Rejoinder—Estimation Issues for Copulas Applied to Marketing Data

Peter J. Danaher, Michael S. Smith,

Rejoinder

Estimation Issues for Copulas Applied to Marketing Data

Peter J. Danaher, Michael S. Smith

Melbourne Business School, University of Melbourne, Carlton, Victoria 3053, Australia
{p.danaher@mbs.edu, mike.smith@mbs.edu}

Estimating copula models using Bayesian methods presents some subtle challenges, ranging from specification of the prior to computational tractability. There is also some debate about what is the most appropriate copula to employ from those available. We address these issues here and conclude by discussing further applications of copula models in marketing.

Key words: Bayesian estimation; discrete copula; Markov chain Monte Carlo; Gaussian copula; media modeling; probability models; website page views

Introduction

We thank George and Jensen (2011; G&J hereafter) for their insightful comments. These help further clarify the ideas presented in our paper (Danaher and Smith 2011) and provide us with an opportunity to elaborate on them in two ways. The first is to further discuss four technical, but important, aspects of Bayesian estimation applied to copulas, as highlighted by G&J. The second is to suggest alternative copulas that have potential for the modeling of marketing data.

Bayesian Estimation

The Prior

Bayesian estimation requires the specification of priors for the parameters in a model, which G&J note includes a matrix Γ for an elliptical copula. Because copulas are scale-free, Γ is restricted to be a correlation matrix via the parameterization in our Equation (14), and this identifies both the Gaussian and t-copulas. However, there is no current consensus in the statistical literature on how best to select a prior for a correlation matrix (Daniels and Pourahmadi 2009). For low-dimensional problems this is not such a problem because most reasonable priors are likely to be dominated by the likelihood. However, for higher-dimensional problems such as in our Example 4, where the dimension is 45, the prior selected for Γ may well meaningfully affect the posterior estimates. G&J correctly point out that by placing flat priors on the strictly upper triangular elements of R (where $\sum^{-1} = R'R$), we have also implied a prior

for Γ. This prior is similar to those for a covariance matrix discussed by Roverato (2000) and Smith and Kohn (2002). Nonetheless, there is no reason why our Algorithm 2 cannot be used with another prior, and it would be worthwhile investigating the impact of different priors for Γ in this way.

Random-Walk Proposal

We implement a random-walk Metropolis-Hastings step for generating the elements of R, r_{ij} for $j > i$, using a proposal distribution with a standard deviation of 0.01 (not variance, as might be inferred from our appendix). We also tried a proposal with a standard deviation of 0.02. Because Γ is a correlation matrix, we find that both these values work well for all our empirical examples, regardless of the scale of the data. However, we agree with G&J that automatic tuning of the random-walk proposal may well improve the mixing and efficiency of the sampling scheme further, and we recommend investigating this in any implementation.

Estimation of Marginal Parameters

Certainly, it is possible to undertake fully Bayesian estimation of a copula model as suggested by G&J. This is where the copula and any marginal parameters are estimated together from their joint posterior distribution. It can be achieved by inserting the following step immediately before Step 1(a) in our Algorithm 2:

New Step: Generate from $\theta_j \mid \{\Theta \backslash \theta_j\}, \{x_{ik}^* \mid k \neq j\}, \Gamma, x$.

This is a reduced conditional posterior distribution, with density as given by Pitt et al. (2006, p. 544) for

the Gaussian copula. The density can be derived similarly for the t-copula, as well as some other copulas. The generation is undertaken by a Metropolis-Hastings step with a proposal based on a normal approximation to the density, which is a common approach in Markov chain Monte Carlo (MCMC) estimation.

However, in our paper we estimate the copula parameters conditional on the marginal distributions, which can be advantageous for two reasons. The first is that many analysts advocate being nonparametric in the margins and parametric in the dependence structure (see, e.g., Oakes 1994, Genest et al. 1995). This corresponds to modeling the margins using their empirical distributions and then estimating the copula parameters conditional on these margins. The second is that even when parametric models for the margins are used (as in all four of our examples), joint estimation is widely observed to produce point estimates that do not change meaningfully. Yet the effort to undertake fully Bayesian estimation is large, both in terms of implementation time and computational demand. For example, because of the increased computational requirements, it would not be practical to estimate our Example 4, which has 45 margins, using current desktop computers, whereas it is practical when conditioning on the margins.

Bayesian Selection

G&J point out that Bayesian methods have the potential to allow for the selection of copulas using Bayes factors. We agree, but we would also like to highlight the importance of another aspect of Bayesian selection. The objective of the paper by Pitt et al. (2006) is to identify zeros in Γ^{-1}, producing a parsimonious Gaussian copula, which is important in high-dimensional problems. The same effect can be obtained using the representation in our Equation (14) by employing the Bayesian selection approach outlined by Smith and Kohn (2002) to identify zeros in the strict upper triangle of R. Estimates are then averages over different configurations of R, each weighted by their posterior probabilities. These types of estimates are called "model averages" in the Bayesian literature and are now widely used in applied analysis (see, e.g., Primiceri 2005, Jacobson and Karlsson 2004).

