Bayesian inference for quantiles of the log-normal distribution

Bayesian inference for quantiles of the log-normal distribution

Aldo Gardini[*], *Carlo Trivisano*[*], *Enrico Fabrizi*[†]

Abstract

The log-normal distribution is very popular for modelling positive right-skewed data and represents a common distributional assumption in many environmental applications. Here we consider the estimation of quantiles of this distribution from a Bayesian perspective. We show that the prior on the variance of the log of the variable is relevant for the properties of the posterior distribution of quantiles. Popular choices for this prior, such as the inverse gamma, lead to posteriors without finite moments. We propose the generalized inverse Gaussian and show that a restriction on the choice of one of its parameters guarantees the existence of posterior moments up to a pre-specified order. In small samples, a careful choice of the prior parameters leads to point and interval estimators of the quantiles with good frequentist properties, outperforming those currently suggested by the frequentist literature. Finally, two real examples from environmental monitoring and occupational health frameworks highlight the improvements of our methodology, especially in a small sample situation.

Keywords: Bessel Functions; Environmental Monitoring; Generalized Inverse Gaussian; Small Samples

1 Introduction

The log-normal is among the most popular distributions for modelling positive, right-skewed continuous data. Its relationship to the normal (if $X \sim \mathcal{N}(\xi, \sigma^2)$ then $Z = \exp(X) \sim \mathcal{LN}(\xi, \sigma^2)$ is said to be log-normally distributed) makes this distribution appealing to applied scientists from various fields (see for instance Kosugi, 1996; May et al., 2000; Ignatov et al., 2000; Limpert et al., 2001; Lawless, 2003; Bengtsson et al., 2005).

In this paper we focus on the estimation of log-normal quantiles, that can be of interest in many applications. Specifically, in environmental monitoring and occupational health analyses it is common to estimate extreme quantiles in the right tail of a skewed distribution from small samples (Bullock and Ignacio, 2006; Gibbons et al., 2009; Krishnamoorthy et al., 2011), or to compare a fixed legal exposure limit to an extreme quantile (or to its upper confidence limit, UCL) estimated from a typically small sample. In this case, the tools available in the current literature can produce unefficient point and interval estimators with poor coverage or low precision and can be significantly improved. The methodology we propose improves current methods, especially in the analysis of small samples.

As a motivating real-data example, the following popular data set (USEPA, 2009; Millard, 2013) is considered. It consists in $n = 8$ chrysene concentrations (ppb) obtained from two background

[*]Dipartimento di Scienze Statistiche, Università degli Studi di Bologna
[†]Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore

wells:

$$19.7,\ 39.2,\ 7.8,\ 12.8,\ 10.2,\ 7.2,\ 16.1,\ 5.7.$$

Background wells data are used to set the tolerance limit for chrysene concentrations in the site. Specifically, the limit equals the 95% UCL of the 95−th percentile. In doing this, as in most of the cases samples from background wells are small, parametric assumptions, and namely the log-normal, are often used.

Statistical techniques for the estimation of this site-specific UCL described in the EPA guidelines (that we will label in section 2.2 as naive) lead to overly conservative, i.e. too large, values for the tolerance limit. This implies wrong evaluations of safety for the site and consequently wrong decisions.

Inference about the log-normal distribution often targets functionals of (ξ, σ^2) such as $\theta_{a,b} = \exp(a\xi + b\sigma^2)$, that include all moments along with the mode and the median. In case of interest on the quantiles, the target functional is $\theta_p = \exp(\xi + \Phi^{-1}(p)\sigma)$ where $\Phi^{-1}(p)$, $p \in (0,1)$ is the quantile function of a $\mathcal{N}(0,1)$ random variable.

Notoriously, naive transformations of estimators of (ξ, σ^2) efficient on the transformed scale do not lead to efficient estimators of functional on the exponential scale such as $\theta_{a,b}$ or θ_p (Finney, 1941). The problem of estimating moments of the log-normal distribution and especially the mean $(a = 1, b = 0.5)$ has a long tradition in both the frequentist and Bayesian literature. For the frequentist approach we may refer to Crow and Shimizu (1988) for a review of early results and to Shen et al. (2006). On the Bayesian side, a key reference is Zellner (1971); Rukhin (1986) proposes a Bayesian estimator with optimal frequentist properties (mean square error); Fabrizi and Trivisano (2012) note that under many popular choices for the prior $p(\sigma^2)$, the posterior for $\theta_{a,b}$, although well defined, has no finite moments, a fact that precludes the use of ordinary loss functions for summarizing the posterior distribution. Adopting a generalized inverse Gaussian (GIG) for $p(\sigma^2)$, they characterize the choices of the hyperparameters that guarantee the existence of posterior moments for $\theta_{a,b}$ and propose to set hyperparameters in order to optimize frequentist properties of point predictors.

The estimation of log-normal quantiles has received little attention so far. In the frequentist literature, Longford (2012) identifies a class of estimators depending on two constants that he determines with the aim of minimizing the frequentist mean square error (MSE); he overlooks relevant inferential problems such as interval estimation. As far as we know, in the Bayesian literature, the problem has not been considered. This paper contributes to fill this gap.

In the first place we adapt the results in Zellner (1971), which are conditional on σ^2; then in line with Fabrizi and Trivisano (2012) we propose a GIG prior for σ^2 and study which values of the hyperparameters lead to posterior distributions of the quantiles with finite moments up to a prespecified order. It turns out that notable special cases of the GIG, such as the inverse gamma, lead to posterior for θ_p without finite moments.

We study analytically the posterior distribution of the quantiles under a conjugate Normal-GIG prior for (ξ, σ^2) (Thabane and Haq, 1999). This posterior is a distribution new to the literature and we label it as $SMNG$: the acronym stays for *a scale-mean mixture of normal distribution assuming a GIG distribution on the scale*. We study the properties of this distribution and obtain formulas for the density, moments and propose algorithms to generate random samples from it. We discuss the choice of hyperparameters of $p(\sigma^2)$ that, a part allowing for the existing of posterior moments, turns out to be relevant in the analysis of small samples. In doing this we aim both at point predictors with small mean square error and posterior probability intervals with good frequentist coverage, proposing two alternative solutions. We find that the two goals are to some extent conflicting and hyperparameters optimal for minimizing the MSE lead to sub-optimal intervals and vice versa. The choice should then be tuned on inferential aims of specific applications. Of course, as the

sample size increases, the choice of hyperparameters, others than the one ruling the existence of posterior moments become irrelevant. Our proposals are motivated by both theoretical results and simulations. Specifically, we obtain probability intervals that reach the nominal frequentist coverage and are shorter than those obtained with the methods currently used.

All the methods presented in this paper have been implemented in `R` functions that are included in the `BayesLN` package (Gardini et al., 2020), in order to simplify the inferential process to practitioners.

The rest of the paper is organized as follows. In section 2 we review published results on log-normal quantiles relevant to our research. In section 3 we obtain the posterior for θ_p when a GIG prior for σ^2 is assumed, while the problem of hyperparameters choice is discussed in section 4. Section 5 introduces a simulation exercise and section 6 two applications of the methodology we propose to small real datasets from the environmental monitoring literature. Section 7 offers concluding remarks.

2 Preliminary results on estimation of quantiles

In this section we review some earlier results on quantile estimation for the log-normal distribution that will be later used in discussing our proposals. Before doing this we set some notation.

Let's assume that a random sample $X_1, \ldots, X_n$ is drawn from a $X \sim \mathcal{N}(\xi, \sigma^2)$; we are interested in estimating $\theta_p = \exp(\xi + \Phi^{-1}(p)\sigma)$, $p \in (0, 1)$, the p-quantile of the log-normal variable $Z = \exp(X)$. Let's denote $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$ and $V^2 = n^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ the two sufficient statistics.

2.1 Non-parametric estimation

Although apparently straightforward, there exist a host of different methods for basic non-parametric estimation of quantiles (see Hyndman and Fan, 1996). As a benchmark for our proposals, the standard `R` function `quantile` will be considered, and specifically the type 7 method based on Gumbel (1939). The estimator of the p-quantile, included between the positions $k-1$ and k in the ordered sample is defined as:

$$\hat{Q}_p^7 = X_{(k-1)} - \left(X_{(k-1)} - X_{(k)}\right) \frac{p_{k+1} - p}{p_{k+1} - p_k}, \tag{1}$$

where $p_k = (k-1)/(n-1)$.

