

The Bayes Premium in an Aggregate Loss Poisson-Lindley Model with Structure Function STSP

Agustin Hernández-Bastida¹ & M. Pilar Fernández-Sánchez¹

¹ Quantitative Methods in Economics, University of Granada, Spain

Correspondence: Agustin Hernández-Bastida, Faculty of Economics, University of Granada, Campus Cartuja s/n 18071, Granada, Spain. Tel: 34-958-242-977. E-mail: bastida@ugr.es

Received: February 3, 2013 Accepted: March 6, 2013 Online Published: March 21, 2013

doi:10.5539/ijsp.v2n2p47

URL: <http://dx.doi.org/10.5539/ijsp.v2n2p47>

Abstract

Many premium calculating problems in actuarial science consider the number of claims, denoted as K , as the variable risk. Traditionally, this random variable is modelled by the Poisson distribution. However, it is well known that automobile insurance portfolios are characterized by zero-inflation (high percentage of zero values in the sample) and overdispersion (the variance is greater than the mean), and the Poisson distribution does not properly reflect the last phenomenon. In this paper we determine the Bayes premium considering that K follows a Poisson-Lindley distribution, with parameter θ_1 in $[0, 1]$, which is a potential alternative to describe these situations. As the structure function for θ_1 we elicit the standardized two-sided power distribution, which is a reasonable alternative to the usual beta distribution. In addition, an aggregate loss model is considered with primary distribution given by the Poisson-Lindley distribution. A Bayesian analysis is developed to obtain the Bayes premium. The conclusion is that the STSP is not an adequate alternative in the problem in question because it is more informative and less dispersed than the Beta distribution.

Keywords: Bayes premium, aggregate loss, Poisson-Lindley model, structure function, standardized two-sided power distribution

1. Introduction

An important issue in insurance theory is that of premium calculation. In particular, determining the Bayes premium is the natural aim when prior information and claim experience are considered (see Klugman, 1992).

In many practical problems of actuarial science, the claim experience is described by a frequency distribution for the number of claims, K , considered as the magnitude risk. Traditionally, this random variable is modelled by the Poisson distribution. However, it is well known that automobile insurance portfolios are characterized by zero-inflation (high percentage of zero values in the sample) and overdispersion (the variance is greater than the mean), and the Poisson distribution does not properly reflect the last phenomenon. For this reason, as Nikoloulopoulos and Karlis (2008) point out, overdispersed models (relative to simple Poisson) are potential alternative to describe these situations. In particular, mixed Poisson distributions are widely considered for the random variable K . For a detailed review of models based on the mixed Poisson distributions see Cohen (1966), Willmot (1986), Grandell (1997), Nadarajah and Kotz (2006a, 2006b) or Antzoulakos and Chadjiconstantinidis (2004), among others.

Some of the advantages of these distributions obtained by mixing are that they are overdispersed, and that they also assign high probabilities at $k = 0$, which is very adequate because the mode of the variable number of claims is often at this value. Sometimes, this situation is managed with zero-inflated distributions (see Angers & Biswas, 2003; Yip & Yau, 2005; Boucher et al., 2007) which are not considered here to avoid introducing another parameter in the problem.

One of the mixed distribution is the Poisson-Lindley distribution, with parameter $\theta_1 \in (0, 1)$, proposed by Sankaran (1971). It has been widely studied by Ghitany et al. (2008) or Ghitany and Al-Mutairi (2009). An extension of this distribution, which includes an additional parameter is suggested in Mahmoudi and Zakerzadeh (2010). Some recent works have pointed out the usefulness of this one-parameter distribution in the Bayes premium calculating problem (see Hernández-Bastida et al., 2011 or Martel-Escobar et al., 2012).

In other practical situations it is considered the aggregate loss model. In actuarial risk theory, the collective risk model, hereafter 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 independent. Note that the independence assumed here is conditional on distribution parameters. There is an extensive body of literature on modelling the risk process, see e.g. Klugman et al. (2004) or McNeil et al. (2005). Estimation of the annual loss distribution by modelling the frequency and severity of losses is a well known actuarial technique. It is also used for modelling solvency requirements in the insurance industry; see e.g. Sandström (2006) or Wüthrich (2006). Then the aggregate loss S is the sum of the individual claim sizes, i.e. $S = \sum_{i=1}^K X_i$, for $K > 0$, and $S = 0$, for $K = 0$.

The present paper aims to develop a Bayesian analysis of the PL model for the number of claims. The structure function for the parameter θ_1 is given by the standardized two-sided power (\mathcal{STSP}) distribution, a reasonable alternative for the beta distribution.

The analysis include two steps. The first step, which is developed in Section 2, considers that the risk variable is the number of claims. In this section the model is presented. The premiums are determined and an application to real data is provided. In the second step, which is developed in Section 3, it is considered an aggregate loss model and the Bayes premiums are also obtained. Both sections show the results obtained from the comparison with the beta distribution, since it is a usual distribution in this problem. A final section draws the main conclusions.

For our purposes, let

$${}_pF_q \left[\{a_1, a_2, \dots, a_p\}; \{b_1, b_2, \dots, b_k\}; z \right] = \sum_{k=0}^{\infty} \frac{\prod_{j=1}^p (a_j)_k z^k}{\prod_{j=1}^q (b_j)_k k!}$$

represent the Gaussian hypergeometric function, and $(\alpha)_j = \alpha(\alpha + 1)\dots(\alpha + j - 1)$ for $j \geq 1$, $(\alpha)_0 = 1$ be the Pochhammer symbol (see <http://functions.wolfram.com>).

2. The Claims Number Model

In this section the variable number of claims, denoted as K , is considered as the variable risk. We determine the Bayes premium, when the \mathcal{STSP} is considered as structure function, and it is compared with the Bayes premium when the structure function is the usual beta distribution. An application to real data is also provided.

2.1 The Model

Let K be the random variable number of claims taking values $\{0, 1, \dots\}$ which is assumed to follow a Poisson-Lindley distribution with a probability mass function (pmf) given by

$$f_{PL}(k|\theta_1) = \theta_1^2 (1 - \theta_1)^k [2 - \theta_1 + (1 - \theta_1)k], \quad (1)$$

for $k = 0, 1, \dots$, and $0 < \theta_1 < 1$, (hereafter the PL model).

It is well known that its moment generating function is given by $M_{PL}(t; \theta_1) = \frac{\theta_1^2(2-\theta_1-e^t+\theta_1 e^t)}{(1-e^t+\theta_1 e^t)^2}$. The first two moments are $E_{PL}[K] = \frac{(2-\theta_1)(1-\theta_1)}{\theta_1}$ and $E_{PL}[K^2] = \frac{(1-\theta_1)(3\theta_1^2-8\theta_1+6)}{\theta_1^2}$, respectively.

