Simulation of Univariate Time Series Using Copulas

Guilherme Pumi[1] **and Sílvia R. C. Lopes**

Federal University of Rio Grande do Sul - Institute of Mathematics.

April, 2010

Abstract

In this work we review and study some problems related to simulation of univariate time series based on some prescribed copula. We give some examples in the context of copula-based semiparametric models as in Chen and Fan (2006). We adapt some results to cover the case of the Brownian motion. Our approach allows one to simulate processes which behaves like a Brownian motion but have arbitrary marginals.

Keywords. Copulas, Monte Carlo Simulations, Markov Chains Simulation, Brownian Motion Simulation.

1 Introduction

The problem of generating stochastic processes with a given structure prescribed in terms of copulas is an important one but still no definitive answer is available. The invariance of copulas by almost sure increasing functions implies that a sample of the copula of a pair (X_{t-1}, X_t) from some stochastic process is theoretically the same as a sample of a pair $\big(f(X_{t-1}), g(X_t)\big)$ for any almost sure increasing functions f and g. Furthermore, to define a stochastic process taking values on $\mathbb{R}$ based on copulas just can't be done without some assumptions in the marginals, simply because a sample from a copula will take values in $[0, 1]$. In order to obtain the sample in the line, one can use the probability integral transform, for instance, and transform the sample from $[0, 1]$ to the whole line (or any other interval). At this point the flexibility of copulas play its role: once we have a sample of the desired copula, the marginal can be arbitrarily chosen in order to transform the sample from $[0, 1]$ to $\mathbb{R}$. As we shall see later, this allow one to separate the dependence structure from the marginals of the process so that one can simulate a sample of a stochastic process that behaves like a Brownian motion but have arbitrary marginals. This separation gives the advantage that one can choose the marginals in order to reflect any desired structure such as fat tails, asymmetries or even infinite variance.

These and other issues related to the problem of simulating univariate time series based on a given structure prescribed via copulas will be treated here. Our work is somewhat inspired by Chen and Fan (2006) in the sense that we shall work in the context of the semiparametric generalized linear model based in copulas presented there. We adapt those ideas in order to simulate univariate Brownian motions and, in principle, any other continuous time Markov process based on copulas.

The paper is organized as follows: in the next section, we briefly review some concepts and results we shall need in the work. Section 3 is devoted to review the copula-based semiparametric generalized linear model of Chen and Fan (2006) with some applications to univariate time series simulation based on copulas. In section 4, by means of an example (the Brownian motion) we show how to extend the ideas in section 3 to include temporal features in the copula used to generate the desired univariate process. Conclusions are reserved to section 5.

2 Some Background

In the following we shall review some facts and results about copulas needed later in the work. Copulas are distribution functions whose marginals are uniformly distributed on $I := [0, 1]$. In the next theorem, whose proof can be found, for instance, throughout Nelsen (2005), we summarize some of the properties of copulas. Except stated otherwise, the measure implicit to phrases like "almost sure", "almost everywhere" and so on will be the Lebesgue measure.

[1]Corresponding Author: guipumi@gmail.com

Theorem 2.1. *Let C be any copula. Let f and g be increasing almost everywhere functions. Then $C_{f(X),g(Y)}(u,v) = C_{X,Y}(u,v)$. Also, for any copula C, the partial derivatives $\frac{\partial}{\partial u}C(u,v)$ and $\frac{\partial}{\partial v}C(u,v)$ exist almost everywhere, are almost everywhere non-negative and bounded by 1.*

The next theorem, known as Sklar's theorem, is the key result for copulas and elucidate the role played by them. See Nelsen (2005) for a proof.

Theorem 2.2 (Sklar's Theorem)**.** *Let X and Y be random variables with marginals F and G, respectively, and joint distribution function H. Then, there exists a copula C such that,*

$$H(x,y) = C\big(F(x), G(y)\big), \quad \forall (x,y) \in \mathbb{R}^2.$$

If the F and G are continuous, then C is unique. Otherwise, C is uniquely determined on $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$*. The converse also holds. Furthermore,*

$$C(u,v) = H\big(F_1^{(-1)}(u), G^{(-1)}(v)\big), \quad \forall (u,v) \in I^2,$$

where $F^{(-1)}$ and $G^{(-1)}$ denote the pseudo-inverse of F and G, respectively.