2.2 Naive estimation

A simple estimator of θ_p can be obtained replacing unknowns with their maximum-likelihood estimators:

$$\hat{\theta}_p = \exp\left\{\bar{X} + \Phi^{-1}(p)\hat{\sigma}\right\}, \tag{2}$$

where $\hat{\sigma} = \sqrt{V^2}$ ($S = \sqrt{nV^2/(n-1)}$ can alternatively be used in small samples). In the same line we can compute the extremes of the confidence intervals. In the two sided case with the fixed confidence level $1 - \alpha$ they are (Gibbons et al., 2009):

$$\left[\exp\left\{\bar{X} + t_{\left(\frac{\alpha}{2}, n-1, k_p\right)} \frac{\hat{\sigma}}{\sqrt{n}}\right\}; \exp\left\{\bar{X} + t_{\left(1-\frac{\alpha}{2}, n-1, k_p\right)} \frac{\hat{\sigma}}{\sqrt{n}}\right\}\right], \tag{3}$$

where $t_{\left(\frac{\alpha}{2}, n-1, k_p\right)}$ is the quantile $\frac{\alpha}{2}$ of a non-central Student's t distribution with $n-1$ degrees of freedom and a non-centrality parameter $k_p = \sqrt{n}\Phi^{-1}(p)$.

2.3 Longford's minimum MSE estimator

Longford (2012) proposes an estimator of θ_p in the form:

$$Q_p = \exp\left\{\bar{X} + b_p\hat{\sigma} + d_p\hat{\sigma}^2\right\}. \tag{4}$$

The values of the constants (b_p, d_p) are determined by minimizing the MSE of the estimator using a Newton-Raphson algorithm. We point out that this estimator has finite expectation only when d_p is negative or when $\sigma^2 < \frac{n-1}{2d_p}$. The same inequality divided by 2 determine the existence condition for the MSE. These conditions are not testable since the variance σ^2 is not known.

2.4 Bayes estimator conditional on σ^2

Conditionally on σ^2, let's assume a normal prior for ξ:

$$\xi|\sigma^2 \sim \mathcal{N}\left(\xi_0, \frac{\sigma^2}{n_0}\right). \tag{5}$$

It can easily be shown that $\xi|\sigma^2, \bar{X}, V^2 \sim \mathcal{N}\left(\xi_1, \frac{\sigma^2}{n_1}\right)$ where $\xi_1 = w\bar{X} + (1-w)\xi_0$, $n_1 = n + n_0$ and $w = \frac{n}{n_1}$. Moreover, if we denote $\eta_p = \log\theta_p = \xi + \Phi^{-1}(p)\sigma$, then we have that:

$$\eta_p|\sigma^2, X \sim \mathcal{N}\left(\bar{\eta}_p, \frac{\sigma^2}{n_1}\right), \tag{6}$$

where $\bar{\eta}_p = \xi_1 + \Phi^{-1}(p)\sigma$.

Zellner (1971) studies a minimum MSE for the log-normal mean $\theta_{1,0.5}$ among those in the form: $\theta^*_{1,0.5} = k \cdot \exp\{\bar{X}\}$, where k is a constant possibly involving σ^2. He assumes the flat prior $p(\xi) \propto 1$, a special case of (5) for $n_0 \to 0$. He finds that the minimum MSE estimator can be obtained as the Bayesian point predictor that minimizes the relative quadratic loss function. We can extend Zellner (1971) result to the estimation of log-normal quantiles.

Theorem 2.1: Among the estimators of the functional θ_p in the class $\theta^*_p = k \cdot \exp\{\bar{X}\}$, the one that minimizes the frequentist MSE is:

$$\hat{\theta}^*_p = \exp\left\{\bar{X} + \sigma\Phi^{-1}(p) - \frac{3\sigma^2}{2n}\right\}. \tag{7}$$

It coincides with the Bayes estimator that minimizes the relative quadratic loss function (conditionally on σ^2).

We note that the predictor minimizing the frequentist mean square error is different from the expectation of $\exp(\eta_p|\bar{X}, \sigma^2)$ that can be derived from (6).

3 The posterior distribution of the log-normal quantiles under a Generalized Inverse Gaussian prior distribution for σ^2

In this paper we assume a generalized inverse Gaussian (GIG) prior for σ^2:

$$p(\sigma^2) \sim GIG(\lambda, \delta, \gamma). \tag{8}$$

The density of the GIG distribution may be written as follows:

$$p(x) = \left(\frac{\gamma}{\delta}\right)^{\lambda} \frac{1}{2K_{\lambda}(\delta\gamma)} x^{\lambda-1} \exp\left\{ -\frac{1}{2}\left(\delta^2 x^{-1} + \gamma^2 x\right)\right\} \mathbf{1}_{\Re^+}, \tag{9}$$

where $K_\nu(y)$ is the Bessel K function of order ν and argument y (Gradshteyn and Ryzhik, 2014). The parameters domain is given by $\delta > 0$, $\gamma \geq 0$ if $\lambda < 0$; $\delta > 0$, $\gamma > 0$ if $\lambda = 0$; $\delta \geq 0$, $\gamma > 0$ if $\lambda > 0$.

The first reason to consider the GIG is that many important distributions may be obtained as special cases. For $\lambda > 0$ and $\gamma > 0$, the gamma distribution emerges as the limit when $\delta \to 0$. The inverse-gamma is obtained when $\lambda < 0$, $\delta > 0$ and $\gamma \to 0$ and an inverse Gaussian distribution is obtained when $\lambda = -\frac{1}{2}$. For more details on the GIG distribution see Jorgensen (1982) or Paolella (2007).

If we couple (5) to (8), the resulting prior for (ξ, σ^2) can be labelled as Normal-GIG (Thabane and Haq, 1999) that enjoys interesting conjugacy properties. A first relevant result is that:

$$\sigma^2 | \bar{X}, V^2 \sim GIG\left(\bar{\lambda}, \sqrt{\delta^2 + nV^2}, \gamma\right), \tag{10}$$

where $\bar{\lambda} = \lambda - \frac{n}{2}$. For a proof see Fabrizi and Trivisano (2012). A second result is:

$$\xi | \bar{X}, V^2 \sim GH\left(\bar{\lambda}, \bar{\gamma}, 0, \bar{\delta}, \bar{\mu}\right), \tag{11}$$

where the GH is the Generalized Hyperbolic distribution introduced by Barndorff-Nielsen (1977) and $\bar{\gamma} = \sqrt{n_1}\gamma$, $\bar{\delta} = \left(\sqrt{n_1}\right)^{-1}\sqrt{\delta^2 + nV^2}$, $\bar{\mu} = \xi_1$. For details on the GH distribution see Bibby and Sørensen (2003). This result is consistent with the conjugacy of the Normal-GIG prior, since the normal prior (5) for ξ is conditioned with respect to σ^2 and it can be shown that marginally the prior on ξ is a GH distribution. The posterior (11) is a direct consequence of Barndorff-Nielsen introduction of the GH as a normal variance-mean mixture where the mixing distribution is GIG. Specifically, if:

$$Y = \mu + \beta W + \sqrt{W} Q,$$

where $\mu, \beta \in \Re$, $Q \sim \mathcal{N}(0,1)$, $W \sim GIG(\lambda, \delta, \gamma)$ with Q and W independent, then the marginal distribution of Y will be GH (i.e., $Y \sim GH(\lambda, \alpha, \beta, \delta, \mu)$, where $\alpha^2 = \beta^2 + \gamma^2$).

Formula (11) is very useful to study the posterior of $\theta_{a,b}$ as $\log\theta_{a,b} = a\xi + b\sigma^2$ is still GH distributed because of closure of GH with respect to linear affine transformations. A similar result would be useful to obtain the posterior of θ_p. In this case we have a different variance-mean mixture that we can label as normal square root mean-variance mixture. Specifically, if we consider:

$$Y' = \mu + \beta\sqrt{W} + \sqrt{W} Q,$$

assuming $W \sim GIG(\lambda, \delta, \gamma)$ independent on Q, the marginal distribution of Y' is not a GH any more. The resulting distribution has never been studied in the literature before. As anticipated in the introduction, we label it as $SMNG$. We can now state the following result.

