

# 两个独立对数正态分布的中位数比的统计推断

## ——基于似然方法、广义枢轴量方法和贝叶斯方法

仲卿照

山东大学数学学院, 山东 济南  
Email: zhongqing@mail.sdu.edu.cn

收稿日期: 2021年1月10日; 录用日期: 2021年2月12日; 发布日期: 2021年2月19日

### 摘要

本文主要研究了基于朴素似然法、两种广义枢轴量法、带Diffuse先验和独立Jeffreys先验的贝叶斯方法对两个独立对数正态分布中位数之比的统计推断。我们通过模拟比较了五种置信区间在覆盖率、平均长度和相对偏度方面的性能。结果表明, 两种广义置信区间和基于Diffuse先验的贝叶斯方法在覆盖率上普遍较优。而基于似然方法和独立Jeffreys先验的置信区间的平均长度较短。此外, 给出了参数的极大似然估计和贝叶斯后验估计。同时, 给出了单边假设检验的广义p值和后验概率比。并以PM2.5的实际数据为例, 说明了该方法的统计推断。

### 关键词

对数正态分布, 极大似然估计, 广义置信区间, 广义P-值, 贝叶斯推断, 数值模拟

# Inferences for the Ratio of Medians of Two Independent Log-Normal Distributions

## —Likelihood-Based, Generalized Pivot Quantity and Bayesian Approaches

Qingzhao Zhong

School of Mathematics, Shandong University, Jinan Shandong  
Email: zhongqing@mail.sdu.edu.cn

Received: Jan. 10<sup>th</sup>, 2021; accepted: Feb. 12<sup>th</sup>, 2021; published: Feb. 19<sup>th</sup>, 2021

### Abstract

In this paper, we concentrate on statistical inferences for the ratio of medians of two independent log-normal distributions based on naive likelihood approach, generalized variable approach with

two generalized pivot quantities (GPQ) and Bayesian approach with Diffuse prior and Independence Jeffreys' prior. We compare the performance of the five confidence intervals in terms of the coverage probabilities, average length and relative bias by using simulations. The results show that the Bayesian approach based on Diffuse prior and two kinds of generalized confidence intervals are generally preferred in terms of coverage probability. However, confidence intervals based on likelihood and Independence Jeffreys' prior have shorter average length. We also give the maximum likelihood estimator and Bayesian posterior estimation of estimated parameter. Simultaneously, the generalized  $p$ -values and posterior probability ratio for a one-sided hypothesis test are proposed. The statistical inference is illustrated using a real data example about PM<sub>2.5</sub>.

## Keywords

Log-Normal, MLE, Generalized Confidence Intervals, Generalized P-Values, Bayesian Inference, Numerical Simulation

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在实际应用中,许多测量或调查的数据往往是带有偏度的,如医学、空气污染、金融、经济等。在科学研究中,这类数据一般假定为对数正态模型。对数正态分布在许多应用中是普遍而有效的,对数正态分布的参数估计是一个非常有趣的问题。关于对数正态分布参数统计推断的研究很多,如均值、多个对数正态分布的公共均值、公共变异系数等。例如,Weerahandi [1]、Tsui 和 Weerahandi [2]介绍了广义变量法,包括假设检验问题的广义置信区间和广义  $p$ -值的构造,成为应用于对数正态分布的有力工具;Tian 和 Wu [3],Behboodian 和 Jafari [4]用广义变量方法对几个对数正态总体的公共均值进行了推断;W. H. Wu 和 H. N. Hsieh [5]构造了关于  $\delta$  对数正态分布均值的广义置信区间;K. Krishnamoorthy, T. Mathew [6]利用广义  $p$ -值和广义置信区间对对数正态分布的均值进行统计推断;J. Hannig, H. Iyer 和 P. Patterson [7]介绍了基准广义置信区间;Thangjai、Niwitpong 和 Niwitpong [8]建立了公共变异系数的区间估计;Zhou 和 Tu [9]用似然法和 bootstrap 法建立了对数正态分布均值比的区间估计。近几十年来,贝叶斯方法在统计推断中得到了广泛的应用,尤其是对数正态分布。K. Aruna Rao 和 J. G. D'Cunha [10]构建了对数正态分布中位数的贝叶斯推断;Maneerat、Niwitpong 和 Niwitpong [11]构建了对数正态分布的单均值和两均值之差的贝叶斯置信区间;Harvey 和 Merwe [12]构建了对数正态分布和二元对数正态分布的均值和方差的贝叶斯置信区间;Zeller [13]讨论了均值对数正态分布的贝叶斯分析与非贝叶斯分析的比较。

然而,带有偏度的对数正态分布通常具有极低的可度量性,这对中位数的影响要小于均值。因此,在这种情况下,中位数是一个比均值更有意义的集中趋势的度量。Rao 和 D'Cunha [8]利用四个先验密度提出了贝叶斯后验估计和贝叶斯可信区间。Zeller [13]讨论了对数正态分布中值的贝叶斯估计。就我们所知,关于两个独立对数正态分布的中位数之比统计推断的研究论文很少。本文利用似然法、广义变量法和贝叶斯方法对两个对数正态总体中位数之比进行了统计推断。此外,我们根据覆盖率、平均长度和相对偏度评估所构建的置信区间的性能。各节内容如下:

第二节主要介绍了对数正态模型,以及构造最大似然估计和置信区间的方法。第三节提出了基于两种不同的广义枢轴量的广义置信区间和广义  $p$ -值的假设检验。在第四节中,我们用 Diffuse 先验和独立

Jeffreys 先验和模拟算法给出了贝叶斯后验点估计、可信区间和后验概率比检验。第五节给出了衡量置信区间优良的数值模拟, 包括讨论和结果。在第六节中, 我们将用一个真实的数据来说明所提出的推断方法。最后的结束语在第七节给出。

## 2. 似然方法

考虑两个独立的对数正态分布, 其参数为  $(\mu_i, \sigma_i^2), i=1, 2$ 。令  $X_{ij}; i=1, 2, j=1, \dots, n_i$  是来自两个独立总体的随机样本, 则有:

$$X_{ij} \sim LN(\mu_i, \sigma_i^2)$$

其概率密度函数为:

$$f(X) = \frac{1}{x_{ij} \sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x_{ij} - \mu_i}{\sigma_i}\right)^2\right\}$$

其中  $x_{ij} > 0, \mu_i \in \mathbb{R}, \sigma_i^2 > 0, i=1, 2, j=1, \dots, n_i$ 。

