Discriminating between Bivariate Birnbaum Saunders and Bivariate Log-Normal distributions

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 five-parameter bivariate Birnbaum-Saunders and bivariate lognormal distributions are two models used to study positive lifetime bivariate data. These distributions exhibit striking similarities, from the surface plot of their joint density functions to the nature of their reliability functions and hazard gradients within certain parameter spaces. Additionally, their marginal distributions share similar shapes. This paper aims to discriminate between these two distributions. To achieve this, we use the difference between the maximized log-likelihood functions as a discrimination statistic. The asymptotic distribution of the test statistic is derived to compute the probability of correct selection (PCS). Moreover, we have also studied the effects of model misspecification on the mode of the marginal hazard function as well as on the components of the hazard gradient. Due to the limitations of the usual likelihood approach, we also propose a suitable modification for the decision rule. Finally, we do a real data analysis.

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

1 Introduction

Kundu et al. (2010) introduced a bivariate Birnbaum–Saunders (BVBS) distribution with five

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

parameters. The joint probability density function of the BVBS distribution is given by:

$$f_{BVBS}(x,y;\Gamma) = \frac{1}{4\alpha_1\alpha_2\beta_1\beta_2}\left[\left(\frac{\beta_1}{x}\right)^{\frac{1}{2}} + \left(\frac{\beta_1}{x}\right)^{\frac{3}{2}}\right]\left[\left(\frac{\beta_2}{y}\right)^{\frac{1}{2}} + \left(\frac{\beta_2}{y}\right)^{\frac{3}{2}}\right]$$
$$\times \exp\left[-\frac{1}{2(1-\rho_1^2)}(u^2+v^2-2\rho_1 uv)\right], \quad (1)$$

where $\Gamma = (\alpha_1, \alpha_2, \beta_1, \beta_2, \rho_1)$; $x, y > 0$; $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$; $-1 < \rho_1 < 1$. Here u and v are defined as:

$$u = \frac{1}{\alpha_1}\left(\sqrt{\frac{x}{\beta_1}} - \sqrt{\frac{\beta_1}{x}}\right), \qquad v = \frac{1}{\alpha_2}\left(\sqrt{\frac{y}{\beta_2}} - \sqrt{\frac{\beta_2}{y}}\right).$$

The density surface of $BVBS(\alpha_1, \alpha_2, \beta_1, \beta_2, \rho_1)$ is unimodal and can take on a variety of shapes (see Kundu et al. (2010)), making it useful for analyzing bivariate lifetime or reliability data. Another well-known five-parameter model is the bivariate lognormal (BVLN) distribution, whose joint probability density function is:

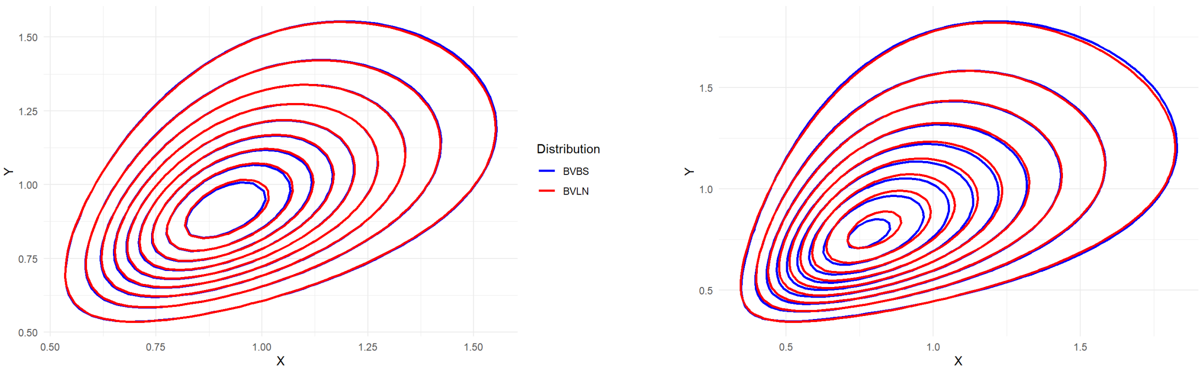
$$f_{BVLN}(x,y;\Sigma) = \frac{1}{2\pi\sigma_1\sigma_2 xy\sqrt{1-\rho_2^2}}\exp\left[-\frac{Q}{2(1-\rho_2^2)}\right], \quad (2)$$

where $\Sigma = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_2)$; $x, y > 0$; $-\infty < \mu_1, \mu_2 < \infty$; $\sigma_1^2, \sigma_2^2 > 0$; $-1 < \rho_2 < 1$, and

$$Q = \left(\frac{\log(x)-\mu_1}{\sigma_1}\right)^2 + \left(\frac{\log(y)-\mu_2}{\sigma_2}\right)^2 - 2\rho_2\left(\frac{\log(x)-\mu_1}{\sigma_1}\right)\left(\frac{\log(y)-\mu_2}{\sigma_2}\right).$$

Both the BVBS and BVLN distributions share the same Gaussian copula, and the latter has been widely applied across several fields. It models the size distribution of aerosol particles and airborne fibers (Schneider and Holst (1983), Ramachandran and Vincent (1999), Baron (2001)), paired PSA values in prostate cancer screening (Ross et al. (2000)), and healthcare cost data before and after interventions (Zhou et al. (2001)). It is also extensively used in insurance modeling.

To illustrate the similarities between these two distributions, we have plotted their contour shapes together. Figure 1 presents contour plots for representative parameter choices. In particular, Figure 1a compares $BVBS(0.25, 0.25, 1, 1, 0.5)$ with its closest BVLN counterpart, while Figure 1b provides a similar comparison for $BVBS(0.4, 0.4, 2, 2, 0.5)$. Each figure presents both distributions on the same plot to highlight their contour similarity under different parameter settings.



(a) $BVBS(0.25, 0.25, 1, 1, 0.5)$ (blue) vs $BVLN(-3.40 \times 10^{-5}, -6.69 \times 10^{-5}, 0.24, 0.24, 0.5)$ (red)

(b) $BVBS(0.4, 0.4, 2, 2, 0.5)$ (blue) vs $BVLN(1.05 \times 10^{-5}, 3.14 \times 10^{-5}, 0.39, 0.39, 0.5)$ (red)

Figure 1: Contour plots demonstrating the similarity between BVBS and BVLN distributions.

The parameters for the BVBS distribution were chosen arbitrarily, whereas the corresponding BVLN parameters were determined by minimizing the two-dimensional Hellinger distance (see Hellinger (1909)) between their probability density functions. Table 1 reports the closest BVLN distributions obtained for selected BVBS parameter settings, along with the associated Hellinger distances. It is evident that when both shape parameters are small (e.g., 0.25), the two distributions are nearly indistinguishable, as indicated by the very small Hellinger's distance. However, as the shape parameters increase, the Hellinger's distance also increases, suggesting greater dissimilarity between the distributions. Notably, even when only one of the shape parameters is set to 1 (with the other held at 0.25), the Hellinger's distance becomes larger compared to the case where both shape parameters are small. These observations motivate the need for appropriate model selection criterion, discussed in the subsequent sections.

