Research Article

Himchan Jeong and Emiliano A. Valdez*

Bayesian credibility premium with GB2 copulas

https://doi.org/10.1515/demo-2020-0009
Received March 11, 2020; accepted June 23, 2020

Abstract: For observations over a period of time, Bayesian credibility premium may be used to predict the value of a response variable for a subject, given previously observed values. In this article, we formulate Bayesian credibility premium under a change of probability measure within the copula framework. Such reformulation is demonstrated using the multivariate generalized beta of the second kind (GB2) distribution. Within this family of GB2 copulas, we are able to derive explicit form of Bayesian credibility premium. Numerical illustrations show the application of these estimators in determining experience-rated insurance premium. We consider generalized Pareto as a special case.

Keywords: hierarchical models, Bayesian posterior, insurance credibility, generalized beta of the second kind (GB2) and its copula, generalized Pareto (GP)

MSC: 62E15, 62F15, 62P05

1 Introduction

For our purpose, a random variable is always well defined on a given probability space and all random variables are continuous. Consider a sequence of random variables $Y_1, \ldots, Y_T$ with the following Bayesian framework:

- $Y_t|\Theta \sim f_{Y_t}(y_t|\theta)$ are independent for $t = 1, \ldots, T$ and
- $\Theta \sim p(\theta)$ is the prior distribution.

It is easy to see that the posterior density of $\Theta|Y_T$ has the expression

$$p(\theta|Y_T = y_T) = \frac{\prod_{t=1}^{T} f_{Y_t}(y_t|\theta)p(\theta)}{\int \prod_{t=1}^{T} f_{Y_t}(y_t|\theta)p(\theta)d\theta}, \quad (1.1)$$

and its posterior mean is given by $E[\Theta|Y_T = y_T] = \int \theta p(\theta|y_T)d\theta$, where $Y_T = (Y_1, \ldots, Y_T)'$, $y_T = (y_1, \ldots, y_T)'$, and $h_T(y_T)$ is the multivariate density function for $Y_T$.

Define a new observation for $t = T + 1$ as $Y_{T+1}$ whereby $Y_{T+1}|\Theta$ is independent of $Y_t|\Theta$ for $t = 1, \ldots, T$. Then we can define the expectation of the new observation, given $\Theta$, as

$$\mu(\theta) = E[Y_{T+1}|\Theta] = \int y_{T+1} f_{Y_{T+1}}(y_{T+1}|\theta)dy_{T+1}, \quad (1.2)$$

*Corresponding Author: Emiliano A. Valdez: Department of Mathematics, University of Connecticut, E-mail: emiliano.valdez@uconn.edu
and the Bayesian posterior mean as
\[ \mathbb{E} [\mu(\Theta) | Y_T = y_T] = \int \mu(\Theta) p(\Theta | y_T) d\Theta \]
in the same manner as [4]. We obtain the following lemma.

**Lemma 1.** Given the above Bayesian specification, the following relation holds:
\[ \mathbb{E} [\mu(\Theta) | Y_T = y_T] = \mathbb{E} [Y_{T+1} | Y_T = y_T]. \] (1.3)

**Proof.** From equations (1.1) and (1.2), we have
\[
\mathbb{E} [\mu(\Theta) | Y_T = y_T] = \int \left( \int y_{T+1} f_{T+1}(y_{T+1} | \Theta) dy_{T+1} \right) \prod_{t=1}^{T} f_T(y_t | \Theta) p(\Theta) d\Theta h_T(y_T) dy_T
\]
\[
= \int \left( \int \frac{T}{h_T(y_T)} \right) y_{T+1} \frac{h_{T+1}(y_{T+1})}{h_T(y_T)} dy_{T+1}
\]
\[= \int y_{T+1} f_{T+1}(y_{T+1} | y_T) dy_{T+1} = \mathbb{E} [Y_{T+1} | Y_T = y_T]. \]

\[ \square \]

As a special case, let us assume the following model specification:
\[ Y_1, \ldots, Y_T, Y_{T+1} | \Theta \overset{i.i.d.}{\sim} N(\Theta, (1-\rho)\sigma^2) \quad \text{and} \quad \Theta \overset{i.i.d.}{\sim} N(\mu, \rho \sigma^2). \]
Then \( Y_T \) has a multivariate normal distribution with a common pairwise correlation of \( \rho \) so that we can write the correlation matrix as
\[
\Sigma_T = \begin{pmatrix}
1 & \rho & \cdots & \rho \\
\rho & 1 & \cdots & \rho \\
\vdots & \vdots & \ddots & \vdots \\
\rho & \rho & \cdots & 1
\end{pmatrix} = (1-\rho) \left( I_T + \frac{\rho}{1-\rho} 1_T 1_T' \right),
\]
where \( 1_T = (1, 1, \ldots, 1) \) is a row vector of 1’s with dimension \( T \) and \( I_T \) is a \( T \times T \) identity matrix. In this case, we can show that all the marginals are also identically distributed as normal with mean \( \mu \) and variance \( \sigma^2 \) and the Bayesian posterior mean has the familiar form of a credibility premium, as shown in [2]:
\[ \mathbb{E} [Y_{T+1} | Y_T = y_T] = \frac{\rho T}{1-\rho + \rho T} \cdot \bar{y} + \frac{1-\rho}{1-\rho + \rho T} \cdot \mu, \] (1.4)
where \( \bar{y} = \frac{1}{T} \sum_{t=1}^{T} y_t \), the average of past values.

The Bayesian specification captures the heterogeneity of each subject in a longitudinal dataset given observed values of \( y_1, \ldots, y_T \). The Bayesian posterior mean in equation (1.3) is the predicted value for the next time period \( T + 1 \) given the historical observations. In the context of insurance, equation (1.4) is called a credibility-weighted premium and is used to calculate the subsequent year’s premium given the past history of claims, for claims that follow the normal distribution.

