Tails of Copulas

Gary G. Venter

Guy Carpenter Instrat

Abstract: The various copulas in the actuarial and statistical literature differ not so much in the degree of association they allow, but rather in which part of the distributions the association is strongest. In property and casualty applications there is interest in copulas that emphasize correlation among large losses, i.e., in the right tails of the distributions. Several copulas that have this characteristic are discussed. In addition, univariate functions of copulas are introduced that describe various aspects of the copulas, including tail concentration. Univariate descriptive functions can be thought of as an intermediate step between the several zero-dimensional measures of association (Kendall, Spearman, Gini, etc.) and the multi-dimensional copula function itself.

Guy Carpenter Instrat Tails of Copulas

Gary G. Venter

Guy Carpenter Instrat

Copulas are just joint distributions of unit uniform variates. They become especially useful when the unit uniform variates are viewed as probabilities from some other distributions. Then the percentiles of the other distributions can be recovered by inverting these probabilities. Copulas thus provide a ready method for simulating correlated variables, e.g., in a DFA model. Quite a few copulas are available for this, and they have differing characteristics that lead to different relationships among the variables generated. This paper reviews several popular copulas, introduces some others, and discusses methods for selecting those that may be most appropriate for a given application. In particular, the behavior of the copulas in the right and left tails is used as a way to distinguish among copulas that produce the same overall correlation.

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

1. General Considerations

Copulas Defined

A joint distribution function $F(x,y)$ can be expressed as a function of $F_X(x)$ and $F_Y(y)$, the individual (or marginal) distribution functions for X and Y, i.e., as $F(x,y) = C(F_X(x),F_Y(y))$. There has to be a function C defined on the unit square that makes this work, because $F_X(x)$ and $F_Y(y)$ are order preserving maps of the real line or some segment of it to the unit interval. This can be defined by $C(u,v) = F(F_X^{-1}(u),F_Y^{-1}(v))$. Then $C(F_X(x),F_Y(y)) = F(F_X^{-1}(F_X(x)),F_Y^{-1}(F_Y(y))) = F(x,y)$. 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.

Copulas work in the multi-variate context as well, but this paper mainly focuses on joint distributions of two variates. A copula is a joint distribution of two 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., $c(u,v) = 1$, $C(u,v) = uv$.

Conditioning with Copulas

The conditional distribution can be defined using copulas. Let $C_1(u,v)$ denote the first partial derivative of $C(u,v)$. 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 that 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, then simulate a value of V from C_1, the conditional distribution of $V|U$.

Correlation

The usual 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 $C(u,v) = uv$, $E[C(u,v)] = 1/4$, so $\tau = 0$. Also, for perfectly correlated variates $U = V$, $E[C(u,v)] = 1/2$, 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.

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:

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

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

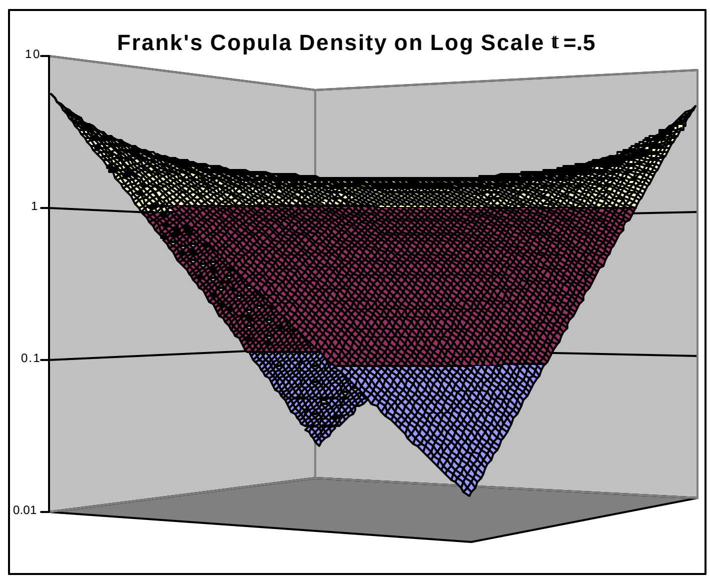
$$c(u,v) = -a g_1(1+g_{u+v}) / (g_u g_v + g_1)^2 \text{ and Kendall's } \tau \text{ of}$$