It is obtained as a mixed Poisson distribution with mixing function given by the Lindley distribution (see Lindley, 1958), whose density function is $f(x|\theta_1) = \frac{\theta_1^2}{\theta_1+1} (x+1) e^{-x\theta_1}$, for $x = 0, 1, \dots$, and $\theta_1 > 0$. Compared with the Poisson distribution, the PL distribution presents three qualities that usually appear in claim data sets. Since it is a mixed Poisson distribution then it is overdispersed. The probability of observing a zero value is higher than under a Poisson distribution with the same mean, termed zero inflated phenomenon, (see Karlis & Xekalaki, 2005 and references therein). Furthermore, for mean values under 2.4, which includes practically all the real data cases, it is straightforward to prove the termed one-deflated phenomenon, i.e., $f_{PL}(1; \theta_1) < f_P(1; \frac{(2-\theta_1)(1-\theta_1)}{\theta_1})$, where f_P is the probability mass function of the Poisson distribution.

Under a bayesian point of view, the parameters of interest of the problem can be estimated using our knowledge about them.

The beta distribution $\mathcal{B}(\theta_1; \alpha, \beta) \propto \theta_1^{\alpha-1} (1 - \theta_1)^{\beta-1}$, $0 < \theta_1 < 1$; $\alpha > 0$ and $\beta > 0$ is a common prior distribution for the parameter θ_1 , where $E_{\mathcal{B}}[\theta_1] = \frac{\alpha}{\alpha+\beta}$ and $Var_{\mathcal{B}}[\theta_1] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$. With $\alpha > 1$ and $\alpha + \beta > 2$ is unimodal and $Mo_{\mathcal{B}}[\theta_1] = \frac{\alpha-1}{\alpha+\beta-2}$. For the application of this model to the premium calculating and operational risk see Hernández-Bastida et al. (2011). It is known that in the PL model the marginal distribution of K , if the $\mathcal{B}(\alpha, \beta)$

prior is considered, is given by

$$m(k|\mathcal{B}) = \frac{A_1(\alpha, \beta, k)B(\alpha + 3, \beta + k)}{B(\alpha, \beta)}, \quad (2)$$

where $A_1(\alpha, \beta, k) = \frac{\alpha + 2\beta + 2 + k(\beta + k + 2)}{\Gamma(\alpha + 2)B(\alpha, \beta)}$. Observe that this marginal distribution is a mixed Poisson distribution with mixing function given by $(\lambda + 1)\mathcal{U}(\alpha + 2, -\beta + 2, \lambda)$, where \mathcal{U} is the Hypergeometric \mathcal{U} .

For the calculation of the moments of the marginal distribution it is useful to observe that they can be written as $E_{m(k|\mathcal{B})}[K^r] = E_{\mathcal{B}}[E_{f_{PL}}[K^r]]$. It also occurs with the rest of the marginal distributions.

A meaningful alternative to the beta distribution is proposed in Van Dorp and Kotz (2002a). Hence, the \mathcal{STSP} distribution is proposed as a choice of prior distribution for θ_1 . Its pdf is given by:

$$\mathcal{STPS}(\theta_1; a, b) = \begin{cases} b \left(\frac{\theta_1}{a}\right)^{b-1}, & 0 \leq \theta_1 \leq a; \\ b \left(\frac{1 - \theta_1}{1 - a}\right)^{b-1}, & a \leq \theta_1 \leq 1. \end{cases} \quad (3)$$

The prior mean and the variance for θ_1 are given by $E_{\mathcal{STSP}}[\theta_1] = \frac{(b-1)a+1}{b+1}$ and $Var_{\mathcal{STSP}}[\theta_1] = \frac{b(2a^2 - 2a + 1) + 2a(1-a)}{(b+2)(b+1)^2}$. When $b > 1$ and $0 < a < 1$ or $a = 0$ or 1 , it is unimodal with mode $Mo_{\mathcal{STSP}}[\theta] = a$ and $\mathcal{STPS}(a) = b$, (see Van Dorp & Kotz, 2002b).

When $b = 2$ the Triangular distribution, denoted as $\mathcal{T}(\theta_1, a)$ (see Johnson, 1997), is obtained. If $b = 1$ the Uniform distribution in $[0, 1]$ is derived (hereafter the \mathcal{U} distribution) and if $a = 1$ it follows the potential distribution (hereafter, the \mathcal{P} distribution).

This paper focuses on the \mathcal{STSP} distribution as an alternative prior distribution to the beta distribution for θ_1 . To make this comparison operative, we establish certain aspects considered essential to the distribution modelling the prior information. The first aspect to consider is that of unimodality with the value of the mode at a , thus completely determining the \mathcal{T} distribution. To rank the comparison we are interested in, the second aspect is set so as to consider unimodal distributions with the same mean. Given a unimodal $\mathcal{STSP}(a, b)$ distribution, $b > 1$, a unimodal $\mathcal{B}(\alpha, \beta)$ distribution with the same mode and mean as the $\mathcal{STSP}(a, b)$ is obtained by considering

$$\alpha = ab - a + 1; \quad \beta = a + b - ab. \quad (4)$$

In the sequel we assume that the parameters for the \mathcal{B} distribution are determined from the expressions in (4).

In accordance with our aim of comparing the three models derived from the three prior distributions for the parameter of the distribution considered, \mathcal{T} , \mathcal{STSP} and \mathcal{B} , we compare the amount of information contained in each of these distributions, considering such quantity to be the discrepancy of each of them with respect to the uniform distribution, defined as (see Shannon, 1948; Kullback, 1959 or Kullback & Leibler, 1951) $D_{KL}[f : \mathcal{U}] = -\int_0^1 f(x)\log f(x)dx$. If the uniform distribution is considered to be the less informative probabilistic model, then the greater the discrepancy $D_{KL}[f : \mathcal{U}]$, the greater the information contained in the f distribution.

The divergence between the \mathcal{B} and the \mathcal{U} distributions is obtained with the following expression, (see Soofi & Retzer, 2002) $D_{KL}[\mathcal{B} : \mathcal{U}] = -\log(B(\alpha, \beta)) + (\alpha - 1)[\psi(\alpha) - \psi(\alpha + \beta)] + (\beta - 1)[\psi(\beta) - \psi(\alpha + \beta)]$, where $\psi(z)$ is the PolyGamma function (see <http://functions.wolfram.com>).

