Biometrics 73, 972–980
September 2017 DOI: 10.1111/biom.12658

Order-Restricted Inference for Means with Missing Values

Heng Wang* and Ping-Shou Zhong**

Department of Statistics and Probability, Michigan State University, Michigan, U.S.A.
**email:* hengwang@msu.edu
***email:* pszhong@msu.edu

Summary. Missing values appear very often in many applications, but the problem of missing values has not received much attention in testing order-restricted alternatives. Under the missing at random (MAR) assumption, we impute the missing values nonparametrically using kernel regression. For data with imputation, the classical likelihood ratio test designed for testing the order-restricted means is no longer applicable since the likelihood does not exist. This article proposes a novel method for constructing test statistics for assessing means with an increasing order or a decreasing order based on jackknife empirical likelihood (JEL) ratio. It is shown that the JEL ratio statistic evaluated under the null hypothesis converges to a chi-bar-square distribution, whose weights depend on missing probabilities and nonparametric imputation. Simulation study shows that the proposed test performs well under various missing scenarios and is robust for normally and nonnormally distributed data. The proposed method is applied to an Alzheimer's disease neuroimaging initiative data set for finding a biomarker for the diagnosis of the Alzheimer's disease.

Key words: Jackknife empirical likelihood; Kernel regression imputation; Missing values; Order-restricted inference.

1. Introduction

Order-restricted hypothesis testing problems are often encountered in many applications such as medical research, clinical trials and animal and plant breeding. In clinical trials, one is often interested in finding biomarkers for the diagnosis of different stages of a disease, so that doctors could prescribe appropriate treatments for patients. For example, Alzheimer's disease (AD) is an irreversible and progressive brain disorder, and is the most common type of dementia. In the ongoing Alzheimer's Disease Neuroimaging Initiative (ADNI) longitudinal study (`http://www.adni-info.org/`), scientists are interested in identifying biomarkers to differentiate three progressive stages of the AD: cognitive normal (CN), late mild cognitive impairment (LMCI), and AD. A biomarker that can be used for the diagnosis should be responsive to different stages of the disease. More specifically, one may wish to test whether the value of a biomarker is increasing or decreasing as the disease progresses. One of the important biomarker candidates in the ADNI study is the cerebral metabolic rates of glucose (CMRglc), a proxy for neuronal activity in AD (Mosconi et al., 2010). The CMRglc is measured by fluorodeoxyglucose (FDG) positron emission tomography (PET) scanning in the ADNI study. A larger FDG value usually indicates a higher level of CMRglc, which is associated with higher neural activity level. The motivation of this article is to determine if CMRglc is an appropriate biomarker for the diagnosis of the progression of AD. To be more specific, assume that y_{ij} is the FDG value for the j-th patient in the i-th group where $i = 1$ for AD, $i = 2$ for LMCI, and $i = 3$ for CN and assume that $E(y_{ij}) = \theta_i$. To find out if CMRglc is an appropriate biomarker, one might be interested in testing the following hypothesis

$$H_0 : \theta_1 = \theta_2 = \theta_3 \quad \text{vs.} \quad H_1 : \theta_1 \le \theta_2 \le \theta_3$$
$$\text{with at least one strict inequality holds.} \qquad (1)$$

The above alternative indicates that the mean of FDG increases as AD gets less severe. Rejecting the null hypothesis means that CMRglc is a good biomarker for AD.

One difficulty in testing the above hypothesis in the ANDI data set is the large number of missing values. Considering the data collected between 2005 and 2007, more than 50% of data were missing for each group (AD, LMCI, and CN). Simply deleting the subjects with missing values will result in a significant loss of data and will lead to inefficient or biased statistical inference (Kim and Shao, 2013). Imputation is a common approach to deal with missing values. Commonly used imputation methods include multiple imputation (Rubin, 1987), hot deck imputation (Fuller and Kim, 2005) and nonparametric kernel regression imputation (Cheng, 1994). In this article, we consider applying the kernel regression to impute the missing values due to its robustness. To the best of our knowledge, no formal method has been developed for testing order restricted hypothesis (1) when part of the data are missing and replaced by imputed values. This motivates us to develop a method for testing the order-restricted hypothesis in (1) for data sets with imputation.

For data without missing values, the hypothesis testing in (1) has been well studied in the literature. Bartholomew (1959a,b, 1961a,b) proposed likelihood ratio tests for the above problem. Under the normality assumption, it was shown that the likelihood ratio test statistic for monotonic alternatives follows a chi-bar-square distribution. Some of recent work on order-restricted inference using likelihood ratio test includes Nettleton (1999, 2009). An excellent review on order-restricted inference can be found in Sivapulle and Sen (2004). However, after the nonparametric imputation, the imputed data do not have a likelihood and hence the parametric likelihood ratio-based methods are not applicable. To relax the parametric distribution assumption in the likelihood ratio test, one could consider using empirical likelihood (EL) method, which is a nonparametric likelihood proposed by Owen (1988, 1990) and does not need any parametric distribution assumption on the data. On one hand, for testing omnibus alternatives without order restrictions, EL has been applied to data with nonparametric imputation, which include Wang and Rao (2002) and Wang and Chen (2009). Both results show that the EL ratios for imputed data are asymptotically distributed as a linear combination of chi-squares rather than a standard chi-square distribution. On the other hand, for testing order restricted alternatives without missing values, El Barmi (1996) ingeniously applied EL method to construct test statistics. Without assuming a specific distribution for data, they have shown that the EL ratio test statistic has the same asymptotic chi-bar-square distribution as the parametric likelihood ratio test statistic. Recently, Davidov et al. (2010, 2014) applied EL to order-restricted semi-parametric inference. However, all the above mentioned procedures either fail to consider order-restricted alternatives or do not consider the problem of missing values. Thus, we are not able to apply the existing procedures to the ANDI study.

A major obstacle that prevents the extension of EL to test order-restricted alternatives for data with nonparametric imputation is the difficulty in obtaining the asymptotic distribution of the test statistic. In order to obtain the asymptotic distribution of the EL-based test statistic for order-restricted alternatives, one of the key steps is to show the asymptotic independence between the projected EL statistics and the event of the projection. However, we find that the asymptotic independence can not be established for the EL-based methods. Therefore, it is difficult to obtain the asymptotic distribution of the EL-based methods.

This article introduces a novel Jackknife EL (JEL) ratio method for testing hypothesis (1). JEL was first introduced by Jing et al. (2009) to overcome the computation difficulty in EL due to the nonlinear estimating equations. In this article, we extend the use of JEL to test order-restricted alternatives for data with nonparametric imputation. We found that, under the null hypothesis, the proposed JEL ratio statistics still have a chi-bar-square distribution. But the weights of the chi-bar-square distribution depend on missing probabilities and nonparametric imputation. The result is valuable in the following two aspects. First, by combining the EL and Jackknife pseudo values, deriving the asymptotic distribution of JEL ratio statistic becomes feasible, which turns out to be asymptotically chi-bar-square distribution; Second, it provides a simple and formal procedure for testing order-restricted hypothesis for data with nonparametric imputation, which has not been considered in the existing literature.