根据数理统计的知识我们不难得到:  $Y_{ij} = \ln(X_{ij}) \sim N(\mu_i, \sigma_i^2)$ , 参数  $\mu_i, \sigma_i^2, i=1, 2$  的极大似然估计 (MLE) 为:

$$\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}, S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$$

对于两个独立对数正态分布, 其中位数之比为:

$$\theta = \frac{\theta_1}{\theta_2} = \frac{e^{\mu_1}}{e^{\mu_2}} = e^{\mu_1 - \mu_2}$$

这里的  $\theta$  就是本文关心的统计量。由极大似然估计的不变性, 我们运用“插入法(Plug-in)”可以得到参数  $\theta$  的一个点估计为:

$$\hat{\theta}_{MLE} = e^{\bar{y}_1 - \bar{y}_2}$$

这里  $\bar{y}_1, \bar{y}_2$  分别是  $\bar{Y}_1, \bar{Y}_2$  的观测值。下面我们考虑参数  $\theta$  的置信区间的构造, 由渐进正态理论和 Delta 方法, 当样本量趋于无穷时, 估计  $\hat{\theta}_{MLE}$  具有渐进正态性, 其渐进均值和方差为:

$$\mu_\theta = e^{\mu_1 - \mu_2}, \sigma_\theta = e^{2(\mu_1 - \mu_2)} \left( \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)$$

由于参数  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$  是未知的, 我们同样可以利用“插入法”, 以  $\bar{y}_1, \bar{y}_2, s_1^2, s_2^2$  来代替未知参数, 即  $\bar{Y}_1, \bar{Y}_2, S_1^2, S_2^2$  的观测值。

因此, 给定显著性水平  $\alpha, (\alpha = 0.05)$ , 则参数  $\theta$  的  $(1-\alpha)100\%$  双边渐进置信区间为:

$$CL_L = \left[ e^{\bar{y}_1 - \bar{y}_2} - z_{\alpha/2} \sqrt{e^{2(\bar{y}_1 - \bar{y}_2)} \left( \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)}, e^{\bar{y}_1 - \bar{y}_2} + z_{\alpha/2} \sqrt{e^{2(\bar{y}_1 - \bar{y}_2)} \left( \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)} \right]$$

其中,  $z_{\alpha/2}$  为标准正态分布的  $\alpha/2$  分位数。

## 3. 广义枢轴量方法

假设  $Y = (Y_1, Y_2, \dots, Y_n)'$  是来自一个总体的随机样本, 该总体取决于参数  $(\theta, \vartheta)$ , 其中  $\theta$  是我们感兴趣的参数而  $\vartheta$  是讨厌参数。广义枢轴量  $R(Y; y, \theta, \vartheta)$  是随机样本  $Y$ , 观测值  $y$ , 和参数  $(\theta, \vartheta)$  的函数, 其具有以下性质:

- $R(Y; y, \theta, \vartheta)$  的分布与未知参数无关;

- 广义枢轴量的观测值  $r = R(y; y, \theta, \mathcal{G})$  不取决于讨厌参数  $\mathcal{G}$ 。

则参数  $\theta$  的广义枢轴量可以定义为:

$$R_{\theta}^1 = \exp\{R_{\mu_1}^1 - R_{\mu_2}^1\}$$

其中  $R_{\mu_i}^1$  定义为:

$$R_{\mu_i}^1 = \bar{y}_i - \frac{\bar{Y}_i - \mu_i}{S_i / \sqrt{n_i}} \frac{s_i}{\sqrt{n_i}} = \bar{y}_i - \frac{Z_i}{\sqrt{n_i}} \frac{s_i \sqrt{n_i - 1}}{U_i}$$

其中  $Z_i, U_i$  相互独立, 为:

$$Z_i = \frac{\bar{Y}_i - \mu_i}{\sigma_i / \sqrt{n_i}} \sim N(0, 1), U_i^2 = \frac{(n_i - 1) S_i^2}{\sigma_i^2} \sim \chi_{n_i - 1}^2$$

不难说明,  $R_{\theta}^1$  满足上述两条性质, 因此, 参数  $\theta$  的  $(1 - \alpha)100\%$  双边广义置信区间为:

$$CI_{G1} = [R_{\theta}^1(\alpha/2), R_{\theta}^1(1 - \alpha/2)]$$

其中,  $R_{\theta}^1(\alpha/2)$  与  $R_{\theta}^1(1 - \alpha/2)$  分别为  $R_{\theta}^1$  的  $100(\alpha/2)$  分位数和  $100(1 - \alpha/2)$  分位数。

除了上面的广义枢轴量的构造方式之外, 我们给出另外一种构造方式:

$$R_{\theta}^2 = \exp\{R_{\mu_1}^2 - R_{\mu_2}^2\}$$

在这种情况下,  $R_{\mu_i}^2$  定义为:

$$R_{\mu_i}^2 = \bar{Y}_i - \frac{S_i}{S_i^*} (\bar{Y}_i^* - \mu_i)$$

其中,  $\bar{Y}_i^*$  和  $S_i^{*2}$  分别为  $\bar{Y}_i$  和  $S_i^2$  的独立复制。则参数  $\theta$  的  $(1 - \alpha)100\%$  双边广义置信区间为:

$$CI_{G2} = [R_{\theta}^2(\alpha/2), R_{\theta}^2(1 - \alpha/2)]$$

其中,  $R_{\theta}^2(\alpha/2)$  与  $R_{\theta}^2(1 - \alpha/2)$  分别为  $R_{\theta}^2$  的  $100(\alpha/2)$  分位数和  $100(1 - \alpha/2)$  分位数。

基于广义枢轴量而构建广义置信区间可以总结为如下算法:

- 给定  $\bar{y}_i, s_i^2, i = 1, 2$ 。
- 从  $j = 1$  至  $m$  循环  
生成随机变量  $Z_i, U_i^2$  并计算  $R_{\theta}^1$ ;  
生成随机变量  $\bar{Y}_i^*, S_i^{*2}$  并计算  $R_{\theta}^2$ ;
- 结束循环;
- 计算  $R_{\theta}^1(\alpha/2)$  与  $R_{\theta}^1(1 - \alpha/2)$  为  $CI_{G1}$  上下界,  $R_{\theta}^2(\alpha/2)$  与  $R_{\theta}^2(1 - \alpha/2)$  为  $CI_{G2}$  上下界。

下面我们考虑假设检验问题:

$$H_0: \theta \leq \theta_0 \text{ v.s. } H_1: \theta > \theta_0$$

接下来, 我们给出广义检验统计量和广义  $p$ -值的概念。广义检验统计量  $T(Y; y, \theta, \mathcal{G})$  是样本  $Y$ , 观测值  $y$ , 和参数  $(\theta, \mathcal{G})$  的函数, 满足如下性质:

- 给定观测值  $y$ ,  $T(Y; y, \theta, \mathcal{G})$  的分布与讨厌参数无关;
- 对于  $Y = y$ ,  $T(Y; y, \theta, \mathcal{G})$  的观测值与未知参数无关;
- 给定  $y, \mathcal{G}$ ,  $P(T(Y; y, \theta, \mathcal{G}) \geq t | \theta)$  是关于  $\theta$  单调的。

如果  $T(Y; y, \theta, \mathcal{G})$  关于  $\theta$  是随机增长的, 则广义  $p$ -值定义为  $P(T(Y; y, \theta, \mathcal{G}) \geq T(y; y, \theta, \mathcal{G}) | \theta_0)$ ; 反之, 如果  $T(Y; y, \theta, \mathcal{G})$  关于  $\theta$  是随机下降的, 则广义  $p$ -值定义为  $P(T(Y; y, \theta, \mathcal{G}) \leq T(y; y, \theta, \mathcal{G}) | \theta_0)$ 。

为了记号的方便, 我们令  $\lambda = \ln(\theta) = \mu_1 - \mu_2$ , 那么假设检验就变成了

$$H_0 : \lambda \leq \lambda_0 \text{ v.s. } H_1 : \lambda > \lambda_0$$

下面我们给出广义检验统计量的具体形式:

$$T(Y; y, \lambda) = R_{\mu_1}^1 - R_{\mu_2}^1 - \lambda$$

不难验证, 按照上述定义的广义检验统计量满足以上的三个条件, 且这时该检验的广义  $p$ -值可以定义为:

$$G_p = P(T(Y; y, \lambda) \leq T(y; y, \lambda) | \lambda = \lambda_0) = P(R_{\mu_1}^1 - R_{\mu_2}^1 \leq \lambda_0) \triangleq P(R_\lambda \leq \lambda_0)$$

如果广义  $p$ -值低于一个显著性水平  $\alpha$ , 则拒绝原假设, 反之, 拒绝备择假设。对于一个复杂的分布, 直接计算广义  $p$ -值是一项十分复杂的工作, 因此我们通常借助模拟的方式来估计广义  $p$ -值。

下面给出计算广义  $p$ -值的算法流程:

- 给定  $\bar{y}_i, s_i^2, i=1, 2$ 。
- 从  $j=1$  至  $M$  循环
  - 生成随机变量  $Z_i \sim N(0,1)$  以及  $U_i^2 \sim \chi_{n_i-1}^2$ ;
  - 计算  $R_\lambda = R_{\mu_1}^1 - R_{\mu_2}^1$ ;
  - 比较  $R_\lambda$  和  $\lambda_0$  的大小, 如果  $R_\lambda$  小于  $\lambda_0$ , 记“1”, 否则记“0”;
- 结束循环
- 广义  $p$ -值  $G_p$  的估计为:

$$\frac{\#\text{“1”的个数}\#}{M}$$

#### 4. 贝叶斯方法

本节我们将利用贝叶斯方法给出待估参数  $\theta$  的点估计、可信区间和后验概率比检验。我们主要考虑两种不同的先验:

- Diffuse 先验:  $\pi^1(\mu_1, \mu_2, \sigma_1, \sigma_2) = 1/(\sigma_1\sigma_2)$
- 独立 Jeffreys 先验:  $\pi^2(\mu_1, \mu_2, \sigma_1, \sigma_2) = 1/(\sigma_1\sigma_2)^2$

根据贝叶斯公式, 基于上述两种先验的后验密度分别为:

$$\begin{aligned} \pi^1(\mu_1, \mu_2, \sigma_1, \sigma_2 | Y) &\propto \pi^1(\mu_1, \mu_2, \sigma_1, \sigma_2) L(\mu_1, \mu_2, \sigma_1, \sigma_2 | Y) \\ &\propto \sigma_1^{-1} \sigma_1^{-n_1} \exp\left\{-\frac{1}{2\sigma_1^2} \left[ (n_1-1)S_1^2 + n_1(\mu_1 - \bar{Y}_1)^2 \right]\right\} \\ &\quad \times \sigma_2^{-1} \sigma_2^{-n_2} \exp\left\{-\frac{1}{2\sigma_2^2} \left[ (n_2-1)S_2^2 + n_2(\mu_2 - \bar{Y}_2)^2 \right]\right\} \\ &\propto N\left(\bar{Y}_1, \frac{\sigma_1^2}{n_1} \mid \mu_1, \sigma_1^2\right) \times \text{IGamma}\left(\frac{n_1-2}{2}, \frac{(n_1-1)S_1^2}{2} \mid \sigma_1^2\right) \\ &\quad \times N\left(\bar{Y}_2, \frac{\sigma_2^2}{n_2} \mid \mu_2, \sigma_2^2\right) \times \text{IGamma}\left(\frac{n_2-2}{2}, \frac{(n_2-1)S_2^2}{2} \mid \sigma_2^2\right) \end{aligned}$$

同样地,

$$\begin{aligned}
\pi^2(\mu_1, \mu_2, \sigma_1, \sigma_2 | Y) &\propto \pi^2(\mu_1, \mu_2, \sigma_1, \sigma_2) L(\mu_1, \mu_2, \sigma_1, \sigma_2 | Y) \\
&\propto \sigma_1^{-2} \sigma_1^{-n_1} \exp\left\{-\frac{1}{2\sigma_1^2} \left[(n_1-1)S_1^2 + n_1(\mu_1 - \bar{Y}_1)^2\right]\right\} \\
&\quad \times \sigma_2^{-2} \sigma_2^{-n_2} \exp\left\{-\frac{1}{2\sigma_2^2} \left[(n_2-1)S_2^2 + n_2(\mu_2 - \bar{Y}_2)^2\right]\right\} \\
&\propto N\left(\bar{Y}_1, \frac{\sigma_1^2}{n_1} \mid \mu_1, \sigma_1^2\right) \times \text{IGamma}\left(\frac{n_1-1}{2}, \frac{(n_1-1)S_1^2}{2} \mid \sigma_1^2\right) \\
&\quad \times N\left(\bar{Y}_2, \frac{\sigma_2^2}{n_2} \mid \mu_2, \sigma_2^2\right) \times \text{IGamma}\left(\frac{n_2-1}{2}, \frac{(n_2-1)S_2^2}{2} \mid \sigma_2^2\right)
\end{aligned}$$

我们不难发现，后验分布其实是四个密度函数的乘积，其中

$$\sigma_i^2 \sim \text{IGamma}\left(\frac{n_i-2}{2}, \frac{(n_i-1)S_i^2}{2}\right), \mu_i \mid \sigma_i^2 \sim N\left(\bar{Y}_i, \frac{\sigma_i^2}{n_i}\right), i=1,2$$

