Revista Colombiana de Estadística - Theoretical Statistics
January 2025, volume 48, issue 1, pp. 25–38
http://doi.org/10.15446/rce.v48n1.114461

Closed-Form Predictive Density Estimation for Bivariate Gamma Distribution With Application in Hydrological Flood Data

Estimación de densidad predictiva en forma cerrada para la distribución Gamma bivariada con aplicación en datos hidrológicos de inundaciones

Abdolnasser Sadeghkhani[a]

Department of Mathematics and Statistics, North Carolina Agricultural and Technical State University, Greensboro, USA

Abstract

Finding closed-form solutions in Bayesian data analysis can be critical and time-saving, as it eliminates the need for computationally expensive techniques like MCMC methods. This paper explores Bayesian analysis with closed-form solutions of the bivariate gamma distribution. We present predictive density estimations under the Kullback-Leibler divergence, utilizing three well-known (non-) informative prior distributions, all analyzable in closed form. We compare these methods through simulation studies and a real-world example, applying them to hydrological flood data.

***Key words*:** Bayes estimation; Bivariate gamma distribution; Hydrological event analysis; Kullback-Leibler divergence; Predictive density estimation.

Resumen

Encontrar soluciones en forma cerrada en el análisis de datos bayesiano puede ser fundamental y ahorrar tiempo, ya que elimina la necesidad de técnicas computacionalmente costosas como los métodos MCMC. Este artículo explora el análisis bayesiano con soluciones en forma cerrada para la distribución gamma bivariada. Presentamos estimaciones de densidad predictiva bajo la divergencia de Kullback-Leibler, utilizando tres distribuciones a priori bien conocidas (informativas y no informativas), todas analizables en forma cerrada. Comparamos estos métodos mediante estudios de simulación y un ejemplo del mundo real, aplicándolos a datos hidrológicos de inundaciones.

***Palabras clave*:** Análisis de eventos hidrológicos; Divergencia de Kullback-Leibler; Distribución gamma bivariada; Estimación bayesiana; Estimación de densidad predictiva.

[a]Ph.D. E-mail: asadeghkhani@ncat.edu

1. Introduction

The univariate gamma distribution is frequently employed to model time-to-events and finds extensive applications across various fields. Research into the bivariate gamma distribution can be traced back to Kibble (1941) and Moran (1969). In recent years, bivariate gamma distributions have garnered attention, particularly in studies of hydrological events such as floods and storms. However, their application extends beyond hydrological events analysis; for instance, (Chatelain et al., 2007) explore their use in image registration and change detection, while (Furman & Landsman, 2005) examine their applicability in insurance.

While the gamma distribution is often suitable for modeling positively skewed distributions in phenomena like rainfall and floods (e.g., Husak et al., 2007), analyzing more complex data requires bivariate gamma models. Flood events, for example, are characterized by factors such as peak, volume, and duration, which are mutually correlated. Hence, multivariate, including bivariate, gamma distributions are essential for modeling such events (see, e.g., Yue, 2001 Nadarajah, 2009, among others).

Note that there are multiple definitions for the bivariate gamma distribution, and the versions used in the literature sometimes define it differently. See Nadarajah & Gupta (2006) and references there for different versions of bivariate gamma distributions.

Let $\mathbf{x} = (x_1, x_2)'$. Consider the following probability density function (pdf)

$$f(\mathbf{x};\, \alpha_1, \alpha_2, \beta_1, \beta_2) \;=\; \frac{x_1^{\alpha_1 - 1}(x_2 - x_1)^{\alpha_2 - 1}\, e^{\frac{-x_1}{\beta_1} + \frac{x_1 - x_2}{\beta_2}}}{\beta_1^{\alpha_1}\, \beta_2^{\alpha_2}\, \Gamma(\alpha_1)\, \Gamma(\alpha_2)}, \qquad 0 < x_1 < x_2 < \infty,\ \alpha_1 > 0,\ \alpha_2 > 0,\ \beta_1 > 0,\ \beta_2 > 0\,, \tag{1}$$

known as the bivariate gamma distribution $\mathrm{BG}(\alpha_1, \alpha_2, \beta_1, \beta_2)$. The joint distribution of $\mathbf{x} = (x_1, x_2)'$ can be viewed as a joint distribution of $(u_1, u_1 + u_2)$, where u_1 and u_2 are two independent gamma distributions $\mathrm{G}(\alpha_i, \beta_i)$ for $i = 1, 2$, with the mean of $\alpha_i \beta_i$ respectively. It is easy to see

$$E(\mathbf{x}) = (\alpha_1\beta_1,\, \alpha_1\beta_1 + \alpha_2\beta_2)'\,,$$

$$Cov(\mathbf{x}) = \begin{pmatrix} \alpha_1\beta_1^2 & \alpha_1\beta_1^2 \\ \alpha_1\beta_1^2 & \alpha_1\beta_1^2 + \alpha_2\beta_2^2 \end{pmatrix}\,,$$

and therefore the correlation is $\rho = \left(\frac{\alpha_1\beta_1^2}{\alpha_1\beta_1^2 + \alpha_2\beta_2^2}\right)^{1/2}$. This is a generalization of the model introduced by Mathal & Moschopoulos (1992) when $\beta_1 = \beta_2 = \beta$, known as the 3-parameter bivariate gamma, with marginal distributions of $\mathrm{G}(\alpha_1, \beta)$, and $\mathrm{G}(\alpha_1 + \alpha_2, \beta)$ respectively, with $\rho = (\frac{\alpha_1}{\alpha_1 + \alpha_2})^{1/2}$.

