STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, December 2025, pp 2977–2999.
Published online in International Academic Press (www.IAPress.org)

Using Miss-specification Effect for Selection between Inverse Weibull and Lognormal Distributions

Mohammad M Saber[1,*], Parnian Habibi[2], Mohammad Hossein Zarinkolah[3], Abdussalam Aljadani[4], Mahmoud M. Mansour[5,6], Mohamed S. Hamed[6,7], Haitham M yousof[6]

[1]*Department of Statistics, Higher Education Center of Eghlid, Eghlid, Iran*
[2]*Department of Applied Mathematics, Tarbiat Modares University, Tehran, Iran*
[3]*Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran*
[4]*Department of Management, College of Business Administration in Yanbu, Taibah University, Al-Madinah Al-Munawarah 41411, Saudi Arabia*
[5]*Department of Management Information Systems, College of Business Administration in Yanbu, Taibah University, Madinah, Saudi Arabia*
[6]*Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Benha University, Egypt*
[7]*Department of Business Administration, Gulf Colleges, Saudi Arabia*

Abstract Two well-known distributions which are very helpful for modelling data in different areas, are the lognormal and inverse Weibull distributions. Choosing between the true or false distribution is substantial and of great importance. In order to determine the correct model, the ratios of biases and mean squared errors will have been computed by performing miss-specified analysis on the mean of these distributions and decision is made by comparing these ratios. To confirm the achieved theoretical results, a simulation study has been done. When the correct model is lognormal, then the miss-specification as the inverse Weibull (IW) model leads to larger values for ratios of biases and mean squared errors, so in this case miss-specification does not have a significant chance in practice. However, when the correct model is IW, there is a big chance for false specifying the lognormal model. Finally, this methodology is applied to determine the true distribution for a real data set of Covid-19 mortality rate in Germany.

Keywords Covid-19 Data, Inverse Weibull, Lognormal, Miss-Specification, Model Selection.

DOI: 10.19139/soic-2310-5070-2343

1. Introduction

Recently, the problem of finding appropriate distribution for a data set has been studied by many statisticians. This subject is called as the model selection which may be performed by using different criteria and methods such as the probability plots [1], goodness-of-fit and hypothesis testing [2], maximum likelihood estimation (MLE) [3], scale invariance [4], Bayesian procedures [5], minimum Kolmogorov distance methods [6], probability of correct selection and other approaches [7, 8, 9].

The effect of mis-specified distribution on some basic estimates is an important aspect. Yu [10, 11] analyzed the effect of miss-specification on the estimation and confidence interval for the 100pth percentile between two normal and extreme value distributions. Pascual [12] performed this research to censored data. Yu [13] evaluated the effect of miss-specification on the precision of selecting significant factors between these two distributions.

*Correspondence to: Mohammad M Saber (Email: mmsaber@eghlid.ac.ir). Department of Statistics, Higher Education Center of Eghlid, Eghlid, Iran

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Jia et al. [14] studied the effect of miss-specification on mean and its efficiency on selection between the Weibull and lognormal models. Their research has been motivated by the gap on the "analysis of the effect of miss-specification on mean and model selection".

Since there are numerous similarities between IW and Weibull distributions, hence another distribution which may arise in the problem of model selection against lognormal is Inverse Weibull (IW). The IW and lognormal are two distributions which selection between them may be crucial and of great importance in a wide range of applications. They are widely used to describe the data in engineering, for instance Akgul et al [15], recommended IW distribution as an alternative of Weibull in modeling the wind speed data. The aim of this article is to perform some study similar to [14], but by using IW instead of Weibul distribution. Some important and logical motives for applying this methodology which remains valid for many other distributions have been listed in [14].

Therefore, for miss of redundancy we refer readers only to [14], for benefits of the method which is applied in this study.

This paper is organized as follows. The influence of unknown lognormal distribution as IW distribution on the lognormal expectation is analyzed in Section 2. In Section 3, the influence on the IW mean has been investigated when the IW distribution is not correct one. Finally, Section 4 is devoted to determine the distribution of a set of real data, corresponding to the Covid-19 mortality rate in Germany by using the method which has been applied in this paper.

2. Miss-specifying lognormal distribution as IW distribution

In this section, we suppose that the correct distribution is lognormal. The mean of lognormal and its estimation under the incorrectly specified IW distribution have been investigated. At the end, the effect of this miss-specification will have been analyzed.

The pdf of the lognormal distribution by μ and σ as its location and scale parameters is

$$f_l(t;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma t} e^{-\frac{(lnt-\mu)^2}{2\sigma^2}},$$

with mean

$$h(\mu,\sigma) = e^{\mu+\frac{\sigma^2}{2}} \tag{1}$$

The MLE of $\theta = (\mu,\sigma)$ based on a random sample $T_1, \ldots, T_n$ would have been obtained by: maximizing the Log-likelihood function

$$L_l = \sum_{i=1}^{n} lnf_l(t_i) = -nln\sqrt{2\pi} - nln\sigma - \sum_{i=1}^{n} \ln t_i - \sum_{i=1}^{n} \frac{(\ln t_i - \mu)^2}{2\sigma^2}.$$

The MLE's are

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \ln t_i,$$

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (\ln t_i - \widehat{\mu})^2.$$

So, the MLE of (1) is:

$$\widehat{h} = e^{\widehat{\mu}+\frac{\widehat{\sigma}^2}{2}}. \tag{2}$$

The Fisher's information matrix is

$$\mathbf{J} = E_l\left(-\frac{\partial^2 \ln L_l}{\partial\theta^2}\right) = \begin{bmatrix} \frac{n}{\sigma^2} & o \\ o & \frac{2n}{\sigma^2} \end{bmatrix}.$$

In addition, the limiting distribution of $\widehat{h}$ is $N\ (h,\ V_{h|l}\ )$, where:

$$V_{h|l}\ = \left(\frac{\partial h}{\partial \mu}, \frac{\partial h}{\partial \sigma}\right) \mathbf{J}^{-1} \left(\frac{\partial h}{\partial \mu}, \frac{\partial h}{\partial \sigma}\right)^{T} = \frac{\sigma^2\ e^{2\mu+\sigma^2}}{n} \left(1 + \frac{\sigma^2}{2}\right) \tag{3}$$

The pdf of IW distribution is

$$f_{IW}(t; \alpha, \lambda) = \lambda \alpha t^{-(\alpha+1)} e^{-\lambda t^{-\alpha}},$$

by α and λ as the shape and scale parameters, with the expected value:

$$g(\alpha, \lambda) = \lambda^{\frac{1}{\alpha}}\ \Gamma\left(1 - \frac{1}{\alpha}\right). \tag{4}$$

Suppose that the IW distribution has been incorrectly considered for the random sample $T_1,\ \ldots, T_n$. The log-likelihood function is then given by:

$$L_{IW} = n(ln\alpha + ln\lambda) - (\alpha + 1) \sum_{i=1}^{n} lnt_i - \lambda \sum_{i=1}^{n} t_i^{-\alpha},$$

where the MLE's of the parameters α and λ can be calculated by maximizing L_{IW}. Differentiating L_{IW} w.r.t. λ and then α leads to solving the following equation:

$$\eta(\alpha) = \frac{1}{\alpha} - \frac{\sum_{i=1}^{n} lnt_i}{n} + \frac{\sum_{i=1}^{n} t_i^{-\alpha} lnt_i}{\sum_{i=1}^{n} t_i^{-\alpha}} = 0,$$

by which the MLE of α say $\widehat{\alpha}$, will have been computed firstly and subsequently have been used for calculating the MLE of λ as

$$\widehat{\lambda} = \left(\frac{1}{n} \sum_{i=1}^{n} t_i^{-\widehat{\alpha}}\right)^{-1}.$$

Moreover, the quasi-MLE (QMLE) of mean for lognormal distribution has the following form:

$$h_q = g\left(\widehat{\alpha}, \widehat{\lambda}\right) = \widehat{\lambda}^{\frac{1}{\widehat{\alpha}}}\ \Gamma\left(1 - \frac{1}{\widehat{\alpha}}\right) \tag{5}$$

The asymptotic distribution of (5) is required for achieving the goal of this article.

Let α^* and λ^* be the values that maximize the supposed logarithm of likelihood $E_l(L_{IW})$ with reversion to μ and σ, i.e.

$$(\alpha^*\ , \lambda^*) = \underset{\mu, \sigma}{argmax}\ E_l(L_{IW})$$

In order to compute $(\alpha^*\ , \lambda^*)$, first we calculate $E_l(L_{IW})$ and afterwards obtain its partial derivatives w.r.t. α and λ. Finally by setting the latter's equal to zero; successively the following calculations have been done:

$$\begin{aligned} E_l\left(L_{IW}\right) &= E_l\left[nln\alpha + nln\lambda - (\alpha + 1) \sum_{i=1}^{n} ln(T_i) - \lambda \sum_{i=1}^{n} T_i^{-\alpha}\right] \\ &= n\left(ln\alpha + ln\lambda - (\alpha + 1) E_l(lnT) - \lambda E_l\left(T^{-\alpha}\right)\right) \end{aligned}$$

then

$$E_l\left(L_{IW}\right) = n\left(ln\alpha + ln\lambda - (\alpha + 1)\mu - \lambda\ e^{-\alpha\mu + \frac{\alpha^2\sigma^2}{2}}\right),$$

and

$$\frac{\partial E_l\left(L_{IW}\right)}{\partial \alpha} = n\left(\frac{1}{\alpha} - \mu - \lambda\, e^{-\alpha\mu + \frac{\alpha^2\sigma^2}{2}}\left(\alpha\sigma^2 - \mu\right)\right) = 0,$$
$$\frac{\partial E_l\left(L_{IW}\right)}{\partial \lambda} = n\left(\frac{1}{\lambda} - e^{-\alpha\mu + \frac{\alpha^2\sigma^2}{2}}\right) = 0,$$

which leads to:

$$\alpha^* = \frac{1}{\sigma},\ \lambda^* = e^{\frac{\mu}{\sigma} - \frac{1}{2}} \tag{6}$$