This paper extends this result in the framework where we have a GB2 copula. In Section 2, we show a construction of the multivariate GB2 and study its properties. In Section 3, we briefly define the concept of copula and introduce the Bayesian credibility premium as an expectation under a change of probability measure. In Section 4, we show that using this change of measure, it becomes straightforward to derive Bayesian credibility premium with GB2 copulas. We consider a numerical illustration in Section 5. Section 6 discusses the generalized Pareto as a special case. Finally, we provide concluding remarks in Section 7.
2 The multivariate GB2 distribution

To construct the multivariate GB2 distribution, consider the following model specification:

\[ Y_i | \Theta \overset{I.i.d.}{\sim} \text{G-Gamma} \left( \psi, \frac{\Gamma(\psi)}{\Gamma(\psi + 1/p)} \Theta, p \right) \quad \text{and} \quad \Theta \sim \text{GI-Gamma} \left( k, \frac{\Gamma(k)}{\Gamma(k - 1/p)} \mu, p \right), \]

where

- G-Gamma \( \left( \psi, \frac{\Gamma(\psi)}{\Gamma(\psi + 1/p)} \Theta, p \right) \) refers to a generalized gamma with density

\[
f_Y(y; \Theta) = \frac{p}{y \Gamma(\psi)} \left( \frac{y \Gamma(\psi + 1/p)}{\Gamma(\psi) \Theta} \right)^{p\psi} \exp \left[ - \frac{\left( y \Gamma(\psi + 1/p) \right)^p}{\Gamma(\psi) \Theta} \right],
\]

and

- GI-Gamma \( \left( k, \frac{\Gamma(k)}{\Gamma(k - 1/p)} \mu, p \right) \) refers to a generalized inverse-gamma with density

\[
p(\theta) = \frac{p}{\Theta^k} \left( \frac{\Gamma(k) \mu}{\Gamma(k - 1/p) \Theta} \right)^{pk} \exp \left[ - \frac{\left( \Gamma(k) \mu \right)^p}{\Gamma(k - 1/p) \Theta} \right],
\]

for \( p > 0 \). This specification can easily lead us to the following multivariate GB2 distribution by integrating out the scale parameter \( \Theta \):

\[
h_T(y_T) = \int_0^\infty \prod_{t=1}^T f_{Y_t}(y_t; \Theta)p(\theta)d\theta = \frac{p^T}{\prod_{t=1}^T \Gamma(y_t^p + c^p)} \prod_{t=1}^T y_t^{p\psi} \sum_{t=1}^T \Gamma(k \mu - p), \quad (2.1)
\]

where \( B(\Psi_T) \) is the multivariate beta function, \( c = \mu \frac{\Gamma(k)}{\Gamma(k - 1/p) \Gamma(\psi + 1/p)} \), and \( \Psi_T = (\psi, \psi, \ldots, \psi, k) \) is a vector of size \( (T + 1) \) with equal first \( T \) elements. Since \( Y_i | \Theta \) are i.i.d., then the marginal distribution is straightforward to determine by setting \( T = 1 \) so that

\[
f_Y(y) = \frac{p}{\psi B(\psi, k)} \left( \frac{\Gamma(k)}{\psi^p + c^p} \right)^{\psi + k}, \quad (2.2)
\]

where \( B(\Psi) = B(\psi, k) \) with \( B(\cdot, \cdot) \) is the beta function. For a random variable \( Y \), with density given in equation (2.2), we can write \( Y_T \sim \text{GB2}(k, c, \psi, p) \). See [12] and [11] for applications of univariate GB2. A similar derivation of the multivariate GB2 has appeared in [17], but the parameterization is different from above.

From the model specification, we can easily deduce the following moment properties:

\[
\mathbb{E}[Y_i | \Theta] = \Theta \quad \text{and} \quad \text{Cov}(Y_i, Y_{i'} | \Theta) = \begin{cases} \Theta^2 \left[ \frac{B(\psi + 2/p, \psi)}{B(\psi + 1/p, \psi + 1/p)} - 1 \right], & \text{for } t = t' \\ 0, & \text{for } t \neq t' \end{cases},
\]

and

\[
\mathbb{E}[\Theta] = \mu \quad \text{and} \quad \text{Var}(\Theta) = \mu^2 \left( \frac{B(k - 2/p, k)}{B(k - 1/p, k - 1/p)} - 1 \right).
\]

These properties lead us to the unconditional mean and variance, respectively, of a GB2 distribution:

\[
\mathbb{E}[Y_i] = \mathbb{E}[\mathbb{E}[Y_i | \Theta]] = \mathbb{E}[\Theta] = \mu,
\]

\[
\text{Var}(Y_i) = \mathbb{E}[\text{Var}(Y_i | \Theta)] + \mathbb{E}[\text{Var}(\Theta)] = \mathbb{E} \left[ \Theta^2 \left( \frac{B(\psi + 2/p, \psi)}{B(\psi + 1/p, \psi + 1/p)} - 1 \right) \right] + \text{Var}(\Theta)
\]

\[
= \mu^2 \left[ \frac{B(k - 2/p, k)}{B(k - 1/p, k - 1/p)} \frac{B(\psi + 2/p, \psi)}{B(\psi + 1/p, \psi + 1/p)} - 1 \right]. \quad (2.3)
\]

The following lemma shows that the multivariate GB2 has a pairwise correlation structure.
Lemma 2. Suppose \((Y_1, \ldots, Y_T)'\) follows a multivariate GB2 distribution as given in (2.1). Then they have a so-called pairwise correlation structure. In other words,

\[
\text{Corr}(Y_t, Y_{t'}) = \begin{cases} 
1, & \text{if } t = t' \\
\rho_{p,k,\psi}, & \text{if } t \neq t'
\end{cases}
\]

where

\[
\rho_{p,k,\psi} = \frac{B(k - 2/p, k) - 1}{B(k - 1/p, k - 1/p) - 1}.
\]

Proof. By definition, \(\text{Corr}(Y_t, Y_{t'}) = 1\). It is straightforward to see that in the case where \(t \neq t'\), we have

\[
\text{Cov}(Y_t, Y_{t'}) = \mathbb{E} \left[ \text{Cov}(Y_t, Y_{t'} | \Theta) + \text{Cov}(\mathbb{E} [Y_t | \Theta], \mathbb{E} [Y_{t'} | \Theta]) \right]
\]

\[
= \text{Cov}(\Theta, \Theta) = \mu^2 \left( \frac{B(k - 2/p, k)}{B(k - 1/p, k - 1/p)} - 1 \right),
\]

and from equation (2.3),

\[
\text{Var}(Y_t) = \text{Var}(Y_{t'}) = \mu^2 \left( \frac{B(k - 2/p, k)}{B(k - 1/p, k - 1/p)} \frac{B(\psi + 2/p, \psi)}{B(\psi + 1/p, \psi + 1/p)} - 1 \right).
\]