The rest of this article is organized as follows. In Section 2, we introduce the basic setup and the problem of interest. JEL ratio statistic is presented in Section 3. The asymptotic null distribution is provided in Section 4. Extensive simulation studies are provided in Section 5. In Section 6, we applied our proposed procedure to the ADNI data set mentioned at the beginning of the introduction. Some concluding remarks are given in Section 7. All the technical details and additional simulation results are relegated to the web-based supplementary material.

2. Basic Setting and Hypothesis of Interest

Let y_{ij} be the univariate quantitative measurement for the j-th subject in the i-th group, where $i = 1, \ldots, k$ and $j = 1, \ldots, n_i$. Let δ_{ij} be the indicator of nonmissingness such that $\delta_{ij} = 1$ if y_{ij} is observed and $\delta_{ij} = 0$ if y_{ij} is missing. In this study, we assume missing at random (MAR), which assumes that, given a d-dimensional covariate $\boldsymbol{x}_{ij}$, y_{ij} is independent of the non-missing indicator δ_{ij}, that is,

$$P(\delta_{ij} = 1|\boldsymbol{x}_{ij}, y_{ij}) = P(\delta_{ij} = 1|\boldsymbol{x}_{ij}) := R_i(\boldsymbol{x}_{ij}).$$

We use the triplet $(y_{ij}, \boldsymbol{x}_{ij}, \delta_{ij})$ to denote all the data we observed for the j-th individual within the i-th group, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$. We impute the missing value nonparametrically using a nonparametric regression (Cheng, 1994; Wang and Rao, 2002; Zhong and Chen, 2014). Specifically, we replace the missing values y_{ij} by the Nadaraya–Watson (NW) estimator $\hat{m}_i(\boldsymbol{x}_{ij})$ for $m_i(\boldsymbol{x}_{ij}) = E(y_{ij}|\boldsymbol{x}_{ij})$, where $\hat{m}_i(\boldsymbol{x}_{ij})$ is defined by

$$\hat{m}_i(\boldsymbol{x}_{ij}) = \frac{\sum_{j'=1, j' \neq j}^{n_i} \delta_{ij'} y_{ij'} K_{h_i}(\boldsymbol{x}_{ij}, \boldsymbol{x}_{ij'})}{\sum_{j'=1, j' \neq j}^{n_i} \delta_{ij'} K_{h_i}(\boldsymbol{x}_{ij}, \boldsymbol{x}_{ij'})}. \quad (2)$$

Here, $K_{h_i}(\boldsymbol{x}_{ij}, \boldsymbol{x}_{ij'}) = h_i^{-d} K\left\{(\boldsymbol{x}_{ij} - \boldsymbol{x}_{ij'})/h_i\right\}$, and K is a kernel function and h_i's are bandwidths. When the dimension of covariates d is large, the NW imputation might be affected by the curse of dimensionality. In this case, we propose to use a nonparametric additive regression model to perform the nonparametric imputation. The details of the extension and some simulation experiments are provided in the web Appendix C of the web-based supplementary material.

Suppose y_{ij}'s are independent identically distributed (IID) random variables generated from some distribution $\mathscr{F}_i$ ($i = 1, \ldots, k$). Assume $\theta_i = E(y_{ij})$ ($j = 1, \ldots, n_i$) is the mean value for the measurements y_{ij}'s from the i-th group. The aim of this article is to test

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_k \quad \text{vs.} \quad H_1 : \theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$$
$$\text{with at least one strict inequality holds,} \quad (3)$$

when the missing responses are imputed using the nonparametric kernel regression.

3. Jackknife EL Ratio Statistics

We consider JEL ratio statistics for testing (3). To this end, we first define the pseudo-values for estimating θ_i's. Based on the observed y_{ij}'s with $\delta_{ij} = 1$ and the imputed $\hat{m}_i(\boldsymbol{x}_{ij})$'s when $\delta_{ij} = 0$ ($i = 1, \ldots, k$; $j = 1, \ldots, n_i$), a consistent estimate of the group mean θ_i is

$$\hat{\theta}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\{\delta_{ij}y_{ij} + (1-\delta_{ij})\hat{m}_i(\boldsymbol{x}_{ij})\} \text{ for } i = 1, \ldots, k. \quad (4)$$

Let $\tilde{y}_{ij}$ be the response after imputation, that is, $\tilde{y}_{ij} = y_{ij}$ if $\delta_{ij} = 1$, and $\tilde{y}_{ij} = \hat{m}_i(\boldsymbol{x}_{ij})$ if $\delta_{ij} = 0$. Then $\hat{\theta}_i = \sum_{j=1}^{n_i}\tilde{y}_{ij}/n_i$ is an average of the response after imputation.

Following Zhong and Chen (2014), we define jackknife pseudo-values ν_{ij}'s as $\nu_{ij} = n_i\hat{\theta}_i - (n_i-1)\hat{\theta}_i^{(-j)}$, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$, where $\hat{\theta}_i^{(-j)}$ is the estimate for θ_i defined by (4) without using the j-th subject in the i-th group. If no missing value exists, it is clear that $\nu_{ij} = \tilde{y}_{ij} = y_{ij}$. However, if missing values exist, ν_{ij} is not always equal to $\tilde{y}_{ij}$, which shows the impact of missing values on the pseudo-values. Specifically, if j corresponds to $\delta_{ij} = 0$, ν_{ij} is asymptotically the same as $\tilde{y}_{ij} = \hat{m}_i(\boldsymbol{x}_{ij})$. But if j corresponds to $\delta_{ij} = 1$, ν_{ij} is asymptotically the same as $y_{ij} + e_n(\boldsymbol{x}_{ij})$ but not the same as $\tilde{y}_{ij} = y_{ij}$, where $e_n(\boldsymbol{x}_{ij}) = \{(n_i-1)G(\boldsymbol{x}_{ij})\}^{-1}\sum_{l=1,l\neq j}^{n_i}(1-\delta_{ij})\delta_{il}K_{h_i}(\boldsymbol{x}_{il}, \boldsymbol{x}_{ij})\{y_{ij} - m_i(\boldsymbol{x}_{ij})\}$ and $G(\boldsymbol{x}) = E\{\delta_{ij'}K_{h_i}(\boldsymbol{x}, \boldsymbol{x}_{ij'})|\boldsymbol{x}\}$. The adjustment term $e_n(\boldsymbol{x}_{ij})$ is important, which makes the pseudo-values ν_{ij}'s different from the imputed responses $\tilde{y}_{ij}$'s.