BVBS Parameters	**Closest BVLN Parameters**	**Hellinger Distance**
(0.25, 0.25, 2, 2, 0.5)	(0.69, 0.69, 0.24, 0.24, 0.50)	3.85×10^{-5}
(0.50, 0.50, 2, 2, 0.5)	(0.69, 0.69, 0.49, 0.49, 0.50)	5.25×10^{-4}
(1.00, 1.00, 2, 2, 0.5)	(0.69, 0.69, 0.90, 0.90, 0.50)	4.72×10^{-3}
(0.25, 1.00, 2, 2, 0.5)	(0.69, 0.69, 0.24, 0.90, 0.50)	2.42×10^{-3}

Table 1: Closest Bivariate Lognormal (BVLN) distribution to a given BVBS based on Hellinger distance.

The problem of determining whether a given data set originates from one of multiple probability distributions has a long history. It was first addressed by Cox (1961), who explored the consequences of incorrect model selection. Subsequent developments were made by Atkinson (1969, 1970), Dyer (1973), and Chen (1980). Since then, the topic has evolved considerably, with contributions such as Dey and Kundu (2009), Ahmad et al. (2017), and Raqab et al. (2018). Recent work includes Abid and Kokonendji (2023), who investigated discrimination between Poisson–Tweedie models; Diyali et al. (2023), who studied lognormal vs. log-logistic discrimination under Type-II censoring; and Basu and Kundu (2023), who analyzed the Birnbaum–Saunders vs. lognormal case under model misspecification. Other studies such as Diyali et al. (2024) extended this to inverse Weibull and inverse Gaussian models.

Discrimination between two distributions is crucial, since the assumed model directly affects estimates of tail probabilities, quantiles, and reliability functions—all of which are highly sensitive to model choice. Several studies have explored model misspecification effects, e.g., Wiens (1999) and Khakifirooz et al. (2020). Basu et al. (2009) demonstrated how model misspecification impacts strength estimation in brittle materials, and Pasari (2018) examined its role in forecasting rare events such as large Himalayan earthquakes. However, only limited work has been done on discrimination between bivariate distributions—most notably by Dey and Kundu (2012) and Ristić et al. (2018). A new addition in the bivariate setup has been done by Rajput et al. (2026).

The key contributions of this paper are threefold: (i) development of asymptotic results for the proposed discrimination criterion; (ii) introduction of a modified approach to address cases where the conventional method fails; and (iii) investigation of the effects of model misspecification on the mode of the marginal hazard function and on the individual components of the bivariate hazard gradient.

The rest of the paper is organized as follows. Section 2 discusses the motivation for the study and highlights why discrimination between the BVBS and BVLN models can be particularly challenging. Section 3 introduces the proposed discrimination criterion and develops its asymptotic properties. Section 4 presents the computation of misspecified parameters along with supporting numerical results. Section 5 describes the modified decision rule and demonstrates its effectiveness through simulation. A real data application is provided in Section 6, and concluding remarks are given in Section 7.

2 Motivation

Despite the apparent similarity between the univariate and bivariate forms of the Birnbaum–Saunders and lognormal distributions, there can be some serious effects of model misspecification. To demonstrate this, we analyze a dataset from Vasey and Thayer (1987) involving 22 participants, where electromyographic (EMG) activity from the left brow region was recorded under two conditions: Stage 1 (X), exposure to relaxing music, and Stage 2 (Y), exposure to music inducing positive affect. Each trial lasted 90 seconds, and the response variable in both stages was the average EMG amplitude (μV).

Subject	1	2	3	4	5	6	7	8	9	10	11
Stage 1	143	142	109	123	276	235	208	267	183	245	324
Stage 2	368	155	167	135	216	386	175	358	193	268	507

Subject	12	13	14	15	16	17	18	19	20	21	22
Stage 1	148	130	119	102	279	244	196	279	167	345	524
Stage 2	378	142	171	94	204	365	168	358	183	238	507

Table 2: Left brow EMG amplitudes from 22 subjects.

Both BVBS and BVLN models were fitted to the data via maximum likelihood estimation (MLE), with initial values obtained from marginal fits and the sample correlation. The resulting estimates are:

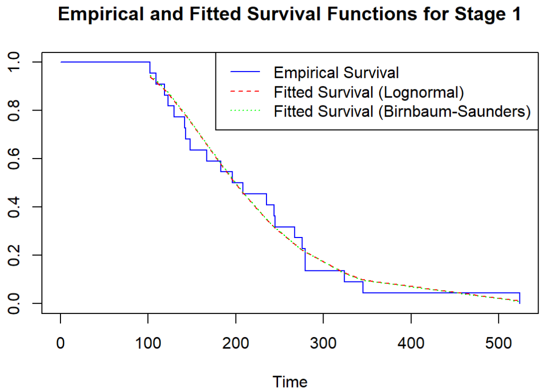
$$\text{BVBS: } \hat{\alpha}_1 = 0.4255,\ \hat{\alpha}_2 = 0.4623,\ \hat{\beta}_1 = 199.60,\ \hat{\beta}_2 = 235.57,\ \hat{\rho}_1 = 0.6925,$$
$$\text{BVLN: } \hat{\mu}_1 = 5.2930,\ \hat{\mu}_2 = 5.4609,\ \hat{\sigma}_1 = 0.4281,\ \hat{\sigma}_2 = 0.4650,\ \hat{\rho}_2 = 0.6896.$$

Marginal adequacy was assessed using Kolmogorov–Smirnov (KS) statistics:

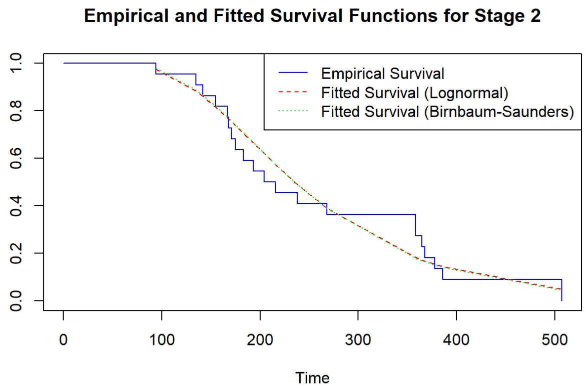
$$\text{BVBS: } D_X = 0.1233\ (p = 0.8912),\ D_Y = 0.1827\ (p = 0.4547),$$
$$\text{BVLN: } D_X = 0.1189\ (p = 0.9149),\ D_Y = 0.1801\ (p = 0.4729),$$

indicating both models fit the marginals well (see Figure 2). However, marginal fit alone does not ensure joint adequacy, which is the main motivation for this study.

To assess joint fit, we applied transformations that should yield bivariate normality if the



(a) Stage 1 (relaxing music).



(b) Stage 2 (positive affect).

