Two-sample density-based empirical likelihood tests for incomplete data in application to a pneumonia study

Albert Vexler, **Jihnhee Yu**

Department of Biostatistics, University at Buffalo, the State University of New York, New York 14214, U.S.A.

Email: avexler@buffalo.edu

ABSTRACT

In clinical trials examining the incidence of pneumonia it is common practice to measure infection via both invasive and non-invasive procedures. In the context of a recently completed randomized trial comparing two treatments the invasive procedure was only utilized in certain scenarios due to the added risk involved, and given that the level of the non-invasive procedure surpassed a given threshold. Hence, what was observed was bivariate data with a pattern of missingness in the invasive variable dependent upon the value of the observed non-invasive observation within a given pair. In order to compare two treatments with bivariate observed data exhibiting this pattern of missingness we developed a semi-parametric methodology utilizing the density-based empirical likelihood approach in order to provide a nonparametric approximation to Neyman-Pearson type test statistics. This novel empirical likelihood approach has both a parametric and nonparametric component. The nonparametric component utilizes the observations for the non-missing cases, while the parametric component is utilized to tackle the case where observations are missing with respect to the invasive variable. The method is illustrated through its application to the actual data obtained in the pneumonia study and is shown to be an efficient and practical method.

Keywords

Censored data; Empirical likelihood; Linear regression; Two sample test; Ventilator-associated pneumonia.

1. Introduction

The primary goal of this article is to propose and examine the method for comparing two treatment groups based on the incomplete bivariate data in the context of the ventilator-associate pneumonia study carried out at University at Buffalo, State University of New York (Scannapieco et al. 2009). The ventilator-associate pneumonia is a nosocomial infection among subjects hospitalized in intensive care units (ICU), who are assisted by mechanical ventilators as they have a difficulty in breathing their own. Investigation of pneumonia incidence in this setting presents a unique demand, that is, practitioners need to maintain the accuracy of diagnosis as much as possible while also minimizing the use of an invasive procedure, such as biopsies, so as not to further compromise a subject's quality-of-life. To deal with this requirement, a typical pneumonia diagnosis consists of two procedures, one is a non-invasive procedure used to check the relevant symptoms of pneumonia, e.g., the clinical pulmonary infection score, or clinical pulmonary infection score (CPIS). The other measure consists of a biopsy to determine micro-organism counts directly from a subject's lung (e.g., bronchoalveolar lavage, or BAL). While not an invasive procedure, CPIS itself consists of intensive symptom measurements such as fever, blood leukocyte counts, tracheal secretions, oxidation index and chest X-ray. The CPIS is known to have a range of sensitivities of 72% to 77% and specificities of 42% to 85% (Koenig and Truwit, 2006) relative to detecting a true pneumonia diagnostis. BAL, an invasive procedure has an estimated sensitivity of 73% and an estimated specificity of 82% (Dupont et al., 2004) in terms of a pneumonia diagnosis. In receiver operating characteristic (ROC) curve analyses, various values of the area under the ROC curve have been reported, e.g., 0.82 in Swoboda et al.

(2006), 0.69 in Huh et.al (2008), and 0.873 in Ramirez et al., (2008). The CPIS is generally available for most of subjects, while the results from a BAL are more often than not incomplete due to the fact that a BAL is in general performed only for subjects with a CPIS value greater than 5. Since both outcomes convey valuable information in terms of the severity of pneumonia related symptoms it is important to consider the joint information contained in both measures. Yu et al. (2010) and Vexler et al. (2010) developed approaches that allowed using both outcomes within a comparative study. To handle the inferential problems, the authors incorporated a nonparametric empirical likelihood approach (EL) methodology (Owen, 1988, 1990, 1991, 2001) in conjunction with the appropriate functional model. The EL method was utilized to provide more flexibility with respect to the departure from the underlying distributional assumptions. The EL method incorporates information or assumptions regarding the parameters and subsequently translates this information into the likelihood estimation process. Thus the method can be used to combine additional information about the parameters of interest (e.g., Qin and Lawless, 1994; Lazar and Mykland, 1998). The distribution free statements of the EL allow the method to be used for analyzing data with potentially complex underlying distributional forms. For example, Qin and Leung (2005) used a semiparametric likelihood approach to estimate the distribution of malaria parasite levels, which is a mixture distribution, where a component of the mixture distribution was a mixture of discrete and continuous distributions. Often, in the context of nonparametric testing for parameters of populations, EL's have exhibited many asymptotic properties that are equivalent to conventional parametric likelihoods (e.g., Lazar and Mykland, 1998). In part this is due to the fact that conventional EL methodology utilizes forms of likelihood functions that are based on distribution functions (e.g., Owen, 2001).

Powerful parametric test statistics are commonly based on joint density functions (e.g., Lehmann and Romano, 2005; Vexler and Wu, 2009; Vexler, Wu and Yu, 2010). In this vein, Vexler and Gurevich (2010) developed an EL methodology that conserves the form of the density function for the likelihood statistics. They proposed density-based EL tests for one sample goodness-of-fit problems, which have an entropy-based structure, provided that null distributions have known forms. To test the distribution of a data set against a set of known distribution (e.g., a normal distribution), the alternative likelihood function is maximized following EL concepts, where empirical constraints regarding the densities are modified in the form of a summation of unknown components based on the set of order statistics (Vexler and Gurevich, 2010; Vexler et al., 2011). The ultimate EL test statistic consists of components based on sample entropy, which makes it possible to investigate the theoretical characteristics via the existing theorems regarding well-known entropy based statistics (e.g., Vasicek, 1976; Dedewicz and Van Der Meulen, 1981). Historically, sample entropy based test statistics have been difficult to adapt to common data analyses due to the problem of finding a favorable value of an integer parameter inherent in the structure of the test statistic. We will explore this issue further in Section 2. In the context of the EL method, Vexler and Gurevich (2010) attended to this issue with respect to their proposed methods. Through a theoretical investigation and an extensive Monte-Carlo study, the authors illustrated that their proposed tests are powerful one sample nonparametric procedures that approximate the parametric likelihood ratio test, when null distributions have known forms. In this article, we use the idea introduced by Vexler and Gurevich (2010) to provide a nonparametric approximation to the Neyman-Pearson type test statistic, that is used in the standard two-sample problem when the null and alternative distributions are completely unknown.

We have two primary objectives in this article. The first objective is to construct a semi-parametric test with an approximate likelihood ratio structure and to carry out an investigation regarding its operating characteristics. The other objective is to adapt the new approach to analyzing the actual clinical trial data with respect to conducting a two-sample comparison based on the observed incomplete bivariate data (CPIS, BAL). In contrast to the earlier work of Yu et al. (2010) and Vexler et al. (2010), this article's focus is on detecting a difference in the joint distributions, as compared to testing for difference in population quantities, e.g., location parameters. In this context, we conclude regarding outcomes of the pneumonia study in a different manner from that of Yu et al. (2010) and Vexler et al. (2010). In addition, it should be noted that oftentimes distribution based EL method-based ratio-type tests show poor characteristics when the tests utilize skewed data sets, e.g., see Vexler et al. (2009). Our proposed approach does not have this deficiency, and as will be shown, demonstrates good behavior given data with small sample sizes as well given data that are skewed.

The rest of this paper has the following structure. We first specify the combined likelihood for the incomplete and complete data. Then, the EL likelihood ratio test is developed for the purpose of comparing two treatment groups relative to the respective bivariate distributions. Second, the direct application to the actual data set will be carried out. Third, through an extensive Monte-Carlo study, we investigate the properties of the likelihood ratio test. Finally, we conclude the article with concluding remarks.

2. Main Results

In this section, we first formalize the data structure and the corresponding likelihood functions. For the development of the relevant likelihood ratio test, we utilize the framework of Yu et al. (2010) as a starting point. Suppose that there are n_1 and n_2 observations for the control and treatment groups, respectively. Let (X_{ij}, Y_{ij}) denote independent bivariate data points, where the observations for the control group are indicated by subscripts $i=1$ and $j=1,\ldots,n_1$ and those for the treatment groups are indicated by subscripts $i=2$ and $j=1,\ldots,n_2$. Note that for both groups Y_{ij} is observed conditional on X_{ij} exceeding a certain threshold value, c_{ij}. Let the observation X_{ij} have unknown density function, $f_{X_i}(x_{ij})$ and let Y_{ij} given X_{ij} have parametric conditional density function $g_{Y_i}(y_{ij} \mid x_{ij}, \theta_i)$, where θ_i is a vector of parameters corresponding to group *i*. In a similar manner to Qin (2000) and Yu et al. (2010), we construct the likelihood function, L, having the form

$$\begin{aligned} L &= \prod_{i=1}^{2} \prod_{j=1, x_{ij} \geq c_{ij}}^{n_i} f_{X_i}(x_{ij}) g_{Y_i}(y_{ij} \mid x_{ij}, \theta_i) \prod_{j=1, x_{ij} < c_{ij}}^{n_i} f_{X_i}(x_{ij}) \\ &= \prod_{i=1}^{2} \prod_{j=1, x_{ij} \geq c_{ij}}^{n_i} g_{Y_i}(y_{ij} \mid x_{ij}, \theta_i) \prod_{j=1}^{n_i} f_{X_i}(x_{ij}), \end{aligned} \tag{2.1}$$

which includes both groups. The corresponding likelihood ratio statistic is given by

$$R = \frac{\prod_{i=1}^{2} \prod_{j=1, x_{ij} \geq c_{ij}}^{n_i} g_{Y_i}(y_{ij} \mid x_{ij}, \theta_i) \prod_{j=1}^{n_i} f_{X_i}(x_{ij})}{\prod_{i=1}^{2} \prod_{j=1, x_{ij} \geq c_{ij}}^{n_i} g_{Y}(y_{ij} \mid x_{ij}, \theta) \prod_{j=1}^{n_i} f_{X}(x_{ij})}, \tag{2.2}$$

where $g_Y(y_{ij} \mid x_{ij}, \theta)$ and $f_X(x_{ij})$ are the densities for the pooled observations. Yu et al. (2010) proposed to replace $\prod_{i=1}^{2} \prod_{j=1}^{n_i} f_{X_i}(x_{ij})$ with $\prod_{i=1}^{2} \prod_{j=1}^{n_i} p_{ij}$. The EL functions for null and alternative hypotheses were obtained by maximizing $\prod_{i=1}^{2} \prod_{j=1}^{n_i} p_{ij}$ subject to constraints related to the respective hypothesis. Then the EL ratio test was derived based on the maximized likelihood functions. Specifically, Yu et al. (2010) used the equality of the two group means as the constraint under the null hypothesis. The maximizations were achieved by the Lagrange multiplier method. Utilizing the EL function allows more generalizability as compared to the direct method based on parametric models due to the fact that the EL method is intrinsically nonparametric.