The result for \(t \neq t'\) immediately follows because \(\text{Corr}(Y_t, Y_{t'}) = \frac{\text{Cov}(Y_t, Y_{t'})}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_{t'})}} = \frac{\text{Cov}(Y_t, Y_{t'})}{\text{Var}(Y_t)} \). \(\Box\)

Since it is known that for any \(a\), \(\lim_{k \to \infty} \frac{\Gamma(k + a)}{\Gamma(k) k^a} = 1\), we have that for any fixed \(p\),

\[
\lim_{k \to \infty} \frac{B(k - 2/p, k)}{B(k - 1/p, k - 1/p)} = \lim_{k \to \infty} \frac{\Gamma(k - 2/p)}{\Gamma(k - 1/p) k^2} = \lim_{k \to \infty} \frac{\Gamma(k)}{\Gamma(k - 1/p) (k - 1/p)^{1/p}} = 1.
\]

This implies \(\lim_{k \to \infty} \text{Cov}(Y_t, Y_{t'}) = \lim_{k \to \infty} \rho_{p,k,\psi} = 0\) if \(t \neq t'\).

The following lemma shows that conditional distribution of each component of a multivariate GB2 random vector remains a member of the GB2 family.

Lemma 3. Suppose \((Y_1, \ldots, Y_T)'\) follows a multivariate GB2 distribution as given in (2.1). Then for any \(j = 1, \ldots, T\),

\[
(Y_j | Y_1, \ldots, Y_{j-1}, Y_{j+1}, \ldots, Y_T) \sim \text{GB2}(k + \psi(T - 1), c^*, \psi, p),
\]

where \(c^* = c \left( 1 + \sum_{t \neq j} (y_t/c)^p \right)^{1/p}\).

Proof. The proof is straightforward by noting that

\[
f(Y_j | Y_1, \ldots, Y_{j-1}, Y_{j+1}, \ldots, Y_T) = \frac{h_T(y_T)}{\int_0^\infty h_T(y_T) dy_j}.
\]

\(\Box\)

3 Change of probability measure with copulas

Copula is a widely used method to model dependency among multivariate observations. It has increased in popularity in recent years because of its widespread applications in several disciplines including, but not limited to, medical science, demography, hydrology, insurance, finance, and engineering. With copulas, one can decompose the marginal distributions and their dependence structure. See [9], [5], [6], [15], and [14].
Copulas are functions that join (or couple) the multivariate distribution functions to their one-dimensional marginal distributions functions. See [8]. Specifically, we have

\[ H_T(y_1, \ldots, y_T) = \mathbb{P}(Y_1 \leq y_1, \ldots, Y_T \leq y_T) = C_T(F_1(y_1), \ldots, F_T(y_T)), \]

where \( F_t(\cdot) \) refers to the marginal distribution associated with \( Y_t \), \( H_T \) is their joint distribution function, and \( C_T \) is the corresponding copula function where subscripts of \( H \) and \( C \) refer to the dimension of the random vector. [16] proved the existence of copulas for every joint distribution function and demonstrated that they are indeed unique if the marginal distribution functions are continuous. It is also sometimes convenient to write this as

\[ H_T(y_1, \ldots, y_T) = C_T(u_1, \ldots, u_T), \]

where \( u_T = (u_1, \ldots, u_T)' \) and \( u_t = F_t(y_t) \) with \( F_t \) the marginal distribution for \( t = 1, \ldots, T \).

Vectors shall be written in bold letters. For example, we shall denote the observed values of \( Y_T \) by \( y_T = (y_1, \ldots, y_T)' \) and similarly for \( Y_{T+1} \) by \( y_{T+1} = (y_1, \ldots, y_T, y_{T+1})' \). For ease of notation, we will also denote the vectors \( u_T = (u_1, \ldots, u_T)' \) and \( u_{T+1} = (u_1, \ldots, u_T, u_{T+1})' \). We shall assume that the densities of the copulas exist and are respectively denoted by

\[ c_T(u_T) = \frac{\partial^T C_T(u_T)}{\partial u_1 \cdots \partial u_T}, \]

and

\[ c_{T+1}(u_{T+1}) = \frac{\partial^{T+1} C_{T+1}(u_{T+1})}{\partial u_1 \cdots \partial u_T \partial u_{T+1}}. \]

Notice that the marginal distribution functions have been denoted by

\[ F_t(y_t) = \mathbb{P}(Y_t \leq y_t) \quad \text{for} \quad t = 1, \ldots, T, T + 1, \]

and if the corresponding density functions exist, we denote them by

\[ f_t(y_t) = \frac{dF_t(y_t)}{dy_t} \quad \text{for} \quad t = 1, \ldots, T, T + 1. \]

Similarly, the multivariate density functions, if they exist, will be respectively denoted by

\[ h_T(y_T) = \frac{\partial^T H_T(y_T)}{\partial y_1 \cdots \partial y_T} \quad \text{and} \quad h_{T+1}(y_{T+1}) = \frac{\partial^{T+1} H_{T+1}(y_{T+1})}{\partial y_1 \cdots \partial y_T \partial y_{T+1}}. \]

In the following theorem, we derive the fundamental building block for deriving the Bayesian credibility premium, \( \mathbb{E} [Y_{T+1}|Y_T = y_T] \), within a copula framework.

**Theorem 1.** Consider the copula model satisfying the assumptions described in this section. The conditional expectation of \( Y_{T+1}|Y_T \) can be expressed in the following manner:

\[ \mathbb{E} [Y_{T+1}|Y_T = y_T] = \int y_{T+1} \frac{c_{T+1}(u_{T+1})}{c_T(u_T)} f_T(y_T) \, dF_{T+1}(y_{T+1}), \]

where \( c_T(u_T) \) and \( c_{T+1}(u_{T+1}) \) are respectively defined in (3.2) and (3.3), and that \( F_{T+1} \) is a known marginal distribution function of \( Y_{T+1} \).