Which Is the Most Appropriate Copula?

The Gaussian Copula

The last issue raised by G&J concerns the possible appropriateness of the Gaussian copula in our Equation (9). In particular, they reiterate the widely reported failure of the Gaussian copula to capture dependence adequately for the extreme events in the financial markets during the global financial crisis of 2008–2009. This is because the Gaussian copula has quite rigid tail dependence properties. For example, the upper and lower asymptotic tail dependence is zero (McNeil et al. 2005, §5.3.1), which makes the Gaussian copula a poor candidate for modeling financial asset returns.

However, this is unlikely to be such an issue in the modeling of marketing data, which is often discrete and where tail dependence is less important. For example, the Gaussian copula is already widely (if often unknowingly) used in marketing in the form of the multivariate probit model (see, e.g., Seetharaman et al. 2005), which is a Gaussian copula model with univariate probit margins. To illustrate this in the bivariate case, consider binary-valued margins, where $\Pr(X_j = 1) = \Phi(-\mu_j)$ and $\Pr(X_j = 0) = \Phi(\mu_j)$, for margins $j = 1, 2$, with the corresponding distribution function

$$F_j(x_j) = \begin{cases} 0 & \text{if } x_j < 0, \\ \Phi(\mu_j) & \text{if } 0 \le x_j < 1, \\ 1 & \text{if } x_j \ge 1. \end{cases}$$

Then, employing the copula model, the joint probability mass function using our Equation (13) is

$$\begin{aligned} f(x) &= \Pr(X_1 = x_1, X_2 = x_2) \\ &= C(F_1(x_1), F_2(x_2)) - C(F_1(x_1 - 1), F_2(x_2)) \\ &\quad - C(F_1(x_1), F_2(x_2 - 1)) + C(F_1(x_1 - 1), F_2(x_2 - 1)). \end{aligned}$$

Now, the Gaussian copula in our Equation (9) has probability mass function

$$\begin{aligned} f(0, 0) &= \Phi_2(\mu_1, \mu_2 \mid \Gamma), \\ f(0, 1) &= \Phi(\mu_1) - \Phi_2(\mu_1, \mu_2 \mid \Gamma), \\ f(1, 0) &= \Phi(\mu_2) - \Phi_2(\mu_1, \mu_2 \mid \Gamma), \\ f(1, 1) &= 1 - \Phi(\mu_1) - \Phi(\mu_2) + \Phi_2(\mu_1, \mu_2 \mid \Gamma), \end{aligned}$$

which is that of a bivariate probit (Ashford and Sowden 1970), where the cutoff points μ_j are often modeled as linear in exogenous variables with coefficient vectors θ_j. Considering higher dimensions results in the multivariate probit model, whereas changing the marginals to ordered probit models results in the multivariate ordered probit. By considering different margins F_j, a much wider range of multivariate models can be constructed, and our Bayesian estimation methodology therefore extends the popular Bayesian MCMC estimation methods for probit models by Albert and Chib (1993) and Chib and Greenberg (1998). Given the estimation simplicity for Gaussian copulas, it seems that they should be the first choice for many copula modeling applications in marketing.

Other Copulas

Nevertheless, what copula choices are there other than the Gaussian? For bivariate problems the Archimedean class has long been widely used. There are many of these, with Nelsen (2006, Chapter 4) providing an introduction. These copulas can be estimated easily with discrete or continuous margins by using either maximum likelihood or Bayesian methods. For example, see Cameron et al. (2004) for an econometric application to bivariate count data. Particularly popular are mixtures of two or three such copulas, which can provide a more flexible dependence structure; see Rodriguez (2007) for a recent financial study that uses a mixture of three Archimedean copulas. However, in higher dimensions, there are currently only three practical alternatives to the Gaussian copula:

(1) *t*-Copula: The first alternative is to use the other popular elliptical copula, the *t*-copula, in our Equation (10). It has greater tail dependence than the Gaussian copula, and Algorithm 2 can be employed to compute Bayesian estimates with minor modification. The degrees of freedom ν can either be fixed, as in our Example 1, or estimated jointly with the correlation matrix Γ_T. We agree with G&J that the *t*-copula should be considered alongside the Gaussian.

(2) Vine copulas: The second alternative is to construct higher-dimensional copulas from nested sequences of bivariate copulas, such as the Archimedean copulas. These nested copulas are called "vines" and are the subject of much current research in statistics, with a recent discussion on estimation via maximum likelihood given by Aas et al. (2009) and via Bayesian methods by Smith et al. (2010b). However, current work only focuses on estimation when the data are continuous, which limits the usefulness of vine copulas for many marketing problems we examine.