Vexler and Gurevich (2010) proposed the density-based EL approach to obtain the likelihood ratio test for one-sample goodness-of-fit problems when the null distribution has a known parametric form. The proposed density-based EL technique directly approximates density-based likelihood functions. In this case, the optimal properties of the new EL ratio test are expected to follow the Neyman-Pearson Lemma (e.g., Lehmann and Romano, 2005; Vexler and Wu, 2009; Vexler, Wu and Yu, 2010). The density-based test statistic then is a function of the order statistics, having a similar structure to that of the class of entropy-based statistics.

2.1. Derivation of the EL ratio.

In this section, we first extend the approach of Vexler and Gurevich (2010) to the two-sample nonparametric comparison. Let $X_{i(j)}$ indicate j-th order statistic for group i. The likelihood function then can be presented as

$$L_1 = \prod_{i=1}^{2}\prod_{j=1}^{n_i} f_{X_i}(x_{ij}) = \prod_{i=1}^{2}\prod_{j=1}^{n_i} f_{X_i}(x_{i(j)}). \tag{2.3}$$

Under H_0 (equal distributions) we assume that both groups are identically distributed as $f_X(x)$.

Thus, the likelihood function under H_0 is given by

$$L_0 = \prod_{i=1}^{2}\prod_{j=1}^{n_i} f_X(x_{ij}) = \prod_{i=1}^{2}\prod_{j=1}^{n_i} f_X(x_{i(j)}). \tag{2.4}$$

By (2.3) and (2.4), the likelihood ratio can be expressed as

$$R = \frac{\prod_{i=1}^{2}\prod_{j=1}^{n_i} f_{X_i}(x_{i(j)})}{\prod_{i=1}^{2}\prod_{j=1}^{n_i} f_X(x_{i(j)})} = \frac{\prod_{i=1}^{2}\prod_{j=1}^{n_i} f_{1ij}}{\prod_{i=1}^{2}\prod_{j=1}^{n_i} f_{0ij}}, \quad f_{1ij} = f_{X_i}(x_{i(j)}),\, f_{0ij} = f_X(x_{i(j)}). \tag{2.5}$$

We begin with an evaluation of the likelihood $L_{11}=\prod_{j=1}^{n_1} f_{11j}$. Following the EL concept, we approximate the likelihood L_{11}, finding values of f_{11j} that maximize L_{11} subject to an empirical constraint. This constraint is naturally associated with the restriction $\int_{-\infty}^{\infty} f_{X_1}(x)dx=1$. To formalize the constraint, we consider the following result shown in Vexler and Gurevich (2010).

Proposition 2.1. *Suppose that independent identically distributed observations* $X_1,\ldots,X_n$ *have the density function,* $f(x)$ *. For all positive integer* $m\le n/2$ *,*

$$\frac{1}{2m}\sum_{j=1}^{n}\int_{X_{(j-m)}}^{X_{(j+m)}} f(x)dx=\int_{X_{(1)}}^{X_{(n)}} f(x)dx-\sum_{k=1}^{m-1}\frac{m-k}{2m}\int_{X_{(n-k)}}^{X_{(n-k+1)}} f(x)dx-\sum_{k=1}^{m-1}\frac{m-k}{2m}\int_{X_{(k)}}^{X_{(k+1)}} f(x)dx, \quad (2.6)$$

where $X_{(j)}=X_{(1)}$, *if* $j\le 1$, *and* $X_{(j)}=X_{(n)}$, *if* $j\ge n$.

One can show that the last two terms in the right-side of (2.6) vanish asymptotically to 0 when $m,n\to\infty$ and $m/n\to 0$ (for details, see Appendix A as well as Lemma 2.1. and its proof that appear in Vexler et al., 2011). Applying Proposition 2.1 to the observations in the control group, we have the inequality

$$\frac{1}{2m}\sum_{j=1}^{n_1}\int_{X_{1(j-m)}}^{X_{1(j+m)}}\frac{f_{X_1}(u)}{f_X(u)} f_X(u)du\le 1. \qquad (2.7)$$

By utilizing the approximate analog to the mean value integration theorem, we can write

$$\int_{X_{1(j-m)}}^{X_{1(j+m)}}\frac{f_{X_1}(u)}{f_X(u)} f_X(u)du\cong\frac{f_{11j}}{f_{01j}}\int_{X_{1(j-m)}}^{X_{1(j+m)}} f_X(u)du.$$

Therefore, approximating (2.7), we can represent (2.7) in the form of

$$\frac{1}{2m}\sum_{j=1}^{n_1}\frac{f_{11j}}{f_{01j}}\int_{X_{1(j-m)}}^{X_{1(j+m)}} f_X(u)du=\frac{1}{2m}\sum_{j=1}^{n_1}\left[F_X(x_{1(j+m)})-F_X(x_{1(j-m)})\right]\frac{f_{11j}}{f_{01j}}\le 1, \qquad (2.8)$$

with an appropriate choice of m, and where $F_X(x)$ denotes the distribution function under H_0. Since F_X can be estimated via the empirical distribution function, an empirical constraint corresponding to (2.7) takes the form

$$\frac{1}{2m}\sum_{j=1}^{n_1}\frac{f_{11j}}{f_{01j}}\Delta_{m1j}\leq 1, \tag{2.9}$$

where

$$\Delta_{mij}=\frac{1}{n_1+n_2}\sum_{k=1}^{2}\sum_{l=1}^{n_k}\left(I(x_{kl}\leq x_{i(j+m)})-I(x_{kl}\leq x_{i(j-m)})\right), \tag{2.10}$$

and I denotes an indicator function that takes the value 1 if the condition in the parenthesis is satisfied, and takes the value 0, otherwise.

The maximization of $\log L_{11}=\sum_{j=1}^{n_1}\log f_{11j}$, the log of the likelihood function, subject to the constraint at (2.9) can be achieved using the method of Lagrange multipliers. We find the stationary points of the Lagrange function given as

$$\sum_{j=1}^{n_1}\log f_{11j}+\lambda_1\left(1-\frac{1}{2m}\sum_{j=1}^{n_1}\frac{f_{11j}}{f_{01j}}\Delta_{m1j}\right), \tag{2.11}$$

where λ_i is the Lagrange multiplier. Taking a partial derivative of (2.11) with respect to f_{11k}, we obtain

$$\frac{1}{f_{11k}}-\frac{1}{2m}\lambda_1\frac{\Delta_{m1k}}{f_{11k}}=0,\quad k=1,\ldots,n_1. \tag{2.12}$$

Summing (2.12) over $k=1,\ldots,n_1$ and using (2.9) produces

$$f_{11k}=f_{01k}\frac{2m}{n_1\Delta_{m1k}}. \tag{2.13}$$

Following a similar scheme described above for group 2, we conclude

$$f_{12k} = f_{02k} \frac{2v}{n_2 \Delta_{v2k}}, \tag{2.14}$$

with integer v and $f_{lik}, \Delta_{mik}, l = 0,1, k = 1,\ldots,n_i, i = 1,2$ defined by (2.5) and (2.10), respectively. Combing the results found at equation (2.13) and (2.14) we arrive at the nonparametric approximation

$$\tilde{R}_{m,v,n_1,n_2} = \prod_{j=1}^{n_1} \frac{2m}{n_1 \Delta_{m1j}} \prod_{j=1}^{n_2} \frac{2v}{n_2 \Delta_{v2j}} \tag{2.15}$$

to the optimal parametric likelihood ratio (2.5). One can show the statistic $\tilde{R}_{m,v,n_1,n_2}$ has a structure that is similar to that of entropy-based test statistics for goodness-of-fit problems, e.g., Vasicek (1976) and Dedewicz, Van Der Meulen (1981) for details. Also, see Vexler and Gurevich (2010) for additional references. In general it may be shown that these statistics have asymptotically optimal properties, e.g., Tusnady (1977) who illustrates this point. In the context of entropy-based statistics, m (or v in our notation) is referred to as an integer parameter. The proper selection of the integer parameter among the possible range of values is an important issue in this setting, since the operating characteristic of the test statistic is directly dependent upon the corresponding value. The current literature related to entropy-based decision making recommends selecting values of the integer parameter utilizing information regarding alternative distributions when sample sizes are finite, e.g., Vasicek (1976) and Mudholkar and Tian (2002, 2004) for examples. As it turns out, EL methodology may be utilized with respect to specifying the integer parameter. A discussion regarding the method for choosing the integer value may be found in Appendix A. Some in-depth discussion is also found in Vexler and Gurevich (2010) in the context of one-sample goodness-of-fit testing problems. Thus, the density-based EL likelihood ratio test statistic takes the form

$$\tilde{R}_{n_1,n_2} = \min_{l_{n_1} \le m \le u_{n_1}} \prod_{j=1}^{n_1} \frac{2m}{n_1 \Delta_{m1j}} \min_{l_{n_2} \le v \le u_{n_2}} \prod_{j=1}^{n_2} \frac{2v}{n_2 \Delta_{v2j}}, \qquad (2.16)$$
$$l_n = n^{0.5+\delta}, u_n = \min(n^{1-\delta}, n/2), \delta \in (0, 0.25)$$

where the bounds l_n, u_n are necessary with respect to following Proposition 2.2, where we need $m/n_1 \to 0$, $v/n_2 \to 0$ in order for the distributional assumptions to be met. The bounds l_n, u_n can be associated with restrictions on integer parameters included in the structure of entropy-based statistic with respect to the asymptotic theorems of Vasicek (1976), Tusnady (1977) and Vexler and Gurevich (2010). Thus, we propose rejecting the null hypothesis, H_0, for large values of $\tilde{R}_{n_1,n_2}$. Note that similar to Canner (1975), we define $\Delta_{lij} = 1/(n_1 + n_2)$, if $\Delta_{lij} = 0$. Proposition 2.2 regarding the consistency of the proposed test directly utilizes these bounds in its proof scheme. In this article, we specify $\delta = 0.1$. It is shown in Section 3 within our extensive Monte Carlo study that the power of the test statistic (2.16) is not associated significantly with values of $\delta \in (0, 0.25)$ under various alternatives. Since the technique mentioned above illustrates that the form of the density based EL statistic is a nonparametric approximation to the optimal parametric likelihood ratio (2.5), it should be anticipated that this approach has good power properties. Notice that, since $I(X > Y) = I(F(X) > F(Y))$ in (2.10) for any distribution function *F*, the null distribution of $\tilde{R}$ is independent with respect to the form of the underlying distributions given H_0. Hence, we can tabulate universal critical values regardless of the distribution of the X_{ij}'s. Thus, the EL ratio test statistic at (2.16) is simple and provides an exact test. Some distinctive characteristics of the proposed test statistic as compared to the typical EL approach (Owen, 2001 for one sample test; Yu et al., 2010 for two-sample test) are summarized in Table 1.

