U.P.B. Sci. Bull., Series A, Vol. 71, Iss. 3, 2009 ISSN 1223-7027

PARAMETRIC ESTIMATION OF CONDITIONAL COPULAS

Mariana CRAIU[1]

Modelarea dependentei într-un numar mare de modele statistice s-a dezvoltat în urma definirii de noi copule. O copula stableste modele de dependenta bidimensionale intre variabile aleatoare ale caror repartitii sunt marginalele bidimensionalei.Posibilitatea introducerii de covariate conduce la variatii ale parametrului de dependenta cu valorile covariatelor.In articolul present studiem estimatii ale parametrilor acestor copule conditionate de covariate,folosind modele liniare.

Copulas have evolved into a popular tool for modeling dependence in a large number of statistical models.Generally, a copula establishes a flexible bridge between marginal distributions ,thus allowing various dependent models to be created.Recently ,an extension of the classical copula construction allows us to incorporate covariates in the model and allows the copula parameter to vary with some of these covariates. In this paper I investigate inference using linear models for conditional copula parameters.

Keywords: Copulas,Conditional Copulas,Clayton Copula,Frank Copula,Linear Models.

Mathematics Subject Classification 2000:62H12,62E20

1. Introduction

A theorem due to Sklar (1959) shows that any n-dimensional joint distribution function may be be decomposed into n marginal distributions ,and a copula ,which completely describes the dependence between the n variables.The copula is a more informative measure of dependence between two (or more) variables than linear correlation ,as when the joint distribution of the variables of interest in nonelliptical

The usual correlation coefficient is no longer sufficient to describe the dependence structure.The crux of the method is the ability to flexibly “couple”fixed marginal continuous distributions into multivariate distribution via a copula function.

[1] Prof., Mathematics Department, University POLITEHNICA of Bucharest, Romania

More precisely,if X,Y are continuous random variables with distribution functions (df) F_X and, respectively, F_Y we specify the joint df using the copula C:[0,1] $\times$ [0,1] $\longrightarrow$ [0,1] such that

$$F_{XY}(F_X^{-1}(u), F_Y^{-1}(v)) = P(X \le F_X^{-1}(u), Y \le F_Y^{-1}(v)) = C(u,v), \qquad (1)$$

Equation (1) illustrates the way in which the copula function “bridges” the marginal and the joint df’s.In general,the copula family depends on a parameter ,say θ . As θ varies,various degrees and types of dependence are represented by the copula C_θ .For instance consider the Clayton copula (Clayton 1978)

$$C_\theta(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\vartheta} \quad , \qquad (2)$$

and Frank’s copula (Frank 1979)

$$C_\theta(u,v) = -\frac{1}{\theta}\ln\left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right] \qquad (3)$$

Many other examples of copulas can be found in the comprehensive text of Nelsen(2006). Recently, Patton (2006) has introduced an extension of the classical copula construction in which one can incorporate the dependence between θ and a covariate Z using the concept of a **conditional copul**a.

Definition 1.1 The conditional copula of (X,Y)|Z=z, where X|Z=z ~$F_{X|Z}$ (.|z) and Y|Z=z ~ $F_{Y|Z}$ (.|z), is the conditional joint distribution function of U = $F_{X|Z}$ (X|z) and V= $F_{Y|Z}$ (Y|z) given Z=z.

The two variables U and V are known as the conditional probability integral transforms of X and Y given Z.It is simple to extend existing results to show that a conditional copula has the properties of an unconditional copula,for each Z=z as shown by Patton (2002).Patton (2006) proves the following general theorem.

Theorem 1.1 Let $F_{X|Z}$(. |z) be the conditional distribution of X|Z=z, $F_{Y|Z}$ (.|z) be the conditional distribution of Y|Z=z, $F_{XY|Z}$ (.|z) be the joint conditional distribution of (X,Y)|Z=z. Assume that $F_{X|Z}$ (x|z) and $F_{Y|Z}$ (y|z) are continuous in x and y for all z.

Then there exists a unique conditional copula C(.|z) such that

$F_{XY|Z}(x, y \mid z) = C(F_{X|Z}(x \mid z), F_{Y|Z}(y \mid z) \mid z)$ and each z.

In the next section we consider a possible specification of the conditional copula model. We also present parametric inference based on a linear model for the conditional copula parameter.In section 3 we study the performance of the proposed model using simulated data.

2. Specification and Parametric Inference for the Conditional Copula Model.

To illustrate the concepts presented here we start with a simple example.It is known that there is a dependence between blood pressure (BP) and body mass index (BMI),as discussed in Strohl et al. (1994).However ,it is possible that the dependence varies with various characteristics of the subjects under study,for instance their age.In other words one can expects that as age increases the dependence becomes stronger between the two continuous variables.Suppose we would like to use a copula model to characterize this dependence between BMI and BP.Evidently ,if we use a single copula parameter model then we would infer an average dependence structure.A more sensible model would allow the copula parameter itself to vary with the age of the subject.In this setting we essentially fit a separate copula for each different age group.Using the concept of a conditional copula one can do this automatically using a parametric linear regression model.

