



Empirical likelihood ratio tests for homogeneity of component lifetime distributions based on system lifetime data

Jingjing Qu¹ · Hon Keung Tony Ng² · Chul Moon¹

Received: 7 February 2023 / Accepted: 20 September 2023 / Published online: 21 October 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

In system reliability, practitioners may be interested in testing the homogeneity of the component lifetime distributions based on system lifetimes from multiple data sources for various reasons, such as identifying the component supplier that provides the most reliable components. In this paper, we develop distribution-free hypothesis testing procedures for the homogeneity of the component lifetime distributions based on system lifetime data when the system structures are known. Several non-parametric testing statistics based on the empirical likelihood method are proposed for testing the homogeneity of two or more component lifetime distributions. The computational approaches to obtain the critical values of the proposed test procedures are provided. The performances of the proposed empirical likelihood ratio test procedures are evaluated and compared to the nonparametric Mann–Whitney U test and some parametric test procedures. The simulation results show that the proposed test procedures provide comparable power performance under different sample sizes and underlying component lifetime distributions, and they are powerful in detecting changes in the shape of the distributions.

Keywords Distribution-free · Nonparametric estimation · System signatures · Weibull distribution

1 Introduction

In reliability studies involving systems, practitioners may be interested in comparing the lifetime characteristics of different systems and the lifetime characteristics of the components that make up those systems. For example, one wants to compare

✉ Hon Keung Tony Ng
tng@bentley.edu

¹ Department of Statistics and Data Science, Southern Methodist University, Dallas, TX 75275-0332, USA

² Department of Mathematical Sciences, Bentley University, Waltham, MA 02452, USA

the lifetime characteristics of single-cell cylindrical dry batteries (i.e., components) when they are used in different devices (i.e., systems). The problem of comparing the lifetime characteristics of the components that make up the systems becomes challenging when the system structures are different, especially when the comparisons can only be made based on system lifetime data. For instance, the lifetimes of single-cell cylindrical dry batteries can be very different when used in high-drain devices (e.g., digital cameras and radio-controlled toys) compared to those used in low-drain devices (e.g., clocks and remote controls). In this case, the life testing experiments can be done only when the batteries are used in the devices. Therefore, if researchers want to obtain information about the lifetime of the cylindrical dry batteries, the only accessible data is the system lifetime data. Another example is when the life test involves fielded systems where data are gathered when systems are deployed to the field after the prototype stage. The information on which component leads to the system failure cannot usually be accessed because the experimenters often do not have the need or capability to measure the failed components one by one or the whole system may be discarded after failure. In this paper, the problem of interest is that, given the system structure and system lifetime data, we test if the components from different systems follow the same distribution. The lack of homogeneity of the component lifetime distributions in two different systems could be a consequence of the working conditions in both systems (i.e., the components can be identically distributed a priori, but they are not identically distributed when they are placed in systems with different structures).

Our study is of practical interest, as it will apply directly to the many situations in which the i.i.d. assumption is deemed reasonable, and the notion of system signature proposed by Samaniego (1985) can be applied. Statistical inference for the lifetime distribution of components based on system lifetimes is of interest and has been developed in the past decades. Early literature on this topic is based on “masked data”, which assumes that only partial information is available on the component failures that lead to the system’s failure. Under this framework, Meilijson (1981) and Bueno (1988) identified the component lifetime distribution estimation based on system failure times together with autopsy statistics on the components in the system. Miyakawa (1984) discussed parametric and nonparametric estimation methods for component reliability in the 2-component series system under competing risks with incomplete data. Boyles and Samaniego (1987) derived the nonparametric maximum likelihood estimate (MLE) of component reliability based on nomination sampling in parallel systems. Usher and Hodgson (1988) explored a general method for estimating component reliabilities from a J -component series system lifetime data. Guess et al. (1991) extended Miyakawa’s work and treated a broader class of estimation problems based on masked data. Given the coherent systems with known system structures described by signature, numerous papers were published for component lifetime distributions based on complete system lifetime data. More recent work includes the paper by Bhattacharya and Samaniego (2010) in which the authors estimated the component reliability from system failure data. Eryilmaz et al. (2011) discussed reliability properties of m -consecutive- k -out-of- n : F systems with exchangeable components. Balakrishnan et al. (2011a, b) developed the exact non-parametric method to measure characteristics of the component lifetime distribution

based on the lifetimes of coherent systems with known signatures for complete and censored data, respectively. Navarro et al. (2012) developed a general method for inference on the scale parameter of the component lifetime distribution from the system lifetimes. Ng et al. (2012) discussed parametric statistical inference for the component lifetime distributions when they follow a proportional hazard rate model and the system signature is known. Chahkandi et al. (2014) constructed prediction intervals for the lifetime of a coherent system with known signature. In contrast, while most existing works dealt with system lifetimes from systems with the same signature, Hall et al. (2015) developed a novel nonparametric estimator of component reliability function by maximizing the combined system likelihood function when the systems have different system signatures. Then, Jin et al. (2017) extended the work by Hall et al. (2015) to the case that the system signatures are unknown. For testing the homogeneity of the component lifetime distributions based on system lifetime data, Zhang et al. (2015) developed a nonparametric test and several parametric tests under the exponential distribution and studied the power performance of these test procedures.

In this paper, we develop the nonparametric test procedures for testing the homogeneity of component lifetime distributions using the empirical likelihood ratio. The empirical likelihood was first proposed by Owen (1988). It is known that the empirical likelihood ratio behaves like an ordinary parametric likelihood ratio (Wilks 1938), such as the asymptotic chi-square distribution for the empirical likelihood ratio test statistic, and it has advantages over other nonparametric methods. The book by Owen (2001) on empirical likelihood provides an excellent theoretical background on empirical likelihood and its application. The idea of empirical likelihood has been applied to different areas of statistics such as time series analysis (Nordman and Lahiri 2014), survival analysis (Zhou 2019), longitudinal data analysis (Nadarajah et al. 2014), and regression analysis (Chen and Keilegom 2009). Zhang (2002) proposed a new parameterization approach to construct goodness-of-fit tests based on empirical likelihood ratio and showed that the proposed tests are more powerful than the Kolmogorov–Smirnov, the Cramér–von Mises, and Anderson–Darling tests. Later on, Zhang (2006) proposed a more powerful nonparametric statistic for testing the homogeneity of statistical distributions based on empirical likelihood ratio. In addition to empirical likelihood ratio tests, another nonparametric statistic for testing the homogeneity of distributions can be considered is the rank-based Mann–Whitney-type statistic (which is also known as U statistic) based on the work by Mann and Whitney (1947) and Wilcoxon (1992). The hypothesis testing using U statistic is a rank-sum test, and it can be used for comparing two unpaired groups of data.

In this paper, based on the empirical likelihood ratio, we propose two nonparametric test procedures for testing the homogeneity of components' lifetime distributions given the system lifetimes and the system signatures. The rest of the paper is organized as follows. In Sect. 2, we introduce the mathematical notations and formulate the homogeneity test as a statistical hypothesis testing problem. We also review the existing parametric and nonparametric testing procedures for the homogeneity of component lifetime distributions based on system-level data. In Sect. 3, we introduce the two proposed test procedures based on the empirical likelihood ratio. Then,

the computational approach to obtain the null distributions of the nonparametric test statistics by means of the Monte Carlo method is described in Sect. 4. In Sect. 5, a numerical example is used to illustrate the methodologies proposed in this paper. Monte Carlo simulation studies are used to evaluate the performances of those parametric and nonparametric test procedures in Sect. 6. Finally, concluding remarks and future research directions are provided in Sect. 7.

2 Tests for homogeneity based on system lifetime data

In this section, we introduce the mathematical notations and formulate the homogeneity test as a statistical hypothesis testing problem. Then, we review several existing parametric and nonparametric test procedures.

Suppose there are two different coherent systems, System 1 and System 2, with n_1 and n_2 i.i.d. components following the same absolutely continuous lifetime distribution, respectively. The system lifetime data are obtained by putting M_1 System 1 and M_2 System 2 on a life test. We denote the system lifetimes for System i as $T_i = (T_{i1}, T_{i2}, \dots, T_{iM_i})$, and the lifetimes of the n_i components in the j -th system of System i (i.e., the system with lifetime T_{ij}) as $X_{ij} = (X_{ij1}, X_{ij2}, \dots, X_{ijn_i})$, $i = 1, 2$, $j = 1, 2, \dots, M_i$. The cumulative distribution function (CDF), survival function (SF), and probability density function (PDF) of the lifetimes of System i (i.e., T_{ij}) are denoted by F_{T_i} , \bar{F}_{T_i} , and f_{T_i} , for $i = 1, 2$, and the CDF, SF, and PDF of the lifetimes of the components in System i (i.e., X_{ijk}) are denoted by F_{X_i} , \bar{F}_{X_i} , and f_{X_i} , for $i = 1, 2$.