The JEL function for θ_i is defined (Jing et al., 2009; Zhong and Chen, 2014) as

$$\mathcal{L}(\theta_i) = \max_{p_{ij}}\left\{\prod_{j=1}^{n_i} p_{ij} : p_{ij} > 0,\ j = 1, \ldots, n_i, \right.$$
$$\left.\sum_{j=1}^{n_i} p_{ij} = 1,\ \sum_{j=1}^{n_i} p_{ij}\nu_{ij} = \theta_i\right\}.$$

Using the standard derivation, for a given θ_i, it can be shown that the maximization of $\mathcal{L}(\theta_i)$ is achieved at $\hat{p}_{ij} = {n_i}^{-1}\{1 + \lambda_i(\theta_i)(\nu_{ij} - \theta_i)\}^{-1}$ for $j = 1, \ldots, n_i$, where $\lambda_i(\theta_i)$'s are Lagrange multipliers satisfying $\sum_{j=1}^{n_i}(\nu_{ij} - \theta_i)/\{1 + \lambda_i(\theta_i)(\nu_{ij} - \theta_i)\} = 0$. Therefore, the log-JEL function for θ_i is

$$\ell_i(\theta_i) = -\sum_{j=1}^{n_i}\log\{1 + \lambda_i(\theta_i)(\nu_{ij} - \theta_i)\} - n_i\log(n_i).$$

Because the k samples are independent, the log-JEL function for $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)^T$ is $\ell(\boldsymbol{\theta}) = \sum_{i=1}^{k}\ell_i(\theta_i)$. Then the log-JEL ratio statistic for testing H_0 against H_1 in equation (3) is

$$\begin{aligned}\log R_n &= \max_{H_0}\ell(\theta_1, \ldots, \theta_k) - \max_{H_1}\ell(\theta_1, \ldots, \theta_k)\\ &= \min_{H_1}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\log\{1 + \lambda_i(\theta_i)(\nu_{ij} - \theta_i)\}\\ &\quad - \min_{H_0}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\log\{1 + \lambda_i(\theta_i)(\nu_{ij} - \theta_i)\}.\end{aligned}$$

It follows that the test statistic for testing H_0 versus H_1 can be defined as $\Lambda_n = -2\log R_n$.

Next, we wish to make the JEL ratio test (JELRT) statistic Λ_n more explicitly computable. Under the null space $H_0 : \theta_1 = \theta_2 = \cdots = \theta_k$, θ_i's have a common value θ_0, which can be estimated by the maximum JEL estimator (MJELE) $\hat{\theta}_0$, where

$$\hat{\theta}_0 = \arg\min_{\theta_0}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\log\{1 + \lambda_i(\theta_0)(\nu_{ij} - \theta_0)\}.$$

The corresponding $\hat{\lambda}_i := \lambda_i(\hat{\theta}_0)$ satisfies $\sum_{j=1}^{n_i}(\nu_{ij} - \hat{\theta}_0)/\{1 + \hat{\lambda}_i(\nu_{ij} - \hat{\theta}_0)\} = 0$. Similar to Qin and Lawless (1995), it can be shown that in a neighborhood of the true value of θ_0, there almost surely exists a unique minimizer $\hat{\theta}_0$ that minimizes the objective function $-\ell(\boldsymbol{\theta})$.

To obtain MJELE under the alternative space $H_1 : \theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$, we introduce Lagrange multipliers $\alpha_1, \ldots, \alpha_{k-1}$ and define a new function $f(\theta_1, \ldots, \theta_k, \alpha_1, \ldots, \alpha_{k-1}) = -\ell(\theta_1, \ldots, \theta_k) - \sum_{i=1}^{k-1}\alpha_i(\theta_{i+1} - \theta_i)$. Then the minimizer θ_i's $(i = 1, \ldots, k)$ and α_j's $(j = 1, \ldots, k-1)$ of the function f satisfy the following Karush–Kuhn–Tucker conditions:

$$-\frac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_i} + \alpha_i - \alpha_{i-1} = 0,\ \ \alpha_j(\theta_{j+1} - \theta_j) = 0,\ \text{ and }\ \alpha_j \geq 0,$$

where $\alpha_0 = 0$ and $\alpha_k = 0$. There are $2k-1$ equations in the above estimating equations. Denote the minimizer for $\theta_1, \ldots, \theta_k$ as $\tilde{\theta}_1, \ldots, \tilde{\theta}_k$, and they are the MJELE's for $\theta_1, \ldots, \theta_k$ under H_1. Similar to the null case, the corresponding solutions for λ_i's under H_1 are defined as $\tilde{\lambda}_i := \lambda_i(\tilde{\theta}_i)$. Therefore, the JEL ratio test statistic for H_0 versus H_1 is

$$\Lambda_n = 2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[\log\{1 + \hat{\lambda}_i(\nu_{ij} - \hat{\theta}_0)\} - \log\{1 + \tilde{\lambda}_i(\nu_{ij} - \tilde{\theta}_i)\}\right].$$

4. Asymptotic Null Distribution of Λ_n

In this section, we present the asymptotic distribution of Λ_n under the H_0. For convenience, we rewrite the alternative hypothesis as $H_1 : \boldsymbol{A\theta} \leq 0$ where $\boldsymbol{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_{k-1})^T$ is a $(k-1)\times k$ matrix and $\mathbf{a}_l = (0, \ldots, 0, 1, -1, 0, \ldots, 0)^T$ $(1 \leq l \leq k-1)$ is a k-dim vector whose l and $l+1$ components are 1 and -1, respectively. Let $\mathscr{F}$ be the collection of all the subsets of $\{1, \ldots, k-1\}$ and π be any set in $\mathscr{F}$ with cardinality $|\pi| \leq k-1$. Define $\boldsymbol{A}(\pi)$ as a submatrix of $\boldsymbol{A}$ with rows determined by the set π. So $\boldsymbol{A}(\pi)$ is a $|\pi|\times k$ matrix. For example, if $k \geq 3$ and $\pi_0 = \{1, 2\}$, $\boldsymbol{A}(\pi_0)$ is defined as $\boldsymbol{A}(\pi_0) = (\mathbf{a}_1, \mathbf{a}_2)^T$. Now, denote the $\tilde{\boldsymbol{\theta}}(\pi)$ as the MJELE of $\boldsymbol{\theta}$ under the constraints $H_1(\pi) : \boldsymbol{A}(\pi)\boldsymbol{\theta} = 0$. Meanwhile, define $\Lambda_n(\pi)$ as the JELRT for testing H_0 versus $H_1(\pi)$. Specifically,

$$\begin{aligned}\Lambda_n(\pi) = 2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\Big[&\log\{1 + \hat{\lambda}_i(\nu_{ij} - \hat{\theta}_0)\}\\ &- \log\{1 + \tilde{\lambda}_i(\pi)(\nu_{ij} - \tilde{\theta}_i(\pi))\}\Big].\end{aligned}$$

According to El Barmi (1996) or Wollan and Dykstra (1986), $\tilde{\boldsymbol{\theta}}$ equals $\tilde{\boldsymbol{\theta}}(\pi)$ for one specific $\pi \in \mathscr{F}$. This implies that, for any t,

$$P(\Lambda_n \geq t) = \sum_{\pi \in \mathscr{F}} P\{\Lambda_n \geq t | \tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\} P\{\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\}$$
$$= \sum_{\pi \in \mathscr{F}} P\{\Lambda_n(\pi) \geq t | \tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\} P\{\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\}, \quad (5)$$

where the second equality is true due to the fact that $\Lambda_n = \Lambda_n(\pi)$ under the condition $\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)$. The equation (5) is the key for obtaining the asymptotic distribution of Λ_n under H_0. Combining equation (5) with the following Theorems 1 and 2, we can obtain the asymptotic distribution of Λ_n.

