943	0.953	0.960	0.966	0.971	0.974	0.977
$a = 0.6 ; c = 9 ; d = 7 ; s_1 = 2.01$											
PDF	0.828	0.868	0.897	0.917	0.933	0.944	0.953	0.960	0.965	0.970	0.973
$a = 0.6 ; c = 9 ; d = 7 ; s_1 = 3.01$											
PDF	0.788	0.834	0.868	0.893	0.912	0.927	0.938	0.947	0.955	0.960	0.965
$a = 0.6 ; c = 9 ; d = 7 ; s_1 = 4.01$											
PDF	0.748	0.799	0.838	0.867	0.890	0.908	0.922	0.933	0.943	0.950	0.956

Las conclusiones de la Tabla 7 son idénticas a las de la Tabla 6, tanto en la lectura de la tabla por columnas como de la comparación del segundo caso de la Tabla 7 con el segundo caso de la Tabla 5. La conclusión respecto a la lectura en columnas de la Tabla 7 se ilustra en la siguiente Figura 4, donde los números (1), (2), (3) y (4) se refieren a los cuatro valores de  $s_1$ , considerados en la tabla.

The conclusions drawn from Table 7 are identical to those for Table 6 in the observation of the columns and in the comparison of the second cases of Tables 7 and 5. The conclusions drawn from the columns in Table 7 are illustrated in Figure 4, where the numbers (1), (2), (3) and (4) refer to the four values for  $s_1$  considered in the Table.

**Figura 4/Figure 4**



Para  $s_1 = 0$ , es decir, suponiendo que la observación disponible es nula, como en el caso anterior hay una notable simplificación en  $f_{PLE}(s_1 | \theta_1, \theta_2)$ , y se obtiene la función de densidad predictiva,

$$p_{PLE}^{(T \cdot G)}(s | 0) = \frac{1}{m_{PLE}(0 | T \cdot G)} \{6J_1(4, 1, 1 | s) - 7J_1(5, 1, 1 | s) + 2J_1(6, 1, 1 | s) + 2sJ_1(4, 3, 2 | s) - sJ_1(5, 3, 3 | s)\}.$$

Finalmente, para  $s = s_1 = 0$ , se obtiene el punto de masa de probabilidad en  $s = 0$ , que aprovechando la notable simplificación en  $f_{PLE}(s | \theta_1, \theta_2)$  y en  $f_{PLE}(s_1 | \theta_1, \theta_2)$ , viene dado por,

$$p_{PLE}^{(T \cdot G)}(0 | 0) = \frac{1}{m_{PLE}(0 | T \cdot G)} \{J_1(6, 0, 0 | 0) - 4J_1(5, 0, 0 | 0) + 4J_1(4, 0, 0 | 0)\}$$

En la siguiente Tabla 8 se recogen valores de la función de distribución predictiva cuando la observación disponible  $s_1$  es 0.

For  $s_1 = 0$ , i.e., when it is assumed that the available observation  $s_1$  is null, as in the previous case, there is a notable simplification in  $f_{PLE}(s_1 | \theta_1, \theta_2)$ , with the predictive density function being obtained.

Finally, for  $s = s_1 = 0$ , there is a notable simplification in  $f_{PLE}(s | \theta_1, \theta_2)$  and in  $f_{PLE}(s_1 | \theta_1, \theta_2)$ , obtaining, the mass point of probability at  $s = 0$ , given by

The next Table 8 shows the values of the predictive distribution function when the available observation  $s_1$  is 0.

**Tabla 8/ Table 8**

$a = 0.4 ; c = 3.5 ; d = 2 ; s_1 = 0$											
S	4.5	5	5.5	6	6.5	7	8	9	10	11	11.5
PDF	0.919	0.927	0.934	0.939	0.944	0.948	0.954	0.959	0.962	0.965	0.966
$a = 0.6 ; c = 9 ; d = 7 ; s_1 = 0$											
S	3	3.5	4	4.5	5	5.5	6	7	8	9	19
PDF	0.903	0.920	0.933	0.943	0.951	0.958	0.963	0.971	0.977	0.981	0.984

Comparando los resultados de la Tabla 8 con los obtenidos en la Tabla 4 se vuelve a poner de manifiesto la disparidad en las medidas de riesgo obtenidas según sea la aproximación considerada en el problema.

**Agradecimientos:** AHB está financiado por el Ministerio de Educación y Ciencia (Project ECO2009-14152).

Comparing the results shown in Table 8 and Table 4, there is an evident difference between the risk measures depending on the approach taken to the problem.