To describe the structure of a coherent system, we consider the system signature introduced by Samaniego (1985). The signature $s_i = (s_{i1}, \dots, s_{in_i})$ of the coherent system i ($i = 1, 2$) is defined as

$$\begin{aligned}
 s_{ik} &= \Pr(\text{system } i \text{ fails on the failure of the } k\text{-th component}) \\
 &= \Pr(T_i = X_{k:n_i}),
 \end{aligned}
 \tag{1}$$

where s_{ik} , $k = 1, 2, \dots, n_i$, are non-negative real numbers in $[0, 1]$ that do not depend on the component lifetime distribution F_{X_i} with $\sum_{k=1}^{n_i} s_{ik} = 1$. From Kochar et al. (1999) (see also, Samaniego 2007), the SF of the system lifetime T_i can be written in terms of the SF of the component lifetime X_{ijk} as

$$\begin{aligned}
 \bar{F}_{T_i}(t) &= \sum_{k=1}^{n_i} s_{ik} \sum_{\ell=0}^{k-1} \binom{n_i}{\ell} [1 - \bar{F}_{X_i}(t)]^\ell [\bar{F}_{X_i}(t)]^{n_i-\ell} \\
 &\triangleq h_i(p_i(t)),
 \end{aligned}
 \tag{2}$$

where $p_i(t) = \bar{F}_{X_i}(t)$ and $h_i(\cdot)$ is a polynomial function that can be obtained from the system signature or the minimal signature. We further denote the inverse function of $h_i(\cdot)$ as

$$h_i^{-1}(q_i(t)) = p_i(t),$$

where $q_i(t) = \bar{F}_{T_i}(t)$. Navarro et al. (2011) stated that the polynomial $h(p)$ is strictly increasing for $p \in (0, 1)$, with $h(0) = 0$ and $h(1) = 1$, hence, its inverse function h^{-1} in $(0, 1)$ exists and is also strictly increasing in $(0, 1)$ with $h^{-1}(0) = 0$ and $h^{-1}(1) = 1$. Moreover, the PDF of the system lifetime T_j can be expressed in terms of $f_{X_i}(t)$ and $p_i(t) = \bar{F}_{X_i}(t)$ as

$$f_{T_i}(t) = \sum_{k=1}^{n_i} s_{ik} \binom{n_i}{k} k f_{X_i}(t) [1 - p_i(t)]^{k-1} [p_i(t)]^{n_i-k}. \tag{3}$$

In addition to the system signatures, Navarro et al. (2007) noted that the SF of the system lifetime T_i can be expressed in terms of the SF of k -component series system lifetimes, $X_{i,1:k} = \min(X_{i1}, X_{i2}, \dots, X_{ik}), k = 1, 2, \dots, n_i$:

$$\bar{F}_{T_i}(t) = \sum_{k=1}^{n_i} a_{ik} \bar{F}_{X_i}^k(t),$$

where $\bar{F}_{X_i}^k(t)$ is the SF of a k -component series system lifetime, for some non-negative and negative integers $a_{i1}, a_{i2}, \dots, a_{in_i}$, that do not depend on the component lifetime distribution with $\sum_{k=1}^{n_i} a_{ik} = 1$ (see also, Ng et al. 2012). The vector $\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{in_i})$ is called the minimal signature of System i . The values of the minimal signature are the coefficients of the polynomial function $h_i(\cdot)$ defined in Eq. (2) (see, for example, Navarro 2022; Navarro and Rubio 2009).

Based on the system lifetime data $T_i, i = 1, 2$, we are interested in testing the homogeneity of the component lifetime distributions, which can be formulated as a hypothesis testing problem as testing

$$\begin{aligned} H_0 &: F_{X_1}(t) = F_{X_2}(t), \quad \text{for all } t \in (0, \infty) \\ \text{versus } H_1 &: F_{X_1}(t) \neq F_{X_2}(t), \quad \text{for some } t \in (0, \infty). \end{aligned} \tag{4}$$

2.1 Parametric test procedures

For comparative purposes, we consider two parametric tests—an asymptotic test under the assumption of exponentially distributed component lifetimes and a parametric likelihood ratio test—for testing the hypotheses in Eq. (4) based on system lifetime data $T_i, i = 1, 2$.

2.1.1 Asymptotic tests under exponentially distributed component lifetimes assumption

We consider the asymptotic parametric test developed in Zhang et al. (2015) based on the assumption that the underlying lifetimes of components follow an exponential distribution. Specifically, we assume that the lifetimes of components from System i

($i = 1, 2$) follow an exponential distribution with scale parameter $\theta_i > 0$ (denoted as $Exp(\theta_i)$) with PDF

$$f_{X_i}(t) = \frac{1}{\theta_i} \exp\left(-\frac{t}{\theta_i}\right), \quad \text{for } t \geq 0. \tag{5}$$

To estimate the scale parameter θ_i using the system lifetime data $T_i, i = 1, 2$, we use the method of moments estimator (MME) (Ng et al. 2012) defined as

$$\tilde{\theta}_i = \frac{\sum_{j=1}^{M_i} T_{ij}}{M_i \Delta_i},$$

where T_{ij} is the j -th system lifetime of System i and $\Delta_i = \sum_{k=1}^{n_i} (a_{ik}/k)$. Ng et al. (2012) showed that the MME $\tilde{\theta}_i$ is an unbiased estimator and they derived the variance of $\tilde{\theta}_i$ as

$$Var(\tilde{\theta}_i) = \frac{\theta_i^2}{M_i} \left(\frac{2\Delta_i^{(2)}}{\Delta_i^2} - 1 \right), \tag{6}$$

where $\Delta_i^{(2)} = \sum_{k=1}^{n_i} (a_{ik}/k^2)$. For other properties of the MME $\tilde{\theta}_i$ and its exact distribution, one can refer to Ng et al. (2012) and Zhang et al. (2015). Under the assumption of exponentially distributed component lifetimes, testing the hypotheses of homogeneity in Eq. (4) is equivalent to testing

$$H_0^* : \theta_1/\theta_2 = 1 \quad \text{versus} \quad H_1^* : \theta_1/\theta_2 \neq 1. \tag{7}$$

To test the hypotheses in Eq. (7), Zhang et al. (2015) developed different test procedures based on the MME $\tilde{\theta}_i, i = 1, 2$ and they recommend the test using the logarithm transformation of $R = \tilde{\theta}_1/\tilde{\theta}_2$ based on the power performance in a simulation study. Therefore, we consider the test statistic

$$Z_L = \frac{\ln R}{\sqrt{Var(\ln R)}} = \frac{\ln \tilde{\theta}_1 - \ln \tilde{\theta}_2}{\sqrt{Var(\ln \tilde{\theta}_1) + Var(\ln \tilde{\theta}_2)}},$$

where $Var(\ln \tilde{\theta}_i)$ can be approximated by the delta method as (see, for example, Section C.2. of Meeker et al. 2022)

$$Var(\ln \tilde{\theta}_i) \approx \frac{Var(\tilde{\theta}_i)}{\tilde{\theta}_i^2} = \frac{1}{M_i} \left(\frac{2\Delta_i^{(2)}}{\Delta_i^2} - 1 \right), \tag{8}$$

for $i = 1, 2$. Note that the approximated $Var(\ln \tilde{\theta}_i)$ can also be obtained using the delta method for approximating moments described in Theorem 5.3.1 and Corollary 5.3.2 of Bickel and Doksum (2001) by considering the system lifetime T_{ij} with PDF

$$f_{T_i}(t) = \sum_{k=1}^{n_i} a_{ik} \frac{k}{\theta_i} \exp\left(-\frac{k}{\theta_i} t\right), \quad t > 0, i = 1, 2,$$

and the function of the sample mean $\bar{T}_i = \sum_{j=1}^{M_i} T_{ij}$ as

$$g(\bar{T}_i) = \ln \left(\frac{\bar{T}_i}{\Delta_i} \right) = \tilde{\theta}_i, \quad i = 1, 2.$$

The details of the approximation based on the results in Bickel and Doksum (2001) are provided in the Supplementary Materials.