The log-likelihood function based on $\mathbf{x}_1, \ldots, \mathbf{x}_n$, from (1) is given by

$$\ell \;=\; (\alpha_1 - 1)\sum_{i=1}^{n} \log(x_{1i}) + (\alpha_2 - 1)\sum_{i=1}^{n} \log(x_{2i} - x_{1i}) - \frac{t_2}{\beta_2} - \frac{t_1}{\beta_1} - \alpha_2 n \log(\beta_2) - \alpha_1 n \log(\beta_1) - n\log(\Gamma(\alpha_1)) - n\log(\Gamma(\alpha_2))\,, \tag{2}$$

where

$$t_1 = t_1(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{i=1}^{n} x_{1i}, \;\; t_2 = t_2(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{i=1}^{n} (x_{2i} - x_{1i})\,. \tag{3}$$

Solving $\partial\ell/\partial\alpha_i$ and $\partial\ell/\partial\beta_i$ for $i = 1, 2$ yields the maximum likelihood estimator (MLE) of the parameters α_i and β_i numerically as follows

$$\begin{aligned}
\partial\ell/\partial\alpha_1 &= \sum_{i=1}^{n} \log(x_{1i}) - n\left(\psi(\alpha_1) + \log(\beta_1)\right)\,, \\
\partial\ell/\partial\alpha_2 &= \sum_{i=1}^{n} \log(x_{2i} - x_{1i}) - n\left(\psi(\alpha_2) + \log(\beta_2)\right)\,, \\
\partial\ell/\partial\beta_1 &= \frac{1}{\beta_1^2} \sum_{i=1}^{n} x_{1i} - n\frac{\alpha_1}{\beta_1}\,, \\
\partial\ell/\partial\beta_2 &= \frac{1}{\beta_1^2} \sum_{i=1}^{n} (x_{2i} - x_{1i}) - n\frac{\alpha_2}{\beta_2}\,,
\end{aligned}$$

where $\psi(\cdot)$, known as the diagamma function, is the derivative of the gamma function, given by $\frac{\Gamma'(\cdot)}{\Gamma(\cdot)}$.

Note that the once $\beta_1 = \beta_2 = \beta$, the corresponding the log-likelihood function, ℓ_0, is given as

$$\begin{aligned}
\ell_0 = {} & (\alpha_1 - 1) \sum_{i=1}^{n} \log(x_{1i}) - \frac{\sum_{i=1}^{n} x_{2i}}{\beta} + (\alpha_2 - 1) \sum_{i=1}^{n} \log(x_{2i} - x_{1i}) \\
& - n\left((\alpha_1 + \alpha_2)\log(\beta) + \log(\Gamma(\alpha_1)) + \log(\Gamma(\alpha_2))\right)\,.
\end{aligned}$$

The MLE of parameters α_i for $i = 1, 2$ and β can be obtained numerically from

$$\partial\ell_0/\partial\alpha_1 = \sum_{i=1}^{n} \log(x_{1i}) - n\left(\psi(\alpha_1) + \log(\beta)\right)\,, \tag{4}$$

$$\partial\ell_0/\partial\alpha_2 = \sum_{i=1}^{n} \log(x_{2i} - x_{1i}) - n\left(\psi(\alpha_2) + \log(\beta)\right)\,, \tag{5}$$

$$\partial\ell_0/\partial\beta = \frac{1}{\beta^2} \sum_{i=1}^{n} x_{2i} - n\frac{\alpha_1 + \alpha_2}{\beta_1}\,, \tag{6}$$

Zhao et al. (2022) showed that the MLEs $\hat{\alpha}_{i\,ml}$, for $i = 1, 2$ and $\hat{\beta}_{ml}$ obtained from Equations (4), (5), and (6) are strongly consistent for α_i and $\beta > 0$, except for $\alpha_2 = 2$, and converge in distribution to a normal distribution as the sample size increases to infinity.

Another essential aspect of statistical analysis is predicting future observations. Instead of making point-by-point or interval predictions, it is common to estimate the density of future random variables. In classical inference, predictions are often obtained using *plug-in density estimation*, where parameter estimates (e.g., MLE) are substituted directly into the density function. In contrast, *posterior predictive density estimation* is a key feature of Bayesian inference, leveraging the posterior distribution of parameters to account for parameter uncertainty. For more information on the application of plug-in density estimation, see Ghosh et al. (2008) and Marchand & Sadeghkhani (2018) for examples. Sadeghkhani & Ahmed (2021) elaborated on these techniques in the context of a gamma distribution with applications in sports data. Undoubtedly, there are numerous applications for estimating future bivariate gamma distributions. For example, we may be interested in estimating the joint distribution of flood peak, duration, or volume. Understanding and estimating the density of future events enriches our comprehension of flood event behavior.

To our knowledge, there have been limited studies on Bayesian predictive density estimation in bivariate gamma distributions, and none of them have pursued closed-form solutions for posterior and predictive distributions. In this paper, we present a closed-form Bayesian statistical inference for the unknown parameters of the bivariate gamma model and closed-form posterior predictive inference, emphasizing cost-effective and efficient calculations, with applications in the analysis of flood data.

The rest of this paper is organized as follows. In Section 2 we introduce three well-known priors raging from non-informative to informative, and we find the closed-form posterior distributions along with the corresponding Bayes estimators. All closed-form posterior predictive density estimations related to the prior distributions introduced in Section 2 are presented in Section 3. Section 4 compares the proposed Bayes and posterior predictive density estimators with a simulations along with a famous hydrological events dataset. Finally, we make some concluding remarks in Section 5.

2. Three Well-Known Priors

In this section, we explore three widely recognized prior distributions, both informative and non-informative, for Model (1). These prior distributions facilitate the derivation of closed-form posterior distributions, aiding in the analysis of the model.