Theorem 3.1: Assuming the priors (5), (8) then it is possible to obtain $\eta_p | \bar{X}, V^2 \sim SMNG(\bar{\lambda}, \bar{\delta}, \bar{\gamma}, \bar{\beta}, \bar{\mu})$, where $\bar{\beta} = \sqrt{n_1}\Phi^{-1}(p)$. As a consequence the main result is that:

$$\theta_p | \bar{X}, V^2 \sim \log SMNG(\bar{\lambda}, \bar{\delta}, \bar{\gamma}, \bar{\beta}, \bar{\mu}). \tag{12}$$

For details on the real-valued $SMNG$ distribution, its density, interpretation of parameters, moments and other properties see the Appendix A in the supplementary material. The $\log -SMNG$ distribution is such that its log transform is $SMNG$ distribution. See the Appendix A for a description of this distribution as well. Here we note that the parameters of the θ_p posterior distribution are consistent with the meaning of the $\log -SMNG$ parameters: a higher posterior sample size n_1 implies lighter tails (smaller and negative λ and bigger γ) and a density that is more peaked around the mode (smaller δ). On the other hand, the asymmetry parameter β is not an actual parameter as it is a function of the specific quantile under consideration through the inverse of the standardized Gaussian cumulative distribution function, and the location parameter $\bar{\mu}$ is given by the conditioned posterior mean ξ_1.

Using theorem A.2 in the Appendix we have that the r-th moment of θ_p is given by:

$$\mathbb{E}\left[\theta_p^r\right] = e^{\mu r} \frac{\left(\frac{\bar{\gamma}}{\sqrt{\bar{\gamma}^2 - r^2}}\right)^{\bar{\lambda}}}{K_\lambda(\bar{\delta}\bar{\gamma})} \sum_{i=0}^{+\infty} \frac{(r\bar{\beta})^i}{i!} \left(\frac{\bar{\delta}}{\sqrt{\bar{\gamma}^2 - r^2}}\right)^{\frac{i}{2}} K_{\bar{\lambda}+\frac{i}{2}}\left(\bar{\delta}\sqrt{\bar{\gamma}^2 - r^2}\right), \tag{13}$$

that is defined if $r < \bar{\gamma}$. We can state the following result.

Theorem 3.2: Assuming the priors (5), (8) then $\theta_p|\bar{X}, V^2$ has finite moments up to the order r if and only if:

$$\gamma > \frac{r}{\sqrt{n + n_0}}. \tag{14}$$

The existence of posterior moments is subjected to a restriction on the parameter γ, that controls the right tail of the distribution. Fabrizi and Trivisano (2012) obtained a parallel results for the posterior distribution of $\theta_{a,b}$. The condition (14) is less restrictive than the one they found; this could be expected as the functional involved in the quantile estimation is characterized by less variability, and consequently lighter tails, because of the presence of σ instead of σ^2. We note that the restriction on γ does not depend on the quantile estimated; moreover it becomes less and less restrictive as n is increases, thus allowing for priors with heavier tails. Condition (14) requires that γ is above a positive threshold. Note that the popular inverse gamma prior on σ^2, a special case of the GIG for $\lambda < 0$, $\delta > 0$, $\gamma \to 0$, does not respect condition (14) thereby leading to a posterior distribution with non-existent moments for finite sample sizes. Similarly, the uniform prior over the range $(0, A)$ for σ (Gelman, 2006) implies that $p(\sigma^2) \propto \frac{1}{\sigma}\mathbf{1}_{(0,A)}$, which may be seen as an approximation to a $Gamma(\frac{1}{2}, \epsilon)$ (where $\epsilon = (4A^2)^{-1}$) truncated at A^2. For $\lambda > 0$, $\gamma > 0$ and $\delta \to 0$, $GIG(\lambda, \delta, \gamma) \to Gamma(\lambda, \gamma^2/2)$. If we let $A \to \infty$, therefore, $p(\sigma) \propto 1$ is equivalent to a GIG prior with $\gamma \to 0$ and thus implies non-existent posterior moments.

The result of theorem 3.2 can be extended to the case in which censored observation are included in the sample (Balakrishnan and Mitra, 2011; Krishnamoorthy et al., 2011), as showed in appendix B of the supplementary material. This is an important finding since the Bayesian framework allows to easily deal with censored data, but the posterior moments of the target functional could be not finite under popular prior settings. However, the investigation of the posterior properties are omitted here since it is beyond the aim of this paper.

Focusing again on the uncensored data setting, we can now give the formulas for two Bayes estimators obtained under alternative loss functions.

Theorem 3.3 (Bayes estimators of θ_p): Given that the posterior distribution for the target functional θ_p is (12), then:

1. the Bayes estimator of the log-normal p-th quantile under the quadratic loss function is:

$$\hat{\theta}_p^{QB} = e^{\xi_1} \frac{\left(\frac{\sqrt{n_1}\gamma}{\sqrt{n_1\gamma^2-1}}\right)^{\bar{\lambda}}}{K_{\bar{\lambda}}(\sqrt{nV^2+\delta^2}\gamma)} \sum_{j=0}^{+\infty} \frac{\bar{\beta}^j}{j!} \left(\frac{\sqrt{nV^2+\delta^2}}{\sqrt{n_1}\sqrt{n_1\gamma^2-1}}\right)^{\frac{j}{2}} \times \\ \times K_{\bar{\lambda}+\frac{j}{2}}\left(\frac{\sqrt{nV^2+\delta^2}\sqrt{n_1\gamma^2-1}}{\sqrt{n_1}}\right), \tag{15}$$

   that exists when $\gamma > \frac{1}{\sqrt{n_1}}$;

2. the Bayes estimator under relative quadratic loss is:

$$\hat{\theta}_p^{RQB} = e^{\xi_1} \left(\frac{\sqrt{n_1\gamma^2-4}}{\sqrt{n_1\gamma^2-1}}\right)^{\bar{\lambda}} \times \\ \times \frac{\sum_{j=0}^{+\infty} \frac{\bar{\beta}^j}{j!} \left(\frac{\sqrt{nV^2+\delta^2}}{\sqrt{n_1}\sqrt{n_1\gamma^2-1}}\right)^{\frac{j}{2}} K_{\bar{\lambda}+\frac{j}{2}}\left(\frac{\sqrt{nV^2+\delta^2}\sqrt{n_1\gamma^2-1}}{\sqrt{n_1}}\right)}{\sum_{j=0}^{+\infty} \frac{\bar{\beta}^j}{j!} \left(\frac{4\sqrt{nV^2+\delta^2}}{\sqrt{n_1}\sqrt{n_1\gamma^2-4}}\right)^{\frac{j}{2}} K_{\bar{\lambda}+\frac{j}{2}}\left(\frac{\sqrt{nV^2+\delta^2}\sqrt{n_1\gamma^2-4}}{\sqrt{n_1}}\right)}; \tag{16}$$

   that exists when $\gamma > \frac{2}{\sqrt{n_1}}$.

Unfortunately, both estimators can only be expressed as infinite sums of Bessel K function. These sums are convergent (see proposition A.2 in the Appendix).

Despite the complicated expressions of the proposed estimators, the developed `R` package `BayesLN` contains the function `LN_quant()` that allows to easily obtain the numerical results.

4 Choice of the hyperparameters for the GIG prior

The choice of hyperparameters, especially those of $p(\sigma^2)$, that we assume within the GIG family, has a considerable impact on posterior inferences when the sample size is small. In this section we propose two alternative choice strategies. First, we consider a *weakly informative choice* of the hyperparameters that always satisfies the moments existence condition; in later sections this choice will be shown to produce posterior probability intervals with good properties from a frequentist perspective. The second choice strategy aims at obtaining Bayesian point predictors with optimal frequentist MSE. In all cases only priors that guarantee the existence of at least the first two moments for the posterior $p(\theta_p|\bar{X}, V^2)$ will be considered.

4.1 A weakly informative choice for the hyperparameters

We start letting $n_0 \to 0$ in (5) that becomes $p(\xi) \propto 1$, i.e. we specify an improper flat prior on the location parameter on the log scale, that leads to $\xi|\bar{X}, \sigma^2 \sim N(\bar{X}, n^{-1}\sigma^2)$. Note that, as a consequence, posterior moments coincide with maximum likelihood estimators.

In this line, our proposal for choosing the hyperparameters vector $(\lambda, \delta, \gamma)$ is to have $E(\sigma^2|\bar{X}, V^2) \cong V^2$.

