Tails of Copulas

Abstract: Actuaries who want to model correlated joint distributions have a choice of quite a few copulas, but little basis for choosing one over another. Methods are provided here to describe the features of different copulas, so that more informed choices can be made.

Copulas differ not so much in the degree of association they provide, but rather in which part of the distributions the association is strongest. Often needed for property and casualty applications are copulas that emphasize correlation among large losses, i.e., in the right tails of the distributions. Several copulas that do this are discussed.

To describe aspects of the copulas, univariate functions of copulas are introduced, for example tail concentration functions. These descriptive functions can be thought of as an intermediate step between correlation coefficients, such as Kendall, Spearman, Gini, etc., which are zero-dimensional measures of association, and the multi-dimensional copula function itself.

The descriptive functions can be used to select copulas having desired characteristics, such as tail concentration, and they can also be used in the fitting process to judge how well the fitted copulas match those aspects of the data.

Tails of Copulas

What are copulas?

Copulas provide a convenient way to express joint distributions of two or more random variables. With a copula you can separate the joint distribution into two contributions: the marginal distributions of each variable by itself, and the copula that combines these into a joint distribution. One basic result is that any joint distribution can be expressed in this manner. Another convenience is that the conditional distributions can be readily expressed using the copula.

Some measures of association depend only on the copula and not on the marginal distributions. Both Kendall's tau and Spearman's rank correlation are examples, but the usual Pearson linear product-moment correlation depends on the marginal distributions. Correlation coefficients measure the overall strength of the association, but give no information about how that varies across the distribution. Through the choice of copula, a good deal of control can be exercised over what parts of the distributions the variables are more strongly associated. One aspect emphasized below is controlling the strength of the relationship in the tails of the distributions. For instance, workers compensation and property losses might be correlated in the extreme tails, but not elsewhere in the distributions, and there are copulas with this kind of behavior.

A previous PCAS example of the use of copulas was provided in Wang [8], who provided details of calculation methods for aggregate distributions, with some examples using copulas.

Technically, copulas are joint distributions of unit uniform variates. In application, the unit uniform variates are viewed as probabilities from some other variates. Then the joint distribution of those variates is produced from those probabilities using their individual inverse distribution functions. Copulas thus provide a ready method for describing joint distributions and simulating correlated variables. Quite a few copulas are available, and

they have differing characteristics that lead to different relationships among the variables generated.

This paper reviews several popular copulas, introduces some others, and also introduces methods for selecting which copulas may be most appropriate for a given application. In particular, the behavior of the copulas in the right and left tails can be used to distinguish among joint distributions that produce the same overall correlation.

The organization of the paper is first to review copula methods in general, then to examine several specific copulas, and finally to look at measures that can be used to identify key characteristics of copulas. An example is provided to illustrate how these measures are applied to some correlated loss data.

1. General Considerations

Copulas − Formal Definition

It would be convenient to be able to express a joint distribution function $F(x,y)$ as a function of $F_X(x)$ and $F_Y(y)$, the individual (or marginal) distribution functions for X and Y, i.e., as $F(x,y) = C(F_X(x),F_Y(y))$. To do this, C can be defined by $C(u,v) = F(F_X^{-1}(u),F_Y^{-1}(v))$. This gives $C(F_X(x),F_Y(y)) = F(F_X^{-1}(F_X(x)),F_Y^{-1}(F_Y(y))) = F(x,y)$. The function $C(u,v)$ is called a copula. For many bivariate distributions, the copula form is the easiest way to express and generate the joint probabilities. It allows a separate description of the individual distributions and their association. Copulas work in the multi-variate context also, but this paper will primarily look at bivariate copulas, especially those defined by a single parameter.

In this context, a copula is a joint distribution of two unit uniform random variates U and V with $C(u,v) = \Pr(U \leq u, V \leq v)$. Also, $c(u,v)$ will be used to denote the corresponding probability density, which is the mixed second partial derivative of $C(u,v)$. The simplest copula is the uniform density for independent draws, i.e., $c(u,v) = 1$, $C(u,v) = uv$. Two other simple copulas are $M(u,v) = \min(u,v)$ and $W(u,v) = (u+v-1)_+$, where the "+" means "zero if negative." A standard result, given for instance by Wang[8], is that for any copula

C, $W(u,v) \leq C(u,v) \leq M(u,v)$. M and W are called the Fréchet upper and lower bounds, respectively.

Conditioning with Copulas

The conditional distribution can be defined using copulas. Let $C_1(u,v)$ denote the derivative of $C(u,v)$ with respect to the first argument. When the joint distribution of X and Y is given by $F(x,y) = C(F_X(x),F_Y(y))$, then the conditional distribution of $Y | X=x$ is given by:

$$F_{Y|X}(y) = C_1(F_X(x),F_Y(y))$$

For example, in the independent case $C(u,v) = uv$, the conditional distribution of V given $U=u$ is $C_1(u,v) = v = \Pr(V<v | U=u)$. This is of course independent of u.

If C_1 is simple enough to invert algebraically, then the simulation of joint probabilities can be done using the derived conditional distribution. That is, first simulate a value of U, say u, then simulate a value of V from C_1, the conditional distribution of $V | U=u$.

Correlation

The linear correlation coefficient based on the covariance of two variates is not preserved by copulas. That is, two pairs of correlated variates with the same copula can have different correlations. However, the Kendall correlation, usually denoted by τ, is a constant of the copula. That is, any correlated variates with the same copula will have the τ of that copula.

There are different ways of defining τ, but the simplest may be $\tau = 4E[C(u,v)] - 1$. For independent variates with $C(u,v) = uv$, $E[C(u,v)] = \frac{1}{4}$ so $\tau = 0$. Also, for perfectly correlated variates $U = V$, $E[C(u,v)] = \frac{1}{2}$, so τ will be 1. Thus the scaling makes τ look like a correlation coefficient. The key measure though is $E[C(u,v)]$, which is a basic constant of a copula and generalizes to the case of several variates. The limiting values are obtained for the Fréchet upper and lower bound copulas, with $\tau = -1$ for W and $\tau = 1$ for M. These copulas thus express complete negative correlation and complete positive correlation, respectively.

2. Some Particular Copulas

Some well-known copulas and a few designed particularly for loss severity distributions are reviewed here.

Frank's Copula

Define $g_z = e^{-az} - 1$. Then Frank's copula with parameter $a \neq 0$ can be expressed as:

$$C(u,v) = -a^{-1}\ln[1 + g_u g_v/g_1], \text{ with conditional distribution}$$

$$C_1(u,v) = [g_u g_v + g_v]/[g_u g_v + g_1]$$

$$c(u,v) = -a g_1(1+g_{u+v})/(g_u g_v + g_1)^2 \text{ and Kendall's } \tau \text{ of}$$

$$\tau(a) = 1 - 4/a + 4/a^2 \int_0^a t/(e^t-1)\, dt$$

For $a<0$ this will give negative values of τ.

C_1 can be inverted, so correlated pairs u,v can be simulated using the conditional distribution. First simulate u and p by random draws on $[0,1]$. Here p is considered a draw from the conditional distribution of $V \mid u$. Since this has distribution function C_1, v can then be found as $v = C_1^{-1}(p \mid u)$. The formula for this, which can be found from the formula for C_1, is:

$$v = -a^{-1}\ln\{1 + p g_1/[1 + g_u(1-p)]\}$$

Once u and v have been simulated, the variables of interest X and Y can be simulated by inverting the marginal distributions, i.e., $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(v)$.

Gumbel Copula

This copula has more probability concentrated in the tails than does Frank's. It is also asymmetric, with more weight in the right tail. It is given by:

$$C(u,v) = \exp\{-[(-\ln u)^a + (-\ln v)^a]^{1/a}\}, \ a \geq 1.$$

$$C_1(u,v) = C(u,v)[(-\ln u)^a + (-\ln v)^a]^{-1+1/a}(-\ln u)^{a-1}/u$$

$$c(u,v) = C(u,v)u^{-1}v^{-1}[(-\ln u)^a + (-\ln v)^a]^{-2+2/a}[(\ln u)(\ln v)]^{a-1}\{1+(a-1)[(-\ln u)^a + (-\ln v)^a]^{-1/a}\}$$

$$\tau(a) = 1 - 1/a$$

Unfortunately, C_1 is not invertible, so another method is needed to simulate variates.

Embrechts, etal.[1] discuss the Gumbel copula and give a procedure to simulate uniform deviates from a general class of copulas to which it belongs. For the Gumbel this procedure starts by simulating two independent uniform deviates u and v, and then solving numerically for $1>s>0$ with $\ln(s)s = a(s-u)$. Then the pair $[\exp(\ln(s)v^{1/a}), \exp(\ln(s)(1-v)^{1/a})]$ will have the Gumbel copula distribution.

Heavy Right Tail Copula and Joint Burr

For some applications actuaries need a copula with less correlation in the left tail, but high correlation in the right tail, i.e., for the large losses. Here is one:

$C(u,v) = u + v - 1 + [(1-u)^{-1/a} + (1-v)^{-1/a} - 1]^{-a}$ $a>0$

$C_1(u,v) = 1 - [(1-u)^{-1/a} + (1-v)^{-1/a} - 1]^{-a-1}(1-u)^{-1-1/a}$

$c(u,v) = (1+1/a)[(1-u)^{-1/a} + (1-v)^{-1/a} - 1]^{-a-2}[(1-u)(1-v)]^{-1-1/a}$

$\tau(a) = 1/(2a+1)$

The conditional distribution given by the derivative $C_1(u,v)$ can be solved in closed form for v, so simulation can be done by conditional distributions as in Frank's copula.

Frees and Valdez [2] show how this copula can arise in the production of joint Pareto distributions through a common mixture process. Generalizing this slightly, a joint Burr distribution is produced when the a parameter of both Burrs is the same as that of the heavy right tail copula.

Given two Burr distributions, $F(x) = 1 - (1 + (x/b)^p)^{-a}$ and $G(y) = 1 - (1 + (y/d)^q)^{-a}$, the joint Burr distribution from the heavy right tail copula is:

$$F(x,y) = 1 - (1 + (x/b)^p)^{-a} - (1 + (y/d)^q)^{-a} + [1 + (x/b)^p + (y/d)^q]^{-a}$$

The conditional distribution of y|X=x is also Burr:

$$F_{Y|X}(y|x) = 1 - [1 + (y/d_x)^q]^{-(a+1)}, \text{ where } d_x = d[1 + (x/b)^{p/q}]$$

6

By analogy to the joint normal, this can be called the joint Burr because the marginal and conditional distributions are all Burr. In practice, the degree of correlation can be set with the a parameter, leaving the p and q parameters to fit the tails, and b and d to set the scales of the two distributions.

The Normal Copula

Useful for its easy simulation method and generalized to multi-dimensions, the normal copula is lighter in the right tail than the Gumbel or HRT, but heavier than the Frank copula. The left tail is similar to the Gumbel.

To define the copula functions, let $N(x;m,v)$ denote the normal distribution function with mean m and variance v, $N(x)$ abbreviate $N(x;0,1)$, and $B(x,y;a)$ denote the bivariate standard normal distribution function with correlation = a. Also let $p(u)$ be the percentile function for the standard normal, so $N(p(u)) = u$. Then with parameter a, which is the normal correlation coefficient:

$C(u,v) = B(p(u),p(v);a)$
$C_1(u,v) = N(p(v);ap(u),1-a^2)$
$c(u,v) = 1/\{(1-a^2)^{0.5}\exp([a^2p(u)^2-2ap(u)p(v)+a^2p(v)^2]/[2(1-a^2)])\}$
$\tau(a) = 2\arcsin(a)/\pi$

The Kendall tau is somewhat less than a. The following table shows a few values.

a	0.15643	0.38268	0.70711	0.92388	0.98769
τ	0.10000	0.25000	0.50000	0.75000	0.90000

Simulation uses the conditional distribution C_1. Simulate $p(u)$ from a standard normal and then $p(v)$ from the conditional normal C_1. The standard normal distribution function can then be applied to these percentiles to get u and v.