For an introduction to the subject, refer to Nelsen (2005). The theory of copulas is also intimate related with the theory of probabilistic metric spaces. See Schweizer and Sklar (2005) for more details in this matter.

In this work we will also assume that the reader is familiar with basic concepts related to Markov Chains and Brownian Motions. The results needed here can be found in most introductory texts in stochastic processes. See, for example, Karlin and Taylor (1975).

3 Copula-based Semiparametric Model

In this section we shall review and extend the ideas in Chen and Fan (2006). The set up we shall use in this section is the following: let $\{Y_t\}_{t=1}^n$ be a sample from a stationary first order Markov process and let F be the true invariant distribution. Suppose that F is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}$. Let C_θ be the true parametric copula associated to the pair (Y_{t-1}, Y_t) up to an unknown parameter θ and suppose that it is absolutely continuous with respect to the Lebesgue measure in I^2. Also suppose that C_θ is neither of Fréchet-Hoeffding's bounds.

In this setting, if one writes $U_t = F(Y_t)$ and consider $\{U_t\}_{t=1}^n$, the transformed process is still a stationary Markov process, the joint distribution of (U_{t-1}, U_t) will be $C_\theta(u,v)$ and the conditional density of $U_t|U_{t-1} = u_0$ will be $f_{U_t|U_{t-1}=u_0}(u) := c_\theta(u_0, u)$. This implies that this setting is consistent with the following semiparametric generalized linear model

$$\varphi_\alpha(U_t) = \psi_\beta(U_{t-1}) + \varepsilon_t, \qquad E(\varepsilon_t|U_{t-1}) = 0, \tag{3.1}$$

where φ_α is an increasing parametric function,

$$\psi_\beta(u) = E\Big(\varphi_\alpha(U_t)\big|U_{t-1} = u\Big) = \int_0^1 \varphi_\alpha(v)c_\theta(u,v)\mathrm{d}u$$

and the conditional density of ε_t given $U_{t-1} = u_0$ is given by

$$f_{\varepsilon_t|U_{t-1}=u_0}(x) = \frac{c_\theta\big(u_0, \varphi_\alpha^{-1}\big(x + \psi_\beta(u_0)\big)\big)}{\frac{d}{dx}\big(\varphi_\alpha\big(x + \psi_\beta(u_0)\big)\big)}.$$

In this set up, if one wants to specify, for instance, an AR(1) process, one can take the Gaussian copula, namely, $C_\rho(u,v) = \Phi_\rho\big(\Phi^{-1}(u), \Phi^{-1}(v)\big)$, where $\Phi_\rho(\cdot)$ is the distribution function of a bivariate normal random variable with correlation ρ and $\Phi(\cdot)$ is the distribution function of a standard normal random variable, and set $\varphi_\alpha(u) = \Phi^{-1}(u)$ and $F(u) = \Phi(u)$. Then one can show that (3.1) becomes

$$Y_t = \rho Y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, 1-\rho^2).$$

One way to simulate the structure such like (3.1) is by adapting the conditional distribution method (see Nelsen (2005)). The algorithm was developed in order to sample from any bivariate distribution given the conditional distributions. Suppose we want to sample from a certain bidimensional copula C. The algorithm is as follows:

1. Generate two independent uniform variates u_0, u_1.
2. Let $k_u(v) = \frac{\partial C(u,v)}{\partial u}$. Set $x = u_0$ and $y = k_{u_0}(u_1)$.
3. The desired pair is (x, y).

Now suppose we want to simulate a univariate sample $x_1, \cdots, x_n$ given that the conditional density of U_t given $U_{t-1} = u$ is $c(u, v)$. We can use the following adaptation of the conditional distribution method:

1. Generate n independents uniform variates $u_1, \cdots, u_n$.
2. Let $k_u(v) = \frac{\partial C(u,v)}{\partial u}$ and set $x_1 = u_1$.
3. For $i = 2, \cdots, n$, set $x_i = k^{-1}_{x_{i-1}}(u_i)$.
4. The desired sample is $x_1, \cdots, x_n$.