以及

$$\sigma_i^2 \sim \text{IGamma}\left(\frac{n_i-1}{2}, \frac{(n_i-1)S_i^2}{2}\right), \mu_i \mid \sigma_i^2 \sim N\left(\bar{Y}_i, \frac{\sigma_i^2}{n_i}\right), i=1,2$$

那么，待估参数  $\theta$  的贝叶斯后验估计为：

$$\hat{\theta}_{BS} = E(e^{\mu_1 - \mu_2} \mid Y)$$

下面我们通过最高后验密度(HPD)的方法构建贝叶斯可信区间，因此，关于  $\pi^1(\mu_1, \mu_2, \sigma_1, \sigma_2 \mid Y)$  和  $\pi^2(\mu_1, \mu_2, \sigma_1, \sigma_2 \mid Y)$  的  $(1-\alpha)100\%$  的双边贝叶斯可信区间为：

$$CI_{BS1} = [L_{BS1}, U_{BS1}], CI_{BS2} = [L_{BS2}, U_{BS2}]$$

其中， $L_{BS1}$  和  $L_{BS2}$  是  $\hat{\theta}_{BS}$  关于后验密度  $\pi^1, \pi^2$  的最短最高后验密度区间的下极限，同样， $U_{BS1}$  和  $U_{BS2}$  是上极限。贝叶斯后验估计的计算本质上是积分的计算，同样是一件复杂的工作，因此，我们需要通过模拟的方法进行估计以及构建可信区间。

下面我们给出构建贝叶斯可信区间和后验估计的算法流程：

- 给定  $\bar{y}_i, s_i^2, i=1,2$ 。
- 从  $j=1$  至  $m$  循环
  - 关于后验  $\pi^1$  生成  $\sigma_i^2 \sim \text{IGamma}(\bullet)$  以及  $\mu_i \mid \sigma_i^2 \sim N(\bullet)$ ；
  - 关于后验  $\pi^2$  生成  $\sigma_i^2 \sim \text{IGamma}(\bullet)$  以及  $\mu_i \mid \sigma_i^2 \sim N(\bullet)$ ；
  - 关于后验  $\pi^1(\bullet \mid Y)$  计算  $\hat{\theta}_{BS1}^j = e^{\mu_1 - \mu_2}$ ；
  - 关于后验  $\pi^2(\bullet \mid Y)$  计算  $\hat{\theta}_{BS2}^j = e^{\mu_1 - \mu_2}$ ；
- 结束循环
- 贝叶斯后验密度估计为：

$$\hat{\theta}_{BS1} = \frac{1}{m} \sum_{j=1}^m \hat{\theta}_{BS1}^j, \hat{\theta}_{BS2} = \frac{1}{m} \sum_{j=1}^m \hat{\theta}_{BS2}^j$$

- 计算  $L_{BS1}, U_{BS1}, L_{BS2}$  以及  $U_{BS2}$ 。
- 我们继续考虑假设检验

$$H_0: \theta \leq \theta_0 \text{ v.s. } H_1: \theta > \theta_0$$

则参数  $\theta$  的后验概率比为:

$$\frac{\alpha_0}{\alpha_1} = \frac{\iint_{\theta \leq \theta_0} \pi^i(\mu_1, \mu_2, \sigma_1, \sigma_2 | Y) d\mu_1 d\mu_2}{\iint_{\theta > \theta_0} \pi^i(\mu_1, \mu_2, \sigma_1, \sigma_2 | Y) d\mu_1 d\mu_2}, \quad i = 1, 2$$

如果后验概率比  $\alpha_0/\alpha_1 > 1$ , 则接受原假设, 如果后验概率比  $\alpha_0/\alpha_1 < 1$ , 则拒绝原假设, 如果  $\alpha_0/\alpha_1 \approx 1$ , 则我们需要更多的后验信息。

接下来给出计算后验概率比的算法流程:

- 给定  $\bar{y}_i, s_i^2, i = 1, 2$ 。
- 从  $j = 1$  至  $m$  循环
  - 关于后验  $\pi^1$  生成  $\sigma_i^2 \sim \text{IGamma}(\bullet)$  以及  $\mu_i | \sigma_i^2 \sim N(\bullet)$ ;
  - 关于后验  $\pi^2$  生成  $\sigma_i^2 \sim \text{IGamma}(\bullet)$  以及  $\mu_i | \sigma_i^2 \sim N(\bullet)$ ;
  - 关于后验  $\pi^1(\bullet | Y)$  计算  $\hat{\theta}_{BS1}^j = e^{\mu_1 - \mu_2}$ ;
  - 关于后验  $\pi^2(\bullet | Y)$  计算  $\hat{\theta}_{BS2}^j = e^{\mu_1 - \mu_2}$ ;
  - 比较  $\hat{\theta}_{BSi}^j$  和  $\theta_0$  大小, 如果  $\hat{\theta}_{BSi}^j \leq \theta_0$ , 令  $a_i^j = 1$ , 否则  $b_i^j = 1, i = 1, 2$ 。
- 结束循环
- 贝叶斯后验概率比为:

$$\frac{\alpha_0}{\alpha_1} = \frac{\sum_{j=1}^m a_i^j}{\sum_{j=1}^m b_i^j}, \quad \text{关于 } \pi^i(\mu_1, \mu_2, \sigma_1, \sigma_2 | Y), \quad i = 1, 2$$

## 5. 数值模拟

在本节中, 我们利用 RStudio 来评估所构建的五种置信区间( $CI_L, CI_{G1}, CI_{G2}, CI_{BS1}$  和  $CI_{BS2}$ )的表现。我们提出以下的指标来衡量区间的表现:

- 覆盖率(CP): 模拟的置信区间覆盖住  $\theta$  的比例;
- 平均长度(AL): 模拟置信区间的平均长度;
- 相对偏度(RB):

$$\text{RB} = \frac{|(\text{upper error rate}) - (\text{lower error rate})|}{|(\text{upper error rate}) + (\text{lower error rate})|}$$