Figure 2: Marginal survival functions under fitted BVBS and BVLN models.

assumed model is correct:

$$\text{BVBS: } g_1(x) = \frac{1}{\alpha_1}\left(\sqrt{\frac{x}{\beta_1}} - \sqrt{\frac{\beta_1}{x}}\right), \quad g_2(y) = \frac{1}{\alpha_2}\left(\sqrt{\frac{y}{\beta_2}} - \sqrt{\frac{\beta_2}{y}}\right),$$

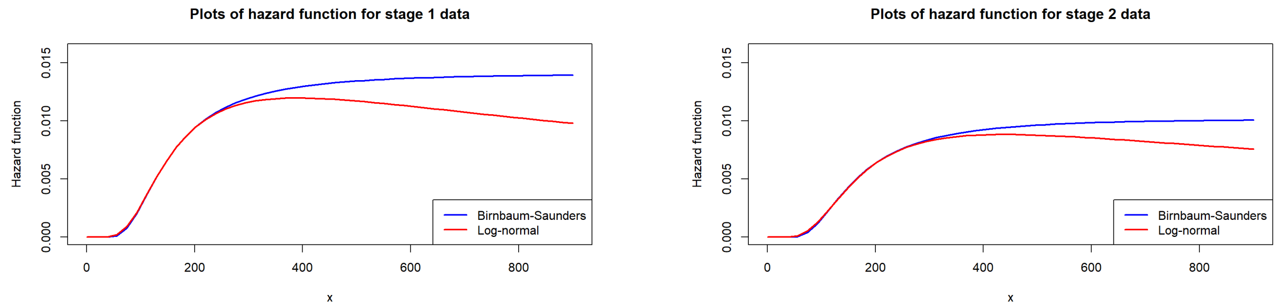
$$\text{BVLN: } g_1(x) = \log(x), \quad g_2(y) = \log(y).$$

Both transformed datasets passed Mardia's skewness and kurtosis tests, and Royston's multivariate Shapiro–Wilk test (p values are obtained in Table 3), again suggesting that both models describe the data well at the joint level.

Model	Skewness	Kurtosis	Shapiro–Wilk
BVBS	0.769	0.228	0.645
BVLN	0.778	0.211	0.643

Table 3: Bivariate normality test p-values for transformed data.

Despite these similarities, the models yield substantially different estimates for reliability-related quantities. Figure 3 compares the fitted hazard functions for each marginal under the two model assumptions. Although both models fit well marginally, substantial differences appear in the estimated modes of the hazard functions (see Table 4) Identifying the mode of the hazard function is often important in survival studies, where it indicates the most critical period for intervention (Gupta et al. (1997)).



(a) Stage 1 data. (b) Stage 2 data.

Figure 3: Comparison of fitted hazard functions under BVBS and BVLN models.

Model	Stage 1	Stage 2
BVBS	1474.54 (722.65, 2488.76)	1369.26 (662.89, 2887.48)
BVLN	389.98 (299.96, 496.11)	437.20 (325.78, 583.52)

Table 4: Estimated modes of marginal hazard functions with 95% confidence intervals.

Simulation experiments further confirm this discrepancy. When data are generated from $BVBS(\alpha, \alpha, 1, 1, 0.5)$ but incorrectly modeled as BVLN, the estimated mode of the marginal hazard differs substantially (see Supplementary Tables S1–S2). To understand the multivariate impact of misspecification, we also examine the bivariate hazard gradient (Johnson and Kotz (1975)) defined as:

$$h(x_1, x_2) = \big(h_1(x_1, x_2), h_2(x_1, x_2)\big), \quad h_i(x_1, x_2) = -\frac{\partial}{\partial x_i} \ln R(x_1, x_2), \tag{3}$$

where $R(x_1, x_2)$ is the bivariate survival function. Comparing the components of $h(x_1, x_2)$ under correct and misspecified models reveals significant differences—for instance, $h_{1_{\mathrm{BVBS}}}(2500, 200) = 0.0140$ versus $h_{1_{\mathrm{BVLN}}}(2500, 200) = 0.0056$, and $h_{2_{\mathrm{BVBS}}}(2.5, 0.2) = 0.006$ versus $h_{2_{\mathrm{BVLN}}}(2.5, 0.2) = 0.003$. These analyses underscore that even when BVBS and BVLN appear similar in shape or marginal behavior, their inferential and reliability implications can differ sharply (see Supplementary Tables S3–S6). Hence, it is crucial to make appropriate model selection.

3 Discrimination criteria and some asymptotic properties

In this section, we first describe the discrimination criterion used to solve the discrimination problem. We then develop its asymptotic distribution, which will help us calculate the probability of correct selection.

3.1 Discrimination criteria

Let $(x, y) = \{(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)\}$ be the random sample. We will develop discrimination procedure on basis of this random sample in this section. We will use the following notations throughout this paper. We will be denoting the set of parameters associated with BVBS as $\Gamma = (\alpha_1, \alpha_2, \beta_1, \beta_2, \rho_1)$ and with that of BVLN as $\Sigma = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_2)$ It is assumed that the data has been obtained from one of the two bivariate distributions: $BVBS(\Gamma)$ or $BVLN(\Sigma)$. Based on the observed sample, let the corresponding log-likelihood function for BVBS and BVLN be denoted by $L_{BVBS}(\Gamma|x, y)$ and $L_{BVLN}(\Sigma|x, y)$, respectively. If $\hat{\Gamma} = (\hat{\alpha_1}, \hat{\alpha_2}, \hat{\beta}_1, \hat{\beta}_2, \hat{\rho}_1)$ and $\hat{\Sigma} = (\hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{\rho}_2)$ denote the corresponding ML-estimators, then the discrimination procedure is as follows: Choose:
a) BVBS if $L_{BVBS}(\hat{\Gamma}|x, y) > L_{BVLN}(\hat{\Sigma}|x, y)$
b)BVLN if $L_{BVLN}(\hat{\Sigma}|x, y) > L_{BVBS}(\hat{\Gamma}|x, y)$

Let $T = L_{BVBS}(\hat{\Gamma}|x, y) - L_{BVLN}(\hat{\Sigma}|x, y)$. T is the difference of maximized log-likelihood and it will be our statistic used for discrimination criterion. This approach was used by Gupta et al. (2002) to discriminate between two overlapping families of distributions. The concept was initially introduced by Cox (1961, 1962) for discriminating between distinct models. Subsequently, Dumonceaux et al. (1973) applied this method to discriminate among various distributions. Bain and Engelhardt (1980) later employed this technique to discriminate between Weibull and gamma distributions. If the data comes from $BVBS(\Gamma^0)$ then probability of correct selection is given as: $PCS_{BVBS} = P(T > 0|\Gamma^0)$ Similarly, if the data comes from $BVLN(\Sigma^0)$ then PCS is given as: $PCS_{BVLN} = P(T < 0|\Sigma^0)$

3.2 Asymptotic Distribution

In this section, we use the following notations. Let $U = (X, Y)$ follow $BVBS(\Gamma^0)$ where $\Gamma^0 = (\alpha_1^0, \alpha_2^0, \beta_1^0, \beta_2^0, \rho_1^0)$. Then for any functions, $f_1(U)$ and $f_2(U)$, $E_{BVBS}(f_1(U)|\Gamma^0)$, $V_{BVBS}(f_1(U)|\Gamma^0)$ and $Cov_{BVBS}(f_2(U), f_1(U)|\Gamma^0)$ 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 $\Sigma^0 = (\mu_1^0, \mu_2^0, \sigma_1^0, \sigma_2^0, \rho_2^0)$ 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 $\phi_1(f_{BVBS}(X, Y; \overline{\Gamma}))$ denotes a function of joint PDF of BVBS for arbitrary parameter set $\overline{\Gamma} = (\overline{\alpha_1}, \overline{\alpha_2}, \overline{\beta_1}, \overline{\beta_2}, \overline{\rho_1})$. Similarly, any function $\phi_2(f_{BVLN}(X, Y; \overline{\Sigma}))$ denotes a function of joint PDF of BVLN for arbitrary parameters $\overline{\Sigma} = (\overline{\mu_1}, \overline{\mu_2}, \overline{\sigma_1}, \overline{\sigma_2}, \overline{\rho_2})$. We have the following two main results:

Theorem 3.1 *If the data comes from $BVBS(\Gamma^0)$ then T is asymptotically normally distributed with mean $E_{BVBS}(T|\Gamma^0)$ and variance $V_{BVBS}(T|\Gamma^0)$.*

To prove the above theorem we need the following lemma :

Lemma 3.1 *If the data comes from $BVBS(\Gamma^0)$ we have the following results as $n \to \infty$:*

a)$\hat{\Gamma} \to \Gamma^0$ a.s. where

$$E_{BVBS}(ln(f_{BVBS}(X, Y; \Gamma^0))) = \max_{\overline{\Gamma}} E_{BVBS}(ln(f_{BVBS}(X, Y; \overline{\Gamma})))$$

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

$$E_{BVBS}(ln(f_{BVLN}(X, Y; \tilde{\Sigma}))) = \max_{\Sigma} E_{BVBS}(ln(f_{BVLN}(X, Y; \Sigma)))$$

c)Let us define

$$T^\star = L_{BVBS}(\Gamma^0) - L_{BVLN}(\tilde{\Sigma})$$

then $\frac{1}{\sqrt{n}}(T - E_{BVBS}(T|\Gamma^0))$ is asymptotically equivalent to $\frac{1}{\sqrt{n}}(T^\star - E_{BVBS}(T^\star))$.

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 $\frac{1}{\sqrt{n}}(T^\star - E_{BVBS}(T^\star))$ is asymptotically normally distributed. Therefore, using part (c) of Lemma 3.1, the result follows immediately.

Similarly, we can establish the asymptotic distribution of T when data is coming from $BVLN(\Sigma^0)$. In this case as well, T is asymptotically normally distributed with mean $E_{BVLN}(T|\Sigma^0)$ and variance $V_{BVLN}(T|\Sigma^0)$. The detailed theorem and its proof are available with the authors and can be provided upon request.

4 Numerical Results

In this section, we perform Monte Carlo (MC) simulations to estimate the Probability of Correct Selection (PCS) and to evaluate the performance of the asymptotic results under varying sample sizes and parameter settings for both distributions. Each experiment is replicated 10,000 times, and PCS is computed as the proportion of correct model selections across all replications.

For the derivation of asymptotic properties of the test statistic T, we require the misspecified parameter estimates under both competing models. These parameters are obtained by maximizing the expected log-likelihood of one model when the data are generated from the other.

To compute $\tilde{\Sigma}$, we maximize

$$E_{BVBS}\big[\ln f_{BVLN}(X,Y;\Sigma)\big]$$

with respect to Σ for fixed Γ^0. Let

$$\psi(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho_2) = -\ln(\sigma_1) - \ln(\sigma_2) - E_{BVBS}[\ln(XY)] - \tfrac{1}{2}\ln\sqrt{1-\rho_2^2} - \tfrac{1}{2(1-\rho_2^2)}E_{BVBS}(Q). \tag{4}$$

Since ψ has no closed form, numerical maximization was carried out using the Nelder–Mead algorithm via the `optim()` function in R, with MLEs from the true BVBS model used as initial values. Convergence was based on relative function changes with `reltol = 1e-8`.

Similarly, to obtain the misspecified parameters $\tilde{\Gamma}$, we maximize

$$E_{BVLN}\big[\ln f_{BVBS}(X,Y;\Gamma)\big]$$

with respect to Γ for fixed Σ^0, where

$$\begin{aligned}\phi(\alpha_1,\alpha_2,\beta_1,\beta_2,\rho_1) = &-\ln(\alpha_1\alpha_2\beta_1\beta_2) - \tfrac{1}{2}\ln(1-\rho_1^2) + E_{BVLN}\Big[\ln\Big(\sqrt{\tfrac{\beta_1}{X}} + \big(\tfrac{X}{\beta_1}\big)^{3/2}\Big)\Big] \\ &+ E_{BVLN}\Big[\ln\Big(\sqrt{\tfrac{\beta_2}{Y}} + \big(\tfrac{Y}{\beta_2}\big)^{3/2}\Big)\Big] - \tfrac{1}{2(1-\rho_1^2)}E_{BVBS}(u^2+v^2-2\rho_1 uv).\end{aligned} \tag{5}$$

As closed-form MLEs are unavailable for the BVBS model, initial values were obtained via a grid search (step size = 0.001), and optimization was performed using the same numerical procedure as above.

4.1 Case 1 : Parent distribution is BVBS

Now we provide values of PCS for given case using first Monte Carlo(MC) simulations and secondly, using asymptotic results discussed in previous sections. In this particular case, we consider four sets of parameters: Set 1: $(0.25, 0.25, 2, 2, 0.5)$, Set 2: $(0.5, 0.5, 2, 2, 0.5)$, Set 3: $(1, 1, 2, 2, 0.5)$, and Set 4: $(2, 2, 2, 2, 0.5)$. For each parameter set and various sample sizes, we generate a sample from BVBS distribution. The reason for fixing equal parameter values with $\rho = 0.5$ is that the correlation coefficient has a negligible effect on the value of PCS, except when it approaches extreme values (towards ± 0.99). We also examined the effect of the correlation coefficient by fixing the other parameters and varying ρ. This observation is further illustrated in see Supplementary Table S8 and S10.

Next, we estimate the maximum likelihood estimates (MLEs) of the unknown parameters and compute the corresponding log-likelihood values, assuming that the data follows either the BVBS or BVLN distribution. True parameter values were used as initial guesses for computing MLE's. Based on the maximized log-likelihood values, we make a decision regarding whether the appropriate distribution has been chosen or not. To assess the performance of our approach, we repeat this process 10,000 times and calculate the proportion of correct selections. This provides an indication of how accurately the correct distribution is identified under different parameter sets and sample sizes. Further, in addition to the PCS values obtained via Monte Carlo simulation and those based on asymptotic results, we also computed the PCS using the parametric bootstrap approach. The corresponding findings are summarized in Table 5. (for different sample size, $n = 20, 50, 80, 100, 250, 500$). As we see, when $\alpha > 1$, the PCS is high and the method for discrimination criteria is working upto expectation. Furthermore, the PCS values obtained through Monte Carlo simulation and those obtained via the parametric bootstrap are largely consistent with each other. Additionally, for sample sizes $n \geq 80$, the PCS values based on asymptotic approximations also exhibit close correspondence with those obtained from Monte Carlo simulation and parametric bootstrap.

However, for the case when $\alpha < 1$ the results so obtained are quite counter intuitive. The behavior of PCS as the sample size increases is quite abrupt. The discrimination method under consideration does not perform well for the range of shape parameters where $\alpha < 1$. This phenomenon can be partially explained by the findings reported in Table 1, which show that the similarity between the two distributions is very high when the shape parameters are less than 1. As a result, the method fails and yields counter-intuitive results with increasing sample size. A similar behavior for this range of shape parameters was also reported in the work of Basu and Kundu (2023), where the authors studied discrimination between the Birnbaum–Saunders and lognormal distributions. This highlights the interesting nature of the problem, as conventional methods struggle to discriminate effectively within this range of shape parameters.