**Proof.** It is clear from the definition of conditional density, that if it exists, we must have

\[ f_{Y_{T+1}|Y_T}(y_{T+1}|y_T) = \frac{h_{T+1}(y_{T+1})}{h_T(y_T)}. \]

The numerator can be written as

\[ h_{T+1}(y_{T+1}) = \frac{\partial^{T+1} C_{T+1}(u_{T+1})}{\partial u_1 \cdots \partial u_T \partial u_{T+1}} \times \prod_{t=1}^{T+1} f_t(y_t) = c_{T+1}(u_{T+1}) \times \prod_{t=1}^{T+1} f_t(y_t), \]
where \( u_{r+1} \) is understood to be evaluated at the respective marginals \( F_t(y_t) \) for \( t = 1, \ldots, T, T+1 \). Similarly, we have

\[
h_T(y_T) = c_T(u_T) \times \prod_{i=1}^{T} f_i(y_i).
\]

From these, we now have

\[
\mathbb{E}[Y_{T+1}|Y_T = y_T] = \int y_{T+1} f_{Y_{T+1}|Y_T}(y_{T+1}|y_T) dy_{T+1} = \int y_{T+1} \frac{h_{T+1}(y_{T+1})}{h_T(y_T)} dy_{T+1}
\]

\[
= \int y_{T+1} c_{T+1}(u_{T+1}) \times \prod_{i=1}^{T+1} f_i(y_i) \frac{c_T(u_T)}{c_T(u_T)} \times \prod_{i=1}^{T} f_i(y_i)
\]

and the result given in (3.4) follows.

Indeed, if we define \( F^{Q}_{T+1}(y_{T+1}) = c_{T+1}(u_{T+1}) f_{T+1}(y_{T+1}) \), we can verify that \( dF^{Q}_{T+1} = f^{Q}_{T+1}(y_{T+1}) \cdot dy_{T+1} \) becomes a probability measure because both \( c_T(u_T) > 0 \) and \( f_{T+1}(y_{T+1}) > 0 \) for any value of \( T \). More specifically, we have

\[
\int dF^{Q}_{T+1} = \int f^{Q}_{T+1}(y_{T+1}) dy_{T+1} = \int c_{T+1}(u_{T+1}) f_{T+1}(y_{T+1}) dy_{T+1}
\]

\[
= \int h_{T+1}(y_{T+1}) \times \prod_{i=1}^{T+1} f_i(y_i) \frac{h_T(y_T)}{c_T(u_T)} \times \prod_{i=1}^{T} f_i(y_i)
\]

In general, we can derive the copula structure based on a random effect framework as follows:

\[
C_T(u_1, \ldots, u_T) = \prod_{i=1}^{T} F_{Y_t|\theta}(F^{-1}_T(u_i)) \pi(\theta) d\theta = \mathbb{E}_\theta \left[ \prod_{i=1}^{T} F_{Y_t|\theta}(F^{-1}_T(u_i)) \right],
\]

where \( F_{Y_t|\theta} \) and \( F_t \) denotes the conditional and marginal distribution functions of \( Y_t \), respectively.

### 4 GB2 and Bayesian credibility premium

Note that when \( Y_t \sim GB2(k, \psi, p) \), we derive the marginal distribution \( F_t \) as follows. By letting \( z = y^p \) and \( v = \frac{z}{z + c^p} \), we have

\[
F_t(y_t) = \int_{0}^{y_t} \frac{p}{y \psi(k)} \frac{y^p c^p}{[y^p + c^p]^\psi k} dy = \int_{0}^{y_t} \frac{1}{z \psi(k)} \frac{z \psi k}{z + c^p} dz
\]

\[
= \int_{0}^{v \psi(k)} \frac{1}{B(\psi, k)} v^{\psi-1}(1 - v)^{k-1} dv = B_{\psi, k} \left( \frac{y_t^p}{c^p + y_t^p} \right),
\]

where \( B_{\psi, k}(t) \) is the distribution function of Beta(\( \psi, k \)) distribution so that \( F^{-1}_t(u_t) = c \left( \frac{B_{\psi, k}^{-1}(u_t)}{1 - B_{\psi, k}^{-1}(u_t)} \right)^{1/p} \).

Therefore, the GB2 copula can be derived in the same manner as follows:

\[
C_{k, \psi}(u_1, \ldots, u_T) = \mathbb{E}_\theta \left[ \prod_{i=1}^{T} F_{Y_t|\theta} \left( c \left( \frac{B_{\psi, k}^{-1}(u_i)}{1 - B_{\psi, k}^{-1}(u_i)} \right)^{1/p} \right) \right] \pi(\theta) d\theta
\]

\[
= \mathbb{E}_\theta \left[ \prod_{i=1}^{T} F_{Y_t|\theta} \left( c \left( \frac{B_{\psi, k}^{-1}(u_i)}{1 - B_{\psi, k}^{-1}(u_i)} \right)^{1/p} \right) \right], \quad (4.1)
\]
and its corresponding density is

\[ c_{k,\psi}(u_1, \ldots, u_T) = \frac{\partial^T}{\partial u_1 \ldots \partial u_T} \int \prod_{t=1}^T F_{Y_t | \theta}(F_t^{-1}(u_t)) \pi(\theta) d\theta \]

\[ = \frac{\int \prod_{t=1}^T f_{Y_t | \theta}(F_t^{-1}(u_t)) \pi(\theta) d\theta}{\prod_{t=1}^T f_t(F_t^{-1}(u_t))} = \frac{\int \prod_{t=1}^T f_{Y_t | \theta}(c \cdot q_t^{-1/p}) \pi(\theta) d\theta}{\prod_{t=1}^T f_t(c \cdot q_t^{-1/p})} = \frac{h_T(c \cdot q_t^{-1/p})}{\prod_{t=1}^T f_t(c \cdot q_t^{-1/p})} \quad (4.2) \]

where \( q_t^{1/p} = F_t^{-1}(u_t)/c = \left( \frac{B_{\psi,k}(u_t)}{1 - B_{\psi,k}(u_t)} \right)^{1/p} \).

