Introduction to Copulas

Mark S. Tenney
Mathematical Finance Company

July 18, 2003

Copyright ©2003 All rights reserved Mark S. Tenney. 4313 Lawrence Street, Alexandria Virginia, 22309. Phone number 703 799 0518.
http://www.mathematical-finance.com

Contents

1 Distributions

We closely follow Chapter 2 of Nelsen [2] and Chapter 2 of Embrechts, Lindskog and McNeil [1].

Definition 1.1 (One Variable Distribution Function). The probability that a random variable is less than or equal to z is F(z). F(z) is between 0 and 1. The variable z is the random outcome and z is called a random variable.

An example is the normal distribution. We note that it need not have mean 0 or variance 1.

Definition 1.2 (Normal Distribution Function). We define the cumulative normal distribution or just distribution function N by

$$N(z; \mu, \sigma^2) = \int_{-\infty}^{z} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

We can write $F(z) = N(z; \mu, \sigma^2)$ to illustrate our earlier notation F.

Definition 1.3 (Multifactor Distribution Function). The joint probability that the i-th random variable is less than or equal to z_i for i=1,...,n is $F(z_1, ..., z_n)$. $F(z_1, ..., z_n)$ is between 0 and 1. The vector $(z_1, ..., z_n)$ is the random outcome and is a random variable.

Definition 1.4 (One factor Marginal Distribution Function). The probability that the i-th random variable is less than or equal to z_i is $F_i(z_i)$. $F(z_i)$ is between 0 and 1. $F(z_i) = \int F(z_1, ..., z_i, .., z_n) dV(i)$ where $dV(i)$ symolizes integration over all variables except z_i.

In addition to the one factor marginals, there are marginals of dimension k for any k between 1 and n-1. All of these are distributions as well.

2 Some Copulas

Definition 2.1 (Bivariate Independent Copula). The Bivariate Independent Copula

$$C(u_1, u_2) = u_1 u_2 \tag{2}$$

The next example is from p25 of Embrechts, Linkskog and McNeil [1]

Definition 2.2 (Bivariate Normal Copula). The Bivariate Normal Copula is

$$C_R^{Ga}(u_1,u_2) = \int_{-\infty}^{N^{-1}(u_1;0,0)} \int_{-\infty}^{N^{-1}(u_2;0,0)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} e^{-\frac{s^2-2R_{12}st+t^2}{2(1-R_{12}^2)}} dsdt \tag{3}$$

The next example is from p26 of Embrechts, Linkskog and McNeil [1]

Definition 2.3 (Bivariate t-Copula). The Bivariate t-Copula is function N by

$$C_{\nu,R}^{t}(u_1,u_2) = \int_{-\infty}^{t_\nu^{-1}(u_1;0,0)} \int_{-\infty}^{t_\nu^{-1}(u_2;0,0)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} (1+\frac{s^2-2R_{12}st+t^2}{\nu(1-R_{12}^2)})^{-(\nu+2)/2} dsdt \tag{4}$$

3 The Problem

C 3.1 (Problem). I have one factor distributions for a collection of variables but don't have a multifactor distribution.

C 3.2 (Goal). I want a multifactor distribution but I want to keep my marginal distributions.

4 Uniform Distributions

Definition 4.1 (Uniform). By this we shall mean a random variable with an equal probability to fall in any subinterval of equal size of the interval 0 to 1.

Theorem 4.1 (Uniform's Distribution). The probability that the outcome of a draw from a uniform is between 0 and u, where u is between 0 and 1 is itself u. So if F is the cumualative density function, then F(u) = u for u between 0 and 1.

5 Transforming to Uniforms

C 5.1 (Solution Step 1). Transform each of the one factor distributions to be uniforms. This is done by setting $u_i = F_i(z_i)$. The random variable u_i is between 0 and 1 because F_i is.

6 Wish

C 6.1 (Wish). We wish we could take the u_i variables and just stick them into different choices for some acceptable joint distribution function.

C 6.2 (Wish Benefits). If our wish comes true, then we can transform the random outcomes $u_i, i = 1, ..., n$ back to z_i by using the inverse of the marginal cumulative distribution functions, $z_i = F^{-1}(u_i)$. Note that $F^{-1}(u_i)$ is defined for u_i from 0 to 1, but the output variable z_i can vary from minus infinity to plus infinity if that is the range of F^{-1}.

C 6.3 (Technical Nit Pick 1). For the general case, $F_i^{-1}(u_i)$ may have multiple values over certain ranges of u. In this case, you can pick any of those choices and get an acceptable joint distribution. Any such choice is called a quasi-inverse function.

C 6.4 (Wish Comes True). We can take the u_i variables and just stick them into different choices for some acceptable joint distribution function called a copula.

7 First Try: Defining Copula from Distribution

Definition 7.1 (Copula). We can take the u_i variables and form them into a joint distribution function $C(u_1, ..., u_n)$ that varies between 0 and 1. We have the interpretation that the joint probability that each u_i is less than or equal to U_i is $C(U_1, ..., U_n)$.

C 7.1 (How do I get a Copula?). We want to have a recipe for a function on n-variables that each are between 0 and 1 for it to be a Copula. We have defined a Copula as a joint probability distribution. Instead we want to define it in terms of a recipe and then have as a theorem that its a joint cumulative probability distribution. So let's start over with our definition of Copula. We start with definitions of special increasing functions. Those let us build our Copula definition from scratch.