α	Method	$n = 20$	$n = 50$	$n = 80$	$n = 100$	$n = 250$	$n = 500$
	MC	0.757	0.677	0.664	0.642	0.626	0.622
0.25	Parametric Bootstrap	0.743	0.662	0.655	0.628	0.615	0.610
	Asymptotic	(0.510)	(0.516)	(0.520)	(0.523)	(0.536)	(0.552)
	MC	0.782	0.735	0.723	0.702	0.771	0.823
0.50	Parametric Bootstrap	0.767	0.722	0.710	0.684	0.759	0.809
	Asymptotic	(0.559)	(0.594)	(0.618)	(0.631)	(0.702)	(0.774)
	MC	0.849	0.844	0.871	0.889	0.957	0.988
1	Parametric Bootstrap	0.833	0.829	0.857	0.873	0.947	0.981
	Asymptotic	(0.679)	(0.770)	(0.825)	(0.852)	(0.950)	(0.990)
	MC	0.933	0.968	0.993	0.999	1.000	1.000
2	Parametric Bootstrap	0.920	0.954	0.987	0.998	1.000	1.000
	Asymptotic	(0.859)	(0.953)	(0.983)	(0.991)	(1.000)	(1.000)

Table 5: Probability of Correct Selection (PCS) based on Monte Carlo simulations, parametric bootstrap, and asymptotic approximation for data from $BVBS(\alpha, \alpha, 2, 2, 0.5)$.

We further looked at PCS values when both the shape parameters are unequal. (results are presented in Supplementary Table S7)

We can see that when both the shape parameters are < 1, then the method is not working in a desired manner, whereas, even if one parameters ≥ 1, then the nature of PCS with sample size is acting in an expected manner. We also observed the effect of correlation coefficient for a fixed value of rest of the parameters.We computed PCS for varying values of ρ_1 when X follows $BVBS(0.5, 1, 2, 2, \rho)$ and the results are presented in see Supplementary Table S8.

We can see that as the value of ρ_1 is moved to extremes then the PCS for initial sample sizes is fairly lower than that of when compared to middle range of ρ_1.

The PCS obtained using the asymptotic distribution for a negative correlation value is approximately the same as that for the corresponding positive correlation. This observation can be explained by the similarity in the values of the misspecified parameters when the data is generated from $BVBS(0.5, 1, 2, 2, 0.75)$ and $BVBS(0.5, 1, 2, 2, -0.75)$. Specifically, the misspecified parameter values are $(0.69, 0.69, 0.48, 0.91, 0.74)$ for the positive correlation and $(0.68, 0.69, 0.49, 0.90, -0.74)$ for the negative correlation. Consequently, the asymptotic mean and variance for both cases are nearly identical, with an asymptotic mean of 0.0102 and an asymptotic variance of 0.0151. This leads to similar PCS values. A consistent pattern was observed for other pairs of positive and corresponding negative correlations as well.

4.2 Case 2 : Parent distribution is BVLN

Similarly, when sample is drawn from BVLN($\mu_1 = \mu_2 = 0, \sigma_1^2 = \sigma_2^2 = \sigma^2, \rho = 0.5$) the PCS so obtained(for different sample size,$n = 20, 50, 80, 100, 250, 500$) is given in Table 6.

In this case there is no counter intuitive results so obtained, yet the PCS is very low for entire parameter space at smaller sample size. The asymptotic correspondence is met for significantly larger sample size and as σ^2 increases the sample size needed to meet asymptotic results decreases(something similar was observed in case of BVBS).

σ^2	Method	$n = 20$	$n = 50$	$n = 80$	$n = 100$	$n = 250$	$n = 500$
	MC	0.258	0.415	0.476	0.512	0.635	0.734
0.25	Parametric Bootstrap	0.246	0.398	0.462	0.498	0.623	0.721
	Asymptotic	(0.561)	(0.596)	(0.621)	(0.634)	(0.707)	(0.779)
	MC	0.316	0.490	0.576	0.628	0.792	0.921
0.50	Parametric Bootstrap	0.302	0.472	0.559	0.610	0.779	0.908
	Asymptotic	(0.606)	(0.665)	(0.705)	(0.726)	(0.829)	(0.911)
	MC	0.382	0.621	0.742	0.789	0.944	0.994
1	Parametric Bootstrap	0.368	0.605	0.728	0.774	0.934	0.988
	Asymptotic	(0.664)	(0.749)	(0.802)	(0.828)	(0.933)	(0.983)
	MC	0.492	0.797	0.894	0.941	0.998	1.000
2	Parametric Bootstrap	0.475	0.785	0.883	0.930	0.996	0.999
	Asymptotic	(0.711)	(0.811)	(0.868)	(0.894)	(0.975)	(0.997)

Table 6: Probability of Correct Selection (PCS) based on Monte Carlo simulations, parametric bootstrap, and asymptotic approximation when data is coming from $BVLN(0, 0, \sigma, \sigma, 0.5)$.

We also obtained results when $\sigma_1 \neq \sigma_2$ in Supplementary Table S9. Something similar could be concluded in this case as well. When even one scale parameter is ≥ 1,the PCS increases. Yet, for initial sample sizes the PCS ≤ 0.50 and this is a serious issue to be resolved.

Effect of ρ for the case X follows $BVLN(0, 0, 0.5, 0.5, \rho)$ are presented in Supplementary Table S10. Unlike the case of BVBS, we see that as value of ρ is moved to extremes, the PCS is fairly high when compared to that of mid range values of ρ.

5 Modified Method

As discussed earlier, the usual method based on the ratio of maximized log-likelihoods does not perform well in certain parameter ranges, particularly when shape parameters are less than one. To address these shortcomings, we propose a modified decision criterion that ensures better decision where the usual method fails.

5.1 Motivation for the Modified Method

The usual approach generally performs well when the shape parameters are greater than one. However, in the case of BVBS, when $\alpha < 1$, the PCS behaves counterintuitively. On the other hand, if at least one shape parameter is ≥ 1, the method functions as expected. Similarly, in the case of BVLN, the PCS falls below 0.5 for small sample sizes when $\sigma^2 < 2$.

These issues indicate that the standard method does not effectively discriminate between the two distributions in certain parameter ranges. To address this, we require an alternative criterion that:

1. Ensures PCS increases with sample size.
2. Guarantees PCS is at least 0.5, as otherwise, a random guess (e.g., a fair coin toss) would be equally effective.