Table 1. *The comparison of the typical EL and density based EL approaches*.

Characteristics	**Owen (1988); Yu, Vexler and Tian (2010)**	**The proposed method**
Construction of the likelihood function	Distribution based	Density based
Usage of Lagrange multipliers method	Yes	Yes
Usage of constraints for maximization	Yes	Yes
Focus of the test	Parameters (e.g., moments)	Overall distributions
Critical value	Asymptotic	Exact
The form of the test statistic	Numeric approach required	No numeric approach

A substantial body of literature has now grown around the asymptotic distribution problems involving Vasicek's entropy estimator and the corresponding test statistics (e.g., Dudewicz and van der Meulen, 1981; van Es, 1992). However, it is generally recognized that the asymptotic distribution of entropy-based type statistics, which depend on estimates of nuisance parameters, are analytically difficult to specify. The details related to this issue is within the literature related to nonparametric tests (e.g., Canner, 1975; Hall and Welsh, 1983; Mudholkar and Tian, 2002, 2004). Hence, we will not attempt to provide further details with respect to an analytical solution for the critical values for our proposed test. We utilize a Monte Carlo approach to determine the necessary critical values.

Now, we introduce a proposition for the purpose of proving the asymptotic consistency of the new test at (2.16). Assume $n_1 / n_2 \to \eta$, as $n_1, n_2 \to \infty$, where $\eta > 0$ is a constant, then we have the following:

Proposition 2.2. Let the expectations $E\left(\log f_{X_k}\left(X_{s1}\right)\right)$ and $E\left(\log f_X\left(X_{s1}\right)\right)$ be finite for all *k*, $k = 1,2$ and *s*, $s = 1,2$. Then

$$\frac{1}{n_1+n_2}\log\left(\tilde{R}_{n_1,n_2}\right)\xrightarrow{p}\tau \text{, as } n_1,n_2\to\infty,$$

where

$$\tau=-\frac{\eta}{1+\eta}E\left(\log\left(\frac{\eta}{1+\eta}+\frac{1}{1+\eta}\frac{f_{X_2}(X_{11})}{f_{X_1}(X_{11})}\right)\right)-\frac{1}{1+\eta}E\left(\log\left(\frac{1}{1+\eta}+\frac{\eta}{1+\eta}\frac{f_{X_1}(X_{21})}{f_{X_2}(X_{21})}\right)\right)$$

and $n_1/n_2\to\eta>0$.

Proof. See Appendix B.

Since under the null hypothesis the ratio $f_{X_1}/f_{X_2}=1$ we have $\tau=0$. Under the alternative hypothesis, H_1, we have $E\left(f_{X_2}(X_{11})/f_{X_1}(X_{11})\right)=E\left(f_{X_1}(X_{21})/f_{X_2}(X_{21})\right)=1$. Hence, we may conclude that

$$\tau\geq-\frac{\eta}{1+\eta}\left(\log\left(\frac{\eta}{1+\eta}+\frac{1}{1+\eta}E\frac{f_{X_2}(X_{11})}{f_{X_1}(X_{11})}\right)\right)-\frac{1}{1+\eta}\left(\log\left(\frac{1}{1+\eta}+\frac{\eta}{1+\eta}E\frac{f_{X_1}(X_{21})}{f_{X_2}(X_{21})}\right)\right)=0,$$

thus further verifying the consistency of the proposed density-based EL test. The bounds $n_1^{0.5+\delta}\leq m\leq n_1^{1-\delta}, n_2^{0.5+\delta}\leq v\leq n_2^{1-\delta}, \delta\in(0,0.25)$ and the occurrence of the minimum values with respect to m and v in the approximation (2.16) are strongly justified in the proof of Proposition 2.2.

2.2. Combining test.

Now, to use all available information regarding effect of treatment, the conditional distribution $g_i(y_{ij}\mid x_{ij},\theta)$ should be involved in the full likelihood function as in (2.1). We assume a simple linear model with respect to the functional relationship between CPIS and BAL values. Note that Pugin et al. (1991), Yu et al. (2010) and Vexler et al. (2010) give detailed discussions of the relationship between CPIS and BAL values that support a linear model assumption. Following the linear model assumption, the conditional distribution $g_i(y_{ij} \mid x_{ij}, \theta)$ is given as

$$g_i(y_{ij} \mid x_{ij}, \theta_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(y_{ij} - \alpha_i - \beta_i x_{ij})^2}{2\sigma_i^{\,2}} \right), \tag{2.17}$$

where α_i, β_i and $\sigma_i^{\,2}$ are the intercept, slope and variance of assumed normally distributed residuals, respectively. Let $\hat{\sigma}^2$ and $\hat{\sigma}_i^{\,2}$ denote the estimators of the pooled variance of the residuals under H_0 and the group i-variance of the residuals under H_1, respectively. The likelihood ratio statistic at (2.2) then can be approximated in a semi-parametric fashion through the combination of (2.17) and (2.16), respectively, and taking the form

$$R = \left(\min_{l_{n_1} \le m \le u_{n_1}} \prod_{j=1}^{n_1} \frac{2m}{n_1 \Delta_{m1j}} \quad \min_{l_{n_1} \le v \le u_{n_1}} \prod_{j=1}^{n_2} \frac{2v}{n_2 \Delta_{v2j}} \right) \left(\frac{\hat{\sigma}^{2(n_1^* + n_2^*)}}{\hat{\sigma}_1^{\,2n_1^*} \hat{\sigma}_2^{\,2n_2^*}} \right), \tag{2.18}$$

where n_1^* and n_2^* are the number of observed Y_{ij} for groups 1 and 2, respectively. Thus, the proposed test is to reject the null hypothesis for large values of the statistic, *R*, at (2.18). To determine the critical values of this test, we remark that the statistic $2\log\left(\hat{\sigma}^{2(n_1^* + n_2^*)}\right) - 2\log\left(\hat{\sigma}_1^{\,2n_1^*} \hat{\sigma}_2^{\,2n_2^*}\right)$ has an asymptotic χ^2 distribution with three degrees-of-freedom. Thus, the critical values C_α of the test statistic *2log(R)* are found using the convolution

$$\int_0^\infty \Pr\left(2\log\left(\tilde{R}_{n_1,n_2}\right) > C_\alpha - u\right) f_{\chi^2_{df=3}}(u)\, du = \alpha\,,$$

where $\tilde{R}_{n_1,n_2}$ at (2.16) is free of the underlying distributions, $f_{\chi^2_{df=3}}(u)$ is a $\chi^2_{df=3}$-density function and α denotes the level of significance. In this case, we require that $E\left|X_{ij}\right|^3 < \infty$ and $n_1^*, n_2^* \to \infty$. We tabulated the critical values for the new test in Tables 2 and 3. Here $\tilde{R}_{n_1,n_2}$ is from (2.16) with $\delta = 0.1$. For each level of significance and sample size, the critical value was obtained based on 50,000 replications. Notice that the critical values are functions of the sample sizes only. In order to examine the performance of the tabulated critical values given small sample sizes we conducted a broad Monte Carlo simulation study. The simulation study, based on the values from Tables 2 and 3, demonstrated excellent Type I error control relative to the desired level, when $n_1, n_2 > 15$.

TABLES 2&3 HERE

3. Simulation study

In order to investigate the performance of the proposed testing method we performed an extensive simulation study. The estimated Type I error and power of the new tests were investigated using various underlying distributions. All tests were conducted at a significance level of 0.05.

3.1. Evaluation of the EL ratio test.

First, we carried out a broad Monte-Carlo simulation study of the test statistic (2.16) for the two sample comparison with X_{ij} alone. Each case is based on 10,000 simulations. The accuracy and power of the proposed method are shown in Tables 4 and 5.

TABLES 4&5 HERE

Regarding the results presented in Table 4, values for X_{ij} were generated from $N(\mu_i, \sigma_{X_i}^2)$.The cases with $\mu_1 = \mu_2$ and $\sigma_{X_1} = \sigma_{X_2}$ examine the Type I error. The results in Table 4 show the excellent Type I error control with even relatively small sample sizes. The power of the test illustrated the ability of our test to differentiate differences between the two groups given that true differences do exist. Since the proposed test compares the distribution of the two groups, the power is demonstrated using the differences in both shape and location parameter. In order to investigate the operating characteristics of the new test with non-normal distributions, the cases with the lognormal distribution are examined with the corresponding results presented in Table 5. Given log-normally distributed data we see the same excellent performance of our test in terms of Type I error control as that obtained under the normal distribution assumption. This experimentally confirms the corresponding theoretical property of the proposed test. Also, the power of the proposed method increases accordingly when the difference between the distributions increases in terms of location and scale changes.

TABLE 6 HERE

Table 6 displays the additional cases of the power investigation of the new test based on statistic (2.16) (with $0<\delta=0.025, 0.05, 0.1, 0.12, 0.15, 0.2<0.25$) as compared to the classical two-sample procedures, namely, the Kolmogorov Smirnov test, the Wilcoxon rank-sum test and the two sample t-test. The power of the proposed test obtained in this Monte Carlo study does not depend significantly on value of $\delta \in (0,0.25)$. In the standard two-sample problem with a constant shift in location (especially when data follow normal distributions with different means and equivalent variances), the classic Wilcoxon test and t-test are known to have good power properties. In these specific cases (see, e.g., (3) in Table 6), the new test provides statistical power that is less efficient as compared to those of the classical procedures. However, the simulation study shows that the new test has high and stable power when detecting nonconstant shift alternatives. In this instance the classical procedures break down completely in terms of the power properties. The proposed test results in a relatively small power loss in comparison to Wilcoxon's rank-sum test and as compared to the t-test when the constant shift is dominant. In the non-constant shift alternative the new test is preferable in terms of large power gains.

3.2. Evaluation of the combined test.

As a next step, we carried out a broad Monte-Carlo simulation for the combined test statistic, 2log of (2.18). We consider the cases with and without thresholds, c_{ij} in (2.1). In cases subject to the thresholds, the variable X_{ij} is completely observed for both groups while the observation of Y_{ij} is depending on the threshold value, c_{ij}.