2.1. Non-Informative Prior

Assuming α_i is known, the non-informative prior $\pi(\beta_i) = 1/\beta_i$, for $i = 1, 2$, and $\pi_1(\beta_1, \beta_2) = \pi(\beta_1)\pi(\beta_2)$, the following lemma provides the joint posterior distribution in Model (1).

Lemma 1. *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be an IID random sample of size* n *from* $\mathrm{BG}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ *in (1), where* $\mathbf{x} = (x_1, x_2)'$*, and* $\pi_1(\beta_1, \beta_2) = \frac{1}{\beta_1\beta_2}$*. Then, the joint distribution of*

$(\beta_1, \beta_2) \mid \mathbf{x}_1, \ldots, \mathbf{x}_n$, provided that $\alpha_i > 1$ for $i = 1, 2$, is given by

$$\pi_1((\beta_1, \beta_2) \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \;=\; \frac{\beta_1^{-(\alpha_1+1)}\beta_2^{-(\alpha_2+1)} t_1^{\alpha_1}(t_2 - t_1)^{\alpha_2} e^{\frac{t_1 - t_2}{\beta_2} - \frac{t_1}{\beta_1}}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\,, \tag{7}$$

where t_1, and t_2 is given in (3).

Proof. From the Bayes' formula, we have

$$\pi_1(\beta_1, \beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \;=\; \frac{\prod_{i=1}^n f(\mathbf{x}_i;\, \alpha_1, \alpha_2, \beta_1, \beta_2)\, \pi_1(\beta_1, \beta_2)}{f(\mathbf{x}_1, \ldots, \mathbf{x}_n)}\,, \tag{8}$$

where the normalization constant is given by

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_n) \;=\; \int_0^\infty\!\!\int_0^\infty \frac{\beta_1^{-(\alpha_1+1)}\beta_2^{-(\alpha_2+1)} k_1^{\alpha_1 - 1} k_2^{\alpha_2 - 1} e^{\frac{t_1 - t_2}{\beta_2} - \frac{t_1}{\beta_1}}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\, d\beta_2\, d\beta_1\,, \tag{9}$$

$$= \; k_1^{\alpha_1 - 1} k_2^{\alpha_2 - 1} t_1^{-\alpha_1}(t_2 - t_1)^{-\alpha_2}\,, \tag{10}$$

where

$$k_1 = k_1(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \prod_{i=1}^n x_{1i}\,, \quad k_2 = k_2(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \prod_{i=1}^n (x_{2i} - x_{1i})\,. \tag{11}$$

Replacing Equation (9) into (8) and simplifying completes the proof. □

Corollary 1. *Under the assumptions of Lemma 1, the marginal posterior distributions along with their expectations are given as*

$$\pi_1(\beta_1 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \;= \frac{t_1^{\alpha_1} e^{-\frac{t_1}{\beta_1}}}{\beta_1^{\alpha_1+1}\,\Gamma(\alpha_1)}\,,$$

$$\pi_1(\beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \;= \frac{(t_2 - t_1)^{\alpha_2} e^{-\frac{t_2 - t_1}{\beta_2}}}{\beta_2^{\alpha_2+1}\,\Gamma(\alpha_2)}\,,$$

$$E(\beta_1 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \;= \frac{t_1}{\alpha_1 - 1}\,,\;\; \alpha_1 > 1\,, \tag{12}$$

$$E(\beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \;= \frac{t_2 - t_1}{\alpha_2 - 1}\,,\;\; \alpha_2 > 1\,. \tag{13}$$

Note that the the expectations obtained in (12) and (13) are the Bayes estimator $\hat{\beta}_{iB1}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ of unknown parameter β_i under the squared error loss (SEL) function

$$L(\hat{\beta}_{iB1}(\mathbf{x}_1, \ldots, \mathbf{x}_n), \beta_i) \;=\; (\hat{\beta}_{iB1}(\mathbf{x}_1, \ldots, \mathbf{x}_n) - \beta_i)^2\,,$$

and minimizes the risk function $E\left(L(\hat{\beta}_{iB1}(\mathbf{x}_1, \ldots, \mathbf{x}_n), \beta_i) \mid \mathbf{x}_1, \ldots, \mathbf{x}_n\right)$. For more information. See, eg., Lehmann & Casella (2006).

2.2. Informative Inverse Gamma Prior

In this setup it is assumed that α_1 and α_2 are known, and two independent β_1 and β_2 follow from the inverse gamma distribution $\mathrm{IG}(a_i, b_i)$ with $\pi(\beta_i) = \frac{e^{-b_i/\beta_i}(b_i/\beta_i)^{a_i}}{\beta_i\,\Gamma(a_i)}$, for $i = 1, 2$, where the hyperparameters $a_i > 0$ (shape) and $b_i > 0$ (scale) are known.

Lemma 2 and Corollary 2 show the joint posterior distribution as well as the marginal distributions with the associated conditional expectations, i.e., the Bayes estimators of the unknown parameter β_i under SEL.

Lemma 2. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be an IID random sample of size n, from $\mathrm{BG}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ in (1), where $\mathbf{x} = (x_1, x_2)'$, and $\pi_2(\beta_1, \beta_2) = \pi(\beta_1)\pi(\beta_2)$, with $\beta_i \sim \mathrm{IG}(a_i, b_i)$. Then the joint distribution of $(\beta_1, \beta_2) \mid \mathbf{x}_1, \ldots, \mathbf{x}_n$ for $i = 1, 2$, and $\alpha_i > 0$ along with other hyper-parameters are known, are given by*

$$\pi_2((\beta_1, \beta_2) \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{(b_1 + t_1)^{a_1+\alpha_1}(b_2 - t_1 + t_2)^{a_2+\alpha_2} e^{-\frac{b_1+t_1}{\beta_1} - \frac{b_2-t_1+t_2}{\beta_2}}}{\beta_1^{a_1+\alpha_1+1}\beta_2^{a_2+\alpha_2+1}\,\Gamma(a_1+\alpha_1)\Gamma(a_2+\alpha_2)}, \tag{14}$$

where t_1, and t_2 are defined in (3).