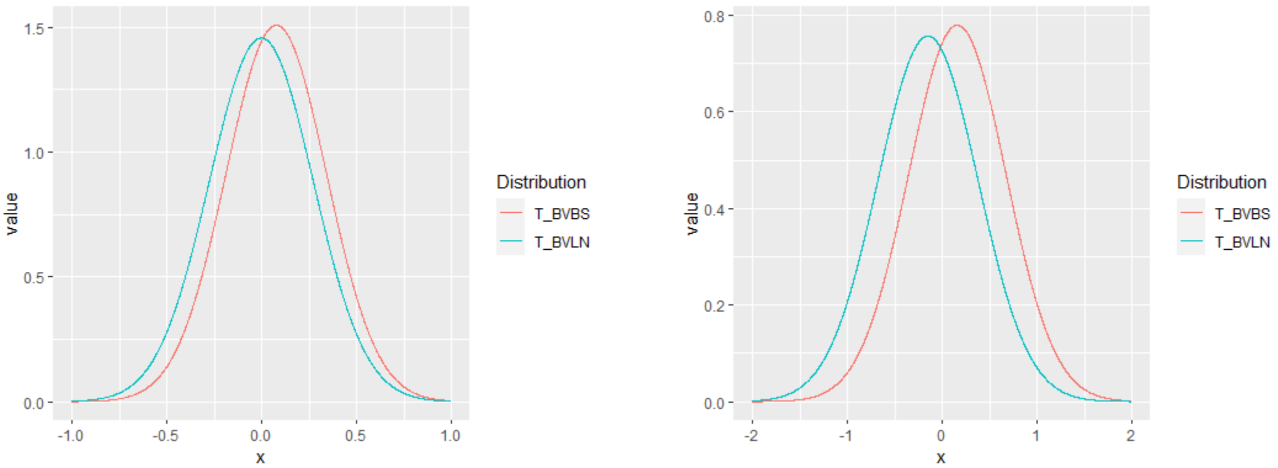
To illustrate the need for modification, we consider the case where data is generated from $BVBS(0.25, 0.25, 2, 2, 0.5)$. The histogram of the relevant statistic T is plotted and compared with the closest matching BVLN distribution (see Figure 4). This suggests that the problem reduces to distinguishing between two overlapping normal distributions with the same variance, say $N(\mu_1^*, \sigma^{*2})$ vs. $N(\mu_2^*, \sigma^{*2})$.

Given an observed sample T^*, the optimal decision criterion should be of the form $T > C$, where C minimizes the total probability of misclassification while ensuring PCS remains at least 0.5. The total misclassification probability is given by:

$$1 - \Phi\left(\frac{C - \mu_2^*}{\sigma^*}\right) + \Phi\left(\frac{C - \mu_1^*}{\sigma^*}\right).$$

Minimizing this expression leads to the choice of cutoff:

$$C = \frac{\mu_1^* + \mu_2^*}{2}.$$



(a) $n = 500$ (b) $n = 1000$

Figure 4: Plot of T when data is coming from $BVBS(0.5, 0.5, 2, 2, 0.5)$ vs when data is coming from closest BVLN.

5.2 Algorithm for Modified Method

Suppose we have a random sample of size n, $\{(x_1, y_1), (x_2, y_2), (x_3, y_3) \ldots (x_n, y_n)\}$ either from $BVBS(\alpha_1, \alpha_2, \beta_1, \beta_2, \rho)$ or $BVLN(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ and we have to choose the parent distribution.We will be using the following notations before we move on to the algorithm for modified method: $\lim_{n\to\infty} \frac{E_{BVBS}(T|\Gamma^0)}{n} = \mu_1^*$ and $\lim_{n\to\infty} \frac{V_{BVBS}(T|\Gamma^0)}{n} = \sigma^{*2}$.Similarly,$\lim_{n\to\infty} \frac{E_{BVLN}(T|\Sigma^0)}{n} = \mu_2^*$ and $\lim_{n\to\infty} \frac{V_{BVLN}(T|\Sigma^0)}{n} = \sigma^{*2}$

Thus, we provide the algorithm for modified method below:

- Step 1: Fit both the distributions to the given sample and find respective MLEs say $\hat{\alpha_1}, \hat{\alpha_2}, \hat{\beta_1}, \hat{\beta_2}, \hat{\rho_1}$ (denote it as $\widehat{\Gamma}$) and $\hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, , \hat{\rho_2}$ (denote it as $\widehat{\Sigma}$).
- Step 2: Find value of statistic T (say T^*).
- Step 3: Further simulate a sample of size n from fitted distribution $BVBS(\widehat{\Gamma})$ and calculate the statistic T (say T_{11}).
- Step 4: To obtain the estimate of μ_1 , repeat step 3 for some k times(1000 iterations) to obtain $(T_{11}, T_{12}, T_{13} ... T_{1k})$.
- Step 5: Based on $T_{11}, T_{12}, T_{13} ... T_{1k}$, obtain estimate of μ_1^* (denote it as $\hat{\mu_1^*}$).
- Step 6: Repeat Step 3 to Step 5, replacing $BVBS(\widehat{\Gamma})$ to $BVLN(\widehat{\Sigma})$ and obtain $\hat{\mu_2^*}$.

- Step 7:Choose Cut-off point as : $C = \frac{\hat{\mu_1^*}+\hat{\mu_2^*}}{2}$.
- Step 8: Choose BVBS IF $T^* > C$, else, we choose BVLN.

5.3 Simulation Results:

Based on the algorithm provided in previous section, we present the PCS obtained using MC simulation under choice of both the distribution as parent distribution. First, we provide PCS so obtained for varying sample sizes and different values of α when data is coming from $BVBS(\alpha, \alpha, 2, 2, 0.5)$ in Table 7.

α	Method	$n = 25$	$n = 50$	$n = 75$	$n = 100$	$n = 200$	$n = 500$
0.25	MC Simulation	0.524	0.548	0.553	0.557	0.562	0.571
	Parametric Bootstrap	0.512	0.538	0.547	0.552	0.560	0.569
0.50	MC Simulation	0.598	0.633	0.657	0.672	0.708	0.788
	Parametric Bootstrap	0.585	0.620	0.643	0.660	0.700	0.778
0.75	MC Simulation	0.633	0.726	0.762	0.784	0.857	0.956
	Parametric Bootstrap	0.620	0.710	0.748	0.770	0.845	0.948

Table 7: PCS based on Monte Carlo (MC) simulations and parametric bootstrap using the modified method when data is coming from $BVBS(\alpha, \alpha, 2, 2, 0.5)$.

Next, we provide PCS so obtained for varying sample sizes and different values of σ^2 when data is coming from $BVLN(0, 0, \sigma, \sigma, 0.5)$ in Table 8.

σ^2	Method	$n = 25$	$n = 50$	$n = 75$	$n = 100$	$n = 200$	$n = 500$
0.25	MC Simulation	0.503	0.519	0.521	0.576	0.612	0.733
	Parametric Bootstrap	0.503	0.510	0.518	0.565	0.603	0.725
0.50	MC Simulation	0.526	0.613	0.619	0.669	0.722	0.855
	Parametric Bootstrap	0.515	0.600	0.610	0.658	0.712	0.847
0.75	MC Simulation	0.567	0.638	0.669	0.712	0.842	0.946
	Parametric Bootstrap	0.555	0.625	0.658	0.700	0.832	0.940

Table 8: PCS based on Monte Carlo (MC) simulations and parametric bootstrap using the modified method when data is coming from $BVLN(0, 0, \sigma, \sigma, 0.5)$.

