Objective Bayesian Estimation of the Mean of Severity and Frequency Distributions

Limor Langbord, Department of Statistics, University of Haifa, Israel

Zinoviy Landsman, Actuarial Research Center, University of Haifa, Israel

Udi E. Makov, Actuarial Research Center, University of Haifa, Israel

May 9, 2018

Abstract

Bayesian estimation, which is often used for the pricing of insurance contracts, relies on prior information and the specification of a loss function. In this study, we explore the intrinsic objective Bayesian point estimators for a quantity of interest based on intrinsic discrepancy loss function - which is an inherent loss function arising only from the underlying distribution or model, without any subjective considerations - and the Jeffrey's prior distribution - which is designed to express absence of information about the quantity of interest. We present the methodology and illustrate it in models which are hitherto partially or fully unexplored in this novel context. In particular, we study the estimation of the mean of the Gamma distribution and the mean of the Poisson distribution, typically used for assessing claim severity and claim frequency, respectively. We compare the performances of the proposed point estimators with the Bayes estimator, which is the posterior mean based on quadratic loss function and Jeffrey's prior distribution.

Keywords: Intrinsic discrepancy loss function, Jeffrey's prior distribution, Intrinsic objective Bayesian point estimator, Intrinsic objective Bayesian risk function.

1 Introduction

The Bayesian approach, which is established as highly relevant to actuarial science (Makov, Smith and Liu 1996, makov 2001), requires a choice of a prior distribution and a loss function which have a direct impact on the outcome of the analysis. Given the sensitivity of the final outcome to these, in the absence of any preference to such a pair, there is a need to establish an objective choice of prior distribution and a loss function. This study proposes Bayesian inference which combines a non-informative prior distribution and an objective loss function, called the intrinsic discrepancy loss function.

Non-informative prior distribution is designed to express lack of information about the quantity of interest. Kass and Wasserman (1996) provided full discussion on the theoretical principles of non-informative prior distributions. Yang and Berger (1998) provided a catalog of non- informative priors and Ghosh (2011) reviewed and compared criteria for the selection of non-informative priors. In this study we adopt the well-known Jeffrey's prior distribution (Jeffrey 1961), which is proportional to the positive square root of the determinant of the Fisher information matrix and is invariant to a one to one transformation of the parameter of interest.

The Loss Function measures the loss incurred by an incorrect estimation of the unknown parameter in the underlying model. The most popular loss function is the quadratic (squared error) loss function which yields the posterior mean as the Bayes estimator. There are two more classical loss functions: the absolute error loss function and the zero-one loss function, which yield the posterior median and the posterior mode as the Bayes estimator, respectively. These classical loss functions are non-invariant (under one to one transformation) and, therefore, provide not-invariant estimators. Several additional loss functions have been suggested in the context of Bayesian inference, including balanced loss functions, general entropy loss function and linex loss function (Payandeh Najafabadi, 2010). Furthermore, the choice of any of these loss functions is typically arbitrary and can be seen as subjective. Robert (1996), Bernardo and Rueda (2002) and Bernardo and Juarez (2003) advocated that the loss function should only depend on the underlying model, without any subjective considerations. Moreover, the loss function should measure the discrepancy between the underlying model based on the true parameter of interest, and the simplified model based on its estimator, instead of the simple discrepancy between the parameter of interest and its estimator. Robert (1996) saw the need for an intrinsic loss function, which derives only from the sampling distribution when no information is available about the utility function of the decision maker. He introduced the intrinsic loss function as non-informative loss function. He proposed two functions: Hellinger distance and kullback-Leibeler divergence. Bernardo and Rueda (2002) and Bernardo and Juarez (2003) further developed the intrinsic loss functions, in the context of hypothesis testing. They introduced an intrinsic discrepancy loss function (IDLF) as an objective loss function which is attractive because of its invariance under reparametrization, and therefore provides invariant Bayes estimators (Bernardo 2005b). Recent works on the IDLF in the context of Bayesian inference include Juarez (2005), Demortier (2009) and Laurent (2010). Juarez (2005) explored the Laplace, Pareto, Inverse-Pareto and Weibull models as well as change-point problems. Demortier (2009) explored this loss function in the context of high energy physics and Laurent (2012) in the context of Bayesian inference on the rate of a Poisson distribution and on the ratio of the rates of two independent Poisson distributions.

The combined use of Jeffrey's prior distribution and the IDLF results in an intrinsic objective Bayesian point estimation, in the sense of being fully dependent on the available data and the underlying model without subjective considerations. The aim of the present study is to investigate this estimator and its approximation (Bernardo 2005b) in two models which have significant relevance to actuarial science for modelling claim severity and claim frequency. As an example of the former we focus on the Gamma model, which is fully unexplored in this context, and for the latter we focus on the Poisson model, which is hitherto only partially explored. The role of the Gamma and the Poisson models in actuarial science is studied by Jorgensen and De Souza (1994), Smyth and Jorgensen (2002), Xacur and Garrido (2015), Garrido, Genest and Schulz (2016), to mention only a few. For Bayesian estimation of the mean parameter of these distributions see Landsman and Makov (1998, 1999), Ntzoufras, Katsis and Karlis (2005), Gschlößl and Czado (2007), among others. Additionally, we illustrate the methodology with simulated data and compare our outcomes with the Bayes estimator, which is the posterior mean based on quadratic loss function and Jeffrey's prior distribution.

This study is organized as follows: Section 2 presents the intrinsic objective Bayesian point estimation methodology, Section 3 explores and illustrates the Gamma model and Section 4 explores and illustrates the Poisson model. Conclusions are given in Section 5.

2 The Intrinsic Objective Bayesian point estimation methodology

Let $\mathbf{x} = \{x_1, x_2, ..., x_n\}$ be a set of conditionally independent and identically distributed observations from an underlying model

$$M = \{p(x|\theta), x \in X, \theta \in \Theta\},$$

where θ is an unknown parameter. The Bayesian approach assumes that θ is a random variable having a prior distribution and that the posterior distribution of θ provides the up-to-date knowledge about θ given $\mathbf{x}$.

In this study we apply Jeffrey's prior distribution, which, for a single parameter model, is

$$p_J(\theta) \propto \sqrt{i_n(\theta)},$$

where $i_n(\theta)$ measures the amount of information about the parameter with respect to the information inherent in the data $\mathbf{x}$,

$$i_n(\theta) = n \cdot i(\theta),$$

where

$$i(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\log p(x|\theta)\right)^2\right].$$

Therefore, the posterior distribution of θ is given by Bayes formula

$$p(\theta|\mathbf{x}) = \frac{p(\mathbf{x}|\theta) \cdot p_J(\theta)}{\int p(\mathbf{x}|\theta) \cdot p_J(\theta)\, d\theta} \propto p(\mathbf{x}|\theta) \cdot \sqrt{n \cdot i(\theta)}. \quad (1)$$

For Bayesian estimation, it is essential to choose an appropriate loss function associated with the estimation problem. In this study we adopt the IDLF.

2.1 The Intrinsic Discrepancy Loss function (IDLF)

The IDLF (Bernardo and Juarez 2003) is defined by

$$\ell_{IDL|\mathbf{x}}\left\{\tilde{\theta}, \theta\right\} = n \cdot \min\left\{k\left(\tilde{\theta}|\theta\right), k\left(\theta|\tilde{\theta}\right)\right\}, \quad (2)$$

where $\tilde{\theta}$ is an estimate of θ , $k\left(\tilde{\theta}|\theta\right) = \int \log \frac{p(x|\theta)}{p\left(x|\tilde{\theta}\right)} \cdot p\left(x|\theta\right) dx$ is the Kullback-Leibeler divergence between $p\left(x|\tilde{\theta}\right)$ and $p\left(x|\theta\right)$, and similarly, $k\left(\theta|\tilde{\theta}\right) = \int \log \frac{p\left(x|\tilde{\theta}\right)}{p(x|\theta)} \cdot p\left(x|\tilde{\theta}\right) dx$ is the Kullback-Leibeler divergence between $p\left(x|\theta\right)$ and $p\left(x|\tilde{\theta}\right)$.