8 2nd Try Defining Copula with Measure Theory

C 8.1 (Measure Theory Definition of Distribution Function). If we start with measure theory, we can define a joint probability distribution function as a set function with certain properties. We define a set function as one that maps sets to the non-negative real numbers. Measure theory has a triple of objects, a sample space of outcomes, a set of subsets of the sample space and a set function from each of these subsets to the real numbers. We require that the probability measure or set function satisfy the following for any set in the collection of subsets:

1. The probability of any set is between 0 and 1 inclusive.
2. The probability of the null set is 0.
3. The probability of the entire sample space is 1.
4. The probability of a countable union of disjoint subsets of the sample space equals the sum of their probabilities.

Using measure theory gets us to cumulative distribution functions that are valid. The cumulative distribution function is the set function defined above. This is the more general approach to increasing set functions. In measure theory, we use the subset relation to define increasing function, instead of a distribution. Measure theory allows us to work with broader sets and to deal with combinations of discrete and continuous probability or point mass probabilities in the middle of continuous ranges. We can do this with increasing functions by using the Riemann-Stieltjes integral. Measure theory uses the Lesbesgue integral which is already "Stieltjified".

C 8.2 (Defining Copula from Measure Theory). We now consider a distribution defined on the unit hypercube in n-dimensions. Such a distribution is a Copula.

C 8.3 (Special Increasing Functions instead of Measure Theory). Rather than go through the complications of measure theory, we use these special increasing functions. If we don't have technical problems we avoid having to use the terminology of measure theory which is more abstract and general. We also get recipes for making joint distribution functions for the continuous outcome case with no point masses without the fuss of measure theory.

9 Special Increasing Functions

C 9.1 (Special Increasing Function One Dimension). In one dimension we need an increasing function that is:

1. Non-negative

2. That starts at 0 and goes to 1 as we vary the input variable from its minimum value, possibly minus infinity, to its maximum value, possibly positive infinity.

C 9.2 (Special Increasing Function Many Dimensions). In two or more dimensions we need:

1. An increasing function that is
2. non-negative
3. and that starts at 0 when all the variables are at the minimum of their range and increases as we increase any one of them holding the others constant.
4. If we get to the maximum of all the variables we want the function to be 1.
5. If we put all the variables but one of them to be at their maximums we want to get a special one

dimensional increasing function.

C 9.3 (Special Increasing Functions are Distribution Functions). Special increasing functions are distribution functions.

10 3rd Try: Defining Copula with Special Increasing Functions

Definition 10.1 (Copula Redefinition). We take a special multivariate increasing function defined on the range of each input variable from 0 to 1. We don't assume these are distribution functions, instead we prove they have the properties of them, i.e. are acceptable for probability theory. If we are measure theorists this means proving there exist random variables which are defined in measure theory terms and which have a special increasing function as its cumulative distribution function.

C 10.1 (The marginals of a Copula are distributions). The marginals of a Copula are distributions.

11 Distribution to Copula

We now assume we are using the increasing function definition of a Copula.

Theorem 11.1 (Sklar's Theorem Part 1). Let $F(z_1, ..., z_n)$ be the joint distribution with margins $F_i(z_i)$, and let $F_i^{-1}(u_i)$ be quasi-inverses, then there exists a copula $C(u_1, ..., u_n)$

$$C(u_1, u_2, ..., u_n) = F^{-1}(F_1^{-1}(u_1), F_2^{-1}(u_2), ..., F_n^{-1}(u_n)) \tag{5}$$

If the F_i are continuous then C is unique.

If the F_i are not continuous, there are some technicalities that relate to what are called sub-copulas and the range of the corresponding variables.

12 Copula to Distribution

We continue to assume we are using the increasing function definition of a Copula.

Theorem 12.1 (Sklar's Theorem Part 2). Let $C(u_1, ..., u_n)$ be a Copula and assume that $F_i(z_i)$ are distribution functions. Then there exists a joint distribution function $F(z_1, ..., z_n)$ given by

$$F(z_1, ..., z_n) = C(F_1(z_1), F_2(z_2), ..., F_n(z_n)) \tag{6}$$

and the $F_i(z_i)$ are the marginal distribution functions.

13 URL's

XIII – 1 Embrechts

Paul Embrechts does research in stochastic finance and insurance.

http://www.math.ethz.ch/~embrechts/

We in part follow Chapter 2 of his paper on Copulas, "Modelling Dependence with Copulas and Applications to Risk Management" [1] is available at the URL:

http://www.math.ethz.ch/~baltes/ftp/copchapter.pdf

XIII – 2 Nelsen

An old URL for Nelsen is:

http://www.cs.bsu.edu/~rnelson/

His new one is:

http://www.lclark.edu/~mathsci/nelsen.html

We in part follow chapter 2 of Nelsen's book [2].

The errata for his book is at

http://www.lclark.edu/~mathsci/errata.pdf

The book's homepage at Springer-Verlag is

http://www.springer-ny.com/detail.tpl?ISBN=0387986235

XIII – 3 Venter

A good introduction to applying copulas to reinsurance is by Gary Venter [3]. This has many good pictures of copulas. This is available at the URL:

http://www.casact.com/coneduc/reinsure/2003/handouts/venter1.pdf

References

[1] Paul Embrechts, Filip Lindskog, and Alexander McNeil. Correlation and dependence in risk management: properties and pitfalls. In M.A.H. Dempster, editor, *Risk Management: Value at Risk and Beyond*. Cambridge University Press, Cambridge University Press, 2002.

[2] Roger B. Nelsen. *An Introduction to Copulas*. Springer-Verlag, New York, 1998.

[3] Gary Venter. Quantifying correlation with copulas. *Guy Carpenter*, 2003.