Furthermore, denote

$$E_l\left(\frac{\partial^2 L_{IW}}{\partial \theta^2}\right) = E_l\begin{bmatrix} \frac{\partial^2 L_{IW}}{\partial \alpha^2} & \frac{\partial^2 L_{IW}}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 L_{IW}}{\partial \alpha \partial \lambda} & \frac{\partial^2 L_{IW}}{\partial \lambda^2} \end{bmatrix}$$

and

$$E_l\left(\frac{\partial L_{IW}}{\partial \alpha}\frac{\partial L_{IW}}{\partial \lambda}\right) = E_l\begin{bmatrix} \left(\frac{\partial L_{IW}}{\partial \alpha}\right)^2 & \frac{\partial L_{IW}}{\partial \alpha}\frac{\partial L_{IW}}{\partial \lambda} \\ \frac{\partial L_{IW}}{\partial \alpha}\frac{\partial L_{IW}}{\partial \lambda} & \left(\frac{\partial L_{IW}}{\partial \lambda}\right)^2 \end{bmatrix}.$$

Since

$$\frac{\partial L_{IW}}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} lnt_i + \lambda \sum_{i=1}^{n} t_i^{-\alpha} lnt_i,$$
$$\frac{\partial L_{IW}}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i^{-\alpha},$$
$$\frac{\partial^2 L_{IW}}{\partial\, \alpha^2} = -\frac{n}{\alpha^2} - \lambda \sum_{i=1}^{n} t_i^{-\alpha} \ln^2 t_i,$$
$$\frac{\partial^2 L_{IW}}{\partial \alpha \partial \lambda} = -\sum_{i=1}^{n} t_i^{-\alpha} lnt_i,$$
$$\frac{\partial^2 L_{IW}}{\partial \lambda^2} = -\frac{n}{\lambda^2},$$

we obtain that:

$$\begin{aligned} E_l\left[\left(\frac{\partial L_{IW}}{\partial \alpha}\right)^2\right] &= \frac{n^2}{\alpha^2} - \frac{2n^2}{\alpha} E_l(lnT) + nE_l\left(ln^2 T\right) + \frac{2n^2\lambda}{\alpha} E_l\left(T^{-\alpha} lnT\right) \\ &+ n(n-1)E_l^2(lnT) - 2n\lambda E_l\left(T^{-\alpha} ln^2 T\right) + n\lambda^2 E_l\left(T^{-2\alpha} \ln^2 T\right) \\ &- 2n\lambda(n-1)E_l(lnT)E_l\left(T^{-\alpha} lnT\right) + n\lambda^2\left(n-1\right)E_l^2(T^{-\alpha} lnT), \\ E_l\left(\frac{\partial L_{IW}}{\partial \alpha}\frac{\partial L_{IW}}{\partial \lambda}\right) &= \frac{n^2}{\alpha\lambda} - \frac{n^2}{\alpha} E_l\left(T^{-\alpha}\right) - n(n-1)E_l\left(T^{-\alpha}\right) E_l(lnT) - \frac{n^2}{\lambda} E_l(lnT) \\ &+ n(n+1)E_l\left(T^{-\alpha} lnt\right) - n\lambda E_l\left(T^{-2\alpha} lnt\right) - \ n\lambda(n-1)E_l\left(T^{-\alpha} lnt\right) E_l\left(T^{-\alpha}\right), \\ E_l\left[\left(\frac{\partial L_{IW}}{\partial \lambda}\right)^2\right] &= \frac{n^2}{\lambda^2} - \frac{2n^2}{\lambda} E_l\left(T^{-\alpha}\right) + nE_l\left(T^{-2\alpha}\right) + n(n-1)E_l^2\left(T^{-\alpha}\right). \end{aligned}$$

After a cumbersome computation we have the following results

$$\begin{aligned}
E_l(lnT)\big|_{\alpha^*,\lambda^*} &= \mu,\\
E_l\left(\ln^2 T\right)\big|_{\alpha^*,\lambda^*} &= \mu^2+\sigma^2,\\
E_l\left(T^{-\alpha}lnT\right)\big|_{\alpha^*,\lambda^*} &= (\mu-\sigma)\, e^{\frac{1}{2}\frac{\mu}{\sigma}},\\
E_l\left(T^{-\alpha}\right)\big|_{\alpha^*,\lambda^*} &= e^{\frac{1}{2}-\frac{\mu}{\sigma}},\\
E_l\left(T^{-\alpha}\ln^2 T\right)\big|_{\alpha^*,\lambda^*} &= \left(\mu^2-2\mu\sigma+2\sigma^2\right)\, e^{\frac{1}{2}-\frac{\mu}{\sigma}},\\
E_l\left(T^{-2\alpha}\right)\big|_{\alpha^*,\lambda^*} &= e^{2(1-\frac{\mu}{\sigma})},\\
E_l\left(T^{-2\alpha}lnT\right)\big|_{\alpha^*,\lambda^*} &= (\mu-2\sigma)e^{2(1-\frac{\mu}{\sigma})},\\
E_l\left(T^{-2\alpha}\ln^2 T\right)\big|_{\alpha^*,\lambda^*} &= \left(\mu^2-4\mu\sigma+5\sigma^2\right)e^{2(1-\frac{\mu}{\sigma})}.
\end{aligned}$$

Now let

$$\mathbf{J}_1 = E_l\left(\frac{\partial^2 L_{IW}}{\partial\theta^2}\right)\Big|_{\alpha^*,\lambda^*}$$

and

$$\mathbf{J}_2 = E_l\left(\frac{\partial L_{IW}}{\partial\alpha}\frac{\partial L_{IW}}{\partial\lambda}\right)|_{\alpha^*,\lambda^*}.$$

Straight computations show that:

$$\mathbf{J}_1 = -n\begin{bmatrix}\mu^2-2\mu\sigma+3\sigma^2 & (\sigma-\mu)e^{\frac{1}{2}-\frac{\mu}{\sigma}}\\ (\sigma-\mu)e^{\frac{1}{2}-\frac{\mu}{\sigma}} & e^{1-2\,\frac{\mu}{\sigma}}\end{bmatrix},$$

$$\mathbf{J}_2 = n(e-1)\begin{bmatrix}(2\sigma-\mu)^2+\frac{e}{e-1}\sigma^2 & (2\sigma-\mu)e^{\frac{1}{2}-\frac{\mu}{\sigma}}\\ (2\sigma-\mu)e^{\frac{1}{2}-\frac{\mu}{\sigma}} & e^{1-2\,\frac{\mu}{\sigma}}\end{bmatrix}.$$

By following the methodology which is recommended in White [16], the limiting distribution of the QMLE h_q becomes

$$h_q \sim N(g\left(\alpha^*,\lambda^*\right), V_{IW|l}),$$

where

$$g\left(\alpha^*,\lambda^*\right) = \Gamma(1-\sigma)e^{\mu-\frac{\sigma}{2}}$$

and

$$V_{IW|l} = \left(\frac{\partial g}{\partial\alpha},\frac{\partial g}{\partial\lambda}\right)\mathbf{J}_1^{-1}\mathbf{J}_2\mathbf{J}_1^{-1}\left(\frac{\partial g}{\partial\alpha},\frac{\partial g}{\partial\lambda}\right)^T\Bigg|_{\alpha^*,\lambda^*}.$$

A cumbersome computation shows the following representation for $V_{IW|l}$:

$$V_{IW|l} = \frac{(2e-1)\sigma^2}{4n}\Gamma^2(1-\sigma)\left\{\left[\frac{1}{2}+\Psi(1-\sigma)-\frac{1}{2e-1}\right]^2+\frac{4e^2-4e}{(2e-1)^2}\right\}e^{2\mu-\sigma}, \tag{7}$$

where

$$\Psi(x) = \frac{d}{dx}(ln\,\Gamma(x)\,)$$

illustrates the derivative of logarithm gamma function.

In order to measure the performance of estimators, two various criteria have been considered and computed; bias and mean squared error (MSE). The bias of the QMLE (h_q) of the correct lognormal mean when incorrect IW model has been purposed is given by:

$$b_{IW|l} = E\left(h_q\right) - h = g\left(\alpha^*, \lambda^*\right) - h = \left[\Gamma(1-\sigma) - e^{\frac{\sigma^2+\sigma}{2}}\right] e^{\mu - \frac{\sigma}{2}}. \tag{8}$$

The corresponding MSE is

$$M_{IW|l} = E\left[\left(h_q - h\right)^2\right]$$

which is determined as follows:

$$M_{IW|l} = V_{IW|l} + \left[E\left(h_q - h\right)^2\right] = V_{IW|l} + \left[\Gamma(1-\sigma) - e^{\frac{\sigma^2+\sigma}{2}}\right]^2 e^{2\mu-\sigma} \tag{9}$$

We evaluate the mis-specification effect on the mean by comparing the bias and MSE of QMLE h_q with those of the MLE $\widehat{h}$ under the lognormal distribution. In other words, the ratio of biases under both models and the ratio of MSE's of them have been computed. If the lognormal distribution was selected, we have $E(\widehat{h}) = h$.Therefore, $rb_{IW|l}$ is the ratio of the bias of (5) to its true value:

$$rb_{IW|l} = \frac{b_{IW|l}}{h} = \Gamma(1-\sigma)e^{-\frac{\sigma^2+\sigma}{2}} - 1. \tag{10}$$

The coefficient of the MSE of h_q to the MSE of $\widehat{h}$ is:

$$rm_{IW|l} = \frac{M_{IW|l}}{V_{h|l}} = \frac{V_{IW|l}}{V_{h|l}} + \frac{n\left[\Gamma(1-\sigma)e^{-\frac{\sigma^2+\sigma}{2}} - 1\right]^2}{\sigma^2\left(1+\frac{\sigma^2}{2}\right)}. \tag{11}$$

By un-biasness of $\widehat{h}$ for the lognormal distribution and therefore, $MSE\left(\widehat{h}\right)$ is just its variance, $V_{h|l}$, given by (3).

Here a simulation has been done like [14]. All programs of this paper have been written in R software and included in the Appendix. The simulation results have been compared with the theoretical ones in Table 1. Figures of this table shows that the theoretic criteria are reliable. Figure 1 presents the heat map based on Table 1.

Table 1. The values of $rb_{IW|l}$ and $rm_{IW|l}$ for miss-specified IW distribution.

