Bayesian analysis of the generalized gamma distribution using non-informative priors

Pedro L. Ramos, Jorge A. Achcar, Fernando A. Moala, Eduardo Ramos & Francisco Louzada

Published online: 22 May 2017.

Submit your article to this journal

View Crossmark data

Taylor & Francis
Taylor & Francis Group

Check for updates

Bayesian analysis of the generalized gamma distribution using non-informative priors

Pedro L. Ramos [a], Jorge A. Achcar[b], Fernando A. Moala[c], Eduardo RAMOS[a] and Francisco Louzada[a]

[a]Institute of Mathematical and Computer Sciences, USP, Sao Carlos, Brazil; [b]Social Medicine Department, Medical School, USP, Ribeirao Preto, Brazil; [c]Department of Statistics, Sao Paulo State University, Presidente Prudente, Brazil

ABSTRACT

The Generalized gamma (GG) distribution plays an important role in statistical analysis. For this distribution, we derive non-informative priors using formal rules, such as Jeffreys prior, maximal data information prior and reference priors. We have shown that these most popular formal rules with natural ordering of parameters, lead to priors with improper posteriors. This problem is overcome by considering a prior averaging approach discussed in Berger et al. [Overall objective priors. Bayesian Analysis. 2015;10(1):189–221]. The obtained hybrid Jeffreys-reference prior is invariant under one-to-one transformations and yields a proper posterior distribution. We obtained good frequentist properties of the proposed prior using a detailed simulation study. Finally, an analysis of the maximum annual discharge of the river Rhine at Lobith is presented.

1. Introduction

Many generalizations for the standard exponential distribution have been proposed in the literature, namely the Weibull, gamma and lognormal distributions which provide more flexibility to describe real data sets. Stacy [1] unified these models using the Generalized gamma (GG) distribution with three parameters. Let T be a non-negative random variable following a GG distribution with probability density function (p.d.f.) given by

$$f(t|\boldsymbol{\theta}) = \frac{\alpha}{\Gamma(\phi)} \mu^{\alpha\phi} t^{\alpha\phi-1} \exp(-(\mu t)^{\alpha}), \tag{1}$$

for $t > 0$ and $\boldsymbol{\theta} = (\alpha, \mu, \phi)$ where $\alpha > 0$ and $\phi > 0$ are shape parameters and $\mu > 0$ is a rate parameter. The cumulative distribution function is given by

$$F(t|\boldsymbol{\theta}) = \int_0^{(\mu t)^{\alpha}} \frac{1}{\Gamma(\phi)} w^{\phi-1} e^{-w} \, dw = \frac{\gamma(\phi, (\mu t)^{\alpha})}{\Gamma(\phi)}, \tag{2}$$

where $\gamma(y, x) = \int_0^x w^{y-1} e^{-w} \, dw$ is the lower incomplete gamma function.

The GG distribution is known by different names such as the Stacy distribution, generalized Weibull and gamma-Weibull. Important probability distributions can be obtained from the GG distribution such as the Weibull distribution ($\phi = 1$), gamma distribution ($\alpha = 1$), lognormal distribution (limit case when $\phi \to \infty$) and the generalized normal distribution ($\alpha = 2$). The generalized normal distribution is also a distribution that includes various known distributions such as half-normal ($\phi =$

$\frac{1}{2}, \mu = 1/\sqrt{2}\sigma$), Rayleigh ($\phi = 1, \mu = 1/\sqrt{2}\sigma$), Maxwell–Boltzmann distribution ($\phi = \frac{3}{2}$) and chi distribution ($\phi = k/2, k = 1, 2, \ldots$). Moreover, the GG distribution is a flexible model for reliability data that provides different forms of the hazard function such as constant, increasing, decreasing, bathtub and unimodal shapes.

This model has been used to describe the air quality in Venice, Italy [2]. Tahai and Meyer [3] analysed journal citations of recent publications using the GG distribution to determine the management journals that have the greatest influence. The GG distribution has also been used to analyse the performance degradation of wireless communication systems and to obtain different techniques for processing synthetic aperture radar images [4]. This model can also be used in likelihood ratio tests as a discrimination method for its sub-models [5] as well as other applications [6–8].

The parameter estimation of the GG distribution has already been discussed. Stacy and Mihram [9] presented estimators for GG distribution parameters based on maximum likelihood (ML). Hager and Bain [10] showed that three nonlinear ML equations are unstable, i.e. the results may depend on the initial values. The method of moments has been discussed to obtain inferences for the parameters of GG distribution [11]. Khodabin and Ahmadabadi [12] compared the method of moments with ML and concluded that maximum likelihood estimators (MLEs) are better. However, Prentice [13] argued that even for sample sizes equal or larger than 400, the approximate normal distribution for $\hat{\phi}$ using the ML theory could not be achieved.

Considering a Bayesian approach, a prior distribution must be assigned. A naive procedure would consider flat priors (such as gamma or uniform distributions with large variances) for the parameters. However, Bernardo [14] argued that the use of simple proper flat priors presumed to be non-informative priors often hide important unwarranted assumptions which may easily dominate or invalidate the statistical analysis and should be strongly discouraged. Chang and Kim [15] derived non-informative priors for GG distribution assuming that ϕ is known. Maswadah et al. [16] also considered that ϕ is known and performed inference for the GG distribution based on order statistics. However, this parameter is non-trivial to be estimated or to be known in any application. An objective Jeffreys prior [17] was presented by Van Noortwijk [6] to estimate the quantiles of the flood of a given river.

In this study, we prove that Jeffreys prior leads to an improper posterior and should not be used. In the literature, there are other objective priors that could also be considered for the GG distribution parameters such as the Maximal Data Information (MDI) prior [18] and the reference prior [14,19–21]. A formal proof is also presented showing that such priors also lead to improper posteriors. This problem was overcome by proposing a hybrid Jeffreys-reference prior that yields a proper posterior distribution. The obtained prior arose considering the prior averaging approach discussed in [22] and is invariant under one-to-one transformations. This new result has a large number of applications, since it enables us to use the GG distribution in practice from the objective Bayesian point of view. A similar analysis was presented by Northrop and Attalides [23] for the generalized Pareto and the generalized extreme value distributions considering the Jeffreys prior and MDI prior. Additionally, as the Bayesian analysis was used improperly by Van Noortwijk [6], the data set related to the annual maximum discharges of the river Rhine at Lobith, Netherlands from 1901 to 1998 was reanalysed.

The remainder of this paper is organized as follows: Section 2 reviews the ML method for the GG distribution. Section 3 presents the Bayesian analysis considering non-informative priors. Section 4 describes a simulation study to compare both approaches. Section 5 presents an analysis of the data set related to the annual maximum discharges of the river Rhine at Lobith. Lastly, some final comments are made in Section 6.

2. Maximum likelihood estimators

Among the classical statistical inference methods, the ML method is usually preferred due to its better asymptotic properties. The MLEs are obtained by maximizing the likelihood function. Let $T_1, \ldots, T_n$

be a random sample where $T \sim GG(\alpha, \mu, \phi)$, the likelihood function for the parameter vector $\boldsymbol{\theta} = (\alpha, \mu, \phi)$ is given by

$$L(\boldsymbol{\theta}; t) = \frac{\alpha^n}{\Gamma(\phi)^n} \mu^{n\alpha\phi} \left\{ \prod_{i=1}^{n} t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha \right\}. \tag{3}$$

The ML estimates of the parameters are obtained by solving the likelihood equations $(\partial/\partial\alpha) \log(L(\boldsymbol{\theta}; t)) = 0$, $(\partial/\partial\mu) \log(L(\boldsymbol{\theta}; t)) = 0$, $(\partial/\partial\phi) \log(L(\boldsymbol{\theta}; t)) = 0$. Therefore, from Equation (3), we have

$$n\psi(\hat{\phi}) = n\hat{\alpha} \log(\hat{\mu}) + \hat{\alpha} \sum_{i=1}^{n} \log(t_i), \tag{4}$$

$$n\hat{\phi} = \hat{\mu}^{\hat{\alpha}} \sum_{i=1}^{n} t_i^{\hat{\alpha}}, \tag{5}$$

$$\frac{n}{\hat{\alpha}} + n\hat{\phi} \log(\mu) + \phi \sum_{i=1}^{n} \log(t_i) = \hat{\mu}^{\hat{\alpha}} \sum_{i=1}^{n} t_i^{\hat{\alpha}} \log(\hat{\mu} t_i), \tag{6}$$

where $\psi(k) = (\partial/\partial k) \log \Gamma(k) = \Gamma'(k)/\Gamma(k)$ is the digamma function. The solutions of Equations (4)–(6) provide the MLEs [9,10]. Numerical methods such as the Newton–Rapshon are required to find the solution of the nonlinear system.

3. Bayesian inference

In this section, we discuss non-informative priors for the parameters of the GG distribution. The proof of whether these priors lead to proper or improper posteriors are available in Appendix 3.

3.1. Jeffreys Prior

The well known non-informative prior introduced by Jeffreys [17] is obtained from the square root of the determinant of the Fisher information matrix $I(\boldsymbol{\theta})$. This prior has been widely used due to its one-to-one invariant property. For the generalized gamma distribution, the information matrix was computed by Harger and Bain [10] and is given by

$$I(\alpha, \mu, \phi) = \begin{bmatrix} \dfrac{1 + 2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2}{\alpha^2} & -\dfrac{1 + \phi\psi(\phi)}{\mu} & -\dfrac{\psi(\phi)}{\alpha} \\ -\dfrac{1 + \phi\psi(\phi)}{\mu} & \dfrac{\phi\alpha^2}{\mu^2} & \dfrac{\alpha}{\mu} \\ -\dfrac{\psi(\phi)}{\alpha} & \dfrac{\alpha}{\mu} & \psi'(\phi) \end{bmatrix}, \tag{7}$$

where $\psi'(k) = (\partial/\partial k)\psi(k)$ is the trigamma function. Van Noortwijk [6] computed the square root of the determinant of $I(\alpha, \mu, \phi)$ to obtain the Jeffreys prior

$$\pi_J(\alpha, \mu, \phi) \propto \frac{\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\mu}. \tag{8}$$

The joint posterior distribution for ϕ, μ and α using the Jeffreys prior is proportional to the product of the likelihood function (3) and prior (8) resulting in

$$p_J(\alpha, \mu, \phi | t) = \frac{1}{d_J(t)} \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \prod_{i=1}^{n} t_i^{\alpha\phi-1} \exp\left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha \right\}, \quad (9)$$

where

$$d_J(t) = \int_{\mathcal{A}} \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \prod_{i=1}^{n} t_i^{\alpha\phi-1} \exp\left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha \right\} d\theta \quad (10)$$

and $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ is the parameter space for θ. The same parameter space $\mathcal{A}$ was assumed throughout the paper.