TABLE 7 HERE

Table 7 shows the results based on the normal and lognormal distributions (columns of Power 1 and Power 2, respectively) as the underlying distribution of X_{ij} and the residuals. The values of the thresholds for the normal and lognormal distributions were selected to be 0 and the median of X_{ij}, respectively, giving rise to an approximately 50% missingness rate under H_0. Note that the treatment group may have higher missingness rates when the difference of the parameters between control and treatment groups increases. For the normal distribution case, the values of Y_{ij} were generated based on the conditional density (2.17) with parameters α_i and β_i and σ_i^2. For the lognormal distribution, the appropriate parameters were used to achieve the semblance of the simulation based on the normal distribution in terms of the magnitude of the location parameters and variances. In Table 7 the Type I error is estimated with respect to the cases with equal parameters for both groups (i.e. $\mu_1 = \mu_2$, $\sigma_{X_1} = \sigma_{X_2}$, $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$, $\sigma_1 = \sigma_2$) . The reasonable Type I error control is shown regardless of the underlying distributions even given small sample sizes. A slight compromise of the accuracy is observed as compared with the simulation results in Tables 4 and 5. This is due to the fact that the test at (2.18) has two components, the parametric and nonparametric pieces. The study illustrates power increases according to the distance between parameters. Yu et al. (2010) proposed the combined test statistic based on the classical EL approach and carried out a similar Monte-Carlo study based on the normal distribution. Comparing the results in Table 7 with those found in Table 2 in Yu et al. (2010) we observe that the Type I error control is comparable, while the power of the proposed method exceeds the performance of the method based on the traditional EL approach. The cases with the lognormal distribution (column of Power 2) also demonstrate the excellent Type I error control and power of the new test. In Yu et al. (2010) it is also noted that the strategy of the combined test based on incomplete bivariate outcomes has increased power when the parametric relationship between two variables has a reasonable functional form. Hence, the proposed strategy is a powerful way to carry out testing using all available observations. Comparing the results of Tables 4 and 5 with Table 7, we see that the combined statistic in the manner of (2.18) has higher values for the estimated power as compared to the statistic based solely on X_{ij} without compromising the accuracy of the study.

4. Application to the pneumonia study

The oral health and ventilator-associate pneumonia study was an institutional randomized clinical trial to determine the effect of chlorhexidine gluconate application (once or twice per day) against the control group on the reduction of oral colonization by pathogens in subjects at an ICU (Scannapieco et al. 2009). In this study we combined the two active oral treatment groups (once per day and twice per day) since there was no dose response difference in terms of the pathogen colony counts, which was the primary outcome. The CPIS values were collected daily until subjects were discharged. If CPIS was greater than 5, BAL was performed. The median frequency o f BAL was 0 (range 0-6, mean=0.65). Eighty-four percent of subjects had no BAL or a single BAL. In the case for which there were multiple BAL's we used the first BAL value and the corresponding CPIS for the purpose of our analysis. If no BAL was performed during the hospital stay, the maximum CPIS is used for the analysis in order to evaluate a possibility of infection. A total of 175 evaluable subjects were accrued (116 subjects for the oral treatment group and 59 subjects for the control group). A total of 24 subjects in the control group and 44 subjects in the oral treatment group had BAL results. The median CPIS was 3 (range: 0~8). The histograms of CPIS scores for the two groups show that their distributions are close to symmetric around same point (Figure 1).

FIGURE 1 HERE

Since a constant shift of distributions in Figure 1 is not evident the Wilcoxon rank-sum test and the two-sample t-test are not sensitive to differences between distributions of CPIS related to potential scale differences between treatment and control. The Wilcoxon rank sum test, two-sample t-test and Kolmogorov-Smirnov test did not find a significant difference between the two groups. The p-values for this study were 0.8754, 0.8625, and 0.9817, respectively. The two-sample Anderson-Darling rank test (Pettitt, 1976) demonstrated the p-value of 0.053 utilizing the CPIS data points. The proposed test statistics is equal to 61.191 with corresponding p-value<0.05 indicating that the oral treatment (CHX) indeed altered the outcome distribution as compared with the control treatment. This is in contrast to case of using standard two-sample comparisons with respect to using CPIS scores only. This is in accordance with the fact that the new test has a much more power to detect the shape differences as demonstrated in Table 5. It should also be noted that entropy-based tests are powerful for detecting a change in scale as well, e.g., Dudewicz and Van Der Meulen (1981). Thus, the new test detected a difference between the distributions that was not detected using the classical procedures. The histograms in Figure 1 display that the distribution functions of CPIS related to the cases and controls have a difference in variances and a nonconstant shift for which the standard procedures break down to recognize adequately (Albers et al., 2001). In order to investigate this issue further we transformed the dataset using the following: $x_{ij} - \sum_{k=1}^{n_i} x_{ik} / n_i$. Under this transformation the classic Wilcoxon rank-sum test and Kolmogorov-Smirnov tests changed their p-values to be more favorable on H_1 , i.e. p-values of 0.074 and 0.003, respectively. In this case, the p-value of the two-sample Anderson-Darling rank test was found to be 0.045. In addition, for example, we depict values of two estimated confidence intervals for sample quantiles based on the centered CPIS datasets using the fact that the p^{th} sample quantile is asymptotically normally distributed. (The confidence interval estimation required to evaluate values of corresponding densities that were easily obtained based on the available computer package, 'density' function in *R*.) Table 8 shows the results of the 95%-confidence interval estimation. It is important to note that the transformation should not affect the accuracy of the test under the null hypothesis.

TABLE 8 HERE

Now, we attend to the combined test results. The median BAL value is 8×10^4 (range: $0 \sim 1.6\times10^8$). Because of the large disparities in BAL, we used the log(1+BAL)-transformation when applied to the final analysis. Also note that some BAL's were performed even if CPIS was less than or equal to 5 due to the fact that a BAL can be independently ordered under physician's discretion. Such irregularity is readily handled by the likelihood function at (2.2). Toward this end, we first construct the likelihood function based on the bivariate observations that adhere to the data collection protocol, and then the components of bivariate observations that did not follow the protocol are added into the likelihood function in the form of (2.1). The 95% critical value corresponding to the specified sample sizes for the data set is 53.960. The proposed likelihood ratio statistic given by 2log of the value of (2.18), provided the value of 62.030, which is above the critical value and corresponds to the p-value of 0.0047. The test result indicates the difference between the treatment groups is significant. Note that, Yu et al. (2010) tried to apply the standard EL technique to compare the cases and controls where their test statistics could not demonstrate a sufficient power to detect the nonconstant shift alternative. Vexler at al. (2009) remarked the classical EL schemes approximately converge to procedures similar to the two sample t-test that cannot detect nonconstant shift alternatives.

5. Concluding remarks

We proposed a two group comparison utilizing a new EL approach. The method was also implemented to incorporate incomplete bivariate observations based on semiparametric approach, which used both a nonparametric component for complete data and a parametric component for incomplete data. The density based EL approach ultimately derived the test statistic that consists of empirical distribution, which have a simple form that yields an efficient method. Since the proposed test statistic approximates the likelihood ratios of joint densities, the approach is closer to the original idea found in the Neyman-Pearson lemma of the likelihood ratio test as compared to the original EL based approach. The extensive Monte-Carlo study shows that our proposed approach provides an accurate and powerful test. Applying our method to an actual clinical trial data set demonstrated the immediate practicality of the proposed approach. The method developed here used a semi-parametric technique, which assumes a linear functional relationship between pairs of bivariate data, such as what was found in terms of the relationship between CPIS and BAL in our illustrative pneumonia study. The proposed method can be extended to incorporate more complicated structures or using a completely nonparametric approach similar to Vexler et al. (2010). These new methods will require a broad investigation based on finite sample sizes, and is the subject of further research.

Acknowledgements

The authors are indebted to Professors Alan D. Hutson, Govind S. Mudholkar and Calyampudi

R. Rao for many helpful discussions and comments. We would like to acknowledge Professor Bradley Efron for providing valuable commentaries. The authors also thank the Associated Editor and the reviewer for their insightful comments that have greatly helped us improve the article.

Appendix A

Here, we discuss the choice of the integer parameters (m, v) in (2.15) that leads to the EL ratio statistic, (2.16). The maximization of the likelihood function under H_1 is carried out for each group separately. For group 1, the likelihood function $L_{11} = \prod_{i=1}^{n_1} f_{X_1}(x_{1j})$ is maximized subject to (2.9). If for some k, we have $(2k)^{-1} \sum_{j=1}^{n} \int_{X_{(j-k)}}^{X_{(j+k)}} f(x)dx \geq 1$, then

$$1 \leq \frac{1}{2k} \sum_{j=1}^{n} \int_{X_{(j-k)}}^{X_{(j+k)}} f(x)dx \leq \int_{X_{(1)}}^{X_{(n)}} f(x)dx \leq \int_{-\infty}^{\infty} f(x)dx$$

by virtue of Proposition 2.1, which is inadmissible. Thus, we restrict values of f_{11j} to satisfy

$$G_k = \frac{1}{2k} \sum_{j=1}^{n_1} \frac{f_{11j}}{f_{01j}} \Delta_{k1j} \leq 1,$$

for all $k \in W$, where W represents a set of suitable values of the integer parameter *m* in (2.15) and G_k approximates $(2k)^{-1} \sum_{j=1}^{n} \int_{X_{(j-k)}}^{X_{(j+k)}} f(x)dx$. Then, we have an inequality,

$$\max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_k}\leq 1,\text{for all}\,k\in W} \prod_{j=1}^{n_1} f_{11j} \leq \max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_{k_0}}\leq 1, k_0\in W} \prod_{j=1}^{n_1} f_{11j}, \qquad \text{(A.1)}$$

for any integer $0 < k_0 \in W$. The inequality of (A.1) is satisfied since the left-hand side has more restrictions than the right-hand side. Then, (A.1) leads to

$$\min_{k_0\in W} \max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_k}\leq 1,\text{for all}\,k\in W} \prod_{j=1}^{n_1} f_{11j} \leq \min_{k_0\in W} \max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_{k_0}}\leq 1} \prod_{j=1}^{n_1} f_{11j},$$

$$\max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_k}\le 1,\,\text{for all}\,k\in W} \prod_{j=1}^{n_1} f_{11j} \le \min_{k_0\in W} \max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_{k_0}}\le 1} \prod_{j=1}^{n_1} f_{11j}. \quad \text{(A.2)}$$

Now, assume that for a fixed $0 < k_0 \in W$

$$G_{k_0} = \frac{1}{2k_0} \sum_{j=1}^{n_1} \frac{f_{11j}}{f_{01j}} \Delta_{k_0 1j} = 1.$$

By virtue of the approximation to the results of Proposition 2.1, we can expect

$$1 = G_{k_0} \approx \int_{x_{1(1)}}^{x_{1(n_1)}} f_{X_1}(x) \approx \frac{1}{2m} \sum_{j=1}^{n_1} \int_{x_{1(j-m)}}^{x_{1(j+m)}} f_{X_1}(x)dx \approx G_m,$$

i.e., if the condition (2.9) holds for a fixed $k_0 \in W$, this condition is expected to be satisfied for different $m \in W$ too. Thus

$$\max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_k}\le 1,\,\text{for all}\,k\in W} \prod_{j=1}^{n_1} f_{11j} \cong \max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_{k_0}}\le 1,\; k_0\in W} \prod_{j=1}^{n_1} f_{11j} \ge \min_{k_0\in W} \max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_{k_0}}\le 1} \prod_{j=1}^{n_1} f_{11j}.$$

This and (A.2) give rise to

$$\max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_k}\le 1\,\text{for all}\,k\in W} \prod_{j=1}^{n_1} f_{11j} \cong \min_{k_0\in W} \max_{f_{111},\ldots,f_{11n_1}:\mathrm{G_{k_0}}\le 1} \prod_{j=1}^{n_1} f_{11j}.$$

Similar approach is applied for group 2.