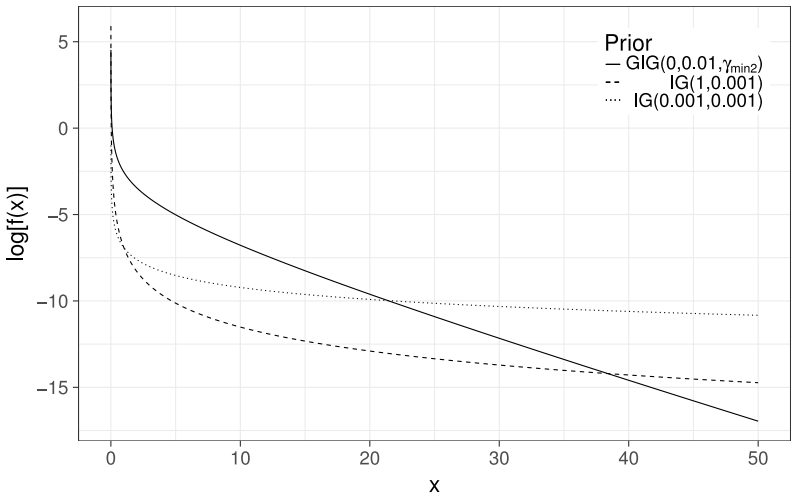

Fig. 1: Log density of the weakly informative GIG distributions proposed and of the most common vague inverse gamma priors.

Using a result from Fabrizi and Trivisano (2012) on the GIG distribution we have that

$$\mathbb{E}\left[\sigma^2|X\right] \cong \frac{\lambda + \sqrt{\lambda^2 + (nV^2 + \delta^2)\,\gamma^2}}{\gamma^2}. \tag{17}$$

By means of a first order expansion of the square root around a value c, implying $\sqrt{c^2+m} \cong |c| + \frac{m}{2|c|}$, we have that

$$\mathbb{E}\left[\sigma^2|X\right] \cong \frac{\left(nV^2 + \delta^2\right)}{-2\lambda + n - 1}. \tag{18}$$

If we set $\lambda = 0$ and $\delta = \varepsilon$ where a ε is a small number close to 0 (e.g. 0.01 or 0.001), then $\mathbb{E}\left[\sigma^2|X\right] \cong V^2$. Note that δ^2 can be be interpreted as the contribution of the prior to sample residual sum of square, so setting it to a small vale makes sense in a weakly informative setting.

The approximation (18) does not depend on γ. In view of (14), we propose to set γ so to guarantee the existence of the first two moments of $p(\theta_p|\bar{X}, V^2)$. For any positive real $\varepsilon > 0$, $\gamma = 2/\sqrt{n} + \varepsilon$ can be suitable. Nonetheless, very small ε can lead to numerical instability, especially for the variance (informally speaking, because of barely existing moments). We then suggest to set $\varepsilon = 1/\sqrt{n}$ leading to:

$$\gamma_{min2} = \frac{3}{\sqrt{n}}. \tag{19}$$

Note that, in view of (14), it is equivalent to imposing the existence of the first three moments of $p(\theta_p|\bar{X}, V^2)$. To sum up, our constrained weakly informative proposal for $(\lambda, \delta, \gamma)$ is given by $(\lambda = 0, \delta = \varepsilon, \gamma = \gamma_{min2})$. From figure 1 we can see how the proposed prior (with $n = 21$ in this case) has higher density than two inverse gammas with popular uninformative choices of the hyperparameters over a wide range of σ^2 values; the density becomes smaller in the right tail of the distribution, a feature that guarantees the existence of posterior moments. To compute the estimates under this prior setting it is required to specify the option `method='weak_inf'` in the `LN_quant()` function included in the `BayesLN` package.

4.2 A choice of hyperparameters minimizing frequentist MSE of Bayes estimators

In this section, in line with Rukhin (1986), we propose a choice of the hyperparameters aimed at minimizing the frequentist MSE of the estimators (15) and (16). Let's consider the first of the two, $\hat{\theta}_p^{QB}$: its frequentist MSE can be written as follows:

$$\begin{aligned}\mathbb{E}\left[\left(\hat{\theta}_p^{QB}-\theta_p\right)^2\right] &= \mathbb{E}\left[e^{2\left(w\bar{X}+(1-w)\xi_0\right)}g(V^2)^2-2\theta_p e^{w_0\bar{X}+(1-w)\xi_0}g(V^2)+\theta_p^2\right] \\ &= \theta_p^2\left[e^{2(1-w)(\xi_0-\xi)+\frac{2w^2\sigma^2}{n}-2\Phi^{-1}(p)\sigma}\times\right. \\ &\quad \times\mathbb{E}\left[\left(g(V^2)-e^{(1-w)(\xi_0-\xi)-\frac{3w^2\sigma^2}{2n}+\Phi^{-1}(p)\sigma}\right)^2\right]+ \\ &\quad \left.+1-e^{-\frac{w_0^2\sigma^2}{n}}\right].\end{aligned} \tag{20}$$

where $g(V^2)$ is implicitly defined by re-writing (15) as $\hat{\theta}_p^{QB}=e^{\xi_1}g(V^2)$. We note that $g(V^2)$ is the only part that depends on the GIG-hyperparameters (λ,δ,γ). As we have five hyperparameters $(\xi_0,n_0,\lambda,\delta,\gamma)$ the optimization problem is clearly over-parametrized.

A first step we take to simplify the problem is to set n_0 as in the weakly informative setting, a choice that protects from mis-specifications of ξ_0. Then the part of the MSE involving (λ,δ,γ) can be written as:

$$\mathbb{E}\left[\left(g(V^2)-\exp\left\{\Phi^{-1}(p)\sigma-\frac{3\sigma^2}{2n}\right\}\right)^2\right] \tag{21}$$

This quantity cannot be treated analytically because of the complicated mathematical expression of $g(V^2)$, involving infinite sums of Bessel K functions. For this reason we consider numerical optimization.

We note that (21) involves the unknown σ^2. When implementing numerical optimization it could be replaced by V^2 but this is not advisable, both to avoid the use of data in the prior specification process and to have the choice influenced by sampling variability that can be substantial in small samples. We propose to replace σ^2 with a guess s_0^2. Of course, the closer the guess is to σ^2 the better it is, but we found that solutions are quite insensitive to s_0^2 unless it is set much greater than the actual σ^2.

A global optimum for the hyperparameters (λ,δ,γ) cannot be found as it should be expected in view of the fact that we have a single functional and three hyperparameters. Fabrizi and Trivisano (2012), using an analytical approximation to the MSE of Bayes estimators of $\theta_{a,b}$, found a solution free of γ, in which λ could be expressed as a function of δ. Unfortunately, their approximation is not viable in our case.

We propose to fix two of the GIG hyperparameters and search the optimal value of the third. In the first place we set $\lambda=0$, again as in the non-informative setting. The reason is technical: this shape parameter appears in the order of the Bessel K functions and numerical optimization algorithms involving it are unstable.

After careful exploration of numerical optimization results, not reported here for brevity, we found that a strategy separating quantiles above and below the median is advisable. Specifically, when

- $p<0.5$: we fix γ to the minimum value that allows for the existence of the first two moments of $p(\theta_p|\bar{X},V^2)$ according to the (19) and find numerically an optimal value for δ;

- $p > 0.5$: fix δ and minimize with respect to γ. Recalling (18), it is possible to specify an informative value of δ, considering it as the contribution of a prior sample to the residual sum of squares. A general proposal could be $\delta = 1$: in most applied problems, values of the variance in the log scale σ^2 are seldom greater than 2, so 1 can be read as a reasonable guess for the size of an hypothetical deviation from the mean when $n = 1$. Of course, if the scale of the problem is totally different, the user can specify alternative values for δ; in any case we have evidence that $\delta = \varepsilon$ (i.e. a very small value) is not a good choice and leads to severely biased Bayes estimators of θ_p in small samples.

Heuristically, searching for optimal γ for quantiles above the median, and optimal δ for those below, is in line with the specialization of these parameters in the GIG distribution: γ rules the right tail of the distribution that is not relevant when $p < 0.5$, while δ is more involved with the general spread of the distribution and is therefore more relevant to the shape of the lower tail.

The case of the median ($p = 0.5$) is peculiar as $\theta_{0.5}$ does not depend on σ. If we try to minimize (21), the solution points in the direction $\lambda \to -\infty$ that implies $\hat{\theta}_p^{QB} \cong \exp(\bar{X})$, the naive estimator. To estimate $\theta_{0.5}$, the Bayes estimator under relative quadratic loss $\hat{\theta}_p^{RQB}$ is much more efficient, although its performances deteriorate quite fast when p moves away from 0.5. We suggest to consider $\hat{\theta}_{0.5}^{RQB}$ when estimating the median. In line with $\theta_{0.5}$ not depending on σ^2, the MSE is quite insensitive to the choice of $(\lambda, \delta, \gamma)$, so the parameters suggested for the weakly informative setting can be used.