The above algorithm, which we shall call adapted conditional distribution method and abbreviate ACD, can be used to generate a wide variety of dependence structures. Here are some examples.

Exemple 3.1. The Frank family of copulas is composed by copulas of the form

$$C(u, v) = -\frac{1}{\theta} \ln\Big(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\Big),$$

for $\theta \in \mathbb{R} \setminus \{0\}$. Limiting cases are $C_{-\infty} = W$, $C_0 = \Pi$ and $C_\infty = M$. Figure 1 to 3 show the time series plot and the autocorrelation function of samples of size 1000 drawn from a Frank copula with parameter $\theta \in \{5, 10, 50\}$ and marginals $N(0, 1)$, t_2, t_7 and Exp(1) using the ACD method. Notice that as the parameter θ increases, the dependence also increases. Also notice the similarities between the normal and t_7 plots.

Exemple 3.2. The Farlie-Gumbel-Morgenstern family, abbreviated FGM, is composed by copulas of the form

$$C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v)),$$

for $\theta \in [-1, 1]$. As special case we have $C_0 = \Pi$. In clear contrast to the relative complexity of the Frank family, the simple form of the copulas in the FGM family do not allow a wide variety of dependence structures. In figure 4 we show the time series plot and the autocorrelation function of samples of size 1000 drawn from the FGM copula with parameter $\theta \in \{-0.2, 0.4\}$ and marginals $N(0, 1)$ and t_2 using the ACD method. Notice that for the FGM family, the time series plot with marginals t_2 and $N(0, 1)$ produce very different processes, in contrast to Example 3.1.

Other examples and more details will be presented in the full paper.

4 Application to Brownian Motion

In this section we will use the ACD method to simulate a one dimensional Brownian Motion from its copula. Let $\{W_s\}_{t>0}$ be a standard Brownian motion. It can be shown that for any time points $0 \le s < t$, the copula associated to (W_s, W_t), called the Brownian copula, is given by

$$C_{s,t}(u, v) = \int_0^u \Phi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(x)}{\sqrt{t - s}}\right) \mathrm{d}x, \tag{4.2}$$