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

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 $V|u$. Since this has distribution function C_1, v can then be found as $v = C_1^{-1}(p|u)$. The formula for this, which can be found from the formula for C_1, is:

$$v = -a^{-1}\ln\{1+pg_1/[1+g_u(1-p)]\}$$



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 $y = F_Y^{-1}(v)$.

The copula density is graphed here. As with many copulas there is a degree of concentration near 0,0 and 1,1. This often weakens considerably in the inversion back to X and Y, which is usually a non-linear transformation. Some of the copulas below have even greater concentration in the tails.

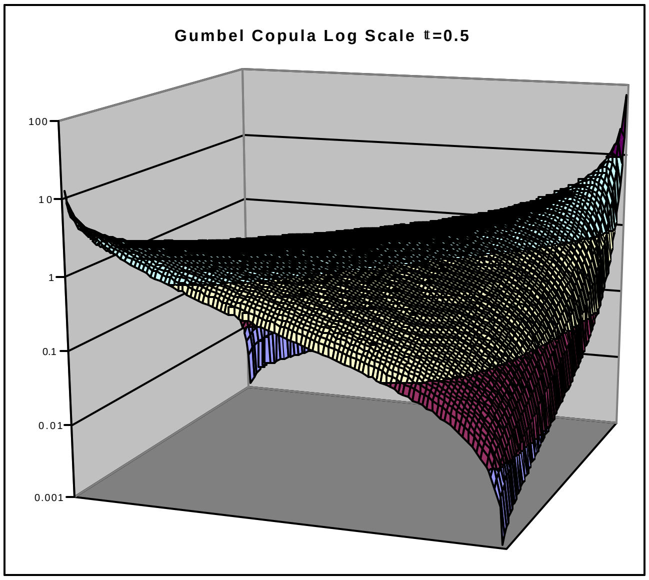
Gumbel Copula

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:

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

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

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



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

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 it. First simulate two independent uniform deviates u and v. Next solve numerically for s>0 with $ue^s = 1 + as$. Then the pair $[\exp(-sv^a), \exp(-s(1-v)^a)]$ will have the Gumbel copula distribution. As shown in this graph, the Gumbel copula has a high degree of tail concentration, with more emphasis in the right tail.

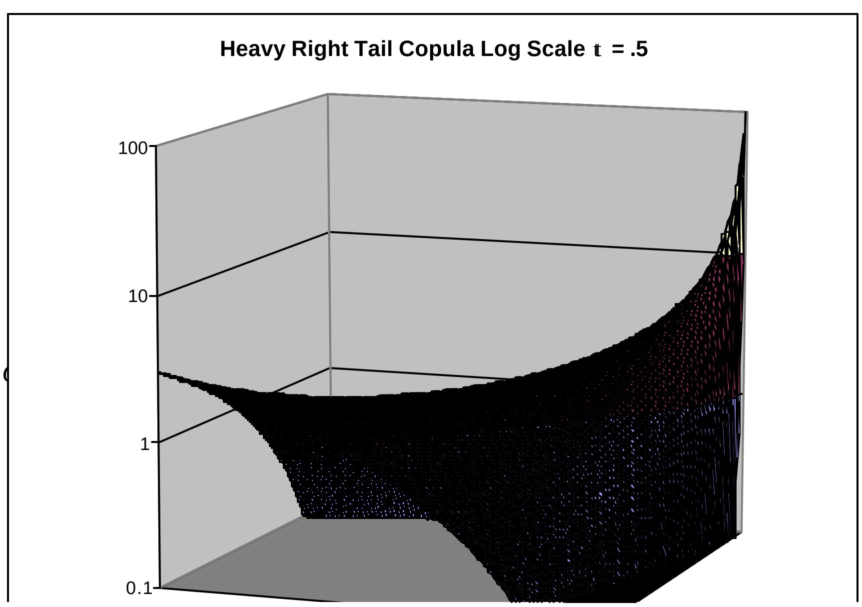
Heavy Right Tail Copula and Joint Burr

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:

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

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

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



[1] Embrechts, McNeil and Strauman

ASTIN Colloquium Papers

$v)]^{-1-1/a}$ $\tau(a) = 1/(2a+1)$

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 $y \mid X=x$ is also Burr:

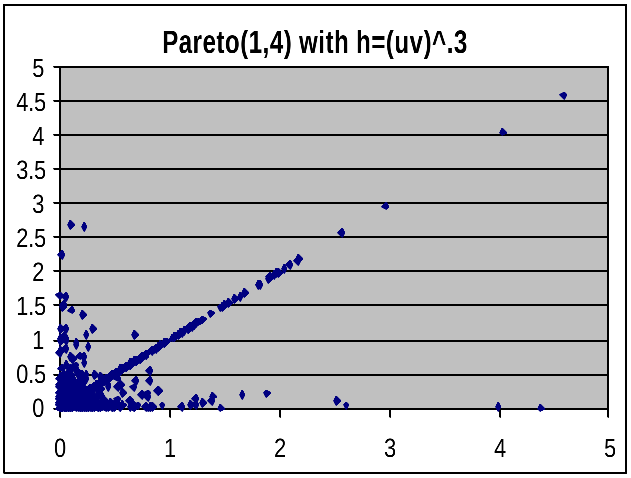
$$F_{Y|X}(y \mid x) = 1 - [1 + (y/d_x)^q]^{-(a+1)}, \text{ where } d_x = d[1 + (x/b)^{p/q}]$$

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.

Kreps' Partial Perfect Correlation Copula Generator

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

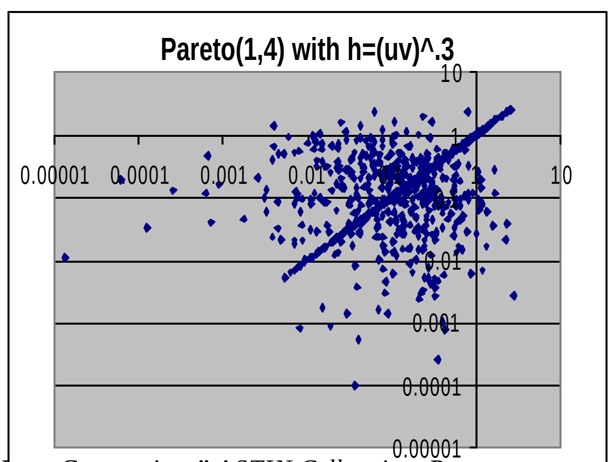
A method for generating copulas that are mixtures of perfectly correlated and independent variates has been developed by Rodney Kreps[3]. This is easier to describe as a simulation procedure, and then look at the copulas.



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 from 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= $F_Y^{-1}(u)$. 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 a controlling how much more.

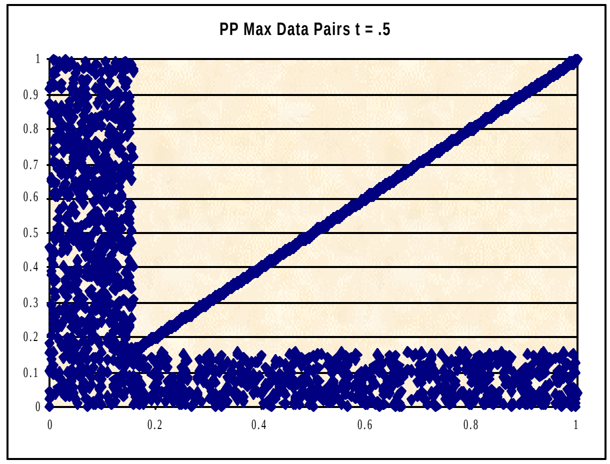
The graphs here illustrate the case where h(u,v) = $(uv)^{0.3}$ and both X and Y are distributed Pareto with $F(x) = 1 - (1 + x)^{-4}$. The correlated and uncorrelated instances clearly show up separately, in either the log or regular scale. For larger values of a, h(u,v) is smaller, so it is less likely that