The numerical optimization procedures required to specify the optimal proposed prior are automatically adopted when the option `method='optimal'` in the `LN_quant()` function included in the `BayesLN` package is used.

5 A simulation exercise

In this section we use a simulation exercise to assess the impact of prior parameter choices when the sample size is small. The aim of the simulation is twofold: first, to assess the frequentist properties of the posterior probability intervals when parameters are chosen according to the *weakly informative* strategy of section 4.1; second, to compare the frequentist MSE of the Bayes point predictor under the *MSE optimal* strategy of 4.2 with relevant alternatives proposed in the literature. For the latter case we evaluated also the performance of the posterior variance as a measure of the estimator uncertainty.

We considered also the Bayesian estimators obtained under inverse gamma priors for σ^2, in particular we simulated the frequentist properties of the credible intervals and point estimators when the classical small parameters inverse gamma $\sigma^2 \sim IG(0.001, 0.001)$ and a more informative setting $\sigma^2 \sim IG(2, 1)$ are assumed. We remark that in both cases the posterior moments of θ_p are not finite.

We note that, in the log-normal estimation context, MSE-optimality implies negative biased estimators, so the objectives of posterior intervals with good frequentist properties and optimality in terms of frequentist MSE are divergent targets when choosing hyperparameters. Of course, as the sample size grows, the impact of the hyperparameters choice become less and less relevant.

In our simulation we generate samples form log-normal distributions with mean 0 in the log-scale, i.e $Z \sim \mathcal{LN}(0, \sigma^2)$ and four different log-scale variances: $\sigma^2 = (0.25, 0.5, 1, 2)$. These values for the variance are the same that Longford (2012) considers. The results related to $n = (11, 21, 51)$ are reported for all the studies, with the exception of the interval estimation case in which the figures have the sample size in the abscissa and more values are considered. The same occurs for the quantile p in the MSE evaluation, the values $p = (0.05, 0.50, 0.95)$ are considered otherwise.

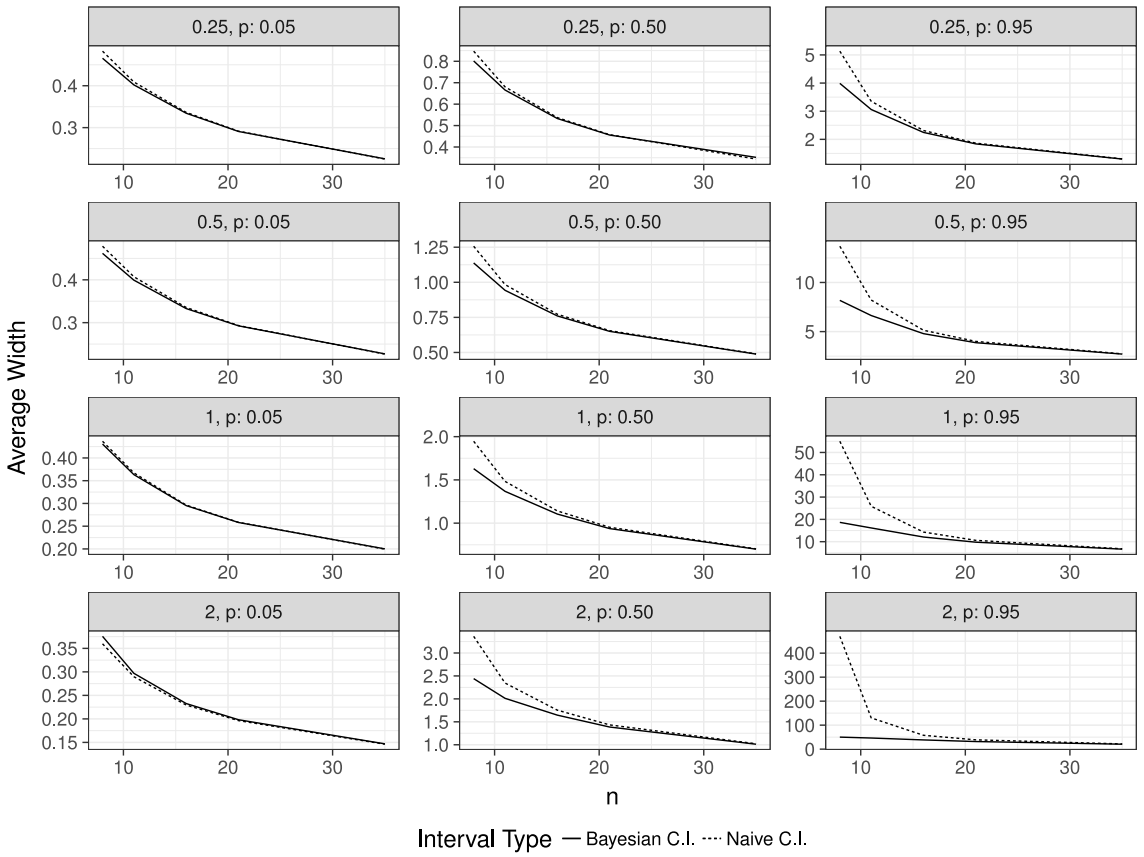

Fig. 2: Average width of the credible intervals obtained under the *weakly informative* choice of the hyperparameters for selected quantiles compared to the average width of the naive frequentist confidence intervals.

All the results we present are based on $B = 50,000$ samples, with the exception of the study about the intervals that is more computationally expensive and $B = 5,000$ replicates are considered. All the `R` code is available.

In figure 2 we explore the average width of the posterior probability intervals based on the *weakly informative* choice of hyperparameters. The intervals are defined by the 0.025 and 0.975 quantiles of the posterior distribution found computing the quantile function. They are compared to the naive intervals obtained by (3). A further comparison between the credible intervals obtained under the proposed GIG prior and the classical inverse gamma prior is reported in the appendix D in the supplementary material. There, the frequentist coverage is reported as well, and the proposed interval shows a better behaviour with respect to the inverse gamma case.

From figure 2, it is apparent that for quantiles in the right tail of the distribution and small sample sizes, the quantiles obtained using the GIG prior under *weakly informative* choice of the hyper-parameters achieve the nominal coverage with significantly narrower intervals; the larger σ, the more we gain in interval width; moreover the gain is apparent also for the median and not only for quantiles in the right tail. For the 0.05 quantile we did not find any relevant difference.

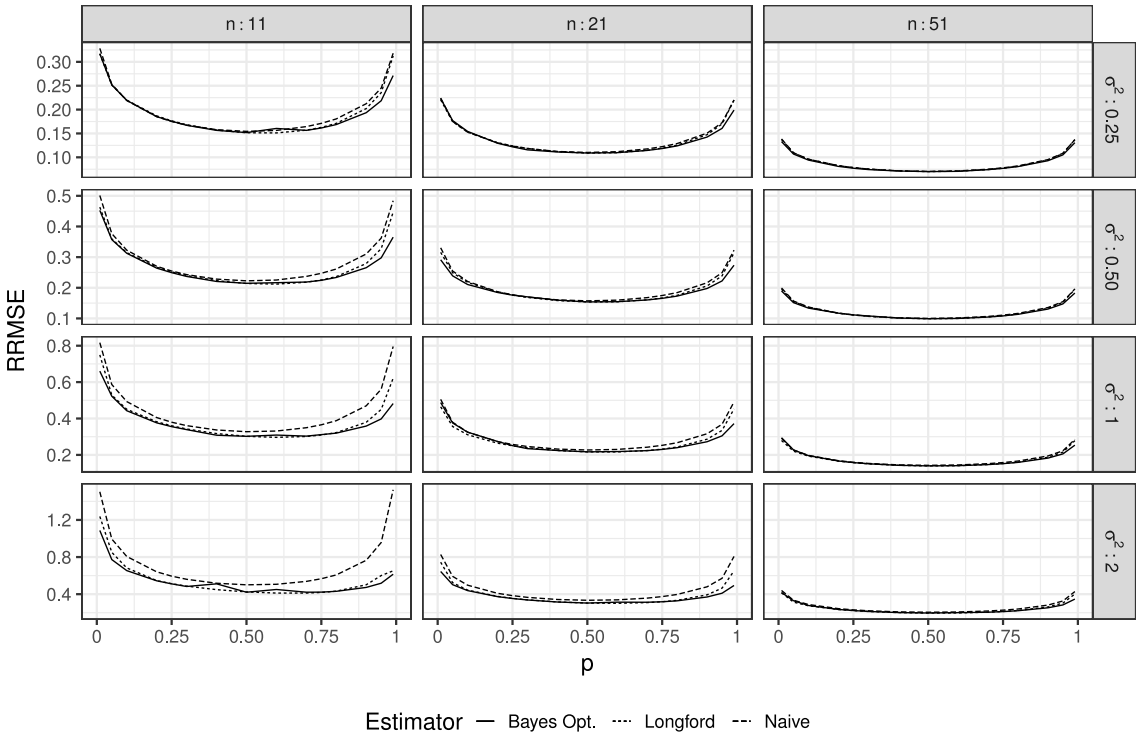

Fig. 3: Relative root mean square error of selected estimators: the Bayes predictor with MSE-minimizing hyper-parameters choice (Bayes Opt.), the predictor (4) (Longford) and the naive predictor (2) (Naive)