Considering a Bayesian analysis under improper priors, it is important to check if these priors lead to improper posterior distributions.

Theorem 3.1: *The posterior (9) is improper, i.e., $d_J(t) = \infty$.*

Proof: See Appendix A.1 ∎

Since the posterior distribution (9) is improper, the results presented by Van Noortwijk [6] probably differ from the expected estimated values. Therefore, in Section 5 we reanalysed the data set related to the annual maximum discharges of the river Rhine at Lobith, Netherlands from 1901 to 1998.

3.2. *Maximal data information prior*

Zellner [18] introduced a procedure to derive a non-informative prior distribution. The aim was to maximize the information from the data in relation to the information known a priori about the parameters. The resulting non-informative prior distribution known as Maximal Data Information prior is defined as

$$\pi_Z(\alpha, \mu, \phi) \propto \exp(H(\alpha, \mu, \phi)), \quad (11)$$

where

$$H(\phi, \mu, \alpha) = \int f(t | \phi, \mu, \alpha) \log f(t | \phi, \mu, \alpha) \, dt \quad (12)$$

is the negative entropy of $f(t | \phi, \mu, \alpha)$, i.e., an information measure of f. From (1) and (12) we have

$$H(\alpha, \mu, \phi) = \int_0^\infty \log\left(\frac{\alpha}{\Gamma(\phi)} \mu^{\alpha\phi} t^{\alpha\phi-1} \exp(-(\mu t)^\alpha) \right) f(t | (\alpha, \mu, \phi) \, dt.$$

After some algebraic manipulations, $H(\alpha, \mu, \phi)$ can be written as

$$H(\alpha, \mu, \phi) = \log(\alpha) - \log(\Gamma(\phi)) + \phi\psi(\phi) + \log(\mu) - \frac{\psi(\phi)}{\alpha} - \phi.$$

Therefore, the MDI prior (11) for the GG distribution (2) is given by

$$\pi_Z(\alpha, \mu, \phi) \propto \frac{\alpha\mu}{\Gamma(\phi)} \exp\left\{ \psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi \right\}. \quad (13)$$

This MDI prior (13) coincides with the MDI prior (when $\alpha = 1$) obtained from the gamma distribution [24].

The joint posterior distribution for ϕ, μ and α obtained using the MDI prior is proportional to the product of the likelihood function (3) and the prior distribution (13) resulting in

$$p_Z(\alpha, \mu, \phi | t) = \frac{1}{d_Z(t)} \frac{\alpha^{n+1} \mu^{n\alpha\phi+1}}{\Gamma(\phi)^{n+1}} \prod_{i=1}^{n} t_i^{\alpha\phi-1} \exp\left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha + \psi(\phi)(\phi - \alpha^{-1}) - \phi \right\}, \quad (14)$$

where

$$d_Z(t) = \int_{\mathcal{A}} \frac{\alpha^{n+1} \mu^{n\alpha\phi+1}}{\Gamma(\phi)^{n+1}} \prod_{i=1}^{n} t_i^{\alpha\phi-1} \exp\left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha + \psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi \right\} d\boldsymbol{\theta}. \quad (15)$$

Theorem 3.2: *The joint posterior density (14) is improper, i.e., $d_Z(t) = \infty$.*

Proof: See Appendix A.2 ∎

3.3. Reference prior

Another well-known class of non-informative priors are reference priors proposed by Bernardo [19] with further developments [14,20,21]. The reference prior is invariant under one-to-one transformation in the parameters and is defined as the prior $\pi(\boldsymbol{\theta})$ that maximizes the expected Kullback–Leibler distance between the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$ and the prior distribution $\pi(\boldsymbol{\theta})$ based on the experimental data. There are many ways to obtain reference priors. The following proposition is useful to obtain these priors for GG distribution.

Proposition 3.3 (Bernardo [14, p. 40, Theorem 14]): *Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ be a vector with the ordered parameters of interest and the data t consist of a random sample of size n from a statistical model $f(t|\boldsymbol{\theta})$, and let $\mathcal{P}$ be the class of all continuous priors with support $\mathcal{A}$. If the posterior distribution of $\boldsymbol{\theta}$ is asymptotically normal with dispersion matrix $V(\hat{\boldsymbol{\theta}}_n)/n$, where $\hat{\boldsymbol{\theta}}_n$ is a consistent estimator of $\boldsymbol{\theta}$, $H(\boldsymbol{\theta}) = V^{-1}(\boldsymbol{\theta})$, V_j is the upper $j \times j$ submatrix of V, $H_j = V_j$ and $h_{j,j}(\boldsymbol{\theta})$ is the lower right element of H_j. Then, if the parameter space of θ_j is independent of $\boldsymbol{\theta}_{-j} = (\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_m)$, for $j = 1, \ldots, m$, and $h_{j,j}(\boldsymbol{\theta})$ are factorized in the form*

$$h_{j,j}^{1/2}(\boldsymbol{\theta}) = f_j(\theta_j)g_j(\boldsymbol{\theta}_{-j}), \quad j = 1, \ldots, m.$$

The reference prior for the ordered parameters $\boldsymbol{\theta}$ is given by

$$\pi_R(\boldsymbol{\theta}) = \pi_R(\boldsymbol{\theta}|\mathcal{P}) = \pi(\theta_j|\theta_1, \ldots, \theta_{j-1}) \times \cdots \times \pi(\theta_2|\theta_1)\pi(\theta_1),$$

where $\pi(\theta_j|\theta_1, \ldots, \theta_{j-1}) = f_j(\theta_j)$, for $j = 1, \ldots, m$ and there is no need for compact approximations, even if the conditional priors are not proper. In other words, the $\boldsymbol{\theta}$-reference prior is given by $\pi_R(\boldsymbol{\theta}) = \prod_{i=1}^{m} f_j(\theta_j)$.

Theorem 3.4: *Let $\boldsymbol{\theta} = (\mu, \phi, \alpha)$ be the ordered parameters of interest for GG distribution, then the $\boldsymbol{\theta}$-reference prior is given by*

$$\pi_R(\mu, \phi, \alpha) \propto \frac{\pi_R(\phi)}{\alpha\mu}, \quad (16)$$

where

$$\pi_R(\phi) \propto \sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}. \quad (17)$$

Proof: See Appendix A.3. ∎

The joint posterior distribution for ϕ, μ and α using the prior distribution (16) is given by

$$p_R(\alpha, \mu, \phi | \boldsymbol{t}) = \frac{1}{d_R(\boldsymbol{t})} \frac{\alpha^{n-1} \pi_R(\phi)}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^{n} t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha \right\}, \qquad (18)$$

where

$$d_R(\boldsymbol{t}) = \int_{\mathcal{A}} \frac{\alpha^{n-1} \pi_R(\phi)}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^{n} t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha \right\} d\boldsymbol{\theta}. \qquad (19)$$

Theorem 3.5: *The posterior density (18) is improper, i.e., $d_R(\boldsymbol{t}) = \infty$.*

Proof: See Appendix A.4. ∎

3.4. Jeffreys/reference prior

Berger et al. [22] suggested initiating with a collection of objective priors and then taking the arithmetic mean or the geometric mean. Furthermore, the authors argued that '... the weights in arithmetic averaging of improper priors are rather arbitrary because the priors have no normalizing constants, whereas geometric averaging is unaffected by normalizing constants'. Thus, considering the geometric mean between the Jeffreys prior (8) and the reference prior (16), the Jeffreys/Reference prior is given by

$$\pi_{JR}(\alpha, \mu, \phi) \propto \frac{\pi_{JR}(\phi)}{\mu \sqrt{\alpha}}, \qquad (20)$$

where

$$\pi_{JR}(\phi) \propto \sqrt[4]{\phi^2 \psi'(\phi)^3 - \psi'(\phi)^2 - \psi'(\phi) - \frac{\psi(\phi)^2(\phi^2\psi'(\phi)^2 - \psi'(\phi) - 1)}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}. \qquad (21)$$

The obtained prior is non-informative as it comes from a geometric mean of the two most used non-informative priors [22]. Moreover, the hybrid Jeffreys-reference prior was constructed as a geometric mean of one-to-one invariant priors, and therefore this prior also has an invariance property under one-to-one transformations.

The joint posterior distribution for ϕ, μ and α using the prior distribution (20) is given by

$$p_{JR}(\alpha, \mu, \phi | \boldsymbol{t}) = \frac{\pi_{JR}(\phi)}{d_{JR}(\boldsymbol{t})} \frac{\alpha^{n-\frac{1}{2}}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^{n} t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha \right\} \qquad (22)$$

where

$$d_{JR}(\boldsymbol{t}) = \int_{\mathcal{A}} \frac{\alpha^{n-1/2} \pi_{JR}(\phi)}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^{n} t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^{n} t_i^\alpha \right\} d\boldsymbol{\theta}. \qquad (23)$$

Theorem 3.6: *The posterior density (18) is proper, i.e., $d_{JR}(\boldsymbol{t}) < \infty$.*

Proof: See Appendix A.5. ∎

The full conditional posterior distributions for ϕ, μ and α are given as follows:

$$p_M(\alpha|\phi, \boldsymbol{t}) \propto \alpha^{n-3/2} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi},$$

$$p_M(\phi|\alpha, \boldsymbol{t}) \propto \pi_{JR}(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi}, \qquad (24)$$

$$p_M(\mu|\phi, \alpha, \boldsymbol{t}) \sim \mathrm{GG}\left(n\phi, \left(\sum_{i=1}^n t_i^\alpha \right)^{1/\alpha}, \alpha \right).$$

These conditional distributions are useful during the use of Monte Carlo Markov Chain (MCMC) methods to simulate samples of parameters of the joint posterior distribution.