μ	σ	$n\|N$ =10000		$M_{IW\|l}$	$rb_{IW\|l}$	$rm_{IW\|l}$
1	0.05	100	simulated	0.0003384081	0.004119081	1.815796
			theoretical	0.000393755	0.004729946	2.123581
1	0.05	30	simulated	0.0007517901	0.003617034	1.236474
			theoretical	0.0009258257	0.004729946	1.497936
4	0.05	30	simulated	0.3044564	0.003537992	1.230158
			theoretical	0.3735048	0.004729946	1.497936
4	0.05	100	simulated	0.134283	0.004181023	1.828597
			theoretical	0.1588521	0.004729946	2.123581

At the moment, we compute the ratios of biases and MSE's ($rb_{IW|l}$ and $rm_{IW|l}$) by Equations (10) and (11), respectively. In Figure 2 and Figure 3, the ratio of biases have been plotted for different values of σ. As Figure 2 and Figure 3 demonstrate $rb_{IW|l}$ is an increasing function of σ and it grows rapidly. In fact, for values of σ which are less than 0.75, the ratio of the biases are less than 1, while for values of σ bigger than 0.95, the $rb_{IW|l}$ s are greater than 10. This indicates that the absolute values of $rb_{IW|l}$ become greater and tend to the true value of the lognormal mean, as σ greats. The $rb_{IW|l}$s are always more than 0, so the mis-specified IW distribution generally overestimate the lognormal mean.

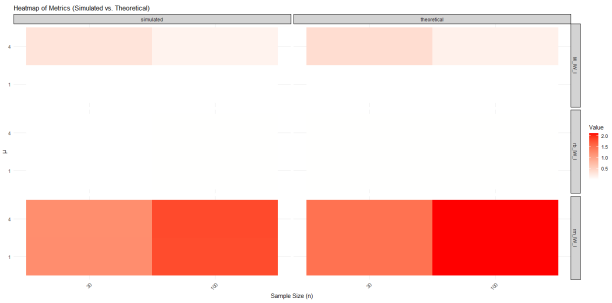

Figure 1. Heat map based on Table 1.

The parameters μ, σ, and n in the simulation study are strategically chosen to assess the robustness of metrics like $M_{IW|l}, rb_{IW|l}$ and $rm_{IW|l}$ under varying conditions, with μ = 1 and μ = 4 representing low and high baseline mortality rates (e.g., controlled vs. severe outbreak scenarios), σ = 0.05 reflecting minimal variability to test estimator performance under idealized low-uncertainty conditions, and sample sizes $n = 30$ (small) and $n = 100$ (larger) evaluating finite-sample behavior versus asymptotic stability. While the fixed σ = 0.05 ensures a baseline analysis of precision, expanding σ to ranges like $\sigma < 0.75$ (low variability, e.g., stable public health interventions) and $\sigma > 0.95$ (high variability, e.g., stochastic surges or heterogeneous populations) would better stress-test estimator reliability in real-world contexts. The data reveal that higher μ (e.g., μ = 4) drastically inflates $M_{IW|l}$ (e.g., 0.304 vs. 0.000338 for μ = 1), highlighting sensitivity to baseline conditions, while larger n reduces $M_{IW|l}$ (e.g., 0.000338 for $n = 100$ vs. 0.000751 for $n = 30$), aligning with improved precision in larger samples. Discrepancies between simulated and theoretical values (e.g., simulated $M_{IW|l}$ =0.134 vs. theoretical 0.1588 for μ = 4, $n = 100$) suggest finite-sample biases or model limitations. To enhance applicability, future simulations should incorporate broader μ ranges, σ thresholds ($\sigma < 0.75$ and $\sigma > 0.95$), and extreme sample sizes ($n = 10$, 1000) to validate estimators across epidemiological extremes, ensuring robustness for pandemic response planning under uncertainty.

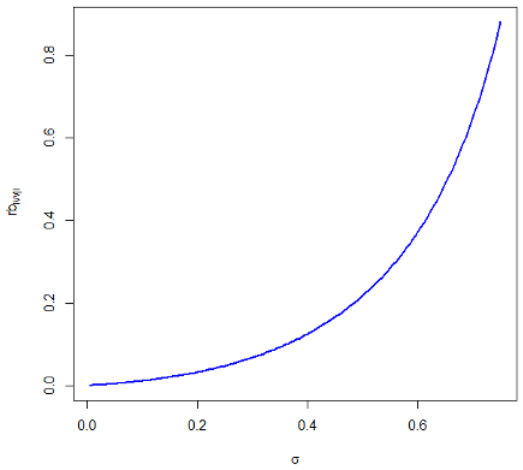

Figure 2. The coefficient of biases with smaller σ.

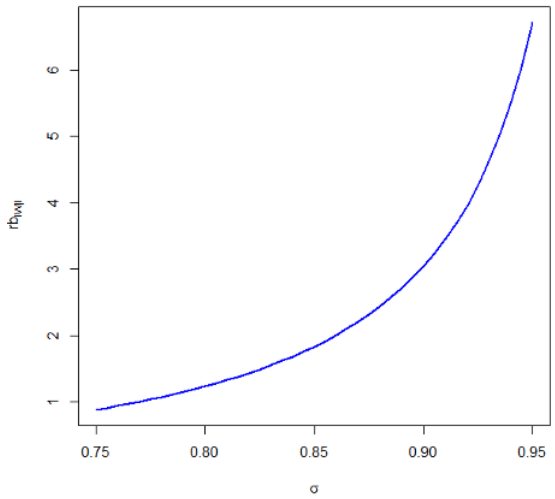

Figure 3. The coefficient of biases with larger σ.

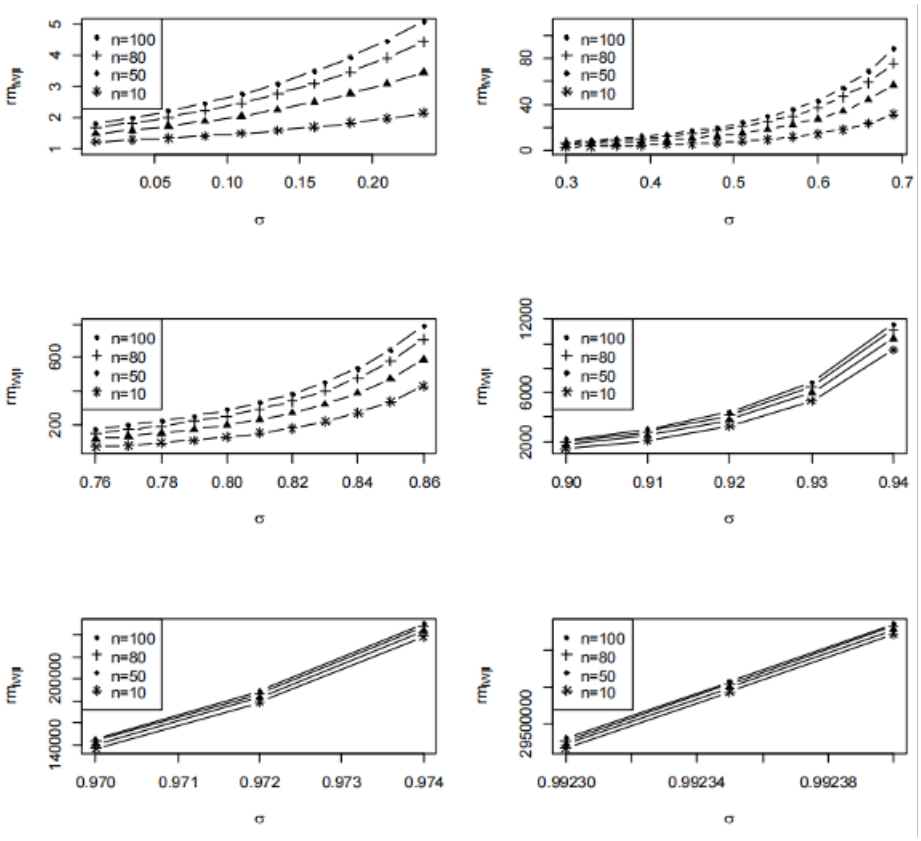

Figure 4. Left: The values of $rm_{IW|l}$ for small σ. Right: The values of $rm_{IW|l}$ for large σ.

The plotted proportion of MSE,s in Figure 4, shows that false specification may occur for small sample sizes and values of σ less than 0.25.

3. Miss-specified lognormal distribution instead of IW distribution

In this section we assume that the correct distribution is IW. In order to analyzing the results of miss-specification, IW mean and some approximations under the wrong assumption of lognormal distribution have been inspected.

Let $X_1, ..., X_n$ be iid with IW(α,λ) distribution. The corresponding MLE's$\widehat{\alpha}$and$\widehat{\lambda}$had been presented in Section 2. The Fisher's matrix of information for parameter $\theta = (\alpha, \lambda)$ is

$$I = E_{IW}(-\frac{\partial^2 \ln L_{IW}}{\partial \theta^2}) = -E_{IW}\left\{\begin{bmatrix} \frac{\partial^2 \ln L_{IW}}{\partial \alpha^2} & \frac{\partial^2 \ln L_{IW}}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 \ln L_{IW}}{\partial \alpha \partial \lambda} & \frac{\partial^2 \ln L_{IW}}{\partial \lambda^2} \end{bmatrix}\right\}$$
$$= \frac{n}{\alpha^2 \lambda^2}\begin{bmatrix} \lambda^2(\psi^{(1)}(1) + (1 + \psi(1) - \ln\lambda)^2) & \alpha\lambda(1 + \psi(1) - \ln\lambda) \\ \alpha\lambda(1 + \psi(1) - \ln\lambda) & \alpha^2 \end{bmatrix},$$

where

$$\psi^{(k)}(y) = \frac{d^{k+1}}{dy^{k+1}} \ln\Gamma(y).$$

The MLE (α, λ) of IW mean had been presented in (5), the limiting distribution of$\widehat{g}$is $N(g, V_{g|IW})$ where g had been presented in (4) and by some calculus we arrive at:

$$V_{g|IW} = \frac{\partial g}{\partial \theta}^T I^{-1} \frac{\partial g}{\partial \theta} = \frac{\Gamma^2(1 - \frac{1}{\alpha})}{n\alpha^2 \psi^{(1)}(1)} \lambda^{\frac{2}{\alpha}} \left\{(1 + \psi(1) - \psi(1 - \frac{1}{\alpha}))^2 + \psi^{(1)}(1)\right\} \tag{12}$$

Now suppose that the lognormal distribution has incorrectly selected to explain the data $X_1, ..., X_n$. Formerly, the MLEs$\widehat{\mu}$and$\widehat{\sigma}$had been presented in Section 2, and the QMLE g_q of the IW mean by (2). The limiting distribution of the (2) will be applied to analyze the findings.