Then, we study the frequentist relative root mean squared error (RRMSE) of the Bayes estimator obtained under the choice of hyperparameters discussed in section 4.2. In the framework of this simulation, in line with Longford (2012), σ^2 has been replaced by S^2 when optimizing (21) to avoid subjective choices, even if we remark that the use of a guess of σ^2 is generally advisable. The aim is that of comparing the sampling RRMSE of (15) with those of alternatives from the frequentist literature such as (4). We note that in this case the predictor is obtained summarizing the posterior distribution using quadratic loss for all quantiles except the median for which we use the relative quadratic loss. Results are summarized in figure 3 where relative root mean square errors are plotted for ease of comparison.

From figure 3 we have that differences between alternative estimators decrease with the sample size, i.e. our choice of the hyper-parameters is less and less important as the sample size grows, as it should be expected. The Bayes predictor with optimal parameter choice has the lowest RRMSE in all cases with the only exception of the plot in bottom left plot (left tail of the distribution). In general, efficiency gains are larger for the right tail of the distribution which is often more relevant in scientific enquiries. The estimator we propose is clearly better than (2) in small samples and a little better of Longford's estimator (see 4) that is the most efficient estimator from the frequentist literature. The results about the posterior mean under inverse gamma priors are omitted here since extremely high RRMSE values are obtained for the quantiles in the right tail, probably due to the

Tab. 1: Monte Carlo standard deviation of the Bayes estimator with MSE optimal prior in the sample space ($\sqrt{V_{MC}}$) and square root of the mean of the posterior variance distribution ($\sqrt{\bar{V}_{QB}}$) at different n, σ^2, p.

	p:	0.05		0.50		0.95	
σ^2	n	$\sqrt{V_{MC}}$	$\sqrt{\bar{V}_{QB}}$	$\sqrt{V_{MC}}$	$\sqrt{\bar{V}_{QB}}$	$\sqrt{V_{MC}}$	$\sqrt{\bar{V}_{QB}}$
	11	0.108	0.104	0.149	0.174	0.484	0.440
0.25	21	0.076	0.075	0.108	0.117	0.358	0.333
	51	0.048	0.048	0.070	0.072	0.234	0.224
	11	0.109	0.101	0.207	0.255	0.905	0.872
0.5	21	0.076	0.074	0.151	0.169	0.683	0.662
	51	0.048	0.048	0.098	0.103	0.457	0.446
	11	0.097	0.086	0.285	0.384	1.884	1.918
1	21	0.066	0.063	0.210	0.249	1.478	1.487
	51	0.042	0.041	0.138	0.148	1.016	1.009
	11	0.075	0.065	0.385	0.619	4.545	4.901
2	21	0.048	0.045	0.288	0.381	3.793	4.008
	51	0.031	0.029	0.193	0.216	2.716	2.764

posterior moments infiniteness.

In table 1 the posterior standard deviations we can associate to $\hat{\theta}^{QB}$ are compared to the frequentist standard error of this predictor estimated empirically from the Monte Carlo experiment for three different quantiles. Posterior standard deviations track the standard errors fairly well. Nonetheless, posterior probability intervals for $\hat{\theta}^{QB}$ attain frequentist coverage below the nominal level because of the frequentist bias. This trade-off between MSE-optimality and bias is in line with results of previous literature (see Shen et al., 2006; Fabrizi and Trivisano, 2012).

6 Applications to real data sets

As anticipated, in different branch of environmental regulatory standard thresholds are frequently compared to percentiles and upper confidence limit, i.e. one sided confidence intervals around percentiles. Moreover, log-normality is often assumed in the analysis of hazardous materials concentration data and monitoring studies are often conducted with small sample sizes, so the methods we illustrated for point and interval estimation with a proper choice of hyper-parameters can be relevant.

In this section, the methodology described in the paper is applied to two real datasets taken from the literature, in which the estimation of extreme quantiles was of practical interest: the first is an application in the environmental health and monitoring framework, whereas the second one will focus on occupational health.

6.1 Example 1: environmental monitoring

In this section we reconsider the motivating example on the chrysene concentrations data from USEPA (2009). As discussed in the introduction, using data from background well at the site in question, the inferential goal is the estimation of the 95-th percentile ($\theta_{0.95}$) of the underlying

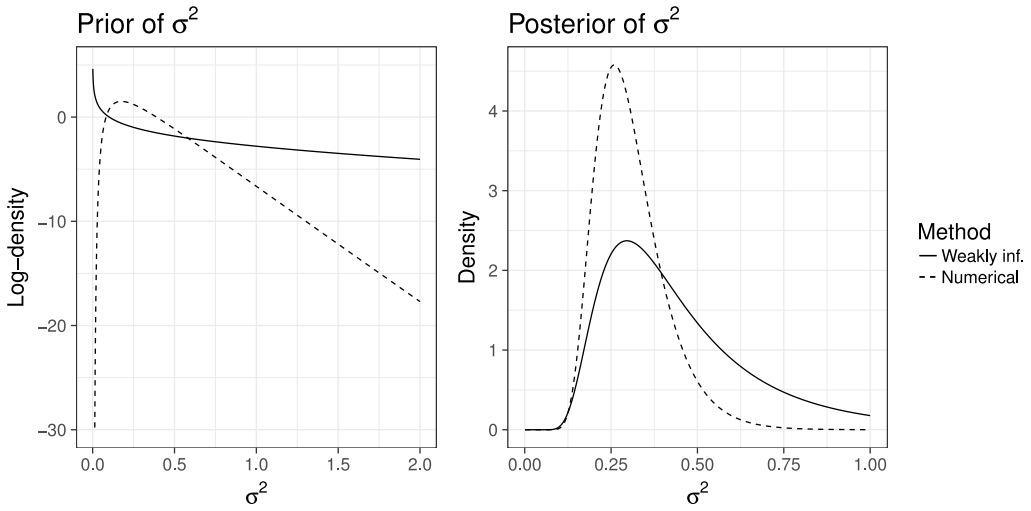

Fig. 4: Left panel: log-density of the prior on σ^2 specified according to the weakly informative choice and the MSE-optimizing ('Numerical') choice. Right panel: posterior density $p(\sigma^2|\bar{X}, V^2)$ under the two different choices for the prior on σ^2

distribution and the associated UCL, i.e. the threshold that includes the 95% of the distribution with the 95% confidence.

The main aim of this analysis is to compare the results obtained with the naive estimation procedures (discussed in section 2.2), which are currently mentioned in the guidelines, to the proposed Bayesian procedures that show better frequentist properties, according to the provided simulation study. More in detail, as point estimator we consider the Bayes estimator under the prior of section 4.2, whereas the weakly informative prior of section 4.1 is used to produce the credible interval to estimate the UCL. We considered also the efficient estimator proposed by Longford and described in section 2.3.

A first interesting point is the comparison between the two considered priors. The MSE-optimizing prior is given by $\sigma^2 \sim GIG(\lambda = 0, \delta = 1, \gamma = 4.61)$, with the last value obtained by numerical optimization. In the left panel of figure 4, this prior is compared to the weakly informative prior $\sigma^2 \sim GIG(\lambda = 0, \delta = 0.01, \gamma = 1.06)$, obtained considering $n = 8$ and $\gamma_{min2} = 3/\sqrt{n}$: we can easily note how the right tail of the weakly-informative prior is heavier, as expected.