The discrepancy of the \mathcal{STSP} with respect to the \mathcal{U} is given by $D_{KL}[\mathcal{STSP} : \mathcal{U}] = \log b - 1 + b^{-1}$, with a minimum at $b = 1$, whose value is 0, and which strictly increases if $b > 1$ (see Van Dorp & Kotz, 2002b, for details). Hence, if $1 < b < 2$ it is verified that $D_{KL}[\mathcal{STSP} : \mathcal{U}] < D_{KL}[\mathcal{T} : \mathcal{U}] \equiv 0.19314718$ and when $b > 2$ the contrary occurs.

From the comparison between the \mathcal{STSP} and the \mathcal{B} with the same mode and mean, it follows that whatever the value of the mode a , and whatever the value $b > 1$, the \mathcal{B} has less information and more dispersion than the \mathcal{STSP} distribution.

The following notation will be useful in the sequel. For the sake of simplicity, the acronym of the distributions are used to indicate the corresponding density functions.

Let ε_1 and ε_2 be integers and let k be a positive integer. We note $E_{STSP} \left[\theta_1^{\varepsilon_1} (1 - \theta_1)^{\varepsilon_2} f(k|\theta_1) \right]$ as $\mathcal{D}(\varepsilon_1, \varepsilon_2|k, a, b)$. The calculation of this expected value is tedious but after a bit of algebra it is obtained (see the Appendix for details).

In the PL model the marginal distribution of K , if the $STSP(a, b)$ prior is considered, is

$$m(k|STSP) = \mathcal{D}(0, 0|k, a, b), \quad k \geq 0. \quad (5)$$

For $k = 0$ the expression is notably simplified and it is obtained that $m(0|STSP) = \frac{6+4b+a[-6+b(2+4b+a(-3+a+b(-2+ab^2)))]}{(b+1)_3}$.

Making $b = 2, b = 1$ and $a = 1$ in expression (5) the marginal distributions $m(k|\mathcal{T})$, $m(k|\mathcal{U})$ and $m(k|\mathcal{P})$ are obtained, respectively.

For the beta and $STSP$ distributions we determine the hyper-parameters with the expected value of the marginal distribution and the value of this distribution at $k = 0$. Making $E_{m(k|\dots)} = \bar{k}$, where \bar{k} represents the sample mean of the data, and $m(0|\dots) = f_0$ where f_0 is the relative frequency of the value $k = 0$, we get two systems of two equations. The hyper-parameters can be obtained by solving the systems. In the particular case of the \mathcal{T} distribution, we can use one of both possibilities since we only need one parameter.

In the real data sets for the problem considered, it is common to find that the frequency of the zero value is extremely high, clearly more than the 50%. Hence, it is worth paying special attention to how the marginal distribution at $k = 0$ behaves.

It is no so difficult to prove that, for instance with the Mathematica package, the $m(0|\mathcal{B})$ and $m(0|STSP)$ distributions, as functions of the hyper-parameters a and b have similar behavior.

For $a \leq 0.54$ fixed, these functions decrease uniformly in b , achieving the maximum value at $b = 1$. This maximum is always under 0.45.

For $a > 0.54$ fixed, these functions increase uniformly in b , whose superior is obtained when b increases indefinitely. It is straightforward to show that $\lim_{b \rightarrow \infty} m(0|\mathcal{B}) = \lim_{b \rightarrow \infty} m(0|STSP) = a^2(2 - a)$, and the limit function only takes values over 0.5 when a is greater than 0.6.

In summary, for the problems in question the range of relevant values for a is given by the interval (0.6, 1).

It is straightforward to prove that $m(0|\mathcal{T})$ varies in the interval (0.234; 0.6), and furthermore, $E_{m(k|\mathcal{T})}[K]$ ranges in the interval (1, 67; 34.8). Accordingly, we will not consider the \mathcal{T} distribution as structure function in data sets which present observed frequency at $k = 0$ greater than 0.6 or sample mean smaller than 1.67.

For the comparison between the $m(k|\mathcal{B})$ and $m(k|STSP)$ distributions we have considered the difference function defined as $Dmarg(k, a, b) = m(k|\mathcal{B}) - m(k|STSP)$ and a complete numeric analysis for fixed values for k has been developed. In each case there have been determined the superior and the inferior in (a, b) for the difference function. It is obtained that,

- When $k \leq 3$, the difference function takes positive and negative values.
- When $k \geq 3$, the difference function always takes positive values.
- In any case the function given by the absolute value of the difference, denoted as $|Dmarg(k, a, b)|$ takes small values. The highest one is 0.033 when $k = 1$ and when $k \geq 3$ it is always under 0.0058.

As a conclusion it is shown that, considering the marginal distribution, it is not direct the difference between the \mathcal{B} and $STSP$ models.

2.2 The Premiums

In this section, for the several specified prior distributions, we examine the collective and Bayes premiums. The collective premium is defined as the expected value of the True Individual premium with respect to the corresponding prior distribution and it is equal to the expected value of the marginal distribution for K . Observe that the collective premium is the appropriate premium when claim experience is not available. It is straightforward to prove that the Bayes premium, which is the expected value of the True Individual premium with respect to the corresponding posterior distribution, is equal to the expected value of predictive function of K . Furthermore, it is the appropriate premium when experience claim is present being the best estimation of the True Individual premium.

In the PL model, the True Individual premium is given by $E[K]$.

Hence, if the $\mathcal{B}(\alpha, \beta)$ or the $STSP(a, b)$ are considered as prior distributions then the corresponding collective premiums are, by direct calculus,

$$CP(\mathcal{B}) = \frac{\beta(\alpha + 2\beta + 1)}{(\alpha - 1)(\alpha + \beta)}, \quad (6)$$

and

$$CP(STSP) = \frac{(2a + 1)b^2 + (3 - 2a)b}{b^2 - 1} - a - 2 + \frac{2b}{(1 - a)^{b-1}} \int_a^1 \frac{(1 - \theta_1)^{b-1}}{\theta_1} d\theta_1, \quad (7)$$

respectively.

When $b = 1$ the $CP(\mathcal{U})$ does not exist and for $a = 1$ it follows that $CP(\mathcal{P}) = \frac{b+3}{b^2-1}$.

In the comparison between the collective premiums for the structure functions \mathcal{B} and $STSP$ with the same mean and mode it is obtained that the first distribution produces greater values for the collective premium, i.e.,

$$CP(\mathcal{B}) \geq CP(STSP). \quad (8)$$

This affirmation is obtained just observing that it is equivalent to $E_{\mathcal{B}}[\theta_1^{-1}] \geq E_{STSP}[\theta_1^{-1}]$ which is equivalent to

$$\frac{b(1-a)}{a(b-1)} - \frac{b}{(1-a)^{b-1}} \int_a^1 \frac{(1-\theta_1)^{b-1}}{\theta_1} d\theta_1 \geq 0.$$

It can be proved that the term in the left hand achieves a minimum at 0.2527.