***Proof*.** The proof is analogous to Lemma 1, and hence it is omitted. □

Corollary 2. *Under the assumptions of Lemma 2, the marginal posterior distributions along with their expectations are given as*

$$\begin{aligned}
\pi_2(\beta_1 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) &= \frac{(\frac{b_1+t_1}{\beta_1})^{\alpha_1+a_1} e^{-\frac{b_1+t_1}{\beta_1}}}{\beta_1\,\Gamma(\alpha_1+a_1)}, \\
\pi_2(\beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) &= \frac{(\frac{b_2+t_2-t_1}{\beta_2})^{\alpha_2+a_2} e^{-\frac{b_2+t_2-t_1}{\beta_2}}}{\beta_2\,\Gamma(\alpha_2+a_2)}, \\
E(\beta_1 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) &= \frac{b_1+t_1}{\alpha_1+a_1-1}, \ \alpha_1+a_1 > 1, \qquad (15) \\
E(\beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) &= \frac{b_2+t_2-t_1}{\alpha_2+a_2-1}, \ \alpha_2+a_2 > 1. \qquad (16)
\end{aligned}$$

The Equations in (12) and (13) are also the Bayes estimators $\hat{\beta}_{iB2}(\mathbf{x})$ of unknown parameters β_i for $i = 1, 2$, under the SEL function.

2.3. Informative Gamma Prior

Suppose the parameter α_i is known, and β_1 and β_2 are independently from $\mathrm{Gam}(a_i, b_i)$, with $\pi(\beta_i) = \frac{\beta_i^{a_i-1} e^{-\frac{\beta_i}{b_i}}}{b_i^{a_i}\,\Gamma(a_i)}$, for $i = 1, 2$. The following lemma and corollary provide the joint posterior distribution as well as the marginal distributions of β_1 and β_2 given $\mathbf{x}$.

Lemma 3. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be an IID random sample of size n, from $\mathrm{BG}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ in (1), where $\mathbf{x} = (x_1, x_2)'$, and $\pi_3(\beta_1, \beta_2) = \pi(\beta_1)\pi(\beta_2)$, with $\beta_i \sim \mathrm{Gam}(a_i, b_i)$, then the joint distribution of $(\beta_1, \beta_2) \mid \mathbf{x}_1, \ldots, \mathbf{x}_n$, for $i = 1, 2$, assuming that α_i and all other hyper-parameters are known, are as follow*

$$\pi_3((\beta_1, \beta_2) \mid \mathbf{x}_1, \ldots, x_n) = \frac{\beta_1^{a_1-\alpha_1-1}\beta_2^{a_2-\alpha_2-1}(b_1t_1)^{\frac{\alpha_1-a_1}{2}}(b_2(t_2-t_1))^{\frac{\alpha_2-a_2}{2}}e^{-\frac{\beta_1}{b_1}-\frac{\beta_2}{b_2}-\frac{t_1}{\beta_1}+\frac{t_1-t_2}{\beta_2}}}{4K_{\alpha_1-a_1}\left(2\sqrt{\frac{t_1}{b_1}}\right)K_{\alpha_2-a_2}\left(2\sqrt{\frac{t_2-t_1}{b_2}}\right)}, \quad (17)$$

where $k_c(z)$ is known as the modified Bessel function of the second kind with order c, and is given as

$$\frac{1}{2^c}\frac{\sqrt{\pi}z^c}{\Gamma(c+\frac{1}{2})}\int_r^\infty (r^2-1)^{c-\frac{1}{2}}\, e^{-zr}\, dr.$$

***Proof*.** Applying Bayes' theorem allows us to calculate the normalization constant under the assumptions of Lemma 3, which is given by

$$\int_0^\infty\int_0^\infty \frac{t_1^{\alpha_1-1}\beta_1^{a_1-\alpha_1-1}\beta_2^{a_2-\alpha_2-1}(t_2-t_1)^{\alpha_2-1}e^{-\frac{\beta_1}{b_1}-\frac{\beta_2}{b_2}-\frac{t_1}{\beta_1}+\frac{t_1-t_2}{\beta_2}}}{b_1^{a_1}b_2^{a_2}\,\Gamma(a_1)\Gamma(\alpha_1)\Gamma(a_2)\Gamma(\alpha_2)}\, d\beta_1\, d\beta_2. \quad (18)$$