Denote $N = \sum_{i=1}^{k} n_i$ as the total of the sample sizes and the proportion as $\rho_i = \lim_{N\to\infty} n_i/N$. Assume that ρ_i is a positive constant.

THEOREM 1. *Under conditions (C1)–(C4) in the web Appendix A of the supplementary material, we have* $\lim_{N\to\infty} P\{\Lambda_n(\pi) \geq t | \tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\} = P(\chi^2_{k-1-|\pi|} \geq t)$ *where* $\chi^2_{k-1-|\pi|}$ *denotes a chi-square distribution with degrees of freedom* $k - 1 - |\pi|$.

The proof of Theorem 1 is provided in the web Appendix A of the supplementary material. Theorem 1 implies that the event $\{\Lambda_n(\pi) \geq t\}$ is asymptotically independent of the event $\{\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\}$. Moreover, the asymptotic distribution of $\Lambda_n(\pi)$ is a chi-square distribution with degrees of freedom $k - 1 - |\pi|$. The results in Theorem 1 have shown to be true for EL ratio statistics for data without missing values (El Barmi, 1996). Our results indicate that such results also hold for JEL ratio statistics for data with nonparametric imputation.

REMARK 1. *The asymptotic distribution of the test statistic* Λ_n *crucially depends on the asymptotic independence between* $\{\Lambda_n(\pi) \geq t\}$ *and* $\{\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\}$. *However, the asymptotic independence does not hold for EL-based test statistics. A detailed discussion of this point is provided in the web Appendix B of the supplementary material.*

Combining Theorem 1 with equation (5), we have

$$\lim_{N\to\infty} P(\Lambda_n \geq t) = \lim_{N\to\infty} \sum_{\pi \in \mathscr{F}} P\{\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\} P(\chi^2_{k-1-|\pi|} \geq t). \quad (6)$$

In the following Theorem 2, we evaluate the asymptotic probability of the event $\{\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\}$. To this end, we define some notation. For $i = 1, \ldots, k$, let $\sigma_i^2 = E\{\sigma_i^2(\boldsymbol{x})/R_i(\boldsymbol{x})\} + \mathrm{Var}\{m_i(\boldsymbol{x})\}$, where $\sigma_i^2(\boldsymbol{x}) = \mathrm{Var}(y_{ij}|\boldsymbol{x})$, $E(y_{ij}|\boldsymbol{x}) = m_i(\boldsymbol{x})$ and $R_i(\boldsymbol{x})$ are the conditional missing probability defined in Section 2.1. Define $\boldsymbol{V} = \mathrm{Diag}(\sigma_1^2/\rho_1, \ldots, \sigma_k^2/\rho_k)$. If $\pi = \emptyset$, define $p_1(\pi) = P\{\mathrm{MVN}(0, \boldsymbol{\Sigma}_1(\pi)) \geq 0\}$ and $p_2(\pi) = 1$, where MVN represents a multivariate normal distribution and $\boldsymbol{\Sigma}_1(\pi) = \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}^T$. For any $\pi \neq \emptyset$ and $\pi \neq \{1, \ldots, k-1\}$, define $p_1(\pi) = P\{\mathrm{MVN}(0, \boldsymbol{\Sigma}_1(\pi)) \geq 0\}$ and $p_2(\pi) = P\{\mathrm{MVN}(0, \boldsymbol{\Sigma}_2(\pi)) \geq 0\}$ where $\boldsymbol{\Sigma}_1(\pi) = \{\boldsymbol{A}(\pi)\boldsymbol{V}\boldsymbol{A}^T(\pi)\}^{-1}$ and $\boldsymbol{\Sigma}_2(\pi) = \boldsymbol{A}(\pi^c)\{\boldsymbol{V} - \boldsymbol{V}\boldsymbol{A}^T(\pi)\boldsymbol{\Sigma}_1(\pi)\boldsymbol{A}(\pi)\boldsymbol{V}\}\boldsymbol{A}^T(\pi^c)$, where π^c is the complement set of π.

THEOREM 2. *Under conditions (C1)–(C4) in the web Appendix A of the supplementary material, for any* $\pi \neq \{1, \ldots, k-1\}$, *we have* $\lim_{N\to\infty} P\{\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\pi)\} = p_1(\pi)p_2(\pi)$, *where* $p_1(\pi)$ *and* $p_2(\pi)$ *are defined above. Moreover, the JEL ratio statistic* Λ_n *has the following asymptotic distribution under* H_0,

$$\lim_{N\to\infty} P(\Lambda_n \geq t) = \sum_{j=0}^{k-1} w_j P(\chi^2_{k-1-j} \geq t), \quad (7)$$

where $w_j = \sum_{|\pi|=j} p_1(\pi)p_2(\pi)$ *for* $0 \leq |\pi| = j < k-1$ *and* $w_{k-1} = 1 - \sum_{j=1}^{k-2} w_j$.

The asymptotic results in Theorem 2 seem to be similar to the results obtained for likelihood ratio statistics and EL ratios for data without imputations in El Barmi (1996). But they are actually different distributions since the weights in the chi-bar-square distribution are different from that in El Barmi (1996). Due to missing values, the weights w_j in equation (7) depend on σ_i^2's, which is a function of the missing probabilities $R_i(\boldsymbol{x})$ and variances of the conditional mean of the imputed values $m_i(\boldsymbol{x})$. For different imputation methods, the asymptotic variances of pseudo values are different. As a result, the asymptotic distributions of the JEL test statistics are different if different imputation methods are applied. Similar phenomena have been found by Wang and Rao (2002), and Wang and Wang (2006). They demonstrated that the asymptotic distributions of EL ratio statistics for testing the omnibus alternatives are different for data with parametric and nonparametric imputation. In the ideal case, if the missing probability is 0, $\sigma_i^2 = \mathrm{Var}(y_{ij})$, then our result is the same as that obtained in El Barmi (1996). Moreover, our result is derived for the JEL ratios rather than the EL ratios. If the EL ratio is used, the asymptotic distribution would be very difficult to obtain as we discussed in Remark 1.

Let us consider a special case with $k = 3$, $n_1 = n_2 = n_3 = n$ and $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$. Then $\boldsymbol{V} = \mathrm{Diag}(3\sigma^2, \ldots, 3\sigma^2)$. If $\pi = \emptyset$, it can be checked that

$$\boldsymbol{A}\boldsymbol{V}\boldsymbol{A}^T = 3\sigma^2 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

As a result, $p_1(\pi) = 1/6$ and $w_0 = 1/6$. If $\pi = \{1\}$ or $\pi = \{2\}$, then it can be easily shown that $p_1(\pi) = p_2(\pi) = 1/2$. Then, $w_1 = 1/4 + 1/4 = 1/2$. Thus, $w_2 = 1 - w_0 - w_1 = 1/3$. In this case, the null distribution of the JEL ratio test statistic follows a chi-bar-square distribution, that is, $\bar{\chi}^2 = (1/6)\chi_2^2 + (1/2)\chi_1^2 + (1/3)\chi_0^2$, where χ_0^2 is the degenerate random variable taking value 0. In this case, we reject the null hypothesis if $\Lambda_n > \bar{\chi}^2_\alpha$, where $\bar{\chi}^2_\alpha$ is the upper α-quantile of $\bar{\chi}^2$ distribution. In general case, the weights w_j's do not have a simple form. The α-quantile of the chi-bar-square distribution can be obtained by using simulation.