Note that if we substitute \( F_t(y_t) \) for \( u_t \), where \( F_t(y_t) \) is a marginal distribution function of GB2, it turns out that \( c_{k,\psi}(u_1, \ldots, u_T) \) becomes \( h_T(y_T) \), a joint density of multivariate GB2 distribution.

By Fubini’s theorem, the bivariate GB2 copula is given as follows:

\[ C_{k,\psi}(u_1, u_2) = \int_0^\infty \int_0^\infty \left( \int \prod_{t=1}^T f_{Y_t | \theta}(y_t) f_{Y_{t-1} | \theta}(y_{t-1}) \pi(\theta) d\theta \right) dy_1 dy_2 \]

\[ = \int_0^\infty \int_0^\infty \left( \int (y_1/c)^p (y_2/c)^p \prod_{t=1}^T (y_t/c)^p \prod_{t=1}^T (y_t/c)^p + 1)^{2p+1} dy_1 dy_2 \right) dy_1 dy_2 \]

where \( v_1 = (y_1/c)^p \) for \( i = 1, 2 \). From the above derivation, we see that the GB2 copula only depends on \( \psi \) and \( k \), but not on \( c \) and \( p \). Figure 1 provides a comparison of the contour plots of bivariate GB2 copulas using different sets of parameters. Both parameters \( \psi \) and \( k \) describe the strength of the relationship as seen in this figure. For example, for a fixed \( k \), a higher value of \( \psi \) implies stronger dependence and for a fixed \( \psi \), a higher value of \( k \) implies weaker dependence.

Now, by applying (4.2), we can get the following result which is a crucial step to derive the Bayesian credibility premium under the multivariate GB2 distribution model.

**Lemma 4.** Suppose \((Y_1, \ldots, Y_{T+1})'\) follows a multivariate GB2 distribution as given in (2.1) and \( c_T \) is the density of GB2 copula. Then the following relationship holds:

\[ f_{Y_{T+1}}(y_{T+1}) \frac{c_{T+1}(u_{T+1})}{c_T(u_T)} \bigg|_{u_T=F(y_T)} = \frac{p}{y_{T+1} B(\psi, k_T)} \frac{c_{T,\psi}^{pkr} y_T^{p\psi}}{(c_{T,\psi}^{pkr} y_T^{p\psi} + m_{T+1} y_T^{p\psi + k_T})}, \]

where \( c_{T,\psi}^{pkr} = (c_T^p + \sum_{t=1}^T y_t^p)^{1/p} \), \( k_T = k + \psi T \).

**Proof.** See appendix for details of the proof. \( \square \)

Based on Theorem 1, it is possible to evaluate the Bayesian credibility premium with GB2 copulas from the following theorem.

**Theorem 2.** Suppose \((Y_1, \ldots, Y_{T+1})'\) follows a multivariate distribution as described in (3.1) where \( C_T \) is given as GB2 copula in (4.1) and the marginal distribution of \( Y_i \) is given as \( F_i \). Then the Bayesian credibility premium
Figure 1: Contours and scatterplots of bivariate GB2 copula with various parameterization
is written as follows:

\[
E[Y_{T+1}|Y_T = y_T] = \int_0^1 F_{T+1}^{-1}(u) \left( \frac{(1 + \eta_T)^{k_t}}{1 - B_{\psi,k}^{-1}(u)} \right)^{\psi + k} \frac{B(k, k_T + \psi)}{B(k_T, k + \psi)} du,
\]

where \( q_t = \left( \frac{B_{\psi,k}^{-1}(F_t(y_t))}{1 - B_{\psi,k}^{-1}(F_t(y_t))} \right) \), \( \eta_T = \sum_{t=1}^T q_t \), and \( k_T = \psi T + k \).

**Proof.** From Theorem 1 and (4.2), we can see that

\[
E[Y_{T+1}|Y_T = y_T] = \int y_{T+1} c_{T+1}(u_{T+1}) \frac{c_T(u_T)}{c_T(y_T)} dF_{T+1}(y_{T+1})
\]

\[
= \int y_{T+1} \left( \frac{1 + \sum_{t=1}^T q_t}{1 + \sum_{t=1}^T q_t} \right)^{k_t} \frac{(1 + q_{T+1})^{\psi + k}}{\Gamma(k_T)\Gamma(k + \psi)} \frac{\Gamma(k)\Gamma(k_T + \psi)}{\Gamma(k_T)\Gamma(k + \psi)} dF_{T+1}(y_{T+1})
\]

\[
= \int y_{T+1} \left( \frac{1 + \eta_T}{1 + \sum_{t=1}^T q_t} \right)^{k_t} \frac{(1 + q_{T+1})^{\psi + k}}{B(k, k_T + \psi)} \frac{B(k, k_T)}{B(k_T, k + \psi)} dF_{T+1}(y_{T+1}).
\]

Since \( 1 + q_{T+1} = \left( \frac{1}{1 - B_{\psi,k}^{-1}(F_{T+1}(y_{T+1}))} \right) \) and \( F_{T+1}(y_{T+1}) \sim U[0, 1] \), (4.4) can be expressed as follows:

\[
E[Y_{T+1}|Y_T = y_T] = \int_0^1 F_{T+1}^{-1}(u) \left( \frac{1 + \eta_T}{1 + \sum_{t=1}^T q_t} \right)^{k_t} \frac{(1 + q_{T+1})^{\psi + k}}{\Gamma(k_T)\Gamma(k + \psi)} \frac{\Gamma(k)\Gamma(k_T + \psi)}{\Gamma(k_T)\Gamma(k + \psi)} \frac{B(k, k_T + \psi)}{B(k, k_T + \psi)} du.
\]

As a special case of Theorem 2, it is possible to derive a nice closed form of Bayesian credibility premium when \( Y_T \) follows a multivariate GB2 distribution as follows.

