Discriminating between Bivariate Weibull and Bivariate Log-Normal distribution under Type-I censoring

Ojasvi Rajput[1*] Debasis Kundu[2†]and Sharmishtha Mitra[3‡]

[1,2,3] *Department of Mathematics & Statistics, Indian Institute of Technology Kanpur, Kanpur-208016, Uttar Pradesh, India*

Abstract

The motivation for this paper comes from a dataset compiled by the US Insurance Services Office, which is dedicated to studying losses and their related Allocated Loss Adjustment Expenses (ALAE). This dataset features Type-I censoring on loss and associated variable ALAE, placing it in the category of Type-I concomitant censoring. The challenge addressed in this paper is selecting the appropriate model for this specific scenario. Five parameter bivariate Weibull and bivariate lognormal distributions are two bivariate distributions for studying positive lifetime bivariate data. These two distributions have striking similarities right from the surface plot of their joint density function to the nature of their reliability function. Along with that, the shape of their marginal distributions are also quite similar. This paper aims to discriminate between these two distributions under bivariate Type-I censoring based on induced order statistics. In this work, the difference between the maximized log-likelihood values of two models is employed as the basis for discriminating between the corresponding distribution functions. The asymptotic distribution of this discrimination statistic is derived and subsequently used to evaluate the probability of correct selection. To assess the performance of the proposed approach, Monte Carlo experiments are carried out. In addition, a modified version of the discrimination rule is proposed. The effect of model misspecification on correlation estimates is also investigated. Finally, an illustrative example based on a real data set is presented.

*Email : ojrj20@iitk.ac.in
†Email : kundu@iitk.ac.in
‡Email : smitra@iitk.ac.in

Keywords: Bivariate Weibull distribution, bivariate lognormal distributions, Frank copula, Likelihood ratio test, asymptotic distribution, probability of correct selection, Monte Carlo simulations.

1 Introduction and motivation

Almetwally et al. (2020) introduced bivariate Weibull (BVWL) distribution, constructed using Farlie-Gumbel-Morgenstern (FGM) copula, whose marginals are Weibull distribution functions.The five-parameter BVWL distribution has several desirable properties, and it can take on different shapes and can therefore be useful while analyzing bivariate survival data. However, one disadvantage of using FGM copula is that this bivariate distribution could be used to model a bivariate data with only weak correlation (i.e., $\frac{-1}{3} \leq \rho \leq \frac{1}{3}$). For this reason and to take into consideration the complete range of correlation coefficients, we would be working with bivariate Weibull distribution constructed using the Frank copula. Another work that discusses the extension of bivariate Weibull using Frank copula and other choices of copula has been discussed in Lee et al. (2011). Form of Frank copula and construction of bivariate Weibull distribution is discussed as following:

$$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right). \tag{1}$$

Thus,

$$c_\theta(u,v) = \frac{d^2}{dudv}C_\theta(u,v) = \frac{\theta(e^\theta-1)e^{\theta(1+u+v)}}{(e^\theta - e^{\theta(1+u)} - e^{\theta(1+v)} + e^{\theta(u+v)})^2}. \tag{2}$$

The joint probability density function of BVWL is given as:

$$f_{BVWL}(x,y;\Gamma) = f_{WE}(x;\alpha_1,\beta_1)f_{WE}(y;\alpha_2,\beta_2)c_\theta(F_{WE}(x;\alpha_1,\beta_1),F_{WE}(y;\alpha_2,\beta_2)), \tag{3}$$

$$x,y>0; \quad \alpha_1,\alpha_2,\beta_1,\beta_2>0; \quad -\infty<\theta<\infty,$$

where $\Gamma = (\alpha_1,\alpha_2,\beta_1,\beta_2,\theta)$, $F_{WE}(z,\alpha,\beta)$ is cumulative distribution function (CDF) of Weibull distribution and $f_{WE}(z,\alpha,\beta)$ is probability density function (PDF) of Weibull distribution i.e. :

$$f_{WE}(z,\alpha,\beta) = (\alpha/\beta)(z/\beta)^{\alpha-1}e^{-(z/\beta)^\alpha}.$$

Another well-known five-parameter distribution is bivariate log normal distribution(BVLN) which has lognormal distribution as their marginals. The joint probability density function of Bivariate Log Normal Distribution (BVLN) is given as following:

$$f_{BVLN}(x, y; \Sigma) = \frac{1}{2\pi\sigma_1\sigma_2 xy\sqrt{1-\rho^2}} exp\Big(- \frac{Q}{2(1-\rho^2)}\Big), \tag{4}$$

where $\Sigma = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ $x, y > 0$; $-\infty < \mu_1, \mu_2 < \infty$; $\sigma_1, \sigma_2 > 0$; $-1 < \rho < 1$ and

$$Q = \Big(\frac{log(x) - \mu_1}{\sigma_1}\Big)^2 + \Big(\frac{log(y) - \mu_2}{\sigma_2}\Big)^2 - 2\rho\Big(\frac{log(x) - \mu_1}{\sigma_1}\Big)\Big(\frac{log(y) - \mu_2}{\sigma_2}\Big).$$

The bivariate lognormal distribution has been widely applied across various fields. It has been used to model the size distribution of aerosol particles and airborne fibers (Schneider and Holst (1983); Ramachandran and Vincent (1999); Baron (2001)) and to analyze paired PSA values in prostate cancer screening studies (Ross et al. (2000)). Zhou et al. (2001) demonstrated its suitability for comparing healthcare costs before and after interventions. Moreover, the bivariate lognormal distribution has found applications in the insurance sector. In a related direction, Khan et al. (2025) introduced a new generalized logistic class of distributions and illustrated its usefulness through flood and earthquake data analyses, where bivariate lognormal models may also be employed.

We can see the closeness between the 100 data points coming from BVWL and the closest BVLN (and the sample correlation under the respective models) in the scatterplot given in Fig. 1. Moreover, for certain range of parameter values, the marginals of the BVWL and BVLN will be very similar. For seeing similarities between Weibull distribution and lognormal distributions see: Kundu and Manglick (2004).

The problem of determining whether a set of observed data comes from one of multiple probability distribution functions has been an old statistical problem. Cox (1961) initially approached this problem in his influential paper, where he explored the consequences of selecting an incorrect model. Since then, numerous studies have focused on distinguishing between two or more distribution functions. For instance, Dey and Kundu (2009) and related sources have worked into this subject, providing further insights and relevant references. Discrimination between two distribution is very important because model assumption ultimately leads us to work with estimating tail probabilities, percentile points, or reliability function and they are very sensitive when it comes to choosing correct model. Several work has been done in studying the effect of model mis-specification, for example, Wiens (1999),Khakifirooz et al. (2020). Only

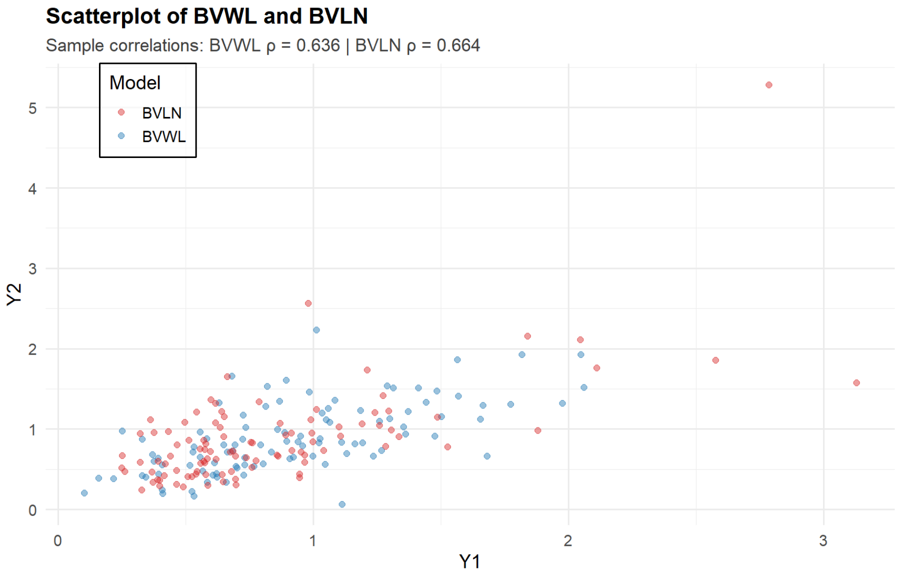

Figure 1: Scatterplot between BVWL$(2, 2, 1, 1, 5)$ and BVLN$(-0.25, -0.25, 0.55, 0.55, 0.60)$.

a few work towards discriminating between bivariate distributions has been done in literature, for example, Dey and Kundu (2012) and Ristić et al. (2018). As we move from a univariate to a bivariate framework, it is important to understand how the correlation structure affects the discrimination process. Additionally, the computation of maximum likelihood estimates (MLE) tends to become more challenging as dimensionality increases. However, no work has been done in discriminating between two bivariate distributions under any kind of censoring setup. When we move from univariate to bivariate setup, the correlation between the two variables holds importance and it is crucial to study the problem of discrimination in bivariate setup. The importance of censoring scheme in bivariate setup followed by importance of discrimination problem is discussed below.