[3] R. Kreps 2000 “A Partially Comonotonic Algorithm For Loss Generation,” *ASTIN Colloquium Papers*

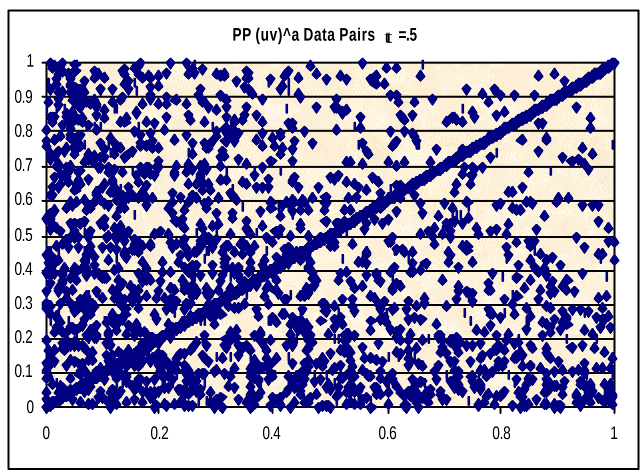
h(u,v) exceeds the random value w and thus less likely that the case u=v will be selected. For small values of a, on the other hand, h(u,v) will be larger, approaching one as a goes to zero. Thus h(u,v)>w is more likely, so u=v will also be more likely.

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 spread-sheets where a logical expression in parentheses will evaluate to zero if the expression is false and one if it is true. Thus (u=v) is one if u=v and zero otherwise, etc.

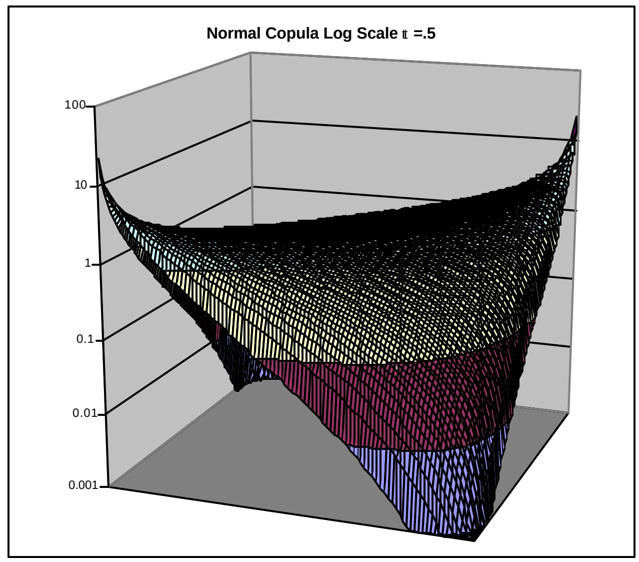
Although Kreps considers more general situations, a relatively simple copula results in the case where h(u,v) breaks out as a product of a univariate function evaluated at u and v, i.e., h(u,v) = h(u)h(v). If we define $H(x) = \int_0^x h(t)dt$, the copula formulas become:



$C(u,v) = uv - H(u)H(v) + H(1)H(\min(u,v))$
$C_1(u,v) = v - h(u)H(v) + H(1)h(u)(v>u)$
$c(u,v) = 1 - h(u)h(v) + H(1)h(u)(u=v)$

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) = (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 above for the case $\tau = ½$.

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.



The graph 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.

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's.

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 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
t	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.