Under the null hypothesis H_0^* in Eq. (7), the test statistic Z_L is asymptotically standard normally distributed. Therefore, the p -value of the asymptotic test based on test statistic Z_L can be calculated as $2(1 - \Phi(|Z_L|))$ where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

2.1.2 Parametric likelihood ratio test

Suppose that the component lifetime distributions for the components in System 1 and System 2 follow the same parametric family of distributions with PDF $f_{X_i}(t; \theta_i)$ $i = 1, 2$ with parameter vector θ_1 and θ_2 , then testing the hypotheses of homogeneity in Eq. (4) is equivalent to testing

$$H_0^{**} : \theta_1 = \theta_2 \quad \text{versus} \quad H_1^{**} : \theta_1 \neq \theta_2. \tag{9}$$

From Eq. (3), the likelihood function based on System i lifetime data $\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iM_i}), i = 1, 2$, is

$$\begin{aligned} L_i(\theta_i | \mathbf{T}_i) &= \prod_{j=1}^{M_i} f_{T_i}(T_{ij}; \theta_i) \\ &= \prod_{j=1}^{M_i} \sum_{k=1}^{n_i} s_{ik} \binom{n_i}{k} k f_{X_i}(T_{ij}; \theta_i) [F_{X_i}(T_{ij}; \theta_i)]^{k-1} [\bar{F}_{X_i}(T_{ij}; \theta_i)]^{n_i-k}. \end{aligned}$$

The MLE of θ_i based on $\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iM_i})$ alone, denoted as $\hat{\theta}_i$, can be obtained by maximizing $L_i(\theta_i | \mathbf{T}_i)$ with respect to θ_i .

Under the null hypothesis $H_0^{**} : \theta_1 = \theta_2 = \theta$ in Eq. (9), the likelihood function based on $(\mathbf{T}_1, \mathbf{T}_2)$ can be expressed as

$$L_{12}(\theta | \mathbf{T}_1, \mathbf{T}_2) = L_1(\theta | \mathbf{T}_1)L_2(\theta | \mathbf{T}_2)$$

and the MLE of θ based on the pooled data $(\mathbf{T}_1, \mathbf{T}_2)$, denoted as $\hat{\theta}$, can be obtained by maximizing $L_{12}(\theta | \mathbf{T}_1, \mathbf{T}_2)$ with respect to θ . The parametric likelihood ratio statistic is defined as

$$\begin{aligned} \lambda_{LR} &= -2 \ln \left[\frac{\sup_{H_0^{**}} L_{12}(\theta | \mathbf{T}_1, \mathbf{T}_2)}{\sup_{H_1^{**}} L_1(\theta_1 | \mathbf{T}_1)L_2(\theta_2 | \mathbf{T}_2)} \right] \\ &= -2 \ln \left[\frac{L_{12}(\hat{\theta} | \mathbf{T}_1, \mathbf{T}_2)}{L_1(\hat{\theta}_1 | \mathbf{T}_1)L_2(\hat{\theta}_2 | \mathbf{T}_2)} \right]. \end{aligned}$$

By the Wilk's theorem (Wilks 1938), the asymptotic distribution of the likelihood ratio statistic λ_{LR} under null hypothesis is chi-square with degrees of freedom ν , where ν is the difference in dimensionality of θ and (θ_1, θ_2) . The p -value of the parametric likelihood ratio test based on test statistic λ_{LR} can be approximated as $\Pr(\chi_\nu^2 > \lambda_{LR})$ where χ_ν^2 is a random variable following the chi-square distribution with degrees of freedom ν .

2.2 Nonparametric Mann–Whitney U Statistic

To test the hypotheses in Eq. (4), Zhang et al. (2015) proposed a nonparametric test procedure based on the Mann–Whitney U test, also called the Mann–Whitney–Wilcoxon test (Mann and Whitney 1947; Wilcoxon 1992), which is a nonparametric test for comparing two component lifetime distributions. Based on the system lifetime data from System 1 and System 2, we define the indicator function between the two systems as

$$D_{j_1, j_2} = \begin{cases} 1, & \text{if } T_{1j_1} < T_{2j_2}, \\ 0, & \text{otherwise,} \end{cases}$$

for $j_1 = 1, 2, \dots, M_1$ and $j_2 = 1, 2, \dots, M_2$. The U statistic for testing the homogeneity of component lifetime distributions is defined as

$$U = \sum_{j_1=1}^{M_1} \sum_{j_2=1}^{M_2} D_{j_1, j_2} = \sum_{j=1}^{M_2} (M_2 - j + 1)R_j,$$

where R_j is the number of ordered observations from sample T_1 which are in between the $(j-1)$ -th and j -th order statistics from sample T_2 . The support of the statistic U is $\{0, 1, \dots, M_1 M_2\}$ and the null hypothesis in Eq. (4) is rejected if U is too large or too small. Here, we reject the null hypothesis in Eq. (4) at α level if $U \leq c_{U1}$ or $U \geq c_{U2}$, where c_{U1} and c_{U2} are critical values and can be determined by $\Pr(U \leq c_{U1} | H_0) \leq \alpha/2$ and $\Pr(U \geq c_{U2} | H_0) \leq \alpha/2$, respectively. The exact p -value can be calculated as presented in Zhang et al. (2015). Since the computation of the exact p -value requires the numerical evaluation of integration which can be computationally intensive, we consider using the Monte Carlo method to obtain the null distribution of U under different scenarios in this paper. The procedure to get the simulated null distribution of U is described in Sect. 4 below.

3 Proposed test procedures based on empirical likelihood ratio

In this section, we develop test procedures based on empirical likelihood ratio to test the hypotheses in Eq. (4) nonparametrically. From Hall et al. (2015), the empirical likelihood function for System i based on T_i only ($i = 1, 2$) can be expressed as

$$L_{T_i}(t) = \binom{M_i}{Y_i(t)} h_i(p_i(t))^{Y_i(t)} [1 - h_i(p_i(t))]^{M_i - Y_i(t)}, \tag{10}$$

where $Y_i(t) = \sum_{j=1}^{M_i} I_{(t, \infty)}(T_{ij})$, $i = 1, 2$, and $I_{(t, \infty)}(T_{ij})$ is the indicator function defined as

$$I_A(b) = \begin{cases} 1, & \text{if } b \in A, \\ 0, & \text{otherwise.} \end{cases}$$

The empirical likelihood function $L_{T_i}(t)$ in Eq. (10) is maximized with respect to $p_i(t) \in (0, 1)$ at

$$\hat{p}_i(t) = h_i^{-1}(1 - \hat{F}_{T_i}(t)), \tag{11}$$

where $\hat{F}_{T_i}(t)$ is the empirical CDF of $F_{T_i}(t)$ defined as

$$\hat{F}_{T_i}(t) = 1 - \frac{1}{M_i} \sum_{j=1}^{M_i} I_{(t, \infty)}(T_{ij}). \tag{12}$$

That is, based on T_i ,

$$\sup\{L_{T_i}(t)\} = \binom{M_i}{Y_i(t)} h_i(\hat{p}_i(t))^{Y_i(t)} [1 - h_i(\hat{p}_i(t))]^{M_i - Y_i(t)}.$$

Under the null hypothesis that $F_{X_1}(t) = F_{X_2}(t)$ (or equivalently $p_1(t) = p_2(t) = p(t)$), we can pool the samples from System 1 and System 2 into an ordered sample of size $M = M_1 + M_2$, denoted as $T^* = (T_{(1)}^* < T_{(2)}^* < \dots < T_{(M)}^*)$. Following Hall et al. (2015), the empirical likelihood function based on the pooled data is

$$L_T^*(t) = \prod_{i=1,2} \binom{M_i}{Y_i(t)} h_i(p(t))^{Y_i(t)} [1 - h_i(p(t))]^{M_i - Y_i(t)}. \tag{13}$$

The nonparametric MLE of $p(t)$, denoted as $\hat{p}(t)$, can be obtained by maximizing $L_T^*(t)$ with respect to $p(t) \in (0, 1)$, i.e.,

$$\sup\{L_T^*(t)\} = \prod_{i=1,2} \binom{M_i}{Y_i(t)} h_i(\hat{p}(t))^{Y_i(t)} [1 - h_i(\hat{p}(t))]^{M_i - Y_i(t)}.$$