To obtain (2.16), we utilized a constraint based on (2.6). The reminder term of the use of (2.6) is

$$-\sum_{l=1}^{m-1}\frac{(m-l)}{2m}\int_{X_{(n-l)}}^{X_{(n-l+1)}} f_X(u)du-\sum_{l=1}^{m-1}\frac{(m-l)}{2m}\int_{X_{(l)}}^{X_{(l+1)}} f_X(u)du=-\sum_{l=1}^{m-1}\frac{(m-l)}{2m}\left(\int_{X_{(n-l)}}^{X_{(n-l+1)}} f_X(u)du+\int_{X_{(l)}}^{X_{(l+1)}} f_X(u)du\right)$$

$$=-\sum_{l=1}^{m-1}\frac{(m-l)}{2m}\left(F_X(X_{(n-l+1)})-F_X(X_{(n-l)})+F_X(X_{(l+1)})-F_X(X_{(l)})\right)$$

$$\approx-\sum_{l=1}^{m-1}\frac{(m-l)}{2m}\left(F_{Xn}(X_{(n-l+1)})-F_{Xn}(X_{(n-l)})+F_{Xn}(X_{(l+1)})-F_{Xn}(X_{(l)})\right)=\frac{(m-1)}{2n},$$

$$F_X(u)=\Pr(X_1<u), F_{Xn}(u)=\frac{1}{n}\sum_{i=1}^{n} I(X_i\le u).$$

Regarding the constraint (2.9), we minimize the influence of the remainder term above, requiring $(m-1)/n_1 \to 0$ as $n_1 \to \infty$, e.g., $W=\{m: m\le n_1^{1-\delta}\}, 1>\delta>0.$ Proposition 2.2 presents the consistency of the proposed test statistic. To prove this proposition, we utilize the lower bound $m \ge n_1^{0.5+\delta}$. Thus, we specify the set $W=\{m: n_1^{0.5+\delta}\le m\le n_1^{1-\delta}\},\ 0<\delta<0.25$ in the definition (2.16). Similar approach is applied for group 2.

Appendix B

Proof of Proposition 2.2

Consider the term from (2.15) that has the form of

$$\log\left(R^1_{mn_1n_2}\right)=\log\prod_{i=1}^{n_1}\frac{2m}{n_1\Delta_{m1i}}=-\sum_{i=1}^{n_1}\log\frac{\Delta_{m1i}}{2m/n_1}. \qquad \text{(B.1)}$$

Note that the definition (2.10) of Δ_{m1i} utilizes the function

$$\frac{1}{n_1+n_2}\left(n_1F_{x_1n_1}(u)+n_2F_{x_2n_2}(u)\right),$$

where $F_{x_ln_l}(u)=n_l^{-1}\sum_{i=1}^{n_l} I(x_{li}\le u), l=1,2$ are the empirical distribution functions.

Rewrite (B.1) as

$$\log\left(R^1_{mn_1n_2}\right) = -\sum_{i=1}^{n_1}\log\frac{F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)}{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)} \tag{B.2}$$

$$+\sum_{i=1}^{n_1}\log\frac{F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)}{\Delta_{m1i}} - \sum_{i=1}^{n_1}\log\frac{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)}{2m/n_1}.$$

where $F^*_{n_1n_2}(u) = (1/(n_1+n_2))\left(n_1F_{x_1}(x)+n_2F_{x_2}(x)\right)$, $F_{x_l}(x), l=1,2$ represent the distribution functions of x_{11} and x_{21}, respectively. We consider the first term in the right side of (B.2)

$$\sum_{i=1}^{n_1}\log\frac{F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)}{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)} = \sum_{i=1}^{n_1}\log\frac{F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)}{x_{1(i+m)}-x_{1(i-m)}} \tag{B.3}$$

$$-\sum_{i=1}^{n_1}\log\frac{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)}{x_{1(i+m)}-x_{1(i-m)}}.$$

Now we will use the proof scheme of Theorem 1 presented by Vasicek (1976) applying some reorganization

$$\frac{1}{n_1+n_2}\sum_{i=1}^{n_1}\log\frac{F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)}{x_{1(i+m)}-x_{1(i-m)}} = \frac{n_1}{n_1+n_2}(2m)^{-1}\sum_{j=1}^{2m}S_j\ , \tag{B.4}$$

where

$$S_j = \sum_{i=1}^{n_1}\log\frac{F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)}{x_{1(i+m)}-x_{1(i-m)}}\left\{F_{x_1n_1}\left(x_{1(i+m)}\right)-F_{x_1n_1}\left(x_{1(i-m)}\right)\right\},\quad i\equiv j \text{ (mod 2m)}\ .$$

Suppose $x_{1(i-m)}$ and $x_{2(i+m)}$ belong to an interval in which

$$f^*_{n_1n_2}(u) = \frac{dF^*_{n_1n_2}(u)}{du} = \frac{n_1}{n_1+n_2}f_{X_1}(u) + \frac{n_2}{n_1+n_2}f_{X_2}(u) > 0$$

and is a continuous function. Then, we can find a point $X_i^* \in \left(x_{1(i-m)}, x_{2(i+m)}\right)$ that satisfies

$$\frac{F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)}{x_{1(i+m)}-x_{1(i-m)}} = f^*_{n_1n_2}\left(X_i^*\right),$$

i.e. we can write

$$S_j = \sum_{i=1}^{n_1} \log\ f^*_{n_1n_2}\left(X_i^*\right)\left\{F_{x_1n_1}\left(x_{1(i+m)}\right)-F_{x_1n_1}\left(x_{1(i-m)}\right)\right\}, \quad i \equiv j \text{ (mod 2m)} .$$

The function $f^*_{n_1n_2}(u)$ approximates the density function

$$f^*(x)=\frac{\eta}{1+\eta} f_{X_1}(x)+\frac{1}{1+\eta} f_{X_2}(x).$$

The condition $n_1/n_2 \to \eta$ implies the inequality $(1-\varepsilon)f^*\left(X_i^*\right)\le f^*_{n_1n_2}\left(X_i^*\right)\le(1+\varepsilon)f^*\left(X_i^*\right)$, for all $\varepsilon>0$ and sufficiently large $n_l, l=1,2$. Then,

$$S_j^{-\varepsilon} \le S_j \le S_j^{\varepsilon},\ S_j^{\varepsilon} = \sum_{i=1}^{n_1} \log\left((1+\varepsilon)f^*\left(X_i^*\right)\right)\left\{F_{x_1n_1}\left(x_{1(i+m)}\right)-F_{x_1n_1}\left(x_{1(i-m)}\right)\right\}, \quad i \equiv j \text{ (mod 2m)} ,$$

for sufficiently large $n_l, l=1,2$. That is, S_j^{ε} is a Stieltjes sum of the function $\log\left((1+\varepsilon)f^*(u)\right)$ with respect to the measure $F_{x_1n_1}$ over the sum of intervals of continuity of $f_{X_1}(u)$ and $f_{X_2}(u)$ in which $f^*(u)>0$.

Because $x_{1(i+m)}-x_{1(i-m)}\to 0$ a.s. uniformly over $m\in\left[n_1^{0.5+\delta}, n_1^{1-\delta}\right]$ and $F_{x_1n_1}(u)\to F_{x_1}(u)$ uniformly over u as $n_1\to\infty$, the arguments mentioned in the proof of Theorem 1 of Vasicek (1976, p. 56) display that S_j^{ε} converges in probability to

$$\int_{-\infty}^{\infty} \log\left((1+\varepsilon)\, f^*(x)\right) dF_{x_1}(x) = E\left(\log\left((1+\varepsilon)\, f^*\left(X_{11}\right)\right)\right)$$

as $n_1\to\infty$, uniformly over $m\in\left[n_1^{0.5+\delta}, n_1^{1-\delta}\right]$ and over j. Therefore, we conclude about

$$E\left(\log\left((1-\varepsilon)\, f^*\left(X_{11}\right)\right)\right)\le(2m)^{-1}\sum_{j=1}^{2m} S_j \le E\left(\log\left((1+\varepsilon)\, f^*\left(X_{11}\right)\right)\right),$$

as $n_1 \to \infty$ uniformly over $m \in \left[n_1^{0.5+\delta}, n_1^{1-\delta} \right]$, and, by (B.4),

$$\frac{1}{n_1+n_2}\sum_{i=1}^{n_1}\log\frac{F_{nk}^*\left(x_{1(i+m)}\right)-F_{nk}^*\left(x_{1(i-m)}\right)}{x_{1(i+m)}-x_{1(i-m)}}\xrightarrow{P}\frac{\eta}{1+\eta}E\left(\log f^*\left(X_{11}\right)\right) \qquad \text{(B.5)}$$

as $n_l \to \infty$, $l=1,2$, $n_1/n_2 \to \eta > 0$, uniformly over $m \in \left[n_1^{0.5+\delta}, n_1^{1-\delta} \right]$.