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:

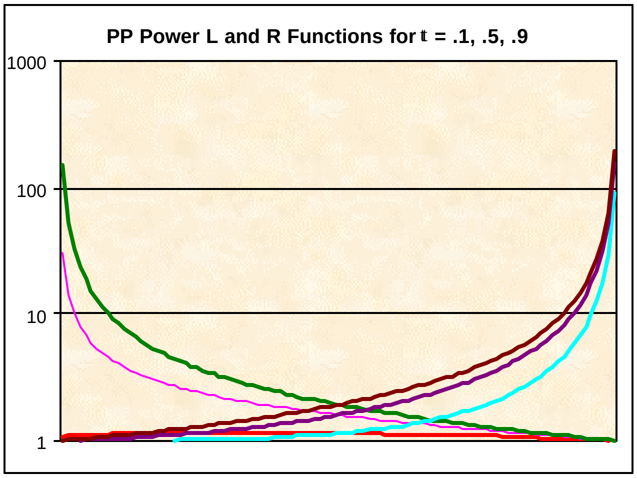
$L(z) = \Pr(U<z,V<z)/z^2$ and $R(z) = \Pr(U>z,V>z)/(1-z)^2$

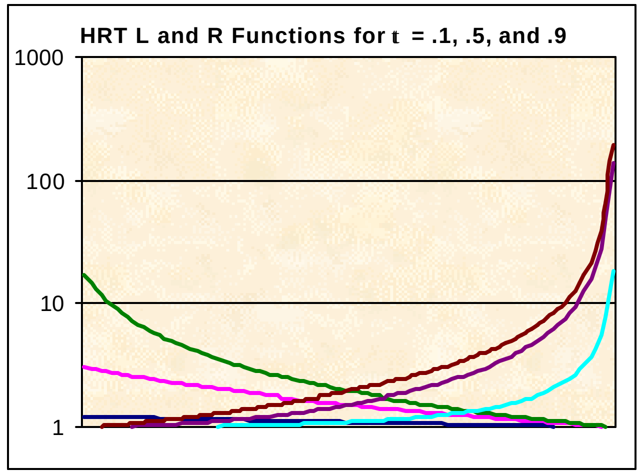
In terms of the copula functions, L(z) is just $C(z,z)/z^2$. To calculate R(z), note that

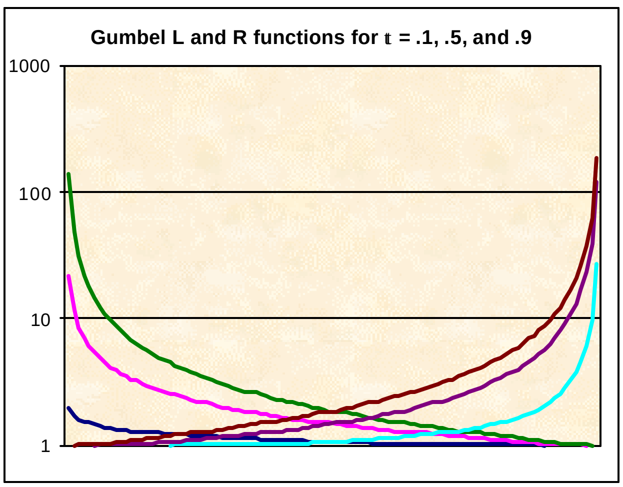
$1 - \Pr(U>z,V>z) = \Pr(U<z) + \Pr(V<z) - \Pr(U<z,V<z) = z + z - C(z,z)$.

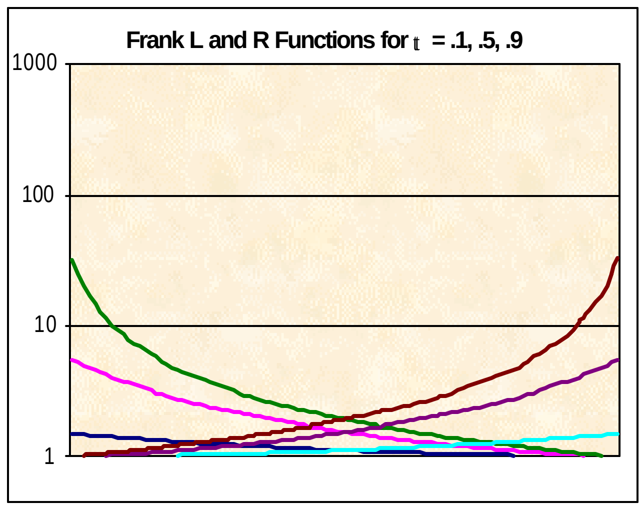
Then R(z) can be calculated by $R(z) = [1 - 2z + C(z,z)]/(1-z)^2$.