4. Simulation analysis

In this section, a simulation study using Monte Carlo methods is presented to compare the efficiency of ML method with our proposed Bayesian approach by computing the Bias and the root-mean-square error (RMSE), given by

$$\mathrm{Bias}_i = \sum_{j=1}^N \frac{\hat{\theta}_{i,j}}{N} - \theta_i, \quad \mathrm{RMSE}_i = \sqrt{\sum_{j=1}^N \frac{(\hat{\theta}_{i,j} - \theta_i)^2}{N}}, \quad \text{for } i = 1, 2, 3, \qquad (25)$$

where $N = 10{,}000$ is the number of estimates obtained throughout the MLE and the posterior modes. The 95% coverage probability of the asymptotic confidence intervals and the Credible Intervals ($\mathrm{CI}_{95\%}$) were also evaluated. Considering this approach, the best estimators will show both Bias and RMSE closer to zero. In addition, for a large number of experiments considering a 95% confidence level, the frequencies of intervals that covered the true values of $\boldsymbol{\theta}$ should be closer to 95%.

To find the ML estimators, the Newton–Raphson method was adopted. In this case, the initial values to start the iterative procedure must be assigned. To ensure a fair comparison, both procedures were under the same conditions (same initial values and samples). The initial values considered were the same values used to generate the samples.

Clearly, the normalizing constant for the marginal posterior densities require two-dimensional integration. Therefore, the MCMC method was considered to obtain the posterior estimates. Since the conditional distributions of α and ϕ were not easily identified, the Metropolis–Hastings algorithm [25] was considered to simulate the posterior quantities. For each simulated data set, 15,500 iterations were performed using MCMC methods. As a burn-in, the first 1000 initial values were discarded, the considered thin was 30 to reduce the correlation among the chains. The Geweke criterion [26] was used to check the convergence of the obtained chains under a 95% confidence level. These values were used to compute the posterior mode estimates, yielding 10,000 estimates for ϕ, μ and α.

The chosen values to perform this procedure were $\boldsymbol{\theta} = ((0.5, 0.5, 3), (2, 1, 0.5), (4, 2, 2), (0.4, 1.5, 5))$ and $n = (50, 100, 200)$. The seed used to generate the random values in the R software was 2016. Table 1 presents the Bias and the RMSE of the estimates obtained through the MLE and the Bayes estimators for 10,000 simulated samples under different values of $\boldsymbol{\theta}$ and n. Table 2 shows the coverage probability with a 95% confidence level.

The Bayes estimators returned estimates with smaller Bias and RMSE than the MLEs, specially for small and moderate sample sizes. For large samples, both estimators returned similar values, i.e. as there is an increase in n, both methodologies behave similarly. Both bias and RMSE have shown to

Table 1. Bias (RMSE) of the ML estimates and the Bayes estimators (posterior mode) for 10,000 samples of sizes $n = (50, 100, 200)$ and different values of θ.

θ	Classical Inference			Bayesian Inference		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0.5$	0.179(0.73)	0.046(0.26)	0.024(0.18)	−0.015(0.23)	0.001(0.19)	0.014(0.16)
$\mu = 0.5$	0.156(0.91)	0.026(0.11)	0.014(0.07)	−0.002(0.08)	0.001(0.07)	0.006(0.05)
$\alpha = 3$	0.487(1.81)	0.240(1.04)	0.145(0.74)	−0.403(0.97)	−0.248(0.79)	−0.209(0.63)
$\phi = 2$	−0.329(1.08)	−0.102(0.93)	−0.115(0.77)	−0.508(0.62)	−0.385(0.53)	0.037(0.63)
$\mu = 1$	0.870(3.71)	1.166(3.85)	0.705(2.96)	−0.431(0.47)	−0.117(0.27)	0.818(2.64)
$\alpha = 0.5$	0.255(0.68)	0.096(0.22)	0.064(0.15)	0.220(0.27)	0.129(0.16)	0.105(0.15)
$\phi = 4$	1.616(4.92)	1.115(4.03)	0.335(2.95)	−1.811(2.07)	−1.621(1.94)	−0.755(1.46)
$\mu = 2$	1.890(4.34)	1.282(3.34)	0.557(2.23)	−0.791(0.86)	−0.562(0.75)	−0.114(0.76)
$\alpha = 2$	0.531(1.64)	0.306(1.08)	0.279(0.79)	0.993(1.33)	0.648(0.96)	0.496(0.75)
$\phi = 0.4$	0.234(0.68)	0.128(0.25)	0.035(0.14)	0.042(0.21)	0.078(0.17)	0.026(0.13)
$\mu = 1.5$	0.249(0.98)	0.105(0.21)	0.032(0.11)	0.031(0.17)	0.055(0.14)	0.022(0.10)
$\alpha = 5$	−0.013(2.07)	−0.468(1.12)	0.088(1.13)	0.502(1.89)	−0.317(1.04)	0.050(1.12)

Table 2. Coverage probability with a 95% confidence level equals the ML estimates and the Bayes estimators (posterior mode) considering 10,000 samples of sizes $n = (50, 100, 200)$ and different values of θ.

θ	Classical	Inference		Bayesian	Inference	
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0.5$	89.30%	92.36%	92.39%	98.68%	97.91%	95.81%
$\mu = 0.5$	90.03%	92.31%	92.53%	99.34%	98.95%	97.61%
$\alpha = 3$	93.89%	96.01%	95.44%	98.64%	97.74%	95.62%
$\phi = 2$	75.62%	84.34%	85.74%	92.78%	94.02%	89.66%
$\mu = 1$	61.48%	71.16%	72.63%	81.16%	87.55%	89.38%
$\alpha = 0.5$	100.00%	99.01%	97.92%	91.97%	93.21%	89.35%
$\phi = 4$	80.92%	82.53%	81.37%	92.25%	92.60%	87.87%
$\mu = 2$	77.12%	78.90%	77.65%	91.79%	92.37%	88.04%
$\alpha = 2$	100.00%	98.82%	97.02%	92.06%	92.38%	87.86%
$\phi = 0.4$	97.66%	99.97%	95.76%	97.40%	95.72%	95.66%
$\mu = 1.5$	97.27%	99.83%	96.09%	98.50%	97.98%	97.85%
$\alpha = 5$	90.11%	91.88%	94.76%	97.45%	95.94%	95.82%

be consistent and asymptotically unbiased for the parameters. However, the CIs of the MLEs using the asymptotic method does not have good coverage probabilities. These results correlate with Prentice [13], i.e. even for large sample sizes, the approximate normal distribution for the parameters using the ML theory could not be achieved. On the other hand, the credible interval based on the Bayes estimators provided excellent coverage probabilities even for small sample sizes. For these reasons, our Bayes estimators (24) should be considered to achieve the parameter estimators of GG distribution.

5. Real data application

As the Bayesian analysis was used improperly by Van Noortwijk [6], the data set related to the annual maximum discharges of the river Rhine at Lobith, Netherlands from 1901 to 1998 was reanalysed. reanalysed. The results presented by Van Noortwijk [6, table 1, pg 6] can be seen in Table 3.

From the credibility intervals of ϕ and α available in Table 3, there is a good indication that the

Table 3. Posterior mean and 95% credibility intervals for ϕ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901–1998.

θ	Mean	$CI_{95\%}(\theta)$
ϕ	1.380	(0.01; 6.00)
$1/\mu$	4936	(⋯ ; ⋯)
α	2.310	(0.01; 6.00)

⋯ Not presented.

numerical techniques did not provide good results. These large credibility intervals were probably influenced due to the improper posterior distribution. In this study, considering a proper posterior distribution, the GG distribution can be used to analyse these data under the same assumptions as Section 4. The posterior summaries obtained using the MCMC methods are given in Table 4.

Considering $CI_{95\%}(\theta)$, there is an indication that $\alpha = 1$ or $\phi = 1$, i.e. the GG distribution may reduce to gamma or the Weibull distribution. The obtained results were compared with the sub-models such as Weibull, gamma and lognormal distributions using Akaike information criterion ($AIC = -2l(\hat{\theta}; x) + 2k$), corrected Akaike information criterion ($AICc = AIC + 2k(k+1)(n-k-1)^{-1}$) and Bayesian information criterion ($BIC = -2l(\hat{\theta}; x) + k\log(n)$), where k is the number of parameters to be fitted and $\hat{\theta}$ is the estimate of θ. The best model is the one which provides the minimum values of these criteria.

Considering any criteria, it can be concluded from the results in Table 5 that among the chosen models, the gamma distribution fit best considering the annual maximum discharges of the river Rhine at Lobith from 1901 to 1998. Moreover, to verify the goodness of fit, Figure 1 shows the survival function adjusted for different distributions of overlapping probability in the empirical function.

Table 4. Posterior mode, standard deviations and 95% credible intervals for ϕ, μ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901–1998.

θ	Mode	SD	$CI_{95\%}(\theta)$
ϕ	2.7704	1.5599	(0.9874; 7.0422)
$1/\mu$	4204.4	1987.0	(1242.9; 8606.8)
α	1.8766	0.584	(1.1583; 3.3692)

Table 5. Results of AIC, AICc and BIC criteria for different probability distributions considering the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901–1998.

