Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

A discussion of ‘statistical inference for nonignorable missing data problems: a selective review’ by Niansheng Tang and Yuanyuan Ju

Kosuke Morikawa & Jae Kwang Kim

To cite this article: Kosuke Morikawa & Jae Kwang Kim (2018) A discussion of ‘statistical inference for nonignorable missing data problems: a selective review’ by Niansheng Tang and Yuanyuan Ju, Statistical Theory and Related Fields, 2:2, 140-140, DOI: 10.1080/24754269.2018.1522995

To link to this article: https://doi.org/10.1080/24754269.2018.1522995

Published online: 26 Sep 2018.

STATISTICAL THEORY AND RELATED FIELDS
2018, VOL. 2, NO. 2, 140
https://doi.org/10.1080/24754269.2018.1522995

Taylor & Francis
Taylor & Francis Group

SHORT COMMUNICATION

Check for updates

A discussion of 'statistical inference for nonignorable missing data problems: a selective review' by Niansheng Tang and Yuanyuan Ju

Kosuke Morikawa[a] and Jae Kwang Kim[b]

[a]Graduate School of Engineering Science, Osaka University, Osaka, Japan; [b]Department of Statistics, Iowa State University, Ames, IA, USA

ARTICLE HISTORY Received 7 September 2018; Accepted 10 September 2018

We congratulate Tang and Ju for their refined review paper on the recent statistical inference of nonignorable missing data. This paper covers various topics in missing data analysis including semiparametric estimation, Bayesian inference, structural equation modelling, sensitivity analysis and model selection, so that it fills the gap between complete and missing data analysis.

We have a comment on the semiparametric estimation of mean functionals. With known γ, they used '$n^{-1}\sum_i \xi_i^a(\gamma)=\theta$' as a calibration condition in Section 3.4, where $\xi_i^a(\gamma)=\delta_i y_i/\pi_i(\gamma)+\{1-\delta_i/\pi_i(\gamma)\}m_0(x_i)$. In a similar way, '$n^{-1}\sum_i \hat{\xi}_i(\gamma)=\theta$' also works as a calibration condition from $E\{\delta_i y_i+(1-\delta_i)m_0(x_i;\gamma)\}=\theta$, where $\hat{\xi}_i(\gamma)=\delta_i y_i+(1-\delta_i)\hat{m}_n^0(x_i;\gamma)$. However, as is pointed out in the paper, γ is generally unknown, and to estimate it, some additional conditions are needed. Therefore, one can apply the idea of Shao and Wang (2016) introduced in the beginning of Section 3 as the calibration conditions, i.e.

$$\sum_{i=1}^n p_i\hat{\xi}_i(\gamma)=\theta,$$

$$\sum_{i=1}^n p_i\mathbb{M}_l(y_i,u_i,\delta_i,\hat{g}_\gamma,\gamma)=0,\quad l=1,\ldots,L,$$

where $\hat{g}_\gamma$ is the nonparametric estimator of the g-function. Let $\hat{\ell}(\theta,\gamma)$ be the empirical likelihood with the calibration conditions. Then, an estimator can be obtained as the maximiser of the empirical likelihood. In addition, according to Qin and Lawless (1994), if θ_0 is the true value, the following profile empirical log-likelihood ratio function for θ would have the asymptotic chi-squared distribution under some regularity conditions:

$$\mathcal{R}=2\{\max_\gamma \ell(\theta_0,\gamma)-\max_{\theta,\gamma}\ell(\theta,\gamma)\}.$$

This estimator has two advantages compared to Zhao, Zhao, and Tang (2013): it does not require (i) the previous knowledge of γ and (ii) correct specification of $g(x)$. Because this estimator does not require estimating the asymptotic variance, it is superior to Shao and Wang (2016)'s method with respect to estimation of the asymptotic variance or the statistical testing.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Kosuke Morikawa, an assistant professor of Engineering Science at Osaka University and a visiting research fellow at Earthquake Research Institute of the University of Tokyo, writes on missing data analysis. He specializes in missing data analysis and spatial data analysis.

Jae-kwang Kim, a professor of Statistics at Iowa State University, is a fellow of American Statistical Association and the recipient of 2015 Gertrude M. Cox award. He majors in survey sampling and missing data analysis.

References

Qin, J., & Lawless, J. (1994). Empirical likelihood and general estimating equations. *Journal of the American Statistical Association*, *22*, 300–325.

Shao, J., & Wang, L. (2016). Semiparametric inverse propensity weighting for nonignorable missing data. *Biometrika*, *103*, 175–187.

Zhao, H., Zhao, P. Y., & Tang, N. S. (2013). Empirical likelihood inference for mean functionals with nonignorably missing response data. *Computational Statistics & Data Analysis*, *66*, 101–116.

CONTACT Jae Kwang Kim jkim@iastate.edu Iowa State University, Snedecor Hall, Ames, IA 50011, USA