**Corollary 1.** Suppose \( Y_1, \ldots, Y_{T+1} \) follows a multivariate distribution as described in (3.1) where \( c_T \) is given as GB2 copula in (4.1) and the marginal distribution of \( Y_i \) is univariate GB2 distribution as in (2.2). Then the Bayesian credibility premium is given as follows:

\[
E[Y_{T+1}|Y_T = y_T] = \left( 1 + \sum_{t=1}^T (y_t/c)^{1/p} B(k, k_T - 1/p) B(k - 1/p, k_T)^{-1} \right) \frac{B(k, k_T + \psi)}{B(k, k_T + \psi)}
\]

**Proof.** From Theorem 1 and Lemma 4, we can see that

\[
f_{T+1}^{\psi}(y_{T+1}) = f_{T+1}(y_{T+1}) \frac{c_{T+1}(u_{T+1})}{c_T(u_T)} = \frac{p}{y_{T+1}} B(y, k_T)^{-1} \frac{c_{T+1}^{p^{k_T p^{\psi}} y_{T+1}}}{(c_{T+1}^{p^{*}})^{p^{*} y_{T+1}}}. \]

This ratio of densities of the copula, \( c_{T+1}(u_{T+1})/c_T(u_T) \), in fact induces a change of probability measure so that in effect, we can write the prediction as the following unconditional expectation

\[
E[Y_{T+1}|Y_T = y_T] = E^D[Y_{T+1}],
\]
under a change of measure $dF^Q_{T+1}(y_{T+1}) = \frac{c_{T+1}(u_{T+1})}{c_T(u_T)} dF_{T+1}(y_{T+1})$. This change of measure allows us to construct an explicit expression for the posterior mean based on

$$Y_{T+1} \sim \text{GB2}(k, c, \psi, p) \quad \text{under } dF_{T+1}(y_{T+1}),$$

and

$$Y_{T+1} \sim \text{GB2}(k_T, c_{T,p}, \psi, p) \quad \text{under } dF^Q_{T+1}(y_{T+1}).$$

Note that if $Y \sim \text{GB2}(k, c, \psi, p)$, then $\mathbb{E}[Y] = \frac{\Gamma(k+1/p)}{\Gamma(k)} c$ from (2.3). Therefore, we can get the following result directly from the definition of $c_{T,p}$:

$$\mathbb{E}[Y_{T+1} | Y_T = y_T] = \mathbb{E}^Q[Y_{T+1}] = \frac{\Gamma(k_T+1/p)}{\Gamma(k_T)} c_{T,p} \mathbb{E}[Y_T]$$

$$= \left(1 + \sum_{t=1}^T (y_t/c)^{1/p}\right) \frac{\Gamma(k_T+1/p)}{\Gamma(k_T)} c_{T,p} \mathbb{E}[Y_T]$$

$$= \left(1 + \sum_{t=1}^T (y_t/c)^{1/p}\right) \frac{B(k_T-1/p, k)}{B(k_T, k-1/p)} \mu,$$

since $c = \frac{\Gamma(k_T+1/p)}{\Gamma(k_T)} c_{T,p} \mu$. \hfill \Box

If the marginal distribution of each does now follow GB2, then the integral in (4.3) might not able to be evaluated analytically. The corollary below provides a nice approximation of the credibility premium based on Monte Carlo method.

**Corollary 2.** A Monte Carlo approximation of Bayesian credibility premium in (4.3) is given as follows:

$$\mathbb{E}[Y_{T+1} | Y_T = y_T] \approx \frac{1}{M} \sum_{m=1}^M F_{T+1}^{-1}(u_{[m]}) (1 + \eta_T)^k \left[1 \left(\frac{1}{1 - B_{T,k}(u_{[m]})} \right) \psi + k \right] B(k_T + \psi) B(k_T, k + \psi)^k$$

where $u_{[m]}$ for $m = 1, \ldots, M$ are random samples generated from $t([0, 1])$.

If $T = 0$, which means the case when there is no history of past claims, then $\eta_T = 0$ and $k_T = k$ so that (4.6) is reduced as follows:

$$\mathbb{E}[Y_{T+1} | Y_T = y_T] \approx \frac{1}{M} \sum_{m=1}^M F_{T+1}^{-1}(u_{[m]}),$$

which is indeed a natural Monte Carlo approximation of the prior mean of $Y_{T+1}$, $\mathbb{E}[Y_{T+1}]$.

## 5 Numerical illustration - insurance ratemaking

For numerical illustration, we consider the case where our primary variable of interest, $y$, is the amount of claim for a portfolio of insurance contracts. In particular, we have a set of random claims for $T = 10$ periods: $y_1, \ldots, y_{10}$. Determining the pure premium based on historical claim experience is the subject of experience rating and credibility. In this case, the pure premium is the Bayesian credibility premium with multivariate GB2 discussed in Corollary 1. This numerical example shows how the choice of parameters affect Bayesian credibility premium with observed claim experience. To fix ideas, we control the prior mean, $\mu$, to be the same as $10$ for all cases, but we vary the values of the corresponding coefficient of variation (CV), which is defined
as \(\sqrt{\text{Var}(Y)/\mathbb{E}[Y]}\). Table 1 shows different combinations of parameters which have the same mean but with different coefficient of variations, respectively.

Table 1: Combinations of parameters for the GB2 posterior means

| CV    | \(\mu\) | \(k\)   | \(\psi\) | \(c\) |
|-------|---------|---------|---------|-------|
| 10%   | 10      | 120.000 | 661.111 | 1.800 |
| 50%   | 10      | 7.000   | 24.000  | 2.500 |
| 90%   | 10      | 7.000   | 1.967   | 30.500|
| 100%  | 10      | 7.000   | 1.500   | 40.000|
| 110%  | 10      | 5.067   | 1.500   | 27.117|
| 150%  | 10      | 3.053   | 1.500   | 13.684|
| 200%  | 10      | 2.500   | 1.500   | 10.000|

Using these combinations of parameters, we assume three scenarios of observed claims, where each scenario is represented by the randomly generated quantiles of incurred claims. The first is a ‘risky’ scenario so that the average of generated quantiles is 56.84%. The second is a ‘normal’ scenario so that the average of generated quantiles is 49.63%. Finally, the last is a ‘safe’ scenario so that the average of generated quantiles is 41.50%.