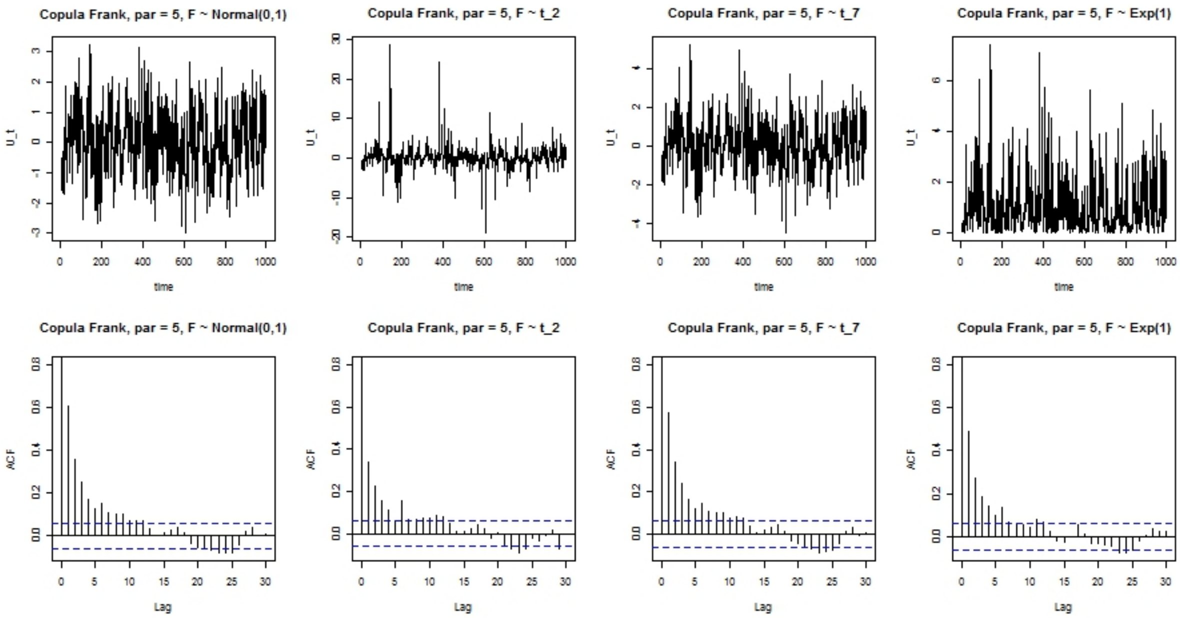

Figure 1: Time series plot and autocorrelation function of a sample generated from the Frank copula with $\theta = 5$ by the ACD method.

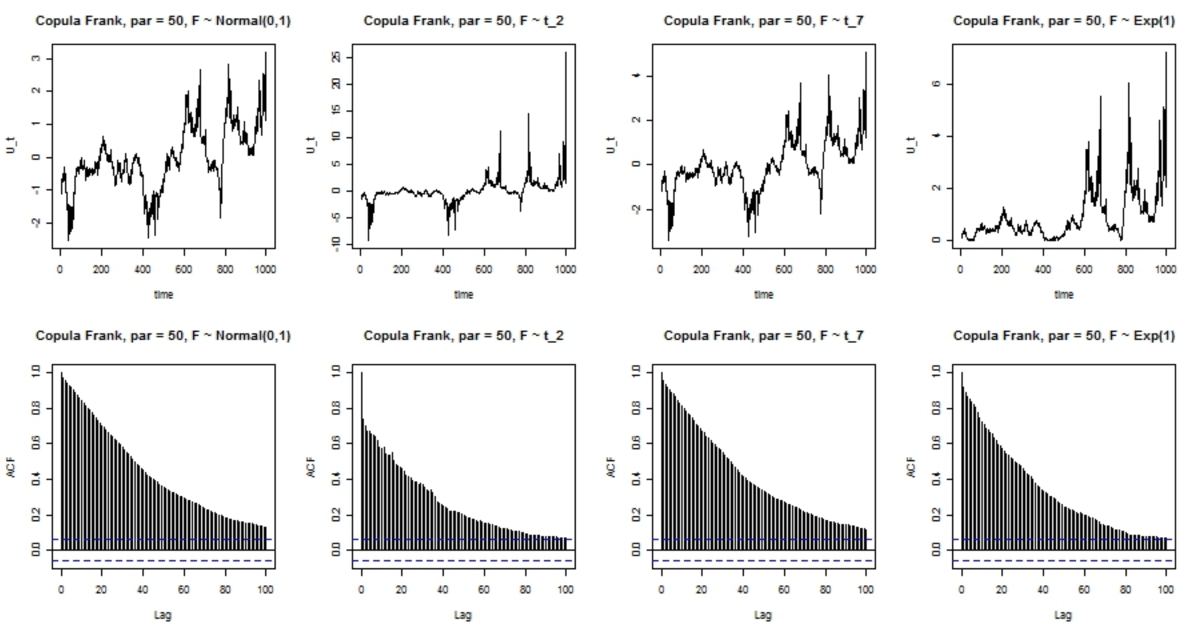

Figure 2: Time series plot and autocorrelation function of a sample generated from the Frank copula with $\theta = 50$ by the ACD method.