As discussed in previous subsections, the standard method based on the maximized log-likelihood fails when the data is generated from $BVBS(\alpha, \alpha, 2, 2, 0.5)$ with $\alpha < 1$ or from $BVLN(0, 0, \sigma, \sigma, 0.5)$ with $\sigma^2 < 1$, as found in Tables 5 and 6. However, as shown in Tables 7

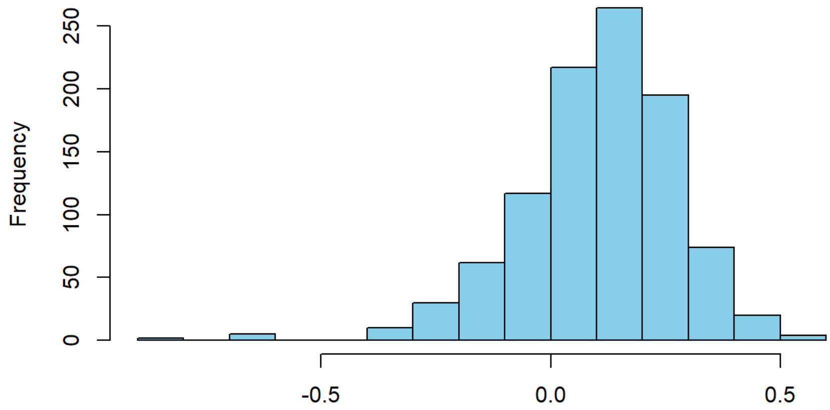


Figure 5: Histogram of bootstrap sample of the discrimination statistic.

and 8, the modified method successfully addresses these shortcomings within the problematic parameter space. The proposed modification performs well even for small sample sizes, ensuring that the probability of correct selection remains at least 0.5 in both cases without producing the counterintuitive results observed with the usual method.

We further examined the effect of correlation coefficient for modified approach as well and the findings are reported Supplementary Table S11 and S12, From Table S11, we observe that for a fixed sample size, the PCS is slightly lower when ρ is near the extremes (± 0.99) for the BVBS model. Conversely, in Table S12, the PCS tends to be slightly higher at these extreme ρ values for the BVLN model, with a similar pattern of PCS increasing as sample size grows. For other values of ρ, the PCS remains relatively similar for both models.

6 Real data analysis

Revisiting the dataset discussed in Section 1, we now determine the more appropriate distribution using our proposed decision criterion. The MLEs for respective distribution were as follows: For the BVBS distribution:

$$\hat{\alpha_1} = 0.4255, \quad \hat{\alpha_2} = 0.4623, \quad \hat{\beta}_1 = 199.6003, \quad \hat{\beta}_2 = 235.5726, \quad \hat{\rho}_1 = 0.6925.$$

For the BVLN distribution:

$$\hat{\mu_1} = 5.2930, \quad \hat{\mu_2} = 5.4609, \quad \hat{\sigma_1} = 0.4281, \quad \hat{\sigma_2} = 0.4650, \quad \hat{\rho}_2 = 0.6896.$$

Since all estimated parameters α_1, α_2, σ_1, and σ_2 are all less than 1, we apply the modified decision rule for model selection. The computed test statistic is

$$T^* = L_{BVBS}(\hat{\Gamma} \mid x, y) - L_{BVLN}(\hat{\Sigma} \mid x, y) = 0.3512,$$

with a cutoff value of $C = 0.0981$. Since $T^* > C$, we select the BVBS distribution as the preferred model.

To evaluate the probability of correct selection in this scenario, we conduct a parametric bootstrap analysis. The histogram of the bootstrap sample for the discrimination statistic is presented in Figure 5. Based on $1,000$ bootstrap replications, it is observed that the probability of correct selection is 0.611.

7 Conclusions

In this paper, we have addressed the classical problem of model discrimination between two non-nested distributions, specifically the five-parameter bivariate Birnbaum-Saunders and bivariate lognormal distributions. We employed the traditional likelihood ratio approach to discriminate between these distributions and observed that, for certain ranges of shape parameters, discrimination becomes challenging. To overcome this issue, we proposed a modified likelihood ratio method, which demonstrates improved performance.

Additionally, we explored the effects of model misspecification on the mode of the hazard gradient of the respective marginals, as well as on both components of the bivariate hazard gradient. If prior information is available, a Bayesian approach could be employed to discriminate between these distributions. Furthermore, this work may be extended for censored observations. Further exploration in higher dimensions and censored settings remains a promising direction for future research.

Acknowledgment

We are grateful to the Associate Editor and the reviewer for their valuable comments and suggestions, which were highly helpful in improving the quality of our work.

Appendix:

Proof of Lemma 3.1(a):

From Strong law of large number, we have:

$$-\frac{1}{n}\sum_{i=1}^{n} logf_{BVBS}(x_i, y_i; \bar{\Gamma}) \rightarrow -E_{BVBS}(logf_{BVBS}(X, Y; \bar{\Gamma})) \text{ wp } 1$$

Also, $-E_{BVBS}(logf_{BVBS}(X, Y; \bar{\Gamma}))$ has a minimum at true value ie: $\Gamma^0 = (\alpha_1^0, \alpha_2^0, \beta_1^0, \beta_2^0, \rho_1^0)$ which are identifiably unique.
Moreover,

$$-\frac{1}{n}\sum_{i=1}^{n} logf_{BVBS}(x_i, y_i; \hat{\Gamma}) = \inf_{\bar{\Gamma}\in\Theta} -\frac{1}{n}\sum_{i=1}^{n} logf_{BVBS}(x_i, y_i; \bar{\Gamma})$$

Using Lemma 2.2 of White (1980), we have, $\hat{\Gamma} \rightarrow \Gamma^0$ a.s.

Proof of Lemma 3.1(b):

Using similar flow of logic as used in Lemma 3.1(a), proof of Lemma 3.1(b) can be established.

Proof of Lemma 3.1(c):

Now for proving statement of Lemma 3.1(c), we need to show that :

$$\frac{1}{\sqrt{n}}[T - E_{BVBS}(T|\Gamma^0) - T^* + E_{BVBS}(T^*)] \xrightarrow{P} 0$$

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

$$\frac{1}{\sqrt{n}}[T - T^*] \xrightarrow{P} 0$$

followed by proving :

$$E_{BVBS}\left(\frac{1}{\sqrt{n}}[T - T^*]\right) \longrightarrow 0$$

Before that let: $h(X_i, Y_i; \chi) = logf_{BVBS}(X_i, Y_i; \Gamma) - logf_{BVLN}(X_i, Y_i; \Sigma)$ where
$\chi = (\alpha_1, \alpha_2, \beta_1, \beta_2, \rho_1, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho_2) = (\chi_1, \chi_2, \chi_3, \chi_4, \chi_5, \chi_6, \chi_7, \chi_8, \chi_9, \chi_{10})$
Now consider,

$$\begin{aligned}\frac{1}{\sqrt{n}}[T - T^*] &= \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}[h(X_i, Y_i; \hat{\chi}) - h(X_i, Y_i; \tilde{\chi})]\right] \\ &= \sum_{j=1}^{10}\left[\frac{1}{n}\sum_{i=1}^{n} h_j^{'}(X_i, Y_i; \tilde{\chi_j})\right]\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j}) + o_p(1)\end{aligned}$$

where $h_j^{'}(.) = \frac{\partial h(.)}{\partial \chi_j}$.Here $\hat{\chi} = (\hat{\alpha_1}, \hat{\alpha_2}, \hat{\beta}_1, \hat{\beta}_2, \hat{\rho_1}, \hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{\rho_2})$ and
$\tilde{\chi} = (\alpha_1^0, \alpha_2^0, \beta_1^0, \beta_2^0, \rho_1^0, \tilde{\mu_1}, \tilde{\mu_2}, \tilde{\sigma_1}, \tilde{\sigma_2}, \tilde{\rho_2})$.$\hat{\chi_j}$ is the j-th component of $\hat{\chi}$ and $\tilde{\chi_j}$ is the j-th component of $\tilde{\chi}$ for $j = 1(1)10$. Thus,using WLLN, we have:

$$\frac{1}{n}\sum_{i=1}^{n} h_j^{'}(X_i, Y_i; \tilde{\chi_j}) \xrightarrow{P} E_{BVBS}(h_j^{'}(X, Y; \chi))\Big|_{\chi=\tilde{\chi}} \tag{6}$$

$$E_{BVBS}(h_j^{'}(X, Y; \chi))\Big|_{\chi=\tilde{\chi}} = \left[\int_0^{\infty}\int_0^{\infty} \frac{\partial}{\partial \chi_j} log\left(\frac{f_{BVBS}(x, y; \Gamma)}{f_{BVLN}(x, y; \Sigma)}\right) f_{BVBS}(x, y; \Gamma^0)\, dx\, dy\right]\Big|_{\chi=\tilde{\chi}} \tag{7}$$

Now for first component of χ Equation (10) being the expectation of score function is 0, under usual regularity conditions that allow us to interchange the operations of integration and differentiation. For last five components, after interchanging integration and differentiation using dominated convergence theorem (DCT), we have:

$$\frac{\partial}{\partial \Sigma_k} E_{BVBS}(logf_{BVLN(X,Y,\Sigma)})\Big|_{\chi=\tilde{\chi}} = 0$$

where Σ_k is k-th component of $\Sigma = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_2)$ Therefore,

$$E_{BVBS}(h_j^{'}(X, Y; \chi))\Big|_{\chi=\tilde{\chi}} = 0 \tag{8}$$

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

$$\frac{1}{\sqrt{n}}[T - T^*] \xrightarrow{P} 0$$

Now,

$$\begin{aligned}
E_{BVBS}\left(\frac{1}{\sqrt{n}}[T - T^*]\right) &= E_{BVBS}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[h(X_i, Y_i; \hat{\chi}) - h(X_i, Y_i; \tilde{\chi})] \\
&\leq \sum_{j=1}^{10} E_{BVBS}\left|\left[\frac{1}{n}\sum_{i=1}^{n} h_j^{'}(X_i, Y_i; \tilde{\chi_j})\right]\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j})\right| + o(1) \\
&\leq \sum_{j=1}^{10}\left[E_{BVBS}[\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j})]^2 E_{BVBS}[\frac{1}{n}\sum_{i=1}^{n} h^{'}(X_i, Y_i; \tilde{\chi_j})]^2\right]^{\frac{1}{2}} + o(1)
\end{aligned}$$

Here $o(1)$ refers to a sequence that tends towards 0. Last step is written using Cauchy-Schwarz inequality. Since $E_{BVBS}[\sqrt{n}(\hat{\chi_j} - \tilde{\chi_j})]^2 < \infty$ (using theorem 3.2 of White (1982)) and using equation (8), we have,

$$E_{BVBS}\left[\frac{1}{n}\sum_{i=1}^{n} h^{'}(X_i, Y_i; \tilde{\chi_j})\right]^2 \longrightarrow 0$$

Thus, $E_{BVBS}\left(\frac{1}{\sqrt{n}}[T - T^*]\right) \longrightarrow 0$
Hence, the 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

Abid, R. and Kokonendji, C. C. (2023). Choice between and within the classes of poisson-tweedie and poisson-exponential-tweedie count models. *Communications in Statistics-Simulation and*

Computation, 52(5):2115–2129.

Ahmad, M. A., Raqab, M. Z., and Kundu, D. (2017). Discriminating between the generalized rayleigh and weibull distributions: Some comparative studies. *Communications in Statistics-Simulation and Computation*, 46(6):4880–4895.

Atkinson, A. C. (1969). A test for discriminating between models. *Biometrika*, 56(2):337–347.

Atkinson, A. C. (1970). A method for discriminating between models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 32(3):323–345.

Bain, L. J. and Engelhardt, M. (1980). Probability of correct selection of weibull versus gamma based on livelihood ratio. *Communications in statistics-theory and methods*, 9(4):375–381.

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

Basu, B., Tiwari, D., Kundu, D., and Prasad, R. (2009). Is weibull distribution the most appropriate statistical strength distribution for brittle materials? *Ceramics International*, 35(1):237–246.

Basu, S. and Kundu, D. (2023). Model misspecification of log-normal and birnbaum-saunders distributions. *Communications in Statistics-Simulation and Computation*, pages 1–21.

Chen, W. (1980). On the tests of separate families of hypotheses with small sample size. *Journal of Statistical Computation and Simulation*, 11(3-4):183–187.

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.

Cox, D. R. (1962). Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society: Series B (Methodological)*, 24(2):406–424.

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.

Diyali, B., Kumar, D., and Singh, S. (2023). Discriminating between log-normal and log-logistic distributions in the presence of type-ii censoring. *Computational Statistics*, pages 1–25.

Diyali, B., Kumar, D., and Singh, S. (2024). Discriminating among inverse weibull, lognormal, and inverse gaussian distributions. *Quality and Reliability Engineering International*, 40(4):1698–1718.

Dumonceaux, R., Antle, C. E., and Haas, G. (1973). Likelihood ratio test for discrimination between two models with unknown location and scale parameters. *Technometrics*, 15(1):19–27.

Dyer, A. R. (1973). Discrimination procedures for separate families of hypotheses. *Journal of the American Statistical Association*, 68(344):970–974.

Gupta, R. C., Kannan, N., and Raychaudhuri, A. (1997). Analysis of lognormal survival data. *Mathematical biosciences*, 139(2):103–115.

Gupta, R. D., Kundu, D., and Manglick, A. (2002). Probability of correct selection of gamma versus ge or weibull versus ge based on likelihood ratio statistic. In *Recent advances in statistical methods*, pages 147–156. World Scientific.

Hellinger, E. (1909). Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. *Journal für die reine und angewandte Mathematik*, 1909(136):210–271.

Johnson, N. L. and Kotz, S. (1975). A vector multivariate hazard rate. *Journal of Multivariate Analysis*, 5(1):53–66.

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.

Kundu, D., Balakrishnan, N., and Jamalizadeh, A. (2010). Bivariate birnbaum–saunders distribution and associated inference. *Journal of multivariate Analysis*, 101(1):113–125.

Pasari, S. (2018). Stochastic modelling of earthquake interoccurrence times in northwest himalaya and adjoining regions. *Geomatics, Natural Hazards and Risk*, 9(1):568–588.

Rajput, O., Kundu, D., and Mitra, S. (2026). Discriminating between bivariate weibull and bivariate log-normal distribution under type-i censoring: O. rajput et al. *Journal of Statistical Theory and Practice*, 20(1):5.

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

Raqab, M. Z., Al-Awadhi, S. A., and Kundu, D. (2018). Discriminating among weibull, log-normal, and log-logistic distributions. *Communications in Statistics-Simulation and Computation*, 47(5):1397–1419.

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.

Vasey, M. W. and Thayer, J. F. (1987). The continuing problem of false positives in repeated measures anova in psychophysiology: A multivariate solution. *Psychophysiology*, 24(4):479–486.

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.

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.