Due to the inversion of the function $h(p)$ may not be explicit, a closed-form solution $\hat{p}(t)$ may not be obtained, and the maximization can be approximated by numerical methods such as the Newton–Raphson method. The nonparametric MLE $\hat{p}(t)$ can be written as

$$\hat{p}(t) = \begin{cases} 1, & t < T_{(1)}^*; \\ q_j, & t \in [T_{(j)}^*, T_{(j+1)}^*); \\ 0, & t \geq T_{(M)}^*. \end{cases}$$

where q_j maximizes the likelihood function in Eq. (13) for t in $[T_{(j)}^*, T_{(j+1)}^*]$, $j = 1, 2, \dots, M - 1$. For the initial values of q_j (denoted as $q_j^{(0)}$) of the iterative maximization algorithm and $t \in [T_{(j)}^*, T_{(j+1)}^*]$, we consider a weighted function of $\hat{p}_i(t)$, $i = 1, 2$, in Eq. (11)

$$q_j^{(0)} = \frac{M_1 \hat{p}_1(t) + M_2 \hat{p}_2(t)}{M}, \quad j = 1, 2, \dots, M - 1.$$

The empirical likelihood ratio at time t is

$$R_T(t) = \frac{\sup\{L_T^*(t) : p_1(t) = p_2(t)\}}{\sup\{L_{T_1}(t)\} \sup\{L_{T_2}(t)\}}.$$

Then, we obtain the log empirical likelihood ratio as

$$\begin{aligned} \ln R_T(t) &= \ln \left\{ \frac{\sup\{L_T^*(t) : p_1(t) = p_2(t)\}}{\sup\{L_{T_1}(t)\} \sup\{L_{T_2}(t)\}} \right\} \\ &= \ln \left\{ \frac{\prod_{i=1,2} \binom{M_i}{Y_i(t)} h_i(\hat{p}(t))^{Y_i(t)} [1 - h_i(\hat{p}(t))]^{M_i - Y_i(t)}}{\prod_{i=1,2} \binom{M_i}{Y_i(t)} h_i(\hat{p}_i(t))^{Y_i(t)} [1 - h_i(\hat{p}_i(t))]^{M_i - Y_i(t)}} \right\} \\ &= \sum_{i=1}^2 \left\{ Y_i(t) \ln \left[\frac{h_i(\hat{p}(t))}{h_i(\hat{p}_i(t))} \right] + (M_i - Y_i(t)) \ln \left[\frac{1 - h_i(\hat{p}(t))}{1 - h_i(\hat{p}_i(t))} \right] \right\}. \end{aligned}$$

Based on the log empirical likelihood ratio, we define

$$\begin{aligned} G_t^2 &= 2 \ln R_T(t) \\ &= 2 \sum_{i=1}^2 \left\{ Y_i(t) \ln \left[\frac{h_i(\hat{p}(t))}{h_i(\hat{p}_i(t))} \right] + (M_i - Y_i(t)) \ln \left[\frac{1 - h_i(\hat{p}(t))}{1 - h_i(\hat{p}_i(t))} \right] \right\}, \end{aligned}$$

which is a function of time t . We consider two functions based on G_t^2 :

$$Z^* = \sup_{t \in (0, \infty)} [G_t^2 \times w(t)], \tag{14}$$

$$Z = \int_0^\infty G_i^2 dw(t), \tag{15}$$

where $w(t)$ is a weight function. Following Zhang (2006), we propose two test statistics for testing the hypotheses in Eq. (4) with different weight functions:

1. Take $w(t) = 1$ in Eq. (14), we obtain the test statistic

$$\begin{aligned} Z_K &\triangleq \sup_{t \in (0, \infty)} [G_t^2] \\ &= \max_{1 \leq k \leq M} \left\{ 2 \sum_{i=1}^2 Y_i(T_{(k)}^*) \ln \left[\frac{h_i(\hat{p}_i(T_{(k)}^*))}{h_i(\hat{p}(T_{(k)}^*))} \right] \right. \\ &\quad \left. + (M_i - Y_i(T_{(k)}^*)) \ln \left[\frac{1 - h_i(\hat{p}_i(T_{(k)}^*))}{1 - h_i(\hat{p}(T_{(k)}^*))} \right] \right\}. \end{aligned}$$

2. We denote the maximum likelihood estimator of component CDF for the pooled data as $\hat{F}_X(t)$ and define $\hat{F}_X(T_{(0)}^*) = 0$. Then, $\hat{F}_X(t) = 1 - \hat{p}(t)$. In Eq. (15), take the weight function

$$dw(t) = \frac{1}{\hat{F}_X(t)(1 - \hat{F}_X(t))} d\hat{F}_X(t).$$

Then, the test statistic can be written as

$$\begin{aligned} Z_A &\triangleq \int_0^\infty G_t^2 \frac{1}{\hat{F}_X(t)(1 - \hat{F}_X(t))} d\hat{F}_X(t) \\ &= 2 \sum_{k=1}^M \frac{\hat{p}(T_{(k)}^*) - \hat{p}(T_{(k-1)}^*)}{\hat{p}(T_{(k)}^*)(1 - \hat{p}(T_{(k)}^*))} \\ &\quad \times \left\{ \sum_{i=1}^2 \left[Y_i(T_{(k)}^*) \ln \frac{h_i(\hat{p}_i(T_{(k)}^*))}{h_i(\hat{p}(T_{(k)}^*))} + (M_i - Y_i(T_{(k)}^*)) \ln \frac{1 - h_i(\hat{p}_i(T_{(k)}^*))}{1 - h_i(\hat{p}(T_{(k)}^*))} \right] \right\}. \end{aligned}$$

Large values of Z_K and Z_A support the alternative hypothesis in Eq. (4), which leads to the rejection of the null hypothesis in Eq. (4).

4 Null distributions of Z_K, Z_A , and U based on Monte Carlo method

Since the distributions of Z_K and Z_A are intractable theoretically in general, even under the null hypothesis, we rely on the Monte Carlo method to obtain the null distributions of Z_K and Z_A . We simulate the lifetimes of M_1 systems for System 1 and the lifetimes of the M_2 systems for System 2 from any component lifetime

Table 1 Simulated 90-th, 95-th, and 99-th percentage points of the null distributions of Z_K and Z_A with $\mathbf{s}_1 = (0, 0, 0, 1)$ (i.e., $n_1 = 4$) and $\mathbf{s}_2 = (1, 0, 0)$ (i.e., $n_2 = 3$) for different sample sizes of $M_1 = M_2$

Sample sizes	Z_K			Z_A		
	90%	95%	99%	90%	95%	99%
$M_1 = M_2 = 10$	4.2826	6.2433	10.0858	4.2609	6.6789	13.4692
$M_1 = M_2 = 15$	5.7377	6.6794	10.2433	4.7174	7.0945	13.4382
$M_1 = M_2 = 20$	7.6502	7.6502	10.7033	5.1977	7.1561	13.2830
$M_1 = M_2 = 30$	6.9730	7.7890	11.4753	5.9456	7.6422	13.4232
$M_1 = M_2 = 50$	7.1904	8.7383	11.8581	6.5791	8.3773	13.6362

distribution F_X with $F_{X_1} = F_{X_2} = F_X$ for given n_1, n_2, M_1, M_2 , and system signatures \mathbf{s}_1 and \mathbf{s}_2 . The statistics Z_K and Z_A are computed from the simulated lifetimes.

In practice, we can simulate the null distributions of Z_K, Z_A , and U by choosing F_X as a distribution that is easy to simulate (e.g., standard exponential distribution, $Exp(1)$) with a large number of Monte Carlo simulations (say, 1,000,000 times). For example, the simulated 90-th, 95-th, and 99-th percentage points for Z_K and Z_A with $\mathbf{s}_1 = (0, 0, 0, 1)$ (i.e., $n_1 = 4$, with minimal signature $\mathbf{a}_1 = (4, -6, 4, -1)$) and $\mathbf{s}_2 = (1, 0, 0)$ (i.e., $n_2 = 3$, with minimal signature $\mathbf{a}_2 = (0, 0, 1)$) based on 1,000,000 simulations with $F_X \sim Exp(1)$ are presented in Table 1.

For the null distribution of the Mann–Whitney U statistic, since U is discrete and the number of possible values of U can be small when the sample sizes M_1 and M_2 are small, the actual percentage when U equals to or less than the critical value might be much larger than the nominal significance level. Therefore, instead of using the exact distribution presented in Zhang et al. (2015), we propose to use a Monte Carlo simulation method to obtain the approximated percentage points of the null distribution and use a randomization procedure to obtain a test procedure based on U with the required significance level. To illustrate the procedure for obtaining the critical values, we consider the following example in which System 1 is a 4-component parallel system, and System 2 is a 3-component series system with system signatures $\mathbf{s}_1 = (0, 0, 0, 1)$ and $\mathbf{s}_2 = (1, 0, 0)$, respectively. In Table 2, we present the critical values of the Mann–Whitney statistic U and the corresponding simulated probability that U is more extreme than or equal to the critical value for different significance levels and sample sizes. The values in Table 2 are generated based on 1,000,000 simulations. For illustrative purposes, we consider the case of $M_1 = M_2 = 10$ to demonstrate how to conduct the hypothesis test at 5% significance level.