Visualizing Copulas

The copula densities can be graphed as surface plots, and these are somewhat informative, but to get a better feeling for what the copulas will do in practice it is helpful to look at the joint distributions they produce from a standard sample distribution. The unit lognormal (where ln(x) is standard normal) is used for this in the contour plots of the joint densities for the copulas defined so far, using τ = .35.
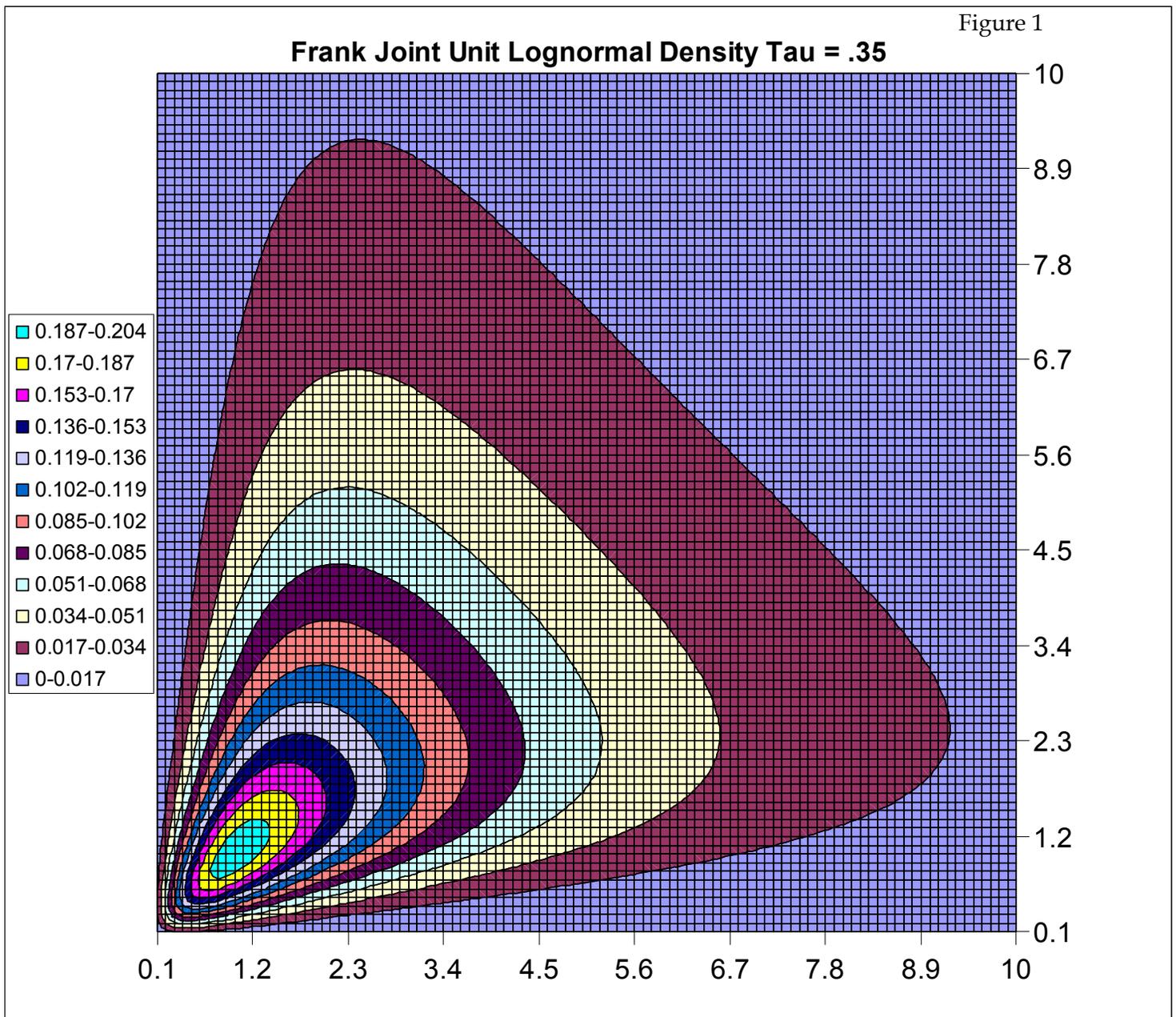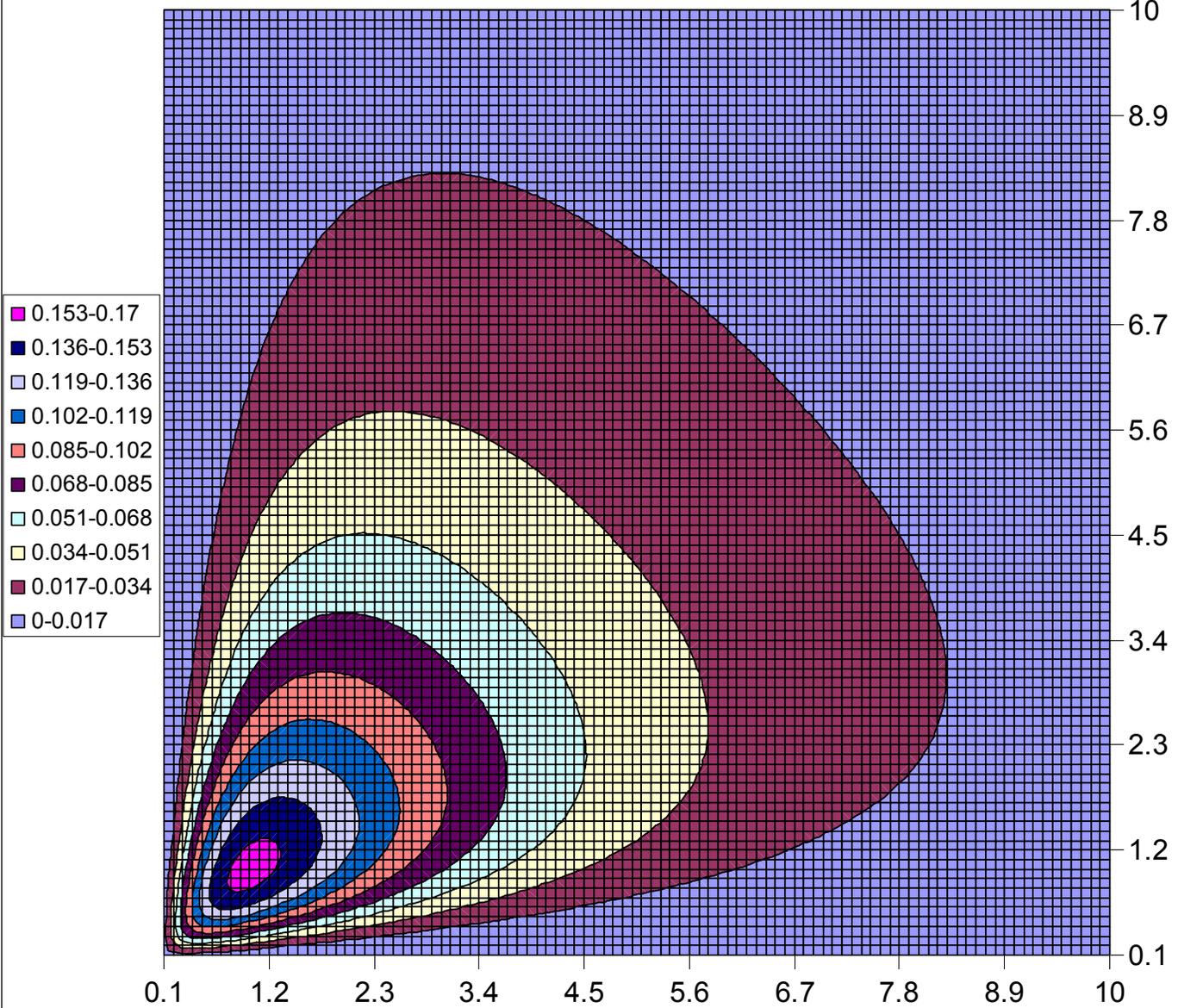
Frank Joint Unit Lognormal Density Tau = .35

Legend:
- 0.187-0.204
- 0.17-0.187
- 0.153-0.17
- 0.136-0.153
- 0.119-0.136
- 0.102-0.119
- 0.085-0.102
- 0.068-0.085
- 0.051-0.068
- 0.034-0.051
- 0.017-0.034
- 0-0.017

Figure 2

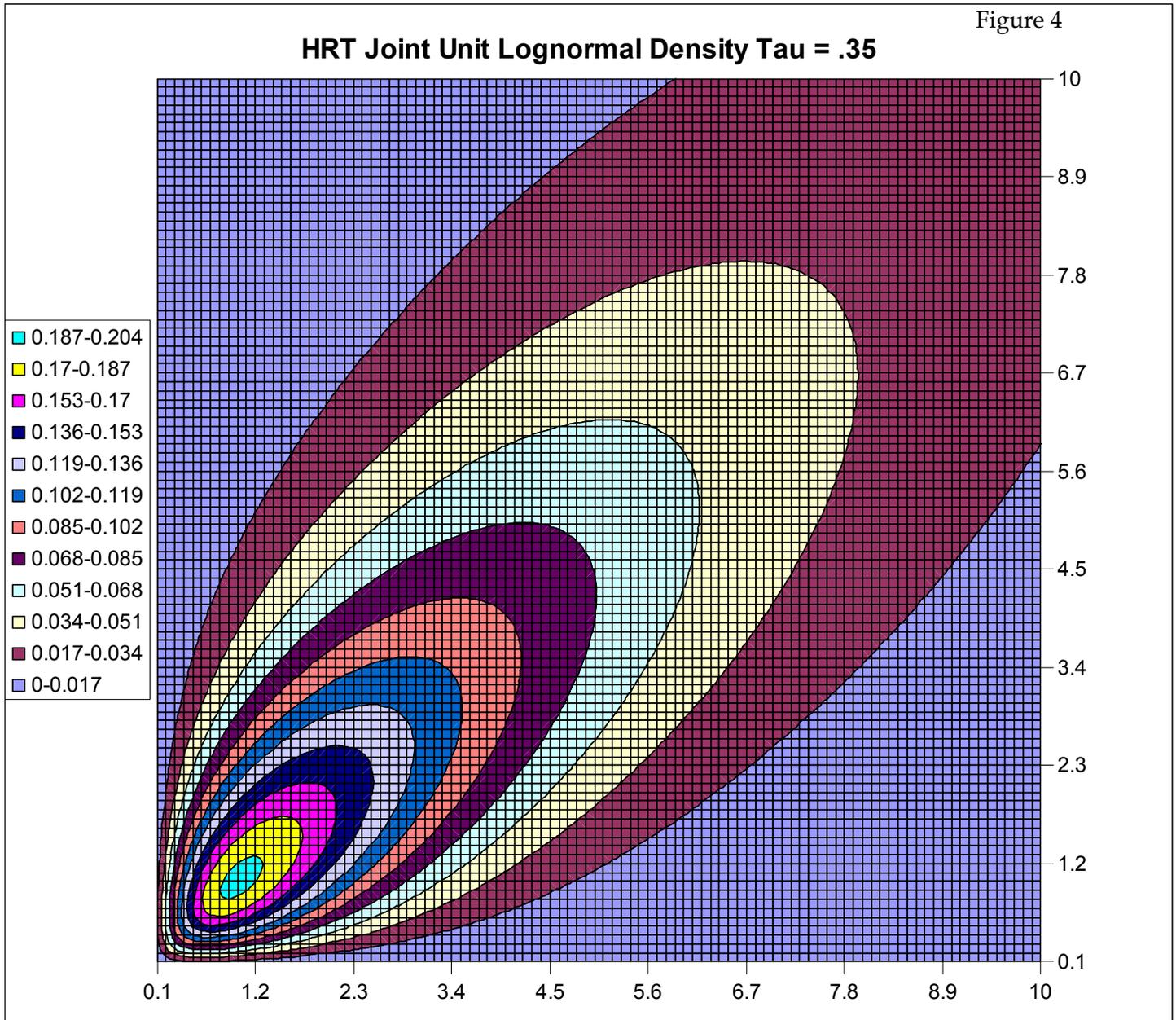Normal Joint Unit Lognormal Density Tau = .35

The normal and Frank copulas graphed in Figures 1 and 2 do not produce a strong relationship between large losses, although the normal shows a slightly stronger relationship, whereas the Frank is stronger around the mode.

Figure 3

Gumbel Joint Unit Lognormal Density Tau = .35
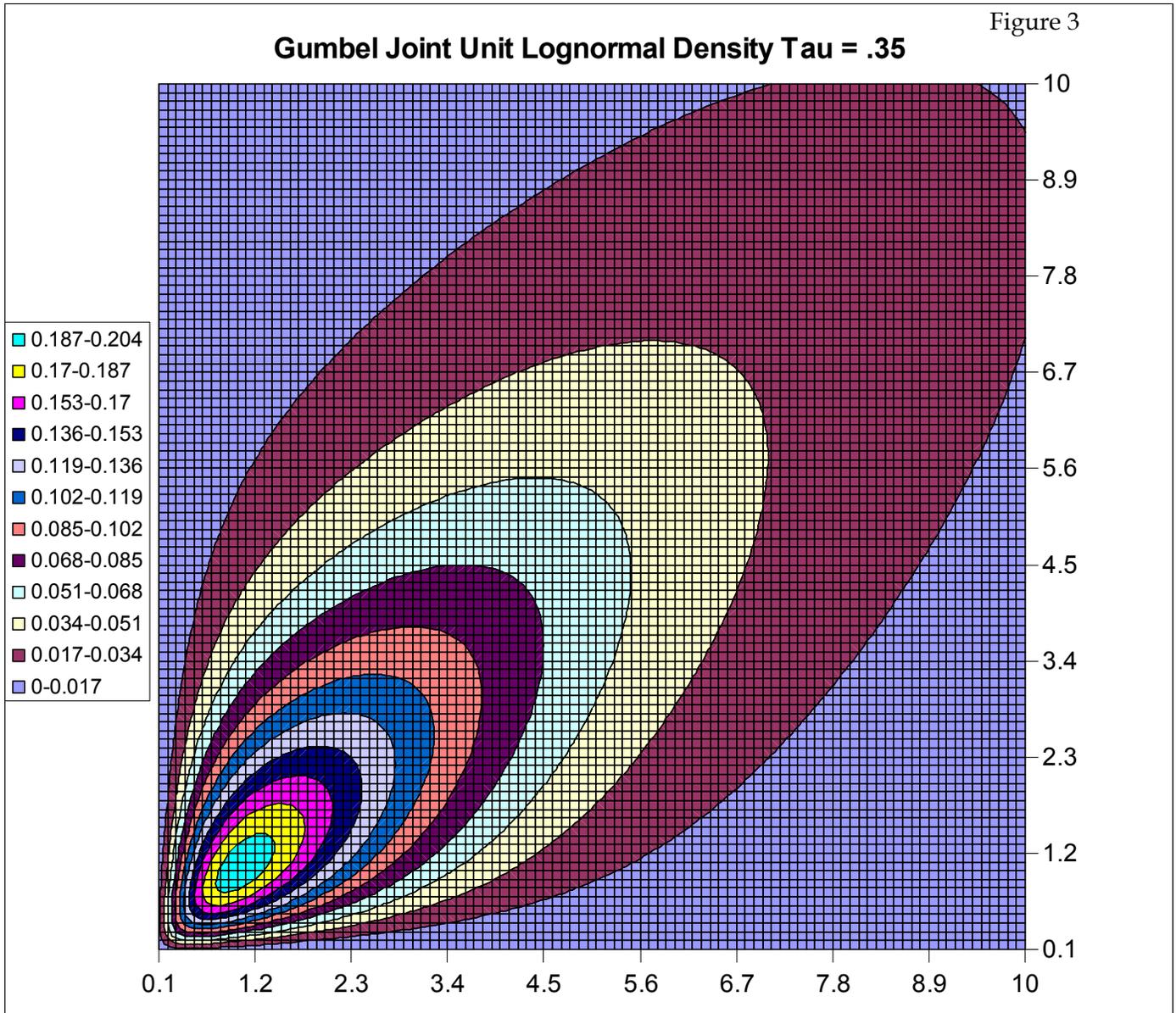
In contrast, the Gumbel copula keeps a strong relationship even for the large losses, as seen in the higher values of the density function in the upper right of Figure 3.
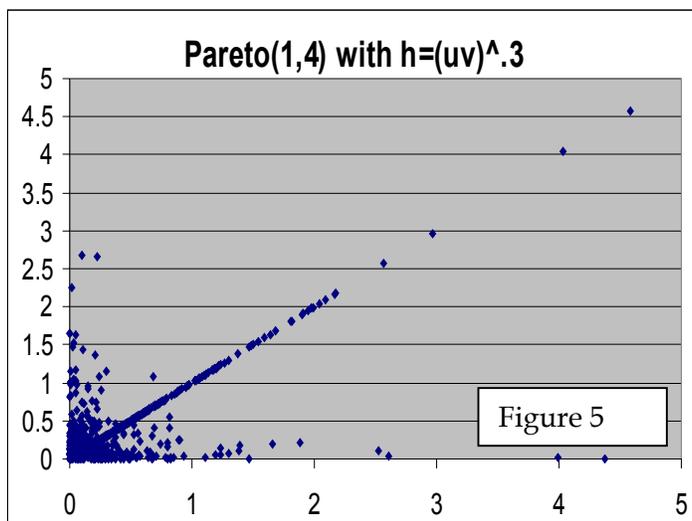
The heavy right tail (HRT) copula is even stronger in right tail correlation than is the Gumbel. While difficult to see in Figure 4, it is also weaker in the left tail. This will be more clear with the tail concentration functions discussed below.