After doing some algebra, one can finally obtain Equation (18) as

$$\frac{4b_1^{-a_1}b_2^{-a_2}\mathrm{k1}^{\alpha_1-1}\mathrm{k2}^{\alpha_2-1}(b_1t_1)^{\frac{a_1-\alpha_1}{2}}K_{\alpha_1-a_1}\left(2\sqrt{\frac{t_1}{b_1}}\right)(b_2(t_2-t_1))^{\frac{a_2-\alpha_2}{2}}K_{\alpha_2-a_2}\left(2\sqrt{\frac{t_2-t_1}{b_2}}\right)}{\Gamma(a_1)\Gamma(\alpha_1)\Gamma(a_2)\Gamma(\alpha_2)},$$

where k_1, and k_2 are given in (11). This completes the proof. □

Corollary 3. *Under the assumptions of Lemma 3, the marginal posterior distributions and the corresponding expectations are given by*

$$\pi_3(\beta_1 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{\beta_1^{a_1-\alpha_1-1}e^{-\frac{\beta_1}{b_1}-\frac{t_1}{\beta_1}}(b_1t_1)^{\frac{\alpha_1-a_1}{2}}}{2K_{\alpha_1-a_1}\left(2\sqrt{\frac{t_1}{b_1}}\right)}, \quad (19)$$

$$\pi_3(\beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{\beta_2^{a_2-\alpha_2-1}e^{-\frac{t_2-t_1}{\beta_2}-\frac{\beta_2}{b_2}}(b_2(t_2-t_1))^{\frac{\alpha_2-a_2}{2}}}{2K_{\alpha_2-a_2}\left(2\sqrt{\frac{t_2-t_1}{b_2}}\right)}, \quad (20)$$

$$E(\beta_1 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{\sqrt{b_1t_1}K_{a_1-\alpha_1+1}\left(2\sqrt{\frac{t_1}{b_1}}\right)}{K_{\alpha_1-a_1}\left(2\sqrt{\frac{x_1}{b_1}}\right)}, \quad (21)$$

$$E(\beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{\sqrt{b_2(t_2-t_1)}K_{a_2-\alpha_2+1}\left(2\sqrt{\frac{t_2-t_1}{b_2}}\right)}{K_{\alpha_2-a_2}\left(2\sqrt{\frac{t_2-t_1}{b_2}}\right)}. \quad (22)$$

The Equations in (12) and (13) are also the Bayes estimator $\hat{\beta}_{iB3}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ of unknown parameter β_i under the SEL function.

Figure 1, illustrates the 3D plot as well the corresponding contour plot of the joint posterior distribution (17) in Lemma 3 with $(\alpha_1, \alpha_2, a_1, a_2, b_1, b_2, t_1, t_2) = (3, 4, 5, 5, 4, 1, 3, 7)$.

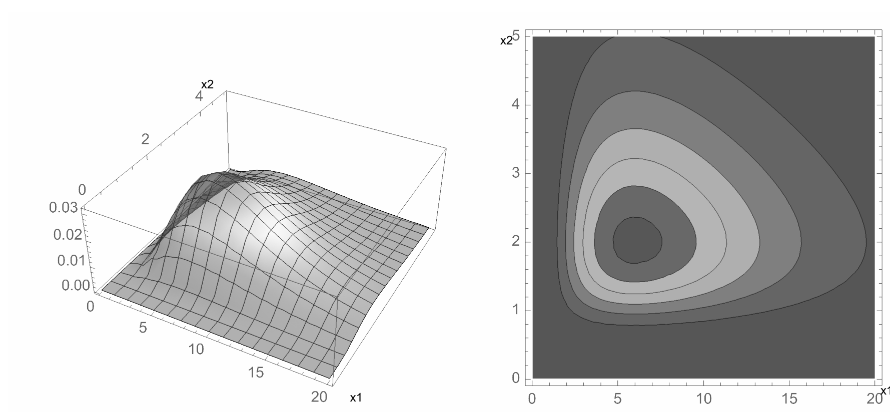

FIGURE 1: The 3D plot (left), and the contour plot (right) of the joint posterior distribution $\pi_3(\beta_1, \beta_2 \mid \mathbf{x})$ as in Lemma 3, for $(\alpha_1, \alpha_2, a_1, a_2, b_1, b_2, t_1, t_2) = (3, 4, 5, 5, 4, 1, 3, 7)$.

3. Closed-Form Predictive Density Estimation

Suppose we are interested in estimating the distribution of future random variable $\mathbf{y} = (y_1, y_2)'$, based on observable $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with $\mathbf{x}_i = (x_{1i}, x_{2i})'$ in Model (1). The posterior predictive distribution of $\mathbf{y}$ given $\mathbf{x}_1, \ldots, \mathbf{x}_n$, is given as (Corcuera & Giummolè, 1999)

$$\hat{g}(\mathbf{y}; \mathbf{x}_1, \ldots, \mathbf{x}_n) = \int_0^\infty \int_0^\infty g(\mathbf{y}; \alpha_1, \alpha_2, \beta_1, \beta_2)\, \pi(\beta_1, \beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n)\, d\beta_1\, d\beta_2 \,, \tag{23}$$

where $\mathbf{y} \sim g$, and α_i $(i = 1, 2)$ is known. Here, we consider three scenarios for the posterior distribution $\pi(\beta_1, \beta_2 \mid \mathbf{x}_1, \ldots, \mathbf{x}_n)$ in Equation (23) corresponding to the (non-)informative priors as introduced in introduced in Section 2.

To assess the proximity of the estimator $\hat{g}(\mathbf{y}; \mathbf{x}_1, \ldots, \mathbf{x}_n)$ in estimating the actual distribution, we apply the Kullback Leibler (KL) divergence (loss function), given as

$$\mathrm{KL}(\hat{g}, g) \;=\; \int_0^\infty \int_0^{y_2} g(\mathbf{y}; \alpha_1, \alpha_2, \beta_1, \beta_2) \log \frac{g(\mathbf{y}; \alpha_1, \alpha_2, \beta_1, \beta_2)}{\hat{g}(\mathbf{y}; \mathbf{x}_1, \ldots, \mathbf{x}_n)}\, dy_1\, dy_2 \,. \tag{24}$$

Suppose $\pi_j(\beta_1, \beta_2)$ for $j = 1, 2, 3$ is the prior distribution given in Subsection 2.1, 2.2, and 2.3, respectively. Theorem 1 finds a closed–form posterior predictive density estimator for future $\mathbf{y}$ in Model (1), under each prior distribution.