From Table 2, for $\alpha = 0.05$, the critical values are $c_{U1} = 0$ and $c_{U2} = 12$ and we have the following probability

$$\Pr(U \leq 0) = 0.3254, \Pr(U \geq 12) = 0.02520.$$

If we use the critical values $c_{U1} = 0$ and $c_{U2} = 12$ directly, the significance level will be $0.3254 + 0.0252 = 0.3506$, which is much higher than the 5% level. Therefore, we use the simulated null distribution with a randomization procedure to control

Table 2 Simulated 0.5-th, 95-th, and 99-th percentage points of the null distributions of U with $s_1 = (0, 0, 0, 1)$ (i.e., $n_1 = 4$) and $s_2 = (1, 0, 0)$ (i.e., $n_2 = 3$) for different sample sizes $M_1 = M_2$

Sample size	Percentage								
	0.5%			2.5%			5%		
$M_1 = M_2$	c_{U1}	$\Pr(U = c_{U1})$	$\Pr(U \leq c_{U1})$	c_{U1}	$\Pr(U = c_{U1})$	$\Pr(U \leq c_{U1})$	c_{U1}	$\Pr(U = c_{U1})$	$\Pr(U \leq c_{U1})$
10	0	0.3254	0.3254	0	0.3254	0.3254	0	0.3254	0.3254
15	0	0.1483	0.1483	0	0.1483	0.1483	0	0.1483	0.1483
20	0	0.0648	0.0648	0	0.0648	0.0648	0	0.0648	0.0648
30	0	0.0114	0.0114	2	0.0110	0.0301	4	0.0149	0.0571
50	8	0.0013	0.0062	17	0.0034	0.0280	23	0.0051	0.0548

Sample size	Percentage								
	0.5%			2.5%			5%		
$M_1 = M_2$	c_{U2}	$\Pr(U = c_{U2})$	$\Pr(U \geq c_{U2})$	c_{U2}	$\Pr(U = c_{U2})$	$\Pr(U \geq c_{U2})$	c_{U2}	$\Pr(U = c_{U2})$	$\Pr(U \geq c_{U2})$
10	10	0.0148	0.05011	12	0.0076	0.02520	16	0.0019	0.005864
15	18	0.0106	0.05556	21	0.0057	0.02953	28	0.0013	0.006212
20	29	0.0068	0.05083	33	0.0039	0.02822	43	0.0009	0.005748
30	56	0.0042	0.05308	64	0.0023	0.02630	81	0.0005	0.005297
50	134	0.0022	0.052092	150	0.0012	0.02553	183	0.0003	0.005162

the significance level of the test based on U . For example, to control $\Pr(\text{reject } H_0 \text{ when } U \text{ is too small} | H_0) \leq 0.025$, we do not reject H_0 if the observed value of U is greater than 0 and less than 12. If the observed value of U is 0, we reject H_0 with probability $0.025/0.3254$, i.e.,

$$\Pr(\text{reject } H_0 \text{ when } U \text{ is too small} | H_0) = \Pr(U \leq 0 | H_0)(0.025/0.3254) = 0.025.$$

Similarly, to control $\Pr(\text{reject } H_0 \text{ when } U \text{ is too large} | H_0) \leq 0.025$, we do not reject H_0 if the observed value of U is less than 12 and reject H_0 if the observed value of U is larger than 12. If the observed value of U is 12, we reject H_0 with probability $(0.025 - \Pr(U > 12) / \Pr(U = 12)) = (0.025 - (\Pr(U \geq 12) - \Pr(U = 12)) / \Pr(U = 12)) = 0.9737$, i.e.,

$$\begin{aligned} &\Pr(\text{reject } H_0 \text{ when } U \text{ is too large} | H_0) \\ &= \Pr(U > 12 | H_0) + \Pr(U = 12 | H_0) \left[\frac{0.025 - \Pr(U > 12 | H_0)}{\Pr(U = 12 | H_0)} \right] \\ &= (\Pr(U \geq 12 | H_0) - \Pr(U = 12 | H_0)) + \Pr(U = 12 | H_0) \times 0.9737 \\ &= 0.025. \end{aligned}$$

Obviously, when $\Pr(U \leq c_{U1}) = 0.025$ and $\Pr(U \geq c_{U2}) = 0.025$, no randomization procedure is needed. Following this procedure, we can control the significance level of the two-sided U test under H_0 for $\alpha = 0.01, 0.05, 0.1$ using the critical values c_{U1}

and c_{U2} and the corresponding values of $\Pr(U \leq c_{U1})$ and $\Pr(U \geq c_{U2})$ when these probabilities are not equal to $\alpha/2$.

5 Illustrative example

To illustrate the test procedures developed in this paper, we analyze a data set based on the example presented in Yang et al. (2016) and Frenkel and Khvatskin (2006). The example given in Yang et al. (2016) and Frenkel and Khvatskin (2006) described the phosphor acid filter system as a real-life prototype of a consecutive 2-out-of- n system. For a consecutive 2-out-of- n system, the system fails when any two adjacent components fail. For illustrative purposes, we consider that System 1 is a consecutive 2-out-of-8 system with system signature $\mathbf{s}_1 = (0, 1/4, 11/28, 2/7, 1/14, 0, 0, 0)$ (minimal signature $\mathbf{a}_1 = (0, 7, -6, 0, 0, 0, 0, 0)$) and component lifetimes follow a Birnbaum–Saunders distribution (Birnbaum and Saunders 1969) with CDF

$$F_X(t; a, b) = \Phi \left[\frac{1}{a} \left(\left(\frac{t}{b} \right)^{1/2} - \left(\frac{b}{t} \right)^{1/2} \right) \right], \quad t > 0,$$

where the shape parameter is $a = 1$ and scale parameter $b = 1$ and System 2 is a 4-component with system signature $\mathbf{s}_2 = (1/4, 1/4, 1/2, 0)$ (minimal signature $\mathbf{a}_2 = (0, 3, -3, 1)$) and component lifetimes follow a Weibull(3, 2) distribution. The system lifetime data for System 1 and System 2 with sample sizes $M_1 = M_2 = 20$ are presented in Table 3. A hypothesis test is conducted to determine if the components from two different systems follow the same lifetime distribution.

The nonparametric MLE of the SF of the component lifetime distribution F_{X_1} based on T_1 from Eq. (12) (denoted as \hat{F}_{X_1}), the nonparametric MLE of the SF of the component lifetime distribution F_{X_2} based on T_2 from Eq. (12) (denoted as \hat{F}_{X_2}), the average of \hat{F}_{X_1} and \hat{F}_{X_2} , and the nonparametric MLE of the component lifetime distribution based on the pooled data (T_1, T_2) under $H_0 : F_{X_1} = F_{X_2}$ by maximizing Eq. (13), are plotted in Fig. 1.

To test the hypotheses in Eq. (4) using the nonparametric Mann–Whitney U test and the two proposed empirical likelihood ratio tests at 1% level of significance for the data set in Table 3, we obtain the critical values based on the procedures described in Sect. 4 with 1,000,000 simulations. The critical values for the empirical likelihood ratio tests based on Z_K and Z_A are 12.6394 and 18.4978, respectively,

Table 3 Simulated lifetime data from System 1 and System 2 with $M_1 = M_2 = 20$

j	1	2	3	4	5	6	7	8	9	10
T_{1j}	0.2598	0.2803	0.3329	0.4172	0.4532	0.459	0.5541	0.5769	0.5842	0.7784
T_{2j}	0.5890	0.6423	0.7774	0.9879	1.0754	1.1200	1.1685	1.2410	1.2412	1.2642
j	11	12	13	14	15	16	17	18	19	20
T_{1j}	0.7917	0.8565	0.8895	0.9186	0.9348	1.1130	1.2049	1.3938	1.4406	1.6351
T_{2j}	1.4164	1.4401	1.4924	1.5415	1.6912	2.0695	2.4374	2.5627	2.6197	2.7791