Figure 4

HRT Joint Unit Lognormal Density Tau = .35

Legend:
- 0.187-0.204
- 0.17-0.187
- 0.153-0.17
- 0.136-0.153
- 0.119-0.136
- 0.102-0.119
- 0.085-0.102
- 0.068-0.085
- 0.051-0.068
- 0.034-0.051
- 0.017-0.034
- 0-0.017

Kreps' Partial Perfect Correlation Copula Generator

A family of copulas has been developed by Rodney Kreps [6]. This is based on a method for generating copulas that are mixtures of perfectly correlated and totally independent variates This is easier to describe as a simulation procedure, and then look at the copulas.

The basic idea is to draw two perfectly correlated deviates in some cases and two uncorrelated deviates otherwise. More specifically, let h(u,v) be a symmetric function of u and v mapping the unit square to the unit interval. To implement the simulation, draw three unit random deviates u, v, and w. If h(u,v) < w, simulate x and y as $F_X^{-1}(u)$ and $F_Y^{-1}(v)$ respectively. Otherwise take the same x but let y = $F_Y^{-1}(u)$ = x. Thus some draws are independent and some are perfectly correlated. The choice of the h function provides a lot of control over how often pairs will be correlated and what parts of the distributions are correlated.

For instance, h can be set to 0 or 1 in some interval like j < u,v < k to provide independence or perfect correlation in that interval, or it could be set to a constant p to provide correlation in 100p% of the cases in that interval. Another choice is h(u,v) = (uv)^a. This creates more correlation for larger values of u and v, with the parameter a controlling how much more.

Pareto(1,4) with h=(uv)^.3

Figure 5

Pareto(1,4) with h=(uv)^.3
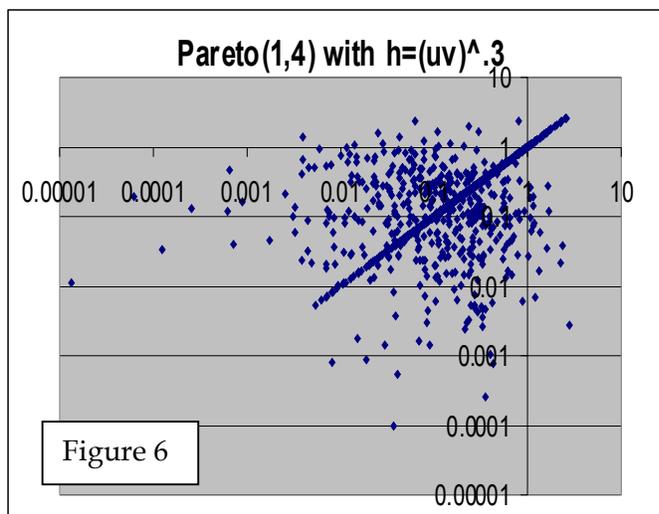
Figure 6

Figures 5 and 6 illustrate simulations in the case where $h(u,v) = (uv)^{0.3}$ and both X and Y are distributed Pareto with $F(x) = 1 - (1 + x)^{-4}$. The correlated and uncorrelated instances clearly show up separately, in either the log or regular scale.

For larger values of a, $h(u,v)$ is smaller, so it is less likely that $h(u,v)$ exceeds the random value w and thus less likely that the case u=v will be selected. For small values of a, on the other hand, $h(u,v)$ will be larger, approaching one as a goes to zero. Thus $h(u,v)>w$ is more likely, so u=v will also be more likely. The partial perfect correlation copula generator thus provides a good deal of flexibility and control over how much correlation is incorporated and where in the distribution it occurs.

To describe the copulas that result, it will be convenient to adopt the notation used in spreadsheets where a logical expression in parentheses will evaluate to zero if the expression is false and to one if it is true. Thus (u=v) is one if u=v and zero otherwise, etc.

Although Kreps considers more general situations, a relatively simple copula results in the case where $h(u,v)$ breaks out as a product of a univariate function evaluated at u and v, i.e., $h(u,v) = h(u)h(v)$. If we define $H(x) = \int_0^x h(t)dt$, the copula formulas become:

Figure 7

$C(u,v) = uv - H(u)H(v) + H(1)H(min(u,v))$
$C_1(u,v) = v - h(u)H(v) + H(1)h(u)(v>u)$
$c(u,v) = 1 - h(u)h(v) + H(1)h(u)(u=v)$
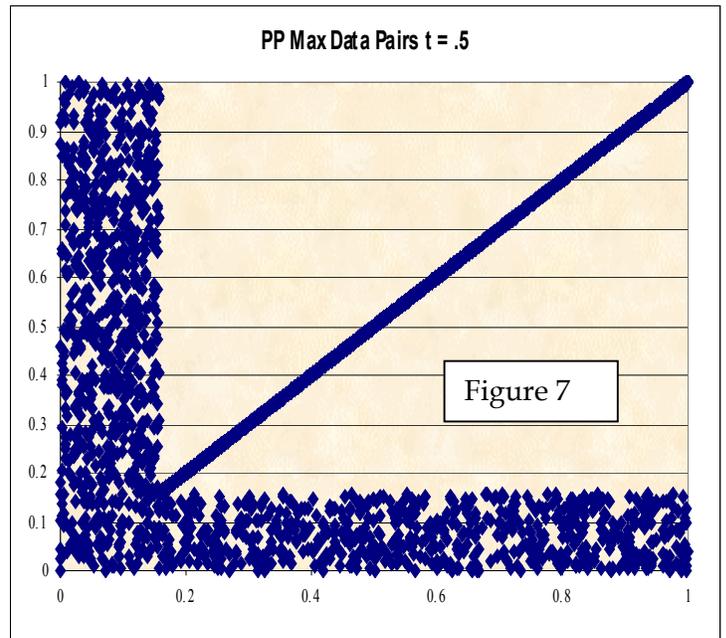
For a concrete example, pick an a between zero and one, and let $h(u) = (u>a)$. Thus if both u and v exceed a, the simulated values of u and v will be identical, and otherwise they will be independent. If $x>a$, $H(x) = \int_a^x dt = x - a$, and if not, $H(x) = 0$. Thus $H(u) =$

13

(u – a)(u>a). Also, H(1) = 1 – a, and H(min(u,v)) = [min(u,v) – a](u>a)(v>a). The copula formulas above can then be computed directly for this h. The Kendall correlation is $\tau(a)$ = $(1 – a)^4$. Sometimes this copula is called PP max, for partial perfect max function. The scatter plot of a simulated sample is graphed in Figure 7 for the case τ = ½.

Another example is to take h(u) = u^a. Then H(u) = $u^{a+1}/(a+1)$, and H(1) = 1/(a+1). Here, $\tau(a)$ = $1/[3(a+1)^4]$ + $8/[(a+1)(a+2)^2(a+3)]$. As a increases, this approaches zero, reflecting the fact that selecting u=v becomes less likely, and at a = 0, τ =1, as this gives the perfect correlation case.
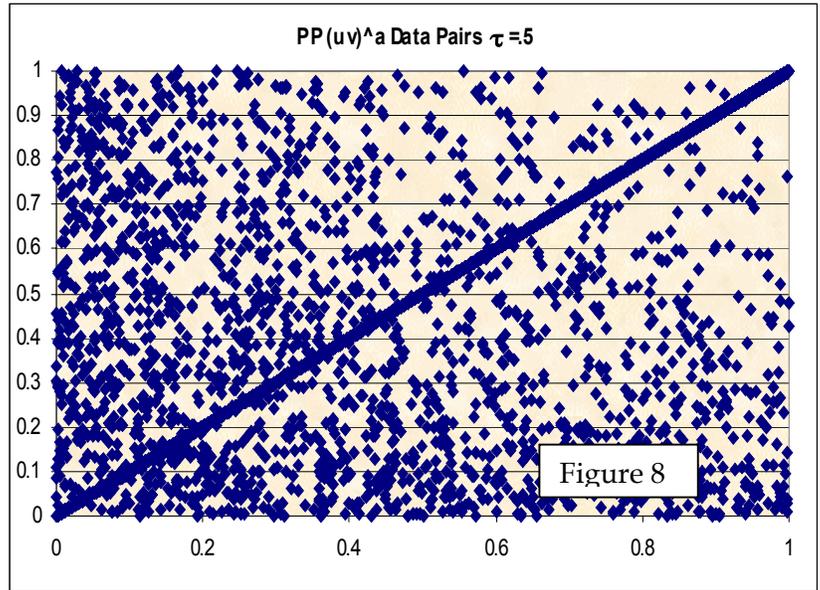

PP (uv)^a Data Pairs τ =.5

Figure 8

Figure 8 shows simulated pairs for the case τ = ½. More correlated pairs occur at higher values of u and v, as can be seen from the growing paucity of independent pairs when going to the upper right.

3. Distinguishing among copulas

A few functions are introduced here to help illustrate different properties that can distinguish the various copulas. These functions can also be approximated from data, and so can be used to assess which copulas more closely capture features of the data.

Tail Concentration Functions

Given a copula, right and left tail concentration functions can be defined with reference to how much probability is in regions near <1,1> and <0,0>. For any z in (0,1) define: L(z) = Pr(U<z,V<z)/z and R(z) = Pr(U>z,V>z)/(1 – z). In terms of the copula functions, L(z) is just C(z,z)/z. To calculate R(z), note that 1 - Pr(U>z,V>z) = Pr(U<z) + Pr(V<z) - Pr(U<z,V<z) = z + z – C(z,z). Then R(z) can be calculated by R(z) = [1 – 2z +C(z,z)]/(1 – z). Also, note that Pr(U<z,V<z) = Pr(U<z|V<z)Pr(V<z). But Pr(V<z) is just z, as copulas are

14

defined with uniform unit marginals, so L(z) = Pr(U<z|V<z) = Pr(V<z|U<z), and similarly R(z) = Pr(U>z|V>z). Joe [4] uses the term "upper tail dependence parameter" for R = R(1) = $\lim_{(z \to 1)}$R(z), and "lower tail dependence parameter" for L = L(0) = $\lim_{(z \to 0)}$L(z).

The left tail function approaches unity for z near 1, so does not distinguish much between copulas there, and similarly for the R function near 0. Thus they can be combined into an LR function which is L below ½ and R above ½.. This is graphed in Figure 9 for the copulas discussed above and for the Clayton copula, a heavy left tailed copula discussed later.