Utilizing a proof similar to that presented above, one can easily show

$$\frac{1}{n_1+n_2}\sum_{i=1}^{n_1}\log\frac{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)}{x_{1(i+m)}-x_{1(i-m)}}\xrightarrow{P}\frac{\eta}{1+\eta}E\left(\log f_{X_1}\left(X_{11}\right)\right) \qquad \text{(B.6)}$$

as $n_l \to \infty, l=1,2$, $n_1/n_2 \to \eta > 0$, uniformly over $m \in \left[n_1^{0.5+\delta}, n_1^{1-\delta} \right]$. The outputs (B.5), (B.6) applied to (B.3) provide

$$\frac{1}{n_1+n_2}\sum_{i=1}^{n_1}\log\frac{F_{n_1n_2}^*\left(x_{1(i+m)}\right)-F_{n_1n_2}^*\left(x_{1(i-m)}\right)}{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)}\xrightarrow{P}\frac{\eta}{1+\eta}E\left(\log f^*\left(X_{11}\right)\right)-\frac{\eta}{1+\eta}E\left(\log f_{X_1}\left(X_{11}\right)\right) \qquad \text{(B.7)}$$

$$=\frac{\eta}{1+\eta}E\left(\log\left(\frac{\eta}{1+\eta}+\frac{1}{1+\eta}\frac{f_{X_2}\left(X_{11}\right)}{f_{X_1}\left(X_{11}\right)}\right)\right).$$

as $n_l \to \infty, l=1,2$, $n_1/n_2 \to \eta > 0$, uniformly over $m \in \left[n_1^{0.5+\delta}, n_1^{1-\delta} \right]$.

Now we considering the term $\sum_{i=1}^{n_1}\log\left(F_{n_1n_2}^*\left(x_{1(i+m)}\right)-F_{n_1n_2}^*\left(x_{1(i-m)}\right)\right)\left(\Delta_{m1i}\right)^{-1}$ of (B.2). By virtue of the Theorem of Kolmogorov (e.g., Serfling, 1980, p. 62), we have

$\Pr\left(\sup_{-\infty<u<\infty}\left|F_{x_1}(u)-F_{x_1n_1}(u)\right|>n_1^{-0.5+\varepsilon}\right)\to 0$ and $\Pr\left(\sup_{-\infty<u<\infty}\left|F_{x_2}(u)-F_{x_2n_2}(u)\right|>n_2^{-0.5+\varepsilon}\right)\to 0$ as $n_l \to \infty, l=1,2$, for each $0<\varepsilon<\delta/4$. Thus $\Pr\left(\sup_{-\infty<u<\infty}\left|F_{x_2}(u)-F_{x_2n_2}(u)\right|>(2n_1/\eta)^{-0.5+\varepsilon}\right)\to 0$ and

$$\Pr\left(\sup_{-\infty<u<\infty}\left|F_{n_1n_2}^*(u)-\frac{1}{n_1+n_2}\left(n_1F_{x_1n_1}(u)+n_2F_{x_2n_2}(u)\right)\right|>n^{-0.5+2\varepsilon}\right)\to 0,$$

$n_l \to \infty, l=1,2$, $n_1/n_2 \to \eta > 0$. Therefore, we can focus on cases when

$$\sup_{-\infty<u<\infty}\left|F^*_{n_1n_2}(u)-\frac{1}{n_1+n_2}\left(n_1F_{x_1n_1}(u)+n_2F_{x_2n_2}(u)\right)\right| \le n^{-0.5+2\varepsilon}.$$

This and the inequality $\Delta_{m1i} \ge 2m/(n_1+n_2)$ obtained by virtue of the definition (2.10), lead to the following results

$$\frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\log\frac{\left(F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)\right)}{\Delta_{m1i}} \le \frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\log\frac{\Delta_{m1i}+n_1^{-0.5+\delta/2}}{\Delta_{m1i}}$$
$$\le \frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\log\left(1+\frac{n_1^{-0.5+\delta/2}}{2m/(n_1+n_2)}\right) \le \frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\frac{n_1^{-0.5+\delta/2}}{2n_1^{0.5+\delta}/(n_1+n_2)} \underset{n_1\to\infty}{\to} 0, \text{ since } m \ge n_1^{0.5+\delta};$$

$$\frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\log\frac{\left(F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)\right)}{\Delta_{m1i}} \ge \frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\log\frac{\Delta_{m1i}-n_1^{-0.5+\delta/2}}{\Delta_{m1i}}$$
$$\ge \frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\log\left(1-\frac{n_1^{-0.5+\delta/2}}{2m/(n_1+n_1)}\right) \ge -\frac{1}{(n_1+n_2)}\sum_{i=1}^{n}\frac{2n_1^{-0.5+\delta/2}}{2n_1^{0.5+\delta}/(n_1+n_1)} \underset{n_1\to\infty}{\to} 0.$$

That concludes

$$\frac{1}{(n_1+n_2)}\sum_{i=1}^{n_1}\log\frac{\left(F^*_{n_1n_2}\left(x_{1(i+m)}\right)-F^*_{n_1n_2}\left(x_{1(i-m)}\right)\right)}{\Delta_{m1i}} \xrightarrow[n_1,n_2\to\infty]{P} 0 \qquad \text{(B.8)}$$

uniformly over $n_1^{0.5+\delta} \le m \le n_1^{1-\delta}$.

Now, we consider the part $-\sum_{i=1}^{n_1}\log\frac{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)}{2m/n_1}$ of (B.2). The proof scheme of Lemma 1 presented in Vasicek (1976, p. 55) shows directly

$$0 \le -\sum_{i=1}^{n_1}\log\frac{F_{x_1}\left(x_{1(i+m)}\right)-F_{x_1}\left(x_{1(i-m)}\right)}{2m/n_1} \xrightarrow[n_1,n_2\to\infty]{P} 0, \qquad \text{(B.9)}$$

uniformly over $m \in \left[n_1^{0.5+\delta}, n_1^{1-\delta}\right]$. Utilizing (B.1), (B.2), (B.7)-(B.9), one can show

$$\log\left(R^1_{mn_1n_2}\right) \xrightarrow{P} -\frac{\eta}{1+\eta}E\left(\log\left(\frac{\eta}{1+\eta}+\frac{1}{1+\eta}\frac{f_{X_2}(X_{11})}{f_{X_1}(X_{11})}\right)\right)$$

uniformly over $m \in \left[n_1^{0.5+\delta}, n_1^{1-\delta} \right]$ as $n_1, n_2 \to \infty$, $n_1 / n_2 \to \eta > 0$. The proof scheme above can be applied to analyze the part

$$\log\left(R^2_{vn_1n_2}\right) = \log \prod_{i=1}^{n_2} \frac{2v}{n_2 \Delta_{v2i}} = -\sum_{i=1}^{n_2} \log \frac{\Delta_{v2i}}{2v / n_2}$$

of (2.15) in order to present the asymptotic result

$$\frac{1}{n_1 + n_2} \log\left(R^2_{vn_1n_2}\right) \underset{n_1, n_2 \to \infty}{\overset{p}{\to}} -\frac{1}{1+\eta} E\left(\log\left(\frac{1}{1+\eta} + \frac{\eta}{1+\eta} \frac{f_{X_1}\left(X_{21}\right)}{f_{X_2}\left(X_{21}\right)} \right) \right)$$

uniformly over $n^{0.5+\delta} \le m \le n^{1-\delta}$. This completes the proof of Proposition 2.2.

Reference

Albers, W., Kallenberg, W. C. M., and Martini, F. (2001). Data-driven rank tests for classes of tail alternatives. *Journal of the American Statistical Association*, 96, 685-696.

Canner, P. L. (1975). A simulation study of one-and two-sample Kolmogorov-Smirnov statistics with a particular weight function. *Journal of the American Statistical Association* 70, 209-211.

Dedewicz, E. J. and Van Der Meulen, E. C. (1981). Entropy-based tests of uniformity. *Journal of the American Statistical Association* 76, 967–974.

Dupont, H., Chalhoub, V., Planteféve, G., Vaumas, C., Kermarrec, N., Paugam-Burtz, C., and Mantz, J.(2004). Variation of infected cell count in bronchoalveolar lavage and timing of ventilator-associated pneumonia. *Journal of Intensive Care Medicine* **30**, 1557–1563.

Hall, P. and Welsh, A. H. (1983). A test for normality based on the empirical characteristic function. *Biometrika* 70, 485-489.

Huh, J. W., Lim, C. M., Koh, Y., Oh, Y. M., Shim, T. S., Lee, S. D., Kim, W. S., Kim, D. S., Kim, W. D. and Hong, S. B. (2008). Diagnostic utility of the soluble triggering receptor expressed on myeloid cells-1 in bronchoalveolar lavage fluid from patients with bilateral lung infiltrates. *Critical Care* **12**, R6.

Koenig, S. M. and Truwit, J. D. (2006). Ventilator-associated pneumonia: diagnosis, treatment, and prevention. *Clinical Microbiology Reviews* **19**, 637–657.

Lazar, N. and Mykland, P.A. (1998). An evaluation of the power and conditionality properties of empirical likelihood. *Biometrika* **85**, 523–534;

Lehmann, E. L and Romano, J. P. (2005). *Testing Statistical Hypotheses.* Third Edition,

Springer,

Mudholkar, G. S. and Tian, L. (2002). An entropy characterization of the inverse Gaussian distribution and related goodness-of-fit test. *Journal of Statistical Planning and Inference* 102, 211–221.

Mudholkar, G. S. and Tian, L. (2004). A test for homogeneity of ordered means of inverse Gaussian populations. *Journal of Statistical Planning and Inference* 118, 37–49.

Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–249.

Owen, A. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics* **18**, 90–120.

Owen, A. (1991). Empirical likelihood for linear models. *The Annals of Statistics* **19**, 1725–1747.

Owen, A. (2001). *Empirical Likelihood.* New York, Chapman & Hall.

Pettitt, A. N. (1976). A two-sample Anderson-Darling rank statistic. Biometrika **63,** 161-168

Qin, J., and Lawless, J. (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics* **22**, 300–325.

Qin, J. (2000). Combining parametric and empirical likelihoods. *Biometrika* 87, 484–490.

Qin, J., Leung, D. (2005). A semi-parametric two component "compound" mixture model and its application to estimating malaria attributable fractions. *Biometrics* **61**, 456–464.

Ramirez, P., Garcia, M. A., Ferrer, M., Aznar, J., Valencia, M., Sahuquillo, J. M., Menéndez, R., Asenjo, M. A. and Torres, A. (2007). Sequential measurements of procalcitonin levels in diagnosing ventilator-associated pneumonia. *The European Respiratory Journal* **31**, 356-362.