其中, “lower error rate” 为  $\theta$  低于区间下极限的次数的比例, 同样 “upper error rate” 可类似定义。在模拟学习中, 我们选择显著性水平  $\alpha = 0.05$ 。对于任何参数配置, 我们分别从两个独立对数正态总体中生成 5000 组随机样本来构建 95% 置信区间。我们选取样本量大小  $(n_1, n_2)$  分布为  $(20, 20), (40, 40), (60, 60), (20, 60), (20, 100), (40, 100), (60, 100), (100, 100)$ 。所有模拟结果均被记录在表 1、表 2 和表 3 中。表 1 中的参数配置为  $(\mu_1, \mu_2) = (1, 1), (\mu_1, \mu_2) = (3, 1)$  以及  $(\sigma_1, \sigma_2) = (0.1, 0.1)$ , 表 2 的参数配置为  $(\mu_1, \mu_2) = (1, 1), (\mu_1, \mu_2) = (3, 1)$  以及  $(\sigma_1, \sigma_2) = (0.1, 0.3)$ , 表 3 的参数配置为  $(\mu_1, \mu_2) = (1, 1), (\mu_1, \mu_2) = (3, 1)$  以及  $(\sigma_1, \sigma_2) = (0.3, 0.3)$ 。

**Table 1.** Simulation results of Coverage Probability, Average Length and Relative Bias of CIs where (Sigma1, Sigma2) = (0.1, 0.1)**表 1.**覆盖率、平均长度和相对偏度的模拟结果

		样本量(均值参数都为 1)							
		(20, 20)	(40, 40)	(60, 60)	(20, 60)	(20, 100)	(40, 100)	(60, 100)	(100, 100)
$CI_L$	CP	0.9446	0.9472	0.9526	0.9358	0.9370	0.9422	0.9492	0.9416
	AL	0.1234	0.0873	0.0714	0.1004	0.0948	0.0730	0.0639	0.0553
	RB	0.1311	0.2273	0.1477	0.1526	0.0349	0.2042	0.1102	0.0205
$CI_{G1}$	CP	0.9586	0.9512	0.9578	0.9454	0.9476	0.9480	0.9526	0.9422
	AL	0.1315	0.0900	0.0729	0.1060	0.1003	0.0749	0.0650	0.0560
	RB	0.0239	0.1121	0.0685	0.0706	0.0781	0.0866	0.0167	0.0345
$CI_{G2}$	CP	0.9586	0.9512	0.9578	0.9454	0.9476	0.9480	0.9526	0.9422
	AL	0.1315	0.0899	0.0728	0.1059	0.1003	0.0750	0.0650	0.0560
	RB	0.0242	0.1475	0.0711	0.0476	0.1069	0.1385	0.0295	0.0865
$CI_{BS1}$	CP	0.9626	0.9524	0.9548	0.9486	0.9520	0.9476	0.9472	0.9410
	AL	0.1336	0.0899	0.0725	0.1071	0.1015	0.7470	0.0645	0.0555
	RB	0.0481	0.1597	0.1239	0.0739	0.1000	0.1450	0.0985	0.0102
$CI_{BS2}$	CP	0.9540	0.9474	0.9550	0.9422	0.9430	0.9446	0.9504	0.9400
	AL	0.1296	0.0888	0.0719	0.1044	0.0988	0.0739	0.0641	0.0553
	RB	0.0522	0.2167	0.1467	0.0796	0.0456	0.1769	0.0484	0.0333
		样本量(均值参数分别为 3 和 1)							
		(20, 20)	(40, 40)	(60, 60)	(20, 60)	(20, 100)	(40, 100)	(60, 100)	(100, 100)
$CI_L$	CP	0.9360	0.9422	0.9476	0.9400	0.9402	0.9454	0.9426	0.9480
	AL	0.9110	0.6469	0.5279	0.7426	0.7034	0.5396	0.4718	0.4094
	RB	0.1313	0.0519	0.0992	0.1467	0.1037	0.1136	0.0244	0.1846
$CI_{G1}$	CP	0.9496	0.9500	0.9510	0.9488	0.9510	0.9482	0.9478	0.9496
	AL	0.9706	0.6671	0.5387	0.7834	0.7447	0.5536	0.4798	0.4143
	RB	0.0480	0.0418	0.0588	0.0902	0.0159	0.0403	0.0530	0.0939
$CI_{G2}$	CP	0.9496	0.9500	0.9510	0.9488	0.9510	0.9482	0.9478	0.9496
	AL	0.9703	0.6674	0.5385	0.7837	0.7444	0.5541	0.4799	0.4144
	RB	0.0159	0.0080	0.0122	0.0469	0.0286	0.0502	0.0498	0.1032
$CI_{BS1}$	CP	0.9550	0.9456	0.9496	0.9510	0.9512	0.9472	0.9440	0.9468
	AL	0.9856	0.6667	0.5365	0.7919	0.7531	0.5520	0.4772	0.4112
	RB	0.0667	0.0221	0.0873	0.1184	0.1311	0.0833	0.0714	0.1654
$CI_{BS2}$	CP	0.9482	0.9432	0.9482	0.9472	0.9742	0.9470	0.9430	0.9468
	AL	0.9560	0.6568	0.5315	0.7724	0.7333	0.5460	0.4733	0.4085
	RB	0.0811	0.0211	0.0733	0.0909	0.0909	0.0491	0.0175	0.1353



**Table 2.** Simulation results of Coverage Probability, Average Length and Relative Bias of CIs where (Sigma1, Sigma2) = (0.1, 0.3)**表 2.** 覆盖率、平均长度和相对偏度的模拟结果