LR Functions for Tau = .35

Figure 9

Legend: Gum, HRT, Frank, Max, Power, Clay, Norm
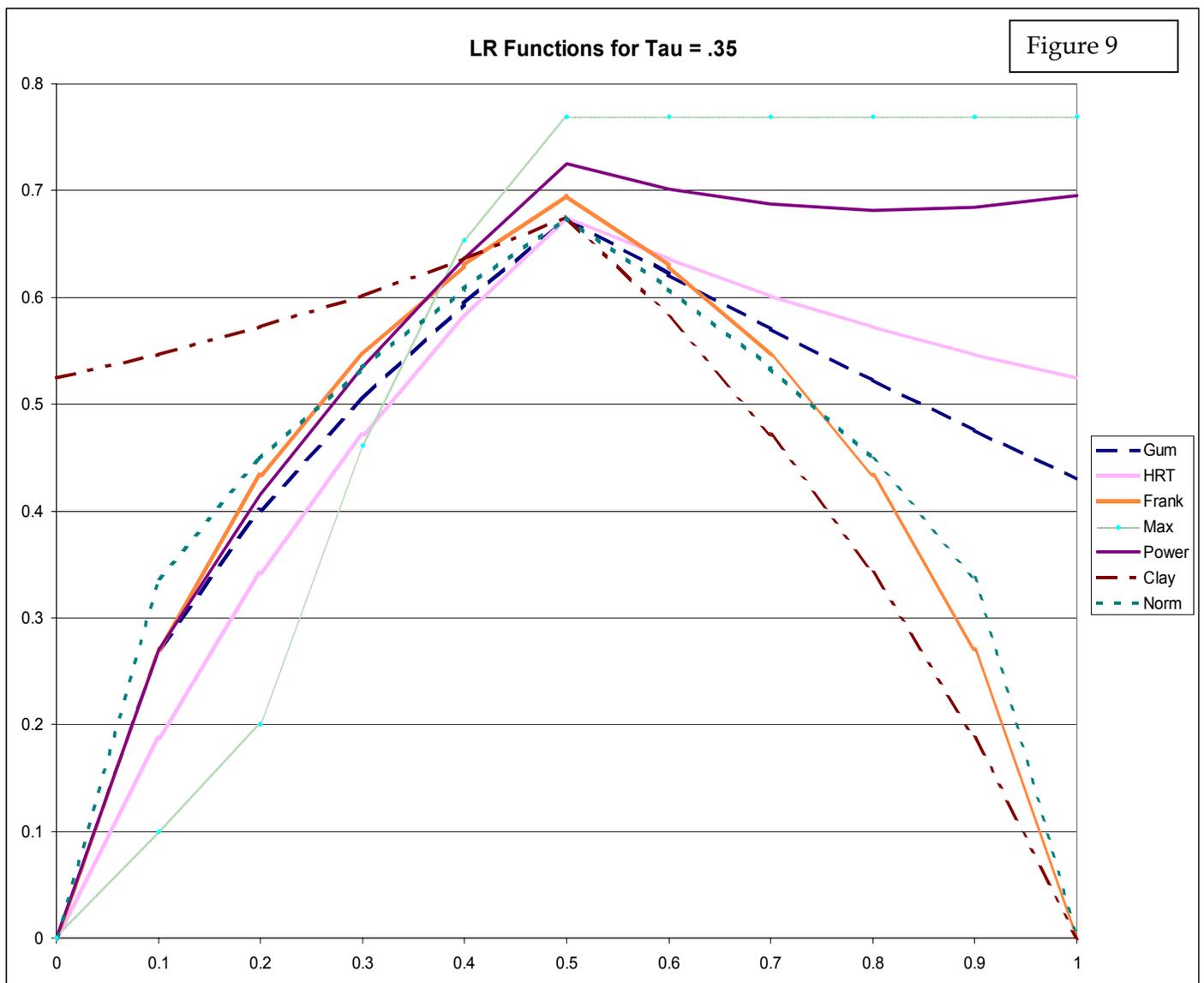
A basic feature that distinguishes copulas in the right tail is whether R=R(1) = 0 or something greater. The Gumbel, HRT, and partial perfect copulas all have R>0. The HRT is heavier in the right tail than the Gumbel, but less so than the partial perfect copulas. The Clayton is the only copula here showing positive left tail dependence. The HRT and PP Max copulas are very lowly dependent in the left tail. In fact, for the PP Max L(z) function the variates are independent in the left tail. Thus for low z, L(z) = 1. The normal and Frank copulas do not show tail dependence in the limits, but away from the extremes the normal shows greater tail concentration than the Frank on both sides.

For the four copulas with R > 0, R is shown below:

	Gumbel	HRT	PP Power	PP Max
R	$2 - 2^{1/a}$	2^{-a}	$1/(1+a)$	$1 - a$

Since R and τ are functions of the same parameter, they can be viewed as functions of each other. Once one is determined, the other is fixed for single-parameter copulas. Figure 10 graphs this relationship.

16
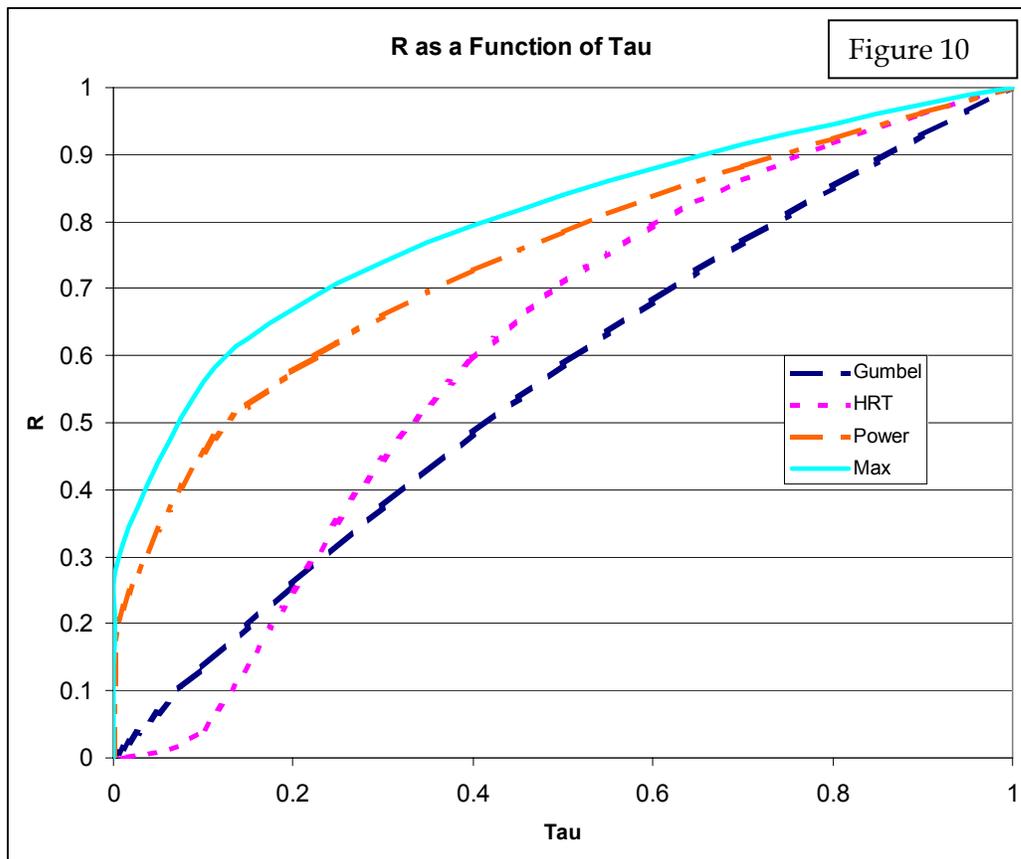
A good starting point for choosing a copula would be to look at the target pair <τ,R> and find which copula is closest. But since for the copulas above R is usually greater than τ, lower values of R would not be matched by any of them. R is somewhat tricky to determine for empirical data, as the far tail values have increasingly less data. Some projection of the lower values of R(z) might be necessary. Also the fitting should look at the R(z) function, not just R.
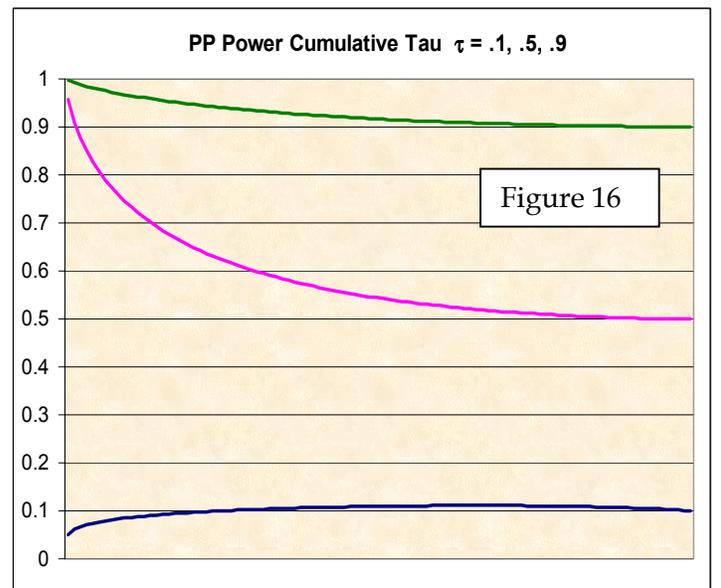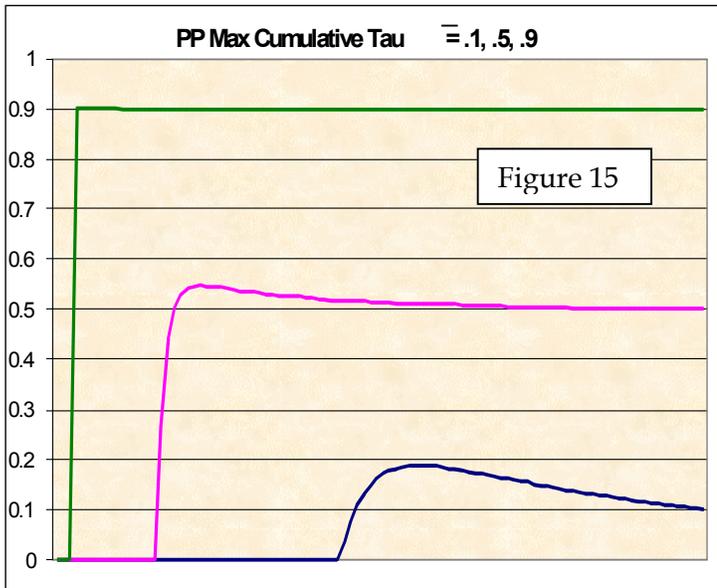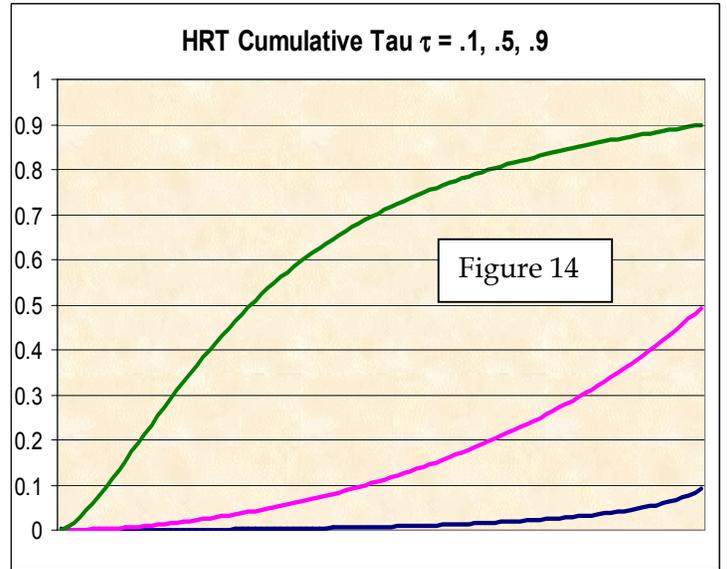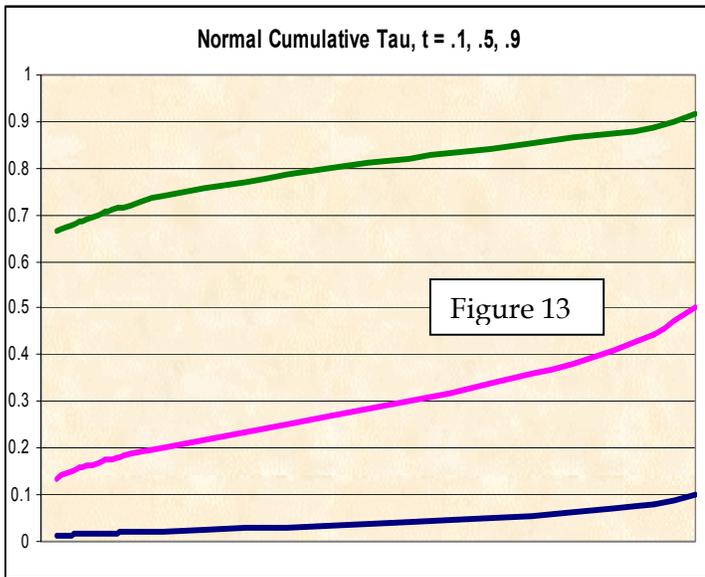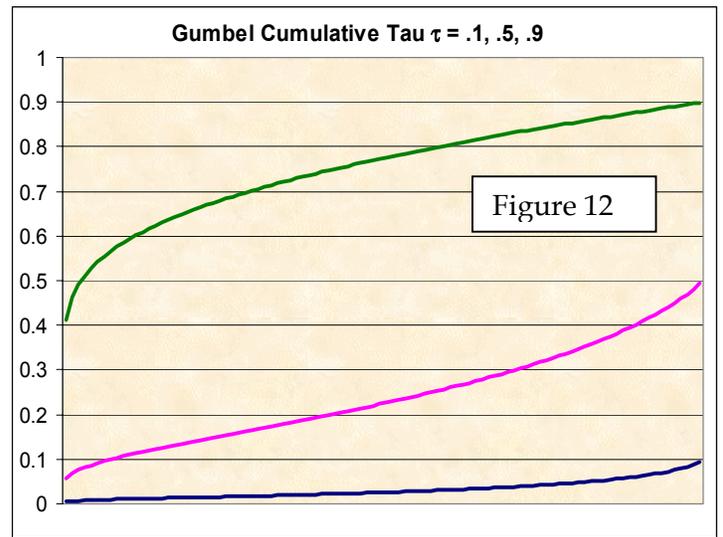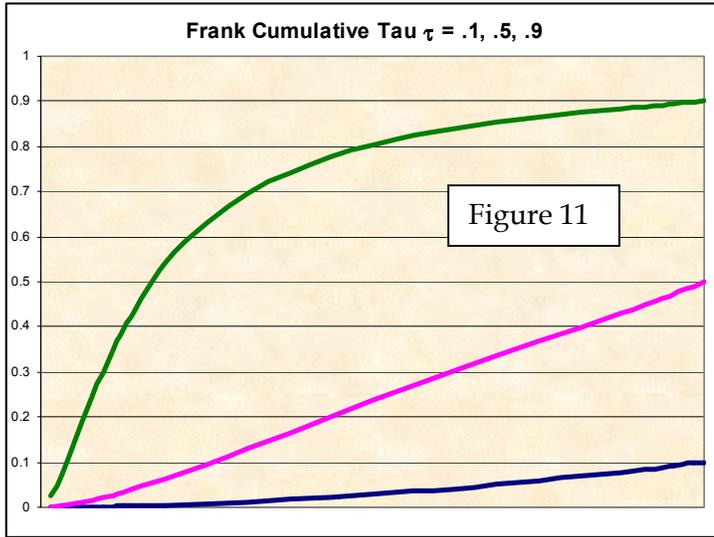
Cumulative Tau

Other descriptive functions can be defined that show different aspects of copulas. The cumulative tau function decomposes the integral defining the Kendall tau. Recall that tau is defined as $-1+4\int_0^1\int_0^1 C(u,v)c(u,v)dvdu$. The cumulative tau can be defined as J(z) = $-1+4\int_0^z\int_0^z C(u,v)c(u,v)dvdu/C(z,z)^2$.