The different choices for the prior on σ^2 are reflected into different posteriors $p(\sigma^2|\bar{X}, V^2)$: if a weakly informative choice is taken, the posterior will be more diffuse, thus giving more weight to large values of the estimand, while the one obtained by means of numerical optimization is more peaked around its mean and light-tail.

Alternative point estimates for the 95-th percentiles are displayed in table 2. We note that the most efficient estimator according to our simulations of section 5, i.e. $\hat{\theta}^{QBo}_{0.95}$ (that indicates the Bayes estimator under MSE-optimizing prior), provides the smallest point estimate of $\theta_{0.95}$; also $\hat{Q}_{0.95}$ produces a value smaller than the naive predictor $\hat{\theta}_{0.95}$. We stress that the latter method is recommended in many operational guidelines in environmental monitoring, although it turned out to be the less efficient in our simulation study of section 5. This ordering of the estimates is consistent with the fact that the most efficient estimators tend to be negatively biased. For the Bayes estimators, the posterior standard deviation was reported as estimate of the estimator standard error;

Tab. 2: Estimates of $\theta_{0.95}$ with different methods for example 1: naive ($\hat{\theta}_{0.95}$), Longford ($\hat{Q}_{0.95}$), our Bayes estimator under MSE-optimizing prior ($\hat{\theta}^{QBo}_{0.95}$). The estimates of the estimator standard error are reported too.

	$\hat{\theta}_{0.95}$	$\hat{Q}_{0.95}$	$\hat{\theta}^{QBo}_{0.95}$
Estimate	34.517	33.696	31.181
S.e.	-	8.056	8.257

whereas the squared root of the minimized MSE was reported for Longford's proposal, even if the author observed a severe underestimation of the true MSE.

As far as it concerns the estimation of the upper confidence limit, we compare our proposal based on the weakly informative prior specification (the one using numerical optimization for γ leads to intervals with frequentist coverage below the nominal level) to the method currently employed in the literature. According to the simulation study of section 5, our methodology leads to intervals with a shorter length on average with respect to the naive method recommended by EPA (see USEPA (2009) 17-17) although preserving the nominal coverage level. Specifically, substantive improvements can be observed with small samples for quantiles in the right tail. A smaller estimated UCL has the relevant consequence of defining a smaller threshold.

From an operational viewpoint, the estimated UCL is used as comparison term for observations coming from the wells to evaluate at the site. The contamination level is then determined according to the amount of water samples that exceeds the estimated threshold. Therefore, the proposed Bayes credible intervals represent a more powerful tool than the currently used method since it leads to a lower probability of rejecting the hypothesis that the site is uncontaminated when actually it is, because of its higher precision.

As it is possible to observe in figure 5, the value of our estimate is 76.195, whereas the estimate by EPA is 90.925: this can have relevant implications when the UCL is used to evaluate the compliance of pollutant's concentration future samples.

Eventually, we note that our estimates can be easily reproduced using the `R` package `BayesLN`, specifically running the following two commands:

```
LN_Quant(x = EPA09, quant = 0.95, method = "optimal", CI = F),
LN_Quant(x = EPA09, quant = 0.95, method = "weak_inf", type_CI = "UCL").
```

In this way, we hope to encourage practitioners to improve the statistical tools used in such delicate analysis.

6.2 Example 2: pollutant exposure assessment

To show a possible use of our methods in the occupational health field, a small dataset from the appendix *IV* of the book by Bullock and Ignacio (2006) is considered. It consists of $n = 10$ observations of exposures (ppm) of the coil feed operator and helper to Methyl Isobutyl Ketone (MIBK) during cleanup:

$$23,\ 42,\ 86,\ 62,\ 34,\ 107,\ 29,\ 65,\ 54,\ 55.$$

For these data, all the most popular normality tests do not reject the hypothesis of log-normality.

Occupational health is a particularly interesting field to apply the developed statistical methodology because Bayesian inference is already used and the log-normal distribution represents the default distributional assumptions for data as it is testified from the web application *Expostats* (Lavoué et al., 2019). Within this widget, largely useful for practitioners, Bayesian methods are employed and an

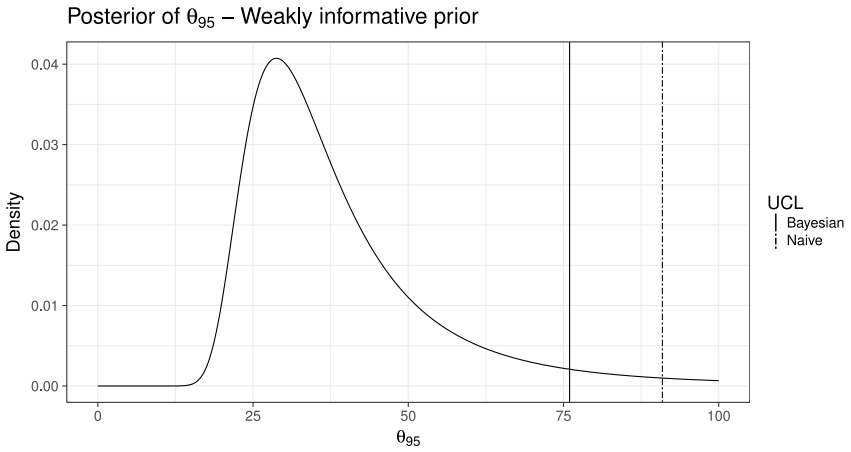

Fig. 5: Posterior distribution of θ_{95} with weakly informative prior and different estimates of the UCL for the data of example 1.

informative prior on σ is suggested; its specification is based on estimated variances $\hat{\sigma}_e^2$ from past exposure studies reported in Kromhout et al. (1993). Specifically, the authors proposed to fit a log-normal distribution on these estimated variances, yielding the prior $\sigma \sim \mathcal{LN}(-0.17, 0.39)$, and thereby $\sigma^2 \sim \mathcal{LN}(-0.34, 1.16)$.

We decided to incorporate the same information but considering our GIG prior setting, in order to obtain a posterior distribution of θ_p with finite moments, since it can be proved that the log-normal prior on σ does not preserve their existence (see appendix C). To do so, we fit a GIG distribution by maximum likelihood on $\hat{\sigma}_e^2$ using the routine `gigFit` from the `GeneralizedHyperbolic` package (Scott, 2018). We obtained the prior $\sigma^2 \sim GIG(0.29, 0.59, 0.98)$ and the existence condition for moments up to the second order is fulfilled for $n > 9$, otherwise we would have replaced the γ obtained in fitting the value of γ with $\gamma_{min2} = 3/\sqrt{n}$ that is in this case slightly lower.

The GIG prior is compared to the histogram of $\hat{\sigma}_e^2$ in figure 6. It can be noted that the higher flexibility of the GIG distribution can be useful in the specification of informative priors, whereas the log-normal distribution shows some difficulties in capturing the peak near to 0 and keep the right weight in the right tail. Therefore, the GIG prior we suggest can provide the basis for the specification of informative priors for the σ^2 parameter in the occupational health context. Moreover, the GIG prior guarantees the existence of the posterior moments for θ_p. Eventually, another appealing property of the proposed prior is the fact that conditions for the existence of the posterior moments can be easily worked out also in presence of censored data that quite common in occupational health. Details about these conditions can be found in appendix B.

For the data introduced at the beginning of this section, let's consider the problem of estimating $\theta_{0.95}$ and the related UCL, in order to compare them to the MIBK short term exposure limit fixed at 75 ppm. The posterior summaries of $\theta_{0.95}$ obtained under the two informative priors described above are reported in table 3. Because of the lighter tail of the log-normal prior, the posterior median obtained under this prior are slightly lower than the ones obtained under the informative GIG prior. Note that, differently from the case of the log-normal prior posterior moments are a well defined and represent meaningful summaries of the posterior distribution, and for this reason displayed in table 3.