Firstly, let us consider this dataset that is available in R, which was collected by the Insurance Services Office. The data used in the present paper were collected by the US Insurance Services Office, and comprise general liability claims randomly chosen from late settlement lags. This dataset focuses on examining losses and their corresponding ALAE (Allocated Loss Adjustment Expense) within a general liability insurance portfolio. For each of the 1500 randomly selected claims the indemnity payment ("loss"), ALAE, and the policy limit were recorded. Here ALAE's are type of insurance company expenses that are specifically attributable to the settlement of individual claims such as lawyers' fees and claims investigation expenses. In this dataset, there are a total of $n = 1,500$ claims, but 34 of these claims have been censored. Specifically, for

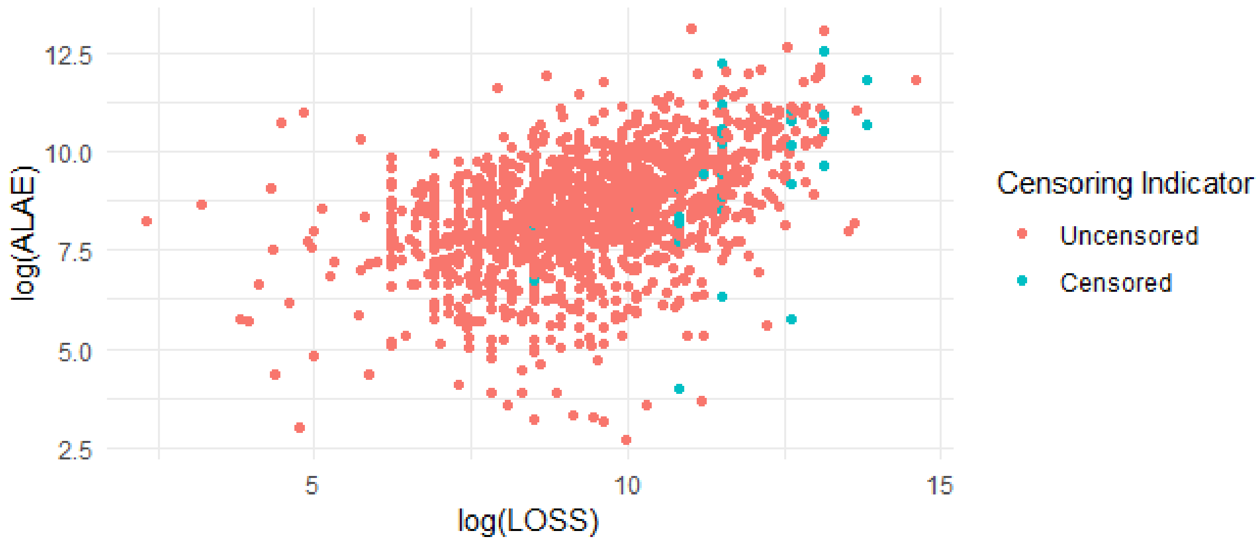

Figure 2: Scatterplot for loss and ALAE (on the log-scale).

each claim, we have information about the observed loss, denoted and the associated ALAE. We denote "ALAE" as X_i and "loss" as Y_i

Additionally, each loss Y_i is linked to a policy limit l_i, which represents the maximum amount that the insurance policy will cover for that particular claim. If the observed loss Y_i exceeds the policy limit l_i, the recorded loss is constrained to l_i.

$$\delta_i = \begin{cases} 1 & \text{if } Y_i \leq \ell_i \\ 0 & \text{if } Y_i > \ell_i \end{cases}$$

for $i = 1, \ldots, n$. A scatterplot of (loss, ALAE) on the log scales is depicted in Figure 2.

Data analysis of this dataset has been discussed in Section 7 where we show that both models can be used to fit this dataset along with possible effect of model misspecification.

A handful of studies have been conducted using this dataset. Frees and Valdez (1998) demonstrated the use of copulas for statistical inference through insurance company data on losses and expenses. They showed how to fit copulas and illustrated their utility in pricing reinsurance contracts and estimating expenses for specified losses. In contrast to our approach, they did not employ a formal bivariate goodness-of-fit test as we did. Instead, they simply compared AIC (Akaike information criterion) values for model selection under a specific choice of marginals (Pareto) and various copula structure. Denuit et al. (2004) proposed a semiparametric modeling strategy to identify the appropriate Archimedean copula that best fits the data. Similarly, Klugman and Parsa (1999) presented a method for fitting bivariate models, assessing goodness-of-fit, and applied this method to the same dataset. Maghsoudi and Bakar (2017) explored several composite models from the Transformed Gamma and Inverse Transformed Gamma families, fitting them to allocated loss adjustment expenses (ALAE).

This dataset motivates our work on discriminating between BVWL and BVLN distribution under bivariate Type-I censoring based on induced order statistics. Now, we formalize our general problem as follows. Suppose (X_i, Y_i), $i = 1, 2, \ldots, n$, is a random sample from an absolutely continuous bivariate population (X, Y). If we order the sample by the Y-variate, and obtain the order statistics, $Y_{1:n}$,$Y_{2:n}$,..., $Y_{n:n}$, for the Y sample, then the X-variate associated with the r-th order statistic $Y_{r:n}$ is called the concomitant of the r-th order statistic, and is denoted by $X_{[r:n]}$. An experimenter might want to investigate how the time until failure is related to a group of associated variables(see Harrell and Sen (1979) for more).

Another domain where concomitants of order statistics come into play is in ranked-set sampling, a concept pioneered by McIntyre (1952). Additionally, concomitants of order statistics have found application in the estimation of parameters for multivariate datasets subjected to some form of type II censoring. Noteworthy examples include the works of Harrell and Sen (1979), Gomes (1984), Gill et al. (1990). For a thorough exploration of these applications, a comprehensive review can be found in Section 9.8 and 11.7 of David and Nagaraja (2004).

To discriminate between the two distributions in this framework, we consider the difference of the maximized log-likelihoods as the discrimination statistic. Since deriving its exact distribution is difficult, we rely on its asymptotic form for inference. One of our main contributions in this paper is development of asymptotic distribution of the test statistic using the idea of Bhattacharyya (1985) and extension of his work to bivariate setup done by Ferenstein (1992). In similar manner, our work can be extended for Type-II censoring based on induced order statistics. This result enables us to compute the probability of correct selection (PCS) by utilizing the properties of the normal distribution, computation of minimum sample size determination, and also working on modified decision rule. Monte Carlo(MC) simulations are performed to study the effectiveness of the proposed method, and it is observed that the usual method is working well. Moreover, we have emphasized the importance of appropriate model selection by studying the effect of model misspecification on the correlation coefficient. Understanding the consequences of choosing an incorrect model is crucial, and we have addressed this by comparing the relative bias in the estimation of the correlation coefficient under respective model assumptions. This is discussed in detail in Section 6.

Some major contributions of this paper deals with how problem of discriminating between

the concerned two distribution shapes up as we move from univariate to bivariate setup under censoring. We have discussed estimation procedures of respective distributions under Type-I concomitant censoring. Extending previous findings, the paper develops asymptotic results for the discrimination criterion, extending from univariate to bivariate setup by building upon the works of Bhattacharyya (1985) and Ferenstein (1992). Furthermore, the paper discusses the effects of model misspecification on the correlation coefficient.
Rest of the paper is organized as follows. In Section 2, we discuss more about the form of data and discrimination procedure that is followed. Required asymptotic properties are presented in Section 3. In Section 4, we present minimum sample size determination and discuss modified approach. Numerical results has been presented in Section 5. We discuss effect of model misspecifications in Section 6. We discuss real life data analysis in Section 7. Finally, we conclude the paper in Section 8.

2 Form of data and Discrimination procedure

Suppose we consider a bivariate random sample $\{(X_i, Y_i),\ i = 1, \ldots, n\}$ consisting of n i.i.d. observations from either the BVWL or BVLN distribution. In practice, however, the full data may not always be available; instead, we may observe a Type-I censored sample of the following form.

Assume that the ordered values of Y_i are observed under a Type-I censoring mechanism, where

$$Y_{1:n} < Y_{2:n} < \cdots < Y_{r:n},$$

with only the first r observations being uncensored, while the remaining $(n - r)$ values are censored at a pre-specified time T. The censoring indicator is defined as

$$\delta_i = \begin{cases} 1 & \text{if } Y_i \leq T, \\ 0 & \text{if } Y_i > T, \end{cases} \qquad i = 1, \ldots, n.$$

For the X-variable, the associated concomitants of the order statistics (see David and Nagaraja (2004)) are observed, denoted by

$$X_{[1:n]},\ X_{[2:n]},\ \ldots,\ X_{[r:n]}.$$

Thus, the observed data can be represented as

$$(X_{[1:n]}, Y_{1:n}),\ (X_{[2:n]}, Y_{2:n}),\ \ldots,\ (X_{[r:n]}, Y_{r:n}). \tag{5}$$

Now, we describe the discrimination procedure based on a random sample discussed above. We will use the following notations throughout this paper. We will be denoting the set of parameters associated with BVWL as $\Gamma = (\alpha_1, \alpha_2, \beta_1, \beta_2, \theta)$ and with that of BVLN as $\Sigma = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$. It is assumed that the above discussed form of data has been generated from one of the two bivariate distributions: $BVWL(\alpha_1, \alpha_2, \beta_1, \beta_2, \theta)$ or $BVLN(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$. Using Lemma 1 of Bhattacharya (1974), loglikelihood function of BVWL can be written as:

$$L_{BVWL}(\Gamma) = \sum_{i=1}^{r} log f_{BVWL}(X_{[i:n]}, Y_{i:n}; \Gamma) + (n-r) log(1 - F_{WE}(T, \alpha_2, \beta_2)).$$

Similarly, we can write loglikelihood function of BVLN as following:

$$L_{BVLN}(\Sigma) = \sum_{i=1}^{r} log f_{BVLN}(X_{[i:n]}, Y_{i:n}; \Sigma) + (n-r) log\left(1 - \Phi\left(\frac{log(T) - \mu_2}{\sigma_2}\right)\right).$$