Let $\mu^*$and σ^* maximize the expected log-likelihood $E_{IW}(\ln L_l)$ with regard to α and λ., i.e.

$$(\mu^*, \sigma^*) = \underset{\alpha,\lambda}{argmax} E_{IW}(\ln L_l).$$

First of all we have

$$E_{IW}(\ln L_l) = -n \ln\sqrt{2\pi} - n \ln\sigma - nE_{IW}(\ln X) - \frac{n}{2\sigma^2} E_{IW}(\ln X - \mu)^2.$$

So,

$$E_{IW}(\ln L_l) = -n \ln\sqrt{2\pi} - n \ln\sigma - \frac{n}{\alpha}(\psi(1) - \ln\lambda) - \frac{nA}{2\sigma^2}$$

where

$$A = \frac{\psi^{(1)}(1)}{\alpha^2} + \left(\frac{1}{\alpha}(\ln\lambda - \psi(1)) - \mu\right)^2.$$

In order to maximize $E_{IW}(\ln L_l)$, its partial derivatives w.r.t. to μ and σ will be set to 0:

$$\frac{\partial E_{IW}(\ln L_l)}{\partial \mu} = \frac{n}{\sigma^2}[\frac{1}{\alpha}(\psi(1) - \ln\lambda) - \mu] = 0 \text{ and } \frac{\partial E_{IW}(\ln L_l)}{\partial \sigma} = -\frac{n}{\sigma}[1 + \frac{A}{\sigma^2}] = 0.$$

So, the values of μ^*andσ^* have been obtained as:

$$\mu^* = \frac{\ln\lambda - \psi(1)}{\alpha}, \qquad \sigma^* = \frac{\sqrt{\psi^{(1)}(1)}}{\alpha}. \tag{13}$$

Similar to the approach of previous section, we have:

$$I_1|_{\mu^*,\sigma^*} = -\frac{n\alpha^2}{\psi^{(1)}(1)}\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix} I_2|_{\mu^*,\sigma^*} = \frac{n\alpha^2}{\psi^{(1)}(1)}\begin{bmatrix}1 & -\frac{\psi^{(2)}(1)}{\left(\psi^{(1)}(1)\right)^{\frac{3}{2}}}\\ -\frac{\psi^{(2)}(1)}{\left(\psi^{(1)}(1)\right)^{\frac{3}{2}}} & \frac{\psi^{(3)}(1)}{\left(\psi^{(1)}(1)\right)^{2}}+2\end{bmatrix}.$$

Now, asymptotic distribution of QMLE g_q is,

$$g_q \sim N(h(\mu^*,\sigma^*), V_{l|IW}$$

where:

$$h(\mu^*,\sigma^*) = e^{\frac{\ln\lambda-\psi(1)}{\alpha}+\frac{\psi^{(1)}(1)}{2\alpha^2}}$$

and

$$V\frac{\partial h}{\partial \mu}\frac{\partial h}{\partial \sigma}_{1}{}^{-1}{}_{2}{}^{-1}\frac{\partial h}{\partial \mu}\frac{\partial h}{\partial \sigma}^{T}_{l|IW}. \tag{14}$$

By substituting (12) in (14) we have

$$\left(V\frac{\psi^{(1)}(1)}{n\alpha^2}^{\frac{2}{\alpha}}\left\{\frac{\psi^{(1)}(1)}{2\alpha^2}\left(\frac{\psi^{(3)}(1)}{2\left(\psi^{(1)}(1)\right)^2}+1\right)-\frac{\psi^{(2)}(1)}{\alpha\psi^{(1)}(1)}+1\right\}^{\frac{\psi^{(1)}(1)}{\alpha^2}-\frac{2\psi(1)}{\alpha}}\right)_{l|IW}. \tag{15}$$

Based on the bias and MSE criterias, the results of the miss-specified problem will have been analyzed. The bias criterion of the (2) of the IW mean versus the incorrect lognormal distribution is:

$$bq^{**}{}^{\frac{1}{\alpha}}\left[e^{\frac{\psi^{(1)}(1)}{2\alpha^2}-\frac{\psi(1)}{\alpha}}-\Gamma(1-\frac{1}{\alpha})\right]_{l|IW}. \tag{16}$$

The MSE of $g_q$is

$$Mq^2 2\frac{2}{\alpha}\left[e^{\frac{\psi^{(1)}(1)}{2\alpha^2}-\frac{\psi(1)}{\alpha}}-\Gamma(1-\frac{1}{\alpha})\right]^2_{l|IW} l|IW_{l|IW\,l|IW}. \tag{17}$$

For simplicity of comparison, (16) and (17) will have been calculated similar to Section 2. Thus:

$$\left[rb\frac{\psi^{(1)}(1)}{2\alpha^2}-\frac{\psi(1)}{\alpha}^{-1}\frac{1}{\alpha}\right]_{l|IW}, \tag{18}$$

$$rm\frac{M_{l|IW}}{\left\{V_{g|IW}=\frac{V_{l|IW}}{V_{g|IW}+\frac{n\alpha^2\psi^{(1)}(1)\left[e^{\frac{\psi^{(1)}(1)}{2\alpha^2}-\frac{\psi(1)}{\alpha}}\Gamma^{-1}(1-\frac{1}{\alpha})-1\right]^2}{(1+\psi(1)-\psi(1-\frac{1}{\alpha}))^2+\psi^{(1)}(1)}}\right\}_{l|IW}}. \tag{19}$$

Similar to Section 2, a simulation study has been employed which again confirms the accuracy of (17)-(19). The results of this simulation have been presented in Table 2. Therefore, employing Equations 10 and 11 in order to computing the ratios of biases and MSE's, say $rb_{l|IW}$and $rm_{l|IW}$, is valid and justifiable. The criteria $rb_{l|IW}$as the ratio of biases have been plotted for different values of α in Figure 6. Figure 1 presents the heat map based on Table 2.

Table 2. The values of $rb_{l|IW}$ and $rm_{l|IW}$ for miss-specified lognormal distribution.

α	λ	n		$M_{l\|IW}$	$rb_{l\|IW}$	$rm_{l\|IW}$
3.5	0.9	10	simulated	0.03292290	−0.00972015	0.9441920
			theoretical	0.03168536	−0.01159774	0.9882228
3.5	0.9	30	simulated	0.01066024	−0.01011047	0.9768248
			theoretical	0.01069926	−0.01159774	1.001085
1.25	0.9	100	simulated	3.198000	−0.4074761	1.004237
			theoretical	3.278238	−0.4148648	1.380777
1.25	0.9	400	simulated	3.092395	−0.4130553	4.086844
			theoretical	3.118075	−0.4148648	5.253269

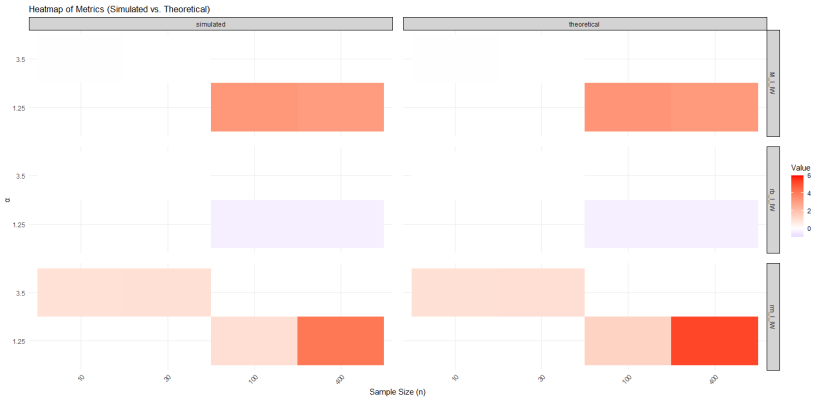

Figure 5. Heat map based on Table 2.

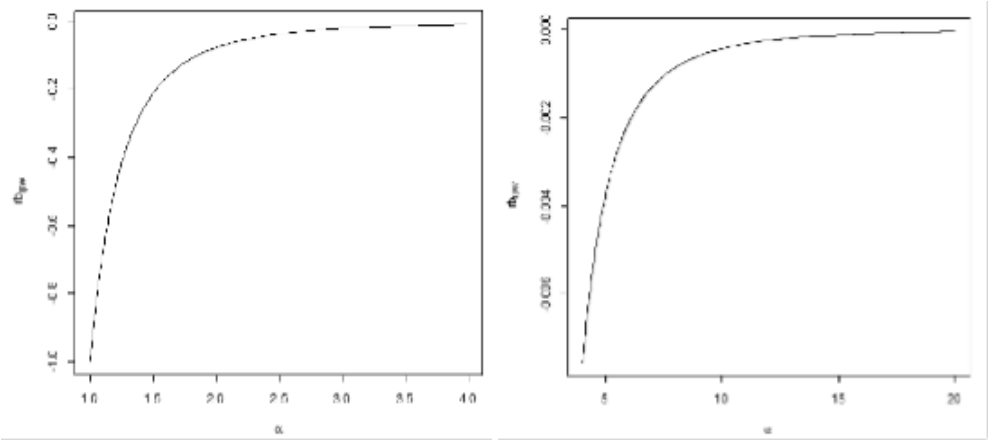

Figure 6. Left: The graph of $rb_{l|IW}$ for small α. Right: The graph of $rb_{l|IW}$ for large α.