		样本量(均值参数都为 1)							
		(20, 20)	(40, 40)	(60, 60)	(20, 60)	(20, 100)	(40, 100)	(60, 100)	(100, 100)
$CI_L$	CP	0.9414	0.9454	0.9410	0.9490	0.9476	0.9434	0.9494	0.9460
	AL	0.2750	0.1951	0.1595	0.1745	0.1464	0.1327	0.1279	0.1238
	RB	0.3311	0.2601	0.1458	0.3020	0.0763	0.1519	0.0514	0.1556
$CI_{G1}$	CP	0.9536	0.9502	0.9464	0.9550	0.9534	0.9452	0.9508	0.9488
	AL	0.2945	0.2017	0.1630	0.1802	0.1511	0.1351	0.1297	0.1255
	RB	0.0862	0.1148	0.0308	0.0933	0.0815	0.0076	0.0855	0.0077
$CI_{G2}$	CP	0.9536	0.9502	0.9464	0.9550	0.9534	0.9452	0.9508	0.9488
	AL	0.2947	0.2016	0.1630	0.1802	0.1510	0.1350	0.1296	0.1255
	RB	0.0603	0.1165	0.0373	0.1022	0.0815	0.0219	0.0894	0.0156
$CI_{BS1}$	CP	0.9558	0.9516	0.9448	0.9542	0.9510	0.9432	0.9494	0.9456
	AL	0.2981	0.2015	0.1620	0.1800	0.1510	0.1341	0.1286	0.1243
	RB	0.1674	0.2149	0.0580	0.2751	0.0531	0.0493	0.9190	0.0441
$CI_{BS2}$	CP	0.9516	0.9476	0.9416	0.9510	0.9500	0.9430	0.9484	0.9406
	AL	0.2895	0.1986	0.1606	0.1774	0.1488	0.1330	0.1278	0.1236
	RB	0.2479	0.1985	0.0548	0.2000	0.0320	0.0877	0.0233	0.0909
		样本量(均值参数分别为 3 和 1)							
		(20, 20)	(40, 40)	(60, 60)	(20, 60)	(20, 100)	(40, 100)	(60, 100)	(100, 100)
$CI_L$	CP	0.9366	0.9434	0.9420	0.9426	0.9500	0.9474	0.9498	0.9496
	AL	2.0281	1.4394	1.1799	1.2914	1.0807	0.9803	0.9444	0.9516
	RB	0.1861	0.3004	0.2276	0.2544	0.0960	0.0875	0.1554	0.1032
$CI_{G1}$	CP	0.9506	0.9518	0.9500	0.9494	0.9564	0.9508	0.9510	0.9518
	AL	2.1724	1.4894	1.2068	1.3343	1.1149	0.9975	0.9582	0.9282
	RB	0.1120	0.1322	0.0569	0.0909	0.1121	0.0909	0.0252	0.0041
$CI_{G2}$	CP	0.9506	0.9518	0.9500	0.9494	0.9564	0.9508	0.9510	0.9518
	AL	2.1724	1.4883	1.2063	1.3330	1.1144	0.9973	0.9587	0.9281
	RB	0.0607	0.0871	0.0480	0.1067	0.0642	0.0813	0.0286	0.0622
$CI_{BS1}$	CP	0.9540	0.9500	0.9478	0.9488	0.9572	0.9500	0.9516	0.9482
	AL	2.2009	1.4862	1.1997	1.3334	1.1151	0.9902	0.9493	0.9197
	RB	0.0783	0.2000	0.1111	0.1719	0.0093	0.0480	0.0248	0.0656
$CI_{BS2}$	CP	0.9494	0.9472	0.9434	0.9454	0.9510	0.9484	0.9474	0.9474
	AL	2.1366	1.4647	1.1886	1.3147	1.0989	0.9826	0.9444	0.9147
	RB	0.1225	0.1894	0.1661	0.1868	0.0041	0.0388	0.0722	0.0951

**Table 3.** Simulation results of Coverage Probability, Average Length and Relative Bias of CIs where (Sigma1, Sigma2) = (0.3, 0.3)

**表 3.** 覆盖率、平均长度和相对偏度的模拟结果

		样本量(均值参数都为 1)							
		(20, 20)	(40, 40)	(60, 60)	(20, 60)	(20, 100)	(40, 100)	(60, 100)	(100, 100)
$CI_L$	CP	0.9434	0.9432	0.9492	0.9466	0.9402	0.9464	0.9420	0.9520
	AL	0.3710	0.2628	0.2145	0.3022	0.2852	0.2200	0.1922	0.1662
	RB	0.3852	0.2887	0.2441	0.1835	0.3378	0.2090	0.2000	0.2333
$CI_{G1}$	CP	0.9570	0.9542	0.9530	0.9586	0.9516	0.9492	0.9470	0.9522
	AL	0.3975	0.2717	0.2194	0.3202	0.3028	0.2263	0.1959	0.1685
	RB	0.0741	0.0169	0.0209	0.1429	0.0947	0.0040	0.0039	0.0087
$CI_{G2}$	CP	0.9570	0.9542	0.9530	0.9586	0.9516	0.9492	0.9470	0.9522
	AL	0.3976	0.2716	0.2192	0.3198	0.3028	0.2263	0.1959	0.1685
	RB	0.0791	0.0131	0.0043	0.1787	0.1074	0.0236	0.0189	0.0209
$CI_{BS1}$	CP	0.9558	0.9494	0.9518	0.9556	0.9540	0.9502	0.9438	0.9494
	AL	0.4020	0.2710	0.2178	0.3226	0.3055	0.2253	0.1944	0.1668
	RB	0.2579	0.1858	0.1286	0.0090	0.2000	0.1165	0.0890	0.1621
$CI_{BS2}$	CP	0.9516	0.9488	0.9508	0.9538	0.9516	0.9496	0.9428	0.9480
	AL	0.3900	0.2671	0.2160	0.3141	0.2973	0.2227	0.1931	0.1659
	RB	0.2314	0.1518	0.1301	0.0216	0.2314	0.0952	0.1399	0.1231
		样本量(均值参数分别为 3 和 1)							
		(20, 20)	(40, 40)	(60, 60)	(20, 60)	(20, 100)	(40, 100)	(60, 100)	(100, 100)
$CI_L$	CP	0.9428	0.9484	0.9498	0.9378	0.9378	0.9424	0.9414	0.9458
	AL	2.7432	1.9434	1.5887	2.2276	2.1139	1.6230	1.4140	1.2275
	RB	0.4266	0.2713	0.2908	0.2219	0.3248	0.3194	0.2696	0.1661
$CI_{G1}$	CP	0.9536	0.9566	0.9540	0.9486	0.9536	0.9518	0.9436	0.9490
	AL	2.9379	2.0071	1.6245	2.3588	2.4667	1.6678	1.4412	1.2440
	RB	0.0667	0.0456	0.0610	0.1000	0.0169	0.0857	0.0786	0.0078
$CI_{G2}$	CP	0.9536	0.9566	0.9540	0.9486	0.9536	0.9518	0.9436	0.9490
	AL	2.9398	2.0096	1.6240	2.3597	2.2432	1.6689	1.4392	1.2449
	RB	0.0776	0.0415	0.1043	0.0817	0.0172	0.0622	0.0426	0.0039
$CI_{BS1}$	CP	0.9548	0.9548	0.9506	0.9482	0.9562	0.9456	0.9444	0.9474
	AL	2.9755	2.0041	1.6131	2.3777	2.2674	1.6627	1.4285	1.2322
	RB	0.3009	0.1239	0.1903	0.0734	0.1781	0.1691	0.0426	0.1103
$CI_{BS2}$	CP	0.9510	0.9520	0.9512	0.9454	0.9472	0.9450	0.9412	0.9480
	AL	2.8830	1.9762	1.5984	2.3174	2.2054	1.6425	1.4182	1.2258
	RB	0.2408	0.0917	0.2377	0.0623	0.1894	0.2145	0.1633	0.1538

## 模拟结果分析

基于表 1、表 2 和表 3:

- 若考虑“覆盖率”，基于两种广义枢轴量以及 Diffuse 先验所构造的置信区间要更优于其他两种方法的置信区间。对于小样本和中等样本，基于 Diffuse 先验的贝叶斯可信区间表现更好，然而，两种广义置信区间更适用于大样本。
- 若考虑“平均长度”，基于似然方法构造的置信区间要明显优于其他几种，但无论参数的配置是怎样的，通常伴随着较低的覆盖率。
- 若考虑“覆盖率”和“平均长度”，两种广义枢轴量构造的广义置信区间的表现几乎一致，但在  $\mu_1 = \mu_2$  的情况下，第二种广义置信区间的“相对偏度”较大，但当  $\mu_1 \neq \mu_2$  时，其“相对偏度”会较低。
- 关于两种贝叶斯方法，基于 Diffuse 先验的贝叶斯可信区间在“覆盖率”之下是表现更好的，但通常区间长度会更长，因此，若要得到更“短”的置信区间，基于独立 Jeffreys 先验的贝叶斯可信区间是更适合的。
- 对于相对较大的  $(\sigma_1, \sigma_2)$  或者  $(\sigma_1, \sigma_2)$  相差较大的情形，我们构造的置信区间通常伴随着明显较宽的区间长度，这种情形在  $\mu_1 \neq \mu_2$  时尤为明显。

总而言之，对于小样本或中等样本，基于 Diffuse 先验的贝叶斯方法值得被推荐，对于大样本，两种广义枢轴量构造的置信区间更适合。但若想要得到较短的区间，也就是在考虑“平均长度”的情形下，似然方法以及基于独立 Jeffreys 先验的贝叶斯方法表现更好。

## 6. 实例分析

大气细颗粒物(PM2.5)是指直径小于或等于 2.5 微米、可以长时间悬浮在空气中的颗粒物。其主要化学成分为碳氢化合物、氮氧化物、硫氧化物、金属离子化合物、卤素化合物。它可以作为反应床，进行各种反应，将初级污染物转化为次级污染物。它们来源广泛，随着时间的推移，空间分布也各不相同。由于其粒径小，重量轻，长期漂浮，可被吸入肺部，对人体健康造成极大危害。PM2.5 主要来自汽车尾气、燃煤和工业废气，在交通发达、人口众多的发达城市尤为严重。

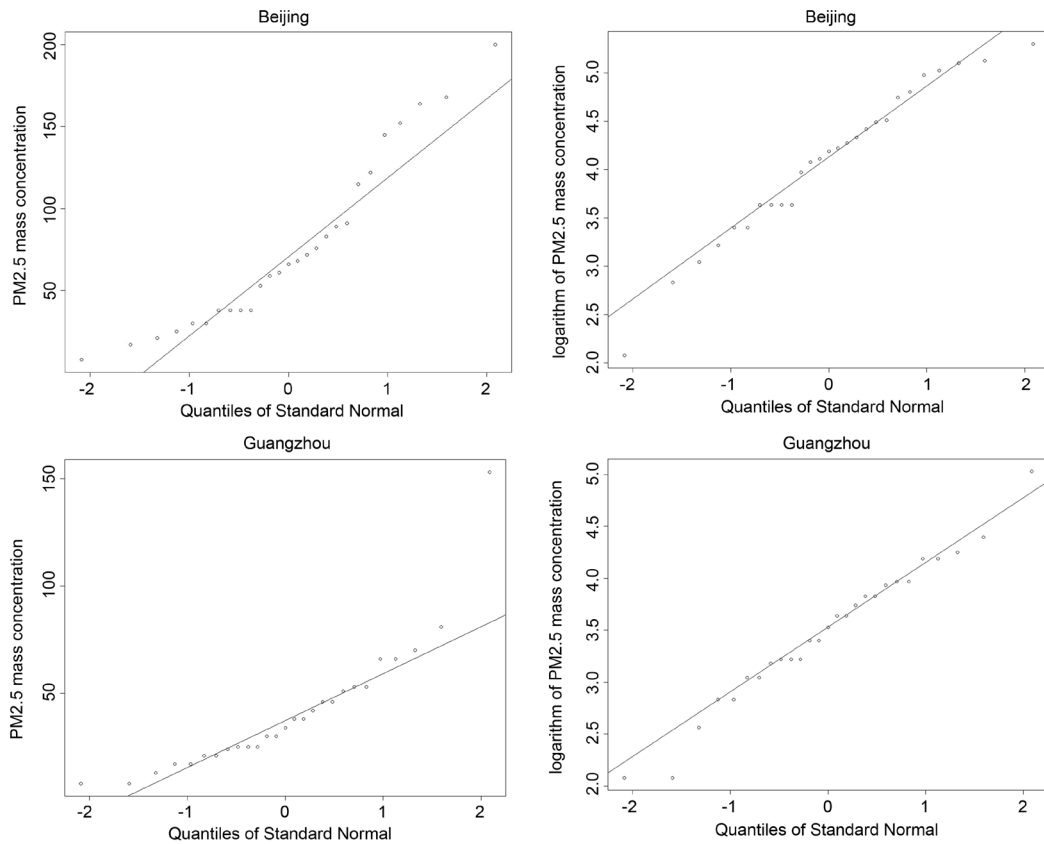
在本例子中，我们研究了北京和广州的 PM2.5 质量浓度，这两个城市在中国都是比较发达的城市，交通流量大，但就经验而言，位于南海沿岸的广州，空气质量总体上比北京好。因此我们会用提出的方法比较两个区域测量的中位数之比。我们在 2020 年 7 月 10 日至 2020 年 10 月 25 日每隔 4 天早上 8:00(北京时间)同步记录 PM2.5 质量浓度实时数据，数据来源：<http://www.young-0.com/airquality/>。

我们记录到的数据如下：

- 北京 PM2.5 ( $\text{ug}/\text{m}^3$ ): 91, 83, 66, 122, 38, 168, 68, 72, 89, 115, 38, 145, 152, 30, 53, 30, 25, 8, 38, 76, 164, 17, 38, 200, 21, 61, 59;

广州 PM2.5 ( $\text{ug}/\text{m}^3$ ): 46, 46, 13, 17, 30, 25, 25, 66, 30, 53, 8, 24, 34, 53, 153, 21, 17, 25, 8, 38, 21, 70, 38, 81, 51, 42, 66。