Here $\Phi(.)$ is CDF of standard normal distribution.
Using the results of Harrell and Sen (1979), we know that based on a Type-I censored data of the form (5), the maximum likelihood estimates of the parameters of the bivariate normal distribution are given by:

$$\hat{\mu_X} = \bar{X} + \left(\frac{S_{XY}}{S_Y^2}\right)\left(\hat{\mu_Y} - \bar{Y}\right), \quad \hat{\sigma_X^2} = S_X^2 + \left(\frac{S_{XY}^2}{S_Y^2}\right)\left(\frac{\hat{\sigma_Y^2}}{S_Y^2} - 1\right), \quad \hat{\rho} = \frac{\hat{\sigma_Y} S_{XY}}{\hat{\sigma_X} S_Y^2},$$

where:

$$\bar{X} = \frac{1}{r}\sum_{i=1}^{r} X_{[i:n]} \quad \bar{Y} = \frac{1}{r}\sum_{i=1}^{r} Y_{i:n} \quad S_X^2 = \frac{1}{r}\sum_{i=1}^{r}(X_{[i:n]} - \bar{X})^2 \quad S_Y^2 = \frac{1}{r}\sum_{i=1}^{r}(Y_{i:n} - \bar{Y})^2$$

$$S_{XY} = \frac{1}{r}\sum_{i=1}^{r}(X_{[i:n]} - \bar{X})(Y_{i:n} - \bar{Y}).$$

The above results helps us in computing maximum likelihood estimators (MLEs) of BVLN denoted as $(\hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{\rho})$. Where as, to find MLEs of BVWL denoted as $(\hat{\alpha_1}, \hat{\alpha_2}, \hat{\beta_1}, \hat{\beta_2}, \hat{\theta})$

were computed using 5-dimensional optimization where initial guesses were given using inference functions for margins method as done in Almetwally et al. (2020). The discrimination procedure is given as following: Choose

a) BVWL if $L_{BVWL}(\hat{\alpha_1}, \hat{\alpha_2}, \hat{\beta}_1, \hat{\beta}_2, \hat{\theta}) > L_{BVLN}(\hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{\rho})$

b) Else, choose BVLN if $L_{BVLN}(\hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{\rho}) > L_{BVWL}(\hat{\alpha_1}, \hat{\alpha_2}, \hat{\beta}_1, \hat{\beta}_2, \hat{\theta})$

Let

$$D_n = L_{BVWL}(\hat{\alpha_1}, \hat{\alpha_2}, \hat{\beta}_1, \hat{\beta}_2, \hat{\theta}) - L_{BVLN}(\hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{\rho}), \tag{6}$$

here D_n is known as Difference of maximized Log-Likelihood and it will be our statistic used for discrimination criteria. If data is coming from $BVWL(\Gamma^0)$, where $\Gamma^0 = (\alpha_1^0, \alpha_2^0, \beta_1^0, \beta_2^0, \theta^0)$, then probability of correct selection is given as: $PCS_{BVWL} = P(D_n > 0|\Gamma^0)$. Similarly, if data is coming from $BVLN(\Sigma^0)$, where $\Sigma^0 = (\mu_1^0, \mu_2^0, \sigma_1^0, \sigma_2^0, \rho^0)$ then PCS is given as:
$PCS_{BVLN} = P(D_n < 0|\Sigma^0)$.

Since limited research exists on discrimination procedures in the bivariate setup and, to the best of our knowledge, no work has been carried out under censoring, there are no established methods available for direct comparison. However, for reference, we employ the "In-and-out-of-sample" (IOS)-based bivariate goodness of fit test discussed in Section 7 for our data analysis. The procedure is as follows: we compute the IOS statistic using equation (21) under each model assumption, namely BVWL and BVLN. If the value of $\mathrm{IOS_{BVWL}} < \mathrm{IOS_{BVLN}}$, the BVWL model is preferred; otherwise, the BVLN model is selected.

3 Asymptotic distribution

In this section, we derive the asymptotic distribution of D_n. This derivation is helpful in calculating the probability of correct selection (PCS) and establishing a correspondence with results obtained through Monte Carlo (MC) simulations. Additionally, this analysis will help experimenter in determining the PCS and the finding minimum sample size required to discriminate between two distributions. In Section 4, we will demonstrate how this asymptotic distribution forms the basis for modified approach.

We will use the following notations. Let $U = (X, Y)$ follow $BVWL(\Gamma^0)$ where $\Gamma^0 = (\alpha_1^0, \alpha_2^0, \beta_1^0, \beta_2^0, \theta^0)$. Then for any functions, $f_1(U)$ and $f_2(U)$, $E_{BVWL}(f_1(U)|\Gamma^0)$, $V_{BVWL}(f_1(U)|\Gamma^0)$ and $Cov_{BVWL}(f_2(U), f_1(U)|\Gamma^0)$ will denote the mean of $f_1(U)$, the variance of $f_1(U)$, and the covariance between $f_1(U)$ and $f_2(U)$, respectively. Similarly, if $U = (X, Y)$ follow $BVLN(\Sigma^0)$, where $\Sigma^0 = (\mu_1^0, \mu_2^0, \sigma_1^0, \sigma_2^0, \rho^0)$ then , for any functions $f_1(U)$ and $f_2(U)$, we define $E_{BVLN}(f_1(U)|\Sigma^0)$, $V_{BVLN}(f_1(U)|\Sigma^0)$ & $Cov_{BVLN}(f_2(U), f_1(U)|\Sigma^0)$ as the mean of $f_1(U)$, the variance of $f_1(U)$, and the covariance between $f_1(U)$ and $f_2(U)$, respectively. Any function $\phi_1(f_{BVWL}(X, Y; \overline{\Gamma}))$ denotes a function of the joint PDF of BVWL for arbitrary parameter set $\overline{\Gamma} = (\overline{\alpha_1}, \overline{\alpha_2}, \overline{\beta_1}, \overline{\beta_2}, \overline{\theta})$. Similarly, any function $\phi_2(f_{BVLN}(X, Y; \overline{\Sigma}))$ denotes a function of joint PDF of BVLN for arbitrary parameters $\overline{\Sigma} = (\overline{\mu_1}, \overline{\mu_2}, \overline{\sigma_1}, \overline{\sigma_2}, \overline{\rho})$. These notations will help in finding mean and variances of D_n and also developing the proof.

3.1 Data is coming from BVLN

Recalling equation (6), we can write,

$$\begin{aligned} D_n =& \sum_{i=1}^{r} log f_{BVWL}(X_{[i:n]}, Y_{i:n}; \hat{\Gamma}) + (n-r) log(1 - F_{WE}(T, \hat{\alpha_2}, \hat{\beta_2})) \\ & - \sum_{i=1}^{r} log f_{BVLN}(X_{[i:n]}, Y_{i:n}; \hat{\Sigma}) + (n-r) log\bigg(1 - \Phi\bigg(\frac{log(T) - \hat{\mu_2}}{\hat{\sigma_2}}\bigg)\bigg). \end{aligned} \tag{7}$$

Therefore, if data is coming from $BVLN(\Sigma^0)$, we have,

$$\begin{aligned} E_{BVLN}(D_n|\Sigma^0) =& E_{BVLN}\Bigg[\sum_{i=1}^{r} log f_{BVWL}(X_{[i:n]}, Y_{i:n}; \hat{\Gamma}) + (n-r) log(1 - F_{WE}(T, \hat{\alpha_2}, \hat{\beta_2})) \\ & - \sum_{i=1}^{r} log f_{BVLN}(X_{[i:n]}, Y_{i:n}; \hat{\Sigma}) + (n-r) log\bigg(1 - \Phi\bigg(\frac{log(T) - \hat{\mu_2}}{\hat{\sigma_2}}\bigg)\bigg)\Bigg]. \end{aligned} \tag{8}$$

Similarly, we have,

$$\begin{aligned} V_{BVLN}(D_n|\Sigma^0) =& V_{BVLN}\Bigg[\sum_{i=1}^{r} log f_{BVWL}(X_{[i:n]}, Y_{i:n}; \hat{\Gamma}) + (n-r) log(1 - F_{WE}(T, \hat{\alpha_2}, \hat{\beta_2})) \\ & - \sum_{i=1}^{r} log f_{BVLN}(X_{[i:n]}, Y_{i:n}; \hat{\Sigma}) + (n-r) log\bigg(1 - \Phi\bigg(\frac{log(T) - \hat{\mu_2}}{\hat{\sigma_2}}\bigg)\bigg)\Bigg]. \end{aligned} \tag{9}$$

Now we present the following theorem which helps us in establishment of asymptotic property of our decision criterion D_n.

Theorem 3.1 *Under the assumption that the censored data are from $BVLN(\Sigma^0)$, the distribution of D_n is approximately normally distributed with mean $E_{BVLN}(D_n|\Sigma^0)$ and variance $V_{BVLN}(D_n|\Sigma^0)$ for large n.*

To prove the above theorem we need the following lemma :

Lemma 3.1 *Under the assumption that the Type-I censored data data is coming from $BVLN(\Sigma^0)$, we have the following results as $n \longrightarrow \infty$:*
(a) $\hat{\Sigma} \to \Sigma^0$ a.s. where

$$\begin{aligned}\Gamma_3(\Sigma^0) &= \max_{\Sigma} \Gamma_3(\Sigma) \text{ here} \\ \Gamma_3(\Sigma) &= \int_0^T \int_0^\infty \ln f_{BVLN}(x,y,\Sigma) f_{BVLN}(x,y,\Sigma^0)\, dx\, dy \\ &+ \ln(1-F_{LN}(T,\mu_2,\sigma_2))(1-F_{LN}(T,\mu_2^0,\sigma_2^0))\end{aligned}$$

(b) $\hat{\Gamma} \to \tilde{\Gamma}$ a.s. where

$$\begin{aligned}\Gamma_4(\tilde{\Gamma}) &= \max_{\Gamma} \Gamma_4(\Gamma) \text{ where} \\ \Gamma_4(\Gamma) &= \int_0^T \int_0^\infty \ln f_{BVWL}(x,y;\Gamma) f_{BVLN}(x,y;\Sigma^0)\, dx\, dy \\ &+ \ln(1-F_{WE}(T,\alpha_2,\beta_2))(1-F_{LN}(T,\mu_2^0,\sigma_2^0))\end{aligned}$$

c) Let us define

$$D_n^\star = \frac{L_{BVWL}(\tilde{\Gamma})}{L_{BVLN}(\Sigma^0)}$$

then $\frac{1}{\sqrt{n}}(D_n - E_{BVLN}(D_n|\Sigma^0))$ is asymptotically equivalent to $\frac{1}{\sqrt{n}}(D_n^\star - E_{BVLN}(D_n^\star))$