The criterion $rb_{l|IW}$ is less than 0 for all values of parameters, which provide this finding that the miss-specified lognormal distribution underestimate the IW mean. Again, this criterion is an increasing function of α which approaches to 0 although it grows very slowly. The $rb_{l|IW}$ is between -1 and 0 for all values of α which imply on this fact that the absolute value of $rb_{l|IW}$ decreases and appears to reach with the actual value of the IW. The $rm_{l|IW}$ shown in Figure 5 is a steadily increasing function of n which states that for large values of sample size n, the effect on the mis-estimated IW mean is more significant. Incorrect specifications are more likely to occur for values greater than 2.5. In short, $rb_{l|IW}$ is determined only by although $rm_{l|IW}$ depends on both α and n. Both

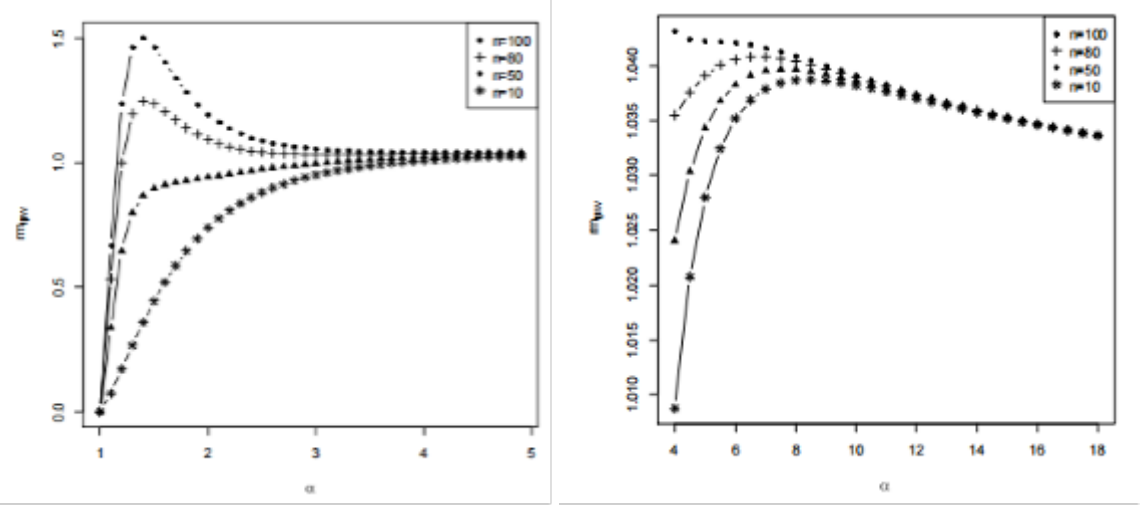

Figure 7. Left: The graph of $rm_{l|IW}$ for small α. Right: The graph of $rm_{l|IW}$ for large α.

$rb_{l|IW}$ and $rm_{l|IW}$ tend to their small values whenever α increases. The false determined lognormal distribution is accepted for large values of parameter α, especially more than 3.

4. Real data analysis

This section is devoted to an illustrative real data analysis. It will have been shown that for the problem of choosing a satisfactory distribution for a data set, various selection methods may suggest different distributions. Furthermore, the model selection, which is proposed in this article, has been applied to determine an appropriate model for this dataset. The data correspond to the mortality rate of Covid-19 virus in Germany in 2021, provided by World Health Organization (https://covid19.who.int/). Given that the data corresponds to the mortality rate of the Covid-19 virus in Germany in 2021, we can expand the interpretation of the Cullen and Frey plot in the context of real-world implications. The mortality rate data you're working with is likely to have some important characteristics due to the nature of the pandemic. Mortality rates for infectious diseases like Covid-19 typically exhibit varying levels of severity across different regions and time periods. The Covid-19 mortality rates can fluctuate due to factors like the emergence of new virus variants, government interventions (lockdowns, vaccination rates), healthcare system capacity, and population age structure.

The summary statistics for Germany's 2021 COVID-19 mortality rate reveal a right-skewed distribution, with a median of 1.23 and a higher mean of 1.8375, indicating that extreme values, likely during pandemic surges, pulled the average upward. The substantial standard deviation (1.438) and wide range (0.3–5.0) underscore significant variability, reflecting periods of both controlled mortality (as low as 0.3) and severe outbreaks (peaking at 5.0). The interquartile range (0.63–2.83) captures moderate variability in the central 50% of data, while the gap between the mean and median, coupled with the high maximum value, highlights skewness driven by outliers. This pattern likely mirrors real dynamics such as regional disparities in healthcare access, variant-driven waves (e.g., Delta), vaccination rollout timing, and fluctuating public health measures. The data emphasize the need for targeted interventions to address uneven mortality outcomes and underscore the value of visualizing trends (e.g., histograms, time-series plots) to better contextualize outliers and inform adaptive policy responses.

Figure 8 below gives the Cullen and Frey plot for the mortality rates. He Cullen and Frey plot helps assess the distribution of the mortality rate data by comparing its skewness (asymmetry) and kurtosis (peaked-ness) against several well-known distributions. If the data has a positive skew (values with higher mortality are more spread out), this indicates that most of the data points (mortality rates) are clustered around lower values, but there is

a tail of higher mortality rates. A negative skew (higher values are clustered) would suggest the opposite: most regions have higher mortality rates, but a few regions experience very low rates. If the data has a kurtosis around 3 (mesokurtic), this will imply that the data follows a distribution like the normal distribution. Kurtosis greater than 3 (leptokurtic) indicates that the data is more peaked, meaning that the mortality rates are clustered around the mean but with a higher frequency of extreme values (outliers). Kurtosis less than 3 (platykurtic) suggests that the mortality rates are more evenly spread out (flatter distribution), with fewer extreme values. It's likely that the data would show positive skewness. This would mean that most of the regions or time periods had relatively lower mortality rates, but a few regions or periods had significantly higher mortality rates. For instance, areas with overwhelmed healthcare systems, or later waves of the pandemic, could have had higher mortality rates due to the surge in cases. Additionally, the inequality in healthcare access, age distribution of the population, and variant impact could contribute to this positive skew, with more deaths in areas with higher vulnerability, poor vaccination coverage, or where hospitals were overwhelmed. A positive skew might reflect the fact that most regions or time periods had low mortality (e.g., after the vaccine rollout), while a few periods (e.g., before the vaccines were available or during the peak of variants like Delta or Omicron) had much higher mortality. The kurtosis of the plot will tell us if the distribution is more concentrated around the mean or has a tail with extreme values. A high kurtosis (greater than 3) could indicate that there were some extreme outliers (periods or regions with extremely high mortality rates compared to the rest), which is often the case in pandemic data. This would suggest a leptokurtic distribution, where most mortality rates were within a moderate range, but a few significant spikes occurred during critical waves of the pandemic (e.g., during the first wave or when new, more virulent variants emerged). A lower kurtosis (less than 3) would suggest that the mortality rates were relatively spread out, with no significant peaks—perhaps indicating that no specific regions or periods had dramatically higher rates. If the data points align with a Log-Normal distribution, it could suggest that the data follows a distribution where small values are very common (low mortality rates), but the higher values (outliers with much higher mortality) occur less frequently but still significantly affect the overall distribution. This is plausible in the context of Covid-19, as the pandemic experienced waves of both low and high mortality rates. Figure 9: The Kernel density estimation plot (top left), the total time in test plot (top right), the box plot (bottom left) and the Violin plot (bottom right) for the mortality rates. The Kernel Density Estimation (KDE) plot gives a smooth estimate of the probability density function for the mortality rates. This plot illustrates the distribution of the data, showing a right-skewed distribution. The peak of the distribution occurs around 1.0, with a long tail extending to higher mortality rates. This indicates that most mortality rates are clustered around the lower end, with a few extreme values at the higher end. The KDE plot highlights the general trend that, while most observations are moderate, there are occasional outliers with much higher mortality rates. This is important for understanding the overall spread and identifying potential high-risk events or regions. The right skew suggests that while the majority of mortality rates are concentrated around moderate levels (1.0), there are significant fluctuations in the data with occasional spikes in mortality, possibly related to specific events, outbreaks, or regions. The total time in test plot clearly indicates that mortality rates are increasing by showing the progression of mortality rates over time or across different categories. This upward trend could be due to a variety of reasons including worsening conditions in the pandemic, system overloads, or other socio-political factors. The plot helps visualize how mortality has evolved and can provide insights into the timing of interventions, impact of public health measures, and potential areas of concern that may require further investigation or intervention.

The Box Plot shows the distribution of mortality rates by illustrating the median, interquartile range (IQR), and the presence of outliers. The median (Q2) is located at approximately 1.0, indicating that the middle of the dataset is concentrated around this value. The box shows the range between Q1 (25th percentile) and Q3 (75th percentile), highlighting the spread of the middle 50% of the data. There is a long whisker on the upper side, indicating the presence of higher mortality rates (outliers) beyond Q3. This suggests that, while most mortality rates are moderate, there are extreme values that significantly differ from the majority. The box plot helps to confirm the skewed distribution identified in the KDE plot. The box plot provides a clear view of the spread of mortality rates, indicating that while the data is fairly concentrated in the lower ranges, the presence of outliers highlights areas where the mortality rate spikes. The whisker and outliers suggest that certain extreme values should be further investigated. The Violin Plot combines the box plot with a kernel density estimate, offering a detailed view of the

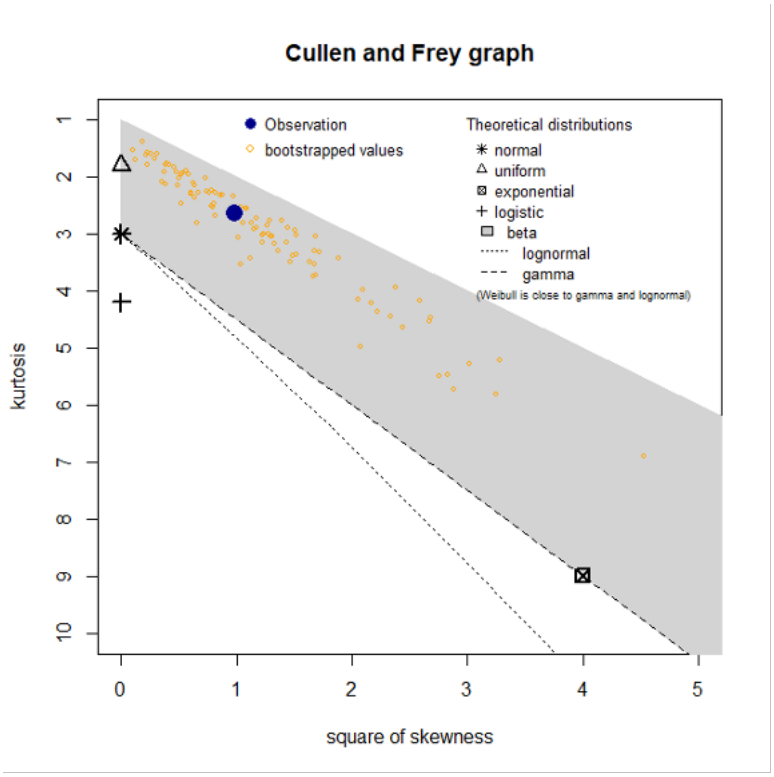

Figure 8. Cullen and Frey plot for the mortality rates.