The full double integral is a probability weighted average of C(u,v), i.e., EC(u,v). To compare to this on the square from (0,0) to (z,z), the partial integral has to be divided by the weights, hence the first power of C(z,z) in the denominator. This quotient will give the average value of C(u,v) in the square from (0,0) to (z,z). This will increase as a function of z for any copula. The second C(z,z) divisor expresses this average relative to C(z,z), i.e., shows how the average C compares to the maximal C in the square. This may or may not increase as a function of z, which makes it a more interesting property of the copula.

The normalization to the range of a correlation with the –1 and 4 is a matter of convenience and familiarity, and gives J(1) = τ. The integration can be done numerically, although for some copulas, formulas are given in Appendix A. The shape of the J function depends on the copula and the tau. It is graphed for several taus for each copula in Figures 11 – 16. All the graphs end up at τ for z=1, but can start off high or low, and can increase or decrease at varying rates.

Some other descriptive functions are discussed in Appendix B.

Figure 11 — Frank Cumulative Tau τ = .1, .5, .9

Figure 12 — Gumbel Cumulative Tau τ = .1, .5, .9

Figure 13 — Normal Cumulative Tau, t = .1, .5, .9

Figure 14 — HRT Cumulative Tau τ = .1, .5, .9

Figure 15 — PP Max Cumulative Tau $\overline{=}$ .1, .5, .9

Figure 16 — PP Power Cumulative Tau τ = .1, .5, .9

18

The notation $S(x) = 1 - F(x)$ is often used to describe the survival function $Pr(X>x)$. The joint survival function $S(x,y) = Pr(X>x, Y>y)$ is not $1 - F(x,y)$, however, as that would be the probability that either $X>x$ or $Y>y$, but not necessarily both. In fact, $S(x,y) = 1 - F_X(x) - F_Y(y) + F(x,y)$, i.e., $Pr(X>x, Y>y) = 1 - [Pr(X<x) + Pr(Y<y)] + Pr(X<x, Y<y)$.

Similarly for a copula $C(u,v) = Pr(U<u, V<v)$ the survival function of the copula, i.e., $C_S(u,v) = Pr(U>u, V>v)$, is $C_S(u,v) = 1 - u - v + C(u,v)$. Since $C(F_X(x),F_Y(y)) = F(x,y)$, we have $C_S(F_X(x),F_Y(y)) = S(x,y)$.

For a copula C, define $C_F(u,v) = C_S(1 - u, 1 - v) = u + v - 1 + C(1 - u, 1 - v)$. Then $C_F(S_X(x),S_Y(y)) = C_S(F_X(x),F_Y(y)) = S(x,y)$. Note that C_S is not a copula as it is zero at (1,1), but C_F is a copula. Call C_F the flipped copula of C. When the flipped copula is applied to the survival functions it gives the joint survival function for the copula. However, the flipped copula can be applied to distribution functions, and then it can have quite different properties than the original copula has. The next copula is an example.

Clayton's Copula

This copula has a heavy concentration of probability near (0,0) so it correlates small losses. It is not intuitively interesting for property-liability claims, but it may have some application.

$C(u,v) = [u^{-1/a} + v^{-1/a} - 1]^{-a} \qquad a>0$

$C_1(u,v) = u^{-1-1/a}[u^{-1/a} + v^{-1/a} - 1]^{-a-1}$

$c(u,v) = (1+1/a)[uv]^{-1-1/a}[u^{-1/a} + v^{-1/a} - 1]^{-a-2}$

$\tau(a) = 1/(2a + 1)$

What is interesting here is that the heavy right tail copula is actually the flipped Clayton copula. The tau is the same for both copulas[1], and the tail concentration functions are swapped. This is actually how the HRT copula was defined, and suggests defining other copulas by flipping known copulas. The copula would have to have some asymmetry to make this worthwhile. One candidate would be Gumbel's copula.

The Flipped Gumbel

Gumbel's copula is heavier in the right tail than the left. Flipping it would produce a copula with the opposite property:

$C(u,v) = u + v - 1 + \exp\{-([-\ln(1-u)]^a + [-\ln(1-v)]^a)^{1/a}\}$, $a \geq 1$.

$C_1(u,v) = 1 - \exp\{-([-\ln(1-u)]^a + [-\ln(1-v)]^a)^{1/a}\}\{[-\ln(1-u)]^a + [-\ln(1-v)]^a\}^{1/a\,-1}[-\ln(1-u)]^{a-1}/[1-u]$

$c(u,v) = (1-u)^{-1}(1-v)^{-1}\{[-\ln(1-u)]^a + [-\ln(1-v)]^a\}^{-2+1/a}[\ln(1-u)\ln(1-v)]^{a-1}$ x

$\qquad [a + \{[-\ln(1-u)]^a + [-\ln(1-v)]^a\}^{1/a} - 1]\exp\{-([-\ln(1-u)]^a + [-\ln(1-v)]^a)^{1/a}\}$

$\tau(a) = 1 - 1/a$

5. Applications

Loss Adjustment Expense

Two recent actuarial papers fit parameters to the joint distribution of loss and loss adjustment expense for a liability line using 1500 claims supplied by Insurance Services Office, Inc. The two studies may or may not have used the same data, but they present scatter plots that are similar. They both use copulas to describe the joint distribution.

There were a couple of methodological differences between the two papers. Frees and Valdez [2] assume Pareto marginals for both distributions, but compare fits for several copulas. Klugman and Parsa [5], on the other hand, compare fits for a number of severity distributions, but select Frank's copula arbitrarily. The papers may have taken different approaches to the censoring of claims by policy limits as well. Klugman and Parsa say

---

[1] Tau for a sample is the average value of sign[(u – x)(v – y)] among all distinct pairs (u,v), (x,y). This value is the same for the flipped pairs (1–u, 1–v), (1–x, 1–y), so tau will be the same for the original and the flipped sample for any copula.

they omit claims for which either loss or expense is zero, so they can get true severity distributions for both. Frees and Valdez probably do this as well.

Frees and Valdez used the K(z) function discussed in Appendix B to select among copulas. Plotting the empirical K(z) against the values from several copulas, they found the Gumbel looked best. The Gumbel also gave the best value for the Akaike information criterion (AIC). Optimizing the AIC is equivalent to finding the copula with the highest maximum likelihood in this case, as all the copulas they tried had one parameter. The best fit they found was produced by the Gumbel copula with a = 1.453. This gives τ = 0.31. Klugman and Parsa estimate the Frank a = 3.07438, which also gives τ = 0.31.

A convenient way to compare heavy-tailed severity fits is to look at the median and the heaviness of the tail, which can be quantified as the smallest positive moment that does not converge. For the Pareto, for example, this moment is just the shape parameter.

If we express the Pareto as $F(x) = 1 - (1+x/b)^{-a}$, then Frees and Valdez find: for loss, a = 1.122, b = 14,036, and for expense, a = 2.118 and b = 14,219. Klugman and Parsa find the best severity fits with the inverse Burr, which can be expressed as $F(x) = (1+(x/b)^{-c})^{-a}$. They estimate[2] for loss, a = 1.046 = c, b = 11,577.7, and for expense, a = 1.57658, b = 10,100.2, c = 0.573534. These parameters are converted to median and tail heaviness (= c for the inverse Burr) below. There is reasonably close agreement among these values except for the tail heaviness for loss expense, for which the divergence is a little greater.

	Loss Median	Loss Tail	Expense Median	Expense Tail
Frees & Valdez	12,000	1.12	5500	2.12
Klugman & Parsa	12,275	1.05	5875	1.58

---

[2] The inverse Burr with a = c they call the inverse paralogistic, which is actually a name I coined some years ago. For the loglogistic, $F(x) = 1-(1+(x/b)^a)^{-1}$, whereas the Pareto has $F(x) = 1-(1+(x/b)^1)^{-a}$, so the combined form $F(x) = 1-(1+(x/b)^a)^{-a}$ could be called the paralogistic. The inverse of a distribution in this context is the distribution of $1/X$ from that distribution, which generates the inverse Burr, inverse paralogistic, etc.

Neither paper looked at the heavy right tail copula. For τ of 0.31, this is not too different from the Gumbel. In fact it is similar to the Gumbel in the right tail and more like the Frank in the left tail. This suggests that the joint Burr discussed above, which is built from the HRT copula, may provide a reasonable approximation to the loss and expense distribution, particularly in the right tail. This could be useful for excess-of-loss reinsurance estimates, especially when data is scarce. Recall that the joint Burr distribution is given by:

$$F(x,y) = 1 - (1+(x/b)^p)^{-a} - (1+(y/d)^q)^{-a} + [1+(x/b)^p + (y/d)^q]^{-a}$$

The a parameter comes from the HRT copula, with $\tau = 1/(1+2a)$. For $\tau = 0.31$, the implied a is 1.11. The tail heaviness factors are ap and aq, so p and q can be estimated from these parameters for this value of a. The tail heaviness can be estimated from available data or industry values could be used. A simple choice given the table above would be to take the loss factor as 1.11, which would give p = 1. A reasonable choice for q might be 1.5. Finally, b and d can be estimated from the respective medians. E.g., for b and p, the median is $b(2^{1/a} - 1)^{1/p}$. For a = 1.11, then, b = (median)$1.15^{1/p}$. The medians from Klugman and Parsa with p = 1 and q = 1.5 give (rounded):
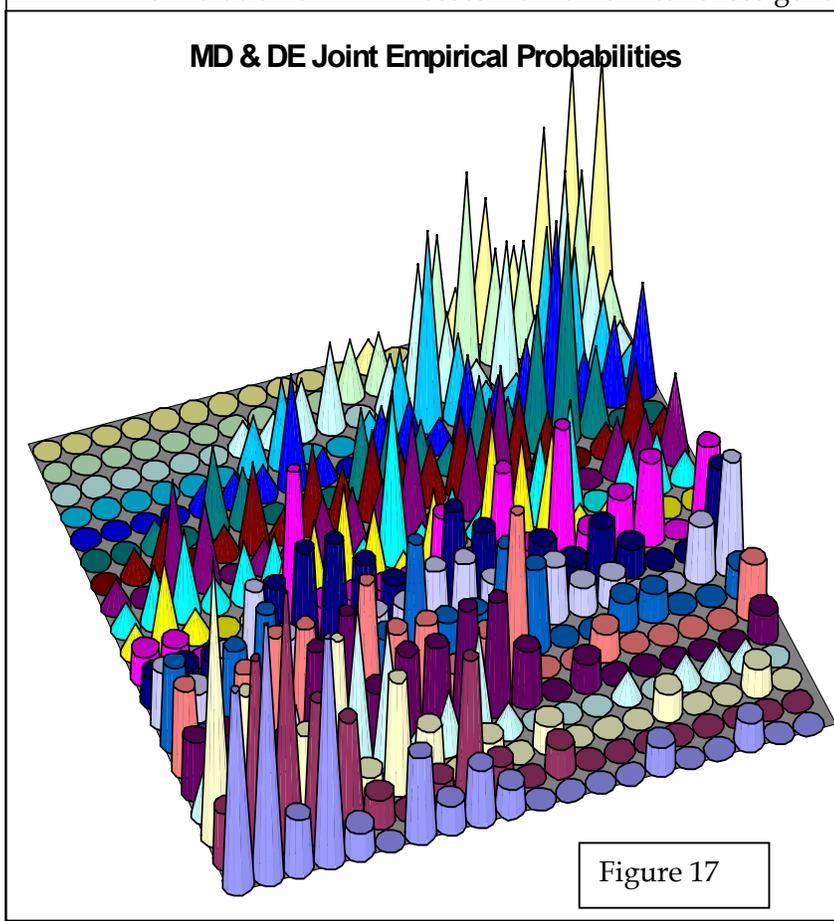
$$F(x,y) = 1 - [1+x/14150]^{-1.11} - [1+(y/6450)^{1.5}]^{-1.11} + [1+x/14150 + (y/6450)^{1.5}]^{-1.11}$$

Given a loss of x, the conditional distribution of loss expense is also Burr:

$$F_{Y|X}(y|x) = 1 - [1+(y/d_x)^{1.5}]^{-2.11} \text{ with } d_x = 6450 + 11x^{2/3}$$

Simulated Hurricane Losses

A simulation of n=727 losses from a hurricane loss generator for a sample data set of