5. Simulation Study

In this section, we present a series of simulation studies designed to demonstrate finite sample performance of the proposed JEL ratio test. Additional simulation results are presented in the web Appendix D of supplementary material, where we compared the proposed method with a test based on the asymptotic normality of $\hat{\theta}_i$'s and a JEL ratio test using parametric imputation.

We simulated independent random variables y_{ij} from normal distribution with mean $\theta_i + ax_{ij}$ and variance 1 for three groups ($i = 1, 2$ and 3) and $j = 1, \ldots, n_i$. The nonmissing indicators δ_{ij} were generated from a Bernoulli distribution with mean p_{ij} where $p_{ij} = \exp(t_{ij})/\{1 + \exp(t_{ij})\}$ and $t_{ij} = b_0 + 2x_{ij}$. The covariates x_{ij} were generated from the standard normal.

First, we considered a balanced case where the sample sizes n_i (i =1, 2, 3) were the same across three treatment groups (i.e., $n_1 = n_2 = n_3 = n$). In addition, we assumed b_0 to be the same among three groups. In the simulation, we chose $n =$ 100, 200, and 400. To check the impact of missing values, two values for b_0 were used. Specifically, $b_0 = 0.5$ and 1 were used. When $b_0 = 0.5$, the missing probability was around 43%. When $b_0 = 1$, the missing probability was around 35%. For the balanced case, according to the special case discussed after Theorem 2, the null distribution of the JEL ratio test statistic follows a chi-bar-square distribution, that is $\bar{\chi}^2 = (1/6)\chi^2_2 + (1/2)\chi^2_1 + (1/3)\chi^2_0$.

For evaluating the empirical size of the proposed test, we set $\theta_1 = \theta_2 = \theta_3 = 0$ under the null hypothesis. For evaluating the power of the test, we designed three alternative scenarios satisfying $\theta_1 < \theta_2 < \theta_3$: scenario A: $\theta_1 = 0, \theta_2 = 0.125, \theta_3 = 0.25$; scenario B: $\theta_1 = 0, \theta_2 = 0.25, \theta_3 = 0.5$; and scenario C: $\theta_1 = 0, \theta_2 = 0.5, \theta_3 = 1$. All the simulation results reported in this section were based on 1000 simulation replicates. The nominal level was 0.05 for all the tests.

To demonstrate the impact of imputation on the existing methods for testing order restricted hypothesis, we compared the proposed method with the likelihood ratio (LR) test (Bartholomew, 1961a) and the EL ratio test (El Barmi, 1996) without considering the missing values. Namely, we applied the LR test and EL ratio test by considering the imputed values as observed data.

Table 1 summarizes the empirical sizes and powers of the proposed test, and the empirical sizes of the LR test and EL ratio test. We observe that the imputation has a big impact on the performance of the LR test and EL ratio test. As we can see from Table 1, the empirical sizes of the LR test and EL ratios test are quite far away from the nominal level 0.05. The results are not surprising because Bartholomew (1961a) and El Barmi (1996)'s methods are designed for data sets without missing values. These results demonstrate that considering imputation is necessary and imputed values can not be treated as the observed values.

On the other hand, we observe that the proposed JEL method controls type I error well around the nominal level for both $b_0 = 0.5$ and $b_0 = 1$. Under the alternatives, the power of JEL method increases as n increasing from $n = 100$ to $n = 400$. The power also increases from scenario A to scenario C, since the distance between the alternatives and null increases. The bandwidths in the simulation are chosen according to $h_i = c_i n_i^{-11/40}$, which satisfies the condition (C2) in the web Appendix A of the supplementary material and $c_i = c\hat{\sigma}_{x_i}$ where $\hat{\sigma}_{x_i}$ is the sample standard deviation of x_{ij}'s within the i-th group. Here, c is a constant between 2 and 4. The proposed method is not very sensitive to the choice of the constant c because the bandwidths do not have a leading order impact on the mean squares of $\hat{\theta}_i$ and $\tilde{\theta}_i$ (for $i = 1, 2$ and 3). The above phenomena are similar to those observed in Zhong and Chen (2014).

Table 1
Empirical sizes and powers for the proposed JEL method for balanced data generated from normal distributions

			Size			JEL power		
	n	a	LR	EL	JEL	A	B	C
$b_0 = 0.5$	100	0.6	0.269	0.314	0.056	0.161	0.337	0.805
	100	0.8	0.234	0.280	0.051	0.146	0.336	0.793
	200	0.6	0.254	0.310	0.049	0.234	0.555	0.976
	200	0.8	0.261	0.317	0.048	0.222	0.520	0.960
	400	0.6	0.142	0.216	0.049	0.361	0.791	1.000
	400	0.8	0.151	0.233	0.054	0.343	0.756	1.000
	n	a	LR	EL	JEL	A	B	C
$b_0 = 1$	100	0.6	0.123	0.208	0.048	0.171	0.369	0.839
	100	0.8	0.161	0.252	0.052	0.156	0.367	0.813
	200	0.6	0.227	0.272	0.057	0.243	0.588	0.986
	200	0.8	0.209	0.262	0.053	0.221	0.584	0.986
	400	0.6	0.152	0.226	0.060	0.409	0.861	1.000
	400	0.8	0.111	0.177	0.054	0.377	0.829	1.000

The empirical powers for the JEL method were evaluated under three scenarios A, B and C defined above. The empirical sizes for the likelihood ratio (LR) test and the EL ratio test are also reported.

To further evaluate the impact of missing values on the classical LR test, we applied the LR test to the complete data with missing values removed from the original data sets. The data were generated according to the same model as that used in Table 1 with $b_0 = 0.5$. We compared the LR test with the proposed JEL test under the null hypothesis for evaluating the empirical sizes. The results were demonstrated using quantile-quantile plots in Figure 1. In the left panel of Figure 1, we plotted the quantile of LR test statistics for data with missing values deleted, versus the theoretical quantile for LR test statistics obtained by Bartholomew (1961a) where missing values were not considered. As we observed from the left panel of Figure 1, these two quantiles were quite far away from each other. These results indicated that the missing values have significant impact on the LR test and we cannot simply delete the missing values. On the other hand, we can see that our proposed JEL method performed reasonably well as demonstrated in right panel of Figure 1.