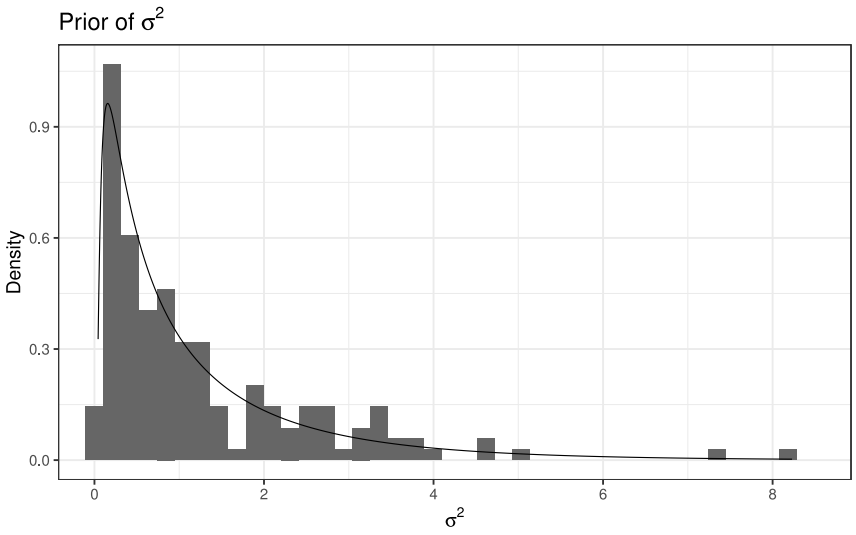

Fig. 6: Prior distribution obtained fitting a GIG distribution on $\hat{\sigma}_e^2$. The histogram of $\hat{\sigma}_e^2$ is reported too.

Tab. 3: Summaries of the posterior distribution of $\theta_{0.95}$ under the GIG prior on σ^2 and a log-normal on σ.

Prior	Post. median	Post. mean	Post. s.d.	UCL
GIG	121.17	132.20	57.00	223.41
Log-normal	120.67	-	-	219.94

7 Concluding remarks

In this paper we have considered the popular log-normal distribution and the problem of estimating its quantiles. We assumed a GIG prior for the variance on the log-scale as the GIG encompasses many popular distributions (e.g. inverse gamma, gamma) as special or limiting cases. We show that the existence of the posterior moments for the log-normal quantiles depend on the choice of one of the GIG hyperparameters and that priors of common use such as *small parameters* inverse gammas lead to posterior with non-existent expectations and not only in small samples.

A careful choice of the GIG hyper-parameters can improve inference for log-normal quantiles with respect to current practice; specifically, it can lead to shorter interval estimators, although keeping frequentist nominal coverage.

Our results can be extended in several directions. In the first place, to the regression case where we assume an heterogeneous population in which for the i-th unit in the population $X_i \sim \mathcal{N}(\mathbf{a}_i^T \beta, \sigma^2)$. The problem of estimating the mean of the $Z = \exp(X)$ conditionally on a given point of the covariate space $\mathbf{a}_0$ has been already considered in Fabrizi and Trivisano (2016); extension to the estimation of quantiles of this distribution can be obtained using the methodology illustrated in this paper. In this way, a quantile regression procedure with an underlying parametric assumption could be implemented. Another interesting development might be represented by the investigation of the result in appendix B about the censored data case, extending the finding to the mean estimation problem too.

8 Supporting information

Additional information for this article is available online. Technical details and theorems proves are contained in appendix A, some results about the censored data case are reported in appendix B, remarks on the log-normal prior for σ are in appendix C whereas additional figures are in appendix D. Moreover, the `R` code required to reproduce the results is provided.

References

N. Balakrishnan and D. Mitra. Likelihood inference for lognormal data with left truncation and right censoring with an illustration. *Journal of Statistical Planning and Inference*, 141(11):3536–3553, 2011.

O. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 353 (1674):401–419, 1977.

M. Bengtsson, A. Ståhlberg, P. Rorsman, and M. Kubista. Gene expression profiling in single cells from the pancreatic islets of langerhans reveals lognormal distribution of mrna levels. *Genome research*, 15(10):1388–1392, 2005.

B. M. Bibby and M. Sørensen. Hyperbolic processes in finance. *Handbook of heavy tailed distributions in finance*, 1:211–248, 2003.

W. H. Bullock and J. S. Ignacio. *A strategy for assessing and managing occupational exposures*. AIHA, 2006.

E. Crow and K. Shimizu. *Lognormal distributions: Theory and practice*. Marcel Decker, 1988.

E. Fabrizi and C. Trivisano. Bayesian estimation of log-normal means with finite quadratic expected loss. *Bayesian Analysis*, 7(4):975–996, 2012.

E. Fabrizi and C. Trivisano. Bayesian conditional mean estimation in log-normal linear regression models with finite quadratic expected loss. *The Scandinavian Journal of Statistics*, 43(4):1064–1077, 2016.

D. Finney. On the distribution of a variate whose logarithm is normally distributed. *Supplement to the Journal of the Royal Statistical Society*, 7(2):155–161, 1941.

A. Gardini, E. Fabrizi, and C. Trivisano. *BayesLN: Bayesian Inference for Log-Normal Data*, 2020. URL `https://CRAN.R-project.org/package=BayesLN`. R package version 0.1.2.

A. Gelman. Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian analysis*, 1(3):515–534, 2006.

R. D. Gibbons, D. K. Bhaumik, and S. Aryal. *Statistical methods for groundwater monitoring*, volume 59. John Wiley & Sons, 2009.

I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Academic press, 2014.

E. Gumbel. La probabilité des hypotheses. *Comptes Rendus de l'Académie des Sciences (Paris)*, 209:645–647, 1939.

R. J. Hyndman and Y. Fan. Sample quantiles in statistical packages. *The American Statistician*, 50 (4):361–365, 1996.

A. Ignatov, B. Holben, and T. Eck. The lognormal distribution as a reference for reporting aerosol optical depth statistics; empirical tests using multi-year, multi-site aeronet sunphotometer data. *Geophysical Research Letters*, 27(20):3333–3336, 2000.

B. Jorgensen. *Statistical properties of the generalized inverse Gaussian distribution*, volume 9. Springer, 1982.

K. Kosugi. Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resources Research*, 32(9):2697–2703, 1996.

K. Krishnamoorthy, A. Mallick, and T. Mathew. Inference for the lognormal mean and quantiles based on samples with left and right type i censoring. *Technometrics*, 53(1):72–83, 2011.

H. Kromhout, E. Symanski, and S. M. Rappaport. A comprehensive evaluation of within-and between-worker components of occupational exposure to chemical agents. *The Annals of occupational hygiene*, 37(3):253–270, 1993.

J. Lavoué, L. Joseph, P. Knott, H. Davies, F. Labrèche, F. Clerc, G. Mater, and T. Kirkham. Expostats: a bayesian toolkit to aid the interpretation of occupational exposure measurements. *Annals of work exposures and health*, 63(3):267–279, 2019.

J. F. Lawless. *Statistical models and methods for lifetime data*. John Wiley & Sons, 2003.

E. Limpert, W. A. Stahel, and M. Abbt. Log-normal distributions across the sciences: Keys and clues. *AIBS Bulletin*, 51(5):341–352, 2001.

N. T. Longford. Small-sample estimators of the quantiles of the normal, log-normal and pareto distributions. *Journal of Statistical Computation and Simulation*, 82(9):1383–1395, 2012.

J. H. May, D. P. Strum, and L. G. Vargas. Fitting the lognormal distribution to surgical procedure times. *Decision Sciences*, 31(1):129–148, 2000.

S. P. Millard. *EnvStats, an R Package for Environmental Statistics*. Wiley Online Library, 2013.

M. S. Paolella. *Intermediate probability: A computational approach*. John Wiley & Sons, 2007.

A. L. Rukhin. Improved estimation in lognormal models. *Journal of the American Statistical Association*, 81(396):1046–1049, 1986.

D. Scott. *GeneralizedHyperbolic: The Generalized Hyperbolic Distribution*, 2018. URL `https://CRAN.R-project.org/package=GeneralizedHyperbolic`. R package version 0.8-4.

H. Shen, L. D. Brown, and H. Zhi. Efficient estimation of log-normal means with application to pharmacokinetic data. *Statistics in medicine*, 25(17):3023–3038, 2006.

L. Thabane and M. S. Haq. Prediction from a normal model using a generalized inverse gaussian prior. *Statistical Papers*, 40(2):175–184, 1999.

USEPA. Statistical analysis of groundwater monitoring data at rcra facilities: Unified guidance. Technical report, Office of Resource Conservation and Recovery, Program Implementation and Information Division, U.S. Environmental Protection Agency, Washington, D.C., 2009.

A. Zellner. Bayesian and non-bayesian analysis of the log-normal distribution and log-normal regression. *Journal of the American Statistical Association*, 66(334):327–330, 1971.