MD & DE Joint Empirical Probabilities

Figure 17

Maryland and Delaware exposures will be used as an example of some of the issues that arise in copula estimation. As the emphasis is on the copula, not the marginal severities, the simulated losses were converted to probabilities by dividing the loss ranks for each state by n+1=728. The probability pairs were grouped into 20 intervals of 5% probability in each state for the graph here. The graph in Figure 17 shows there is a positive relationship between the loss probabilities for the two states, with some degree of concentration near (0,0) and (1,1). This is given in table form in Appendix C. A scatter plot of the empirical probabilities is shown in Figure 18.
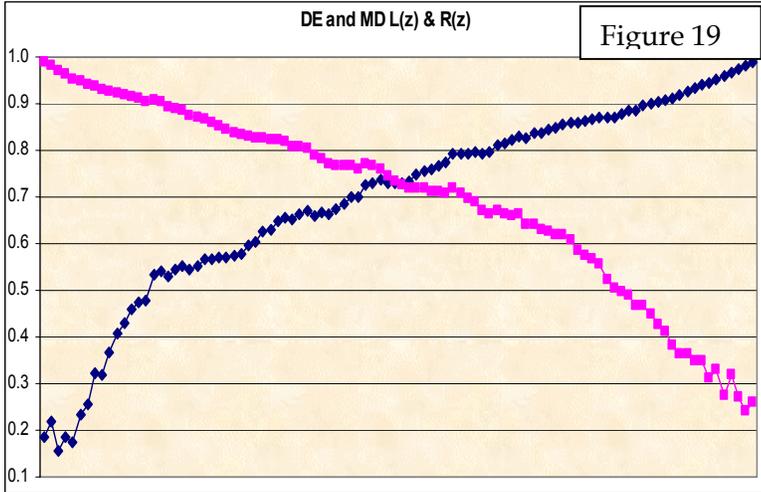

DE vs. MD copula

Figure 18
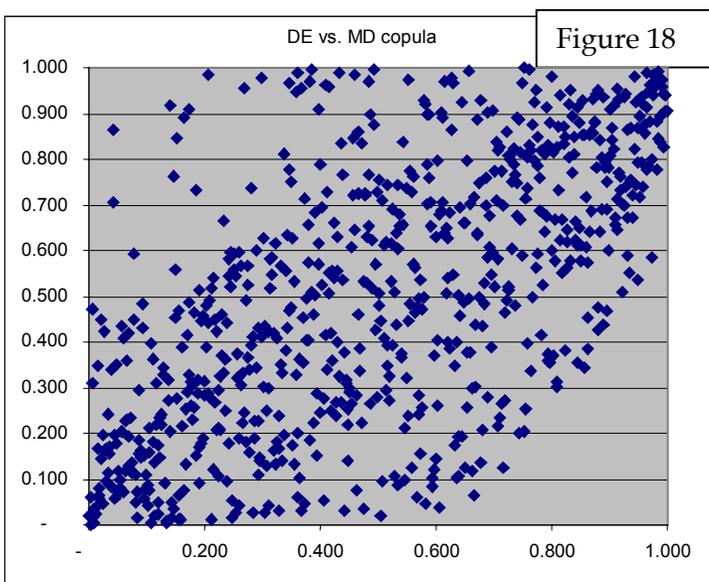
The usual estimate for the Kendall tau is to compute the average value over all pairs of observations (u_i, v_i), (u_j, v_j), i<j of sign$[(u_i - u_j)(v_i - v_j)]$. In this case the estimate is $\tau = .4545$.

An empirical copula can also be built at each point by counting the other points that are less in both states. As there are

n–1 other pairs, the count divided by n–1 can be taken as an estimate of the copula at that point. For this data, the maximum empirical copula value is 0.9821 and the average is 0.36363. Four times this less 1 is another estimate of tau, and this also is 0.4545.

Empirical L and R functions can be computed similarly. An estimate for L(z) can be obtained as C(z,z)/z where C(z,z) is computed as the proportion of pairs with u and v both less than z. Then with this C, R is estimated by by R(z) = [1 – 2z +C(z,z)]/(1 – z). These functions are graphed in Figure 19. The limiting values L and R are problematic immediately, in that they appear to be positive, yet much less than tau. All the copulas reviewed above have either R=0 or R>τ. The tails are fairly symmetrical, which poses additional fitting difficulties for single-tailed copulas like the HRT, PP Max, and Clayton. The Frank and normal copulas are thus likely to fit best, even though they are too light in the extreme tails.

DE and MD L(z) & R(z)   Figure 19

An empirical cumulative tau can also be calculated. For each z, the empirical C(u,v) can be computed for each (u,v) pair with both u and v less than z. Then the average of these values estimates the average copula in the square from (0,0) to (z,z). This divided by C(z,z), times four less one, is the estimate of J(z).


DE and MD J(z)   Figure 20
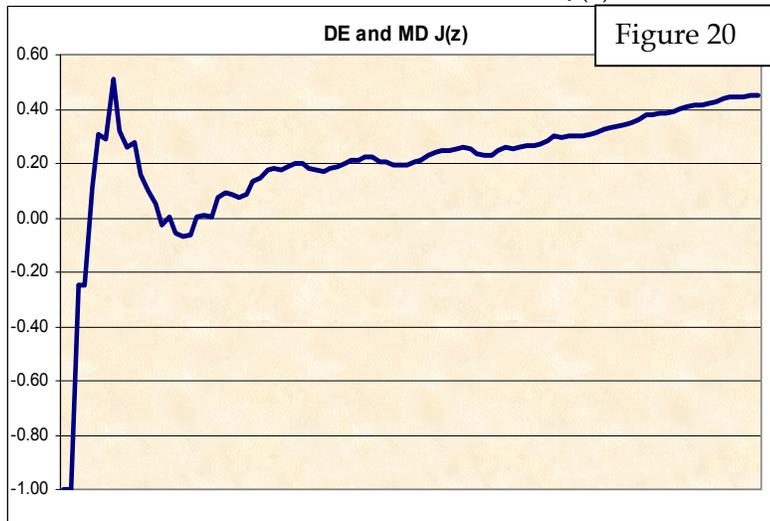
Its graph in Figure 20 is not like the J(z) for any of the copulas for small values of z, but the empirical calculation is based on few points when z is small. For larger z it is most similar to the almost linear J of the Frank copula.
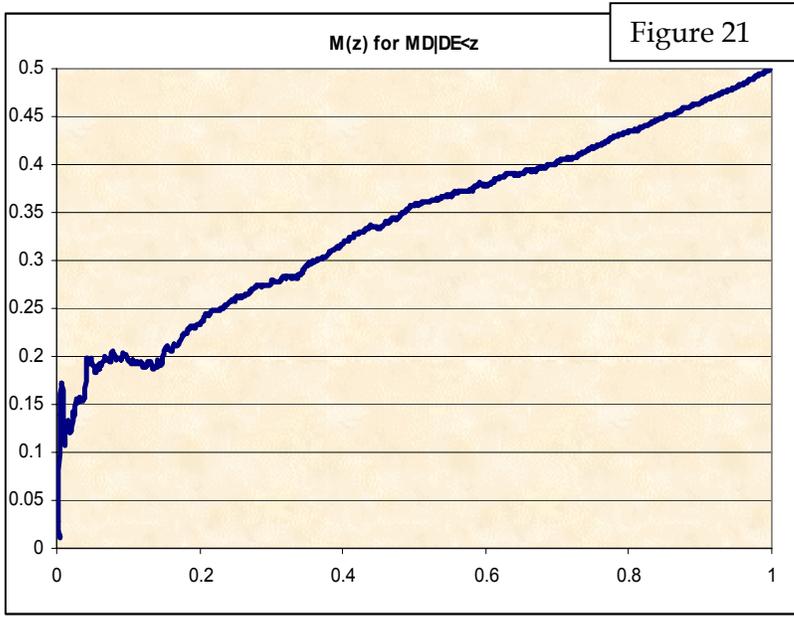
Figure 21

The M(z) function discussed in Appendix B can be calculated either for DE|MD or MD|DE. Figure 21 graphs MD|DE. It is most like the M function for the normal copula.

The descriptive functions thus suggest that the normal and Frank copulas should provide the best fits to this data, but they will be light in the tails.

Figure 22
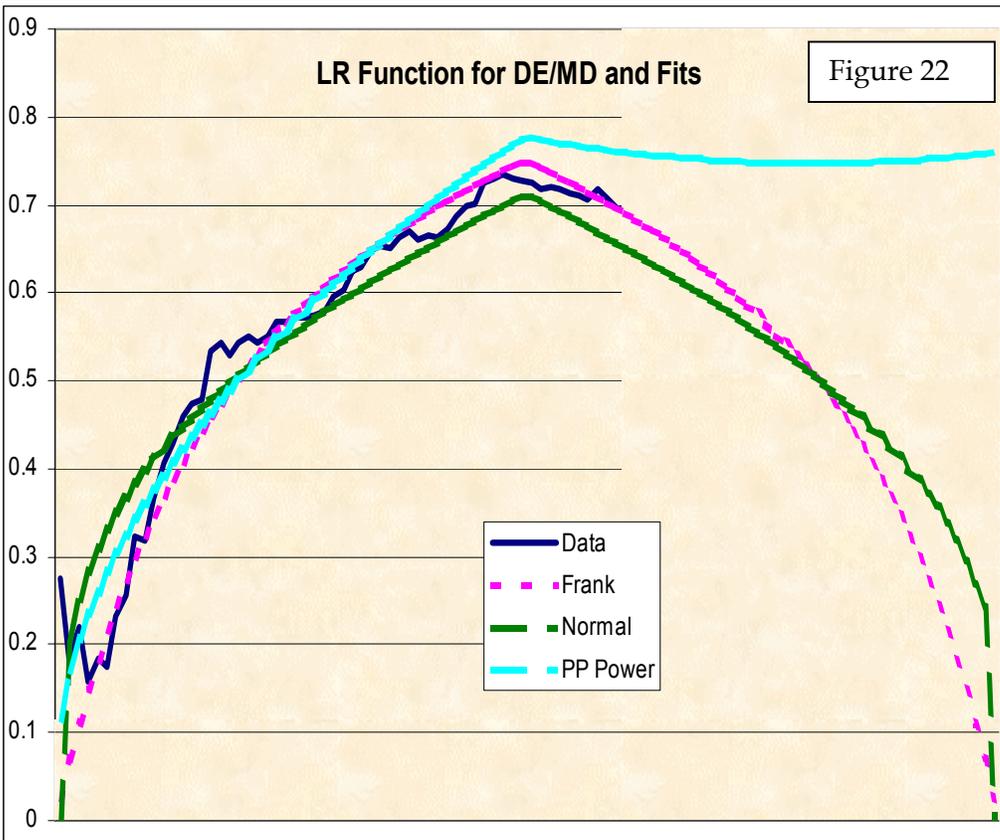
Maximum likelihood estimation of the parameter was performed for several of these copulas. The parameter and the maximal likelihood are shown below. As all the copulas here have a single pa-rameter, the ordering of the likelihood function is the same as those from the various information criteria like AIC, etc.

	HRT	Gumbel	Frank	Normal	Flipped Gumbel
Parameter	0.968	1.67	4.92	0.624	1.68
Ln Likelihood	124	157	183	176	161
Tau	0.34	0.40	0.45	0.43	0.40

The partial perfect copulas are difficult to estimate by MLE, as it is rare to have observations with exactly equal marginal probabilities. Nonetheless these copulas may be reasonable as scenario generators. An alternative is to estimate the parameter by matching tau. For the PP Power copula this gives a = 0.314. However for this data some of the descriptive functions seem to make this copula unlikely.

The likelihood function favors the Frank copula in this case. Some of the descriptive functions are graphed for the fit and the data for this copula and in some cases some other copulas in Figures 22 and 23. The L and R functions are combined in the graph above. R(z) is shown for z>0.5, and L(z) for z < 0.5. The Frank copula looks like a close fit all along except in the tails, where the normal is a little better. The PP Power appears to be too heavy in the right tail for this data.

Figure 23
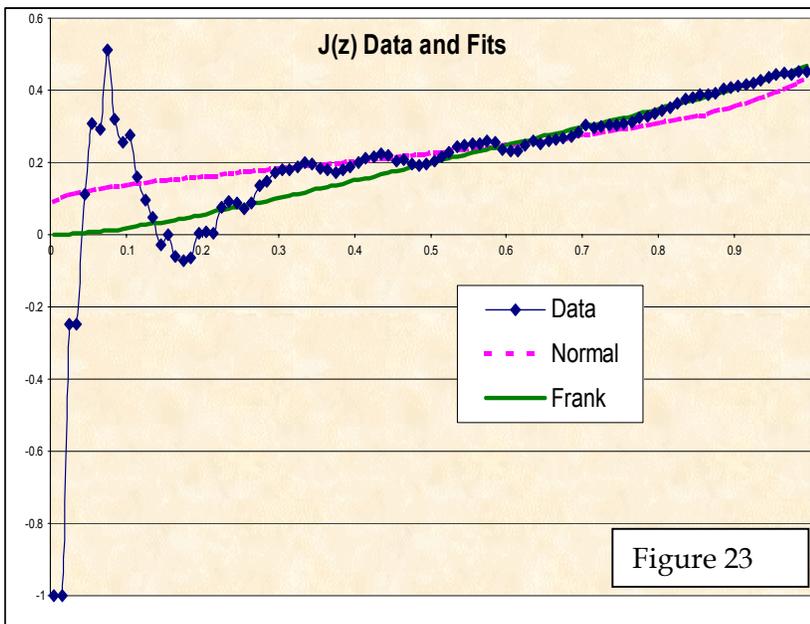
The graph at left shows the J(z) function for the data and the normal and Frank copulas. The two copulas provide quite different fits to this data, but it is a subjective matter as to which is better, with the Frank probably having the edge for its close fit for z>0.5. The Frank copula has a lower sum of squared errors, but this disappears if the first two points (at –1) are omitted.

Even though the Frank copula provides the best fit according to the likelihood function, there are fitting problems in the tails. Somewhat heavier-tailed copulas with strength in both tails would be useful here. See Appendix D for an example. Another alternative would be to use the Frank copula but model the extreme events separately.