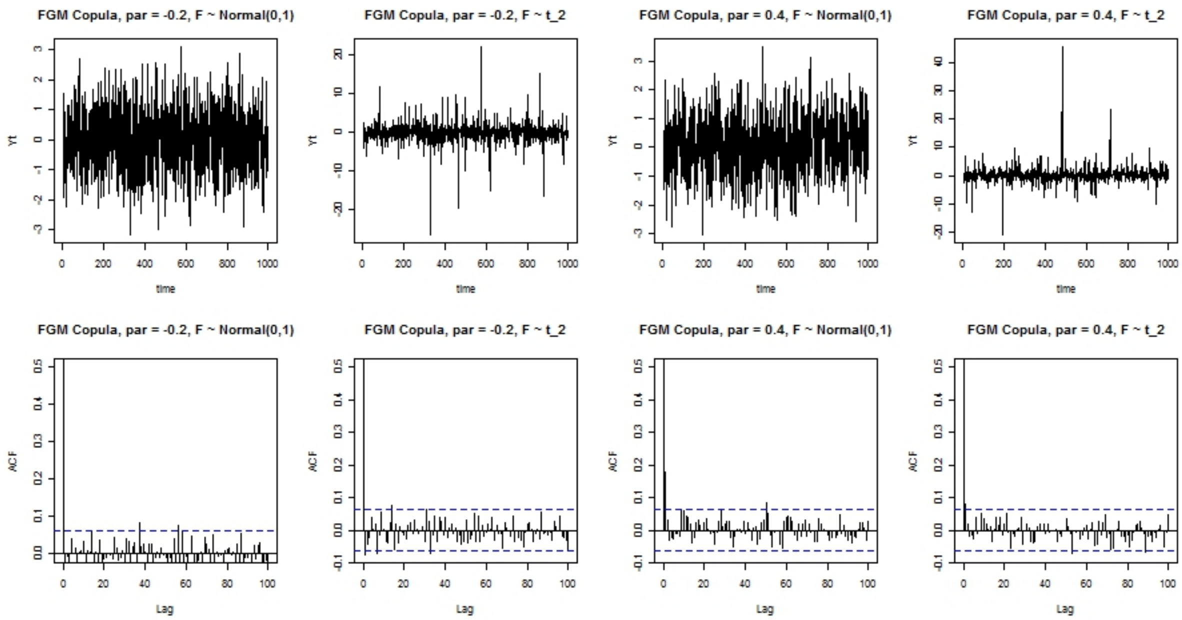

Figure 3: Time series plot and autocorrelation function of of a sample generated from the FGM copula with $\theta \in \{-0.2, 0.4\}$ and marginals $N(0,1)$ and t_2 by the ACD method.

so that

$$k_u^{-1}(v; s, t) = \Phi\left(\sqrt{1 - \frac{s}{t}}\Phi^{-1}(v) + \sqrt{\frac{s}{t}}\Phi^{-1}(u)\right).$$

In this case, for any $0 \leq s < t$ we can further adapt the ACD method in order to incorporate the time feature of the Brownian copula and simulate a Brownian motion by means of $C_{s,t}$ alone. The ACD algorithm become the following:

1. Let $0 \leq t_1 < \cdots < t_n$ be the grid of time points where the sample of the Brownian motion will be drawn. Generate n independent uniform variates $u_1, \cdots, u_n$.
2. Set $x_1 = u_1$.
3. For $i = 2, \cdots, n$ set $x_i = k_{x_{i-1}}^{-1}(u_i; t_{i-1}, t_i)$.
4. The desired sample is $x_1, \cdots, x_n$.

Notice that the ACD should give us a series with the temporal dependence and common behavior of a Brownian motion, but the sample will be, evidently, limited to $[0, 1]$. In order to simulate an usual Brownian motion one need to apply the probability integral transform and write $w_i = \Phi^{-1}(x_i)$, then w_i will be a "true" Brownian motion. One can also write $y_i = G^{-1}(x_i)$ for some distribution function G other than Gaussian in order to get a stochastic process with the behavior of a Brownian motion but different marginal distributions. We illustrate this in Figure 5, where we show the time series plot of a realization of the Brownian copula (4.2) with six different marginals distribution: $U(0,1)$ (the copula itself), $N(0,1)$, t_2, Exp(1), Cauchy(1) and χ^2_5. We choose to start at 1 and the time grid is equally spaced with a step of 0.1. Notice that all the plots show similar probabilistic behavior, but the scales and ranges differs, depending on how heavy are the tails of the distribution. However, the only "true" Brownian motion is the one with $N(0,1)$ marginals.