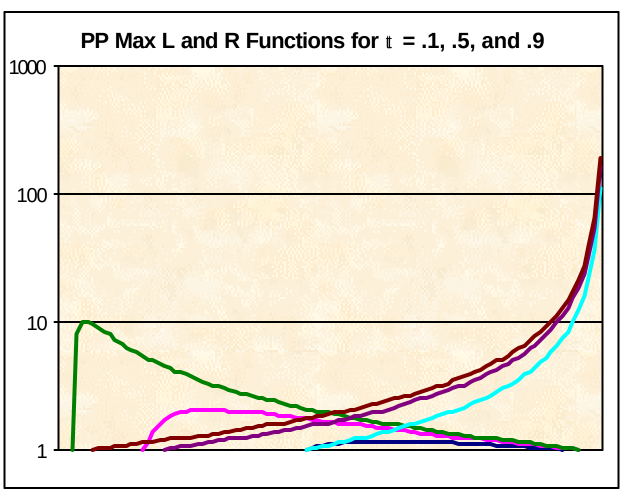
These functions are graphed for a few copulas below. As L(1) = R(0) = 1, the L and R functions are easily distinguished on the graphs. Also, the higher correlations have higher tail concentrations.

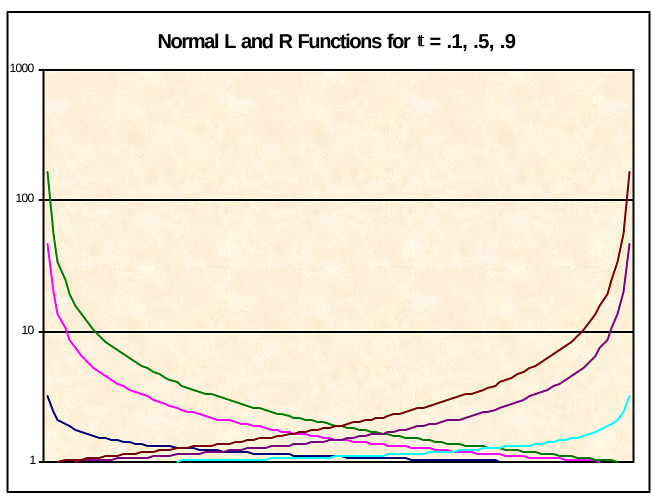












The Gumbel has fairly high concentrations in both tails, but the right tail is heavier. The relative strength of the left tail gets less as the correlation decreases. The heavy right tail copula has about the same right tail concentration as the Gumbel, but it has less in the left tail. The Frank copula is basically symmetrical between the left and right tails. Both tails are less concentrated than the Gumbel but the left tail concentration is about the same as in the heavy right tail copula. The PP Max L(z) function is somewhat different, as the variates are independent in the left tail. Thus for low z, L(z) = 1. The lower the correlation, the longer the L function stays at 1. Note also that the right tail function is strong even for low overall correlation. The PP Power copula has functions similar to the Gumbel, although for low correlation the L function is nearly flat, and the R function is high for even low τ. The tails of the normal copula show more concentration than did the Frank, but usually not as strong as the Gumbel, particularly in the right tail.