The Bayes premium is determined and compared for the different prior distributions. By direct calculus,

$$BP(k|\mathcal{B}) = \frac{C_3(k, \alpha, \beta)(\beta + k)}{(\alpha + 1)(\alpha + \beta + k + 3)[\alpha + 2\beta + 2 + k(\beta + k + 2)]}, \quad (9)$$

where

$$C_3(k, \alpha, \beta) = 4(\alpha + \beta + k + 3)(\beta + k + 1) + (\alpha + 1)_2 + k(\beta + k + 1)(\alpha + 2\beta + 2k + 5).$$

Using (17) for the $STSP$,

$$BP(k|STSP) = \frac{2\mathcal{D}(-1, 1|k, a, b) - \mathcal{D}(0, 1|k, a, b)}{m(k|STSP)}. \quad (10)$$

For the comparison of the Bayes premium we made a complete numerical analysis for the difference function $Dbp(k, a, b) = BP(k|\mathcal{B}) - BP(k|STSP)$ for integers of k , $a \geq 0.6$ and $b > 1$. The conclusions are clear, when there is experience of "non sinistrality", i.e., $k = 0$, the difference function takes values positives and negatives although the last ones predomine.

However, when any claims are declared, i.e., $k \geq 1$, the difference function is always positive. That means that, compared with the $STSP$ distribution, the \mathcal{B} distribution penalizes more severely the additional claim declaration. Table 1 shows the highest values which have been observed in the function $\frac{Dbp(k, a, b)}{BP(k|STSP)} \times 100$, and it indicates the level of additional penalization given by the \mathcal{B} as structure function.

Table 1. The highest observed values in b for the function $\frac{Dbp(k, a, b)}{BP(k|STSP)} \times 100$ for the indicated values of k and a

| k | a | | | |
|---|-------|-------|-------|-------|
| | 0.6 | 0.7 | 0.8 | 0.9 |
| 1 | 17.32 | 23.20 | 27.27 | 27.93 |
| 2 | 36.67 | 42.54 | 45.10 | 42.78 |
| 3 | 51.66 | 56.47 | 56.99 | 52.28 |
| 4 | 63.58 | 66.95 | 65.62 | 58.99 |
| 5 | 73.41 | 75.21 | 72.23 | 64.12 |

It is clear that the differences between the premiums are remarkable, and the conclusion in the model for the number of claims is clear: If the essential aspects of the prior information are the unimodality, with the value of the mode,

and the mean value, then the \mathcal{B} distribution is a better choice than the $STSP$ distribution. The \mathcal{B} distribution is less informative than the other distribution and also produces higher values for the premiums i.e., so it results more conservative from the point of view of the insurance firm.

2.3 An Application to Real Data

To address the issues described previously, we develop the analysis applied to a real data set which is well known in studies of automobile insurance claim calculation. These data, taken from Klugman et al. (2004) concern the number of automobile liability policies in Germany during the years 1960-61. The sample mean is 0.1442 and the sample variance is 0.1639. Accordingly, the sample dispersion index is 1.136. The Binomial Negative distribution (hereafter, NB), which is another mixed Poisson distribution allowing for overdispersion, and the Poisson distribution are compared with the Poisson-Lindley model to data set. Table 2 shows fits from these distributions. For comparative and illustrative purposes, all the usual measures, such as p-value, -Loglikelihood, the Akaike Information criterio (AIC) and the Bayesian Information Criteio (BIC) are used to compared the models. As it is known, a model with a minimum BIC value is preferred.

Table 2. Fitting of automobile claim data

| Number of claims | Observed frequency | Fitting distribution (expected frequency) | | |
|------------------|--------------------|---|-------------------|-------------------|
| | | Poisson | Neg.Binomial | Poisson-Lindley |
| 0 | 20592 | 20420 | 20596.8 | 20612.10 |
| 1 | 2651 | 2945.10 | 2631.03 | 2604.39 |
| 2 | 297 | 212.37 | 318.37 | 326.21 |
| 3 | 41 | 10.21 | 37.81 | 40.56 |
| 4 | 7 | 0.37 ^a | 4.45 ^a | 5.01 ^a |
| 5 | 0 | 0.01 ^a | 0.52 ^a | 0.62 ^a |
| 6+ | 1 | 0.00 ^a | 0.06 ^a | 0.07 ^a |
| $\chi^d(d.f)$ | | 203.87(2) | 5.083(1) | 4.40(2) |
| p-value | | < 1% | 0.08 | 0.22 |
| -Loglikelihood | | 10297.8 | 10223.4 | 10223.9 |
| AIC | | 20597.7 | 20450.8 | 20449.8 |
| BIC | | 20605.8 | 20467.8 | 20457.8 |

^a Expected frequencies have been combined for the calculation of χ^2 .

Table 2 shows that the PL model performs very well in fitting the distribution with respect to the Poisson distribution and provides a fit as good as that of the biparametric Negative Binomial model. Based on the AIC and the BIC the PL is the preferred model. Furthermore, taking into account the Ockham’s razor principle, it is simpler than the Negative Binomial and therefore it might appear to be more preferable than a less complex model.

The next step is to choose the prior distributions. As the sample mean is under 1.67, the Triangular distributions is not considered here as possible structure function.

For the \mathcal{B} and $STSP$ distributions we solve the systems of equations proposed in the previous section, and we obtain a pair of values which let us to the $\mathcal{B}(8, 1)$ or the $STSP(1, 8)$, which are the same distribution. The following figure shows the prior, marginal and posterior distributions for a $\mathcal{B}(8, 1)$.

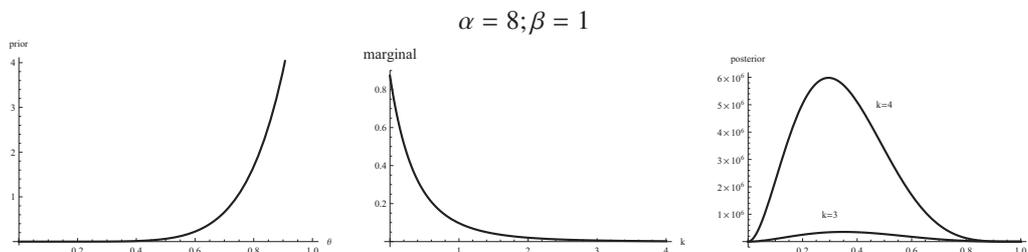


Figure 1. Prior, marginal and posterior distributions

In this case, the collective premium, which is given by the expected value of the marginal distribution is equal to 0.1746. The Bayes premium and the variation rates, denoted as TV and calculated as $\frac{BP(k|\dots) - BP(k-1|\dots)}{BP(k-1|\dots)} \times 100$, are in Table 3.