Theorem 1. *The posterior predictive density estimator $\hat{g}_j(\mathbf{y};\mathbf{x}_1,\ldots,\mathbf{x}_n)$ $(j = 1,2,3)$ of future $\mathbf{y} = (y_1, y_2)' \sim \text{BG}(\alpha_1, \alpha_2, \beta_1, \beta_2)$, under the KL divergence (loss function), based on observing $\mathbf{x}_1,\ldots,\mathbf{x}_n$ IID from $\text{BG}(\alpha_1, \alpha_2, \beta_1, \beta_2)$, and postulated prior $\pi_j(\beta_1, \beta_2)$, provided that α_i, for $i = 1, 2$ and other hyperparameters are known, are as follows*

1. *If $\pi_1(\beta_1, \beta_2) = \pi(\beta_1)\,\pi(\beta_2)$, with $\pi(\beta_i) = 1/\beta_i$, for $i = 1, 2$, then*

$$\hat{g}_1(y;\mathbf{x_1},\ldots,\mathbf{x_n}) = \frac{\Gamma(2\alpha_1)\Gamma(2\alpha_2)l_1^{\alpha_1-1}y^{\alpha_1-1}(l_2-l_1)^{\alpha_2}}{\Gamma(\alpha_1)^2\Gamma(\alpha_2)^2}(l_1+y_1)^{-2\alpha_1}(y_2-y_1)^{\alpha_2-1} \times(-l_1+l_2-y_1+y_2)^{-2\alpha_2}. \tag{25}$$

2. *if $\pi_2(\beta_1, \beta_2) = \pi(\beta_1)\pi(\beta_2)$, with $\beta_i \sim \text{IG}(a_i, b_i)$, then*

$$\hat{g}_2(y;\mathbf{x_1},\ldots,\mathbf{x_n}) = \frac{y_1^{\alpha_1-1}\beta_1^{-a_1-2\alpha_1-1}\beta_2^{-a_2-2\alpha_2-1}(y_2-y_1)^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(a_1+\alpha_1)\Gamma(a_2+\alpha_2)} \times\left(\frac{1}{b_1+t_1}\right)^{-a_1-\alpha_1}\left(\frac{1}{b_2-t_1+t_2}\right)^{-a_2-\alpha_2} e^{-\frac{b_1+t_1+y_1}{\beta_1}-\frac{b_2-t_1+t_2-y_1+y_2}{\beta_2}}. \tag{26}$$

3. *if $\pi_3(\beta_1, \beta_2) = \pi(\beta_1)\pi(\beta_2)$, with $\beta_i \sim \text{Gam}(a_i, b_i)$, then*

$$\hat{g}_3(y;\mathbf{x_1},\ldots,\mathbf{x_n}) = \frac{y_1^{\alpha_1-1}\beta_1^{a_1-2\alpha_1-1}\beta_2^{a_2-2\alpha_2-1}(y_2-y_1)^{\alpha_2-1}(b_1t_1)^{\frac{\alpha_1-a_1}{2}}(b_2(t_2-t_1))^{\frac{\alpha_2-a_2}{2}}}{4\Gamma(\alpha_1)\Gamma(\alpha_2)K_{\alpha_1-a_1}\left(2\sqrt{\frac{t_1}{b_1}}\right)K_{\alpha_2-a_2}\left(2\sqrt{\frac{t_2-t_1}{b_2}}\right)} \times e^{-\frac{\beta_1}{b_1}-\frac{\beta_2}{b_2}+\frac{t_1-t_2+y_1-y_2}{\beta_2}-\frac{t_1+y_1}{\beta_1}}. \tag{27}$$

Note that $y_2 > y_1$, and t_1, and t_2 are given in (3).

***Proof*.** By substituting the posterior distributions presented in the Lemmas 1, 2, and 3 in (23), and after doing some calculations, the posterior predictive density estimators in (25), (26), and (27) can be obtained, respectively. □

4. Simulation and Real Data Example

In this section, we investigate the correlation coefficient, ML and Bayesian estimates of the parameters of Model (1) by simulating the bivariate gamma distribution. Additionally, we explore predictive density estimates for future bivariate random variables. To illustrate the practical application of our proposed estimators, we analyze a real dataset concerning hydrological events.

4.1. Simulation Study

In this simulation, we generated data from the bivariate gamma distribution $\text{BG}(\alpha_1 = 3, \alpha_2 = 5, \beta_1 = 2, \beta_2 = 1)$ with a correlation of 0.86. The simulation was repeated $N = 1000$ times for different sample sizes $n = 25, 50, 100, 250$. In the Maximum Likelihood (ML) method, all parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ were treated as unknown and estimated. In contrast, the Bayesian method assumed α_1 and α_2 were known, and only β_1 and β_2 were estimated using non-informative priors. The Mean Squared Errors (MSEs) are presented in parentheses next to each corresponding estimator.

As shown in Table 1, the Bayesian estimators demonstrate slightly better performance compared to the ML estimators for smaller sample sizes ($n = 25, 50$), as evidenced by the lower MSE values. This improvement is consistent with the Bayesian approach's ability to incorporate prior information and reduce variability in parameter estimation. As the sample size increases ($n = 100, 250$), the Bayesian and ML estimators converge, with both showing increasingly smaller MSE values. This convergence indicates that, with larger sample sizes, both methods provide similar accuracy in estimating the parameters β_1 and β_2, reflecting the consistency and efficiency of these estimators as n becomes large.

Table 1: Simulation results for different sample sizes: ML and Bayes estimators.