In information theory, $k\left(\tilde{\theta}|\theta\right)$ is a measure of the information loss when $p\left(x|\tilde{\theta}\right)$ is used to approximate $p\left(x|\theta\right)$. Thus, $k\left(\cdot|\cdot\right)$ is smaller the more accurate is the approximation. Clearly, the IDLF measures the loss caused by incorrect estimation, relying solely on the nature of the underlying.

The IDLF, which is an attractive loss function mainly because it provides an invariant Bayes estimator, has the following properties (Bernardo 2005b):

- Symmetry
- Non-negativity: $\ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\} \geq 0$, and $\ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\} = 0$ if and only if, $p\left(x|\tilde{\theta}\right) = p\left(x|\theta\right)$ almost everywhere.
- Invariance: if $y = y\left(x\right)$ is one to one transformation and $p\left(y|\theta\right)$ is the probability density of y induced by $p\left(x|\theta\right)$, then $\ell_{IDL|y}\left\{\tilde{\theta}, \theta\right\} = \ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\}$.
- Additivity for independent observation: if $\mathbf{x} = \{x_1, x_2, ..., x_n\}$ and the x_j's are independent observations, then $\ell_{IDL|\mathbf{x}}\left\{\tilde{\theta}, \theta\right\} = \sum \ell_{IDL|x_j}\left\{\tilde{\theta}, \theta\right\}$. If they are also identically distributed, then $\ell_{IDL|\mathbf{x}}\left\{\tilde{\theta}, \theta\right\} = n \cdot \ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\}$.
- Consistently marginalization: if $t = t\left(\mathbf{x}\right) \in T$ is sufficient for the underlying model, then $\ell_{IDL|t}\left\{\tilde{\theta}, \theta\right\} = \ell_{IDL|\mathbf{x}}\left\{\tilde{\theta}, \theta\right\}$.

2.2 The Intrinsic Objective Bayesian point estimator

From Bayesian perspective it is well known that the point estimator of the quantity of interest is the value which minimizes the posterior expected loss.

Let $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$ is the intrinsic objective Bayesian risk function, given by

$$R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)=E\left(\ell_{IDL|\mathbf{x}}\left\{\tilde{\theta},\theta\right\}|\mathbf{x}\right)=\int\ell_{IDL|\mathbf{x}}\left\{\tilde{\theta},\theta\right\}\cdot p\left(\theta|\mathbf{x}\right)d\theta, \tag{3}$$

where $\ell_{IDL|\mathbf{x}}\left\{\tilde{\theta},\theta\right\}$ is the IDLF according to (2) and $p\left(\theta|\mathbf{x}\right)$ is the posterior distribution of θ which is derived from Jeffrey's prior according to (1).

Clearly, the intrinsic objective Bayesian point estimator of θ, θ_{IE}, is derived from

$$\theta_{IE}=\arg\min_{\tilde{\theta}\in\Theta}R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right). \tag{4}$$

2.3 The Approximate Intrinsic Objective Bayesian point estimator

Bernardo (2005b) provided an asymptotic approximation formula to θ_{IE} using a *reference parameterization* $\phi=\phi\left(\theta\right)$, the indefinite integral obtained by

$$\phi\left(\theta\right)=\int\sqrt{i_n\left(\theta\right)}d\theta, \tag{5}$$

where $\phi=\phi\left(\theta\right)$ provides a variance stabilizing transformation. So, the estimator $\tilde{\phi}_n=\phi\left(\tilde{\theta}_n\right)=\tilde{\phi}_n\left(\mathbf{x}\right)$ behaves asymptotically as $normal\left(\phi,1/n\right)$, where $\tilde{\theta}_n$ is sufficient and consistent estimator of θ, if exists. Bernardo showed that the approximate intrinsic objective Bayesian point estimator of ϕ, ϕ_{AIE}, is $E_{\theta|\mathbf{x}}\left(\phi\left(\theta\right)\right)$, which is determined by

$$E_{\theta|\mathbf{x}}\left(\phi\left(\theta\right)\right)=\int\phi\left(\theta\right)\cdot p\left(\theta|\mathbf{x}\right)d\theta=\phi_{AIE}. \tag{6}$$

Accordingly, the approximate intrinsic objective Bayesian point estimator of θ is given by

$$\theta_{AIE}=\theta\left(\phi_{AIE}\right), \tag{7}$$

where $\theta\left(\phi\right)$ is the inverse function of $\phi\left(\theta\right)$.

In the next two sections, we explore the intrinsic objective Bayesian point estimation for the mean of the Gamma and Poisson models.

3 Gamma model

In this section, we explore the Gamma distribution, which is very often used for modeling claim size or severity, and propose the intrinsic objective Bayesian point estimator of the mean claim size, based on the intrinsic objective Bayesian point estimation for θ. The Gamma model has not yet been explored in the context of this study, except for the special case of the exponential distribution (Bernardo 2005b).

Consider $\mathbf{x} = \{x_1, x_2, ..., x_n\}$, a set of conditionally i.i.d claim severity observations from a Gamma model with unknown rate parameter θ and known shape parameter λ, following the probability density function ($Gamma\,(\lambda, \theta)$)

$$g\,(x|\lambda, \theta) = \frac{\exp\,(-\theta x) \cdot x^{\lambda-1} \cdot \theta^{\lambda}}{\Gamma\,(\lambda)}, \quad x \geq 0, \lambda > 0, \theta > 0. \tag{8}$$

Theorem 1 *Let* $\mathbf{x} = \{x_1, x_2, ..., x_n\}$ *be a random sample from a Gamma model with unknown rate parameter* θ *and known shape parameter* λ*. Then the intrinsic discrepancy loss function (IDLF) of the Gamma model, associated with estimating* θ *by* $\tilde{\theta}$*, is*

$$\ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\} = n\lambda \cdot \min\left\{\frac{\tilde{\theta}}{\theta} - 1 - \log\frac{\tilde{\theta}}{\theta}, \frac{\theta}{\tilde{\theta}} - 1 - \log\frac{\theta}{\tilde{\theta}}\right\},$$

or,

$$\ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\} = \left\{\begin{array}{ll} n\lambda \cdot \left(\frac{\tilde{\theta}}{\theta} - 1 - \log\frac{\tilde{\theta}}{\theta}\right) & \theta \geq \tilde{\theta} \\ n\lambda \cdot \left(\frac{\theta}{\tilde{\theta}} - 1 - \log\frac{\theta}{\tilde{\theta}}\right) & \theta \leq \tilde{\theta} \end{array}\right. . \tag{9}$$

The proof is given in Appendix 6.1 .

Figure 1 shows the curve of the IDLF of $Gamma\,(\lambda = 0.1, \theta)$ model where $n = 10$ and $\theta = 0.05, 0.1, 0.2$. As can be seen, the IDLF is an asymmetric function, penalizing underestimation more than overestimation.

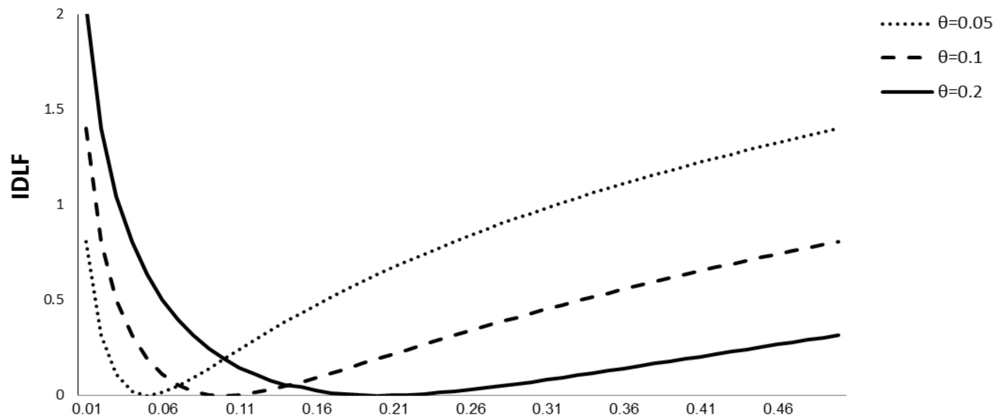

Figure 1: $\tilde{\theta} \mapsto \ell_{IDL|\mathbf{x}} \left\{\tilde{\theta}, \theta\right\}$ for $n = 10$ and $\theta = 0.05, 0.1, 0.2$.