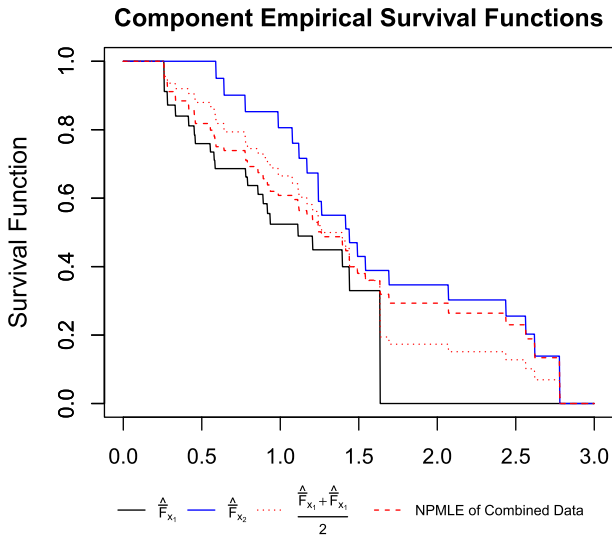


Fig. 1 The nonparametric MLE of the SF of the component lifetime distribution F_{X_1} based on T_1 from Eq. (12) (denoted as \hat{F}_{X_1}), the nonparametric MLE of the SF of the component lifetime distribution F_{X_2} based on T_2 from Eq. (12) (denoted as \hat{F}_{X_2}), the average of \hat{F}_{X_1} and \hat{F}_{X_2} , and the nonparametric MLE of the component lifetime distribution based on the pooled data (T_1, T_2) under $H_0 : F_{X_1} = F_{X_2}$ by maximizing Eq. (13)

and the 0.5 and 99.5 percentile of the Mann–Whitney U statistic are 172 and 314, respectively.

For the data presented in Table 3, we can compute the test statistics $Z_K = 13.5608$, $Z_A = 22.0917$, and $U = 334$. We observe that all these test statistics are larger than their corresponding critical value at a 1% level of significance. Therefore, we reject the null hypothesis in Eq. (4) at 1% level with p -value of 2×10^{-6} based on test statistics Z_A and Z_K and p -value of 0.01 based on the test statistic U . These results agree with our expectation since the component lifetimes in System 1 are simulated from a Birnbaum–Saunders distribution, and the component lifetimes in System 2 are simulated from a Weibull distribution.

6 Monte Carlo simulation studies

In this section, Monte Carlo simulation studies are used to evaluate the performances of those parametric and nonparametric test procedures described in Sects. 2 and 3 for testing the hypotheses in Eq. (4). In these simulation studies, we consider simulating the system lifetimes for the systems with component lifetimes following the statistical distributions listed below:

1. Exponential distribution: Exponential distribution with mean θ has PDF in Eq. (5), denoted as $Exp(\theta)$.
2. Gamma distribution: Gamma distribution with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$, denoted as $Gamma(\alpha, \beta)$, has PDF

$$f_X(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t}, \quad \text{for } t \geq 0, \tag{16}$$

where $\Gamma(\alpha) = \int_0^\infty x^\alpha \exp(-x) dx$ is the gamma function.

3. Weibull distribution: Weibull distribution with scale parameter $\lambda > 0$ and shape parameter $\gamma > 0$, denoted as $Weibull(\lambda, \gamma)$, has PDF

$$f_X(t) = \frac{\gamma}{\lambda} \left(\frac{t}{\lambda}\right)^{\gamma-1} \exp\left[-\left(\frac{t}{\lambda}\right)^\gamma\right] \quad \text{for } t > 0. \tag{17}$$

4. Lognormal distribution: Lognormal distribution with scale parameter $\exp(\mu) > 0$ and shape parameter $\sigma > 0$, denoted as $Lognormal(\mu, \sigma^2)$, has PDF

$$f_X(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln t - \mu)^2}{2\sigma^2}\right], \quad \text{for } t > 0. \tag{18}$$

In the first simulation study, we conduct the parametric tests to examine how the Type-I error rates vary when the underlying distributions are misspecified. We consider applying the asymptotic parametric test under exponentially distributed component lifetimes assumption presented in Sect. 2.1.1 to cases when the component lifetimes follow different statistical distributions described above. The PDFs of the distributions considered in the first simulation study are plotted

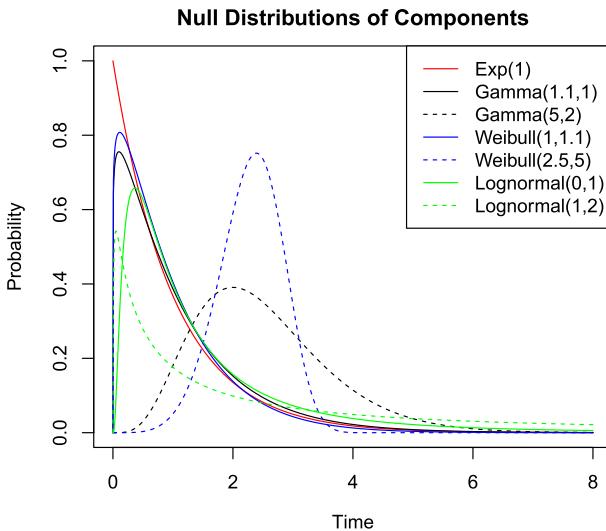


Fig. 2 PDFs of the distributions considered in the Monte Carlo simulation studies

in Fig. 2. We consider that System 1 is a 4-component parallel system and System 2 is a 3-component series system with system signatures $s_1 = (0, 0, 0, 1)$ and $s_2 = (1, 0, 0)$, respectively, with different sample sizes $M_1 = M_2 = 10, 15, 20, 30$, and 50. The simulated rejection rates of the asymptotic parametric test based on 10,000 simulations under the null hypothesis that the components in System 1 and System 2 have the same distribution with 5% level of significance are presented in Table 4.

Table 4 shows that the asymptotic parametric test developed under the exponential assumption inflates the simulated Type-I error rates when the underlying component lifetime distributions deviate from the exponential, such as *Gamma*(5, 2), *Weibull*(2.5, 5), and *Lognormal*(1, 2) (see Fig. 2). On the other hand, the simulated Type-I error rates are close to the nominal level of 5% when the underlying component lifetime distributions are exponential or similar to the exponential distribution, such as *Gamma*(1.1, 1), *Weibull*(1, 1.1) and *Lognormal*(0, 1) (see Fig. 2). These results indicate that those parametric tests for homogeneity of component lifetime distributions may not be appropriate, especially when the underlying distribution is unknown.

In the second simulation study, we evaluate the power performances of the proposed empirical likelihood ratio tests Z_K and Z_A , and the Mann–Whitney U test and compare them with the parametric tests by assuming that the underlying distributions agree with the distributions the data are generated from.

For the parametric likelihood ratio test with a two-parameter distribution, we assume that both F_{X_1} and F_{X_2} are from the same class of distributions and the two parameters are unknown, but one of the parameters in F_{X_1} and F_{X_2} are the same under the alternative hypothesis. For example, in the simulation where the component lifetimes follow the Weibull distribution, F_{X_1} and F_{X_2} are *Weibull*(λ_1, γ_1) and *Weibull*(λ_2, γ_2), respectively, we use the hypotheses

Table 4 Simulated rejection rates of the asymptotic parametric test under the exponentially distributed component lifetimes assumption when the null hypothesis is true at 5% significant level for different underlying component lifetime distributions and different sample sizes

True Dist.	Sample size				
	$M_1 = M_2 = 10$	$M_1 = M_2 = 15$	$M_1 = M_2 = 20$	$M_1 = M_2 = 30$	$M_1 = M_2 = 50$
<i>Exp</i> (1)	0.0531	0.0527	0.0491	0.0524	0.0482
<i>Exp</i> (0.1)	0.0531	0.0523	0.0499	0.0525	0.0492
<i>Gamma</i> (5, 2)	0.9707	0.9998	1.0000	1.0000	1.0000
<i>Gamma</i> (1.1, 1)	0.0465	0.0487	0.0546	0.0547	0.0617
<i>Gamma</i> (1.5, 1)	0.0949	0.1374	0.1918	0.2949	0.5014
<i>Gamma</i> (1.1, 2)	0.0465	0.0487	0.0546	0.0547	0.0617
<i>Weibull</i> (1, 1.1)	0.0474	0.0535	0.0639	0.0730	0.1058
<i>Weibull</i> (2.5, 1.5)	0.2466	0.4061	0.5563	0.7691	0.9621
<i>Weibull</i> (2.5, 5)	1.0000	1.0000	1.0000	1.0000	1.0000
<i>Lognormal</i> (0, 1)	0.0406	0.0437	0.0457	0.0461	0.0490
<i>Lognormal</i> (1, 2)	0.8821	0.9448	0.9740	0.9930	0.9993

$$H_0^{**} : \lambda_1 = \lambda_2 \text{ and } \gamma_1 = \gamma_2$$

$$\text{versus } H_1^{**} : \lambda_1 \neq \lambda_2 \text{ and } \gamma_1 = \gamma_2.$$