Scannapieco, F. A. ,Yu, J., Raghavendran, K., Vacanti, A., Owens, S. I., Wood, K., and Mylotte, J. M. (2009). A randomized trial of chlorhexidine gluconate on oral bacterial pathogens in mechanically ventilated patients. *Critical care* **13**, R117.

Swoboda, S. M., Dixon, T. and Lipsett, P. A. (2006).Can the clinical pulmonary infection score impact ICU antibiotic days? *Surgical Infections* **7**, 331-339.

Tusnady, G. (1977). On Asymptotically Optimal Tests. *The Annals of Statistics* **5**, 385–393.

Van Es, B. (1992). Estimating Functionals Related to a Density by a Class of Statistics Based on Spacings. *Scandinavian Journal of Statistics* 19, 61–72.

Vasicek, O. (1976) A Test for Normality Based on Sample Entropy. *Journal of the Royal Statistical Society, Series B* 38, 54–59.

Vexler, A. and Gurevich, G. (2010). Empirical likelihood ratios applied to goodness-of-fit tests based on sample entropy. *Computational Statistics & Data Analysis* 54, 531–545.

Vexler, A., Liu, S., Kang, L. and Hutson, A. D. (2009). Modifications of the Empirical Likelihood Interval Estimation with Improved Coverage Probabilities. *Communications in Statistics (Simulation and Computation)* 38, 2171–2183.

Vexler A. and Wu, C. (2009). An Optimal Retrospective Change Point Detection Policy. *Scandinavian Journal of Statistics* 36, 542–558.

Vexler A., Wu C., and Yu, K. F. (2010). Optimal hypothesis testing: from semi to fully Bayes factors. *Metrika* 71, 125–138.

Vexler, A., Yu, J., Tian, L., and Liu, S. (2010). Two-sample nonparametric likelihood inference based on incomplete data with an application to a pneumonia study. *Biometrical Journal* 52, 348–361.

Vexler, A., Shan, G., Kim, S., Tsai, W-M., Tian, L. and Hutson, A. D. (2011). An empirical likelihood ratio based goodness-of-fit test for Inverse Gaussian distributions. *Journal of Statistical Planning and Inference.* In Press. doi:10.1016/j.jspi.2010.12.024

Yu, J., Vexler, A., and Tian, L. (2010). Analyzing Incomplete Data Subject to a Threshold using Empirical Likelihood Methods: An Application to a Pneumonia Risk Study in an ICU Setting. *Biometrics* **66**, 123–130.

Table 2. *The critical values* C_α *for the combined two-sample test with the test statistic 2log(R), where R is defined in (2.18), for various sample sizes* n_1, n_2. *Thus,* $\Pr\{2\log(R) > C_\alpha\} = \alpha$ *, under the null hypothesis.*

n_1	α	n_2								
		10	15	20	25	30	35	40	50	60

10	0.01	30.695								
	0.03	27.697								
	0.05	25.621								
	0.1	23.198								
15	0.01	32.014	33.317							
	0.03	28.999	30.291							
	0.05	26.956	28.272							
	0.1	24.535	25.894							
20	0.01	33.342	34.724	35.991						
	0.03	30.389	31.683	33.018						
	0.05	28.383	29.647	30.952						
	0.1	25.951	27.263	28.551						
25	0.01	34.690	35.966	37.256	38.599					
	0.03	31.684	32.926	34.221	35.551					
	0.05	29.763	30.913	32.209	33.551					
	0.1	27.308	28.528	29.850	31.183					
30	0.01	35.989	37.312	38.537	39.791	41.123				
	0.03	32.984	34.226	35.518	36.750	38.038				
	0.05	31.100	32.255	33.477	34.780	36.040				
	0.1	28.604	29.828	31.095	32.430	33.691				
35	0.01	36.153	37.453	38.624	39.968	41.271	41.520			
	0.03	33.089	34.382	35.639	36.921	38.194	38.403			
	0.05	31.260	32.449	33.658	34.938	36.240	36.499			
	0.1	28.760	29.994	31.273	32.589	33.882	34.099			
40	0.01	37.509	38.738	40.018	41.381	42.535	42.716	43.976		
	0.03	34.399	35.654	36.942	38.284	39.444	39.651	40.920		
	0.05	32.537	33.715	34.973	36.302	37.453	37.724	38.929		
	0.1	30.011	31.265	32.547	33.886	35.111	35.326	36.553		
50	0.01	38.923	40.139	41.383	42.684	43.965	44.108	45.425	46.799	
	0.03	35.771	37.015	38.259	39.577	40.836	41.066	42.343	43.687	
	0.05	33.968	35.142	36.345	37.668	38.892	39.144	40.383	41.776	
	0.1	31.412	32.613	33.883	35.240	36.495	36.711	37.990	39.379	
60	0.01	41.349	42.527	43.821	45.188	46.430	46.605	47.801	49.346	51.678
	0.03	38.232	39.453	40.760	42.047	43.281	43.495	44.700	46.178	48.564
	0.05	36.452	37.580	38.844	40.134	41.368	41.577	42.781	44.273	46.634
	0.1	33.889	35.075	36.368	37.692	38.949	39.162	40.417	41.873	44.258
70	0.01	42.649	43.946	45.136	46.505	47.675	47.933	49.260	50.576	53.030
	0.03	39.503	40.816	42.072	43.360	44.548	44.835	46.061	47.451	49.916
	0.05	37.696	38.974	40.183	41.467	42.655	42.928	44.176	45.550	48.022
	0.1	35.149	36.411	37.700	38.993	40.241	40.519	41.791	43.161	45.621
80	0.01	43.921	45.191	46.516	47.761	49.125	49.267	50.502	51.836	54.325
	0.03	40.751	42.062	43.406	44.659	45.938	46.117	47.371	48.812	51.215
	0.05	38.989	40.227	41.478	42.779	44.053	44.246	45.480	46.897	49.318

	0.1	36.418	37.664	38.975	40.301	41.613	41.811	43.084	44.492	46.929
90	0.01	45.235	46.529	47.792	49.118	49.125	50.553	51.821	53.210	55.652
	0.03	41.982	43.346	44.641	45.970	45.938	47.370	48.663	50.059	52.522
	0.05	40.204	41.544	42.802	44.100	44.053	45.515	46.812	48.191	50.667
	0.1	37.657	38.936	40.284	41.592	41.613	43.085	44.379	45.814	48.258
100	0.01	46.550	47.753	49.100	50.369	51.622	51.794	53.069	54.513	56.869
	0.03	43.272	44.574	45.931	47.208	48.475	48.614	49.921	51.369	53.741
	0.05	41.522	42.770	44.097	45.347	46.620	46.766	48.070	49.522	51.875
	0.1	38.943	40.226	41.532	42.824	44.138	44.370	45.620	47.077	49.491
110	0.01	47.699	49.074	50.296	51.706	52.984	53.150	54.332	55.673	58.203
	0.03	44.539	45.885	47.133	48.478	49.767	49.952	51.188	52.553	55.061
	0.05	42.822	44.090	45.349	46.633	47.905	48.089	49.337	50.702	53.209
	0.1	40.264	41.520	42.792	44.118	45.417	45.614	46.893	48.318	50.811
130	0.01	50.199	51.475	52.851	54.171	55.423	55.668	56.910	58.319	60.723
	0.03	46.942	48.335	49.705	50.941	52.226	52.474	53.719	55.119	57.519
	0.05	45.221	46.544	47.886	49.087	50.394	50.638	51.878	53.310	55.712
	0.1	42.631	43.972	45.300	46.572	47.896	48.181	49.409	50.884	53.292
150	0.01	51.614	52.836	54.268	55.490	56.808	57.021	58.281	59.702	62.141
	0.03	48.266	49.645	51.002	52.329	53.559	53.816	55.018	56.489	58.950
	0.05	46.550	47.893	49.172	50.491	51.740	51.988	53.189	54.660	57.174
	0.1	43.969	45.299	46.621	47.926	49.243	49.524	50.770	52.234	54.746
200	0.01	56.594	57.752	59.211	60.456	61.827	62.074	63.329	64.730	67.309
	0.03	53.237	54.537	56.011	57.265	58.553	58.823	60.084	61.514	64.050
	0.05	51.516	52.781	54.229	55.474	56.761	56.992	58.270	59.716	62.248
	0.1	48.955	50.184	51.625	52.928	54.213	54.493	55.775	57.235	59.811

Table 3. *Continued from Table 2. The critical values for the combined test with the test statistic 2log(R).*

n_1	α	n_2							
		70	80	90	100	110	130	150	200
70	0.01	54.373							
	0.03	51.258							
	0.05	49.334							
	0.1	46.945							
80	0.01	55.661	56.950						
	0.03	52.581	53.834						
	0.05	50.695	51.981						
	0.1	48.292	49.615						
90	0.01	56.970	58.338	59.671					
	0.03	53.859	55.217	56.477					
	0.05	51.998	53.338	54.678					
	0.1	49.621	50.944	52.271					

100	0.01	58.240	59.561	60.969	62.262				
	0.03	55.138	56.450	57.818	59.089				
	0.05	53.274	54.617	55.965	57.272				
	0.1	50.873	52.232	53.574	54.857				
110	0.01	59.553	60.879	62.248	63.511	64.856			
	0.03	56.382	57.752	59.079	60.370	61.725			
	0.05	54.568	55.881	57.254	58.547	59.907			
	0.1	52.168	53.513	54.857	56.130	57.499			
130	0.01	62.142	63.459	64.764	66.142	67.435	70.038		
	0.03	58.965	60.329	61.605	62.948	64.251	66.841		
	0.05	57.156	58.511	59.785	61.149	62.452	65.041		
	0.1	54.736	56.083	57.389	58.718	60.037	62.611		
150	0.01	63.500	64.915	66.191	67.594	68.868	71.415	72.808	
	0.03	60.316	61.707	63.006	64.358	65.642	68.193	69.639	
	0.05	58.523	59.935	61.188	62.598	63.866	66.418	67.851	
	0.1	56.103	57.495	58.787	60.158	61.480	64.014	65.459	
200	0.01	68.725	70.005	71.303	72.649	73.953	76.641	78.148	83.286
	0.03	65.413	66.728	68.104	69.407	70.752	73.358	74.891	80.001
	0.05	63.636	64.963	66.325	67.646	69.008	71.617	73.128	78.247
	0.1	61.183	62.544	63.903	65.225	66.591	69.185	70.680	75.810