(3) Skew *t*-copulas: These are a direct extension of the ideas presented in our study and follow the spirit of G&J's suggestions on how further extensions might proceed. The idea is that the latent variables x^* in our §2.5 are not elliptically distributed, but follow a slightly more flexible distribution called a "skew *t*." This is an extension of the *t* distribution, where there are some additional parameters to control the level of skew in each dimension. When used to construct a copula, this translates into a more flexible dependence structure, including differing levels of upper and lower tail dependence. Moreover, the copula nests both the Gaussian and *t*-copulas. Smith et al. (2010a) outline how to extend the ideas here to allow for Bayesian estimation of the skew *t*-copula and, by extension, the symmetric *t*-copula. They show using a 15-dimensional website exposure example that the skew *t*-copula can provide a more accurate fit than the *t*-copula, resulting in an improved measurement of total audience reach. However, a major downside is that estimation is substantially more complex and challenging to implement than for elliptical copulas.

Further Marketing Applications of Copulas

Since the time of our original study, we have embarked on further applications of copulas in marketing. One is an extension of the website sales and duration analysis presented in our Example 1 to include purchase incidence and page views. This results in a four-dimensional stochastic model where all four margins are different, an ideal situation to implement a copula model. Another application is a multivariate survival model for time to redemption for a set of three promotional offers sent to consumers via their cell phones.

We are also heartened to observe that several presenters at the recent Marketing Science conference in Cologne, Germany are using copula models. This underscores the versatility of this powerful tool.

References

Aas, K., C. Czado, A. Frigessi, H. Bakken. 2009. Pair-copula constructions of multiple dependence. *Insurance: Math. Econom.* **44**(2) 182–198.

Albert, J. H., S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88**(422) 669–679.

Ashford, J. R., R. R. Sowden. 1970. Multi-variate probit analysis. *Biometrics* **26**(3) 535–546.

Cameron, A. C., T. Li, P. K. Trivedi, D. M. Zimmer. 2004. Modelling the difference in counted outcomes using bivariate copula models with application to mismeasured counts. *Econometrics J.* **7**(2) 566–584.

Chib, S., E. Greenberg. 1998. Analysis of multivariate probit models. *Biometrika* **85**(2) 347–361.

Danaher, P. J., M. S. Smith. 2011. Modeling multivariate distributions using copulas: Applications in marketing. *Marketing Sci.* **30**(1) 4–21.

Daniels, M. J., M. Pourahmadi. 2009. Modeling covariance matrices via partial correlations. *J. Multivariate Anal.* **100**(10) 2352–2363.

Genest, C., K. Ghoudi, L.-P. Rivest. 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**(3) 543–552.

George, E. I., S. T. Jensen. 2011. A latent variable perspective of copula modeling. *Marketing Sci.* **30**(1) 22–24.

Jacobson, T., S. Karlsson. 2004. Finding good predictors for inflation: A Bayesian model averaging approach. *J. Forecasting* **23**(7) 479–496.

McNeil, A. J., R. Frey, P. Embrechts. 2005. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, Princeton, NJ.

Nelsen, R. B. 2006. *An Introduction to Copulas*, 2nd ed., Springer Series in Statistics. Springer, New York.

Oakes, D. 1994. Multivariate survival distributions. *J. Nonparametric Statist.* **3**(3 & 4) 343–354.

Pitt, M., D. Chan, R. Kohn. 2006. Efficient Bayesian inference for Gaussian copula regression models. *Biometrika* **93**(3) 537–554.

Primiceri, G. E. 2005. Time varying structural vector autoregressions and monetary policy. *Rev. Econom. Stud.* **72**(3) 821–852.

Rodriguez, J. C. 2007. Measuring financial contagion: A copula approach. *J. Empirical Finance* **14**(3) 401–423.

Roverato, A. 2000. Cholesky decomposition of a hyper inverse Wishart matrix. *Biometrika* **87**(1) 99–112.

Seetharaman, P. B., S. Chib, A. Ainslie, P. Boatwright, T. Chan, S. Gupta, N. Mehta, V. Rao, A. Strijnev. 2005. Models of multi-category choice behavior. *Marketing Lett.* **16**(3/4) 239–254.

Smith, M., R. Kohn. 2002. Parsimonious covariance matrix estimation for longitudinal data. *J. Amer. Statist. Assoc.* **97**(December) 1141–1153.

Smith, M., Q. Gan, R. Kohn. 2010a. Modeling dependence using skew *t* copulas: Bayesian inference and applications. *J. Appl. Econometrics*, ePub ahead of print October 12.

Smith, M., A. Min, C. Almeida, C. Czado. 2010b. Modeling longitudinal data using a pair-copula decomposition of serial dependence. *J. Amer. Statist. Assoc.*, ePub ahead of print September 8.