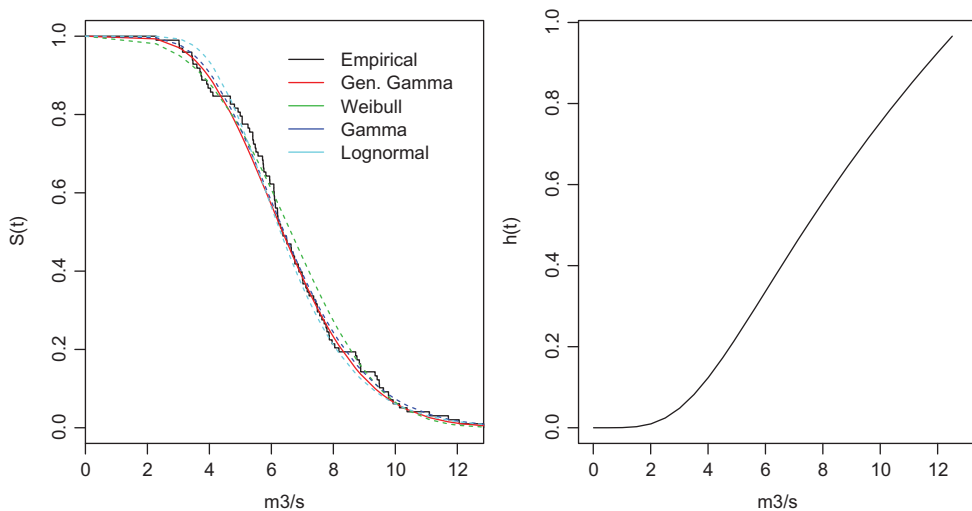| Criteria	G. Gamma	Weibull	Gamma	Lognormal
AIC	429.33	430.11	**426.52**	429.81
AICc	429.59	430.23	**426.65**	429.93
BIC	437.09	435.28	**431.69**	434.98

Figure 1. Survival function fitted by the empirical and by different probability distributions considering the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998 and the hazard function fitted by a GG distribution.

Table 6. Posterior mean and 90% credibility intervals for ϕ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901–1998.

Distribution	River discharge	Cl$_{90\%}$
G. gamma (Van Noortwijk)	15,150	(12,950; 16,950)
G. gamma (Our Approach)	15,448	(10,556; 22,444)
G. gamma (Classical Inference)	14,780	(12,699; 16,542)
Two-parameter gamma	15,598	(14,370; 17,583)

Van Noortwijk [6] argued that '... the Dutch river dikes have to withstand water levels and discharges with an average return period of up to 1250 years, where a downstream water level can be determined on the basis of the upstream discharge by using a river flow simulation model'. The main aim was to find the annual maximum river discharge in which the probability of exceedance is 1/1250 per year. Table 6 presents the discharge with a probability of exceedance of 1/1250 and the 90% uncertainty interval for the GG distribution (Van Noortwijk, our results and the classical inference) and the two-parameter gamma distribution. To evaluate the Bayes estimators of the two-parameter gamma distribution, we considered a posterior distribution obtained with the reference prior (see Berger et al. [22], p. 199, eq. 14).

The improper posterior produced an underestimated annual maximum discharge. The difference between the Van Noortwijk estimate and ours was 420 m^3/s. Hence, the Dutch river dikes will have to withstand water levels and discharges of up to 1570 m^3/s. The ML estimators of the GG distribution also returned an underestimated value for the River maximum discharge. On the other hand, the results obtained from the gamma distribution are similar to those obtained from the GG distribution in our approach. Our results clearly showed that the gamma distribution should be used to estimate the annual maximum discharges of the river Rhine at Lobith.

6. Discussion

In this paper, we introduced a Bayesian analysis for the GG distribution considering non-informative priors. An interesting aspect of our findings is that the Jeffreys, MDI and the reference priors for GG parameters failed to yield a proper posterior distribution and should not be used in the Bayesian analysis.

To overcome this problem, an alternative prior based on Jeffreys general rule and the reference prior was proposed considering the prior averaging approach discussed in [22]. The obtained prior is invariant under one-to-one transformations and yields a proper posterior distribution. The simulation study showed that the Bayes estimators returned estimates with smaller Bias and RMSE than the MLEs, specially for small and moderate sample sizes. Additionally, the credible interval based on the Bayes estimators provided very good coverage probabilities even for small sample sizes. On the other hand, even for large sample sizes, the approximate normal distribution for the parameters using the ML theory could not be achieved.

As the Bayesian analysis was used improperly by Van Noortwijk [6], the data set related to the annual maximum discharges of the river Rhine at Lobith, Netherlands from 1901 to 1998 was reanalysed and its annual maximum river discharge with 1/1250 probability of exceedance was computed properly.

In conclusion, our Bayes estimators should be considered to achieve the GG distribution parameter estimators. One possible extension of our work is to consider an objective Bayesian analysis in regression modelling where the responses follow GG distribution. Another extension is to explore other formal rules to construct priors based on probability matching priors [27].

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Pedro L. Ramos   http://orcid.org/0000-0002-5387-2457

References

[1] Stacy EW. A generalization of the gamma distribution. Ann Math Stat. 1962;33(3):1187–1192.
[2] Marani A, Lavagnini I, Buttazzoni C. Statistical study of air pollutant concentrations via generalized gamma distributions. J Air Pollut Control Assoc. 1986;36(11):1250–1254.
[3] Tahai A, Meyer MJ. A revealed preference study of management journals' direct influences. Strateg Manage J. 1999;20(3):279–296.
[4] Li HC, Hong W, Wu YR, Fan PZ. On the empirical-statistical modeling of sar images with generalized gamma distribution. IEEE J Sel Top Signal Process. 2011;5(3):386–397.
[5] Lawless JF. Statistical models and methods for lifetime data. Vol. 362. Floboken, New Jersey: Wiley; 2011.
[6] Van Noortwijk JM. Bayes estimates of flood quantiles using the generalised gamma distribution. Syst Bayesian Reliab. 2001;5:351–374.
[7] Ramos PL, Louzada F, Ramos E. An efficient, closed-form map estimator for nakagami-m fading parameter. IEEE Commun Lett. 2016;20(11):2328–2331.
[8] Cox C, Chu H, Schneider MF. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. Stat Med. 2007;26(23):4352–4374.
[9] Stacy EW, Mihram GA. Parameter estimation for a generalized gamma distribution. Technometrics. 1965;7(3):349–358.
[10] Hager HW, Bain LJ. Inferential procedures for the generalized gamma distribution. J Am Stat Assoc. 1970;65(332):1601–1609.
[11] Huang PH, Hwang TY. On new moment estimation of parameters of the generalized gamma distribution using it's characterization. Taiwanese J Math. 2006;10(4):pp 1083.
[12] Khodabin M, Ahmadabadi A. Some properties of generalized gamma distribution. Math Sci. 2010;4(1):9–28.
[13] Prentice RL. A log gamma model and its maximum likelihood estimation. Biometrika. 1974;61(3):539–544.
[14] Bernardo JM. Reference analysis. Handbook Stat. 2005;25:17–90.
[15] Chang IH, Kim BH. Non-informative priors in the generalized gamma stress–strength systems. IIE Trans. 2011;43(11):797–804.
[16] Maswadah M, Seham AM, Ahsanullah M. Bayesian inference on the generalized gamma distribution based on generalized order statistics. J Statist Theory Appl. 2013;12(4):356–377.
[17] Jeffreys H. Theory of probability. Oxford: Oxford University Press; 1939.
[18] Zellner A. Maximal data information prior distributions. In: Aykac A, Brumat C, editors. New Methods in the applications of Bayesian Methods. Amsterdam: North-Holland; 1977. p. 211–232.
[19] Bernardo JM. Reference posterior distributions for Bayesian inference. J R Statist Soc Ser B (Methodol). 1979;41:113–147.
[20] Berger JO, Bernardo JM. Estimating a product of means: Bayesian analysis with reference priors. J Am Stat Assoc. 1989;84(405):200–207.
[21] Berger JO, Bernardo JM. Ordered group reference priors with application to the multinomial problem. Biometrika. 1992;79(1):25–37.
[22] Berger JO, Bernardo JM, Sun D, et al. Overall objective priors. Bayesian Anal. 2015;10(1):189–221.
[23] Northrop P, Attalides N. Posterior propriety in Bayesian extreme value analyses using reference priors. Stat Sin. 2016;26(2).
[24] Moala FA, Ramos PL, Achcar JA. Bayesian inference for two-parameter gamma distribution assuming different noninformative priors. Revista Colombiana de Estadística. 2013;36(2):321–338.
[25] Gamerman D, Lopes HF. Markov chain monte carlo: stochastic simulation for Bayesian inference. Boca Raton, Florida: CRC Press; 2006.
[26] Geweke J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo JM, Berger J, Dawid AP, Smith JFM, editors. Bayesian Statistics 4. Oxford: Oxford University Press; 1992. p. 169–193.
[27] Tibshirani R. Noninformative priors for one parameter of many. Biometrika. 1989;76(3):604–608.
[28] Abramowitz M, Stegun IA. Handbook of mathematical functions. 10th ed. Washington, DC: NBS; 1972.
[29] Folland GB. Real analysis: modern techniques and their applications. 2nd ed. New York: Wiley; 1999.

Appendix 1. Definitions and auxiliary propositions

The following propositions are useful to prove the results related to the posterior distribution. Let $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ denote the *extended real number line* with the usual order ($\geq$), $\mathbb{R}^+$ are strictly positive numbers and $\mathbb{R}_0^+$ denote the positive numbers including 0.

Definition A.1: Let $g : \mathcal{U} \to \bar{\mathbb{R}}_0^+$ and $h : \mathcal{U} \to \bar{\mathbb{R}}_0^+$, where $\mathcal{U} \subset \mathbb{R}$ and $m \in \mathbb{N}$. We say that $g(x) \propto h(x)$ if there exists $c_0 \in \mathbb{R}^+$ and $c_1 \in \mathbb{R}^+$ such that $c_0\, h(x) \leq g(x) \leq c_1\, h(x)$ for every $x \in \mathcal{U}$.

Definition A.2: Let $a \in \bar{\mathbb{R}}$, $g : \mathcal{U} \to \mathbb{R}^+$ and $h : \mathcal{U} \to \mathbb{R}^+$, where $\mathcal{U} \subset \mathbb{R}$. We say that $g(x) \underset{x \to a}{\propto} h(x)$ if

$$\liminf_{x \to a} \frac{g(x)}{h(x)} > 0 \quad \text{and} \quad \limsup_{x \to a} \frac{g(x)}{h(x)} < \infty.$$

The meaning of the relations $g(x) \propto_{x \to a^+} h(x)$ and $g(x) \propto_{x \to a^-} h(x)$ for $a \in \mathbb{R}$ are defined analogously. Note that, if for some $c \in \mathbb{R}^+$ we have $\lim_{x \to a}(g(x)/h(x)) = c$, then $g(x) \propto_{x \to a} h(x)$.