Table 3. Bayes premiums and variation rates

| k | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----|-------|--------|--------|--------|-------|--------|--------|--------|
| BP | 0.143 | 0.317 | 0.504 | 0.696 | 0.891 | 1.087 | 1.284 | 1.484 |
| TV% | | 58.940 | 38.064 | 27.909 | 22.30 | 18.250 | 18.200 | 15.527 |

3. The Collective Risk Model

This section develops an aggregate loss model and determines the Bayes premium. A *STSP* distribution is considered as an alternative to the usual \mathcal{B} distribution as structure function for the parameter of the primary distribution whereas a gamma distribution is considered for the secondary one.

3.1 The Model

Let X_i be the random variable size of a single claim assumed to follow an Exponential distribution with parameter $\theta_2 > 0$, i.e., $f_E(x|\theta_2) = \theta_2 e^{-\theta_2 x}$, $x > 0$ (hereafter the model *E*). Its moment generating function is $M_2^E(t; \theta_2) = \frac{\theta_2}{\theta_2 - t}$ and the expected value and the variance are $E_E[X] = \frac{1}{\theta_2}$ and $V_E[X] = \frac{1}{\theta_2^2}$, respectively.

If the primary distribution is a Poisson-Lindley and the secondary is an Exponential, (the *crmPLE*), it is verified that:

(i) The probability density function of the random variable aggregate claim, when $s > 0$, is given by,

$$f_{PLE}(s|\theta_1, \theta_2) = \theta_1^2 (1 - \theta_1) \theta_2 (3 - 2\theta_1 + (1 - \theta_1)^2 \theta_2 s) e^{-\theta_1 \theta_2 s}, \tag{11}$$

while

$$f_{PLE}(0|\theta_1, \theta_2) = \theta_1^2 (2 - \theta_1), \tag{12}$$

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

(ii) The moment generating function of S is given by $M_3^{PLE}(t; \theta_1, \theta_2) = \frac{\theta_1^2 [(2 - \theta_1)t^2 + (\theta_1 \theta_2 - 3\theta_2)t + \theta_2^2]}{(\theta_1 \theta_2 - t)^2}$ and the mean and the variance are $E_{PLE}(S) = \frac{\theta_1^2 - 3\theta_1 + 2}{\theta_1 \theta_2}$ and $V_{PLE}(S) = \frac{-\theta_1^4 + 4\theta_1^3 - 5\theta_1^2 + 3\theta_1 + 2}{\theta_1^2 \theta_2^2}$, respectively.

The natural choice of a prior pdf for θ_2 is the Gamma density $\mathcal{G}(\theta_2; c, d) = \frac{d^c}{\Gamma(c)} \cdot \theta_2^{c-1} \cdot \exp(-d\theta_2)$, $\theta_2 > 0$, $c, d > 0$. We consider values of c such that $c > 1$, a necessary and sufficient condition for the existence of the inverse moment. The prior mean and variance for θ_2 are given by $E_{\mathcal{G}}[\theta_2] = \frac{c}{d}$ and $Var_{\mathcal{G}}[\theta_2] = \frac{c}{d^2}$. The corresponding prior mode for the parameter θ_2 is $Mo_{\mathcal{G}}[\theta_2] = \frac{c-1}{d}$.

The following notation and formulas are used throughout this paper.

Let $\epsilon_1, \epsilon_2, \epsilon_3$ be integers and let k be a positive integer. Let s be a real positive number. If there is no possibility of misunderstanding and for the sake of simplicity of the notation, the hyper-parameters will be omitted. We note $E_{STSP \cdot \mathcal{G}}[\theta_1^{\epsilon_1} (1 - \theta_1)^{\epsilon_2} \theta_2^{\epsilon_3} e^{-s\theta_1 \theta_2}]$ as $\mathcal{J}(\epsilon_1, \epsilon_2, \epsilon_3|s)$. This expected value is calculated in a detailed way in the Appendix.

We note $E_{\mathcal{B} \cdot \mathcal{G}}[\theta_1^{\epsilon_1} \theta_2^{\epsilon_2} f_{PLE}(s|\theta_1, \theta_2)]$ as $I(\epsilon_1, \epsilon_2|s)$. This expected value is obtained as a linear combination of hypergeometric functions (see the Appendix for details).

In the *crmPLE* the marginal distribution of S , if independent $\mathcal{B}(\alpha, \beta)$ and $\mathcal{G}(c, d)$ priors are considered, is given by a linear combination of hypergeometric functions when $s > 0$

$$m(s|\mathcal{B} \cdot \mathcal{G}) = I(0, 0|s), \tag{13}$$

using (9'), and when $s = 0$, it follows that $m(0|\mathcal{B} \cdot \mathcal{G}) = m(0|\mathcal{B})$.

In addition, in the *crmPLE* the marginal distribution of S , if independent *STSP*(a, b) and $\mathcal{G}(c, d)$ priors are considered, is given by

$$m(s|STSP \cdot \mathcal{G}) = 3\mathcal{J}(2, 1, 1|s) - 2\mathcal{J}(3, 1, 1|s) + s\mathcal{J}(2, 3, 2|s), \quad s > 0 \tag{14}$$

and, when $s = 0$, it follows that, $m(0|STSP \cdot \mathcal{G}) = m(0|STSP)$.

As it was indicated in section 2, the real data sets in the problems in question present a high percentage of values at $s = 0$. Accordingly, it is interesting to analyze the values for the marginal distributions at $s = 0$. The analysis is reduced to that made in the previous section.

From the comparison between the marginal distributions $m(s|\mathcal{B} \cdot \mathcal{G})$ and $m(s|STSP \cdot \mathcal{G})$ when $s > 0$, the numeric analysis show that there exist inequalities in both directions.

By assuming $b = 1$ in expressions (14) the marginal distribution for the prior $\mathcal{U} \cdot \mathcal{G}$ distribution is obtained and it is irrespective of the hyper-parameters a and b .