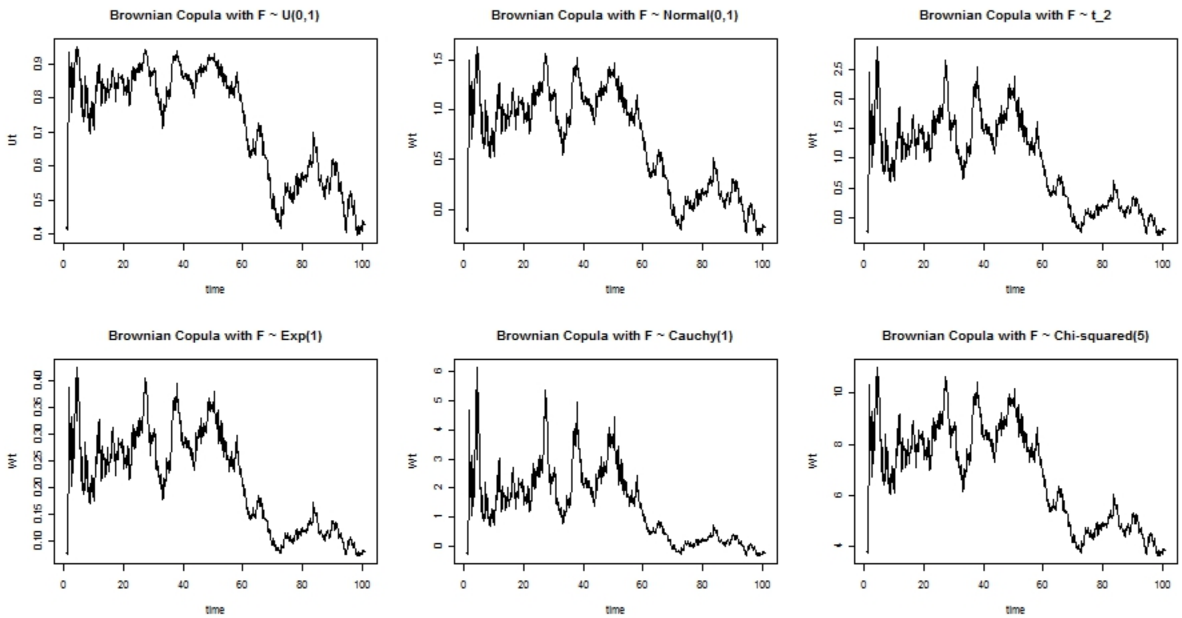

Figure 4: Time series plot of a sample generated by the ACD method from the Brownian copula with different marginals.

5 Conclusions

In this work we discussed some methods for simulating univariate time series based on bidimensional copulas. We analyzed an adaptation of the conditional distribution method, which we called ACD and applied it to cases where the sample comes from a Markov Chain and the copulas are compatible with some generalized linear model problem. The algorithm arises naturally in the study of copula-based time series in the context of Chen and Fan (2006). We also presented an adaptation for the case where the copula also depends on temporal features. With this adaptation we could generate time series whose probabilistic behavior is similar to a Brownian motion, but the marginal distributions are non-Gaussian and can be arbitraily chosen.

Acknowledgements

Guilherme Pumi was supported by CNPq-Brazil. Sílvia R.C. Lopes' research was partially supported by CNPq-Brazil, by CAPES-Brazil, by INCT *em Matemática* and also by Pronex *Probabilidade e Processos Estocásticos* - E-26/170.008/2008 -APQ1.

References

1. Chen, X. and Fan, Y. (2006) "Estimation of Copula-Based Semiparametric Time Series Models". *Journal of Econometrics,* **130**, 307-335.

2. Karlin, S. and Taylor, H. M. (1975). *A First Course in Stochastic Processes.* San Diego: Academic Press, 2nd edition.

3. Nelsen, R. B. (2005). *An Introduction to Copulas.* New York: Springer Verlag, 2nd edition.

4. Schweizer, B. and Sklar, A. (2005). *Probabilistic Metric Spaces.* Mineola: Dover Publications.