Bocconi

Università Commerciale Luigi Bocconi

Department of Decision Sciences

Statistics Seminar

Dichotomization invariant log-mean linear parameterization for discrete graphical models of marginal independence

Alberto Roverato

Università di Bologna

Thursday, 18th October 2012

12:30pm Room 3-E4-SR03 Via Rontgen 1 Milano

Abstract

Graphical models of marginal independence use a graph where every vertex is associated with a variable and missing edges encode marginal independence relationships according to a given Markov property. The probability distribution of a set of discrete variables is characterized by the associated probability table, but defining a suitable parameterization for these models is not straightforward. A basic requirement for the flexible implementation of marginal constraints is that interaction terms involving a subset of variables satisfy upward compatibility, that is they should reflect a property of the corresponding marginal distribution. Upward compatibility means invariance with respect to marginalization but, for discrete variables with arbitrary number of levels, a stronger invariance property may be required. Collapsing two or more levels of a discrete variables into a single level can be regarded as as a special kind of marginalization and invariance with respect to this operation is an useful feature for a parameterization.

We extend the Log-Mean Linear (LML) parameterization introduced by Roverato, Lupparelli and La Rocca (2011, arXiv:1109.6239) for binary data to discrete variables with arbitrary number of levels and show that also in this case it can be used to parameterize graphical models of marginal independence. Furthermore, we show that the LML parameterization satisfies a stronger version of upward compatibility that we call dichotomization invariance. As a consequence, the LML parameterization allows one to simultaneously represent marginal independencies among variables and marginal independencies that only appear when certain levels are collapsed into a single one. This feature is useful in several applied contexts, such as genetic association studies. Furthermore, it provides a natural way to reduce the parameter count by means of substantive constraints that give additional insight on the dependence structure of variables.

Department of Decision Sciences

Via Röntgen 1 - 20136Milano

Tel. 02 5836.5632
Fax 02 5836.5630