As shown in the first simulation study, the asymptotic parametric test developed under the exponentially distributed components assumption may not be appropriate for distributions other than exponential distribution; we only consider the asymptotic parametric test when the data are generated from an exponential distribution. In this simulation study, we consider the following settings for the system structures of System 1 and System 2 (Navarro et al. 2007):

- [S1] System 1: 4-component parallel system with system signature $\mathbf{s}_1 = (0, 0, 0, 1)$ (minimal signature $\mathbf{a}_1 = (4, -6, 4, -1)$);
System 2: 3-component series system with system signature $\mathbf{s}_2 = (1, 0, 0)$ (minimal signature $\mathbf{a}_2 = (0, 0, 1)$).
- [S2] System 1: 3 components system with system signature $\mathbf{s}_1 = (0, 2/3, 1/3)$ (minimal signature $\mathbf{a}_1 = (1, 1, -1)$);
System 2: 4 components system with system signature $\mathbf{s}_2 = (1/4, 1/4, 1/2, 0)$ (minimal signature $\mathbf{a}_2 = (0, 3, -3, 1)$).
- [S3] System 1: 3 components system with system signature $\mathbf{s}_1 = (0, 2/3, 1/3)$ (minimal signature $\mathbf{a}_1 = (1, 1, -1)$);
System 2: 4 components system with system signature $\mathbf{s}_2 = (0, 1/2, 1/4, 1/4)$ (minimal signature $\mathbf{a}_2 = (1, 0, 1, -1)$).

For comparative purposes, we also consider the nonparametric test procedures based on complete component-level data, i.e., all $n_k M_k$ component lifetimes are observed for System k . This is equivalent to considering System 1 and System 2 be 1-component systems such that the lifetime of the component is equal to the lifetime of the system. Since this is an ideal scenario with complete information on the component lifetimes, we denote this case as “Full data”. Note that the power performance of the test procedures based on complete component-level data can be served as a benchmark for the power comparisons since this is the case that all the $n_1 M_1 + n_2 M_2$ component lifetimes are observed. The critical values of nonparametric tests are obtained from the simulation mentioned in Sect. 4 based on 1,000,000 simulations.

The component lifetimes are generated from the following distribution settings:

- [D1] Exponential distributions with changes in the scale parameter: System 1: $Exp(\theta_1)$ with $\theta_1 = 1$ (i.e., $\ln \theta_1 = 0$);
System 2: $Exp(\theta_2)$ with $\ln \theta_2$ varies from -1.6 to 1.6 with increment 0.1 (denoted by -1.6 (0.1) 1.6);
- [D2] Weibull distributions with changes in the shape parameter: System 1: $Weibull(\lambda_1, \gamma_1)$ with $\lambda_1 = 1$ and $\gamma_1 = 1$;
System 2: $Weibull(\lambda_2, \gamma_2)$ with $\lambda_2 = 1$ and γ_2 varies from 0.5 to 1.5 with increment 0.1 (denoted by 0.5 (0.1) 1.5);

- [D3]** Lognormal distributions with changes in the standard deviation on the log-scale (i.e., change in the shape parameter): System 1: $Lognormal(\mu_1, \sigma_1)$ with $\mu_1 = 0$ and $\sigma_1 = 2$;
System 2: $Lognormal(\mu_2, \sigma_2)$ with $\mu_2 = 0$ and σ_2 varies from 1 to 3 with increment 0.1 (denoted by 1 (0.1) 3).
- [D4]** Weibull distributions with changes in the scale parameter: System 1: $Weibull(\lambda_1, \gamma_1)$ with $\lambda_1 = 2.5$ and $\gamma_1 = 5$;
System 2: $Weibull(\lambda_2, \gamma_2)$ with λ_2 varies from 1.5 to 3.5 with increment 0.1 (denoted by 1.5 (0.1) 3.5) and $\gamma_2 = 5$;
- [D5]** Gamma distribution with changes in the shape parameter: System 1: $Gamma(\alpha_1, \beta_1)$ with $\alpha_1 = 5$ and $\beta_1 = 2$;
System 2: $Gamma(\alpha_2, \beta_2)$ with α_2 varies from 3 to 7 with increment 0.2 (denoted by 3 (0.2) 7) and $\beta_2 = 2$;
- [D6]** Gamma distribution with changes in the rate parameter: System 1: $Gamma(\alpha_1, \beta_1)$ with $\alpha_1 = 5$ and $\beta_1 = 2$;
System 2: $Gamma(\alpha_2, \beta_2)$ with $\alpha_2 = 5$ and β_2 varies from 1 to 3 with increment 0.1 (denoted by 1 (0.1) 3);
- [D7]** Lognormal distributions with changes in the mean on the log-scale (i.e., change in the scale parameter): System 1: $Lognormal(\mu_1, \sigma_1)$ with $\mu_1 = 0$ and $\sigma_1 = 1$;
System 2: $Lognormal(\mu_2, \sigma_2)$ with μ_2 varies from -1.6 to 1.6 with increment 0.1 (denoted by -1.6 (0.1) 1.6) and $\sigma_2 = 1$.

The rejection rates with significance level 5% under different settings are estimated based on 10,000 simulations. For the sake of saving space, we present here the simulated power curves for setting [D1] with [S1] in Figure 3, the simulated power curves for setting [D2] with [S1] in Figure 4, the simulated power curves for setting [D3] with [S1], and the simulated power curves for settings [D4] with [S1], [S2], and [S3] in Figures 6–8 in the Supplementary Materials, respectively. We present the simulated power curves for each figure under the sample sizes $M_1 = M_2 = 10, 15, 20, 30,$ and 50 . The simulated power curves for other settings, including [D5]–[D7] with system structures [S1]–[S3] are presented in Figures 11–25 in the Supplementary Materials.

The simulated power curves centered at the simulated rejection rates under the null hypothesis (i.e., $F_{X_1} = F_{X_2}$), which are expected to be close to the nominal significance level of 5%. When the differences between the parameters increase (i.e., moving away from the center), we expect the simulated power values to increase. The closer the power values to one, the better the performance of the test procedure. As mentioned above, the simulated power curves with “Full data” can serve as benchmarks for comparisons as the power values based on complete component-level data are larger than those based on system-level data. Moreover, we expect that the power values for the parametric likelihood ratio tests under the correct model specification of underlying distributions are better than those of nonparametric tests.

From Figures 3 and 4 in the Supplementary Materials, we observe that for the extreme setting [S1] with System 1 being a parallel system and System 2 being a series system, when the sample sizes are small (say, $M_1 = M_2 \leq 20$), the power

values are low when the mean lifetime of components in System 1 is smaller than the mean lifetime of components in System 2 (i.e., the left-hand side of the power curve) for the nonparametric tests. As pointed out by Zhang et al. (2015), it is due to the nature of the problem since we are comparing the worst component in one system to the best component in another system to determine if the lifetime characteristics of the components are the same.

From Figures 3–8 in the Supplementary Materials, the power values of the nonparametric tests with complete component-level data and the parametric tests under the correct specification of the underlying component lifetime distribution are larger than the power values of the nonparametric tests based on system-level data. We observe that the proposed empirical likelihood ratio tests provide comparative power performance in most cases compared to the Mann–Whitney U test. After considering the upper bound of the Monte Carlo error $\sqrt{(0.5)(1 - 0.5)/10,000} = 0.005$, the proposed empirical likelihood ratio tests are more powerful for small sample sizes on the right-hand side. Between the two empirical likelihood ratio tests, the test based on Z_A tends to have better power performance than the test based on Z_K in most cases.

For comparative purposes, the differences between the simulated power values of the test procedures based on U statistic and Z_K , and the test procedures based on U statistic and Z_A for Weibull distribution with $M_1 = M_2 = 30$ and for lognormal distribution with different sample sizes are plotted in Figures 9 and 10 in the Supplementary Materials, respectively. Negative values of the differences indicate that the proposed tests based on Z_A and Z_K provide better power performance than the Mann–Whitney U test. We also included the lines for plus and minus three Monte Carlo errors ($\pm 3MCE$) to indicate if the differences are significant.

When the mean lifetime of the component lifetime distribution of System 2 is larger than the mean lifetime of the component lifetime distribution of System 1, the proposed empirical likelihood ratio test based on Z_A provides better power values among the three nonparametric tests considered here. The advantage of the test based on Z_A is more significant when the two systems are extremely different (i.e., setting [S1]). When the mean lifetime of the component lifetime distribution of System 2 is smaller than the mean lifetime of the component lifetime distribution of System 1, the Mann–Whitney U test provides better power values among the three nonparametric tests considered here.