data's distribution and density. The wide part of the violin near 1.0 indicates that most mortality rates are clustered around this value. The long tail extending to the right confirms that a smaller proportion of mortality rates are higher, creating a skewed distribution. The boxplot overlay within the violin plot clearly shows the median, Q1, and Q3, helping to reinforce the information from the box plot. This plot is particularly useful for understanding the density of the data, as it provides both the shape of the distribution and a summary of key statistical measures. The violin plot effectively highlights the distribution and density of the data. The rightward skew is more apparent here, showing that while most mortality rates are lower (around 1.0), there are a small number of regions or times with significantly higher mortality rates. The quartile lines within the violin plot make it easy to identify the central tendency and spread, and the density plot shows where data points are most concentrated.

Various selection methods discussed in Introduction are applied to choose right distribution between the IW and lognormal distributions for mentioned data. Results of applying some of these techniques containing Kolmogorov-Smirnov (KS) test, MLE and scale invariant (SI) methods to Covid19 data has been presented in Table 3. Also, the graphs of Q-Q plot of this data for lognormal and IW distribution has been shown in Figure 10. As this table demonstrates, Kolmogorov-Smirnov test and MLE methods recommend lognormal distribution while two other methods yield in selecting IW distribution.

Table 3. Different methods for selecting appropriate distribution, the estimated parameters in MLE methods are $\widehat{\mu} = 2.5$, $\widehat{\sigma} = 0.19$, $\widehat{\alpha} = 1.17$ and $\widehat{\lambda} = 0.12$.

	MSE	KS		MLE	SI
		statistics	p-value	Log(likelihood)	
lognormal	0.84	0.16	0.35	-99.54	-100.41
IW	0.77	0.29	0.24	-102.03	-98.15

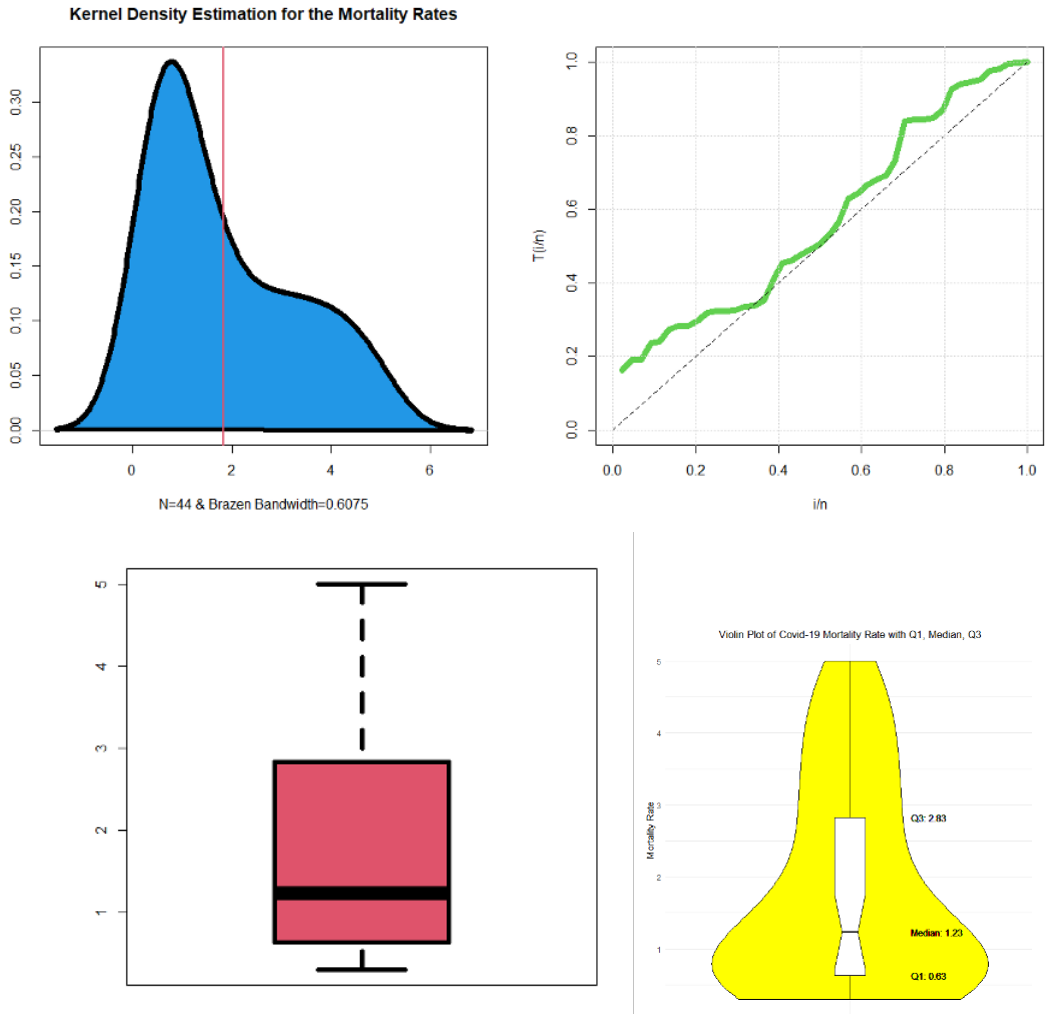

Figure 9. The Kernel density estimation plot (top left), the total time in test plot (top right), the box plot (bottom left) and the Violin plot (bottom right) for the mortality rates.

Now, we investigate about the impact of miss-specified distribution on the mean. Henceforth, we consider two cases as follows.

(i) The lognormal distribution is the correct model versus IW.
(ii) The IW distribution is the correct model versus lognormal.

The values of four criteria $rb_{IW|l}$, $rm_{IW|l}$, $rb_{l|IW}$ and $rm_{l|IW}$ have been computed by using the Equations (10), (11), (18) and (19), respectively. By the figures of Table 4, $\left|rb_{IW|l}\right| < \left|rb_{l|IW}\right|$ and $rm_{l|IW} < rm_{IW|l}$ which together by the results of Sections 2 and 3, yields in the conclusion that IW model is better for this dataset may be more logical.

The analysis also shows the influence of incorrect model selection on the mean estimation under different conditions. When the correct model is IW, the miss-specified model (lognormal) results in relatively smaller errors in the mean, particularly for large values of the shape parameter, where $rb_{l|IW}$ tends toward zero and $rm_{l|IW}$ remains below 1.5. When the correct model is lognormal, however, selecting IW results in larger errors in the mean,

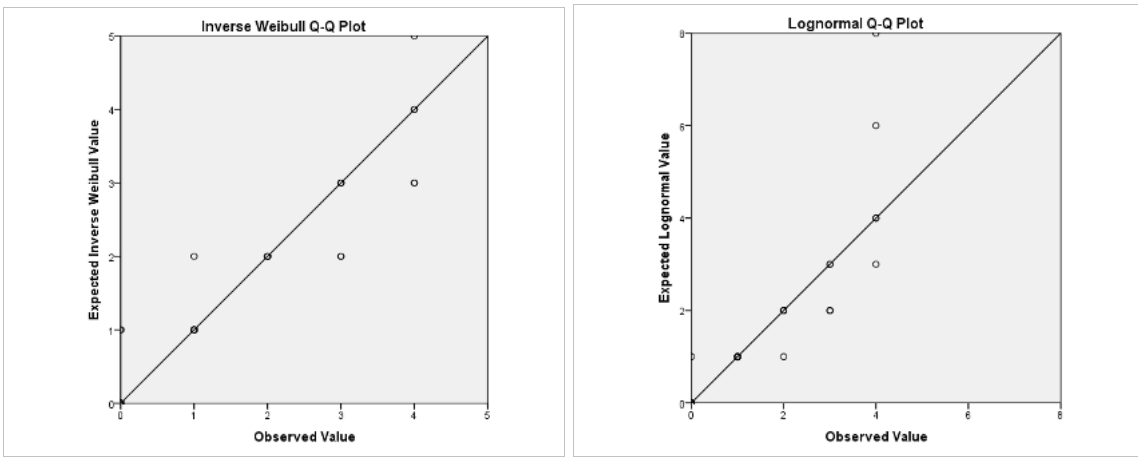

Figure 10. The Q-Q plot of data for two distributions.

Table 4. Criteria $rb_{IW|l}$, $rm_{IW|l}$, $rb_{l|IW}$ and $rm_{l|IW}$.