The following proposition is a direct consequence of the above definition.

Proposition A.3: *For $a \in \bar{\mathbb{R}}$ and $r \in \mathbb{R}$, let $f_1(x) \propto_{x \to a} f_2(x)$ and $g_1(x) \propto_{x \to a} g_2(x)$ then the following hold*

$$f_1(x)g_1(x) \underset{x \to a}{\propto} f_2(x)g_2(x) \quad \text{and} \quad f_1(x)^r \underset{x \to a}{\propto} f_2(x)^r.$$

The following proposition relates Definition A.1 and Definition A.2.

Proposition A.4: *Let $g : (a, b) \to \mathbb{R}^+$ and $h : (a, b) \to \mathbb{R}^+$ be continuous functions on $(a, b) \subset \mathbb{R}$, where $a \in \bar{\mathbb{R}}$ and $b \in \bar{\mathbb{R}}$. Then $g(x) \propto h(x)$ if and only if $g(x) \propto_{x \to a} h(x)$ and $g(x) \propto_{x \to b} h(x)$.*

Note that if $g : (a, b) \to \mathbb{R}^+$ and $h : (a, b) \to \mathbb{R}^+$ are continuous functions on $(a, b) \subset \mathbb{R}$, then by continuity it follows that $\lim_{x \to c}(g(x)/h(x)) = g(c)/h(c) > 0$. Therefore $g(x) \propto_{x \to c} h(x)$ for every $c \in (a, b)$. This fact and the Proposition A.4 imply directly the following.

Proposition A.5: *Let $g : (a, b) \to \mathbb{R}^+$ and $h : (a, b) \to \mathbb{R}^+$ be continuous functions in $(a, b) \subset \mathbb{R}$, where $a \in \bar{\mathbb{R}}$ and $b \in \bar{\mathbb{R}}$, and let $c \in (a, b)$. Then if $g(x) \propto_{x \to a} h(x)$ or $g(x) \propto_{x \to b} h(x)$ we have respectively that*

$$\int_a^c g(t)\, dt \propto \int_a^c h(t)\, dt \quad \text{or} \quad \int_c^b g(t)\, dt \propto \int_c^b h(t)\, dt.$$

Appendix 2. Useful proportionalities

The following proportionalities are also useful to prove results related to the posterior distribution.

Proposition A.6: *The following results hold*

$$\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} \underset{\phi \to 0^+}{\propto} 1 \quad \text{and} \quad \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} \underset{\phi \to \infty}{\propto} \frac{1}{\phi}.$$

Proof: Let us present the proof for the first case. Considering the recurrence relation $\psi'(\phi) = \frac{1}{\phi^2} + \psi'(\phi + 1)$, we have

$$\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} = \sqrt{\frac{1}{\phi^2} + 2\psi'(\phi + 1) + \phi^2 \psi'(\phi + 1)^2 - \frac{1}{\phi^2} - \psi'(\phi) - 1}$$

$$= \sqrt{\psi'(\phi + 1) - 1 + \phi^2 \psi'(\phi + 1)^2}.$$

Therefore

$$\lim_{\phi \to 0^+} \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} = \sqrt{\psi'(1) - 1} = \sqrt{\frac{\pi^2}{6} - 1} > 0. \tag{A1}$$

For the second case, note that [28, p. 260]

$$\psi'(\phi) = \frac{1}{\phi} + \frac{1}{2\phi^2} + \frac{1}{6\phi^3} + o\left(\frac{1}{\phi^3}\right).$$

Hence, it follows that

$$\phi^2 \left(\frac{1}{\phi} + \frac{1}{2\phi^2} + \frac{1}{6\phi^3} + o\left(\frac{1}{\phi^3}\right) \right)^2 - \frac{1}{\phi} - \frac{1}{2\phi^2} - \frac{1}{6\phi^3} - o\left(\frac{1}{\phi^3}\right) - 1$$

$$= \phi^2 \left(\frac{1}{\phi^2} + \frac{1}{\phi^3} + \frac{7}{12\phi^4} + o\left(\frac{1}{\phi^4}\right) \right) - \frac{1}{\phi} - \frac{1}{2\phi^2} + o\left(\frac{1}{\phi^2}\right) - 1$$

$$= 1 + \frac{1}{\phi} + \frac{7}{12\phi^2} + o\left(\frac{1}{\phi^2}\right) - \frac{1}{\phi} - \frac{1}{2\phi^2} + o\left(\frac{1}{\phi^2}\right) - 1 = \frac{1}{12\phi^2} + o\left(\frac{1}{\phi^2}\right).$$

Therefore,

$$\lim_{\phi \to \infty} \frac{\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\phi^{-1}} = \lim_{\phi \to \infty} \sqrt{\frac{1}{12} + o(1)} = \sqrt{\frac{1}{12}}.$$

∎

Proposition A.7: *The following results hold:*

$$\sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}} \underset{\phi \to 0^+}{\propto} \frac{1}{\sqrt{\phi}} \quad \text{and} \quad \sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}} \underset{\phi \to \infty}{\propto} \frac{1}{\phi}.$$

Proof: By the recurrence relations

$$\psi(\phi) = -\frac{1}{\phi} + \psi(\phi + 1) \quad \text{and} \quad \psi'(\phi) = \frac{1}{\phi^2} + \psi'(\phi + 1).$$

It follows that

$$2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1$$

$$= 2\left(-\frac{1}{\phi} + \psi(\phi + 1)\right) + \phi\left(\frac{1}{\phi^2} + \psi'(\phi + 1)\right) + \phi\left(\frac{1}{\phi^2} - \frac{2}{\phi}\psi(\phi + 1) + \psi(\phi + 1)^2\right) + 1$$

$$= 1 + \phi\left(\psi(\phi + 1)^2 + \psi'(\phi + 1)\right).$$

Therefore,

$$\sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}$$

$$= \sqrt{\frac{(\phi^{-2} + \psi'(\phi + 1))(1 + \phi(\psi(\phi + 1)^2 + \psi'(\phi + 1))) - (\phi^{-1} + \psi(\phi + 1))^2}{1 + \phi(\psi(\phi + 1)^2 + \psi'(\phi + 1))}}$$

$$= \sqrt{\frac{1}{\phi}} \sqrt{\frac{2\psi(\phi + 1) + \psi(\phi + 1)^2 + \psi'(\phi + 1) + \phi L(\phi)}{1 + \phi(\psi(\phi + 1)^2 + \psi'(\phi + 1))}},$$

where

$$L(\phi) = (\psi'(\phi + 1) - \psi(\phi + 1)^2) + \phi^2 \psi'(\phi + 1)(\psi(\phi + 1)^2 + \psi'(\phi + 1)).$$

Since $\lim_{\phi \to 0^+} L(\phi) < \infty$, we have

$$\lim_{\phi \to 0} \frac{\sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}}{\frac{1}{\sqrt{\phi}}} = \sqrt{2\psi'(1) + \psi(1)^2 + \psi'(1)}$$

$$= \sqrt{\frac{\pi^2}{6} - \gamma^2 - 2\gamma} > 0,$$

which proves the first proportionality. Now, by Abramowitz [28], we have

$$\psi(\phi) = \log(\phi) - \frac{1}{2\phi} - \frac{1}{12\phi^2} + o\left(\frac{1}{\phi^2}\right) \quad \text{and} \quad \psi'(\phi) = \frac{1}{\phi} + \frac{1}{2\phi^2} + o\left(\frac{1}{\phi^2}\right),$$

where it follows directly that

$$\psi(\phi)^2 = \log(\phi)^2 - \frac{\log(\phi)}{\phi} + o\left(\frac{1}{\phi}\right).$$

Therefore

$$2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1 = \phi\log(\phi)^2 + \log(\phi) + 2 + o(1),$$

and

$$\sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}$$

$$= \sqrt{\frac{(\frac{1}{\phi} + \frac{1}{2\phi^2} + o(\frac{1}{\phi^2}))(\phi\log(\phi)^2 + \log(\phi) + 2 + o(1)) - \log(\phi)^2 + \frac{\log(\phi)}{\phi} + o(\frac{1}{\phi})}{\phi\log(\phi)^2 + \log(\phi) + 2 + o(1)}}$$

$$= \sqrt{\frac{\frac{1}{\phi}(\log(\phi)^2 + o(\log(\phi)^2))}{\phi(\log(\phi)^2 + o(\log(\phi)^2))}} = \frac{1}{\phi}\sqrt{\frac{1 + o(1)}{1 + o(1)}}.$$

Thus

$$\lim_{\phi \to \infty} = \frac{\sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}}{\frac{1}{\phi}} = 1$$

which proves the second proportionality. ∎

Proposition A.8: *The following results hold:*

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \underset{\phi \to 0^+}{\propto} \phi^{n-1} \quad \text{and} \quad \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \underset{\phi \to \infty}{\propto} n^{n\phi}\phi^{(n-1)/2}.$$

Proof: Considering the recurrence relation $\Gamma(z) = (1/z)\Gamma(z+1)$ it follows that $\lim_{z \to 0^+}(\Gamma(z)/(1/z)) = \Gamma(1) = 1$. Therefore

$$\Gamma(z) \underset{\phi \to 0^+}{\propto} \frac{1}{z} \tag{A2}$$

and

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \underset{\phi \to 0^+}{\propto} \frac{\frac{1}{n\phi}}{\frac{1}{\phi^n}} \underset{\phi \to 0^+}{\propto} \phi^{n-1}. \tag{A3}$$

Now, considering Stirling's approximation for gamma function

$$\lim_{z \to \infty} \frac{\Gamma(z)}{\sqrt{2\pi}\, z^{z-1/2}\, e^{-z}} = 1 \Rightarrow \lim_{z \to \infty} \frac{\Gamma(z)}{z^{z-\frac{1}{2}}\, e^{-z}} = \sqrt{2\pi} \Rightarrow \Gamma(z) \underset{z \to \infty}{\propto} z^{z-1/2}\, e^{-z}.$$

Then, by Proposition A.3 we have

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \underset{\phi \to \infty}{\propto} \frac{(n\phi)^{n\phi - 1/2}\, e^{-n\phi}}{(\phi^{\phi - 1/2}\, e^{-\phi})^n} = n^{-1/2} n^{n\phi}\phi^{(n-1)/2} \underset{\phi \to \infty}{\propto} n^{n\phi}\phi^{(n-1)/2}.$$

∎

Proposition A.9: *Let* $p(\alpha) = \log((1/n)\sum_{i=1}^n t_i^\alpha / \sqrt[n]{\prod_{i=1}^n t_i^\alpha})$, *for* $t_1, t_2, \ldots, t_n$ *positive and not all equal, then* $p(\alpha) > 0$ *and the following hold*

$$p(\alpha) \underset{\alpha \to 0^+}{\propto} \alpha^2 \quad \text{and} \quad p(\alpha) \underset{\alpha \to \infty}{\propto} \alpha.$$

Proof: Note that $(1/n)\sum_{i=1}^n t_i^\alpha / \sqrt[n]{\prod_{i=1}^n t_i^\alpha} > 1 \Rightarrow p(\alpha) > 0$ by the arithmetic-geometric inequality.

