Journal of Classification 35:1-4 (2018)
DOI: 10.1007/s00357-018-9254-1

Editorial

The first paper in this issue of the *Journal of Classification* by Fionn Murtagh, Michael Orlov, and Boris Mirkin addresses the ever tricky issue of evaluating the quality of research produced by individuals. The approach by Murtagh and coauthors avoids many of the issues that are leveled against common metrics of productivity (like impact factor for instance, see Brumback, 2009; Brischoux and Cook, 2009; Rossner, Epps, and Hill, 2007) by taking a more holistic approach to assessing a researcher's contributions. Personally, I believe that this is one of first efforts of what promises to be many forays into taking a broad approach at tackling the issue of the contribution(s) of an individual researcher.

In the next paper, Richard Payne and Bani Mallick provide a two-stage Metroplis-Hastings algorithm for Bayesian classification. As mentioned in the last issue, Bayesian analysis has been somewhat of a rarity in the *Journal of Classification*; however, along with Ligtvoet (2017), this article is the second paper in two issues that addresses Bayesian techniques, and it is a welcome addition to the panoply of methods normally covered. The goal of Payne and Mallick's procedure is to reduce the computational burden when the number of observations is much greater than the number of variables (e.g., so called "tall" data). Their solution consists of the best kind of solution: one that is obvious upon seeing it but heretofore not proposed by anyone else (e.g., not as obvious as we might think). At its simplest, they provide approximations of the likelihood and only provide the full estimate if the proposed solution is likely to be accepted. While Bayesian statistics is not my primary area of research, this reminds me of several strategies found in the computational optimization literature where various solutions are pruned during the search process as to ease the overall computational burden (see Brusco and Stahl, 2005, for example).

Wenxin Zhu, Yunyan Song, and Yingyuan Xiao offer up a modification to support vector machine + (SVM+) in the third paper. Much like Bayesian statistics, there has been a dearth of papers on SVM in the *Journal of Classification*; however, this is the second paper in as many issues (with Nunez, Gonzalez-Abril, and Angulo, 2017, being the first). Specifically, they add a pinball loss function, which sounds like something that would be more appropriate in a Roger Daltrey lyric, but merely gets its name from an appearance that resembles the trajectory of a pinball (or a billiard ball, for that matter). Zhu et al. show that incorporating the pinball loss function into SVM+ acts as a desensitization method for mitigating the effect of noise on the outcome and is also more stable than the traditional SVM+.

The fourth paper, by Carolina Euan, Hernando Ombao, and Joaquin Ortega, addresses the growing need for new methods for clustering time series (see Michael and Melnykov, 2016; McNeish and Harring, 2017, for recent examples). Many of the recent approaches have used finite mixture modeling as a basis for clustering time series. Here, Euan et al. take a different tack and approach the problem from a more traditional hierarchical clustering approach[1]. In the process, they use the total variation distance between two distributions to determine the similarities between the spectral densities. After the distances are computed, the spectral densities with the smallest total variation distance are merged and then the spectral density of the merged clusters is updated—this operates similar to a traditional hierarchical clustering algorism, such as average linkage. I think this method is promising and has many potential avenues for expanding beyond what is presented in the current manuscript.

In the next paper, Vincenzo Spinelli provides a supervised clustering method based on convex sets (e.g., box hulls) that are used to develop a set of hypergraphs that drive the final classification result. One interesting feature of this approach is that the proposed clusters are formed based on homogeneity conditions and not a similarity measure. For me, the most promising aspect of this work is potential uses alluded to by Spinelli in the conclusion. Namely, the ability to use the aspects of the points surrounding a point of interest to lend themselves to creating a measure of reliability—that would be a contribution that would help push the entire field of clustering and classification in an interesting direction.

As the analysis of binary data seems to be becoming more pervasive in recent years (see Gudicha, Tekle, and Vermunt, 2016; Steinley, Brusco, Hoffman, and Sher, 2017; Sulis and Porcu, 2017; van der Palm, van der Ark, and Vermunt, 2016 for a few examples), Hailemichael Worku and Mark de Rooij provide an extremely nice model for analyzing multiple binary response variables. Personally, I have found the literature for modeling multivariate binary response data sparse and unsatisfying, and Worku and de Rooij have provided a succinct model that is flexible and able to take advantage of shared information among the binary responses in multivariate space. The incorporation of biplots provides an indispensable visual interpretation system and the potential for dimension reduction really makes this model appealing, and its potential for application to a wide range of applications is great. For instance, I think that this modeling can aid in some of the questions that some (Boorsboom and Cramer, 2013) are attempting to answer with network analysis of psychopathology.

[1] Even though hierarchical clustering is one of the first types of cluster analysis, it still sees many advances, year-in year-out (see Martinez-Perez, 2016, for a recent example).