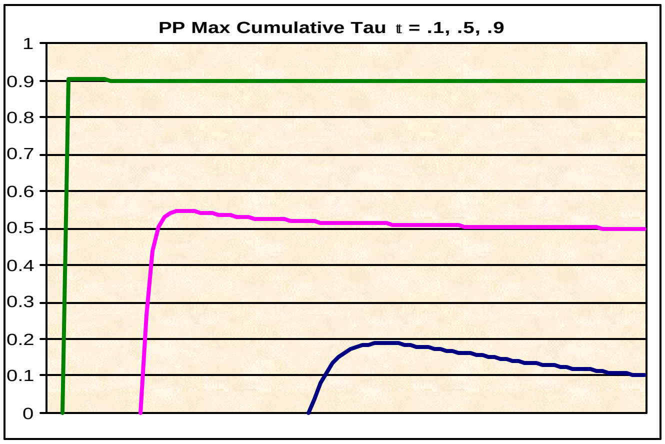
Cumulative Tau

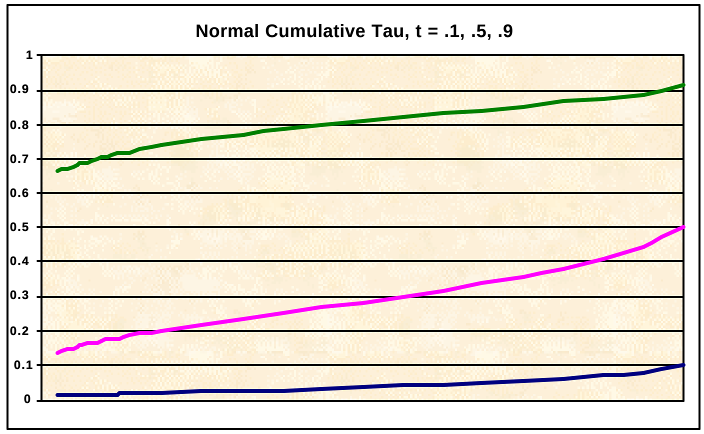
Recall that tau is defined as $-1+4\int_0^1\int_0^1 C(u,v)c(u,v)\,dv\,du$. A cumulative tau can be defined as $J(z) = -1+4\int_0^z\int_0^z C(u,v)c(u,v)\,dv\,du \,/\, C(z,z)^2$.

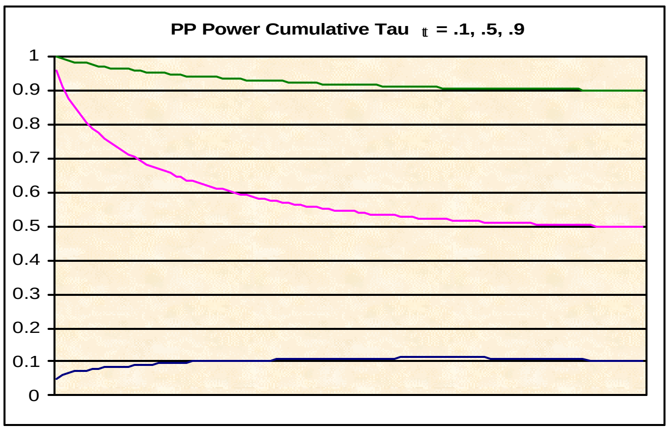
The full double integral is a probability weighted average of C(u,v), i.e., EC(u,v). To compare to this, the partial integral has to be divided by the weights, hence the first C(z,z) in the denominator. This quotient will give the average value of C(u,v) in the square from (0,0) to (z,z). This will increase as a function of z for any copula.

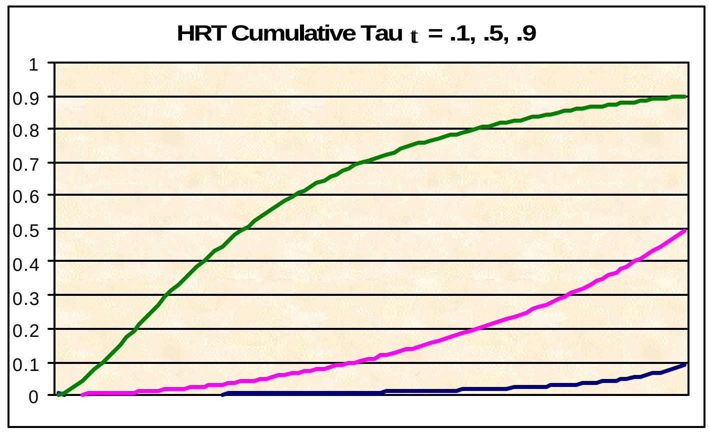
The second C(z,z) divisor expresses this average relative to C(z,z), i.e., shows how the average C compares to the maximal C in the square. This may or may not increase as a function of z, which makes it a more interesting property of the copula.