Now, let $u_i = t_i/\sqrt[n]{\prod_{i=1}^n t_i}$, $i = 1, \ldots, n$ and $u_m = \max\{u_1, \ldots, u_n\}$. Since $t_1, \ldots, t_n$ are not all equal then $u_m > 1$ and

$$\lim_{\alpha \to \infty} \frac{p(\alpha)}{\alpha} = \lim_{\alpha \to \infty} \frac{1}{\alpha} \log \left(\sum_{i=1}^n u_i^\alpha \right) - \frac{\log(n)}{\alpha} = \lim_{\alpha \to \infty} \log \left\{ \left(\sum_{i=1}^n u_i^\alpha \right)^{1/\alpha} \right\}$$

$$= \lim_{\alpha \to \infty} \log \left(u_m \left(\sum_{i=1}^n (u_i/u_m)^\alpha \right)^{1/\alpha} \right).$$

Moreover, $1 \leq \sum_{i=1}^n (u_i/u_m)^\alpha \leq n$, then $1 \leq (\sum_{i=1}^n (u_i/u_m)^\alpha)^{1/\alpha} \leq n^{1/\alpha}$, which implies

$$\lim_{\alpha \to \infty} \left(\sum_{i=1}^n (u_i/u_m)^\alpha \right)^{1/\alpha} = 1$$

and

$$\lim_{\alpha \to \infty} \frac{p(\alpha)}{\alpha} = \lim_{\alpha \to \infty} \log \left(u_m \left(\sum_{i=1}^n (u_i/u_m)^\alpha \right)^{1/\alpha} \right) = \log(u_m) > 0 \qquad (A4)$$

which proves the first result.

Now, $\sum_{i=1}^n \log(u_i) = \log(\frac{\prod_{i=1}^n t_i}{\prod_{i=1}^n t_i}) = \log(1) = 0$ and

$$\lim_{a \to 0^+} \frac{p(\alpha)}{\alpha^2} = \frac{\log\left(\frac{1}{n}\sum_{i=1}^n u_i^\alpha\right)}{\alpha^2} \overset{L'h}{=} \lim_{\alpha \to 0^+} \frac{\frac{\sum_{i=1}^n \log(u_i)u_i^\alpha}{\sum_{i=1}^n u_i^\alpha}}{2\alpha}$$

$$\overset{L'h}{=} \frac{1}{2} \lim_{\alpha \to 0^+} \frac{\left(\sum_{i=1}^n \log(u_i)^2 u_i^\alpha\right)\left(\sum_{i=1}^n u_i^\alpha\right) - \left(\sum_{i=1}^n \log(u_i)u_i^\alpha\right)^2}{\left(\sum_{i=1}^n u_i^\alpha\right)^2}$$

$$= \frac{1}{2} \frac{n \sum_{i=1}^n \log(u_i)^2 - \left(\sum_{i=1}^n \log(u_i)\right)^2}{n^2} = \frac{1}{2} \frac{\sum_{i=1}^n \log(u_i)^2}{n} > 0.$$

Note that $\frac{1}{2}(\sum_{i=1}^n \log(u_i)^2/n) \neq 0$ since, otherwise, we would have $\log(u_i) = 0 \Leftrightarrow u_i = 1, \forall i$ and would imply that t_i are all equal, which contradicts the hypothesis. Hence $p(\alpha) \propto_{\alpha \to 0^+} \alpha^2$ which proves the second result. ∎

Proposition A.10: Let $q(\alpha) = \log(\sum_{i=1}^n t_i^\alpha / \sqrt[n]{\prod_{i=1}^n t_i^\alpha})$, for $t_1, t_2, \ldots, t_n$ positive and not all equal, then $q(\alpha) > 0$ and the following hold

$$q(\alpha) \underset{\alpha \to 0^+}{\propto} 1 \quad \text{and} \quad q(\alpha) \underset{\alpha \to \infty}{\propto} \alpha.$$

Proof: Note that $\sum_{i=1}^n t_i^\alpha / \sqrt[n]{\prod_{i=1}^n t_i^\alpha} > n \Rightarrow q(\alpha) > 0$ by the arithmetic-geometric inequality. Since $q(\alpha) = \log(n) + p(\alpha)$ and by Proposition A.9 $\lim_{\alpha \to 0^+} p(\alpha) = 0$ it follows that

$$\lim_{\alpha \to 0^+} q(\alpha) = \log(n) > 0. \qquad (A5)$$

which proves the first proportionality.

Analogously, from $q(\alpha) = \log(n) + p(\alpha)$ and Proposition A.9 it follows that $q(\alpha) \propto_{\alpha \to \infty} \alpha$, hence the second proportionality is proved. ∎

Proposition A.11: Let $q(\alpha)$ be the same defined in Proposition A.10, then the following results are valid for $k \in \mathbb{R}^+$ and $r \in \mathbb{R}^+$.

$$\gamma(k, rq(\alpha)) \underset{\alpha \to 0^+}{\propto} 1 \quad \text{and} \quad \gamma(k, rq(\alpha)) \underset{\alpha \to \infty}{\propto} 1. \qquad (A6)$$

Proof: From (A5) and the continuity of incomplete gamma function in $\mathbb{R}^+ \times \mathbb{R}^+$ we have

$$\lim_{\alpha \to 0^+} \gamma(k, rq(\alpha)) = \gamma(k, r\log(n)) \Rightarrow \gamma(k, rq(\alpha)) \underset{\alpha \to 0^+}{\propto} 1. \tag{A7}$$

Now, from the definition of lower incomplete gamma function, it follows directly that $\lim_{x \to \infty} \gamma(y, x) = \Gamma(y)$ for $y > 0$. But, since $q(\alpha) \underset{\alpha \to \infty}{\propto} \alpha$, we have $\lim_{\alpha \to \infty} q(\alpha) = \infty$. Therefore

$$\lim_{\alpha \to \infty} \gamma(k, rq(\alpha)) = \Gamma(k) \Rightarrow \gamma(k, rq(\alpha)) \underset{\alpha \to 0^+}{\propto} 1. \tag{A8}$$

∎

Proposition A.12: *Let $p(\alpha)$ be the same defined in Proposition A.9 and let $t_m = \max\{t_1, \ldots, t_n\}$. Then the following results are valid for $k \in \mathbb{R}^+$ and $r \in \mathbb{R}^+$,*

$$\Gamma(k, rp(\alpha)) \underset{\alpha \to 0^+}{\propto} 1 \quad and \quad \Gamma(k, rp(\alpha)) \underset{\alpha \to \infty}{\propto} \alpha^{k-1} e^{-r\log(t_m/\sqrt[n]{\prod_{i=1}^n t_i})\alpha} \tag{A9}$$

where $\Gamma(y, x) = \int_x^\infty w^{y-1} e^{-w} \, dw$ is the upper incomplete gamma function.

Proof: From the definition of upper incomplete gamma function, it follows directly that $\lim_{x \to 0^+} \Gamma(y, x) = \Gamma(y)$ for $y > 0$. However, as $p(\alpha) \underset{\alpha \to 0^+}{\propto} \alpha^2$, we have $\lim_{\alpha \to 0^+} p(\alpha) = 0$. Therefore

$$\lim_{\alpha \to 0^+} \Gamma(k, rp(\alpha)) = \Gamma(k) \Rightarrow \Gamma(k, rp(\alpha)) \underset{\alpha \to 0^+}{\propto} 1. \tag{A10}$$

Now, by L'hospital rule and the definition of upper incomplete gamma function,

$$\lim_{x \to \infty} \frac{\Gamma(s, x)}{x^{s-1} e^{-x}} = 1.$$

However, $p(\alpha) \underset{\alpha \to \infty}{\propto} \alpha^2$ which implies $\lim_{\alpha \to \infty} p(\alpha) = \infty$ and

$$\lim_{\alpha \to \infty} \frac{\Gamma(k, rp(\alpha))}{(rp(\alpha))^{k-1} e^{-rp(\alpha)}} = 1 \Rightarrow \Gamma(k, rp(\alpha)) \underset{\alpha \to \infty}{\propto} p(\alpha)^{k-1} e^{-rp(\alpha)}.$$

Moreover, let c be the number of t_i equal to t_m for $i = 1, \ldots, n$, then

$$\lim_{\alpha \to \infty} p(\alpha) - \log\left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}}\right)\alpha = \lim_{\alpha \to \infty} \log\left(\left(\frac{t_1}{t_m}\right)^\alpha + \cdots + \left(\frac{t_n}{t_m}\right)^\alpha\right) = \log(c).$$