Proof of Lemma 3.1: See in appendix
Proof of Theorem 3.1: Using Central limit theorem and part (b) of Lemma 3.1, it follows that $\frac{1}{\sqrt{n}}(D_n^\star - E_{BVLN}(D_n^\star))$ is asymptotically normally distributed. Therefore, using part (c) of Lemma 3.1, the result immediately follows.
It could be seen that values of $\tilde{\Gamma}$ will depend on T and Σ^0. We will be calling them as misspecified parameters. Let us denote

$$AM_{BVLN} = \lim_{n\to\infty} \frac{E_{BVLN}(D_n|\Sigma^0)}{n} \text{ and } AV_{BVLN} = \lim_{n\to\infty} \frac{V_{BVLN}(D_n|\Sigma^0)}{n} \tag{10}$$

Here, AM_{BVLN} and AV_{BVLN} is the asymptotic mean and asymptotic variance of D_n , when data is coming from bivariate log-normal distribution. Now we provide expressions for the asymptotic mean and asymptotic variance as follows. From the results of Theorem 1 in Ferenstein (1992), the expressions for the asymptotic mean and variance depend on the censoring time T, which is determined by the p-th quantile of the parent distribution. Hence, we have,

$$AM_{BVLN}(p) = \int_0^T \int_0^\infty \ln \frac{f_{BVWL}(x,y,\tilde{\Gamma})}{f_{BVLN}(x,y,\Sigma^0)} f_{BVLN}(x,y,\Sigma^0)\, dx\, dy + (1-p)h_1(T). \tag{11}$$

where

$$h_1(T) = ln \frac{(1 - F_{WE}(T, \tilde{\alpha_2}, \tilde{\beta_2}))}{(1 - F_{LN}(T, \mu_2^0, \sigma_2^0))}. \tag{12}$$

Similarly, we obtain,

$$AV_{BVLN}(p) = \tau + \frac{p(1-p)b^2}{f_{LN}(T, \mu_2^0, \sigma_2^0)}, \tag{13}$$

where

$$\begin{aligned}\tau = &\int_0^T \int_0^\infty \left(\ln \frac{f_{BVWL}(x,y,\tilde{\Gamma})}{f_{BVLN}(x,y,\Sigma^0)} \right)^2 f_{BVLN}(x,y,\Sigma^0)\, dx\, dy \\ &- p^{-1} \left(\int_0^T \int_0^\infty \ln \frac{f_{BVWL}(x,y,\tilde{\Gamma})}{f_{BVLN}(x,y,\Sigma^0)} f_{BVLN}(x,y,\Sigma^0)\, dx\, dy \right)^2,\end{aligned}$$

and

$$\begin{aligned}b = &-\int_0^\infty \ln \frac{f_{BVWL}(x,T,\tilde{\Gamma})}{f_{BVLN}(x,T,\Sigma^0)} f_{BVLN}(x,T,\Sigma^0)\, dx \\ &+ p^{-1} f_{LN}(T, \mu_2^0, \sigma_2^0) \int_0^T \int_0^\infty \ln \frac{f_{BVWL}(x,y,\tilde{\Gamma})}{f_{BVLN}(x,y,\Sigma^0)} f_{BVLN}(x,y,\Sigma^0)\, dx\, dy - (1-p)h_1^{'}(T).\end{aligned}$$

where $h_1^{'}(T)$ is derivative of $h_1(T)$ as defined in equation (12).

3.2 Data is coming from BVWL

As discussed in previous subsection, if data is coming from $BVWL(\Gamma^0)$, we have

$$\begin{aligned}E_{BVWL}(D_n|\Gamma^0) = &E_{BVWL}\Bigg[\sum_{i=1}^r log f_{BVWL}(X_{[i:n]}, Y_{i:n}; \hat{\Gamma}) + (n-r) log(1 - F_{WE}(T, \hat{\alpha_2}, \hat{\beta_2})) \\ &- \sum_{i=1}^r log f_{BVLN}(X_{[i:n]}, Y_{i:n}; \hat{\Sigma}) + (n-r) log\left(1 - \Phi\left(\frac{log(T) - \hat{\mu_2}}{\hat{\sigma_2}}\right)\right)\Bigg].\end{aligned} \tag{14}$$

Similarly, we have,

$$V_{BVWL}(D_n|\Gamma^0) = V_{BVWL}\Bigg[\sum_{i=1}^{r} log f_{BVWL}(X_{[i:n]}, Y_{i:n}; \hat{\Gamma}) + (n-r) log(1 - F_{WE}(T, \hat{\alpha_2}, \hat{\beta_2}))$$
$$- \sum_{i=1}^{r} log f_{BVLN}(X_{[i:n]}, Y_{i:n}; \hat{\Sigma}) + (n-r) log\Bigg(1 - \Phi\Bigg(\frac{log(T) - \hat{\mu_2}}{\hat{\sigma_2}}\Bigg)\Bigg)\Bigg]. \quad (15)$$

Theorem 3.2 *Under the assumption that the censored data are from $BVWL(\Gamma^0)$, the distribution of D_n is approximately normally distributed with mean $E_{BVWL}(D_n|\Gamma^0)$ and variance $V_{BVWL}(D_n|\Gamma^0)$ for large n.*

To prove the above theorem we need the following lemma :

Lemma 3.2 *Under the assumption that the Type-I censored data data is coming from $BVWL(\Gamma^0)$, we have the following results as $n \longrightarrow \infty$:*
(a) $\hat{\Gamma} \rightarrow \Gamma^0$ a.s. where

$$\begin{aligned}
\Gamma_3(\Gamma^0) &= \max_{\Gamma} \Gamma_3(\Gamma) \text{ here} \\
\Gamma_3(\Gamma) &= \int_0^T \int_0^\infty \ln f_{BVWL}(x, y, \Gamma) f_{BVLN}(x, y, \Gamma^0)\, dx\, dy \\
&+ \ln(1 - F_{WE}(T, \alpha_2, \beta_2))(1 - F_{WE}(T, \alpha_2^0, \beta_2^0))
\end{aligned}$$

(b) $\hat{\Sigma} \rightarrow \tilde{\Sigma}$ a.s. where

$$\begin{aligned}
\Gamma_4(\tilde{\Sigma}) &= \max_{\Sigma} \Gamma_4(\Sigma) \text{ where} \\
\Gamma_4(\Sigma) &= \int_0^T \int_0^\infty \ln f_{BVLN}(x, y; \Sigma) f_{BVWL}(x, y; \Gamma^0)\, dx\, dy \\
&+ \ln(1 - F_{LN}(T, \mu_2, \sigma_2))(1 - F_{WE}(T, \alpha_2^0, \beta_2^0))
\end{aligned}$$

c) Let us define

$$D_{n\star} = \frac{L_{BVWL}(\Gamma^0)}{L_{BVLN}(\tilde{\Sigma})}$$

then $\frac{1}{\sqrt{n}}(D_n - E_{BVWL}(D_n|\Gamma^0))$ is asymptotically equivalent to $\frac{1}{\sqrt{n}}(D_{n\star} - E_{BVWL}(D_{n\star}))$

Proof of Lemma 3.2: Same as proof of lemma 3.1.
Proof of Theorem 3.2: Along the same lines as the Proof of Theorem 3.1 it also follows using Lemma 3.2.

It could be seen that values of $\tilde{\Sigma}$ will depend on T and Γ^0. We will be calling them as misspecified parameters. Let us denote

$$AM_{BVWL} = \lim_{n\to\infty} \frac{E_{BVWL}(D_n|\Gamma^0)}{n} \text{ and } AV_{BVWL} = \lim_{n\to\infty} \frac{V_{BVWL}(D_n|\Gamma^0)}{n} \tag{16}$$

Here, AM_{BVWL} and AV_{BVWL} is the asymptotic mean and asymptotic variance of D_n , when data is coming from bivariate Weibull distribution. Now we provide expressions for asymptotic mean and asymptotic variance. Similarly as $AM_{BVLN}(p)$ and $AV_{BVLN}(p)$, $AM_{BVWL}(p)$ and $AV_{BVWL}(p)$ can be obtained using Theorem 1 of Ferenstein (1992) and they are as follows:

$$AM_{BVWL}(p) = -\int_0^T \int_0^\infty \ln \frac{f_{BVLN}(x,y,\tilde{\Sigma})}{f_{BVWL}(x,y,\Gamma^0)} f_{BVWL}(x,y,\Gamma^0)\,dx\,dy - (1-p)h_2(T), \tag{17}$$

where

$$h_2(T) = ln \frac{(1 - F_{LN}(T, \tilde{\mu_2}, \tilde{\sigma_2})}{(1 - F_{WE}(T, \alpha_2^0, \beta_2^0))}. \tag{18}$$

Similarly, we obtain,

$$AV_{BVWL}(p) = \tau + \frac{p(1-p)b^2}{f_{WE}(T, \alpha_2^0, \beta_2^0)}, \tag{19}$$

where

$$\begin{aligned}\tau = &\int_0^T \int_0^\infty \left(\ln \frac{f_{BVLN}(x,y,\tilde{\Sigma})}{f_{BVWL}(x,y,\Gamma^0)} \right)^2 f_{BVWL}(x,y,\Gamma^0)\,dx\,dy \\ &- p^{-1} \left(\int_0^T \int_0^\infty \ln \frac{f_{BVLN}(x,y,\tilde{\Sigma})}{f_{BVWL}(x,y,\Gamma^0)} f_{BVWL}(x,y,\Gamma^0)\,dx\,dy \right)^2,\end{aligned}$$

and

$$\begin{aligned}b = &\int_0^\infty \ln \frac{f_{BVLN}(x,y,\tilde{\Sigma})}{f_{BVWL}(x,T,\Gamma^0)} f_{BVWL}(x,T,\Gamma^0)\,dx \\ &- p^{-1} f_{WE}(T, \alpha_2^0, \beta_2^0) \int_0^T \int_0^\infty \ln \frac{f_{BVLN}(x,y,\tilde{\Sigma})}{f_{BVWL}(x,y,\Gamma^0)} f_{BVWL}(x,y,\Gamma^0)\,dx\,dy + (1-p)h_2'(T).\end{aligned}$$

where $h_2'(T)$ is derivative of $h_2(T)$ as defined in equation (18).