3.1 The Intrinsic objective Bayesian point estimator of θ

The intrinsic objective Bayesian point estimator of θ in the Gamma model is the value which minimizes the $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$ of the Gamma model. This is provided in the following theorem.

Theorem 2 *Let $\mathbf{x} = \{x_1, x_2, ..., x_n\}$ be a random sample from a Gamma model with unknown rate parameter θ and known shape parameter λ. Then the intrinsic objective Bayesian risk function of the Gamma model, associated with estimating θ by $\tilde{\theta}$, is*

$$\begin{aligned} R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right) &= n\lambda \cdot \left[\frac{n\lambda}{\tilde{\theta}t} \cdot G_{(n\lambda+1,t)}\left(\tilde{\theta}\right) + \frac{\tilde{\theta}t}{\Gamma(n\lambda)} \cdot \Gamma\left(n\lambda - 1, \tilde{\theta}t\right)\right. \\ &\quad -1 + \log\tilde{\theta} \cdot \left(2G_{(n\lambda,t)}\left(\tilde{\theta}\right) - 1\right) \\ &\quad + \log t \cdot \left(2G_{(n\lambda,1)}\left(t\tilde{\theta}\right) - 1\right) \\ &\quad \left. + \frac{\psi\left(n\lambda, \tilde{\theta}t\right) \cdot \Gamma\left(n\lambda, \tilde{\theta}t\right) - \psi^*\left(n\lambda, \tilde{\theta}t\right) \cdot \gamma\left(n\lambda, \tilde{\theta}t\right)}{\Gamma(n\lambda)}\right], \end{aligned} \tag{10}$$

where $t=\sum_{i=1}^{n}x_i$, $G_{(a,b)}(x)$ *is Gamma cumulative distribution function with shape parameter* a *and rate parameter* b, $\Gamma(\alpha,\eta)$ *is the upper incomplete gamma function with shape parameter* α, $\gamma(\alpha,\eta)$ *is the lower incomplete gamma function with shape parameter* α, $\psi(\alpha,\eta)$ *is the upper incomplete digamma function defined as* $\frac{\partial}{\partial\alpha}\ln\Gamma(\alpha,\eta)$ *and* $\psi^*(\alpha,\eta)$ *is the lower incomplete digamma function defined as* $\frac{\partial}{\partial\alpha}\ln\gamma(\alpha,\eta)$.

Proof. The intrinsic objective Bayesian risk function is obtained from (3) and (9). Noting that $g(\theta|n\lambda,t)$ is the posterior Gamma density function with respect to θ, with shape parameter $n\lambda$ and rate parameter t, and that $p_J(\theta)\propto\sqrt{\frac{n\lambda}{\theta^2}}$ (Yang and Berger 1998). Thus,

$$\begin{aligned} R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right) &= \int_{o}^{\tilde{\theta}} n\lambda\left(\frac{\theta}{\tilde{\theta}}-1-\log\frac{\theta}{\tilde{\theta}}\right)\cdot g(\theta|n\lambda,t)\,d\theta \\ &\quad+\int_{\tilde{\theta}}^{\infty} n\lambda\left(\frac{\tilde{\theta}}{\theta}-1-\log\frac{\tilde{\theta}}{\theta}\right)\cdot g(\theta|n\lambda,t)\,d\theta, \end{aligned} \tag{11}$$

which leads to

$$\begin{aligned} R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right) &= n\lambda\cdot\Bigg[\frac{n\lambda}{\tilde{\theta}t}\cdot G_{(n\lambda+1,t)}\left(\tilde{\theta}\right)+\frac{\tilde{\theta}t}{\Gamma(n\lambda)}\cdot\Gamma\left(n\lambda-1,\tilde{\theta}t\right) \\ &\quad -1+\log\tilde{\theta}\cdot\left(2G_{(n\lambda,t)}\left(\tilde{\theta}\right)-1\right) \\ &\quad +\int_{\tilde{\theta}}^{\infty}\log\theta\cdot g(\theta|n\lambda,t)\,d\theta \\ &\quad -\int_{o}^{\tilde{\theta}}\log\theta\cdot g(\theta|n\lambda,t)\,d\theta\Bigg]. \end{aligned}$$

Let us notice that

$$\begin{aligned}
\int_{\tilde{\theta}}^{\infty} \log\theta \cdot g(\theta|n\lambda, t)\, d\theta &= \frac{t^{n\lambda}}{\Gamma(n\lambda)} \int_{\tilde{\theta}}^{\infty} \log\theta \cdot \exp(-t\theta) \cdot \theta^{n\lambda-1} d\theta = \\
&= [\Gamma(n\lambda)]^{-1} \int_{z=t\tilde{\theta}}^{\infty} \log z \cdot \exp(-z) \cdot z^{n\lambda-1} dz - \\
&\quad - \log t \cdot \left[1 - G_{(n\lambda,1)}\left(t\tilde{\theta}\right)\right] \\
&= [\Gamma(n\lambda)]^{-1} \cdot \psi\left(n\lambda, t\tilde{\theta}\right) \cdot \Gamma\left(n\lambda, t\tilde{\theta}\right) \\
&\quad - \log t \cdot \left[1 - G_{(n\lambda,1)}\left(t\tilde{\theta}\right)\right],
\end{aligned}$$

In a similar way,

$$\begin{aligned}
\int_{o}^{\tilde{\theta}} \log\theta \cdot g(\theta|n\lambda, t)\, d\theta &= [\Gamma(n\lambda)]^{-1} \cdot \psi^{*}\left(n\lambda, t\tilde{\theta}\right) \cdot \gamma\left(n\lambda, t\tilde{\theta}\right) \\
&\quad - \log t \cdot G_{(n\lambda,1)}\left(t\tilde{\theta}\right).
\end{aligned}$$

■

As can be seen, the analytical expression of the intrinsic objective Bayesian risk function of the Gamma model is intractable so that there is a need for numerical integration in order to find θ_{IE} according to (4). In the next section, we find an approximating formula for θ_{IE}.

3.2 The Approximate Intrinsic Objective Bayesian point estimator of θ

Theorem 3 *Let* $\mathbf{x} = \{x_1, x_2, ..., x_n\}$ *be a random sample from Gamma model with unknown rate parameter* θ *and known shape parameter* λ*. Then the approximate intrinsic objective Bayesian point estimator of* θ*,* θ_{AIE}*, is*

$$\theta_{AIE} = \frac{\exp(\psi(n\lambda))}{t},$$

where $\psi(\alpha)$ *is the digamma function with shape parameter* α*.*

Proof. The approximate intrinsic objective Bayesian point estimator of θ is obtained from (7), where $\phi(\theta)$ is the reference parameterization, which is achieved by

$$\phi(\theta) = \int \frac{\sqrt{n\lambda}}{\theta} d\theta = \sqrt{n\lambda} \cdot \log\theta,$$

and ϕ_{AIE} ,according to (6), is given by

$$\begin{aligned} \phi_{AIE} &= \int_o^\infty \sqrt{n\lambda} \cdot \log\theta \cdot g(\theta|n\lambda, t)\, d\theta \qquad (12) \\ &= \frac{\sqrt{n\lambda}}{\Gamma(n\lambda)} \cdot \int_o^\infty \log\theta \cdot \exp(-t\theta) \cdot \theta^{n\lambda-1} \cdot t^{n\lambda} d\theta \\ &= \frac{\sqrt{n\lambda}}{\Gamma(n\lambda)} \cdot \left(\frac{\partial}{\partial(n\lambda)} \Gamma(n\lambda) - \Gamma(n\lambda) \cdot \log t \right) \\ &= \sqrt{n\lambda} \cdot [\psi(n\lambda) - \log t], \end{aligned}$$