Therefore

$$\lim_{\alpha \to \infty} \frac{p(\alpha)^{k-1} e^{-rp(\alpha)}}{\alpha^{k-1} e^{-r\log(t_m/\sqrt[n]{\prod_{i=1}^n t_i})\alpha}} = \lim_{\alpha \to \infty} \left(\frac{p(\alpha)}{\alpha}\right)^{k-1} e^{-r(p(\alpha) - \log(t_m/\sqrt[n]{\prod_{i=1}^n t_i})\alpha)}$$

$$= 1^{k-1} \times e^{-r\log(c)} = c^{-r} > 0.$$

This implies that $p(\alpha)^{k-1} e^{-rp(\alpha)} \underset{\alpha \to \infty}{\propto} \alpha^{k-1} e^{-r\log(t_m/\sqrt[n]{\prod_{i=1}^n t_i})\alpha}$. Finally

$$\Gamma(k, rp(\alpha)) \underset{\alpha \to \infty}{\propto} \alpha^{k-1} e^{-r\log(t_m/\sqrt[n]{\prod_{i=1}^n t_i})\alpha}.$$

∎

Appendix 3. Proof of the theorems

Theorem A.13: *Let the prior be factored in the form $\pi(\alpha, \mu, \phi) \propto \pi(\phi)\pi(\alpha)\pi(\mu)$, where $\pi(\alpha) \propto \alpha^q$, $\pi(\mu) \propto (1/\mu)$ and $\pi(\phi) \underset{\phi \to 0^+}{\propto} \phi^r$, $s \in \mathbb{R}$ and $r \in \mathbb{R}$. If $q \geq r$ then*

$$d(t) = \frac{\pi(\phi, \alpha, \mu)\alpha^n}{\Gamma(\phi)^n} \mu^{n\alpha\phi} \left\{\prod_{i=1}^n t_i^{\alpha\phi-1}\right\} \exp\left\{-\mu^\alpha \sum_{i=1}^n t_i^\alpha\right\} = \infty \tag{A11}$$

Proof: Since $(\pi(\phi, \alpha, \mu)\alpha^n/\Gamma(\phi)^n)\mu^{n\alpha\phi}\{\prod_{i=1}^n t_i^{\alpha\phi-1}\} \exp\{-\mu^\alpha \sum_{i=1}^n t_i^\alpha\} \geq 0$ by Tonelli theorem (see Folland [29]) we have

$$d(t) = \int_{\mathcal{A}} \frac{\pi(\phi, \alpha, \mu)\alpha^n}{\Gamma(\phi)^n} \mu^{n\alpha\phi} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\theta$$

$$= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\pi(\phi, \alpha, \mu)\alpha^n}{\Gamma(\phi)^n} \mu^{n\alpha\phi} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu \, d\phi \, d\alpha$$

$$\propto \int_0^\infty \int_0^\infty \int_0^\infty \alpha^{n+q} \frac{\pi(\phi)}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu \, d\phi \, d\alpha$$

$$\geq \int_0^\infty \int_0^1 \alpha^{n+q-1} \pi(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha$$

using the Propositions A.5 and A.8 we have

$$d(t) \geq \int_0^\infty \int_0^1 \alpha^{n+q-1} \pi(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha$$

$$\propto \int_0^\infty \int_0^1 \alpha^{n+q-1} \times \phi^r \times \phi^{n-1} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi \, d\alpha = \int_0^\infty \alpha^{n+q-1} \int_0^1 \phi^{n+r-1} e^{-n\phi q(\alpha)} d\phi \, d\alpha$$

$$= \int_0^\infty \alpha^{n+q-1} \frac{\gamma(n+r, nq(\alpha))}{(nq(\alpha))^{n+r}} d\alpha \geq \int_1^\infty \alpha^{n+q-1} \frac{\gamma(n+r, nq(\alpha))}{(nq(\alpha))^{n+r}} d\alpha$$

$$\propto \int_1^\infty \alpha^{n+q-1} \frac{1}{q(\alpha)^{n+r}} \gamma(n+r, nq(\alpha)) \, d\alpha.$$

Now, from Propositions A.5, A.10 and A.11

$$\int_1^\infty \alpha^{n+q-1} \frac{1}{q(\alpha)^{n+r}} \gamma(n+r, nq(\alpha)) \, d\alpha \propto \int_1^\infty \alpha^{n+q-1} \frac{1}{\alpha^{n+r}} \times 1 \, d\alpha \propto \int_1^\infty \alpha^{q-r-1} \, d\alpha. \tag{A12}$$

Consequently $d(t) = \infty$ for $q \geq r$. ∎

A.1 Proof of Theorem 3.1

The Jeffrey prior is given by $\pi_J(\alpha, \mu, \phi) \propto \pi_J(\phi)/\mu$ where $\pi_J(\phi) \propto \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}$. Then, by Proposition A.6 we have

$$\pi_J(\phi) \underset{\phi \to 0^+}{\propto} 1 \quad \text{and} \quad \pi_J(\phi) \underset{\phi \to \infty}{\propto} \frac{1}{\phi}.$$

Thus $\pi(\alpha) \propto 1$, $\pi(\mu) \propto (1/\mu)$. Since $q = 0$ and $r = 0$, we have $q \geq r$ and by Theorem A.13 $d_J(t) = \infty$.

A.2 Proof of Theorem 3.2

Since $(\alpha^{n+1}/\Gamma(\phi)^{n+1})\mu^{n\alpha\phi+1} \exp\{-\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi)(\phi - 1/\alpha) - \phi\}\{\prod_{i=1}^n t_i^{\alpha\phi-1}\} \geq 0$, by Tonelli theorem

$$d_Z(t) = \int_{\mathcal{A}} \frac{\alpha^{n+1} \mu^{n\alpha\phi+1}}{\Gamma(\phi)^{n+1}} \prod_{i=1}^n t_i^{\alpha\phi-1} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi \right\} d\theta$$

$$= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^{n+1} \mu^{n\alpha\phi+1}}{\Gamma(\phi)^{n+1}} \prod_{i=1}^n t_i^{\alpha\phi-1} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi \right\} d\mu \, d\phi \, d\alpha$$

$$= \int_0^\infty \int_0^\infty \alpha^n \frac{\Gamma(n\phi + \frac{2}{\alpha})}{\Gamma(\phi)^{n+1}} \frac{\prod_{i=1}^n t_i^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi+2/\alpha}} \exp\left(\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi \right) d\phi \, d\alpha.$$

Let α be fixed. By the digamma recurrence relation $\psi(\phi) = -1/\phi + \psi(\phi+1)$ we have that $\lim_{\phi\to 0^+}(\psi(\phi)/1/\phi) = -1$. Therefore

$$\lim_{\phi\to 0^+} \frac{\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi}{\frac{1}{\phi}} = \lim_{\phi\to 0^+} \frac{\psi(\phi)}{\frac{1}{\phi}}\left(\phi - \frac{1}{\alpha}\right) - \phi^2 = -1 \times \left(0 - \frac{1}{\alpha}\right) - 0 = \frac{1}{\alpha} > 0.$$

Since the above limit is positive by the basic properties of limit, there is a $b > 0$ such that $\psi(\phi)(\phi - 1/\alpha) - \phi > 0$ for all $\phi \in (0, b]$. However, this also implies $\psi(\phi)(\phi - 1/\alpha) - \phi \propto_{\phi\to 0^+}(1/\phi)$, which by Proposition A.4 follows that $\psi(\phi)(\phi - 1/\alpha) - \phi \propto 1/\phi$ in $(0, b]$. That is, there exists $c > 0$ such that $\psi(\phi)(\phi - 1/\alpha) - \phi \geq c(1/\phi)$ for all $\phi \in (0, b]$. Note that

$$\Gamma\left(n\phi + \frac{2}{\alpha}\right) \underset{\phi\to 0^+}{\propto} \Gamma\left(\frac{2}{\alpha}\right) \underset{\phi\to 0^+}{\propto} 1 \quad \text{and} \quad \frac{(\prod_{i=1}^n t_i)^{\alpha\phi - 1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi + \frac{2}{\alpha}}} \underset{\phi\to 0^+}{\propto} \frac{(\prod_{i=1}^n t_i t)^{-1}}{(\sum_{i=1}^n t_i^\alpha)^{2/\alpha}} \underset{\phi\to 0^+}{\propto} 1.$$

Moreover, from Equation (A2), $(1/\Gamma(\phi)^n)\propto_{\phi\to 0^+}\phi^n$. Therefore, we have

$$\int_0^\infty \alpha^n \frac{\Gamma(n\phi + \frac{2}{\alpha})}{\Gamma(\phi)^{n+1}} \frac{\prod_{i=1}^n t_i^{\alpha\phi - 1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi + 2/\alpha}} \exp\left(\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi\right) d\phi \propto$$

$$\int_0^b 1 \times \phi^n \times 1 \times \exp\left(\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi\right) d\phi \geq \int_0^b \phi^n \times \exp\left(\frac{c}{\phi}\right) d\phi \overset{(\phi = \frac{1}{x})}{=} \int_{1/b}^\infty \frac{e^{cx}}{x^{n+2}} dx = \infty$$

and

$$d_Z(t) = \int_0^\infty \int_0^\infty \alpha^n \frac{\Gamma(n\phi + \frac{2}{\alpha})}{\Gamma(\phi)^{n+1}} \frac{\prod_{i=1}^n t_i^{\alpha\phi - 1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi + 2/\alpha}} \exp\left(\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi\right) d\phi\, d\alpha = \int_0^\infty \infty\, d\alpha = \infty$$

A.3   Proof of Theorem 3.4

Let $\boldsymbol{\theta} = (\mu, \phi, \alpha)$, where the parameter space of θ_j is independent of $\boldsymbol{\theta}_{-j}$ for $j = 1,2,3$. Then, after some algebraic manipulations in Equation (7), we have

$$h_{1,1}^{1/2}(\boldsymbol{\theta}) = \frac{\alpha}{\mu}\sqrt{\frac{\phi^2\psi'(\phi)^2 - \psi'(\phi) - 1}{\psi'(\phi)\left(1 + 2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2\right) - \psi(\phi)^2}} = f_1(\mu)g_1(\phi)g_1(\alpha)$$

where $f_1(\mu) = 1/\mu$, $g_1(\phi) = \sqrt{\phi^2\psi'(\phi)^2 - \psi'(\phi) - 1/\psi'(\phi)(1 + 2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2) - \psi(\phi)^2}$ and $g_1(\alpha) = \alpha$.

$$h_{2,2}^{1/2}(\boldsymbol{\theta}) = \sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}} = g_2(\mu)f_2(\phi)g_2(\alpha),$$

where $g_2(\mu) = 1, f_2(\phi) = \sqrt{\psi'(\phi) - \psi(\phi)^2/2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}$ and $g_2(\alpha) = 1$.

$$h_{3,3}^{1/2}(\boldsymbol{\theta}) = \frac{\sqrt{1 + 2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2}}{\alpha} = g_3(\alpha)g_3(\mu)f_3(\phi),$$

where $g_3(\mu) = 1, g_3(\phi) = \sqrt{1 + 2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2}$ and $f_3(\alpha) = 1/\alpha$.