Define the two variables whose dependence we are interested in X,Y and the covariate Z.Suppose we decide that the copula best summarizes the dependence between X and Y is Clayton's copula (2) but in which we allow the linear dependence between Z and θ as $\theta = \beta Z$ .Besides the parameters of the marginal models for X and Y one is interested also in estimating the parameter β in order to better understand the dependence structure.

For the purpose of this paper ,we assume that the marginal distributions' parameters are known.Without loss of generality we will assume that the marginal distributions are Uniform (0,1).Therefore,the likelihood function of β can be written using the copula density corresponding to (2) which is

$$c_\theta(u,v) \approx (1+\theta)u_1^{-\theta-1}u_2^{-1-\theta}\left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{-\frac{1+2\theta}{\theta}} \tag{4}$$

If we assume that the observations are $(x_i, y_i, z_i)_{1 \le i \le n}$, then

$$L(\beta) = \prod\left[(1+\beta z_i)x_i^{-\beta z_i-1}y_i^{-\beta z\ -1}\left(x_i^{-\beta z_i-1}+y_i^{-\beta z_i-1}-1\right)^{-\frac{1+2\beta z}{\beta z_i}}\right] \tag{5}$$

Similarly ,the density of Frank's copula is

$$c_\theta(u,v) \approx \frac{\theta e^{-\theta(u+v)}\left(1-e^{-\theta}\right)}{\left(e^{-\theta}+e^{-\theta(u+v)}-e^{-\theta u}-e^{-\theta v}\right)^2} \tag{6}$$

Which in turn implies that the likelihood is

$$L(\beta)=\prod_1^n\left[\frac{\beta z_i e^{-\beta z_i (x_i+y_i)_i}\left(1-e^{-\beta z_i}\right)}{\left(e^{-\beta z_i}+e^{-\beta z_i(x_i+y_i)}-e^{-\beta z_i x_i}-e^{-\beta z y_{i_i}}\right)^2}\right] \tag{7}$$

The maximization of (5) and (7) is performed using standard optimization methods like Newton –Raphson.In the next section we perform a simulation study to asses the performance of the proposed approach.

3.Simulation Experiments

We performed a simulation study in which each marginal distribution is Uniform (0,1),the covariate Z~ Gamma (2,1)and the linear dependence between the copula
parameter and the covariate Z is $\theta = \beta$ Z. Evidently, any other linear model can be used. We investigate the performance of the MLE in this case using simulations with sample sizes n =20,50,100,200 and β = 0.5; 2 for models (5) and (7).
.In Table 1 we report the results obtained after 200 replications.The numbers in each cell represent the mean and, between brackets,the standard deviation of the MLE for
We report the same findings for the Franf's copula model in Table 2.

Table 1.

Simulation results in the case of Clayton's copula .The numbers in each cell represent the mean and,between brackets the standard deviation of the MLE for β .

β \ n	20	50	100	200
0,5	0,567 (0,265)	0.523(0.150)	0.503 (0.101)	0.508 (0.079)
2	2.132 (0.562)	1.995(0.328)	2.028(0.247)	2.015(0.166)

Table 2.

Simulation results in the case of Frank's copula.The numbers in each cell represent the mean and,between brackets the standard deviation of the MLE for β .

β \ n	20	50	100	200
0.5	0.599(0.473)	0.526 (0.344)	0.487 (0.232)	0.501 (0.202)
2	2.068 (0.797)	2.116 (0.517)	2.041 (0.353)	2.015 (0.241)

One can see that the Clayton dependence is stronger for the same value of β . The precision in estimating β depends on the copula family used. In our simulation the variance of the maximum likelihood estimator is larger under Frank's model.

4.Discussion

We illustrate one possible way to incorporate covariates with copula models using the concept of a conditional copula.The model we choose to illustrate the method is a linear one but other models are evidently possible.This raises an important question for further research: how can we select the appropriate link between the copula parameter and the covariates as well as how can we select from a set of covariates the ones that indeed influence the copula parameter.These issues will be addressed in a future communication.

R E F E R E N C E S

[1]. *D. Clayton*, A model for association in bivariate life tables and its application in epidemiological studies of familiar tendency in chronic disease incidence. Biometrika 65,141-151, 1978

[2]. *M.J. Frank*, On the simultaneous assciativity of f(x,y). Aequationes Math. 19,194-226, 1979.

[3]. *R.B. Nelsen*, An introduction to Copulas (2nd,ed.) Springer-Verlag, 2006.

[4]. *A.J. Patton*, Applications of copula theory in econometrics. PhDDissertation, 2002. University of California.

[5]. *A.J. Patton*, Modelling asymmetric exchange rate dependence.International Econom.Rev.47,527-556, 2006.

[6]. *K. Strohl, ,R. Novak,W.Singer,C. Cahan,s.a.*, Insulin levels,blood pressure and sleep apnea.Sleep 17(7),614-618, 1994.