4 Determination of sample size and modified approach

In this section, we begin by presenting a procedure to determine the minimum sample size required for discriminating between the two distribution functions, and subsequently introduce a modified decision rule.

Modified approach is useful as it minimises the total probability of misclassification and hence leads to an overall increase in probability of correct selection.

4.1 Calculation of Minimum Sample Size

We now turn to the problem of determining the minimum sample size required to guarantee a prescribed level of protection in the model discrimination problem.

Consider that the data are generated from a $BVLN(\Sigma^0)$ distribution, where the censoring threshold T is chosen as the p^{th} quantile of the marginal distribution $LN(\mu_1^0, \sigma_1^0)$. Here, the parameter $p \in (0, 1)$ controls the censoring level: smaller p corresponds to heavier censoring , while larger p corresponds to lighter censoring.

Since D_n is asymptotically normal with mean $E_{BVLN}(D_n \mid \Sigma^0)$ and variance $V_{BVLN}(D_n \mid \Sigma^0)$, the probability of correct selection (PCS) under the $BVLN$ model is

$$PCS_{BVLN} = P(D_n \leq 0).$$

Analogously, when the data follow a $BVWL(\Gamma^0)$ distribution, the PCS is given by

$$PCS_{BVWL} = P(D_n \geq 0).$$

To formalize the sample size requirement, fix a censoring level p and a desired protection level $\alpha^* \in (0, 1)$. The protection level α^* specifies the minimum PCS that must be guaranteed regardless of whether the true distribution is BVLN or BVWL. That is, both conditions

$$\Phi\left(\frac{-n \cdot AM_{BVLN}(p)}{\sqrt{n \cdot AV_{BVLN}(p)}}\right) = \alpha^*, \quad \text{and} \quad \Phi\left(\frac{n \cdot AM_{BVWL}(p)}{\sqrt{n \cdot AV_{BVWL}(p)}}\right) = \alpha^*, \tag{20}$$

must hold, where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution.

Solving (20) leads to the required sample sizes under the two competing models:

$$n_1 = \frac{z_{\alpha^*}^2\, AV_{BVLN}(p)}{\big(AM_{BVLN}(p)\big)^2}, \qquad n_2 = \frac{z_{\alpha^*}^2\, AV_{BVWL}(p)}{\big(AM_{BVWL}(p)\big)^2},$$

where z_{α^*} is the α^*-quantile of the standard normal distribution. The overall minimum sample size is then taken as

$$n = \max\{n_1, n_2\}.$$

Table 1 reports the analytically computed values of $AM_{BVLN}(p)$, $AV_{BVLN}(p)$, $AM_{BVWL}(p)$, and $AV_{BVWL}(p)$ for several censoring levels p. These values are then substituted into the above formula to calculate the required minimum sample sizes, shown in Table 2, for different choices of α^* and p.

p	0.40	0.50	0.60	0.70	0.80
$AM_{BVLN}(p)$	-0.0290	-0.0403	-0.0530	-0.0690	-0.0886
$AV_{BVLN}(p)$	0.0471	0.0824	0.0920	0.1203	0.1626
$AM_{BVWL}(p)$	0.0456	0.0660	0.0892	0.1173	0.2423
$AV_{BVWL}(p)$	0.1778	0.2651	0.3654	0.4923	1.0847

Table 1: Values of $AM_{BVLN}(p)$, $AV_{BVLN}(p)$, $AM_{BVWL}(p)$, and $AV_{BVWL}(p)$ for different censoring quantiles p.

$\alpha^* \downarrow \;\; p \rightarrow$	0.40	0.50	0.60	0.70	0.80
0.90	141	100	76	59	34
0.95	232	165	125	97	57
0.99	463	330	249	194	113

Table 2: Minimum sample size required to achieve protection level α^* for different censoring quantiles p.

Table 2 reveals a clear trend: for larger values of p (corresponding to lighter censoring), discrimination between BVLN and BVWL becomes easier, as reflected in the smaller required sample sizes. In contrast, when p is small (heavier censoring) and the protection level α^* is high, considerably larger sample sizes are necessary to achieve required discrimination.

4.2 Modified Decision Rule

So far, the discrimination rule has relied solely on the sign of D_n (i.e., choosing BVWL if $D_n > 0$ and BVLN if $D_n < 0$). This simple rule is widely used in the literature, but it does not necessarily maximize the overall probability of correct selection (PCS).

Motivated by the fact that the asymptotic distribution of D_n is normal under both models, we now introduce a modified rule:

$$\text{Select BVWL if } D_n > C^*,$$

where the constant C^* is chosen to minimize the total probability of misclassification. Specifically, C^* is determined as the minimizer of

$$1 - \Phi\left(\frac{C^* - \mu_2^*}{\sigma_2^*}\right) + \Phi\left(\frac{C^* - \mu_1^*}{\sigma_1^*}\right),$$

with (μ_1^*, σ_1^*) denoting the mean and variance under BVWL, and (μ_2^*, σ_2^*) the corresponding quantities under BVLN.

The constant C^* can be obtained by solving a quadratic equation involving these parameters. For practical use, Table 3 provides the computed values of C^* for various censoring levels p.

p	0.40	0.50	0.60	0.70	0.80
C^*	0.2405	0.2953	0.3218	0.3598	0.4740

Table 3: Optimal cutoff values C^* for the modified decision rule under different censoring quantiles p.

5 Numerical Results

In this section, we present simulation results to examine the performance of probability of Correct Selection (PCS) calculations based on asymptotic distributions derived in Section 3. The objective is to show to trend of the values of these PCS across various sample sizes and censoring time T. We explore different sample sizes ($n = 25, 50, 75, 100, 150$) and various censoring time which is governed by value of p-th quantile of marginal Y ($p = 0.8, 0.7, 0.6, 0.5, 0.4$).

PCS values are computed through Monte Carlo(MC) simulation and compared with those derived from asymptotic results in Section 3 along with IOS method. For a given n and T, a sample of size n is generated either from a BVWL distribution with parameters $(2, 2, 1, 1, 5)$

or from a BVLN distribution with parameters $(0, 0, 1, 1, 0.5)$. D_n is then computed based on the Type-I censored sample. The condition $D_n \leq 0$ or $D_n > 0$ is checked, and the process is replicated 10,000 times to obtain the PCS. PCS values are also computed based on the asymptotic results from Section 4, and the outcomes are presented in Tables 4 and 6.

Additionally, PCS values are computed using the modified method proposed in Section 6, incorporating the C^* values reported in Table 3. The modified method is also based on 10,000 replications, and the results are presented in Tables 5 and 7.

We can conclude from the tables that, as expected, the PCS increases with increasing T for fixed n and with increasing n for fixed T. When both T and n are large, the asymptotic results align well with simulated results. A comparison between Tables 4 and 6 with Tables 5 and 7 reveals that the modified method performs slightly better than the classical one. Specifically, the total PCS (PCS under BVVWL + PCS under BVLN) is larger for the modified method than the classical method.

5.1 Parent distribution is BVWL

First, we generate samples from $BVWL(2, 2, 1, 1, 5)$ and for varying values of censoring time T. PCS is obtained for increasing sample size using both MC simulations as well as asymptotic distribution. The results are shown in Table 4.

Method ↓	$n = 25$	$n = 50$	$n = 75$	$n = 100$	$n = 150$
$T = 0.71$					
MC	0.606	0.743	0.835	0.866	0.919
AS	0.707	0.779	0.827	0.862	0.909
IOS	0.572	0.631	0.741	0.802	0.865
$T = 0.83$					
MC	0.645	0.780	0.855	0.922	0.965
AS	0.757	0.838	0.887	0.919	0.956
IOS	0.590	0.675	0.775	0.838	0.873
$T = 0.95$					
MC	0.669	0.832	0.908	0.953	0.989
AS	0.769	0.851	0.899	0.930	0.964
IOS	0.615	0.707	0.798	0.856	0.898
$T = 1.09$					
MC	0.713	0.886	0.927	0.970	0.997
AS	0.798	0.881	0.926	0.952	0.979
IOS	0.645	0.749	0.826	0.874	0.919
$T = 1.26$					
MC	0.751	0.903	0.954	0.988	0.998
AS	0.877	0.950	0.978	0.990	0.997
IOS	0.673	0.768	0.859	0.882	0.941

Table 4: Values of PCS using MC simulation, asymptotic results, and IOS method.

We can see as value of T increases, PCS increases. Moreover PCS increases along with sample size which is intuitive. We can also observe the correspondence between asymptotic results with that of value of PCS obtained using MC simulation when sample size is moderate. However, the PCS obtained from IOS method is fairly low in comparison to other two methods. Using the modified approach, we computed PCS shown in Table 5.

Method ↓	$n = 25$	$n = 50$	$n = 75$	$n = 100$	$n = 150$
$T = 0.71$					
MC	0.587	0.703	0.819	0.873	0.912
AS	0.665	0.753	0.808	0.847	0.899
IOS	0.572	0.631	0.741	0.802	0.865
$T = 0.83$					
MC	0.603	0.751	0.849	0.909	0.956
AS	0.700	0.795	0.851	0.889	0.936
IOS	0.590	0.675	0.775	0.838	0.873
$T = 0.95$					
MC	0.628	0.823	0.893	0.948	0.982
AS	0.736	0.833	0.888	0.922	0.961
IOS	0.615	0.707	0.798	0.856	0.898
$T = 1.09$					
MC	0.695	0.861	0.918	0.969	0.989
AS	0.768	0.866	0.917	0.947	0.977
IOS	0.645	0.749	0.826	0.874	0.919
$T = 1.26$					
MC	0.726	0.898	0.951	0.984	0.996
AS	0.858	0.943	0.975	0.988	0.997
IOS	0.673	0.768	0.859	0.882	0.941

Table 5: PCS obtained by MC simulation, asymptotic results (based on modified approach), and IOS GOF method.

We can see that in this case the PCS so obtained is less than the usual method but the overall probability of correct selection(see the comparison of Table 4 and 6 with Table 5 and 7) is greater than the usual method.