Next, we considered simulations with unbalanced samples and different missing probabilities among three samples. First, we considered six unbalanced combinations for (n_1, n_2, n_3): (100, 100, 200), (200, 100, 100), (100, 100, 400), (400, 100, 100), (100, 200, 300), and (300, 200, 100). Secondly, different missing probabilities were considered across three groups. For a higher missing probability setting, the b_0's for the three groups were 0.25, 0.5, and 1, so the missing probabilities were around 46, 43, and 35%, respectively. For a lower missing probability setting, the b_0's were 0.5, 1, and 2, so the missing probabilities for the three groups were around 43, 35, and

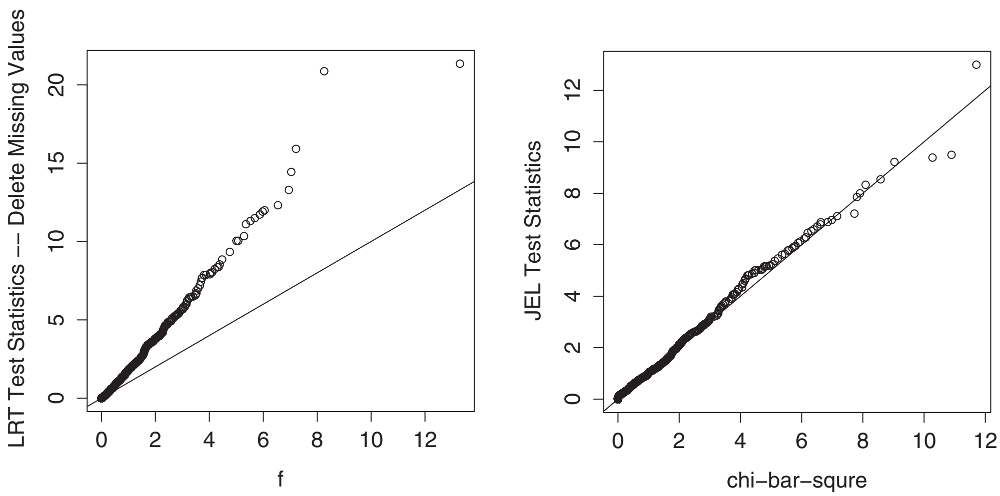

Figure 1. Quantile-quantile (QQ) plots of LR test with the missing values deleted (left panel) and the QQ plot of the proposed JEL test for data with imputation (right panel).

23%, respectively. The alternatives scenarios were the same as the alternatives in the balanced case. Namely, we considered three scenarios A, B, and C for evaluating the power of the proposed test.

Because neither LR test nor EL ratio test maintained the nominal level correctly in the balanced situation, we only considered JEL ratio test in the simulations with unbalanced samples. The asymptotic distribution of the proposed JEL ratio test was derived in Theorem 2, but the weights w_j do not have a simple expression. Therefore, we used computational method to estimate the weights w_j. The bandwidths were chosen similarly to the balanced cases.

The simulation results for unbalanced cases are summarized in Table 2. We can observe that the empirical type I errors are well controlled under the null hypothesis. As the differences between null and alternatives getting large, the powers of the proposed test increased to 1. This demonstrates the consistency of the proposed test.

To investigate the robustness of the proposed procedure in terms of distributions of responses, we also considered chi-square distributions in our simulation studies. Similar to the normal case, we considered the balanced and unbalanced cases. In both situations, we used the similar simulation setups as the normal cases except that we changed the distribution for generating the response. More specifically, the responses were generated using the model $y_{ij} = (\theta_i + ax_{ij} + \varepsilon_{ij})/\sqrt{8}$. Here, x_{ij} are independent standard normal, which are independent of ε_{ij}, and ε_{ij} are centralized chi-square distributed with degrees of freedom 4.

Table 3 summarizes the simulation results for balanced case with data generated from chi-square distributions. The proposed method was also compared with the LR test and the EL ratio test. Similar to the observation in Table 1, the empirical type I errors of the LR test and the EL ratio test were not well controlled. However, the proposed method can still control the type I error well even the distribution is not normally distributed any more. We can see that as the sample size increased, the power of the proposed test increased. As the alternatives deviated from the null, the powers of the proposed test increased.

The simulation results for chi-square distributed random variables with unbalanced samples are summarized in Table 4. The patterns in Table 4 are very similar to the patterns in Table 2. However, it is worth pointing out that the pro-

Table 2

Empirical sizes and powers of JEL ratio test for order restricted hypothesis with unbalanced samples and different missing probabilities among groups

	Higher missing prob.				Lower missing prob.			
(n_1, n_2, n_3)	Size	A	B	C	Size	A	B	C
(100, 100, 200)	0.050	0.340	0.772	1.000	0.055	0.343	0.802	1.000
(200, 100, 100)	0.058	0.415	0.847	1.000	0.051	0.402	0.846	1.000
(100, 100, 400)	0.056	0.348	0.804	1.000	0.047	0.401	0.853	1.000
(400, 100, 100)	0.073	0.535	0.935	1.000	0.055	0.530	0.943	1.000
(100, 200, 300)	0.048	0.378	0.812	1.000	0.042	0.425	0.857	1.000
(300, 200, 100)	0.064	0.473	0.927	1.000	0.049	0.460	0.920	1.000

The data are normally distributed in this table.

Table 3
Empirical sizes and powers for the proposed JEL method for balanced data generated from chi-square distributions

			Size			JEL power		
	n	a	LR	EL	JEL	A	B	C
$b_0 = 0.5$	100	0.6	0.276	0.285	0.054	0.104	0.207	0.501
	100	0.8	0.295	0.301	0.069	0.110	0.190	0.495
	200	0.6	0.342	0.339	0.058	0.131	0.318	0.775
	200	0.8	0.332	0.347	0.063	0.120	0.297	0.742
	400	0.6	0.236	0.220	0.056	0.206	0.505	0.932
	400	0.8	0.232	0.226	0.052	0.193	0.471	0.935
	n	a	LR	EL	JEL	A	B	C
$b_0 = 1$	100	0.6	0.246	0.252	0.053	0.104	0.213	0.539
	100	0.8	0.242	0.258	0.053	0.114	0.189	0.506
	200	0.6	0.296	0.296	0.062	0.158	0.319	0.800
	200	0.8	0.273	0.295	0.048	0.154	0.316	0.794
	400	0.6	0.243	0.235	0.060	0.216	0.519	0.968
	400	0.8	0.193	0.185	0.050	0.230	0.515	0.973

The empirical powers for the proposed JEL method were evaluated under three scenarios A, B and C defined above. The empirical sizes for the likelihood ratio (LR) test and the EL ratio test are also reported.

posed method can maintain the type I error correctly under all the combinations of sample sizes with chi-square distributed random variables.

6. An Application to ADNI Study

In this section, we apply our proposed method to the motivation example given in the introduction part. AD is a common type of dementia, which decreases people's memory, thinking, and behavior abilities. ADNI is a longitudinal study initialed in 2004 (Mueller, et al., 2005). One of the goals of this study is to find out a biomarker for the diagnosis of the AD. The data set we used includes 661 subjects enrolled between 2005 and 2007 with patients at different stages including AD, late mild cognitive impairment (LMCI), as well as CN elderly controls. The number of subjects in AD, LMCI, and CN are, respectively, 204, 318, and 139. More details about the ADNI study can be found at `http://adni.loni.usc.edu`.