**Acknowledgements:** AHB is funded by Ministerio de Educación y Ciencia (Project ECO2009-14152).



## 5. Referencias bibliográficas/References

- BIS (2005). *Basel II: International convergence of capital measurement and capital standard: A revised framework*. Bank for International Settlements (BIS), www.bis.org.
- Carrillo-Menéndez, S. y Suarez, A. (2006). "Effective measurement of the operational risk", *Estabilidad financiera*, Banco de España, 11, pp. 61-90.
- Chavez-Demoulin, V.; Embrechts, P. y Neslehová, J. (2006). "Quantitative models for operational risk: Extremes, dependence and aggregation", *Journal of Banking and Finance*, 30 (10), pp. 2635-2658.
- Cruz, M.G. (2002). *Modelling, measuring and hedging operational risk*. UK: John Wiley and Sons.
- Ghitany, M.E. y Al-Mutairi, D.K. (2008). "Estimation methods for the discrete Poisson-Lindley distribution", *Journal of Statistical Computation and Simulation*, 79 (3), pp. 279-287.
- Ghitany, M.E.; Al-Mutairi, D.K. y Nadarajah, S. (2009). "Zero-truncated Poisson-Lindley distribution and its application", *Mathematics and Computers in Simulation*, 79 (1), pp.1-9.
- Grandell, J. (1997). *Mixed poisson processes*. New York: Chapman and Hall.
- Hernández-Bastida, A.; Fernández-Sánchez, M.P y Gómez-Déniz, E (2009). "The net bayes premium with dependence between the risk profiles", *Insurance:Mathematics and Economics*, 45, pp. 247-254.
- Hernández-Bastida, A.; Fernández-Sánchez, M.P. y Gómez-Déniz, E. (2010). "Collective risk model: Poisson-Lindley and exponential distributions for bayes premium and operational risk", *Journal of Statistical Computation and Simulation*, to appear.
- Johnson, D. (1997). "The triangular distribution as a proxy for the beta distribution in risk analysis", *Journal of the Royal Stat. Society, Series D (The Statistician)*, 46 (3), pp. 388-398.
- Karlis, D. y Xekalaki, E. (2005). "Mixed poisson distributions", *International Statistical Review*, 73, pp. 35-58.
- Klugman, S.A.; Panjer, H.H. and Willmot, G.E. (2004). *Loss models: From data to decision*. New York: Willey and Sons.
- Kozubowski, T.J. and Panoska, A.K. (2005). "A mixed bivariate distribution with exponential and geometric marginals", *Journal of Statistical Planning and Inference*, 234, pp. 501-520.
- Lambrigger, D.D.; Shevchenko, P.V. and Wüthrich, M.V. (2007). "The quantification of operational risk using internal data, relevant external data and expert opinions", *The Journal of Operational Risk*, 2-3, pp. 3-27.
- Lindley, D.V. (1958). "Fiducial distributions and Bayes's theorem", *Journal of the Royal Stat. Soc. Series B*, 1; pp. 102-107.
- McNeil, A.; Frey, R. y Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques and tools*. New Jersey (USA): Princeton University Press.
- Nadarajah, S. y Kotz, S. (2006a). "Compound mixed Poisson distributions I", *Scandinavian Actuarial Journal*, 3, pp. 141-162.

- Nadarajah, S. y Kotz, S. (2006b). "Compound mixed Poisson distributions II", *Scandinavian Actuarial Journal*, 3, pp. 163-181.
- Nikoloulopoulos, A.K. y Karlis, D. (2008). "On modelling count data: a comparison of some well-known discrete distributions", *Journal of Statistical Computation and Simulation*, 78 (3), pp. 437-457.
- Sandström, A. (2006). *Solvency: Models, assessment and regulation*. Boca Raton (FL): Chapman & Hall/CRC.
- Sankaran, M. (1971). "The discrete Poisson-Lindley distribution", *Biometrics*, 26 (1), pp. 145-149.
- Shevchenko, P.V. (2008). "Estimation of operational risk capital charge under parameter uncertainty", *The Journal of Operational Risk*, 1 (3), pp. 51-63.
- Shevchenko, P.V. y Wüthrich, M.V. (2006). "The structural modelling of operational risk via bayesian inference: Combining loss data with expert opinions", *The Journal of Operational Risk*, 1 (3), pp. 3-26.
- Wüthrich, M.V. (2006). "Premium liability risks: Modelling small claims", *Bulletin of the Swiss Association of Actuaries*, 1, pp. 27-38.