5.2 Parent distribution is BVLN

In this case, data is generated from $\text{BVLN}(0, 0, 1, 1, 0.5)$ and similar exercise is executed as done for previous case.

Method →	$n = 25$	$n = 50$	$n = 75$	$n = 100$	$n = 150$
$T = 0.77$					
MC	0.697	0.745	0.801	0.834	0.893
AS	0.748	0.828	0.876	0.909	0.934
IOS	0.663	0.712	0.782	0.826	0.878
$T = 0.99$					
MC	0.741	0.842	0.874	0.917	0.952
AS	0.758	0.839	0.888	0.920	0.957
IOS	0.702	0.789	0.845	0.888	0.932
$T = 1.28$					
MC	0.802	0.882	0.932	0.946	0.984
AS	0.809	0.891	0.935	0.959	0.983
IOS	0.771	0.846	0.898	0.926	0.963
$T = 1.68$					
MC	0.830	0.927	0.963	0.970	0.990
AS	0.840	0.920	0.957	0.976	0.992
IOS	0.789	0.871	0.926	0.953	0.977
$T = 2.32$					
MC	0.862	0.946	0.965	0.987	0.997
AS	0.864	0.939	0.971	0.986	0.996
IOS	0.821	0.903	0.941	0.969	0.989

Table 6: PCS obtained by MC simulation, asymptotic results, and IOS GOF method.

Similar conclusions could also be made for this case. From Table 6, it is evident that the PCS values tend to increase with increasing T as well as with larger sample sizes, which aligns with asymptotic results. The asymptotic results closely approximate the PCS values obtained via Monte Carlo (MC) simulations, especially when the sample size is moderate to large. In contrast, the IOS based PCS values remain consistently lower than those from the other two methods across all scenarios considered.

Method →	$n = 25$	$n = 50$	$n = 75$	$n = 100$	$n = 150$
$T = 0.77$					
MC	0.732	0.804	0.824	0.849	0.921
AS	0.813	0.864	0.900	0.926	0.957
IOS	0.663	0.712	0.782	0.826	0.878
$T = 0.99$					
MC	0.808	0.870	0.902	0.940	0.988
AS	0.817	0.872	0.908	0.934	0.964
IOS	0.702	0.789	0.845	0.888	0.932
$T = 1.28$					
MC	0.857	0.923	0.957	0.973	0.988
AS	0.861	0.917	0.949	0.968	0.987
IOS	0.771	0.846	0.898	0.926	0.963
$T = 1.68$					
MC	0.865	0.942	0.977	0.985	0.992
AS	0.885	0.939	0.967	0.981	0.994
IOS	0.789	0.871	0.926	0.953	0.977
$T = 2.32$					
MC	0.915	0.960	0.987	0.994	0.998
AS	0.908	0.957	0.979	0.989	0.997
IOS	0.821	0.903	0.941	0.969	0.989

Table 7: PCS obtained by MC simulation, asymptotic results (based on modified approach), and IOS GOF method.

6 Effect of Model misspecification

Now we focus briefly on the effects of model misspecification. To examine the effect of model misspecification on the estimation of the correlation coefficient, let us consider a censored sample of the form as described in equation (5). Assume that the data actually follows a bivariate lognormal (BVLN) distribution, and our goal is to estimate Spearman's correlation coefficient. We are particularly interested in understanding the consequences if an experimenter mistakenly assumes the data follows a bivariate Weibull (BVWL) distribution instead, and estimates the correlation coefficient under this incorrect model. In such a scenario, we aim to investigate how severe the impact of this model misspecification is on the estimated correlation coefficient. Suppose, data is coming from BVLN(Σ^0) and we find MLE of ρ assuming data is indeed coming from BVLN(say $\hat{\rho_1}$) and then assuming wrong model i.e. BVWL(say $\hat{\rho_2}$). Then, Relative bias under true model is given as:

$$RB_1 = \frac{\rho_0 - \hat{\rho_1}}{\rho_0},$$

where as that of under wrong model is given as:

$$RB_2 = \frac{\rho_0 - \hat{\rho_2}}{\rho_0}.$$

Firstly, we observe the effect of model misspecification under complete sample when data is coming from $BVWL(2, 2, 1, 1, \theta)$. Results are shown in Table 8.

$\rho \downarrow$ n $\rightarrow$	25	50	75	100
0.08	-0.06433	-0.05875	-0.00357	-0.00214
	(0.19059)	(0.20875)	(0.21393)	(0.22595)
0.30	0.05706	0.04039	0.01196	0.00552
	(0.09631)	(0.14668)	(0.15192)	(0.15290)
0.65	0.00938	0.00209	0.00108	0.00107
	(0.09395)	(0.09460)	(0.10506)	(0.11305)

Table 8: Value above the parenthesis denotes average value of Relative bias under true model where as that value in the parenthesis denotes that of under wrong model(BVLN in this case).

We can see that as sample size increases the relative bias for true model tends towards zero while that under wrong model gets further away from zero which is expected. Moreover, as value of true ρ_0 decreases, value of relative bias under wrong model gets larger. Further, we considered effect of model misspecification under censored sample for varying values of censoring time T and for fixed value of θ and the results can be found in Table 9.

$T \downarrow$ n $\rightarrow$	25	50	75	100
0.71	0.06522	0.05942	0.02126	0.00672
	(0.11631)	(0.15768)	(0.16892)	(0.16292)
0.95	0.05926	0.04729	0.01786	0.00552
	(0.09631)	(0.14668)	(0.15192)	(0.15290)
1.26	0.04321	0.04129	0.01586	0.00481
	(0.08426)	(0.11426)	(0.14292)	(0.15106)

Table 9: Value above the parenthesis denotes average value of Relative bias under true model where as that value in the parenthesis denotes that of under wrong model(BVLN in this case).

We changed the choice of parent distribution to BVLN and did the similar exercise. The required results can be found in Tables 10 and 11

$\rho \downarrow$ n $\rightarrow$	25	50	75	100
0.08	0.09880	0.07055	0.02667	0.01293
	(-0.15895)	(-0.21298)	(-0.22710)	(-0.23595)
0.30	-0.01372	-0.00312	0.01206	0.01105
	(-0.10492)	(-0.11524)	(-0.12688)	(-0.13717)
0.65	0.00898	0.00643	0.00408	-0.00073
	(-0.03762)	(-0.06466)	(-0.07169)	(-0.07882)

Table 10: Value above the parenthesis denotes average value of Relative bias under true model where as that value in the parenthesis denotes that of under wrong model(BVWL in this case).

$T \downarrow$ n $\rightarrow$	25	50	75	100
0.77	0.01271	0.00944	0.00608	-0.00163
	(-0.04772)	(-0.07216)	(-0.07460)	(-0.08012)
1.28	0.00998	0.00843	0.00558	-0.00103
	(-0.03962)	(-0.06466)	(-0.07369)	(-0.07882)
2.32	0.008715	0.00717	0.00516	-0.00013
	(-0.03762)	(-0.06266)	(-0.07169)	(-0.07782)

Table 11: Value above the parenthesis denotes average value of Relative bias under true model where as that value in the parenthesis denotes that of under wrong model(BVWL in this case).

Thus, we see there is observable effect of model misspecification on correlation coefficient and it is important that appropriate model selection is made.

7 Real data analysis

Here, we consider the data set that has been presented in Section 1. We start with checking if BVLN distribution could be used to fit the given bivariate data set or not. To do so we implemented “In-and-out-of-sample" (IOS) goodness of fit test as introduced in Presnell and Boos (2004)(see Zhang et al. (2016),Mei (2016) for more). The methodology for the given goodness of fit test in this case can be stated as:

Step 1: Start with i^{th} component of loglikelihood, say $l(\theta; x_i, y_i)$, where θ is the vector of parameters for the given distribution. Let $\hat{\theta}$ denotes maximum likelihood estimator(MLE) of θ based on complete sample. In-sample loglikelihood is given as $\sum_{i=1}^{n} l(\hat{\theta}; x_i, y_i)$

Step 2: Let $\hat{\theta_{(-i)}}$ denotes MLE when i^{th} sample is deleted. Out-of-sample loglikelihood is given as $\sum_{i=1}^{n} l(\hat{\theta_{(-i)}}; x_i, y_i)$

Step 3: IOS statistic is given as:

$$IOS = \sum_{i=1}^{n} (l(\hat{\theta}; x_i, y_i) - l(\hat{\theta_{(-i)}}; x_i, y_i). \tag{21}$$

Finally with help of parametric bootstrap p values decision is made.

In this scenario, value of MLE's(in case of BVLN) were as follows(respective standard errors within parentheses): $\hat{\mu_1} = 9.391(0.052)$, $\hat{\mu_2} = 8.528(0.046)$, $\hat{\sigma_1} = 1.667(0.042)$, $\hat{\sigma_2} = 1.422(0.039)$,
$\hat{\rho} = 0.443(0.028)$. The value of test statistic so obtained was 3.53 along with p-value as 0.37. Similarly, we fit BVWL distribution and values of MLE's were reported as following (respective standard errors within parentheses): $\hat{\alpha_1} = 0.459(0.026)$, $\hat{\alpha_2} = 0.718(0.053)$, $\hat{\beta}_1 = 8090.422(1788.572)$, $\hat{\beta}_2 = 9694.095(1713.236)$ and $\hat{\theta} = 2.74(1.010)$ and the value of test statistic so obtained was 3.91 along with p-value as 0.29. Therefore, we can conclude that both the distribution fit the given dataset. Now, even though both the models fit the given data set, we focus on why there is a need for appropriate model selection. Firstly, value of Spearman's correlation under BVLN model(respective standard errors within parentheses) is 0.44(0.028) where as that under BVWL model is 0.40(0.100). Infact, we saw in Section 6 that there is serious effect of model misspecification on correlation coefficient. Moreover, we also computed value of Stress-Strength parameter $P(X > Y)$. For the case of BVLN the value of stress strength parameter(respective standard errors within parentheses)=0.302(0.082) where as that under BVWL model=0.570(0.116). Therefore, we see that in spite of the similarities appropriate model selection is needed. To evaluate the adequacy of the competing dependence models, a predictive check was carried out

using the real dataset. For each uncensored observation (x_i, y_i), one data pair was omitted at a time, and the remaining observations were used to estimate the model parameters under both the bivariate Weibull with Frank copula (BVWL) and the bivariate lognormal (BVLN) formulations. The conditional expectations $E[X|Y]$ and $E[Y|X]$ were then computed for the omitted pair, and the absolute prediction errors $|x_i - \widehat{E}[X|Y_i]|$ and $|y_i - \widehat{E}[Y|X_i]|$ were recorded. The overall predictive performance was summarized by the mean absolute error (MAE) over all uncensored observations.