since

$$
\begin{aligned}
\frac{\partial}{\partial(n\lambda)}\Gamma(n\lambda) &= \frac{\partial}{\partial(n\lambda)}\int\limits_{o}^{\infty}\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}d\theta \\
&= \int\limits_{o}^{\infty}\frac{\partial}{\partial(n\lambda)}\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}d\theta \\
&= \int\limits_{o}^{\infty}\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}\cdot(\log\theta+\log t)\,d\theta \\
&= \int\limits_{o}^{\infty}\log\theta\cdot\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}d\theta + \\
&\quad +\int\limits_{o}^{\infty}\log t\cdot\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}d\theta \\
&= \int\limits_{o}^{\infty}\log\theta\cdot\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}d\theta + \\
&\quad +\Gamma(n\lambda)\cdot\log t\cdot\int\limits_{o}^{\infty}\frac{\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}}{\Gamma(n\lambda)}d\theta \\
&= \int\limits_{o}^{\infty}\log\theta\cdot\exp(-t\theta)\cdot\theta^{n\lambda-1}\cdot t^{n\lambda}d\theta+\Gamma(n\lambda)\cdot\log t.
\end{aligned}
$$

Let us notice that the inverse function of $\phi(\theta)$ is

$$\theta(\phi)=\exp\left(\frac{\phi}{\sqrt{n\lambda}}\right), \tag{13}$$

so substituting (12) into (13) leads to θ_{AIE}. ■

3.3 The Intrinsic Objective Bayesian point estimation of the mean value for Gamma model

An important actuarial problem is the evaluation of the mean claim size of an insurance contract. This section explores the intrinsic objective Bayesian point estimator and the approximate intrinsic objective Bayesian point estimator of the mean value of the Gamma model. In this context, we will explore both cases when λ is either known or unknown.

3.3.1 Known shape parameter

The intrinsic objective Bayesian point estimation leads to invariant estimators under monotonic transformations. In the Gamma model, the mean value is one to one transformation of the rate parameter θ. Therefore, the intrinsic objective Bayesian point estimator of the mean value, μ_{IE}, is

$$\mu_{IE} = \frac{\lambda}{\theta_{IE}}, \tag{14}$$

where θ_{IE} is derived from (4).

Subsequently, the approximate intrinsic objective Bayesian point estimator of the mean value, μ_{AIE}, is

$$\mu_{AIE} = \frac{\lambda}{\theta_{AIE}} = \frac{\lambda \cdot t}{\exp\left(\psi\left(n\lambda\right)\right)}. \tag{15}$$

The Bayes estimator is derived as follows,

$$p\left(\mu|\mathbf{x}\right) \propto \exp\left(-\frac{\lambda}{\mu}t\right) \cdot \left(\frac{\lambda}{\mu}\right)^{n\lambda} \cdot \frac{1}{\mu},$$

where the Jeffrey's prior is $p_J\left(\mu\right) \propto \frac{1}{\mu}$ and $t = \sum_{i=1}^{n} x_i$. Clearly, $\mu|\mathbf{x} \sim Inv-gamma\left(n\lambda, \lambda t\right)$ and the Bayes estimator is given by

$$\mu_{Bays} = \frac{\lambda \cdot t}{n\lambda - 1}. \tag{16}$$

Simulation study For numerical illustrations of the three estimators (14) ,(15) and (16), we generated 2000 samples of size $n = 3, 4, 5, 7, 10, 15, 20, 30$ and 50 from $Gamma\left(\lambda = 0.4, \theta = 0.2\right)$ model with mean $\mu = 2$.

In order to compare these estimators we computed three measures for each sample size:

- The mean of estimate which is the average of 2000 estimates.
- The MSE of the 2000 estimates.

- The Mean Intrinsic Risk (MIR) of 2000 estimates defined as

$$MIR\left(\mu_E\right) = \frac{1}{2000 \cdot n} \cdot \sum_{i=1}^{2000} R_{IOR}\left(\mu_{E_i}|\mathbf{x}\right),$$

where μ_{E_i} is the estimator obtained from the i replication, $i = 1, 2, ..., 2000$.

Remark 1 *Both the IDLF and the Intrinsic estimators are invariant under one to one transformations. Therefore,* $MIR\left(\mu_{IE}\right) = MIR\left(\theta_{IE}\right)$ *and* $MIR\left(\mu_{AIE}\right) = MIR\left(\theta_{AIE}\right)$ *but* $MIR\left(\mu_{Bays}\right) \neq MIR\left(\theta_{Bays}\right)$ *(see Appendix 6.2).*

Table 1 presents the performance of these three estimators with respect to the mean estimate, MSE and MIR, for each sample size.

Table 1: Mean Estimate, MSE and MIR of μ_{IE} , μ_{AIE} and μ_{Bays} for the Gamma model(given shape parameter) with $\mu = 2$, calculated from 2,000 simulated for different sample sizes.

		μ_{Bays}	μ_{AIE}	μ_{IE}
	Mean Estimate	11.97	3.196	2.907
$n = 3$	MSE	216.3	9.767	7.722
	MIR	0.3302	0.14898	0.14807
	Mean Estimate	5.287	2.796	2.631
$n = 4$	MSE	27.32	5.257	4.492
	MIR	0.1614	0.1099	0.1094
	Mean Estimate	3.94	2.581	2.472
$n = 5$	MSE	11.59	3.702	3.308
	MIR	0.1121	0.0872	0.0869
	Mean Estimate	3.133	2.434	2.371
$n = 7$	MSE	4.693	2.244	2.089
	MIR	0.0718	0.0621	0.0619
	Mean Estimate	2.702	2.308	2.273
$n = 10$	MSE	2.310	1.421	1.362
	MIR	0.04757	0.04353	0.04349
	Mean Estimate	2.394	2.173	2.155
$n = 15$	MSE	1.106	0.814	0.795
	MIR	0.03088	0.02928	0.02927
	Mean Estimate	2.301	2.146	2.135
$n = 20$	MSE	0.699	0.550	0.541
	MIR	0.02269	0.02214	0.02214
	Mean Estimate	2.189	2.094	2.088
$n = 30$	MSE	0.436	0.375	0.371
	MIR	0.0153	0.01496	0.01496
	Mean Estimate	2.11	2.056	2.053
$n = 50$	MSE	0.238	0.218	0.217
	MIR	0.00923	0.00913	0.00913

We can see from Table 1 that the quality of the two proposed estimators, μ_{IE} and μ_{AIE}, for the mean value of the Gamma model are superior to the Bayes estimator, μ_{Bays}, for all samples size. We note that for small samples there is a large difference between the proposed estimators and the Bayes estimator. Furthermore, we note that the MSE and MIR are smaller for the μ_{IE} and μ_{AIE} than for μ_{Bays}, especially for small samples. Finally, while the

Bayes estimator is limited to $n\lambda > 1$, the two proposed estimators are not constrained. Moreover, from Table 1, we can see that the performances of μ_{IE} is slightly better than the performances of μ_{AIE} and that this difference is reduced as the sample size increases.

We have experimented with a range of λ's values and the results are quite similar, lending support to the advantage of the IDLF. We note that these conclusions hold for other values of λ tested.

3.3.2 Unknown shape parameter

This section explores the realistic case of an unknown shape parameter. We investigate this case by using the Maximum-likelihood estimation (mle) method. We exploit a property of the exponential dispersion model (Bar-Lev and Landsman 2006) stating that μ and λ are orthogonal. consequently, the intrinsic objective Bayesian point estimator, the approximate intrinsic objective Bayesian point estimator and the Bayes estimator of the mean value for Gamma model, when the shape parameter is unknown, can be approximated by replacing in section 3.3.1 λ by it's $\hat{\lambda}_{mle}$.