In the question of the specification of the hyper-parameters different sources of information can be considered in order to help us to choose them on a reasonable basis: the unimodality and the value of the mode are usually a strong intuitive aspect; the consideration, separately, of the PL and E models whose composition leads to the crm model; the value of $m(s|\dots)$ at $s = 0$ and, finally, it can be useful to calculate some moments of the marginal distribution $m(s|\dots)$. For the calculus of the moments for the marginal distribution is useful to know that $E_{m(s|\mathcal{B} \cdot \mathcal{G})}[S^r] = E_{\mathcal{B} \cdot \mathcal{G}}[E_{f_{PLE}}[S^r]]$, which is also true for the other marginal distributions.

3.2 The Premiums

In this section we examine the Risk Net premium, which is the expected value of the likelihood, the collective premium and the Bayes premium.

In the crm.PLE model, the Risk Net premium is $E_{PLE}[S] = E_{PL}[K] \cdot E_E[X]$.

Then, the collective premiums if independence between θ_1 and θ_2 is assumed, are given by $CP(\mathcal{B} \cdot \mathcal{G}) = CP(\mathcal{B}) \cdot CP(\mathcal{G})$ and $CP(STSP \cdot \mathcal{G}) = CP(STSP) \cdot CP(\mathcal{G})$, where $CP(\mathcal{G}) = E_{\mathcal{G}}[\theta_2^{-1}] = \frac{d}{c-1}$.

If $b = 1$ (\mathcal{U} distribution), the collective premium does not exist.

The comparison of the collective premiums is deduce directly from the inequality in the model for the number of claims, specifically $CP(\mathcal{B} \cdot \mathcal{G}) \geq CP(STSP \cdot \mathcal{G})$.

Finally, we determine the Bayes premiums. Hence, the Bayes premium for the $\mathcal{B} \cdot \mathcal{G}$ distribution is obtained by dividing linear combinations of hypergeometric functions,

$$BP(s|\mathcal{B} \cdot \mathcal{G}) = \frac{2I(-1, -1; s) - 3I(0, -1; s) + I(1, -1; s)}{I(0, 0; s)} \quad (15)$$

Assuming $s = 0$ on the right side of this expression we obtain the Bayes Premium in the point of discontinuity, which is given by $BP(0|\mathcal{B} \cdot \mathcal{G}) = BP(0|\mathcal{B}) \cdot CP(\mathcal{G})$.

The Bayes Premium for the $STSP \cdot \mathcal{G}$ prior distribution is given by the following expressions:

For $s > 0$,

$$BP(s|STSP \cdot \mathcal{G}) = \frac{1}{m(s|STSP \cdot \mathcal{G})} \{6\mathcal{J}(1, 2, 0|s) - 7\mathcal{J}(2, 2, 0|s) + 2s\mathcal{J}(1, 4, 1|s) + 2\mathcal{J}(3, 2, 0|s) - s\mathcal{J}(2, 4, 1|s)\}; \quad (16)$$

and for $s = 0$, $BP(0|STSP \cdot \mathcal{G}) = BP(0|STSP) \cdot CP(\mathcal{G})$.

Making $b = 1$, the Bayes premium for the prior $\mathcal{U} \cdot \mathcal{G}$ distribution is obtained.

In the comparison for the Bayes premium we consider, separately, the cases $s = 0$ and $s > 0$.

When $s = 0$, the comparison between the Bayes premium is reduced to the comparison between $BP(0|\mathcal{B})$ and $BP(0|STSP)$ made in the previous section.

For $s > 0$, we have made a wide numerical analysis for the values $a = 0.6, 0.7, 0.8, 0.9$ and $b = 2, 7, 10, 25$. Several gamma distributions have been considered. The conclusions are similar in the following way: it is obtained positive and negative values for the variation rates, defined as the quotient $\frac{BP(s|\mathcal{B} \cdot \mathcal{G}) - BP(s|STSP \cdot \mathcal{G})}{BP(s|STSP \cdot \mathcal{G})} \times 100$ with the hyper-parameter for the beta distribution given by (4). The variation rates indicate that none of the premiums is systematically greater or slower than the other. Furthermore, it is shown that sometimes it is possible to find dramatic differences between the premiums. As an illustration, Table 4 shows the variation rates for the Bayes premium for the $\mathcal{G}(4, 6)$.

Table 4. Variation rates for Bayes premiums in the crmPLE model for the indicated values of a, b and s , when the hyper-parameter $(c, d) = (4, 6)$

| a | 0.6 | | | | 0.7 | | | | |
|------|--------|-------|----------|---------|--------|---------|--------|--------|----|
| | b | 2 | 7 | 15 | 25 | 2 | 7 | 15 | 25 |
| s | | | | | | | | | |
| 0.1 | 56.7 | 117.3 | 134.6 | 137.3 | 66.7 | 244.7 | 475.3 | 679.0 | |
| 5.1 | 44.4 | 119.9 | 132.8 | 128.4 | 43.3 | 182.9 | 325.0 | 412.0 | |
| 10.1 | 38.0 | 119.1 | 132.3 | 125.6 | 427.0 | 161.8 | 289.2 | 358.3 | |
| 15.1 | 33.9 | 117.4 | 131.6 | 123.9 | 29.0 | 141.9 | 272.3 | 334.8 | |
| 20.1 | 31.1 | 115.8 | 130.9 | 122.7 | 25.7 | 136.1 | 262.2 | 321.4 | |
| 25.1 | 29.0 | 114.2 | 130.7 | 121.8 | 23.5 | 135.1 | 255.4 | 312.7 | |
| a | 0.8 | | | | 0.9 | | | | |
| | b | 2 | 7 | 15 | 25 | 2 | 7 | 15 | 25 |
| s | | | | | | | | | |
| 0.1 | 71.5 | 734.8 | -829.9.6 | -478.0 | 68.5 | -5164.7 | -306.9 | -232.6 | |
| 5.1 | 38.9 | 289.2 | 5760.2 | -750.8 | 31.9 | 416.9 | -419.4 | -261.8 | |
| 10.1 | -210.7 | 216.5 | 3645.4 | -978.4 | -283.1 | 249.5 | -508.9 | 276.8 | |
| 15.1 | 212.4 | 184.5 | 1880.8 | -1172.7 | -78.4 | 193.6 | -583.8 | -286.0 | |
| 20.1 | 19.8 | 166.1 | 1427.1 | -1341.2 | 138.0 | 165.2 | 647.8 | -292.3 | |
| 25.1 | 18.0 | 154.0 | 1218.1 | -1488.8 | 11.8 | 147.8 | -703.4 | 296.8 | |

4. Discussion

The aim of this paper is to determine the Bayes and collective premiums, in an aggregate loss Poisson-Lindley model. A natural choice for the structure function is the Beta distribution. In this paper, taking into account the unimodality with mode value and mean value as essentials aspects of the prior information, it is studied the possibility of considering the STSP distribution as alternative to the structure function. The conclusion is that the STSP is not an adequate alternative in the problem in question because it is more informative and less dispersed than the Beta distribution. Comparing the marginal distribution it is not easy to distinguish it from the Beta distribution. However, we obtain premium values totally different.