Table 4. *The Monte Carlo rejection probabilities of the test statistic (2.16) for the two sample comparison, where* $\{X_{1j} \sim N(0,\sigma_1^2),\ j=1,...,n_1\}$ *and* $\{X_{2j} \sim N(\mu,\sigma_2^2),\ j=1,...,n_2\}$. *The case of* $X_1 \sim N(0,1)$ *and* $X_2 \sim N(0,1)$ *corresponds to the Type I error with the significance level 0.05 test (the 4*th *column). The rest of cases are the powers at the significance level 0.05 corresponding to parameter differences for each cell in the table. The Monte Carlo study was based on 10,000 generations of samples for each case.*

σ_1, σ_2	n_1	n_2	μ				
			0	0.1	0.2	0.35	0.5
1, 1	20	20	0.053	0.058	0.082	0.153	0.261
	25	50	0.050	0.067	0.094	0.213	0.387
	50	50	0.046	0.065	0.128	0.301	0.563
	100	100	0.046	0.082	0.204	0.516	0.843
1, 1.5	20	20	0.184	0.183	0.209	0.255	0.338
	25	50	0.434	0.451	0.474	0.538	0.628
	50	50	0.528	0.539	0.580	0.659	0.767
	100	100	0.868	0.874	0.900	0.941	0.975
1, 2	20	20	0.495	0.499	0.508	0.535	0.585
	25	50	0.893	0.886	0.896	0.907	0.924
	50	50	0.962	0.960	0.963	0.973	0.979
	100	100	1.000	1.000	1.000	1.000	1.000

Table 5. *The Monte Carlo rejection probabilities of the test statistic (2.16) for the two sample comparison, where* $\{X_{1j} \sim Lognormal(0,\sigma_1^2),\ j=1,...,n_1\}$ *and* $\{X_{2j} \sim Lognormal(\mu,\sigma_2^2),\ j=1,...,n_2\}$. *The case of* $X_1 \sim Lognormal(0,\sigma_1^2)$ *and* $X_2 \sim Lognormal(0,\sigma_1^2)$ *corresponds to the Type I error with the significance level 0.05 test. The rest of cases are the powers at the significance level 0.05 corresponding to parameter differences for each cell in the table. The Monte Carlo study was based on 10,000 generations of samples for each case.*

σ_1,σ_2	n_1	n_2	μ				
			0	0.1	0.2	0.35	0.5
1, 1	20	20	0.048	0.085	0.185	0.494	0.792
	25	50	0.051	0.082	0.263	0.698	0.937
	50	50	0.048	0.144	0.481	0.935	0.998
	100	100	0.052	0.260	0.814	0.999	1.000
1, 2	20	20	0.646	0.468	0.293	0.126	0.119
	25	50	0.963	0.857	0.648	0.289	0.213
	50	50	0.992	0.946	0.772	0.333	0.265
	100	100	1.000	1.000	0.983	0.617	0.533

Table 6. *The Monte Carlo powers (at the significance level 0.05) of the following tests: (2.16) with different values of δ, Kolmogorov Smirnov (KS) test, Wilcoxon rank sum test, and t-test. For each sample sizes* (n_1,n_2) ***,*** *the cases 1)-5) display simulation studies based on samples from*
1) $X_1 \sim N(0,1), X_2 \sim Unif[-1,1]$*;* 2) $X_1 \sim Exp(1), X_2 \sim LogNorm(0,1)$ *;*
3) $X_1 \sim N(0,1), X_2 \sim N(1.5,1)$; 4) $X_1 \sim N(0,1), X_2 \sim N(0,1.5^2)$;
5) $X_1 \sim Beta(0.7,1), X_2 \sim Exp(2)$. *The Monte Carlo study was based on 10,000 generations of samples for each case.*

Design	n_1	n_2	Proposed test						KS	Wilcoxon	t
			δ								
			0.025	0.05	0.1	0.12	0.15	0.2			
1)											
	45	45	0.953	0.949	0.957	0.949	0.953	0.958	0.146	0.056	0.053
	15	25	0.285	0.314	0.278	0.273	0.261	0.258	0.090	0.063	0.050
	25	15	0.407	0.367	0.372	0.348	0.353	0.342	0.062	0.043	0.049
	15	15	0.239	0.239	0.215	0.224	0.202	0.196	0.035	0.044	0.047
2)											
	45	45	0.587	0.587	0.596	0.583	0.585	0.579	0.339	0.494	0.481
	10	10	0.132	0.132	0.132	0.126	0.129	0.131	0.118	0.126	0.079
3)											
	25	15	0.976	0.978	0.982	0.978	0.981	0.980	0.993	0.991	0.994
	15	15	0.937	0.939	0.947	0.951	0.953	0.952	0.885	0.970	0.979
4)											
	45	45	0.503	0.499	0.488	0.453	0.443	0.414	0.129	0.052	0.050
	35	35	0.383	0.389	0.355	0.350	0.338	0.305	0.073	0.052	0.049
5)											
	45	45	0.511	0.470	0.450	0.445	0.444	0.446	0.063	0.058	0.132

Table 7. *The Monte Carlo rejection probabilities (at the significance level 0.05) based on the combined test statistic 2log(R), where R is defined in (2.18), for various sample sizes* n_1, n_2 *. The selected parameters correspond to the model (2.17) where the subscript* i *indicates group* i *. The results in columns entailed as Power-I and Power-II are based on the normal and lognormal distributions, respectively. Threshold 'No' indicates that the simulation was carried out without thresholds. The thresholds for the normal and lognormal distributions are equal to 0 and the median of* $\{X_{1j}, j = 1, ..., n_1\}$ *, respectively. For each sample size and Power-I (normal distribution), row 9 depicts the simulation results based on the parameters from the pneumonia data (* $\sigma_{X_1} = 1.8$ *,* $\sigma_{X_2} = 1.9$ *,* $\sigma_1 = 4.8, \sigma_2 = 5.4$ *) and row 10 depicts the cases based on the smaller variances (* $\sigma_{X_1} = 0.15$ *,* $\sigma_{X_2} = 0.15, \sigma_1 = 1, \sigma_2 = 1$ *). For the lognormal distribution (Power-II), appropriate parameters were used to achieve the semblance of the simulations based on the normal distribution's parameters. The Monte Carlo study was based on 10,000 generations of samples for each case.*

n_1	n_2	Threshold	μ_1, μ_2	$\alpha_1, \alpha_2, \beta_1, \beta_2$	Power-I	Power-II
50	50	No	0, 0	1,2,2,2	0.039	0.035
		No	0,0.1	1,1,2,2.1	0.089	0.126
		No	0,0.2	1,1.2,2,2.2	0.565	0.751
		No	0,0.3	1,1.3,2,2.3	0.947	0.992
		Yes	0,0	1,1,2,2	0.039	0.043
		Yes	0,0.1	1,1,2,2.1	0.070	0.128
		Yes	0,0.2	1,1.2,2,2.2	0.526	0.732
		Yes	0,0.3	1,1.3,2,2.3	0.922	0.988
		Yes	4.98,5.03	9.69,11.41,0.12,-0.06	0.094	0.067
		Yes	4.98,5.03	9.69,11.41,0.12,-0.06	0.905	0.683
50	100	No	0, 0	1,2,2,2	0.037	0.036
		No	0,0.1	1,1,2,2.1	0.096	0.151
		No	0,0.2	1,1.2,2,2.2	0.717	0.869
		No	0,0.3	1,1.3,2,2.3	0.987	0.999
		Yes	0,0	1,1,2,2	0.040	0.039
		Yes	0,0.1	1,1,2,2.1	0.073	0.144
		Yes	0,0.2	1,1.2,2,2.2	0.662	0.850
		Yes	0,0.3	1,1.3,2,2.3	0.973	0.998
		Yes	4.98,5.03	9.69,11.41,0.12,-0.06	0.124	0.067
		Yes	4.98,5.03	9.69,11.41,0.12,-0.06	0.976	0.793
100	100	No	0, 0	1,2,2,2	0.037	0.032
		No	0,0.1	1,1,2,2.1	0.137	0.270
		No	0,0.2	1,1.2,2,2.2	0.906	0.987
		No	0,0.3	1,1.3,2,2.3	1.000	1.000
		Yes	0,0	1,1,2,2	0.037	0.037
		Yes	0,0.1	1,1,2,2.1	0.102	0.244
		Yes	0,0.2	1,1.2,2,2.2	0.870	0.979

Yes	0,0.3	1,1.3,2,2.3	0.999	1.000
Yes	4.98,5.03	9.69,11.41,0.12,-0.06	0.168	0.086
Yes	4.98,5.03	9.69,11.41,0.12,-0.06	0.998	0.954

Table 8. *The 95% confidence interval estimation of 35th and 60th quantiles of distribution functions related to CPIS observations based on the centered data (CPIS-mean(CPIS)) that correspond to the oral treatment group (CHX) and the control group, respectively.*

	35th quantile			60th quantile		
	Quantile value	95% Lower bound	95% Upper bound	Quantile value	95% Lower bound	95% Upper bound
CHX	-1.0345	-1.4890	-0.5799	-0.0345	-0.4602	0.3913
Control	0.0169	-0.5199	0.5538	1.0169	0.4571	1.5767

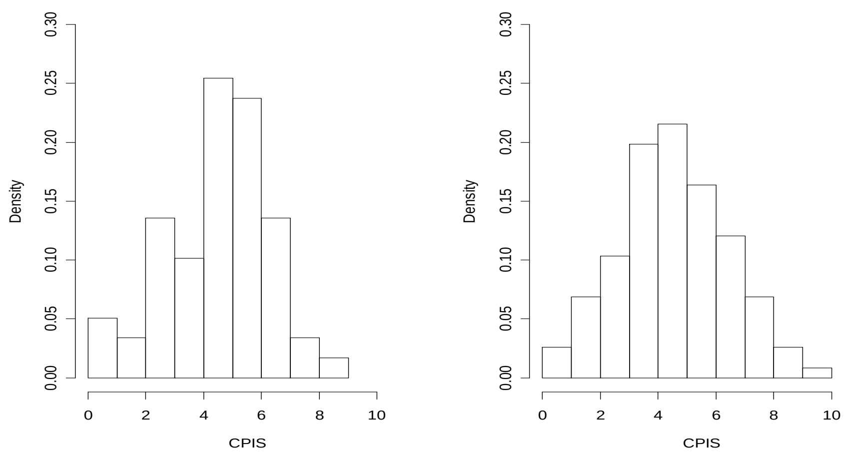

Figure 1. The histograms of CPIS. The figure in the right-hand side corresponds to the oral treatment group, and the figure in the left-hand side corresponds to the control group.