The FDG-PET scanning provides brain imaging data that measure patients' CMRglc. It has been a widely used technique for early AD diagnosis for more than two decades (Mosconi, et al., 2010; Johnson, et al., 2012). Our interest is to test the significance of the monotonic changes among the means of the FDG values for three groups. Specifically, let y_{ij} be the FDG value for the j-th patient in the i-th group where $i = 1$ for AD, $i = 2$ for LMCI, and $i = 3$ for CN. Assume that $E(y_{ij}) = \theta_i$. To confirm if CMRglc is an appropriate biomarker, as discussed in the introduction, we are interested in testing the following hypothesis

$$H_0 : \theta_1 = \theta_2 = \theta_3 \quad \text{vs.} \quad H_1 : \theta_1 \leq \theta_2 \leq \theta_3$$
$$\text{with at least one strict inequality holds.} \qquad (8)$$

The missing percentages for the FDG-PET scanning data at the three stages (AD, LMCI, and CN) are around 79.4, 53.0, and 63.3%, respectively. The missingness is due to various reasons. For example, some participants do not agree on the consent of the PET scan, or have some specific exclusion (e.g., a history of radiation therapy) to the PET scan. To impute the missing values, we consider using the magnetic resonance imaging (MRI) data and age information as covariates. MRI is a principle component of the ADNI study that is important for tracking the progression of AD, and MRI data were observed for almost every patient.

A total of three volumes of brain regions measured by MRI was utilized, which include the volumes of fusiform, hippocampus, and the whole brain. These three brain region volumes and age information were the candidate covariates. In order to find the covariates associated with the missingness, we performed a logistic regression using the above six covariates and the binary nonmissing indicators δ_{ij}'s as responses. A stepwise forward logistic regression was used to select the best model. The final selected model with the smallest AIC (AIC = 867.34) contained two covariates: the volumes of hippocampus and fusiform regions. Therefore, we used the volumes of hippocampus and fusiform areas as the covariates for the nonparametric kernel imputation. More details about the stepwise forward logistic regression are included in web-based Appendix E. The covariate Age was not selected into the imputation model. This might be partially due to the fact that the volumes of the hippocampus and fusiform areas have significant linear relationship with Age. The imputation model was chosen based on the data set and information avail-

Table 4
Empirical sizes and powers of JEL ratio test for order restricted hypothesis with unbalanced samples and different missing probabilities among groups

	Higher missing prob.				Lower missing prob.			
(n_1, n_2, n_3)	Size	A	B	C	Size	A	B	C
(100, 100, 200)	0.057	0.178	0.285	0.613	0.053	0.196	0.326	0.654
(200, 100, 100)	0.051	0.201	0.312	0.669	0.045	0.187	0.359	0.746
(100, 100, 400)	0.044	0.207	0.319	0.648	0.054	0.208	0.347	0.711
(400, 100, 100)	0.070	0.214	0.389	0.822	0.043	0.224	0.374	0.865
(100, 200, 300)	0.054	0.222	0.374	0.663	0.052	0.192	0.355	0.741
(300, 200, 100)	0.048	0.207	0.366	0.774	0.045	0.199	0.369	0.796

The data are chi-square distributed in this table.

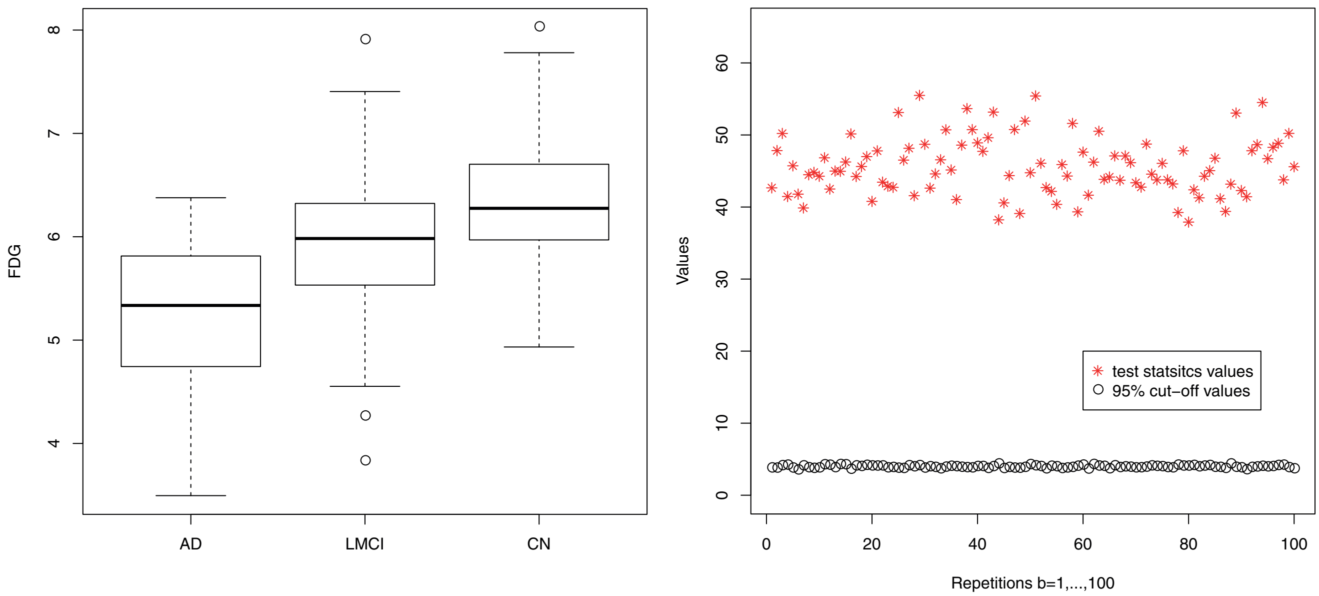

Figure 2. Left panel: Side-by-side boxplots of the FDG values for patients that at different stages of the AD: AD, late mild cognitive impairment (LMCI) and CN. Right panel: Test statistics values and the cutoff values among 100 repetitions.

able to us. If more information related to the missingness (e.g., participant's consent for the PET scan) is available, it could be incorporated to improve the imputation model.

To obtain some insight about the means θ_1, θ_2, and θ_3, we provide the box plots for values of y_{ij} for patients at three different stages in left panel of Figure 2. It can be clearly observed from the box plots that the means among three groups change monotonically.