Sample Size	$\hat{\alpha}_1$ (ML)	$\hat{\alpha}_2$ (ML)	$\hat{\beta}_1$ (ML)	$\hat{\beta}_1$ (Bayes)	$\hat{\beta}_2$ (ML)	$\hat{\beta}_2$ (Bayes)
25	3.390 (1.172)	5.678 (3.860)	1.915 (0.332)	1.925 (0.320)	0.961 (0.084)	0.972 (0.078)
50	3.185 (0.413)	5.287 (1.241)	1.959 (0.162)	1.964 (0.157)	0.984 (0.042)	0.989 (0.040)
100	3.048 (0.160)	5.139 (0.514)	2.002 (0.080)	2.005 (0.078)	0.989 (0.020)	0.991 (0.019)
250	3.026 (0.069)	5.081 (0.202)	1.996 (0.035)	1.997 (0.034)	0.991 (0.008)	0.992 (0.008)

In addition, Table 2 provides the plug-in and posterior predictive density estimators (with respect to noninformative prior) for the simulated data with $n = 250$. The plug-in density estimator is obtained by plugging in the MLE (or other Bayesian estimators) of the parameters into Model (1), resulting in $\hat{f}(\mathbf{y}; \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2)$, where $\hat{\alpha}_i$ and $\hat{\beta}_i$ can be referenced from Table 1.

Table 2: Plug-in and posterior predictive density estimators based on simulation.

Type	Predictive Distribution
Plug-in $\hat{f}$	$0.001\, y_1^{2.164}\, e^{-1.162\,(y_2-y_1)-0.507y_1}\, (y_2 - y_1)^{4.848}$
Posterior predictive density $\hat{g}_1$	$\frac{5.230995309040582\times 10^{441} y_1^2 (y_1-y_2)^4}{(y_1+143.03)^{78}(y_2-y_1+133.24)^{130}}$

4.2. Real Example

Table 5 reports the flood data available from the HYDAT CD (Canada, 1998). Q represents the flood peaks (in m^3/s), and R is the flood volume ratio (in day m^3/s) for the Madawaska basin located in Québec, Canada, covering the period from 1919 to 1995. The corresponding scatter plot is presented in Figure 2.

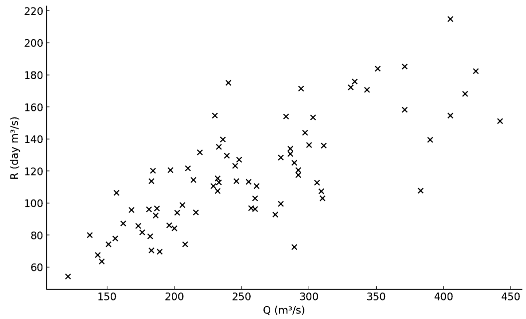

FIGURE 2: Scatter plot of flood data.

The mean vector and the covariance matrix for the data are $(118.227, 254.74)'$ and

$$\begin{pmatrix} 1162.68 & 1956.66 \\ 1956.66 & 5756.69 \end{pmatrix},$$

respectively. The correlation matrix

$$\begin{pmatrix} 1 & 0.756 \\ 0.756 & 1 \end{pmatrix},$$

confirms a fairly strong positive correlation between R and Q.

Zhao et al. (2022) proposed a 3-parameter bivariate gamma model (i.e., Model (1) with $\beta_1 = \beta_2 = \beta$) for the flood data. However, based on the AIC values, it is evident that our proposed four-parameter model $\mathrm{BG}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ is a better fit for this data. The AIC for the three-parameter model was 1596.868, while the AIC for our four-parameter model was 1587.561.

This significant difference in AIC indicates that the additional flexibility provided by allowing β_1 and β_2 to vary independently in the four-parameter model is justified by the data. The four-parameter model better captures the relationship between R and Q, leading to an improved fit despite the increased complexity. The lower AIC shows that this model balances model complexity and goodness-of-fit more effectively, confirming that the extra parameter provides a meaningful improvement.

The ML and Bayes estimators (based on noninformative prior) for the parameters are presented in Table 3.

TABLE 3: ML and Bayes estimators for the flood data.

Model	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
Four-Parameter (ML)	12.291	6.481	9.619	21.064
Four-Parameter (Bayes)	*	*	9.506	20.910
Three-Parameter (ML)	7.898	8.725	15.325	-
Three-Parameter (Bayes)	*	*	15.201	-

Additionally, Table 4 displays the plug-in and posterior predictive density estimators for $\mathbf{y} = (y_1, y_2)'$, where y_1 denotes the future random variable representing the ratio of flood volume to duration (R), and y_2 represents the future random variable flood peak (Q). Different priors, as per Theorem 1, are considered, and the hyperparameters for Bayes 2 and 3 are set accordingly.

Table 4: Plug-in and posterior predictive density estimators for the flood data.

Type	Predictive Distribution
Plug-in $\hat{f}$	$2.1496 \times 10^{-28} y_1^{6.897} e^{0.0652(y_1-y_2)-0.0652y_1} (y_2 - y_1)^{7.725}$
Posterior predictive density $\hat{g}_1$	$\frac{1.1413\times 10^{589} y_1^{6.8976} (y_2-y_1)^{7.725}}{(y_2+2331.22)^{182.849}}$

5. Conclusions

In summary, we introduced a closed-form Bayesian inference approach for a bivariate gamma model, specifically designed to analyze the joint distribution of two positively correlated random variables with gamma marginals. We explored various predictive density estimation techniques, including plug-in and posterior predictive methods, and applied these techniques to model hydrological flood data using flood peaks (Q) and the flood volume ratio (R). These methods allowed us to estimate the density of the future joint distribution of these variables effectively.