The proposed nonparametric tests outperform when the shape of the component lifetime distributions are different (i.e., settings [D2] and [D3]). For example, from the top left panel of Figures 9 and 10 in the Supplementary Materials show that the proposed empirical likelihood ratio tests provide better power performance than the Mann–Whitney U test in detecting changes in the shape parameters. Despite the alteration of the shape parameter in setting [D5] (see Figures 11, 17, and 23 in the Supplementary Materials, the proposed tests do not outperform the U test, which may be due to minor deviations in the shape of the gamma distributions.

7 Concluding remarks

In this manuscript, we studied the problem of testing the homogeneity of component lifetime distributions based on system-level data, which can be applied to many practical situations in life testing procedures involving systems with known structures. We showed that those existing parametric test procedures might suffer from the inflation of Type-I error rates when the underlying probability distributions of the component lifetimes are misspecified. To address this issue, we focus on developing nonparametric statistical test procedures for the homogeneity of component lifetime distributions based on system-level data. We proposed two empirical likelihood ratio tests based on the empirical likelihood ratio and the nonparametric estimation of component lifetime distributions. We provided the computational algorithms for obtaining the null distributions of the test statistics using the Monte Carlo method. Our simulation results show that the proposed nonparametric procedures provide comparative power values with those existing tests. These proposed test procedures have advantages in power performance when the two systems are very different. The computer programs to execute the test procedures presented in this manuscript are written in R (R Core Team 2022), and they are available from the authors upon request.

For future research, since censoring and truncation are common in life testing procedures, one can consider the extensions of the test procedures discussed in this paper to the cases in which some system lifetimes are censored. On the other hand, in this work, the system structures are assumed to be known. This assumption may not be realistic in some situations in which the systems of interest are black boxes. Therefore, developing test procedures for the homogeneity problem will be interesting when complete information about the system structure is unavailable. In this situation, we can consider that auxiliary data, such as the number of failed components at the time of system failure, is available along with system lifetime (Jin et al. 2017).

Another possible research direction is to investigate using the homogeneity tests for detecting the dependence between the component lifetimes. Note that the homogeneity tests of component lifetime distributions based on system lifetime data studied here can also be used to detect the dependence between the components. However, when the null hypothesis in Eq. (4) is rejected, it may be due to the difference between the lifetime distributions of the components in the two systems or the dependence between the component lifetimes in the two systems, and these effects may not be distinguishable based on the current test procedures. It will be interesting to develop statistical procedures to detect the lack of independence between the component lifetimes using the homogeneity tests.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00180-023-01421-w>.

Acknowledgements The authors would like to thank the Editor, the Associate Editor, and the referees for their valuable comments, which helped to improve the quality of this article. We especially thank the Associate Editor for assisting us in addressing the reviewers' comments.

References

- Balakrishnan N, Ng HKT, Navarro J (2011) Exact nonparametric inference for component lifetime distribution based on lifetime data from systems with known signatures. *J Nonparametric Stat* 23(3):741–752
- Balakrishnan N, Ng HKT, Navarro J (2011) Linear inference for type-II censored lifetime data of reliability systems with known signatures. *IEEE Trans Reliab* 60(2):426–440
- Bhattacharya D, Samaniego FJ (2010) Estimating component characteristics from system failure-time data. *Naval Res Logist (NRL)* 57(4):380–389
- Bickel PJ, Doksum KA (2001) *Mathematical statistics—basic ideas and selected topics*, vol I, 2nd edn. Prentice-Hall, Upper Saddle River
- Birnbaum ZW, Saunders SC (1969) A new family of life distributions. *J Appl Probab* 6(2):319–327
- Boyles R, Samaniego F (1987) On estimating component reliability for systems with random redundancy levels. *IEEE Trans Reliab* 36(4):403–407
- Bueno VC (1988) A note on the component lifetime estimation of a multistate monotone system through the system lifetime. *Adv Appl Probab* 20(3):686–689
- Chahkandi M, Ahmadi J, Baratpour S (2014) Non-parametric prediction intervals for the lifetime of coherent systems. *Stat Pap* 55(4):1019–1034
- Chen SX, Keilegom IV (2009) A review on empirical likelihood methods for regression. *TEST* 18:415–447
- Eryilmaz S, Koutras MV, Triantafyllou IS (2011) Signature based analysis of m-consecutive-k-out-of-n: F systems with exchangeable components. *Naval Res Logist (NRL)* 58(4):344–354
- Frenkel I, Khvatskin L (2006) Cost-effective maintenance with preventive replacement of oldest components. *W SKRÓCIE*, p 37
- Guess FM, Usher JS, Hodgson TJ (1991) Estimating system and component reliabilities under partial information on cause of failure. *J Stat Plan Inference* 29(1–2):75–85
- Hall P, Jin Y, Samaniego FJ (2015) Nonparametric estimation of component reliability based on lifetime data from systems of varying design. *Statistica Sinica* 1313–1335
- Jin Y, Hall PG, Jiang J, Samaniego FJ (2017) Estimating component reliability based on failure time data from a system of unknown design. *Statistica Sinica* 479–499
- Kocher S, Mukerjee H, Samaniego FJ (1999) The “signature” of a coherent system and its application to comparisons among systems. *Naval Res Logist (NRL)* 46(5):507–523
- Mann HB, Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat* 50–60
- Meeker WQ, Escobar LA, Pascual FG (2022) *Statistical methods for reliability data*, 2nd edn. Wiley, New York
- Meilijson I (1981) Estimation of the lifetime distribution of the parts from the autopsy statistics of the machine. *J Appl Probab* 18(4):829–838
- Miyakawa M (1984) Analysis of incomplete data in competing risks model. *IEEE Trans Reliab* 33(4):293–296
- Nadarajah T, Variyath AM, Loredó-Ostí JC (2014) Empirical likelihood based longitudinal data analysis. *Open J Stat* 10:611–639
- Navarro J (2022) *Introduction to system reliability theory*. Springer, Cham
- Navarro J, Rubio R (2009) Computations of signatures of coherent systems with five components. *Commun Stat Simul Comput* 39(1):68–84
- Navarro J, Ruiz JM, Sandoval CJ (2007) Properties of coherent systems with dependent components. *Commun Stat Theory Methods* 36(1):175–191
- Navarro J, Samaniego FJ, Balakrishnan N (2011) Signature-based representations for the reliability of systems with heterogeneous components. *J Appl Probab* 48(3):856–867
- Navarro J, Ng HKT, Balakrishnan N (2012) Parametric inference for component distributions from lifetimes of systems with dependent components. *Naval Res Logist (NRL)* 59(7):487–496
- Ng HKT, Navarro J, Balakrishnan N (2012) Parametric inference from system lifetime data under a proportional hazard rate model. *Metrika* 75(3):367–388
- Nordman DJ, Lahiri SN (2014) A review of empirical likelihood methods for time series. *J Stat Plan Inference* 155:1–18
- Owen AB (1988) Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* 75(2):237–249

- Owen AB (2001) Empirical likelihood. CRC Press, Boca Raton
- R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Samaniego FJ (1985) On closure of the IFR class under formation of coherent systems. *IEEE Trans Reliab* 34(1):69–72
- Samaniego FJ (2007) System signatures and their applications in engineering reliability, vol 110. Springer, Berlin
- Usher JS, Hodgson TJ (1988) Maximum likelihood analysis of component reliability using masked system life-test data. *IEEE Trans Reliab* 37(5):550–555
- Wilcoxon F (1992) Individual comparisons by ranking methods. *Breakthroughs in statistics*. Springer, Berlin, pp 196–202
- Wilks SS (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann Math Stat* 9(1):60–62
- Yang Y, Ng HKT, Balakrishnan N (2016) A stochastic expectation-maximization algorithm for the analysis of system lifetime data with known signature. *Comput Stat* 31(2):609–641
- Zhang J (2002) Powerful goodness-of-fit tests based on the likelihood ratio. *J R Stat Soc Ser B (Stat Methodol)* 64(2):281–294
- Zhang J (2006) Powerful two-sample tests based on the likelihood ratio. *Technometrics* 48(1):95–103
- Zhang J, Ng HKT, Balakrishnan N (2015) Tests for homogeneity of distributions of component lifetimes from system lifetime data with known system signatures. *Naval Res Logist (NRL)* 62(7):550–563
- Zhou M (2019) Empirical likelihood method in survival analysis. Chapman and Hall/CRC, Boca Raton

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.