Acknowledgements

AHB is funded by Ministerio de Educacion y Ciencia (project ECO2009-14152).

References

- Angers, J. F., & Biswas, A. (2003). A Bayesian analysis of zero-inflated generalized Poisson model. *Computational Statistics and Data Analysis*, 42, 37-46.
- Antzoulakos, D., & Chadjiconstantinidis, S. (2004). On Mixed and Compound Mixed Poisson Distributions. *Scandinavian Actuarial Journal*, 3, 161-188. <http://dx.doi.org/10.1080/03461230110106525>
- Boucher, J. P., Denuit, M., & Guillen, M. (2007). Risk classification for claims counts: a comparative analysis of various zero-inflated mixed models. *North American Actuarial Journal*, 11(4), 110-131.
- Cohen, A. C. (1966). A note on certain discrete mixed distributions. *Biometrics*, 22, 566-572.
- Ghitany, M. E., & Al-Mutairi, D. K. (2009). Estimation methods for the discrete Poisson-Lindley distribution. *Journal of Statistical Computation and Simulation*, 79(1), 1-9. <http://dx.doi.org/10.1080/00949650701550259>
- Ghitany, M. E., Al-Mutairi, D. K., & Nadarajah, S. (2008). Zero-truncated Poisson-Lindley distribution and its applications. *Mathematics and Computers in Simulation*, 79(3), 279-287. <http://dx.doi.org/10.1016/j.matcom.2007.11.021>
- Grandell, J. (1997). *Mixed Poisson Processes*. New York: Chapman and Hall.
- Hernández-Bastida, A., Fernández-Sánchez, M. P., & Gómez-Déniz, E. (2011). Collective risk model: Poisson-Lindley and Exponential distributions for Bayes premium and operational risk. *Journal of Statistical Computation and Simulation*, 81(6), 759-778. <http://dx.doi.org/10.1080/00949650903486609>
- Johnson, D. (1997). The Triangular distribution as a proxy for the Beta distribution in risk analysis. *Journal of the Royal Statistical Society, Series D*, 46(3), 388-398. <http://www.jstor.org/stable/2988573>

- Karlis, D., & Xekalaki, E. (2005). Mixed Poisson distributions. *International Statistical Review*, 73, 35-58.
- Klugman, S. A. (1992). *Bayesian Statistics in Actuarial Science*. Kluwer Academic Press.
- Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2004). *Loss Models: From Data to Decision*. New Jersey: Wiley Interscience.
- Kullback, S. (1959). *Information Theory and Statistics*. New York: Kluwer Academic Press.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematics and Statistics*, 22, 79-86.
- Lindley, D. V. (1958). Fiducial Distributions and Bayes's Theorem. *Journal of the Royal Statistical Society, Serie B, 1*, 102-107. <http://www.jstor.org/stable/2983909>
- Mahmoudi, E., & Zakerzadeh, H. (2010). Generalized Poisson-Lindley distribution. *Communications in Statistics-Theory and Methods*, 39, 1785-1798. <http://dx.doi.org/10.1080/03610920902898514>
- Martel-Escobar, M., Hernández-Bastida, A., & Vázquez-Polo, F. J. (2012). On the independence between risk profiles in the compound collective risk actuarial model. *Mathematics and Computers in Simulation*. <http://dx.doi.org/10.1016/j.matcom.2012.01003>
- McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.
- Nadarajah, S., & Kotz, S. (2006a). Compound mixed Poisson distributions I. *Scandinavian Actuarial Journal*, 3, 141-162. <http://dx.doi.org/10.1080/03461230600783384>
- Nadarajah, S., & Kotz, S. (2006b). Compound mixed Poisson distributions II. *Scandinavian Actuarial Journal*, 3, 163-181. <http://dx.doi.org/10.1080/0346123060071525>
- Nikoloulopoulos, A. K., & Karlis, D. (2008). On modelling count data: a comparison of some well-known discrete distributions. *Journal of Statistical Computation and Simulation*, 78, 437-457. <http://dx.doi.org/10.1080/10629360601010760>
- Sandström, A. (2006). *Solvency: Models, Assessment and Regulation*. Boca Raton: Chapman & Hall/CRC.
- Sankaran, M. (1971). The Discrete Poisson-Lindley Distribution. *Biometrics*, 26, 145-149.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
- Soofi, E. S., & Retzer, J. J. (2002). Information indices: unification and applications. *Journal of Econometrics*, 107, 17-40.
- Van Dorp, J. R., & Kotz, S. (2002a). A novel extension of the triangular distribution and its parameter estimation. *Journal of the Royal Statistical Society. Series D*, 51(1), 63-79. <http://www.jstor.org/stable/3650391>
- Van Dorp, J. R., & Kotz, S. (2002b). The Standard Two-Sided Power distribution and its properties: with applications in financial engineering. *The American Statistician*, 56(2), 90-99. Retrieved from <http://search.proquest.com/docview/228481293>
- Willmot, G. (1986). Mixed compound Poisson Distributions. *Astin Bulletin*, 16, 59-79.
- Wüthrich, M. V. (2006). Premium Liability Risks: Modelling Small Claims. *Bull. Swiss. Assoc. of Actuar*, 1, 27-38.
- Yip, K. C. H., & Yau, K. K. W. (2005). On modelling claim frequency data in general insurance with extra zeros. *Insurance: Mathematics and Economics*, 36, 153-163.

Appendix

Let ϵ_1, ϵ_2 and ϵ_3 be integers, let k be a positive integer and let s be a positive number.