After assuming three scenarios based on the quantiles of the observed claims, we convert the quantile vectors to the observed claims under GB2 distribution for each set of parameters. We know that the ‘weight factor’ part of the GB2 credibility premium of the pure premium is given as

\[
\left(1 + \sum_{t=1}^{T} (y_t/c)^p \right)^{1/p} \frac{\beta_{k-1/2}(p,k)}{\beta(k; k-1/p)}
\]

from Corollary 1. Therefore, the value of weight factor is determined both by the parameters and the observed claim values. Note that even though we have the same quantiles of generated claims, the generated claim amount can still vary along with the assumed set of parameters.

Table 2 shows the result of the weight factors under all scenarios and parameter assumptions. Here, \(w_H\) refers to the weight factors under the ‘risky’ scenario, \(w_M\) to the weight factors under the ‘normal’ scenario, and \(w_L\) to the weight factors under the ‘safe’ scenario, respectively.

Table 2: Weight factors of the GB2 posterior mean under all scenarios and parameter assumptions

| CV    | \(w_H\) | \(w_M\) | \(w_L\) |
|-------|---------|---------|---------|
| 10%   | 1.0353 | 0.9972  | 0.9685  |
| 50%   | 1.1996 | 0.9977  | 0.8508  |
| 90%   | 1.2828 | 1.0051  | 0.7963  |
| 100%  | 1.2933 | 1.0069  | 0.7919  |
| 110%  | 1.3633 | 1.0109  | 0.7551  |
| 150%  | 1.5340 | 1.0195  | 0.6767  |
| 200%  | 1.6358 | 1.0211  | 0.6319  |

This numerical illustration presents some very intuitively interesting results. From Table 2, we can infer that as we have higher coefficient of variation, the impact of credibility weighing factor increases. Moreover, for a policyholder with relatively higher claim experience, the resulting credibility weighing factor is greater than 1, which implies a penalty to policyholders with unfavorable claim experience. On the other hand, for a
of the degree of correlation between pairs of GP random variables. Larger values of \( k \) imply less correlated variables.

Under the GP model specification, from (4.5), the Bayesian posterior mean is given as

\[
\mathbb{E} [Y_{T+1} | Y_T = y_T] = \left( 1 + \sum_{t=1}^{T} \frac{y_t}{c} \right) \frac{k-1}{kT-1} \mu = \frac{(k-1) \mu + \psi T \bar{Y}}{\psi T + k - 1} = \frac{\psi T}{\psi T + k - 1} \cdot \bar{Y} + \frac{k-1}{\psi T + k - 1} \cdot \mu,
\]
which can be directly derived from the GB2 framework by letting $p = 1$. Interestingly, this has the form of a weighted average of prior mean $\mu$ and the average of previous observations, as shown in [1], [10], and [3]. Note that this is not at all surprising and is a natural result because it is well known that if $Y_t|\Theta$ follows a distribution that belongs to the exponential family (Gamma distribution clearly belongs to the exponential family), then the posterior mean is exactly a linear combination of prior mean and sample mean of previous observations. For details of such result, see [7].

It is well known that gamma distribution is a choice for modeling the severity component of property and casualty insurance claims. The usual model specification entertained by many insurers is as follows:

$$Y_t \overset{i.i.d.}{\sim} \text{Gamma}(\psi, \mu/\psi),$$

so that $E[Y_t] = \mu$ and $\text{Var}(Y_t) = \mu^2/\psi$. We may regard this formulation as a limiting case of multivariate GP distribution because as $k \to \infty$,

$$\text{Cov}(Y_t, Y_{t'}) \to 0 \text{ and } \Theta \to \mu.$$

An insurance company may wish to use the multivariate GP distribution as a predictive claims model by carefully calibrating the value of $k$.

### 7 Conclusion

As stated in [13], the concept of Bayesian shrinkage estimation is not limited to a normally distributed random variable or a member of the exponential family. This article extends the literature by developing explicit forms of Bayesian credibility premium within the family of GB2 copulas. The development is based on a new concept of using change of probability measure for copulas; this result is stated in Theorem 1. This theorem is the fundamental foundation for developing the explicit forms. Such credibility premium are very useful in actuarial science and insurance for experience-based ratemaking, where contractholders may be rewarded or penalized depending on their own claim experience. The concepts in this paper can be readily applied and extended in several ways. First, note that it is possible to use regression function $g(x|\beta)$ as prior mean, instead of the grand mean $\mu$. Second, credibility premium with GB2 copulas, but with non-GB2 marginals, can be obtained so long as the random variables have continuous support as shown in Theorem 2 and Corollary 2. Finally, as in many diverse applications, the proposed credibility premium can have a wide ranging applications for various multiparameter inference and regression problems.

**Acknowledgements:** We appreciate helpful feedback provided by two anonymous referees. Himchan Jeong was supported by James C. Hickman Doctoral Stipend funded by the Society of Actuaries (SOA). Emiliano A. Valdez was supported by CAE Research Grant on Applying Data Mining Techniques in Actuarial Science funded by SOA.

### A Appendix

#### A.1 Proof of Lemma 4

First, we plug $F_t(y_t)$ in $u_t$ as $F_t(y_t) = B_{\psi, k} \left( \frac{y_t^p}{y_t^p + c^p} \right)$ so that

$$q_t = B_{\psi, k}^{-1}(F_t(y_t))/[1 - B_{\psi, k}^{-1}(F_t(y_t))] = (y_t/c)^p.$$
We now have
\[
c_T(u_T) \bigg|_{u=F(y)} = \frac{\Gamma(k) \Gamma((\psi T + k))}{\Gamma(\psi + k)^T} \times \frac{\prod_{t=1}^{T} (1 + (y_t/c)^p)^{\psi^k}}{(1 + \sum_{t=1}^{T} (y_t/c)^p)^{\psi^T}}.
\]