Therefore, the intrinsic objective Bayesian point estimator of the mean value is

$$\mu_{IE} = \frac{\hat{\lambda}_{mle}}{\theta_{IE}}, \tag{17}$$

where θ_{IE} is derived from (4).

Subsequently, the approximate intrinsic objective Bayesian point estimator of the mean value is

$$\mu_{AIE} = \frac{\hat{\lambda}_{mle} \cdot t}{\exp\left(\psi\left(n\hat{\lambda}_{mle}\right)\right)}. \tag{18}$$

And the Bayes estimator is

$$\mu_{Bays} = \frac{\hat{\lambda}_{mle} \cdot t}{n\hat{\lambda}_{mle} - 1}. \tag{19}$$

Simulation study The numerical illustrations of these three estimators (17) ,(18) and (19), for the mean parameter of $Gamma\,(\lambda, \theta)$ model with unknown rate parameter and unknown shape parameter is based on the same data used in section 3.3.1. The $\hat{\lambda}_{mle}$ is obtained numerically.

Table 2 presents the performance of these three estimators in terms of the mean estimate, MSE and MIR, for each sample size.

Table 2: Mean Estimate, MSE and MIR of μ_{IE} , μ_{AIE} and μ_{Bays} for the Gamma model (unknown shape parameter) with $\mu = 2$, calculated from 2,000 simulated for different sample sizes.

		μ_{Bays}	μ_{AIE}	μ_{IE}
	Mean Estimate		2.923	2.666
$n = 3$	MSE		9.685	6.845
	MIR		0.1488	0.1481
	Mean Estimate		2.674	2.518
$n = 4$	MSE		5.385	4.368
	MIR		0.1103	0.1099
	Mean Estimate		2.509	2.404
$n = 5$	MSE		3.774	3.273
	MIR		0.0878	0.0876
	Mean Estimate		2.382	2.326
$n = 7$	MSE		2.135	1.989
	MIR		0.0624	0.0623
	Mean Estimate	2.669	2.287	2.254
$n = 10$	MSE	2.368	1.403	1.343
	MIR	0.04762	0.04376	0.04373
	Mean Estimate	2.377	2.164	2.147
$n = 15$	MSE	1.093	0.804	0.786
	MIR	0.0309	0.0294	0.0294
	Mean Estimate	2.294	2.142	2.131
$n = 20$	MSE	0.695	0.547	0.538
	MIR	0.0231	0.0222	0.0222
	Mean Estimate	2.187	2.092	2.086
$n = 30$	MSE	0.436	0.374	0.371
	MIR	0.01534	0.01498	0.01498
	Mean Estimate	2.11	2.056	2.053
$n = 50$	MSE	0.238	0.218	0.217
	MIR	0.00926	0.00914	0.00914

Notice that the Bayes estimator is not available for $n = 3, 4, 5, 7$ since $n\hat{\lambda}_{mle} \leqslant 1$. Evidently, the two proposed estimators given in (17) ,(18) are superior to the Bayes estimator, as for the case of known shape parameter.

4 Poisson model

In this section, we explore the Poisson model which is very often used for modeling claim frequency, and find the intrinsic objective Bayesian point estimation of the mean claim frequency.

Consider $\mathbf{x} = \{x_1, x_2, ..., x_n\}$, a set of conditionally i.i.d claim frequency observations from a Poisson model with unknown mean parameter,θ, following the probability function

$$p(x|\theta) = \frac{\exp(-\theta) \cdot \theta^x}{x!}, \theta > 0, x = 0, 1, 2, ...$$

Bernardo (2011) obtained the IDLF of Poisson model as follows

$$\ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\} = n \cdot \min\left\{\theta - \tilde{\theta} + \tilde{\theta} \cdot \log\frac{\tilde{\theta}}{\theta}, \tilde{\theta} - \theta + \theta \cdot \log\frac{\theta}{\tilde{\theta}}\right\},$$

or,

$$\ell_{IDL|x}\left\{\tilde{\theta}, \theta\right\} = \begin{cases} n \cdot \left(\theta - \tilde{\theta} + \tilde{\theta} \cdot \log\frac{\tilde{\theta}}{\theta}\right) & \theta \geq \tilde{\theta} \\ n \cdot \left(\tilde{\theta} - \theta + \theta \cdot \log\frac{\theta}{\tilde{\theta}}\right) & \theta \leq \tilde{\theta} \end{cases}. \tag{20}$$

Figure 2 shows the curve of this IDLF where $n = 1$ and $\theta = 2, 5, 10$. As can be seen, the IDLF is an asymmetric function, penalizing underestimation more than overestimation.

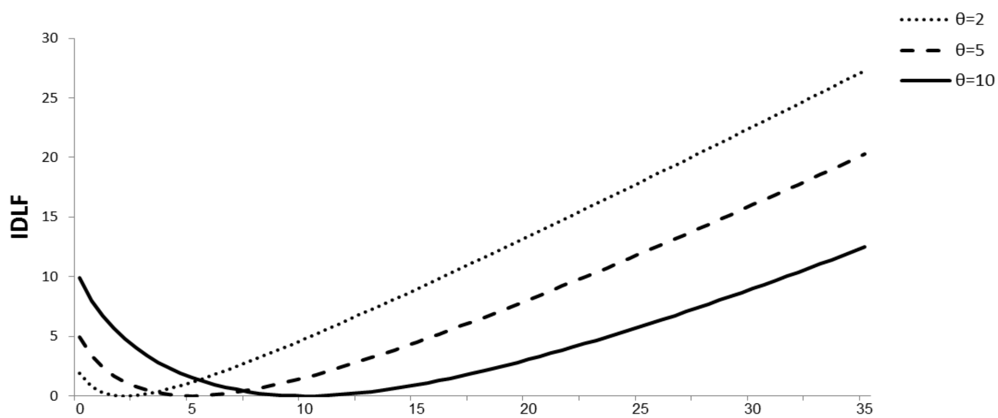

Figure 2: $\tilde{\theta} \mapsto \ell_{IDL|\mathbf{x}} \left\{ \tilde{\theta}, \theta \right\}$ for $n = 1$ and $\theta = 2, 5, 10$.

Additionally, Bernardo (2011) provided approximation to $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$ of Poisson model. In the next sections, we provide an exact expression of $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$ and an approximating formula for θ_{IE} for this model.

4.1 The Intrinsic Objective Bayesian point estimator of θ

The intrinsic objective Bayesian point estimator of θ is the value which minimizes the $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$. The exact expression of $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$ in the Poisson model is provided in the following theorem.

Theorem 4 *Let $\mathbf{x} = \{x_1, x_2, ..., x_n\}$ be a random sample from Poisson model with unknown mean parameter θ. Then the intrinsic objective Bayesian risk function of the Poisson model, associated with estimating θ by $\tilde{\theta}$, is*

$$\begin{aligned} R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right) &= n\tilde{\theta} \cdot \left[2G_{(t+0.5,n)}\left(\tilde{\theta}\right) - 1 + \log\tilde{\theta} \cdot \left(1 - G_{(t+0.5,n)}\left(\tilde{\theta}\right)\right)\right] \\ &+ (t+0.5) \cdot \left[1 - 2G_{(t+1.5,n)}\left(\tilde{\theta}\right) - \log\tilde{\theta} \cdot G_{(t+1.5,n)}\left(\tilde{\theta}\right)\right] \\ &- \log n \cdot \left[(t+0.5) \cdot G_{(t+1.5,1)}\left(n\tilde{\theta}\right) + n\tilde{\theta} \cdot \left(1 - G_{(t+0.5,1)}\left(n\tilde{\theta}\right)\right)\right] \\ &+ \left[\Gamma\left(t+1.5\right)\right]^{-1} \cdot \psi^*\left(t+1.5, n\tilde{\theta}\right) \cdot \gamma\left(t+1.5, n\tilde{\theta}\right) \\ &- \left[\Gamma\left(t+0.5\right)\right]^{-1} \cdot n\tilde{\theta} \cdot \psi\left(t+0.5, n\tilde{\theta}\right) \cdot \Gamma\left(t+0.5, n\tilde{\theta}\right), \end{aligned}$$

where $t = \sum_{i=1}^{n} x_i$, $G_{(a,b)}(x)$ *is Gamma cumulative distribution function with shape parameter* a *and rate parameter* b, $\Gamma(\alpha, \eta)$ *is the upper incomplete gamma function with shape parameter* α, $\gamma(\alpha, \eta)$ *is the lower incomplete gamma function with shape parameter* α, $\psi(\alpha, \eta)$ *is the upper incomplete digamma function defined as* $\frac{\partial}{\partial \alpha} \ln \Gamma(\alpha, \eta)$ *and* $\psi^*(\alpha, \eta)$ *is the lower incomplete digamma function defined as* $\frac{\partial}{\partial \alpha} \ln \gamma(\alpha, \eta)$.