Following the Proposition 3.3, for the ordered parameters (μ, ϕ, α) the conditional reference prior are

$$\pi(\alpha|\phi, \mu) \propto f_3(\alpha) \propto \frac{1}{\alpha}, \quad \pi(\phi|\mu) \propto f_2(\phi) \propto \sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}$$

and

$$\pi(\mu) \propto f_1(\mu) \propto \frac{1}{\mu}.$$

Therefore, the joint $\boldsymbol{\theta}$-reference prior is given by

$$\pi_R(\mu, \phi, \alpha) \propto \pi(\alpha|\phi, \mu)\pi(\phi|\mu)\pi(\mu) \propto \frac{1}{\alpha\mu}\sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1}}.$$

A.4 Proof of Theorem 3.5

The reference prior is given by $\pi_R(\alpha, \mu, \phi) \propto \frac{\pi_R(\phi)}{\mu\alpha}$ where by Proposition A.7 we have

$$\pi_R(\phi) \underset{\phi \to 0^+}{\propto} \frac{1}{\sqrt{\phi}} \quad \text{and} \quad \pi_R(\phi) \underset{\phi \to \infty}{\propto} \frac{1}{\phi}.$$

Since $(\alpha^{n-1}\pi_R(\phi)/\Gamma(\phi)^n)\mu^{n\alpha\phi-1}\{\prod_{i=1}^n t_i^{\alpha\phi-1}\}\exp\{-\mu^\alpha \sum_{i=1}^n t_i^\alpha\} \geq 0$, by Tonelli theorem

$$
\begin{aligned}
d_R(t) &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^{n-1}\pi_R(\phi)}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu \, d\phi \, d\alpha \\
&= \int_0^\infty \int_0^\infty \alpha^{n-2} \frac{\pi_R(\phi)\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha \geq \int_0^1 \int_1^\infty \alpha^{n-2} \frac{\pi_R(\phi)\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha \\
&\propto \int_0^1 \int_1^\infty \alpha^{n-2} \frac{\Gamma(n\phi)}{\phi\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha \propto \int_0^1 \alpha^{n-2} \int_1^\infty \phi^{(n-1)/2-1} n^{n\phi} \frac{(\prod_{i=1}^n t_i^\alpha)^\phi}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha \\
&= \int_0^1 \alpha^{n-2} \int_1^\infty \phi^{(n-1)/2-1} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{(1/n)\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi \, d\alpha. = \int_0^1 \alpha^{n-2} \int_1^\infty \phi^{(n-1)/2-1} e^{-n\mathrm{p}(\alpha)\phi} d\phi \, d\alpha \\
&= \int_0^1 \alpha^{n-2} \frac{\Gamma((n-1)/2, n\mathrm{p}(\alpha))}{(n\mathrm{p}(\alpha))^{(n-1)/2}} d\alpha \propto \int_0^1 \alpha^{n-2} \frac{1}{(n\alpha^2)^{(n-1)/2}} d\alpha \propto \int_0^1 \alpha^{-1} d\alpha = \infty.
\end{aligned}
$$

A.5 Proof of Theorem 3.6

The Jeffreys/Reference prior is given by $\pi_{JR}(\alpha, \mu, \phi) \propto \frac{\pi_{JR}(\phi)}{\mu\sqrt{\alpha}}$ where $\pi_{JR}(\phi) = \sqrt{\pi_J(\phi)\pi_R(\phi)}$ and by Propositions A.6 and A.7 we have

$$\pi_{JR}(\phi) \underset{\phi \to 0^+}{\propto} \frac{1}{\sqrt[4]{\phi}} \quad \text{and} \quad \pi_{JR}(\phi) \underset{\phi \to \infty}{\propto} \frac{1}{\phi}.$$

Since $\alpha^{n-1/2}(\pi_{JR}(\phi)/\Gamma(\phi)^n)\mu^{n\alpha\phi-1}\{\prod_{i=1}^n t_i^{\alpha\phi-1}\}\exp\{-\mu^\alpha \sum_{i=1}^n t_i^\alpha\} \geq 0$, by Tonelli theorem we have

$$
\begin{aligned}
d_{JR}(t) &= \int_{\mathcal{A}} \alpha^{n-1/2} \frac{\pi_{JR}(\phi)}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\theta \\
&= \int_0^\infty \int_0^\infty \int_0^\infty \alpha^{n-1/2}\pi_{JR}(\phi) \frac{1}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu \, d\phi \, d\alpha \\
&= \int_0^\infty \int_0^\infty \alpha^{n-3/2}\pi_{JR}(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha \, d\phi \, d\alpha \\
&= s_1(t) + s_2(t) + s_3(t) + s_4(t),
\end{aligned}
$$

where $s_1(t)$, $s_2(t)$, $s_3(t)$ and $s_4(t)$ are presented below

$$
\begin{aligned}
s_1(t) &= \int_0^1 \int_0^1 \alpha^{n-3/2}\pi_{JR}(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \, d\alpha \\
&\propto \int_0^1 \int_0^1 \alpha^{n-3/2} \times \phi^{-1/4} \times \phi^{n-1} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi \, d\alpha = \int_0^1 \alpha^{n-3/2} \int_0^1 \phi^{n-1/4-1} e^{-n\phi\mathrm{q}(\alpha)} d\phi \, d\alpha \\
&= \int_0^1 \alpha^{n-3/2} \frac{1}{(n\mathrm{q}(\alpha))^{n-1/4}} \gamma\left(n-\frac{1}{4}, n\mathrm{q}(\alpha) \right) d\alpha \propto \int_0^1 \alpha^{n-3/2} \frac{1}{1^{n-1/4}} \times 1 \, d\alpha = \int_0^1 \alpha^{n-3/2} d\alpha < \infty.
\end{aligned}
$$

$$s_2(t) = \int_1^\infty \int_0^1 \alpha^{n-3/2} \pi_{JR}(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} \, d\phi \, d\alpha$$

$$\propto \int_1^\infty \int_0^1 \alpha^{n-3/2} \times \phi^{-1/4} \times \phi^{n-1} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi \, d\alpha = \int_1^\infty \alpha^{n-3/2} \int_0^1 \phi^{n-1/4-1} e^{-n\phi q(\alpha)} \, d\phi \, d\alpha$$

$$= \int_1^\infty \alpha^{n-3/2} \frac{1}{(nq(\alpha))^{n-1/4}} \gamma\left(n - \frac{1}{4}, nq(\alpha) \right) d\alpha \propto \int_1^\infty \alpha^{n-3/2} \frac{1}{\alpha^{n-1/4}} \times 1 \, d\alpha = \int_1^\infty \alpha^{-5/4} \, d\alpha < \infty.$$

$$s_3(t) = \int_0^1 \int_1^\infty \alpha^{n-\frac{3}{2}} \pi_{JR}(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} \, d\phi \, d\alpha$$

$$\propto \int_0^1 \int_1^\infty \alpha^{n-\frac{3}{2}} \times \frac{1}{\phi} \times n^{n\phi} \phi^{\frac{n-1}{2}} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi \, d\alpha = \int_0^1 \alpha^{n-3/2} \int_1^\infty \phi^{(n-1)/2-1} e^{-n\phi p(\alpha)} \, d\phi \, d\alpha$$

$$= \int_0^1 a^{n-3/2} \frac{1}{(np(\alpha))^{(n-1)/2}} \Gamma\left(\frac{n-1}{2}, np(\alpha) \right) d\alpha \propto \int_0^1 \alpha^{n-3/2} \frac{1}{(\alpha^2)^{(n-1)/2}} \times 1 \, d\alpha = \int_0^1 \alpha^{-1/2} \, d\alpha < \infty.$$

$$s_4(t) = \int_1^\infty \int_1^\infty \alpha^{n-\frac{3}{2}} \pi_{JR}(\phi) \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} \, d\phi \, d\alpha$$

$$\propto \int_1^\infty \int_1^\infty \alpha^{n-3/2} \times \frac{1}{\phi} \times n^{n\phi} \phi^{(n-1)/2} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi \, d\alpha = \int_1^\infty \alpha^{n-3/2} \int_1^\infty \phi^{(n-1)/2-1} e^{-n\phi p(\alpha)} \, d\phi \, d\alpha$$

$$= \int_1^\infty a^{n-3/2} \frac{1}{(np(\alpha))^{(n-1)/2}} \Gamma\left(\frac{n-1}{2}, np(\alpha) \right) d\alpha$$

$$\propto \int_1^\infty \alpha^{n-3/2} \frac{1}{\alpha^{(n-1)/2}} \times \alpha^{(n-1)/2-1} e^{-n\log(t_m/\sqrt[n]{\prod_{i=1}^n t_i})\alpha} \, d\alpha = \int_1^\infty \alpha^{-5/2} e^{-L(t)\alpha} \, d\alpha < \infty.$$

where $L(t) = n \log(\frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}}) > 0$. Therefore, we have $d_{JR}(t) = s_1(t) + s_2(t) + s_3(t) + s_4(t) < \infty.$