6. Conclusion

Copulas provide a convenient way to model and simulate correlated variates. Several copulas with varying shapes are available for modeling these relationships. Shape differences among copulas can be discerned with the descriptive functions. These can be used both in fitting copulas to data and in applying informed judgment to select a copula for a given application.

Statisticians have identified a fair number of copulas – e.g., see [7]. The use of the descriptive functions provides an avenue for researching their properties. There may also be more descriptive functions that can reveal other aspects of a copula. For instance, the J and M functions looked at average probabilities between 0 and z. Mirror functions could look at the same probabilities between z and 1, analogous to the way that R mirrors L. It would also be possible to define more functions over non-rectangular parts of the unit square, such as the region where C(u,v) is less than z, as in the K function, or sections like u and v both less than z.

This paper focused on bivariate copulas but many of the concepts can be generalized to the multi-variate case. The descriptive functions have multi-variate analogs except for M(z) which would have to be done pairwise. Only the normal and partial perfect copulas fully generalize to multi-variate forms that allow specification of all pairwise correlations, but there are other multivariate copulas – e.g., see [4].

In summary, actuaries now have a number of copulas to chose among and a number of techniques for refining that choice, yet more copulas and more techniques could still be worth uncovering.

Appendix A – J(z)

For a copula with distribution function C(u,v) define:

$I(z) = \int_0^z \int_0^z C(u,v)c(u,v)\,dv\,du$. Then J can be expressed as:

$J(z) = 4I(z)/C(z,z)^2 - 1$.

For the following distributions the formula for $4I(z)$ is given.

Gumbel:

$(2-1/a)\exp[2^{1+1/a}\ln(z)] - 4(-\ln(z))^a(1-1/a)\int_y^\infty e^{-2w}\,w^{-a}\,dw$, where $y = -2^{1/a}\ln(z)$

Heavy Right Tail:

$8z - 8 + 4(2y-1)^{-a} + [4a(1-z)^2 + 2(1+(2y-1)^{-2a})(a+1)]/[2a+1] + 8a\int_1^y (w+y-1)^{-a-1}w^{-a}\,dw$,

where $y = (1-z)^{-1/a}$.

Partial Perfect Max

$z^4 + (z>a)(a^4 - 4a^3 + 2(1 + 2z)a^2 - 4az + 2z^2 - z^4)$

Partial Perfect Power

$z^4 + 4(a+1)^{-2}[(y^4 - 2y^3/3 + y^2/2)(a+1)^{-2} + z^{a+3}(a^2+3a+4)(a+2)^{-1}(a+3)^{-1} - z^{2(a+2)}(a^2+2a+2)(a+2)^{-2}]$,

where $y = z^{a+1}$.

Clayton

$y^{-b}(b+1-b/y)(b+2)/(b+1)$, where $b = 2a$ and $y = 2/z^{1/a} - 1$.

BB1

$2(1+(ac)^{-1})y^2 - 4(a+1)y^{2+a}/[ac(a+2)] - 4x(ac)^{-1}\int_0^y (w^{-a}-1)^{-c}\,w^{1+a}\,[(ac+1)w^{-a}-a-1]\,dw$, where $x = 2(z^{-a}-1)^c$, and $y = (1+x^{1/c})^{-1/a}$.
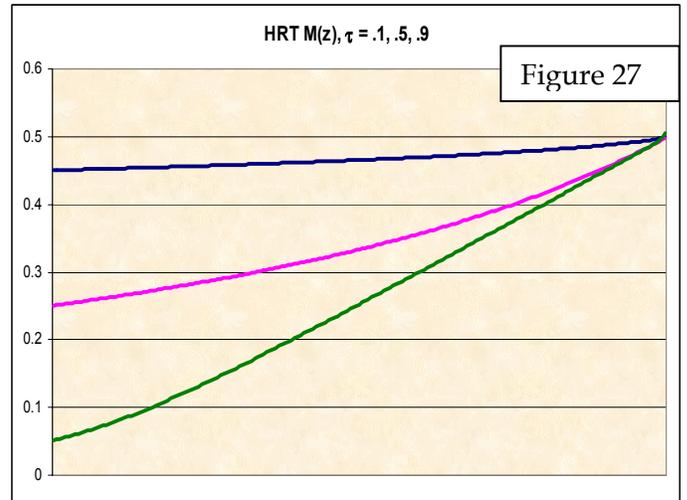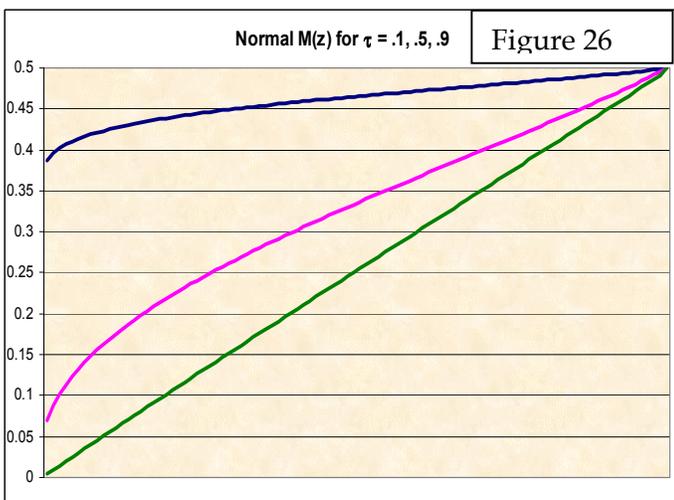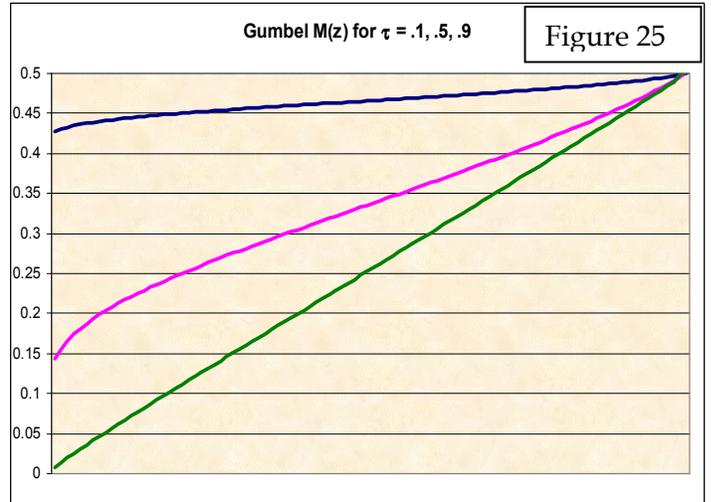
Appendix B – Other Descriptive Functions

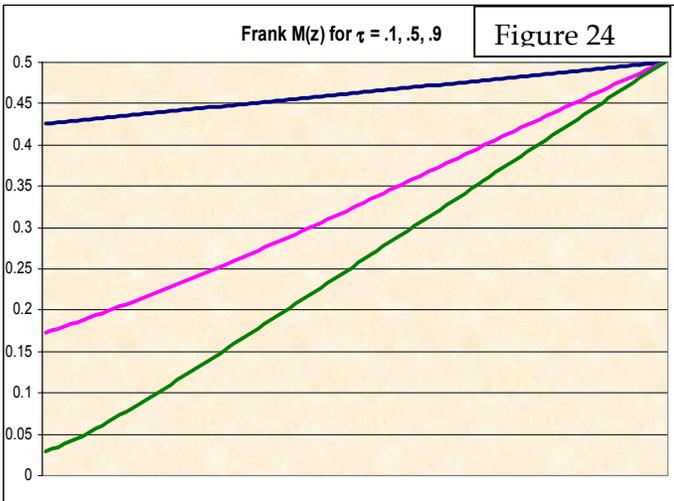## *Cumulative Conditional Mean*

A function of interest is the conditional expected value of V|U=z. However this is often difficult to estimate from data, as there are usually not too many values of V for any given value of U. So a related function is chosen: the expected value of V given U<z. Let

$$M(z) = E(V\,|\,U<z) = \int_0^z\!\int_0^1 v c(u,v)\,dv\,du\,/\,z$$

Since E(V) = ½, every copula will have M(1) = ½ so the differences in M among copulas will be for lower values of z and the shape of the curve approaching z = 1.

Often the integral has to be done numerically, but for a few copulas it is done explicitly at the end of this appendix. Graphs of this function for several copulas are shown in Figures 24 – 29. For this function, the lower τ is, the closer the values stay to ½.

Frank M(z) for τ = .1, .5, .9   Figure 24

Gumbel M(z) for τ = .1, .5, .9   Figure 25

Normal M(z) for τ = .1, .5, .9   Figure 26

HRT M(z), τ = .1, .5, .9   Figure 27

Figure 28 — PP Power M(z), τ = .1, .5, .9

Figure 29 — PP Max M(z), τ = .1, .5, .9
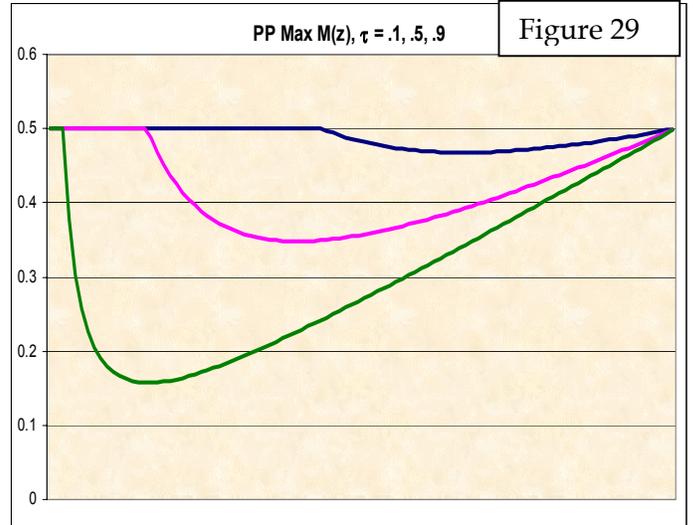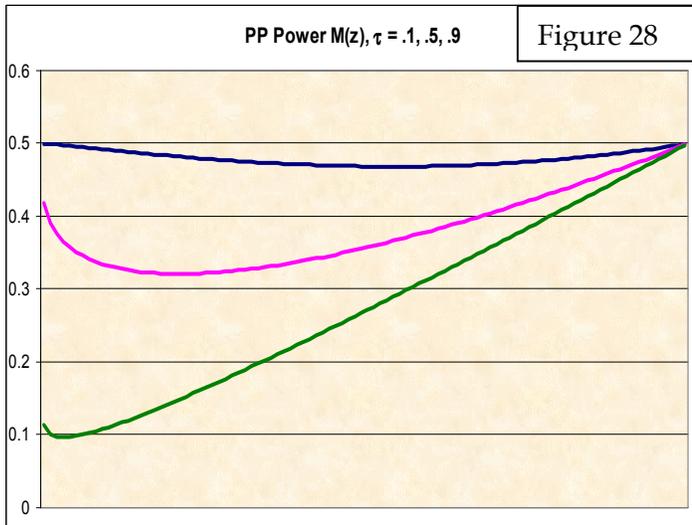
Copula Distribution Function

Genest and Rivest [3] define a function K(z) which is basically $Pr(C(u,v)<z)$. It is the area of the unit square in which $Pr(C(u,v)<z)$. An empirical K(z) can be calculated for any z as the proportion of empirical values of C(u,v) that are less than z. Although C(u,v) approaches one as u and v approach one, it is possible that C is low for most values of u and v, which would make K(z) high for most z's. Or C could grow fairly quickly through lower values of u and v, which would tend to make K(z) smaller.

Genest and Rivest show how to calculate K for a number of copulas. In particular,

Copula	K(z)
Gumbel	$z(1 - \ln z^{1/a})$
Frank	$z + a^{-1}(1-e^{az})\ln[(1-e^{-az})/(1-e^{-a})]$

Hurricane Application

M(z) for the hurricane data and the Frank and normal copulas is graphed in Figure 30. The normal copula is the one with the better fit for small events, and the Frank fits better in the middle of the range.

A scatter plot of the empirical K percentiles as a function of the Frank K percentiles (often called the QQ plot) is shown in Figure 31, along with the line x=y. The values are very

close to the line. This supports the fit, but as K(0) = 0 and K(1) = 1 for any copula, empirical or parametric, fit problems in the tails are difficult to discern with this function.