Proof. The intrinsic objective Bayesian risk function is obtained from (3) and (20). Noting that $g(\theta|t + 0.5, n)$ is the posterior Gamma density function with respect to θ, with shape parameter $t + 0.5$ and rate parameter n , and that $p_J(\theta) \propto \sqrt{\frac{n}{\theta}}$ (Yang and Berger 1998). Thus,

$$
\begin{aligned}
R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right) &= \int_{o}^{\tilde{\theta}} n\left(\tilde{\theta} - \theta + \theta \cdot \log \frac{\theta}{\tilde{\theta}}\right) \cdot g(\theta|t + 0.5, n)\, d\theta + \\
&\quad \int_{\tilde{\theta}}^{\infty} n\left(\theta - \tilde{\theta} + \tilde{\theta} \cdot \log \frac{\tilde{\theta}}{\theta}\right) \cdot g(\theta|t + 0.5, n)\, d\theta,
\end{aligned}
$$

which leads to

$$
\begin{aligned}
R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right) &= n\tilde{\theta} \cdot G_{(t+0.5,n)}\left(\tilde{\theta}\right) - (t + 0.5) \cdot G_{(t+1.5,n)}\left(\tilde{\theta}\right) \\
&\quad + (t + 0.5) \cdot \int_{o}^{\tilde{\theta}} \log \theta \cdot g(\theta|t + 1.5, n)\, d\theta \\
&\quad - \log \tilde{\theta} \cdot (t + 0.5) \cdot G_{(t+1.5,n)}\left(\tilde{\theta}\right) \\
&\quad + (t + 0.5) \cdot \left[1 - G_{(t+1.5,n)}\left(\tilde{\theta}\right)\right] \\
&\quad - n\tilde{\theta} \cdot \left[1 - G_{(t+0.5,n)}\left(\tilde{\theta}\right)\right] + \\
&\quad + n\tilde{\theta} \cdot \log \tilde{\theta} \cdot \left[1 - G_{(t+0.5,n)}\left(\tilde{\theta}\right)\right] \\
&\quad - n\tilde{\theta} \cdot \int_{\tilde{\theta}}^{\infty} \log \theta \cdot g(\theta|t + 0.5, n)\, d\theta.
\end{aligned}
$$

Let us notice that

$$\begin{aligned}
\int_{\tilde{\theta}}^{\infty} \log\theta \cdot g\left(\theta|t+0.5,n\right)d\theta &= \frac{n^{t+0.5}}{\Gamma\left(t+0.5\right)}\int_{\tilde{\theta}}^{\infty} \log\theta \cdot \exp\left(-n\theta\right)\cdot \theta^{t-0.5}d\theta = \\
&= \left[\Gamma\left(t+0.5\right)\right]^{-1}\int_{z=n\tilde{\theta}}^{\infty} \log z \cdot \exp\left(-z\right)\cdot z^{t-0.5}dz \\
&\quad -\log n \cdot \left[1-G_{(t+0.5,1)}\left(n\tilde{\theta}\right)\right] \\
&= \left[\Gamma\left(t+0.5\right)\right]^{-1}\cdot \psi\left(t+0.5,n\tilde{\theta}\right)\cdot \Gamma\left(t+0.5,n\tilde{\theta}\right) \\
&\quad -\log n \cdot \left[1-G_{(t+0.5,1)}\left(n\tilde{\theta}\right)\right],
\end{aligned}$$

In a similar way,

$$\begin{aligned}
\int_{o}^{\tilde{\theta}} \log\theta \cdot g\left(\theta|t+1.5,n\right)d\theta &= \left[\Gamma\left(t+1.5\right)\right]^{-1}\cdot \psi^{*}\left(t+1.5,n\tilde{\theta}\right)\cdot \gamma\left(t+1.5,n\tilde{\theta}\right) \\
&\quad -\log n \cdot \left[1-G_{(t+1.5,1)}\left(n\tilde{\theta}\right)\right].
\end{aligned}$$

■

As can be seen, the analytical expression of $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$ for Poisson model is intractable so that there is a need for numerical integration in order to find θ_{IE} according to (4). In the next section, we find an approximating formula for θ_{IE}.

4.2 The Approximate Intrinsic Objective Bayesian point estimator of θ

Theorem 5 *Let* $\mathbf{x} = \{x_1, x_2, ..., x_n\}$ *be a random sample from Poisson model with unknown mean parameter* θ*. Then the approximate intrinsic objective Bayesian point estimator of* θ*,* θ_{AIE}*, is*

$$\theta_{AIE} = \frac{\pi}{n}\cdot\left(B\left(t+0.5,0.5\right)\right)^{-2}, \tag{21}$$

where $B\left(a,b\right)$ *is the beta function.*

Proof. The approximate intrinsic objective Bayesian point estimator of θ is obtained from (7), where $\phi(\theta)$ is the reference parameterization, which is given by

$$\phi(\theta) = \int \sqrt{\frac{n}{\theta}} d\theta = 2\sqrt{n \cdot \theta},$$

and ϕ_{AIE} according to (6), is given by

$$\begin{aligned}
\phi_{AIE} &= \int_o^\infty 2\sqrt{n \cdot \theta} \cdot g(t + 0.5, n)\, d\theta && (22) \\
&= \int_o^\infty 2\sqrt{n \cdot \theta} \cdot \frac{\exp(-n\theta) \cdot \theta^{t+0.5-1} \cdot n^{t+0.5}}{\Gamma(t+0.5)} d\theta \\
&= \frac{2 \cdot \Gamma(t+1)}{\Gamma(t+0.5)} \cdot \int_o^\infty \frac{\exp(-n\theta) \cdot \theta^{t} \cdot n^{t+1}}{\Gamma(t+1)} d\theta \\
&= \frac{2 \cdot \Gamma(t+1)}{\Gamma(t+0.5)} = \frac{2 \cdot \Gamma(t+1) \cdot \Gamma(0.5)}{\Gamma(t+0.5) \cdot \Gamma(0.5)} \\
&= \frac{2 \cdot \Gamma(0.5)}{B(t+0.5, 0.5)} = \frac{2 \cdot \sqrt{\pi}}{B(t+0.5, 0.5)}.
\end{aligned}$$

Let us notice that the inverse function of $\phi(\theta)$ is

$$\theta(\phi) = \left(\frac{\phi}{2\sqrt{n}}\right)^2 \tag{23}$$

so substituting (22) into (23) leads to θ_{AIE}. ■

4.3 Simulation study

In this section we compare the performance of θ_{IE} and θ_{AIE} with the Bayes estimator, $\theta_{Bayes} = \frac{t+0.5}{n}$, based on quadratic loss function and Jeffrey's prior.

We generated 2000 samples of size $n = 3, 4, 5, 7, 10, 15, 20, 30$ and 50 from Poisson model with $\theta = 2$.

