Tamás Rudas, Renáta Németh

Rejoinder: On the Application of Discrete Marginal Graphical Models

THE FOLLOWING TEXT IS THAT OF THE ORGINIAL SUBMISSION.
THE FINAL PUBLISHED VERSION WILL BE SLIGHTLY DIFFERENT.

We are grateful to all discussants for their insightful contributions. There is very little in their comments with which we disagree, so can keep our reply brief.

We thank Antonio Forcina for his comments and we certainly agree with him that cutting edge research, including his current paper (Forcina, 2012) has more to say on the topic than presented here. Our aim was to show a relatively simple part of the results in this field, that we hoped could be used directly and is intuitive enough to have the potential of becoming part of the standard tools of social scientists.

We also agree about his cautionary comments of careless causal interpretation of models with arrows. We see a growing number of papers calling these models causal models. This practice has the danger of inducing applications to establish causality, as opposed to the undisputed use: to estimate certain parameters related to causality, if one knows that the causal relationship exists. We only wanted to add one, we were hoping intuitive enough, warning: a 'causal' model may fit, but also its 'opposite' may fit. On the other hand, we hope that DAGs, chain graph models and their marginal log-linear parameterizations will continue to be helpful in exploring the causal structure – if that exists.

The comments given by Robin Evans mention some of the ongoing research activities in the area of graphical modeling. We do agree with him that ADMGs represent an interesting and potentially useful option to associate a hypothetical generating mechanism with data that exhibit conditional independencies. Whether practitioners choose one or another interpretation, will depend, ultimately, on how useful they find the possible interpretations in their substantive research. We believe, this applies equally to the competing interpretations of chain graph models, for those, to whom the graph is the starting point of the modeling process (see, e.g., Drton, 2009 or Rudas, Bergsma, Németh, 2010), and to the choice from among several Markov-equivalent models, for those, to whom the conditional independencies are the starting point the modeling process (see, e.g., Lauritzen, 1996).

We were delighted to see that Nanny Wermuth considers marginal modeling a powerful tool. Whether marginal models are too powerful to discuss properties of DAGs, depends on if one prefers using *several* simple and specific descriptions or *one,* of which many models, usually discussed individually, are special cases.

Professor Wermuth uses the word regression to describe situations, when the distribution of a variable conditioned on some other variables depends only on a subset of the latter. We agree, that, in a general sense, these are regression-type situations, in particular, if one only considers the research question. If, however, also the levels of measurement and distributional properties of the variables involved are taken into account, this general regression-type problem has to be clearly distinguished from linear regression analysis. So we agree, that if some of the variables are responses to others, one is looking at

a regression-type problem, but when the variables are categorical, we propose to fit marginal log-linear models and to estimate marginal log-linear parameters. One advantage of this approach is that it provides a way to handle the non-standard cases of vanishing lower-order and existing higher-order interactions. This is likely to become important, among other settings, when faithfulness (see, e.g., Meek, 1995) is studied in the discrete case.

Whether or not the models we name path models deserve that name, depends on which of the properties of the original path modeling approach one wishes to keep in different settings. One thing seems sure: if one wants to define path models for categorical data, the idea of *linear* dependence cannot be sustained. We just cannot associate any meaning with saying that someone's occupational status (as a categorical variable) is a linear function of his or her gender. Therefore, the semi-parametric approach of marginal log-linear parameters is suggested.

Claudia Tarantola, Ioannis Ntzoufras and Monia Lupparelli present a very welcome addition to our analysis from a Bayesian perspective. From the point of view of a scientist, who is primarily interested in analyzing a data set, it is always reassuring to see, when two analyses, based on such different principles, as frequentist and Bayesian analyses are, result in comparable, perhaps even similar results. In practice, the choice of the analysis by a scientist will be the result of a comparison of what she thinks to be the 'right' analysis in similar situations and of what she thinks the right results 'should be' for the actual data set to be analyzed. In other words, it will be a mixture of knowledge and expectations.

It is, obviously, the methodologist's responsibility is to help the researcher to acquire correct knowledge with respect to the assumptions behind and properties of the models used. The field of graphical models is rapidly growing, and we all take inspiration from the work done by others. In this context we are researchers, but when we talk about our models to substantive scientists, social scientists in this case, we become the methodologists. In this role, we want to emphasize, again, that models for Gaussian and categorical data are fundamentally different, because of the fundamentally different association structure allowed in these two cases. Any modeling of covariances, including covariance graph models, is only appropriate for Gaussian data (see, e.g., Dempster, 1972). On the other hand, the MLL parameterization was introduced by Bergsma and Rudas (2002) with the explicit goal of also modeling the higher order interactions which may only be present in categorical data.

Bergsma, Wicher P., and Tamás Rudas. 2002. "Marginal Models for Categorical Data." *Annals of Statistics* 30(1):140–59.

Dempster, Arthur P. 1972. "Covariance selection." *Biometrics* 3(1):157-175.

Drton, Mathias. 2009. "Discrete Chain Graph Models." *Bernoulli* 15(3):736–53.

Forcina, Antonio. 2012. "Smoothness of Conditional Independence Models for Discrete Data." *Journal of Multivariate Analysis* 106(1):49–56.

Lauritzen, Steffen L. 1996. *Graphical Models*. Oxford: Clarendon Press.

Meek, Cristopher. 1995. "Strong Completeness and Faithfulness in Belief Networks". Pp. 411–418 in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, edited by S. Hanks and P. Besnard, Montreal, Canada: Morgan Kaufmann.

Rudas, Tamás, Wicher P. Bergsma, and Renáta Németh. 2010. "Marginal Log-Linear Parameterization of Conditional Independence Models." *Biometrika* 97(4):1006–12.