To formally test the hypothesis in equation (8), we applied the proposed JEL ratio test. We randomly sampled around 86% of the individuals ($n_1 = 120, n_2 = 275$, and $n_3 = 176$), respectively, from the AD, LMCI, and CN groups and performed the proposed JEL ratio test on each sampled data set. We repeated the above procedure for 100 times, and obtained the corresponding test statistic values $\Lambda_n^{(b)}$ for $b = 1, \ldots, 100$. The test statistics values $\Lambda_n^{(b)}$ ranged between 38.22 and 55.49. The average and standard deviation of the 100 test statistics $\Lambda_n^{(b)}$ were, respectively, 45.61 and 3.96. To check the significance, the chi-bar-square distributions under the null were generated using simulation. Specifically, for each sampled data set, the pseudo-values were calculated, and σ_i^2's were estimated based on the variances of the pseudo-values v_{ij}'s. Then the chi-bar-square weights w_j's can be obtained using numerical integration. The upper 5% quantile of the chi-bar-square distributions ranged between 3.29 and 4.18. We illustrate the test statistics and cutoff values in the right panel of Figure 2. From the plot, it is easy to see that all the test statistic values are above the cutoff points. Therefore, JEL ratio test rejects the null hypothesis for all the 100 random samples. Based on the simulation results in Tables 1–3, the proposed method maintained the type I error under various missing probabilities but less powerful as the missing probability increased. In this real data application, the null hypotheses were rejected for all random samples even when the AD group had missing probability as high as 79.4%. This might indicate that the monotone relationship among θ_i's was strong, and CMRglc was a good biomarker for distinguishing different stages of AD.

7. Concluding Remarks

In this article, we studied the problem of testing order restricted means for data with nonparametric imputation. We employed the jackknife EL ratio statistics to construct the test statistics. The asymptotic distribution of the test statistic was derived under the null hypothesis. It turns out the asymptotic null distribution follows a chi-bar-square distribution, which is very easy for implementation. The proposed method is robust to the underlying distribution of the data. No normality assumption is needed. The major contribution of this article is on proposing a formal procedure for testing order restricted means when part of the data are missing and imputed using a nonparametric kernel regression. To the best of our knowledge, no formal procedure in the literature considers the problem of missing values for testing order restricted means. Our proposed procedure bridges this important gap. Our simulation studies have demonstrated that the proposed procedure is valid for various normally and nonnormally distributed data, and is able to accommodate data with nonparametric imputation.

We considered a nonparametric imputation method in this article. As an alternative, one could consider a parametric imputation method. The asymptotic distribution of the JEL test statistic with parametric imputation could be derived similarly. The JEL method with parametric imputation works well if the parametric imputation model (i.e., the conditional expectation of y given $\boldsymbol{x}$) is correctly specified. However, the derived asymptotic null distribution may not be valid for JEL method with an incorrect imputation model. This is because the asymptotic behaviors of jackknife pseudo values based on an incorrect imputation model are different from that based on the correct imputation model. Some simulation results for illustrating the JEL method with parametric imputation are given in the web-based Appendix D.

Inverse weighting method (e.g., Kim and Shao, 2013) is another popular method for handling missing values. If the propensity score function is known or estimated, the EL ratio test statistic with calibration based on the inverse weighted estimating equations may be applied to the problems considered in this article. However, due to the focus of this article is on data with imputation, we will investigate the inverse weighting method for the order restricted inference in a future project.

8. Supplementary Material

Web Appendices referenced in Sections 1–7 and the R code for implementing the proposed method are available with this article at the *Biometrics* website on Wiley Online Library.

Acknowledgments

We are very grateful to the editor, an associate editor, and two referees for their insightful and constructive comments on this manuscript, which has helped improve the quality of the article.

References

Bartholomew, D. J. (1959a). A test of homogeneity for ordered alternatives. *Biometrika* **46**, 36–48.

Bartholomew, D. J. (1959b). A test of homogeneity for ordered alternatives. II. *Biometrika* **46**, 328–335.

Bartholomew, D. J. (1961a). A test of homogeneity of means under restricted alternatives. *Journal of Royal Statistical Society Series B* **23**, 239–281.

Bartholomew, D. J. (1961b). Ordered tests in the analysis of variance. *Biometrika* **48**, 325–332.

Cheng, P. E. (1994). Nonparametric estimation of mean functionals with data missing at random. *Journal of American Statistical Association* **89**, 81–87.

Davidov, O., Fokianos, K., and Iliopoulos, G. (2010), Order-restricted semiparametric inference for the power bias model. *Biometrics* **66**, 549–557.

Davidov, O., Fokianos, K., and George Iliopoulos, G. (2014), Semiparametric inference for the two-way layout under order restrictions. *Scandinavian Journal of Statistics* **41**, 622–638.

El Barmi, H. (1996). Empirical likelihood ratio test for or against a set of inequality constraints. *Journal of Statistical Planning and Inference* **55**, 191–204.

Estevez-Gonzalez, A., Kulisevsky, J., Boltes, A., Otermin, P., and Garcia-Sanchez, C. (2003). Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging. *International Journal of Geriatric Psychiatry* **18**, 1021–1028.

Fuller, W. A. and Kim, J. K. (2005). Hot deck imputation for the response model. *Survey Methodology* **31**, 139–149.

Jing, B. Y., Yuan, J., and Zhou, W. (2009). Jackknife empirical likelihood. *Journal of American Statistical Association* **104**, 1224–1232.

Johnson, K. A., Fox, N. C., Sperling, R. A., and Klunk, W. E. (2012). Brain imaging in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine* **2**, a006213.

Kim, J. K. and Shao, J. (2013). *Statistical Methods for Handling Incomplete Data*. Boca Raton, FL: Chapman and Hall/CRC.

Mosconi, L., Berti, V., Glodzik, L., Pupi, A., De Santi, S., and de Leon, M. J. (2010). Pre-clinical detection of Alzheimer's disease using FDG-PET, with or without amyloid imaging. *Journal of Alzheimer's Disease* **20**, 843–854.

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W. et al. (2005). Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer's and Dementia* **1**, 55–66.

Nettleton, D. (1999). Order-restricted hypothesis testing in a variation of the normal mixture model. *Canadian Journal of Statistics* **27**, 383–394.

Nettleton, D. (2009). Testing for the supremacy of a multinomial cell probability. *Journal of American Statistical Association* **104**, 1052–1059.

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

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

Qin, J. and Lawless, J. F. (1995). Estimating equations, empirical likelihood and constraints on parameters. *Canadian Journal of Statistics* **23**, 145–159.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons, Inc.

Silvapulle, M. and Sen, P. K. (2004). *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*. Hoboken, NJ: Wiley.

Wang, D. and Chen, S. X. (2009). Empirical likelihood for estimating equations with missing values. *The Annals of Statistics* **37**, 490–517.

Wang, Q. and Rao, J. N. K. (2002). Empirical likelihood-based inference under imputation for missing response data. *The Annals of Statistics* **30**, 896–924.

Wang, L. and Wang, Q. (2006). Empirical likelihood for parametric model under imputation for missing data. *Journal of Statistics and Management Systems* **9**, 1–13.

Wollan, P. C. and Dykstra, R. L. (1986). Conditional tests with an order restriction as a null hypothesis. *Advances in Order Restricted Statistical Inference* **37**, 279–295.

Zhong, P. S. and Chen, S. (2014). Jackknife empirical likelihood method for inference with regression imputation. *Journal of Multivariate Analysis* **129**, 193–205.

Received June 2015. *Revised December* 2016.
Accepted December 2016.