Table 12 shows that the total prediction error was consistently smaller for the BVLN model in both directions. Specifically, the MAE for $E[X|Y]$ was lower under BVLN compared to BVWL, and a similar pattern was observed for $E[Y|X]$. This indicates that the bivariate lognormal model provides better predictive capability for the present data, suggesting that its underlying dependence structure more appropriately captures the relationship between the two variables.

Prediction Direction	**BVLN (MAE)**	**BVWL (MAE)**
$E[X\|Y]$	8878.77	8893.31
$E[Y\|X]$	30438.29	31103.62

Table 12: Observed Mean Absolute Error (MAE) for uncensored data under both models.

To choose the preferred model, we first compare the log-likelihood values obtained under both models. The log-likelihood for the BVWL model is given by $L_{\text{BVWL}} = -31932.07$, whereas for the BVLN model it is $L_{\text{BVLN}} = -31435.62$. Hence, the corresponding value of the statistic is $D_n = -496.45$. Based on this value, the BVLN model is preferred.

Furthermore, when comparing the IOS statistics, the value obtained under the BVLN model (3.53) is lower than that under the BVWL model (3.91). This again indicates that the BVLN model provides a better fit to the data.

Moreover, if we look at the marginals for "ALAE" variable, then loglikelihood under Weibull distribution i.e. $L(ALAE)_{WE} = -16934.12$ and that under log-normal, $L(ALAE)_{LN} = -16535.2$. Similarly, when we look at "Loss" variable, we have,$L(Loss)_{WE} = -15097.03$ and that under log-normal, $L(Loss)_{LN} = -15051.57$. Thus from marginals, as well, we have evidence that preferred model is BVLN.

Moreover, we artificially increased percentage of censored observation and sign of D_n is observed and the results are stated in Table 13.

degree of censoring(%)	D_n	Preferred model
5	<0	BVLN
10	<0	BVLN
15	<0	BVLN
20	<0	BVLN

Table 13: Preferred model on basis of discrimination criterion for different degree of censoring.

8 Conclusion

In this paper we have considered discrimination between BVWL and BVLN distributions under bivariate Type-I censoring based on induced order statistics. For motivational purpose, we have presented some simulation results to reflect some effects of model misspecification on correlation coefficient. The discrimination statistic utilized in this study is the difference between the maximized log-likelihood values. We have also derived the asymptotic distribution of the discrimination statistic. To further investigate the performance of this method, we have conducted Monte Carlo (MC) simulations, and computed values of PCS by Monte Carlo (MC) simulations as well as asymptotic results. The method works in a satisfactorily way. Note that we could have chosen a different copula instead of the Frank copula, which captures the complete range of the correlation coefficient. Similar work can be extended using other copula choices. It would be intriguing to see how this problem of discrimination shapes up as we move to higher dimension. More work is needed in this direction.

Acknowledgments

The authors gratefully acknowledge the valuable comments and suggestions of the reviewer and the Associate Editor, which have significantly improved the clarity and quality of this manuscript.

Appendix

Before beginning the proof, we will derive some asymptotic results, regarding number of failure (d), that are required to draw the correspondence from results of Ferenstein (1992). Suppose n units, with independent life times and common CDF $F_X(x)$, are placed on a life-test. Let d

denote the number of units that failed on or before a pre-fixed time t_c, then it can be stated that d follows $Binomial(n,p)$ where $p = F_X(T)$. Therefore, we can conclude that

$$\frac{d}{n} \xrightarrow{P} p \implies \frac{d}{n} = p + o_p(1) \tag{22}$$

where $\xrightarrow{P}$ denotes convergence in probability and $o_p(1)$ means converges to zero in probability.

Proof of Lemma 3.1:

The basic idea of the proof follows results of Lemma 1 and Lemma 2 of Ferenstein (1992).

To use those results let us define

$$g(x,y;\Gamma,\Sigma^0) = \ln \frac{f_{BVWL}(x,y;\Gamma)}{f_{BVLN}(x,y;\Sigma^0)}$$

and

$$h(x;\alpha_2,\beta_2,\mu_2^0,\sigma_2^0)) = \ln \frac{(1-F_{WE}(x;\alpha_2,\beta_2))}{(1-F_{LN}(x;\mu_2^0,\sigma_2^0))}$$

Let $\chi = (\Gamma,\Sigma^0))$ and $\mathcal{B}$ be a compact neighbourhood of χ. Thus, using equation (22) and Lemma 2 of Ferenstein (1992), we have,

$$\frac{1}{n}\Big(L_{BVWL}(\Gamma) - L_{BVLN}(\Sigma^0)\Big) \longrightarrow \Gamma_1(\Gamma) - \Gamma_2(\Sigma^0) \tag{23}$$

with probability 1 and uniformly in $\chi \in \mathcal{B}$. Here,

$$\begin{aligned}\Gamma_1(\Gamma) - \Gamma_2(\Sigma^0) = \int_0^T \int_0^\infty \ln\left(\frac{f_{BVWL}(x,y;\Gamma)}{f_{BVLN}(x,y;\Sigma^0)}\right) f_{BVLN}(x,y;\Sigma^0)\,dx\,dy \\ +q\ln\left(\frac{(1-F_{WE}(T;\alpha_2,\beta_2))}{q}\right)\end{aligned}$$

Moreover, we have,

$$L(\hat{\Gamma}) = \sup_{(\Gamma)\in\mathbb{R}^5} \frac{1}{n}\Big(L_{BVWl}(\Gamma) - L_{BVLN}(\Sigma^0)\Big)$$

Using Lemma 2.2 of White (1980), we have, $\hat{\Gamma} \to \tilde{\Gamma}$ a.s.

Proof of Lemma 3.1(c):

We have,

$$\frac{D_n}{n} = \frac{1}{n}\Bigg[\sum_{i=1}^{d} g(x_i, y_{(i)}, \hat{\chi}) + (n-d)h(T, \hat{\chi})\Bigg]$$

and

$$\frac{D_*}{n} = \frac{1}{n}\Bigg[\sum_{i=1}^{d} g(x_i, y_{(i)}), \tilde{\chi}) + (n-d)h(T, \tilde{\chi})\Bigg]$$

here $\hat{\chi} = (\hat{\Gamma}, \hat{\Sigma})^T$and $\tilde{\chi} = (\tilde{\Gamma}, \Sigma^0)^T$ To prove Lemma 3.1(c), we need to show that :

$$\frac{1}{\sqrt{n}}[D_n - E_{LN}(D_n|\Sigma^0) - D_* + E_{BVLN}(D_*)] \xrightarrow{P} 0$$

In order to do so let us focus on the establishment of the statement:

$$\frac{1}{\sqrt{n}}[D_n - D_*] \xrightarrow{P} 0$$

followed by proving :

$$E_{BVLN}\Bigg(\frac{1}{\sqrt{n}}[D_n - D_*]\Bigg) \longrightarrow 0$$

$$\begin{aligned}
\frac{1}{\sqrt{n}}\left[D_n - D_*\right] &= \sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^{d}\left[g(x_i, y_{(i)}; \hat{\chi}) - g(x_i, y_{(i)}; \tilde{\chi})\right] + \left(\frac{n-d}{n}\right)\left[h(T;\hat{\chi}) - h(T;\tilde{\chi})\right]\right\} \\
&= \sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^{n}\left[g(x_i, y_i; \hat{\chi}) - g(x_i, y_i; \tilde{\chi})\right].\mathbf{1}_{y_i \le T} + \left(\frac{n-d}{n}\right)\left[h(T;\hat{\chi}) - h(T;\tilde{\chi})\right]\right\} \\
&= \sum_{j=1}^{10}\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j}) \times \left[\frac{1}{n}\sum_{i=1}^{n} g_j^{'}(x_i, y_i; \tilde{\chi_j}).\mathbf{1}_{y_i \le T} + \left(1 - \frac{d}{n}\right)h_j^{'}(T;\tilde{\chi_j})\right] + o_p(1) \\
&= \sum_{j=1}^{10}\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j}) \times \left[\frac{1}{n}\sum_{i=1}^{n} g_j^{'}(x_i, y_i; \tilde{\chi_j}).\mathbf{1}_{y_i \le T} + \Big(1 - p - o_p(1)\Big)h_j^{'}(T;\tilde{\chi_j})\right] + o_p(1)
\end{aligned}$$

Second last step appears using multivariate Taylor's expansion of $\left\{\frac{1}{n}\sum_{i=1}^{n}\left[g(x_i,y_i;\hat{\chi})\right].\mathbf{1}_{y_i\le T}+\left(\frac{n-d}{n}\right)\left[h(T;\hat{\chi})\right]\right\}$ around $\tilde{\chi}$. Here $o_p(1)$ means converges to zero in probability, where $g_j^{'}=\frac{\partial}{\partial\chi_j}g(\cdot)$, $h_j^{'}=\frac{\partial}{\partial\chi_j}h(\cdot)$, $\hat{\chi_j}$ is the j-th component of $\hat{\chi}$, and $\tilde{\chi_j}$ is the j-th component of $\tilde{\chi}$ for $j=1,\ldots,10$.Here total number of parameters involved are 10. Here $\sqrt{n}(\hat{\chi_j}-\tilde{\chi_j})$ converges to a normal distribution with zero mean and some finite variance(using Theorem 3.2 of White (1982)), therefore using Slutsky's theorem, we can write last step as:

$$\frac{1}{\sqrt{n}}\left[D_n-D_*\right]=\sum_{j=1}^{10}\left[\frac{1}{n}\sum_{i=1}^{n}g_j^{'}(x_i,y_i;\tilde{\chi_j}).\mathbf{1}_{y_i\le T}+\Big(1-p\Big)h_j^{'}(T;\tilde{\chi_j})\right]\sqrt{n}(\hat{\chi_j}-\tilde{\chi_j})+o_p(1)$$