Figure 30

Figure 31
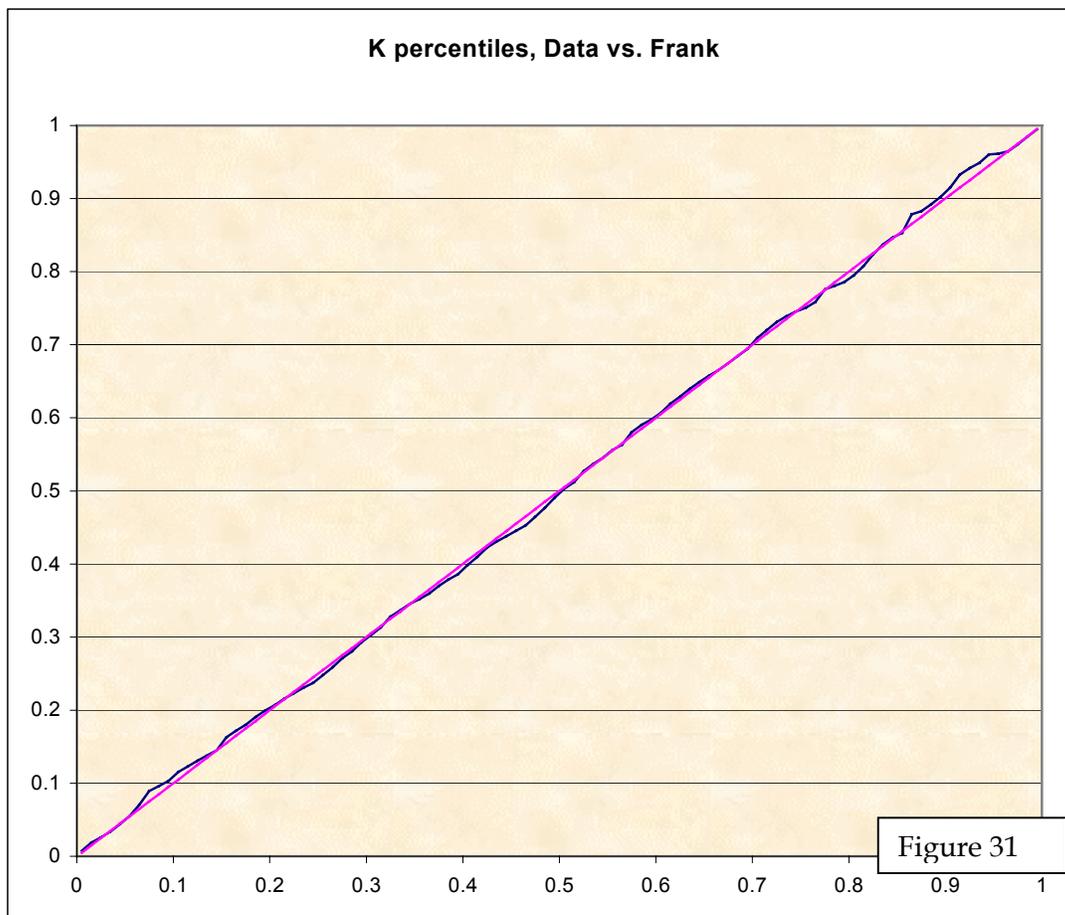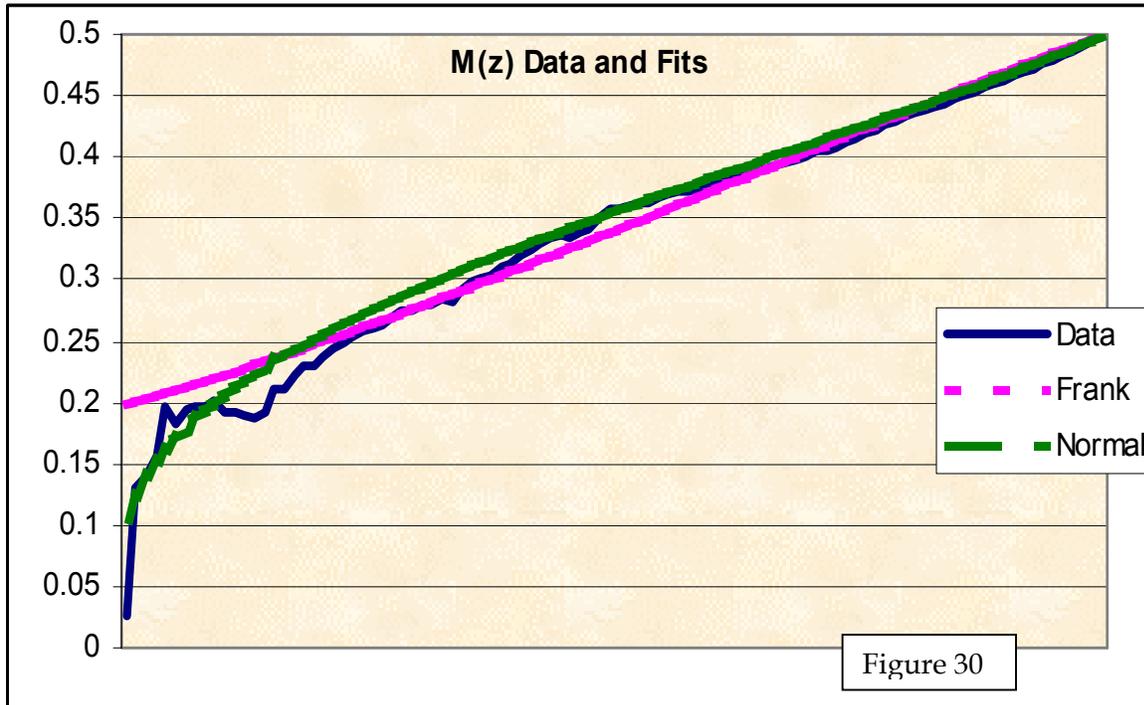
M(z) Formulas

<u>Partial Perfect Maximum</u>

$M(z) = \frac{1}{2} - \frac{1}{2} (z > a)(1 - a)(1 - z)(z - a)/z$

<u>Partial Perfect Power</u>

$M(z) = \frac{1}{2} + (z^{a+1} - z^{a})/[(a+1)(a+2)]$

Appendix C – Delaware and Maryland Probabilities by Range

Range Upper Limits - Maryland

	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
0.05	7	9	2	8	1	0	3	1	2	1	0	0	0	0	1	0	0	1	0	0
0.1	2	7	9	5	4	1	0	1	5	1	0	1	0	0	0	0	0	0	0	0
0.15	10	5	2	3	6	1	4	2	0	1	0	1	0	0	0	1	0	0	1	0
0.2	2	4	2	3	2	6	5	1	2	4	1	0	0	0	1	0	1	1	1	0
0.25	2	2	2	1	4	5	2	3	3	4	4	2	0	1	0	0	0	0	0	1
0.3	3	0	3	4	4	0	5	3	3	1	2	6	0	0	1	0	0	0	0	2
0.35	3	0	3	4	1	2	2	2	5	2	2	4	3	0	0	1	1	0	0	1
0.4	2	1	2	4	2	1	1	2	2	3	3	2	2	2	1	1	0	0	2	4
0.45	1	1	1	0	3	4	4	3	2	0	4	3	1	2	1	2	1	0	0	3
0.5	1	1	0	0	1	6	1	1	1	1	3	1	4	1	5	1	2	3	0	3
0.55	1	3	1	0	1	2	3	4	3	1	3	0	5	4	4	1	1	0	0	0
0.6	1	2	3	1	2	2	0	1	1	6	1	2	1	1	3	4	0	2	2	1
0.65	1	0	4	3	0	1	1	2	2	2	4	0	3	5	1	1	0	2	2	3
0.7	0	1	2	0	1	3	1	3	2	3	2	2	3	3	3	2	0	1	3	1
0.75	0	0	1	0	3	2	0	0	0	3	3	2	1	3	2	5	8	3	0	0
0.8	0	0	0	0	1	1	1	4	1	0	2	2	1	3	2	3	7	2	3	4
0.85	0	0	0	0	0	0	2	3	0	0	1	3	7	3	1	1	6	5	3	1
0.9	0	0	0	0	0	0	1	1	2	3	0	2	5	3	1	5	2	7	4	1
0.95	0	0	0	0	0	0	0	0	0	0	2	2	1	5	7	4	4	3	6	2
1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	5	3	7	9	9

Appendix D – Joe's BB1 Copula

Several examples of two-parameter bivariate copulas are provided by Joe [4]. One that has a closed form for tau and can be heavy in both tails he labels BB1. It is a generalization of the Gumbel and Clayton copulas.

$$C(u,v) = \{1+[(u^{-a} - 1)^c + (v^{-a} - 1)^c]^{1/c}\}^{-1/a} \ , \ a > 0, c \geq 1$$

$$C_1(u,v) = \{1+[(u^{-a} - 1)^c + (v^{-a} - 1)^c]^{1/c}\}^{-1/a - 1}[(u^{-a} - 1)^c + (v^{-a} - 1)^c]^{1/c - 1}(u^{-a} - 1)^{c - 1}u^{-a - 1}$$

$$c(u,v) = \{1+[(u^{-a} - 1)^c + (v^{-a} - 1)^c]^{1/c}\}^{-1/a - 2}[(u^{-a} - 1)^c + (v^{-a} - 1)^c]^{2/c - 2} \times$$

$$\{ac + 1 + a(c- 1) \, [(u^{-a} - 1)^c + (v^{-a} - 1)^c]^{- 1/c}\} \, (u^{-a} - 1)^{c- 1}u^{-a- 1} (v^{-a} - 1)^{c- 1}v^{-a- 1}$$

$$\tau = 1 - 2/[c(a+2)]$$

$$R(1) = 2 - 2^{1/c} \ , \ L(0) = 2^{- 1/(ac)}$$

The Gumbel is the limiting case $a \to 0$. The Clayton arises when $c = 1$, but here the a parameter is the reciprocal of the Clayton a parameter in the text.

With R, L, and τ all closed form it is possible to find a and c to set two of them and then see what the third is. Not all combinations are possible. Figure 32 graphs L as a function of R for several values of τ. For each τ, there is an inverse relationship between R and L. Either can get as low as needed, approaching zero, for any value of τ, but then the other becomes large. Each becomes somewhat higher than τ if the other one is low. Higher τ allows higher R and L. The left tails appear to be somewhat heavier than the right tails, so flipping this copula could be useful for some applications.

This copula was not all that useful for the hurricane data, in that it is so heavy-tailed. The MLE log-likelihood was 170, which was not as good as some other copulas. The a and c were 0.386 and 1.434, which gave $\tau = 0.415$, $L = 0.286$, and $R = 0.379$. So the τ was a little low and the tail parameters higher than the data would suggest.

BB1 L as a Function of R, tau = .1, .2, …, .9

Figure 32

tau = .7

tau = .5

tau = .3
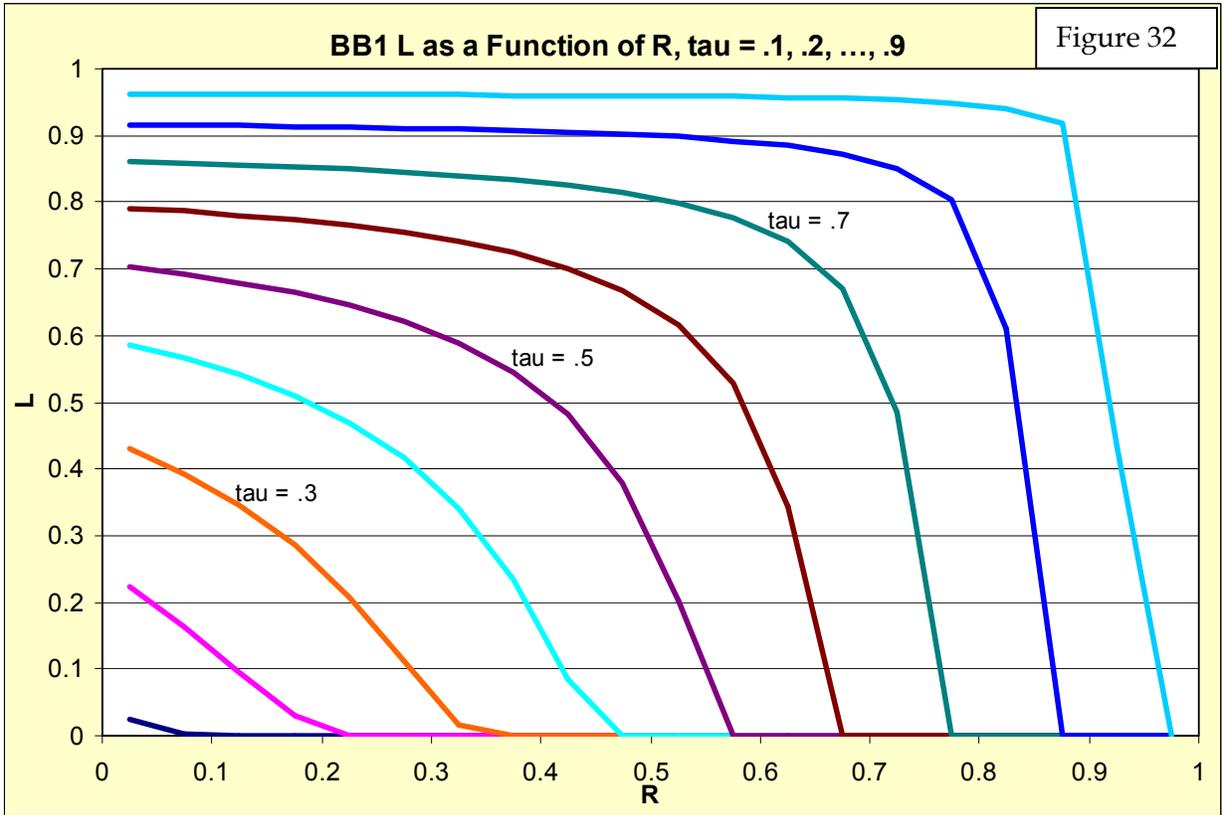
References

1. Embrechts, Lindskog, and McNeil *Modelling Dependence with Copulas and Applications to Risk Management* at http://www.risklab.ch/Papers.html#MTLindskog

2. Frees, EW and Valdez, EA, "Understanding Relationships Using Copulas," *North American Actuarial Journal*, vol.2, #1, pp. 1-25

3. Genest, C and Rivest, L 1993 "Statistical Inference Procedures for Bivariate Archimedean Copulas," *Journal of the American Statistical Association* 88: 1034–1043

4. Joe, H, *Multivariate Models and Dependence Concepts,* Chapman and Hall, 1997

5. Klugman, SA and Parsa, R, "Fitting Bivariate Loss Distributions with Copulas," *Insurance: Mathematics and Economics* 24, pp. 139-148, 1999

6. Kreps, R "A Partially Comonotonic Algorithm For Loss Generation," *ASTIN Colloquium Papers* 2000

7. Nelson, R, *An Introduction to Copulas,* Springer Lecture Notes in Statistics, 1999

8. Wang, S, "Aggregation of Correlated Risk Portfolios: Models and Algorithms," 1999 *PCAS LXXXV*