In the penultimate paper of the issue, Ehsan Bokhari and Lawrence Hubert provide a reanalysis of the MacArthur Violence Risk Assessment, showing the critical need for cross-validation when fitting models. In this particular case, Bokhari and Hubert show the extreme difficulty in predicting violence and the likelihood that the original work by Monahan on assessing violence is overestimating its predictive capability. The example is illustrative, but it speaks more to the ongoing crisis of replicability that has been plaguing psychology, medicine, and other fields (see Pashler and Wagenmakers, 2012, for a review). Steinley and Brusco (2011) provide a similar argument for the dangers of fitting too complex of a model in the context of identifying unknown groups—namely, the more complex the model, the more difficult it will be to replicate. As Bokhari and Hubert indicate, there is much to be gained from conducting cross-validation on flexible models, such as the decision tree analysis they conducted. Such an analysis would have benefited the work of Steinley and Brusco, rather than their default to the simpler model. In fact, it is likely that the field would be better off if we began conducting cross-validation on all models that we fit to data.

The final paper of the issue, by Francesco Bartolucci, Giorgio Montanari, and Silvia Pandolfi also uses an application to motivate their primary development: item selection in latent class analysis (LCA). Item selection has long been an issue when looking for unobserved (e.g., latent groups), both in hierarchical clustering (De Soete, 1986; De Soete, DeSarbo, and Carroll, 1985), nonhierarchical clustering (Steinley and Brusco, 2008; Steinley, Brusco, and Henson, 2012), mixture modeling (Raftery and Dean, 2006), and latent class analysis (Dean and Raftery, 2010; Fop, Smart, and Murphy, 2017). In their treatment, Bartolucci et al. extend the standard LCA model to be able to handle missing data as well as integrate an item selection model to determine which items do not contribute to the overall class structure. The decision rule for discarding items is to treat the cluster solution as "true" and then discard items if their absence does not change that solution. The benefit of this procedure is that it is computationally efficient and easy to implement. Further, I believe that it can be extended in interesting directions when one considers the possibility that some items in the original set could be providing an initial "true" solution that is contaminated by the influence of poor items.

Douglas Steinley

References

BOORSBOOM, D., and CRAMER, A.O.J. (2013). "Network Analysis: An Integrative Approach to the Structure of Psychopathology", *Annual Review of Clinical Psychology, 9,* 91–121.

BRISCHOUX, F. (2009). “Juniors Seek an End to the Impact Factor Race”, *Bioscience, 59,* 238–239.

BRUMBACK, R.A. (2009). “Impact Factor Wars: Episode V – The Empire Strikes Back”, *Journal of Child Neurology, 24,* 260–262.

BRUSCO, M.J., and STAHL, S. (2005). “Brand-and-Bound Applications in Combinatorial Data Analysis”, New York: Springer.

GUDICHA, D.W., TEKLE, F.B., and VERMUNT, J.K. (2016). “Power and Sample Size Computation for Wald Tests in Latent Class Models”, *Journal of Classification, 33,* 30–51.

LIGTVOET, R. (2017). “Exact One-Sided Bayes Factors for 2 by 2 Contingency Tables”, *Journal of Classification, 34,* 465–472.

MARTINEZ-PEREZ, A. (2016). “On the Properties of α-Unchaining Single Linkage Hierarchical Clustering”, *Journal of Classification, 33,* 118–140.

MCNEISH, D., and HARRING, J.R. (2017). “The Effect of Model Misspecification on Growth Mixture Model Class Enumeration”, *Journal of Classification, 34,* 223–248.

MICHAEL, S., and MELNYKOV, V. (2016). “Finite Mixture Modeling of Gaussian Regression Time Series with Applications to Dendrochronology”, *Journal of Classification, 33,* 412–441.

NUNEZ, H., GONZALEZ-ABRIL, L., and ANGULO, C. (2017). “Improving SVM Classification on Imbalanced Datasets by Introducing New Bias”, *Journal of Classification*, 34, 427–443.

PASHLER, H., and WAGENMAKERS, E.J. (2012). “Editors’ Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence?”, *Perspectives on Psychological Science, 7,* 528–530.

ROSSNER, M., EPPS, H.V., and HILL, E. (2007). “Show Me the Data”, *The Journal of Cell Biology, 179,* 1091–1092.

STEINLEY, D., and BRUSCO, M.J. (2011b). “K-means Clustering and Model-Based Clustering: Reply to McLachlan and Vermunt”, *Psychological Methods, 16,* 89–92.

STEINLEY, D., HOFFMAN, M., BRUSCO, M.J., and SHER, K.J. (2017). “A Method for Making Inferences in Network Analysis: Comment on Forbes, Wright, Markon, and Krueger (2017)”, *Journal of Abnormal Psychology, 126*, 1000–1010.

SULIS, I., and PORCU, M. (2017). “Handling Missing Data in Item Response Theory. Assessing the Accuracy of Multiple Imputation Procedure Based on Latent Class Analysis”, *Journal of Classification, 34,* 327–359.

VAN DER PALM, D.W., VAN DER ARK, L.A., and VERMUNT, J.K. (2016). “Divisive Latent Class Modeling as a Density Estimation Method for Categorical Data”, *Journal of Classification, 33,* 52–72.