Now, using the weak law of large numbers, equation (8) and the fact that $h^{'}(\cdot)$ is a continuous function, we can write

$$\left[\frac{1}{n}\sum_{i=1}^{n}g_j^{'}(x_i,y_i;\tilde{\chi_j}).\mathbf{1}_{y_i\le T}+(1-p)h_j^{'}(T;\tilde{\chi_j})\right]\xrightarrow{P}E_{BVLN}\Big(g_j^{'}(X,Y);\chi).\mathbf{1}_{Y\le T}+(1-p)h_j^{'}(T;\chi)\Big)\Big|_{\chi=\tilde{\chi}} \tag{24}$$

Also, using the similar arguments as used in Fearn and Nebenzahl (1991), it follows that:

$$\left|\frac{1}{n}\sum_{i=1}^{n}g_j^{'}(X_i,Y_i;\tilde{\chi_j}).\mathbf{1}_{Y_i\le T}+(1-p)h_j^{'}(T;\tilde{\chi_j})\right|\text{is uniformly bounded} \tag{25}$$

Therefore, using dominated convergence theorem (DCT) to interchange the operations of integration and differentiation, in right hand term of equation (24), we have,

$$\begin{aligned}&E_{BVLN}\Big(g_j^{'}(X,Y;\chi).\mathbf{1}_{Y\le T}+(1-p)h_j^{'}(T;\chi)\Big)\Big|_{\chi=\tilde{\chi}}\\ &=\frac{\partial}{\partial\chi_j}\Big[\int_0^T\int_0^\infty\ln\left(\frac{f_{BVWL}(x,y;\Gamma)}{f_{BVLN}(x,y;\Sigma^0)}\right)f_{BVLN}(x,y;\Sigma^0)\,dx\,dy+q\ln\left(\frac{(1-F_{WE}(T;\alpha_2,\beta_2))}{q}\right)\Big]\Big|_{\chi=\tilde{\chi}}\\ &=0\end{aligned} \tag{26}$$

Using equation (26) and the fact that $\sqrt{n}(\hat{\gamma_j}-\tilde{\gamma_j})$ converges to a normal distribution with zero mean and some finite variance(using Theorem 3.2 of White (1982)), we have

$$\frac{1}{\sqrt{n}}[D_n-D_*]\xrightarrow{P}0$$

Now let us consider the expectation term,

$$
\begin{aligned}
& E_{BVLN}\left(\frac{1}{\sqrt{n}}\left[D_n - D_*\right]\right) \\
= & E_{BVLN}\left|\sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^{d}\left[g(X_i, Y_{(i)}; \hat{\chi}) - g(X_i, Y_{(i}; \tilde{\chi})\right] + \left(\frac{n-d}{n}\right)\left[h(T; \hat{\chi}) - h(T; \tilde{\chi})\right]\right\}\right| \\
\leq & \sum_{j=1}^{10} E_{BVLN}\left|\left[\frac{1}{n}\sum_{i=1}^{n} g_j^{'}(X_i, Y_i; \tilde{\chi_j}).\mathbf{1}_{Y_i \leq T} + (1-p)h_j^{'}(T; \tilde{\chi_j})\right]\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j})\right| + o(1) \\
\leq & \sum_{j=1}^{10}\left[E_{BVLN}[\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j})]^2 \; E_{BVLN}\left[\frac{1}{n}\sum_{i=1}^{n} g_j^{'}(X_i, Y_i; \tilde{\chi_j}).\mathbf{1}_{Y_i \leq T} + (1-p)h_j^{'}(t_c; \tilde{\chi_j})\right]^2\right]^{\frac{1}{2}} + o(1)
\end{aligned}
$$

Last step is written using Cauchy-Schwarz inequality.

Since $E_{BVLN}[\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j})]^2 < \infty$ for all $j = 1 \dots 10$ (using theorem 3.2 of White (1982)) and using equation (25) and (26), we have,

$$E_{BVLN}\left[\frac{1}{n}\sum_{i=1}^{n} g_j^{'}(X_i, Y_i; \tilde{\chi_j}).\mathbf{1}_{Y_i \leq T} + (1-p)h_j^{'}(T; \tilde{\chi_j})\right]^2 \longrightarrow 0$$

Thus, $E_{BVLN}\left(\frac{1}{\sqrt{n}}[D_n - D_*]\right) \longrightarrow 0$

Hence result follows.

Statements and Declarations:

- Funding:No funding was received for conducting this study

- Conflict of interest:The authors have no competing interests to declare that are relevant to the content of this article.

References

Almetwally, E. M., Muhammed, H. Z., and El-Sherpieny, E.-S. A. (2020). Bivariate weibull distribution: properties and different methods of estimation. *Annals of Data Science*, 7:163–193.

Baron, P. A. (2001). Measurement of airborne fibers a review. *Industrial Health*, 39(2):39–50.

Bhattacharya, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics. *The Annals of Statistics*, pages 1034–1039.

Bhattacharyya, G. K. (1985). The asymptotics of maximum likelihood and related estimators based on type ii censored data. *Journal of the American Statistical Association*, 80(390):398–404.

Cox, D. R. (1961). Tests of separate families of hypotheses. In *Proceedings of the fourth Berkeley symposium on Mathematical Statistics and Probability*, volume 1, pages 105–123.

David, H. A. and Nagaraja, H. N. (2004). *Order statistics*. John Wiley & Sons.

Denuit, M., Purcaru, O., Van Keilegom, I., et al. (2004). Bivariate archimedean copula modelling for loss-alae data in non-life insurance. *IS Discussion Papers*, 423.

Dey, A. K. and Kundu, D. (2009). Discriminating among the log-normal, weibull, and generalized exponential distributions. *IEEE Transactions on Reliability*, 58(3):416–424.

Dey, A. K. and Kundu, D. (2012). Discriminating between the bivariate generalized exponential and bivariate weibull distributions. *Chilean Journal of Statistics*, 3(1):93–110.

Fearn, D. and Nebenzahl, E. (1991). On the maximum likelihood ratio method of deciding between the weibull and gamma distributions. *Communications in Statistics-Theory and Methods*, 20(2):579–593.

Ferenstein, E. Z. (1992). Asymptotic properties of the maximum likelihood estimators from type ii censored samples from bivariate distributions. *Demonstratio Mathematica*, 25(3):691–700.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *North American actuarial journal*, 2(1):1–25.

Gill, P., Tiku, M., and Vaughan, D. C. (1990). Inference problems in life testing under multivariate normality. *Journal of Applied Statistics*, 17(1):133–147.

Gomes, M. I. (1984). Concomitants in a multidimensional extreme model. In *Statistical Extremes and Applications*, pages 353–364. Springer.

Harrell, F. E. and Sen, P. K. (1979). Statistical inference for censored bivariate normal distributions based on induced order statistics. *Biometrika*, 66(2):293–298.

Khakifirooz, M., Tseng, S. T., and Fathi, M. (2020). Model misspecification of generalized gamma distribution for accelerated lifetime-censored data. *Technometrics*, 62(3):357–370.

Khan, S., Jamal, F., Tahir, M. H., and Elhassanein, A. (2025). A new generalized logistic class of distributions: Properties and applications on flood and earthquake data sets with bivariate extension. *SCOPUA Journal of Applied Statistical Research*, 1(2).

Klugman, S. A. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: mathematics and economics*, 24(1-2):139–148.

Kundu, D. and Manglick, A. (2004). Discriminating between the weibull and log-normal distributions. *Naval Research Logistics (NRL)*, 51(6):893–905.

Lee, E.-J., Kim, C.-H., and Lee, S.-H. (2011). Life expectancy estimate with bivariate weibull distribution using archimedean copula. *International Journal of Biometrics and Bioinformatics*, 5(3):149–161.

Maghsoudi, M. and Bakar, S. A. A. (2017). In search of best fitted composite model to the alae data set with transformed gamma and inversed transformed gamma families. In *AIP Conference Proceedings*, volume 1842. AIP Publishing.

McIntyre, G. (1952). A method for unbiased selective sampling, using ranked sets. *Australian journal of agricultural research*, 3(4):385–390.

Mei, M. (2016). A goodness-of-fit test for semi-parametric copula models of right-censored bivariate survival times.

Presnell, B. and Boos, D. D. (2004). The ios test for model misspecification. *Journal of the American Statistical Association*, 99(465):216–227.

Ramachandran, G. and Vincent, J. H. (1999). A bayesian approach to retrospective exposure assessment. *Applied occupational and environmental hygiene*, 14(8):547–557.

Ristić, M. M., Popović, B. V., Zografos, K., and Balakrishnan, N. (2018). Discrimination among bivariate beta-generated distributions. *Statistics*, 52(2):303–320.

Ross, K. S., Carter, H. B., Pearson, J. D., and Guess, H. A. (2000). Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. *Jama*, 284(11):1399–1405.

Schneider, T. and Holst, E. (1983). Man-made mineral fibre size distributions utilizing unbiased and fibre length biased counting methods and the bivariate log-normal distribution. *Journal of Aerosol Science*, 14(2):139–146.

White, H. (1980). Nonlinear regression on cross-section data. *Econometrica: Journal of the Econometric Society*, pages 721–746.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, pages 1–25.

Wiens, B. L. (1999). When log-normal and gamma models give different results: a case study. *The American Statistician*, 53(2):89–93.

Zhang, S., Okhrin, O., Zhou, Q. M., and Song, P. X.-K. (2016). Goodness-of-fit test for specification of semiparametric copula dependence models. *Journal of Econometrics*, 193(1):215–233.

Zhou, X.-H., Li, C., Gao, S., and Tierney, W. M. (2001). Methods for testing equality of means of health care costs in a paired design study. *Statistics in Medicine*, 20(11):1703–1720.