A.1.-

$$\begin{aligned} \mathcal{D}(\epsilon_1, \epsilon_2|k, a, b) &= E_{STSP} \left[\theta_1^{\epsilon_1} (1 - \theta_1)^{\epsilon_2} f(k|\theta_1) \right] \\ &= b a^{\epsilon_1+3} \left\{ C_1(k) + \sum_{j=0}^{k+\epsilon_2} (-1)^j \binom{k+\epsilon_2}{j} a^j \mathcal{D}_1(j, k) \right\} + b (1 - a)^{k+\epsilon_2+1} \left\{ C_2(k) + \sum_{j=0}^{\epsilon_1+2} (-1)^j \binom{\epsilon_1+2}{j} (1 - a)^j \mathcal{D}_2(j, k) \right\} \end{aligned}$$

where

$$C_1(k) = \frac{(-a)^{k+\epsilon_2+1} k}{b + \epsilon_1 + \epsilon_2 + k + 3};$$

$$C_2(k) = \frac{-(a - 1)^{\epsilon_1+3}}{b + \epsilon_1 + \epsilon_2 + k + 3};$$

$$\mathcal{D}_1(j, k) = \frac{2}{b + \epsilon_1 + 2 + j} - \frac{a}{b + \epsilon_1 + 3 + j} + \frac{k(k + \epsilon_2 + 1)}{(k + \epsilon_2 + 1 - j)(b + \epsilon_1 + 2 + j)},$$

and

$$\mathcal{D}_2(j, k) = \frac{\epsilon_1 + 3 - 2j}{(\epsilon_1 + 3 - j)(b + \epsilon_2 + k + j)} + \frac{k(1 - a)}{(b + \epsilon_2 + k + 1 + j)}.$$

A.2.-

$$\mathcal{J}(\epsilon_1, \epsilon_2, \epsilon_3|s) \equiv E_{STSP-G} \left[\theta_1^{\epsilon_1} (1 - \theta_1)^{\epsilon_2} \theta_2^{\epsilon_3} e^{-s\theta_1\theta_2} \right], \tag{1'}$$

and can be written as the following addition:

$$\mathcal{J}(\epsilon_1, \epsilon_2, \epsilon_3|s) = \frac{bd^c \Gamma(c + \epsilon_3)}{\Gamma(c)} \left[\frac{\mathcal{J}_1(\epsilon_1, \epsilon_2, \epsilon_3|s)}{a^{b-1}} + \frac{\mathcal{J}_2(\epsilon_1, \epsilon_2, \epsilon_3|s)}{(1 - a)^{b-1}} \right], \tag{2'}$$

where

$$\mathcal{J}_1(\epsilon_1, \epsilon_2, \epsilon_3|s) = \int_0^a \frac{\theta_1^{b-1+\epsilon_1} (1 - \theta_1)^{\epsilon_2}}{(s\theta_1 + d)^{c+\epsilon_3}} d\theta_1 \tag{3'}$$

and

$$\mathcal{J}_2(\epsilon_1, \epsilon_2, \epsilon_3|s) = \int_a^1 \frac{\theta_1^{\epsilon_1} (1 - \theta_1)^{b-1+\epsilon_2}}{(s\theta_1 + d)^{c+\epsilon_3}} d\theta_1. \tag{4'}$$

When $s = 0$ the integrals in (3) and (4) are immediate and given by $\mathcal{J}_1(\epsilon_1, \epsilon_2, \epsilon_3|0) = \frac{a^{b+\epsilon_1}}{d^{c+\epsilon_3}} \sum_{j=0}^{\epsilon_2} \binom{\epsilon_2}{j} \frac{(-1)^j a^j}{b + \epsilon_1 + j}$ and

$$\mathcal{J}_2(\epsilon_1, \epsilon_2, \epsilon_3|0) = \frac{(1 - a)^{b+\epsilon_2}}{d^{c+\epsilon_3}} \sum_{j=0}^{\epsilon_1} \binom{\epsilon_1}{j} \frac{(-1)^j (1 - a)^j}{b + \epsilon_2 + j}.$$

A.3.- From the integral representation of the Kummer confluent hypergeometric function (see <http://functions.wolfram.com/07.20.07.0001.01>) and the integral representation of ${}_2F_1(\dots)$ (see <http://functions.wolfram.com/07.23.07.00003.01>), the following results are derived

$$\begin{aligned} \int_0^1 e^{t\theta} \theta^{\epsilon_1-1} (1 - \theta)^{\epsilon_2-1} d\theta &= B(\epsilon_1, \epsilon_2) {}_1F_1(\epsilon_1; \epsilon_1 + \epsilon_2; t); \\ I_1(\epsilon_1, \epsilon_2, \epsilon_3) \equiv E_{B-G} \left[\theta_1^{\epsilon_1} (1 - \theta_1)^{\epsilon_2} \theta_2^{\epsilon_3} \right] &= \frac{B(\alpha + \epsilon_1, \beta + \epsilon_2) \Gamma(c + \epsilon_3)}{B(\alpha, \beta) \Gamma(c) d^{\epsilon_3}}, \end{aligned} \tag{5'}$$

The integral

$$I_2(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4|s) \equiv E_{B-G} \left[\theta_1^{\epsilon_1} (1 - \theta_1)^{\epsilon_2} \theta_2^{\epsilon_3} e^{-\epsilon_4\theta_2 d} e^{-s\theta_1\theta_2} \right] \tag{6'}$$

is a hypergeometric gaussian function. In particular,

$$I_2(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 | s) = \frac{B(\alpha + \epsilon_1, \beta + \epsilon_2) \Gamma(c + \epsilon_3)}{B(\alpha, \beta) \Gamma(c) d^{\epsilon_3} (\epsilon_4 + 1)^{c + \epsilon_3}} \times {}_2F_1\left(\{\alpha + \epsilon_1; c + \epsilon_3\}; \{\alpha + \beta + \epsilon_1 + \epsilon_2\}; \frac{-s}{d(\epsilon_4 + 1)}\right). \quad (7')$$

The integral

$$I(\epsilon_1, \epsilon_2 | s) \equiv E\left[\theta_1^{\epsilon_1} \theta_2^{\epsilon_2} f_{PLE}(s | \theta_1, \theta_2) | \mathcal{B} \cdot \mathcal{G}\right] \quad (8')$$

is a linear combination of hypergeometric functions, giving the discontinuity at $s = 0$ in $f_{PLE}(s | \theta_1, \theta_2)$. In particular, we have

$$I(\epsilon_1, \epsilon_2 | s) = 3I_2(\epsilon_1 + 2, 1, \epsilon_2 + 1, 0, s) - 2I_2(\epsilon_1 + 3, 1, \epsilon_2 + 1, 0, s) + sI_2(\epsilon_1 + 2, 3, \epsilon_2 + 2, 0, s), \quad s > 0 \quad (9')$$

$$I(\epsilon_1, \epsilon_2 | 0) = 2I_1(\epsilon_1 + 2, 0, \epsilon_2) - I_1(\epsilon_1 + 3, 0, \epsilon_2). \quad (10')$$

The I function will be useful in the analysis of the crmPLE with the \mathcal{B} prior distribution. The I_1 and I_2 functions are auxiliary functions for the I function.

Notice that all the functions introduced above are perfectly calculable, for example with a software package such as Mathematica. Henceforth, all the magnitudes of interest for our purposes are written in terms of these functions.