Correct model	Wrong but selected model	$rb_{l\|IW}$	$rm_{l\|IW}$
IW	lognormal	-1.24	0.45
lognormal	IW	0.004	1.98

with $rb_{l|IW}$ increasing rapidly as the scale parameter σ grows. This suggests that choosing IW when lognormal is correct can lead to significant errors in the data mean. Below, we provide some recommendation in this regard:

1. When choosing between competing distributions, use a comprehensive set of goodness-of-fit measures to evaluate the appropriateness of the models. In particular, rely on criteria like Mean Squared Error (MSE), Kolmogorov-Smirnov (KS) statistics, and log-likelihood values. These measures provide insights into how well a distribution fits the data and can help prevent the selection of a mis-specified model. This study shows that while the lognormal distribution may initially appear to fit better according to some fit statistics (MSE, KS), choosing the IW distribution might be more appropriate for certain datasets. Using multiple fit criteria will reduce the risk of overfitting and help select the most accurate model for the data.
2. Always assess the impact of miss-specification on the estimated mean, especially when there are uncertainties in model selection. Use specialized criteria like rb and rm (defined in Equations 10, 11, 19, and 20) to examine how choosing a wrong distribution (i.e., mis-specified model) influences central tendency estimates such as the mean and variance. This analysis has demonstrated that incorrect model selection, especially when choosing IW over lognormal or vice versa, can lead to different bias levels in mean estimation. In particular, mis-specifying the IW distribution when the true model is lognormal tends to produce significant errors, whereas the opposite (lognormal misspecification for IW) has less impact. Understanding these biases is critical for making robust statistical inferences.
3. Perform simulation studies alongside real data analysis to validate the selection of the correct distribution model. Simulations will allow you to test how well different models perform across a variety of scenarios, including different parameter settings. The theoretical results from this study were supported by simulation studies, which is a powerful method to assess model selection and its effects on data. By generating synthetic datasets and testing model performance in controlled settings, researchers can identify the most reliable distribution and understand how different models behave under different parameter conditions.

Conclusion, discussions and limitations

In this article, we have discussed the problem of model selection between lognormal and IW distributions, for a given data set. The influence of wrong choosing for any of these distributions when another distribution is the correct one, has been studied. Four criteria $rb_{IW|l}$, $rm_{IW|l}$, $rb_{l|IW}$ and $rm_{l|IW}$ were applied in order to determine the influence of miss-specification on the mean. The theoretical results have been approved by a simulation study. In the case of correct IW distribution, interestingly the criteria $rb_{l|IW}$ and $rm_{l|IW}$ does not depend on the scale parameter λ. For all values of the parameter α, it has been observed that $-1 < rb_{l|IW} < 0$ and $rm_{l|IW} < 1.5$, but when α increases we observe that $rb_{l|IW}$ and $rm_{l|IW}$ ten to 0 and 1, respectively. So, in the case of right IW distribution, false choosing of lognormal distribution may occur with no strange estimate in data mean. This phenomenon may happen more for large values of α, by the fact that $rb_{l|IW}$ and $rm_{l|IW}$ are very close to 0 and 1, respectively. For the other case which the true distribution is lognormal, although $rb_{IW|l}$ and $rm_{IW|l}$ do not depend on the location parameter μ, but the concluding results are completely different from the previous situation. In this case only for $\sigma < 0.75$, the observed values of $rb_{IW|l}$ is less than 1. For larger values of σ, $rb_{IW|l}$ rapidly increases, while the criterion $rm_{IW|l}$ is very large for all values of σ. In this case, choosing IW distribution leads to a very inappropriate estimate of data mean. Therefore, as the final result of paper we can claim that wrong choosing of lognormal distribution instead of IW distribution may occur with a high chance, but the converse does not hold. In this article, the problem of model selection between IW and lognormal distributions have been discussed. Same studies can be done by methodology of the present paper for model selection between other distributions for them miss-specification may occur. For instance, model selection between Weibull and IW distributions may be purpose of future work.

The proposed methodology for selecting between IW and lognormal distributions relies on comparing bias ratios and MSE under misspecification. However, a critical limitation arises when the true data-generating process deviates significantly from both distributions. If the data follows a third distribution (e.g., Gamma, Weibull, or a multimodal mixture), neither the IW nor lognormal model may adequately capture the underlying structure. The bias and MSE ratios could misleadingly favor one misspecified model over another, leading to incorrect conclusions. The theoretical framework assumes the true model is either IW or lognormal. real data, such as pandemic mortality rates, often exhibit complexities (e.g., heterogeneity, outliers, or temporal shifts) that violate these parametric assumptions. For example, Covid-19 mortality data might reflect varying transmission dynamics, healthcare capacity, or policy interventions, which neither distribution alone can fully model. The current simulation study focuses on idealized scenarios where the true model is IW or lognormal. If the data deviates from both, the method's performance (e.g., bias ratios) may degrade, and the decision rule (e.g., selecting IW when lognormal is misspecified) could become unreliable.

To address these limitations, future work should explore extensions to mixture models and non-parametric alternatives. Mixture models, such as combining IW and lognormal components, could flexibly capture heterogeneous data structures by allowing subpopulations to follow different distributions. For example, a mixture model might disentangle mortality patterns in vaccinated versus unvaccinated groups or regions with varying healthcare capacities, providing a more nuanced representation of the data. Similarly, Bayesian model averaging could integrate IW and lognormal fits, weighting their contributions based on posterior probabilities to avoid rigid model selection. Non-parametric methods, such as kernel density estimation (KDE) or Bayesian non-parametric approaches (e.g., Dirichlet process mixtures), offer further flexibility by relaxing parametric assumptions entirely. These methods adapt to the data's complexity without imposing a fixed distributional form, making them ideal for modeling unknown or irregular patterns. Hybrid frameworks, such as semi-parametric models that combine parametric components (e.g., IW/lognormal for the bulk of the data) with non-parametric tails, could also balance interpretability and flexibility, particularly for capturing extreme values or outliers. Additionally, integrating machine learning techniques (e.g., random forests or neural networks) could help identify latent patterns in the data, guiding the choice of parametric models or signaling the need for non-parametric alternatives. While these extensions introduce computational and interpretability challenges, they would significantly enhance

the methodology's robustness, enabling reliable inference even when data deviates from standard parametric assumptions. Cross-validation or bootstrapping could further validate these approaches, ensuring their applicability to complex, real scenarios like pandemic modeling.

Finally, to determine whether the IW or lognormal distribution better fits the data, use the following criteria based on the ratio of biases (rb) and ratio of mean squared errors (rm) under misspecification:

Select IW if:

$rm_{IW|l} < 1.5$: The IW model's MSE is sufficiently smaller than the lognormal model's, indicating better predictive accuracy.

$rb_{IW|l} > -0.5$: The bias direction aligns with IW assumptions (e.g., tail behavior or hazard rate characteristics).

Select Lognormal if:

$rm_{IW|l} \geq 1.5$: The lognormal model's MSE is comparable or superior, suggesting IW misspecification.

$rb_{IW|l} \leq -0.5$: The bias diverges significantly from IW expectations, favoring lognormal.

Appendix

```
si1$<$-digamma(1);si11$<$-trigamma(1);si12$<$-psigamma(1,2);si13$<$-psigamma(1,3)
Vh1$<$-function(n,mu,sigma){
a1$<$-(sigma^2/n)*(1+sigma^2/2)
a3$<$-exp(2*mu+sigma^2)
G$<$-a1*a3
return(G)            }
Viw1$<$-function(n,mu,sigma){
a1$<$-(2*exp(1)-1)*sigma^2*gamma(1-sigma)^2/(4*n)
a2$<$-(.5+digamma(1-sigma)-1/(2*exp(1)-1))^2+(4*exp(1)^2-4*exp(1))/((2*exp(1)-1)^2)
a3$<$-exp(2*mu-sigma)
G$<$-a1*a2*a3
return(G)    }
Viw1m$<$-function(n,mu,sigma){
a11$<$-mu^2-2*mu*sigma+3*sigma^2
a12$<$-(sigma-mu)*exp(.5-mu/sigma)
a22$<$-exp(1-2*mu/sigma)
J1$<$--n*matrix(c(a11,a12,a12,a22),2,2)
b11$<$-(2*sigma-mu)^2+sigma^2*exp(1)/(exp(1)-1)
b12$<$-(2*sigma-mu)*exp(.5-mu/sigma)
b22$<$-exp(1-2*mu/sigma)
J2$<$-n*(exp(1)-1)*matrix(c(b11,b12,b12,b22),2,2)
c1$<$-exp(mu-sigma/2)*sigma^2*(digamma(1-sigma)*gamma(1-sigma)...
+gamma(1-sigma)*(.5-mu/sigma))
c2$<$-gamma(1-sigma)*sigma*exp(mu-sigma/2+.5-mu/sigma)
mo$<$-matrix(c(c1,c2),1,2)
G$<$-mo%*%solve(J1)%*%J2%*%solve(J1)%*%t(mo)
return(G)          }
biw1$<$-function(mu,sigma){
a1$<$-gamma(1-sigma)-exp((sigma^2+sigma)/2)
a3$<$-exp(mu-sigma/2)
G$<$-a1*a3
return(G)        }
Miw1$<$-function(n,mu,sigma){
a1$<$-Viw1(n,mu,sigma)
```

```
a3$<$-biw1(mu,sigma)^2
G$<$-a1+a3
return(G)            }
rbiw1$<$-function(sigma){
a1$<$-gamma(1-sigma)*exp(-(sigma^2+sigma)/2)
G$<$-a1-1
return(G)       }
rMiw1$<$-function(n,sigma){
a1$<$-Miw1(n,1,sigma)
a3$<$-Vh1(n,1,sigma)
G$<$-a1/a3
return(G)        }

##Fig 1,2
x$<$-seq(0,.75,.005)
x1$<$-x[x!=round(x)]
y$<$-rbiw1(x1)
plot(x1,y,'l',ylab=expression(paste('rb'[' IW|l']')),xlab=expression(sigma))
x$<$-seq(.75,.95,.005)
x1$<$-x[x!=round(x)]
y$<$-rbiw1(x1)
plot(x1,y,'l',ylab=expression(paste('rb'['IW|l']')),xlab=expression(sigma))
## Fig 3
x$<$-seq(.01,.25,.025)
y10$<$-rMiw1(10,x)
y50$<$-rMiw1(50,x)
y80$<$-rMiw1(80,x)
y100$<$-rMiw1(100,x)
plot(x,seq(1,5,length=length(x)),' n',ylab=expression(paste(' rm'[' IW| l']'))...
,xlab=expression(sigma),pch=20)
lines(x,y100,' b',pch=20)
lines(x,y80,' b',pch=3)
lines(x,y50,' b',pch=17)
lines(x,y10,' b',pch=8)
legend(' topleft',legend=c(' n=100',' n=80',' n=50',' n=10'),pch=c(20,3,18,8))
###
si1$<$-digamma(1);si11$<$-trigamma(1);si12$<$-psigamma(1,2);si13$<$-psigamma(1,3)
Vgw$<$-function(n,a,l){
e1$<$-(gamma(1-1/a)/a)^2
e2$<$-l^(2/a)/(n*si11)
d1$<$-e1*e2
d2$<$-(1+si1-digamma(1-1/a))^2+si11
d3$<$-d1*d2
return(d3)       }
Vlw$<$-function(n,a,l){
d1$<$-(si11*l^(2/a))/(n*a^2)
e1$<$-si11/(2*a^2)
e2$<$-si13/(2*si11^2)
d2$<$-e1*(1+e2)
d3$<$-(-si12/(a*si11))+1
```

```
d4$<$-exp((-2*si1/a)+(si11/(a^2)))
d5$<$-d1*(d2+d3)*d4
return(d5)          }
Mlw$<$-function(n,a,l){
d1$<$-exp((-si1/a)+(si11/(2*a^2)))-gamma(1-1/a)
d2$<$-l^(2/a)*d1^2
d3$<$-Vlw(n,a,l)
d4$<$-d3+d2
return(d4)         }
rblw$<$-function(a){
d1$<$-exp((-si1/a)+(si11/(2*a^2)))
d2$<$-1/gamma(1-1/a)
d3$<$-d1*d2-1
return(d3)        }
rmlw$<$-function(n,a) Mlw(n,a,1)/Vgw(n,a,1)
##Fig 4
x$<$-seq(1.000001,4,.005)
y$<$-rblw(x)
plot(x,y,' l',ylab=expression(paste(' rb'[' l| IW'])),xlab=expression(alpha))
x$<$-seq(4,20,.005)
y$<$-rblw(x)
plot(x,y,' l',ylab=expression(paste(' rb'[' l| IW'])),xlab=expression(alpha))
##Fig 5
x$<$-seq(1.001,5,.1)
y10$<$-rmlw(10,x)
y50$<$-rmlw(50,x)
y80$<$-rmlw(80,x)
y100$<$-rmlw(100,x)
plot(x,y100,' n',ylab=expression(paste(' rm'[' l| IW'])),xlab=expression(alpha),...
pch=20)
lines(x,y100,' b',pch=20)
lines(x,y80,' b',pch=3)
lines(x,y50,' b',pch=17)
lines(x,y10,' b',pch=8)
legend(' topright',legend=c(' n=100',' n=80',' n=50',' n=10'),pch=c(20,3,18,8))
x$<$-seq(4,18,.5)
y10$<$-rmlw(10,x)
y50$<$-rmlw(50,x)
y80$<$-rmlw(80,x)
y100$<$-rmlw(100,x)
n$<$-length(x)
y0$<$-c(min(y10,y100),y10[2:(n-1)],max(y10,y100))
plot(x,y0,' n',ylab=expression(paste(' rm'[' l| IW'])),xlab=expression(alpha),pch=20)
lines(x,y100,' b',pch=20)
lines(x,y80,' b',pch=3)
lines(x,y50,' b',pch=17)
lines(x,y10,' b',pch=8)
legend(' topright',legend=c(' n=100',' n=80',' n=50',' n=10'),pch=c(20,3,18,8))
```