The main contributions of this work include providing a Bayesian framework for the bivariate gamma model that incorporates closed-form posterior distributions, which can facilitate computational efficiency and analytical tractability. Additionally, the study demonstrated the advantages of the Bayesian approach in predictive density estimation, where posterior predictive distributions offer a more comprehensive way to account for parameter uncertainty compared to traditional methods.

Moreover, we proposed prior distributions for two parameters, β_1 and β_2, while assuming that α_1 and α_2 were known. The choice to focus on β_1 and β_2 was driven by the challenge of estimating all four parameters in the bivariate gamma model simultaneously. In practice, estimating all four parameters would require substantial data or additional prior information, especially in cases where the marginals are highly correlated. Moreover, while the MLE approach can estimate all four parameters, it does not provide a way to incorporate uncertainty into predictions as the Bayesian method does through posterior predictive distributions. This trade-off between parameter estimation and uncertainty quantification is a critical aspect of choosing the appropriate method, and further research could explore alternative ways to handle these challenges within the Bayesian framework.

Looking ahead, future work could extend this approach by exploring multivariate gamma distributions that generalize the bivariate case, allowing for the modeling of multiple correlated gamma-distributed variables simultaneously. This extension would be valuable in applications where more than two correlated components are present, such as multivariate flood analysis across multiple locations. Additionally, the development of efficient computational methods for parameter estimation in such multivariate settings, possibly through hierarchical Bayesian models, would be a promising avenue for further research.

TABLE 5: Flood data of the Madawaska basin in Québec, Canada from 1919 to 1995: R represents the flood volume ratio (in day m^3/s), and Q denotes the flood peaks (in m^3/s).

R	Q	R	Q	R	Q
120.57	292	74.336	208	72.528	289
63.635	146	70.436	183	128.461	279
103.028	260	99.617	279	80.067	137
135.93	311	107.35	309	110.633	261
87.267	162	93.931	202	112.743	306
214.885	405	113.81	183	131.77	219
121.808	210	84.193	200	125.14	289
129.528	239	171.5	294	185.191	371
123.339	245	69.817	189	110.731	229
175.188	240	172.177	331	98.742	206
106.426	157	120.163	184	92.917	275
130.815	286	154.686	230	135.197	233
183.887	351	77.882	156	95.65	168
170.667	343	114.536	214	153.448	303
136.319	300	67.617	143	107.592	232
79.216	182	54.197	121	92.253	186
85.825	173	117.459	292	168.239	416
113.684	246	127.073	248	143.877	297
158.333	371	151.2	442	96.226	260
139.744	236	175.954	334	102.956	310
107.844	383	74.368	151	120.562	197
154.064	283	139.619	390	154.67	405
81.772	176	96.179	181	113.	233
96.711	187	94.119	216	86.256	196
182.397	424	113.254	255	96.895	257
115.383	232	134.015	286		

[Received: February 2024 — Accepted: October 2024]

References

Canada, E. (1998), 'HYDAT CD-ROM Version 98-1.05.8: Surface water and sediment data', Ottawa: Environment Canada, Water Survey Division. Accessed: June 2024.

Chatelain, F., Tourneret, J. Y., Inglada, J. & Ferrari, A. (2007), 'Bivariate gamma distributions for image registration and change detection', *IEEE Transactions on Image Processing* **16**(7), 1796–1806.

Corcuera, J. M. & Giummolè, F. (1999), 'A generalized bayes rule for prediction', *Scandinavian Journal of Statistics* **26**(2), 265–279.

Furman, E. & Landsman, Z. (2005), 'Risk capital decomposition for a multivariate dependent gamma portfolio', *Insurance: Mathematics and Economics* **37**(3), 635–649.

Ghosh, M., Mergel, V. & Datta, G. S. (2008), 'Estimation, prediction and the Stein phenomenon under divergence loss', *Journal of Multivariate Analysis* **99**, 1941–1961.

Husak, G. J., Michaelsen, J. & Funk, C. (2007), 'Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications', *International Journal of Climatology: A Journal of the Royal Meteorological Society* **27**(7), 935–944.

Kibble, W. F. (1941), 'A two-variate gamma type distribution', *Sankhya: The Indian Journal of Statistics* pp. 137–150.

Lehmann, E. L. & Casella, G. (2006), *Theory of Point Estimation*, Springer Science & Business Media.

Marchand, E. & Sadeghkhani, A. (2018), 'On Predictive Density Estimation with Additional Information'. Preprint. https://arxiv.org/abs/1805.08961

Mathal, A. M. & Moschopoulos, P. G. (1992), 'A form of multivariate gamma distribution', *Annals of the Institute of Statistical Mathematics* **44**(1), 97–106.

Moran, P. A. P. (1969), 'Statistical inference with bivariate gamma distributions', *Biometrika* **56**(3), 627–634.

Nadarajah, S. (2009), 'A bivariate distribution with gamma and beta marginals with application to drought data', *Journal of Applied Statistics* **36**(3), 277–301.

Nadarajah, S. & Gupta, A. K. (2006), 'Some bivariate gamma distributions', *Applied Mathematics Letters* **19**(8), 767–774.

Sadeghkhani, A. & Ahmed, S. E. (2021), 'Predicting the Scoring Time in Hockey', *Journal of Statistical Theory and Practice* **15**(2), 1–14.

Yue, S. (2001), 'A bivariate gamma distribution for use in multivariate flood frequency analysis', *Hydrological Processes* **15**(6), 1033–1045.

Zhao, J., Jang, Y. H. & Kim, H. M. (2022), 'Closed-form and Bias-corrected Estimators for the Bivariate Gamma Distribution', *Journal of Multivariate Analysis* **105009**.