图 1 显示了两个城市的原始数据和经过对数变换之后数据的 QQ 图(其中，左上和左下两张图是原始数据，右上和右下是经过对数变换的)。这些图表明 PM2.5 质量浓度的分布是正偏态的，但经对数变换后的图像近似对称，这直观地说明了对数变换后的数据具有正态性。经对数变换后两市的正态性 Shapiro-Wilk 检验的 p-值分别为 0.5283 和 0.8618，两市原始数据的 p-值分别为 0.02777 和 0.0002015。因此，有充分的证据表明北京和广州的 PM2.5 质量浓度分别服从对数正态分布。



**Figure 1.** QQ-plots of PM2.5 mass concentration in Beijing and Guangzhou  
**图 1.** 北京和广州两市 PM2.5 浓度的 QQ 图

在表 4 中，我们给出了 PM2.5 质量浓度数据对数变换后的汇总统计。在表 5 中，我们根据提出的方法给出了  $\theta$  的 95% 双边置信区间。基于似然法  $\theta$  的点估计分别为 1.8057、基于贝叶斯后验密度  $\pi^1, \pi^2$  的点估计分别为 1.8533、1.8475。现在考虑假设检验问题，令  $\theta_0 = 1$ ，我们分别给出广义 p-值 0.0038，以及基于后验密度  $\pi^1, \pi^2$  的后验概率比 0.004419446 和 0.002004008。因此，我们需要拒绝原假设，接受备择假设。这一结果可以解释为 2020 年 7 月 10 日至 10 月 25 日广州 PM2.5 质量浓度中值低于北京。这也意味着这段时间广州的空气质量比北京好，这与事实相符。

**Table 4.** The summary statistics of log-transformed data both in Beijing and Guangzhou

**表 4.** PM2.5 质量浓度数据对数变换后的统计量汇总

城市	样本量	均值	标准差
北京	27	4.0813	0.7836
广州	27	3.4903	0.6840

**Table 5.** The 95% two-sided confidence intervals for theta

**表 5.** 待估参数  $\theta$  的 95% 的双边置信区间

方法	置信区间	区间长度
$CI_L$	[1.0973, 2.5142]	1.4169
$CI_{G1}$	[1.2225, 2.6700]	1.4475

## Continued

$CI_{G_2}$	[1.1893, 2.7036]	1.5143
$CI_{BS1}$	[1.0786, 2.6151]	1.5364
$CI_{BS2}$	[1.1314, 2.5558]	1.4244

## 7. 结束语

本文的主要任务是构造基于似然、广义枢轴量和贝叶斯方法的置信区间。通过数值模拟，我们发现基于两种广义枢轴量和贝叶斯后验密度  $\pi^1$  的置信区间在覆盖率方面较好，而基于似然的置信区间由于渐近理论在大样本中表现较好。从广义枢轴量的性质来看，它的构造不是唯一的，由此得出广义置信区间的构造也不是唯一的。实际上，目前我们还不知道哪一种方法是构造广义置信区间的最优方法。与贝叶斯方法类似，如何选择先验也是一个永恒的话题。因此，这需要进一步的研究。

## 致 谢

非常感谢为本文做出贡献的老师、同学们还有朋友们，以及感谢为本文提供数据的网站：<http://www.young-0.com/airquality/>。

## 参考文献

- [1] Weerahandi, S. (1993) Generalized Confidence Intervals. *Journal of the American Statistical Association*, **88**, 899-905. <https://doi.org/10.1080/01621459.1993.10476355>
- [2] Tsui, K.W. and Weerahandi, S. (1989) Generalized p-Values in Significance Testing of Hypotheses in the Presence of Nuisance Parameters. *Journal of the American Statistical Association*, **84**, 602-607. <https://doi.org/10.1080/01621459.1989.10478810>
- [3] Tian, L. and Wu, J. (2007) Inferences on the Common Mean of Several Log-Normal Populations: The Generalized Variable Approach. *Biometrical Journal*, **49**, 944-951. <https://doi.org/10.1002/bimj.200710391>
- [4] Wu, W.H. and Hsieh, H.N. (2014) Generalized Confidence Interval Estimation for the Mean of Delta-Lognormal Distribution: An Application to New Zealand Trawl Survey Data. *Journal of Applied Statistics*, **41**, 1471-1485. <https://doi.org/10.1080/02664763.2014.881780>
- [5] Krishnamoorthy, K. and Mathew, T. (2003) Inferences on the Means of Lognormal Distributions Using Generalized p-Values and Generalized Confidence Intervals. *Journal of Statistical Planning and Inference*, **115**, 103-121. [https://doi.org/10.1016/S0378-3758\(02\)00153-2](https://doi.org/10.1016/S0378-3758(02)00153-2)
- [6] Behboodian, J. and Jafari, A.A. (2006) Generalized Inference for the Common Mean of Several Lognormal Populations. *Journal of Statistical Theory and Applications*, 240-259.
- [7] Hannig, J., Iyer, H. and Patterson, P. (2006) Fiducial Generalized Confidence Intervals. *Journal of the American Statistical Association*, **101**, 254-269. <https://doi.org/10.1198/016214505000000736>
- [8] Thangjai, W., Nimitpong, S.A. and Niwitpong, S. (2020) Confidence Intervals for the Common Coefficient of Variation of Rainfall in Thailand. *PeerJ*, **8**, e10004. <https://doi.org/10.7717/peerj.10004>
- [9] Zhou, X.H. and Tu, W. (2000) Interval Estimation for the Ratio in Means of Log-Normally Distributed Medical Costs with Zero Values. *Computational Statistics & Data Analysis*, **35**, 201-210. [https://doi.org/10.1016/S0167-9473\(00\)00009-8](https://doi.org/10.1016/S0167-9473(00)00009-8)
- [10] Rao, K.A. and D'Cunha, J.G. (2016) Bayesian Inference for Median of the Lognormal Distribution. *Journal of Modern Applied Statistical Methods*, **15**, 526-535. <https://doi.org/10.22237/jmasm/1478003400>
- [11] Maneerat, P., Nimitpong, S.A. and Niwitpong, S. (2019) Bayesian Confidence Intervals for a Single Mean and the Difference between Two Means of Delta-Lognormal Distributions. *Communications in Statistics-Simulation and Computation*. <https://doi.org/10.1080/03610918.2019.1616095>
- [12] Harvey, T. and van der Merwe, A.J. (2012) Bayesian Confidence Intervals for Means and Variances of Lognormal and Bivariate Lognormal Distributions. *Journal of Statistical Planning Inference*, **142**, 1294-1309.

<https://doi.org/10.1016/j.jspi.2011.12.006>

- [13] Zellner, A. (1971) Bayesian and Non-Bayesian Analysis of the Log-Normal Distribution and Log-Normal Regression. *Journal of the American Statistical Association*, **66**, 327-330. <https://doi.org/10.1080/01621459.1971.10482263>