It follows that we can simplify \(c_{T+1}(u_{T+1})/c_T(u_T)\bigg|_{u=F(y)}\) as follows:
\[
c_{T+1}(u_{T+1})/c_T(u_T) \bigg|_{u=F(y)} = \frac{\Gamma(k) \Gamma((\psi + k))}{\Gamma(\psi + k)^T} \times \frac{\prod_{t=1}^{T} (1 + (y_t/c)^p)^{\psi^k}}{(1 + \sum_{t=1}^{T} (y_t/c)^p)^{\psi^T+k}}
\]
\[
\times \frac{\Gamma(k) \Gamma((\psi + k_T))}{\Gamma(\psi + k_T)^T} \times \frac{(1 + (y_{T+1}/c)^p)^{\psi^k}}{(1 + \sum_{t=1}^{T} (y_t/c)^p)^{\psi^T+k_T}} \times (1 + \sum_{t=1}^{T} (y_t/c)^p)^{\psi^T}
\]
\[
= \frac{\Gamma(k) \Gamma((\psi + k_T))}{\Gamma(\psi + k_T)^T} \times \frac{(c^p + y_{T+1}^p)^{\psi^k}}{(c^p + \sum_{t=1}^{T} y_t^p)^{\psi^T+k_T}} \times \frac{(c^p + \sum_{t=1}^{T} y_t^p)^{\psi^T+k_T}}{(c^p + \sum_{t=1}^{T} y_t^p)^{\psi^T}} \times \frac{c_{T,p}^{*} p^{k_T}}{c^{p^k}}
\]
where \(c_{T,p}^{*} = (c^p + \sum_{t=1}^{T} y_t^p)^{1/p}\) and \(k_T = k + \psi T\).

Note that \(f_{T+1}(y_{T+1})\) is given as
\[
f_{T+1}(y_{T+1}) = \int_{0}^{\infty} f_{Y_{T+1}}(y_{T+1} | \theta) p(\theta) d\theta
\]
\[
= \frac{p}{y_{T+1} B(\psi, k)} \frac{c^{p^k} y_{T+1}^{p^\psi}}{(c^p + y_{T+1}^p)^{\psi^k}} = \frac{p}{y_{T+1} \Gamma(\psi) \Gamma(k)} \frac{c^{p^k} y_{T+1}^{p^\psi}}{(c^p + y_{T+1}^p)^{\psi^k}}.
\]

Thus, we have
\[
f_{T+1}(y_{T+1}) c_{T+1}(u_{T+1})/c_T(u_T) \bigg|_{u=F(y)} = \frac{p}{y_{T+1} \Gamma(\psi) \Gamma(k)} \frac{c^{p^k} y_{T+1}^{p^\psi}}{(c^p + y_{T+1}^p)^{\psi^k}} \times \frac{\Gamma(k) \Gamma((\psi + k_T))}{\Gamma(\psi + k_T)^T} \times \frac{(c^p + y_{T+1}^p)^{\psi^k}}{(c^p + \sum_{t=1}^{T} y_t^p)^{\psi^T+k_T}} \times \frac{c_{T,p}^{*} p^{k_T}}{c^{p^k}}
\]
\[
= \frac{p}{y_{T+1} \Gamma(\psi) \Gamma(k)} \frac{c^{p^k} y_{T+1}^{p^\psi}}{(c^p + y_{T+1}^p)^{\psi^k}} \times \frac{\Gamma(k) \Gamma((\psi + k_T))}{\Gamma(\psi + k_T)^T} \times \frac{(c^p + y_{T+1}^p)^{\psi^k}}{(c^p + \sum_{t=1}^{T} y_t^p)^{\psi^T+k_T}} \times \frac{c_{T,p}^{*} p^{k_T}}{c^{p^k}}
\]
\[
= \frac{p}{y_{T+1} B(\psi, k_T)} \frac{c^{p^k} y_{T+1}^{p^\psi}}{(c^p + y_{T+1}^p)^{\psi^k}} \times \frac{\Gamma(k) \Gamma((\psi + k_T))}{\Gamma(\psi + k_T)^T} \times \frac{(c^p + y_{T+1}^p)^{\psi^k}}{(c^p + \sum_{t=1}^{T} y_t^p)^{\psi^T+k_T}} \times \frac{c_{T,p}^{*} p^{k_T}}{c^{p^k}}
\]
Electronic copy available at: https://ssrn.com/abstract=3373377
References

[1] Bailey, A. L. (1950). Credibility procedures: Laplace’s generalization of Bayes’ rule and the combination of collateral knowledge with observed data. *Proceeding of the Casualty Actuarial Society* 37(67), pp. 7–23.

[2] Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer, New York.

[3] Bühlmann, H. (1967). Experience rating and credibility. *Astin Bull.* 4(3), 199–207.

[4] Bühlmann, H. and A. Gisler (2005). *A Course in Credibility Theory and its Applications*. Springer, Berlin.

[5] Frees, E. W. and E. A. Valdez (1998). Understanding relationships using copulas. *N. Am. Actuar. J.* 2(1), 1–25.

[6] Hougaard, P., B. Harvald, and N. V. Holm (1992). Measuring the similarities between the lifetimes of adult danish twins born between 1881–1930. *J. Amer. Statist. Assoc.* 87(417), 17–24.

[7] Jewell, W. S. (1974). Credible means are exact bayesian for exponential families. *Astin Bull.* 8(1), 77–90.

[8] Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Chapman and Hall, London.

[9] Li, D. X. (2000). On default correlation: a copula function approach. *J. Fixed Income* 9(4), 43–54.

[10] Mayerson, A. L. (1964). A bayesian view of credibility. *Proceeding of the Casualty Actuarial Society* 51(95), pp. 85–104.

[11] McDonald, J. B. and R. M. Bookstaber (1991). Option pricing for generalized distributions. *Comm. Statist. Theory Methods* 20(12), 4053–4068.

[12] McDonald, J. B. and R. J. Butler (1990). Regression models for positive random variables. *J. Econometrics* 43(1-2), 227–251.

[13] Morris, C. N. (1983). Parametric empirical Bayes inference: theory and applications. *J. Amer. Statist. Assoc.* 78(381), 47–65.

[14] Salvadori, G. and C. De Michele (2007). On the use of copulas in hydrology: theory and practice. *J. Hydrol. Eng.* 12(4), 369–380.

[15] Shih, J. H. and T. A. Louis (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* 51(4), 1384–1399.

[16] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* 8, 229–231.

[17] Yang, X., E. W. Frees, and Z. Zhang (2011). A generalized beta copula with applications in modeling multivariate long-tailed data. *Insurance Math. Econom.* 49(2), 265–284.