Heat map#1:

```
# Load required libraries
library(ggplot2)
library(reshape2)

# Manually input the data
data <- data.frame(
  mu = c(1, 1, 1, 1, 4, 4, 4, 4),
  sigma = rep(0.05, 8),
  n = c(100, 100, 30, 30, 30, 30, 100, 100),
  Type = rep(c("simulated", "theoretical"), 4),
  M_IW_l = c(0.0003384081, 0.000393755, 0.0007517901, 0.0009258257,
             0.3044564, 0.3735048, 0.134283, 0.1588521),
  rb_IW_l = c(0.004119081, 0.004729946, 0.003617034, 0.004729946,
              0.003537992, 0.004729946, 0.004181023, 0.004729946),
  rm_IW_l = c(1.815796, 2.123581, 1.236474, 1.497936,
              1.230158, 1.497936, 1.828597, 2.123581)
)

# Reshape to long format
data_long <- melt(data,
                  id.vars = c("mu", "sigma", "n", "Type"),
                  variable.name = "Metric",
                  value.name = "Value")

# Create heatmap
ggplot(data_long, aes(x = factor(n), y = factor(mu), fill = Value)) +
  geom_tile() +
  facet_grid(Metric ~Type, scales = "free") +  # Separate plots for metrics/types
  scale_fill_gradient2(low = "blue", mid = "white", high = "red",...
  midpoint = 0) +  # Color scale
  labs(
    title = "Heatmap of Metrics (Simulated vs. Theoretical)",
    x = "Sample Size (n)",
    y = "\U{3bc}",
    fill = "Value"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    strip.background = element_rect(fill = "lightgray")
  )

Heat map#2:
# Load required libraries
library(ggplot2)
library(reshape2)

# Manually input the data
data <- data.frame(
  alpha = c(3.5, 3.5, 3.5, 3.5, 1.25, 1.25, 1.25, 1.25),
```

```
  lambda = rep(0.9, 8),
  n = c(10, 10, 30, 30, 100, 100, 400, 400),
  Type = rep(c("simulated", "theoretical"), 4),
  M_l_IW = c(0.03292290, 0.03168536, 0.01066024, 0.01069926,
             3.198000, 3.278238, 3.092395, 3.118075),
  rb_l_IW = c(-0.00972015, -0.01159774, -0.01011047, -0.01159774,
              -0.4074761, -0.4148648, -0.4130553, -0.4148648),
  rm_l_IW = c(0.9441920, 0.9882228, 0.9768248, 1.001085,
              1.004237, 1.380777, 4.086844, 5.253269)
)

# Reshape to long format
data_long <- melt(data,
                  id.vars = c("alpha", "lambda", "n", "Type"),
                  variable.name = "Metric",
                  value.name = "Value")

# Create heatmap
ggplot(data_long, aes(x = factor(n), y = factor(alpha), fill = Value)) +
  geom_tile() +
  facet_grid(Metric ~Type, scales = "free") +  # Separate plots by metric and type
  scale_fill_gradient2(
    low = "blue", mid = "white", high = "red", midpoint = 0,
    limits = c(-1, 6)  # Adjust based on your data range
  ) +
  labs(
    title = "Heatmap of Metrics (Simulated vs. Theoretical)",
    x = "Sample Size (n)",
    y = "\U{3b1}",
    fill = "Value"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    strip.background = element_rect(fill = "lightgray"),
    panel.spacing = unit(1, "lines")  # Space between facets
  )

# Subset data for rb_l_IW
rb_data <- subset(data_long, Metric == "rb_l_IW")

# Plot with divergent color scale
ggplot(rb_data, aes(x = factor(n), y = factor(alpha), fill = Value)) +
  geom_tile() +
  facet_wrap(~Type) +
  scale_fill_gradient2(low = "blue", mid = "white", high = "red", midpoint = 0)+
  labs(title = "Bias Comparison (rb_l_IW)", x = "n", y = "\U{3b1}") +
  theme_minimal()
```

REFERENCES

1. W. B. Nelson, Applied Life Data Analysis, vol. 577. John Wiley and Sons, 2005.
2. J. Massey, and J. Frank, The kolmogorov-smirnov test for goodness of fit, Journal of the American statistical Association, vol. 46, no. 253, pp. 68-78, 1951.
3. W. G. Strupczewski, H. T. Mitosek, K. Kochanek, V. P. Singh, and S. Weglarczyk, Probability of correct selection from lognormal and convective diffusion models based on the likelihood ratio. Stochastic Environmental Research and Risk Assessment, vol. 20, no. 3, pp. 152-163, 2006.
4. C. P. Quesenberry, and J. Kent. Selecting among probability distributions used in reliability, Technometrics, vol. 24, no. 1, pp. 59-65, 1982.
5. S. K. Upadhyay, and M. Peshwani, Choice between weibull and lognormal models: A simulation based bayesian study, Communications in Statistics-Theory and Methods, vol. 32, no. 2, pp. 381-405, 2003.
6. R. Pakyari, Discriminating between generalized exponential, geometric extreme exponential and weibull distributions. Journal of Statistical Computation and Simulation, vol. 80, no. 12, pp. 1403-1412, 2010.
7. D. N. Prabhakar Murthy, M. Bulmer, and J. A. Eccleston, Weibull model selection for reliability modelling, Reliability Engineering and System Safety, vol. 86, no. 3, pp. 257-267, 2004.
8. M. Nourbakhsh, Y. Mehrali, A. Jamalizadeh, and Gh. Yari, On a selection of Weibull distribution, Communications in Statistics-Theory and Methods, vol. 44, no. 8, pp. 1640-1652, 2015.
9. A. Barabadi, Reliability model selection and validation using weibull probability plot-a case study, Electric Power Systems Research, vol. 101, pp. 96-101, 2013.
10. H. F. Yu, Miss-specification analysis between normal and extreme value distributions for a linear regression model, Communications in StatisticsTheory and Methods, vol. 36, no. 3, pp. 499-521, 2007.
11. H. F. Yu, The effect of miss-specification between the lognormal and weibull distributions on the interval estimation of a quantile for complete data. Communications in Statistics-Theory and Methods, vol. 41, no. 9, 1617-1635, 2012.
12. F. G. Pascual, Maximum likelihood estimation under misspecified lognormal and weibull distributions, Communications in StatisticsSimulation and Computation, vol. 34, no. 3, pp. 503-524, 2005.
13. H. F. Yu, Miss-specification analysis between normal and extreme value distributions for a screening experiment, Computers and Industrial Engineering, vol. 56, no. 4, pp. 1657-1667, 2009.
14. X. Jia, S. Nadarajah, and B. Guo, The effect of miss-specification on mean and selection between the Weibull and lognormal modelsm, Physica A: Statistical Mechanics and its Applications, vol. 492, pp. 1875-1891, 2017.
15. F. G. Akgul, B. Senoglu, and T. Arslan, An alternative distribution to Weibull for modeling the wind speed data: Inverse Weibull distribution. Energy conversion and management, vol. 114, pp. 234-240, 2016.
16. H. White, Maximum likelihood estimation of misspecified models, Econometrica: Journal of the Econometric Society, pp. 1-25, 1982.