Table 3 presents the performance of these three estimators with respect to the mean estimate, MSE and MIR, for each sample size.

Table 3: Mean Estimate, MSE and MIR of θ_{IE} , θ_{AIE} and θ_{Bays} for the Poisson model with parameter $\theta = 2$, calculated from 2,000 simulated for different sample sizes.

		θ_{Bays}	θ_{AIE}	θ_{IE}
	Mean Estimate	2.175	2.093	2.096
$n = 3$	MSE	0.6957	0.6722	0.6721
	MIR	0.1475	0.14604	0.14603
	Mean Estimate	2.129	2.068	2.070
$n = 4$	MSE	0.5152	0.5021	0.5020
	MIR	0.1128	0.11195	0.11194
	Mean Estimate	2.117	2.067	2.069
$n = 5$	MSE	0.4193	0.4099	0.4098
	MIR	0.0913	0.0908	0.0908
	Mean Estimate	2.069	2.033	2.034
$n = 7$	MSE	0.2984	0.2947	0.2946
	MIR	0.0662	0.0659	0.0659
	Mean Estimate	2.051	2.026	2.027
$n = 10$	MSE	0.1981	0.1962	0.1961
	MIR	0.0470	0.0468	0.0468
	Mean Estimate	2.035	2.0182	2.0186
$n = 15$	MSE	0.1384	0.1339	0.1339
	MIR	0.0317	0.0316	0.0316
	Mean Estimate	2.026	2.0135	2.0137
$n = 20$	MSE	0.0954	0.0949	0.0949
	MIR	0.02394	0.0239	0.0239
	Mean Estimate	2.019	2.011	2.011
$n = 30$	MSE	0.0684	0.0681	0.0681
	MIR	0.01609	0.01608	0.016
	Mean Estimate	2.004	1.998	1.998
$n = 50$	MSE	0.04031	0.0403	0.0403
	MIR	0.0097	0.0097	0.0097

We can seen from Table 3 that the intrinsic objective Bayesian point estimator and the approximate intrinsic objective Bayesian point estimator for the mean value of the Poisson model are superior to the Bayes estimator, for all sample sizes. We note that the θ_{IE} and the θ_{AIE} are better estimators then θ_{Bays} not only in terms of the mean but also in terms of MSE and MIR, for all sample sizes and especially for small samples. Furthermore, we note that the performance of the θ_{IE} and θ_{AIE} are almost identical.

5 Conclusions

In this paper, we have studied the objective Bayesian estimation for the mean Gamma and Poisson distributions, typically used to model claim severity and frequency. We assumed absence of prior knowledge about the unknown parameter of interest and the loss function and adopted the Jeffrey's prior distribution and the IDLF. The combined use of Jeffrey's prior distribution and the IDLF results in an objective and invariant Bayesian point estimator of the mean. We derived the intrinsic objective Bayesian point estimator and its approximation for both models. A numerical study was provided to illustrate the performance of these proposed estimators. Their superiority over the Bayes estimator is demonstrated regardless whether the exact estimator or the approximate estimator are use. The supremacy of the IDLF over the mean squared error loss function stems from the fact that the latter penalizes erroneous $\hat{\theta}$ only as a function of the distance between $\hat{\theta}$ and θ. In contrast, the IDLF penalizes $\hat{\theta}$ with respect to its impact on the divergence between $p\left(x|\hat{\theta}\right)$ and $p\left(x|\theta\right)$, thus taking into account the mathematical nature of the underlying model.

Acknowledgments

The Financial support of the Zimmerman Foundation for the study of Banking and Finance at the University of Haifa is greatly acknowledged.

References

[1] Bar-Lev, S.K. and Landsman, Z. (2006). Exponential dispersion models: Second-order minimax estimation of the mean for unknown dispersion parameter. *Journal of Statistical Planning and Inference* **136**, 3837-3851.

[2] Berger, J.O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors. *Journal of the American Statistical Association* **84**, 200-207.

[3] Berger, J.O. and Bernardo, J. M. (1992a). Ordered group reference priors with applications to a multinomial problem. *Biometrika* **79**, 25-37.

[4] Berger, J.O. and Bernardo, J. M. (1992b). Reference priors in a variance components problem. *Bayesian Analysis in Statistics and Econometrics* (P.K. Goel and N.S. Yyengar, eds.) Berline: Springer, 323-340.

[5] Berger, J.O. and Bernardo, J. M. (1992c). On the development of reference priors. *Bayesian Statistics* **4**. Oxford: University Press, 35-60.

[6] Berger, J. O., Bernardo, J. M. and Sun, D. (2009). The formal definition of reference priors. *The Annals of Statistics* **37**, 905-938.

[7] Berger, J. O., Bernardo, J. M. and Sun, D. (2013). Overall objective priors. *Bayesian Analysis* **10,**189-221.

[8] Bernardo, J. M. (1979b). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society* **B 41**, 113-147.

[9] Bernardo, J. M. (2005a). Reference analysis. *Handbook of Statistics* **25**. Amsterdam: Elsevier, pp. 17-90.

[10] Bernardo, J. M. (2005b). Intrinsic credible regions: An objective Bayesian approach to interval estimation. *Test* **14**, 317-384.

[11] Bernardo, J. M. (2007). Objective Bayesian point and region estimation in location-scale models. *Sort* **31**, 3-44.

[12] Bernardo, J. M. (2011). Bayes and discovery: Objective Bayesian hypothesis testing. *PHY- STAT 2011*, Geneva: CERN, 27-49.

[13] Bernardo, J. M. and Juarez, M. A. (2003). Intrinsic Estimation. *Bayesian Statistics* **7**, 465-476.

[14] Bernardo, J. M. and Rueda, R. (2002). Bayesian hypothesis testing: A reference approach. *International Statistical Review* **70**, 351-372.

[15] Datta, G.S. and Ghosh, M. (1996). On the invariance of noninformative priors. *The Annals of Statistics* **24**, 141-159.

[16] Demortier, L. (2009). Search Procedures in High Energy Physics. The Rockefeller University.

[17] Garrido, J., Genest, C. and Schulz, J. (2016). Generalized linear models for dependent frequency and severity of insurance claims. *Insurance: Mathematics and Economics*, **70**, 205-215.

[18] Ghosh, M. (2011). Objective priors: An introduction for frequentists. *Statistical Science* **26**, 187-202.

[19] Gradshteyn, I.S. and Ryzhik, I.M. (1965). *Table of Integrals, Series, and Products*. Academic Press. New York and London.

[20] Gschlößl, S. and Czado, C. (2007). Spatial modelling of claim frequency and claim size in non-life insurance. *Scandinavian Actuarial Journal*, **3**, 202-225.

[21] Jeffreys, H. (1961). *Theory of Probability. 3rd ed.* London: Oxford University Press.

[22] Jorgensen, B. (1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society* **B 49**, 127-162.

[23] Jorgensen, B. (1997). *The Theory of Dispersion Models*. Chapman & Hall, London.

[24] Jorgensen, B. and Paes De Souza, M. C. (1994). Fitting Tweedie's compound Poisson model to insurance claims data. *Scandinavian Actuarial Journal*, **1**, 69-93.

[25] Juarez, M. A. (2005). Objective Bayes Estimation and Hypothesis Testing: the reference-intrinsic approach. University of Warwick.

[26] Kass, R.E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical* **91**, 1343-1370.

[27] Landsman, Z. (2001). Second-order minimax estimation of the mean value for exponential dispersion models. *Journal of Statistical Planning and Inference* **98**, 57-71.

[28] Landsman, Z. and Makov, U.E. (1998). Exponential dispersion models and credibility. *Scandinavian Actuarial Journal,* **1**, 89-96.

[29] Landsman, Z. and Makov, U.E. (1999). Credibility evaluations for exponential dispersion families. *Insurance: Mathematics and Economics,* **24**, 33-39.

[30] Liseo, B. (1993). Elimination of nuisance parameters with reference priors. *Biometrika* **80**, 295-304.

[31] Laurent, S. (2012). Intrinsic Bayesian inference on a Poisson rate and on the ratio of two Poisson rates. *Journal of Statistical Planning and Inference* **142(9)**, 2656-2671.

[32] Makov, U.E., Smith, A.F.M. and Liu, Y.-H. (1996). Bayesian Methods in Actuarial Science. *The Statistician* **45(4)**, 503-515.

[33] Makov, U.E. (2001). Principal applications of Bayesian Methods in Actuarial Science: A Perspective. *North American Actuarial Journal* **5(4)**, 53-73.

[34] Ntzoufras, I., Katsis, A. and Karlis, D. (2005). Bayesian assessment of the distribution of insurance claim counts using reversible jump MCMC. *North American Actuarial Journal*, **9(3)**, 90-108.

[35] Payandeh Najafabadi, A. T. (2010). A new approach to the credibility formula. *Insurance: Mathematics and Economics* **46(2)**, 334–38.

[36] Robert, C.P. (1996). Intrinsic Losses. *Theory and Decision* **40**, 191-214.

[37] Smyth, G. K. and Jorgensen, B. (2002). Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling. *ASTIN Bulletin: The Journal of the IAA*, **32(1)**, 143-157.

[38] Sun, D. and Berger, J. O. (1998). Reference priors under partial information. *Biometrika* **85**, 55-71.

[39] Xacur, O. A. Q. and Garrido, J. (2015). Generalised linear models for aggregate claims: to Tweedie or not?. *European Actuarial Journal*, **5(1)**, 181-202.

[40] Yang, Rand Berger, J.O. (1996). A catalog of noninformative priors. *Tech. Rep.,Dept. Statistics*, Purdue University, USA.

6 Appendix

6.1 *Proof of theorem 1*

The IDLF of Gamma model associated with estimating θ by $\tilde{\theta}$ is derived from

$$\ell_{IDL|x}\left\{\tilde{\theta},\theta\right\} = n\cdot\min\left\{k\left(\tilde{\theta}|\theta\right), k\left(\theta|\tilde{\theta}\right)\right\},$$

where

$$k\left(\tilde{\theta}|\theta\right) = \int \log\frac{g(x|\lambda,\theta)}{g\left(x|\lambda,\tilde{\theta}\right)}\cdot g\left(x|\lambda,\theta\right)dx = \int_0^\infty \log\frac{\exp(-\theta x)\cdot\theta^\lambda}{\exp\left(-\tilde{\theta}x\right)\cdot\tilde{\theta}^\lambda}\cdot\frac{\exp(-\theta x)\cdot x^{\lambda-1}\cdot\theta^\lambda}{\Gamma(\lambda)}dx$$

$$= \int_0^\infty\left(\log\exp\left(\left(\tilde{\theta}-\theta\right)\cdot x\right)+\lambda\cdot\log\frac{\theta}{\tilde{\theta}}\right)\cdot\frac{\exp(-\theta x)\cdot x^{\lambda-1}\cdot\theta^\lambda}{\Gamma(\lambda)}dx =$$

$$= \int_0^\infty\left(\tilde{\theta}-\theta\right)\cdot x\cdot\frac{\exp(-\theta x)\cdot x^{\lambda-1}\cdot\theta^\lambda}{\Gamma(\lambda)}dx + \lambda\cdot\log\frac{\theta}{\tilde{\theta}}\int_0^\infty\frac{\exp(-\theta x)\cdot x^{\lambda-1}\cdot\theta^\lambda}{\Gamma(\lambda)}dx =$$

$$= \left(\tilde{\theta}-\theta\right)\cdot E\left(X\right)+\lambda\cdot\log\frac{\theta}{\tilde{\theta}} =$$

$$= \left(\tilde{\theta}-\theta\right)\cdot\frac{\lambda}{\theta}+\lambda\cdot\log\frac{\theta}{\tilde{\theta}} = \lambda\left(\frac{\tilde{\theta}}{\theta}-1+\log\frac{\theta}{\tilde{\theta}}\right).$$

and, in a similar way,

$$k\left(\theta|\tilde{\theta}\right) = \int\log\frac{g\left(x|\lambda,\tilde{\theta}\right)}{g(x|\lambda,\theta)}\cdot g\left(x|\lambda,\tilde{\theta}\right)dx = \lambda\left(\frac{\theta}{\tilde{\theta}}-1-\log\frac{\theta}{\tilde{\theta}}\right).$$

6.2 *Proof of remark 1*

Let us show that $R_{IOR}\left(\tilde{\mu}|\mathbf{x}\right) = R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$.The IDLF is invariant under one to one transformations, therefore we get

$$
\begin{aligned}
R_{IOR}\left(\tilde{\mu}|\mathbf{x}\right) &= \int_{o}^{\tilde{\mu}} n\lambda\left(\frac{\mu}{\tilde{\mu}}-1-\log\frac{\mu}{\tilde{\mu}}\right)\cdot Inv-g\left(\mu|n\lambda,\lambda t\right)d\mu \\
&\quad+\int_{\tilde{\mu}}^{\infty} n\lambda\left(\frac{\tilde{\mu}}{\mu}-1-\log\frac{\tilde{\mu}}{\mu}\right)\cdot Inv-g\left(\mu|n\lambda,\lambda t\right)d\mu \\
&= \int_{o}^{\tilde{\mu}} n\lambda\left(\frac{\mu}{\tilde{\mu}}-1-\log\frac{\mu}{\tilde{\mu}}\right)\cdot\frac{\exp\left(-\frac{\lambda}{\mu}t\right)\cdot\mu^{-n\lambda-1}\cdot(\lambda t)^{n\lambda}}{\Gamma\left(n\lambda\right)}d\mu \\
&\quad+\int_{\tilde{\mu}}^{\infty} n\lambda\left(\frac{\tilde{\mu}}{\mu}-1-\log\frac{\tilde{\mu}}{\mu}\right)\cdot\frac{\exp\left(-\frac{\lambda}{\mu}t\right)\cdot\mu^{-n\lambda-1}\cdot(\lambda t)^{n\lambda}}{\Gamma\left(n\lambda\right)}d\mu
\end{aligned}
$$

Let us define a new variable $\theta=\frac{\lambda}{\mu}$, then

$$
\int_{o}^{\tilde{\mu}} n\lambda\left(\frac{\mu}{\tilde{\mu}}-1-\log\frac{\mu}{\tilde{\mu}}\right)\cdot\frac{\exp\left(-\frac{\lambda}{\mu}t\right)\cdot\mu^{-n\lambda-1}\cdot(\lambda t)^{n\lambda}}{\Gamma\left(n\lambda\right)}d\mu=\int_{\tilde{\theta}}^{\infty} n\lambda\left(\frac{\tilde{\theta}}{\theta}-1-\log\frac{\tilde{\theta}}{\theta}\right)\cdot g\left(\theta|n\lambda,t\right)d\theta
$$

and

$$
\int_{\tilde{\mu}}^{\infty} n\lambda\left(\frac{\tilde{\mu}}{\mu}-1-\log\frac{\tilde{\mu}}{\mu}\right)\cdot\frac{\exp\left(-\frac{\lambda}{\mu}t\right)\cdot\mu^{-n\lambda-1}\cdot(\lambda t)^{n\lambda}}{\Gamma\left(n\lambda\right)}d\mu=\int_{o}^{\tilde{\theta}} n\lambda\left(\frac{\theta}{\tilde{\theta}}-1-\log\frac{\theta}{\tilde{\theta}}\right)\cdot g\left(\theta|n\lambda,t\right)d\theta
$$

thus, $R_{IOR}\left(\tilde{\mu}|\mathbf{x}\right)$ equals to $R_{IOR}\left(\tilde{\theta}|\mathbf{x}\right)$ from (11).
Therefore, $MIR\left(\mu_{IE}\right)=MIR\left(\theta_{IE}\right)$.