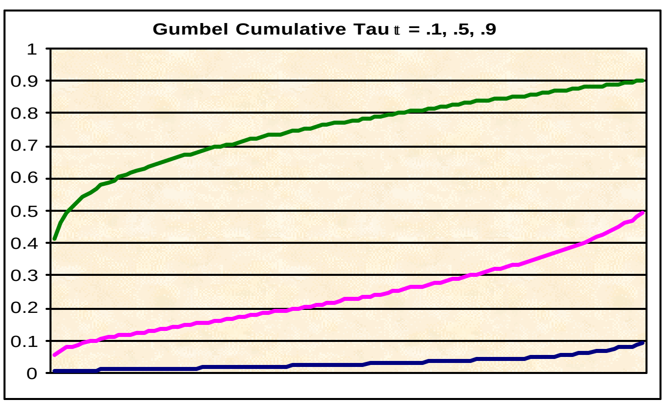
The normalization to the range of a correlation is a matter of convenience and familiarity, and gives J(1) = τ. The integration can be done numerically, although for some copulas, formulas are given in Appendix 1. Looking at the graphs shows that for each correlation, the shape of the J function varies noticeably from one copula to another. This provides a way of narrowing down the choice of copulas given data or other criteria. All the graphs end up at τ for z=1, but can start off with high or low correlation, and can increase or de-

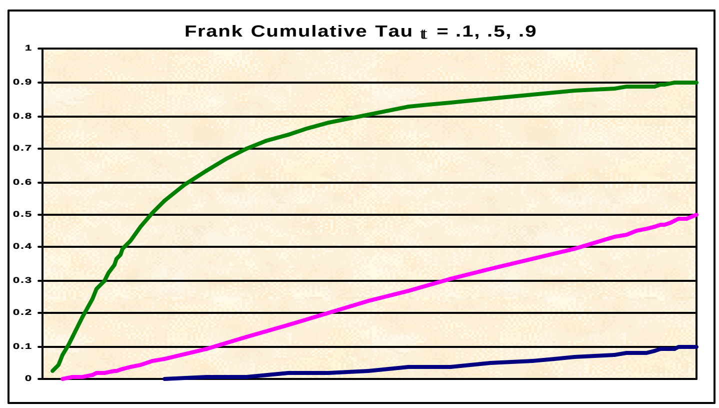












crease at varying rates.

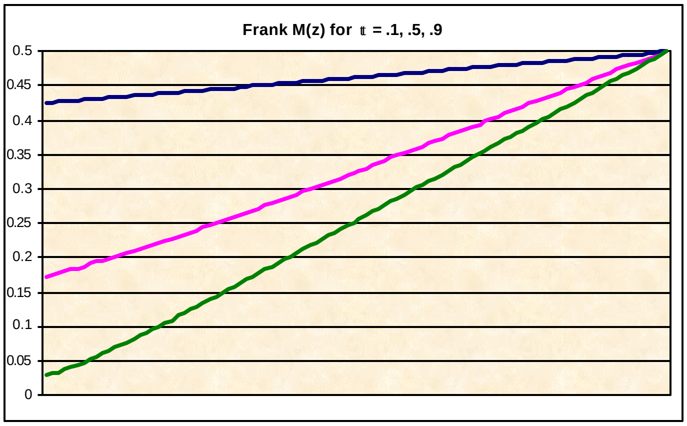
Cumulative Conditional Mean

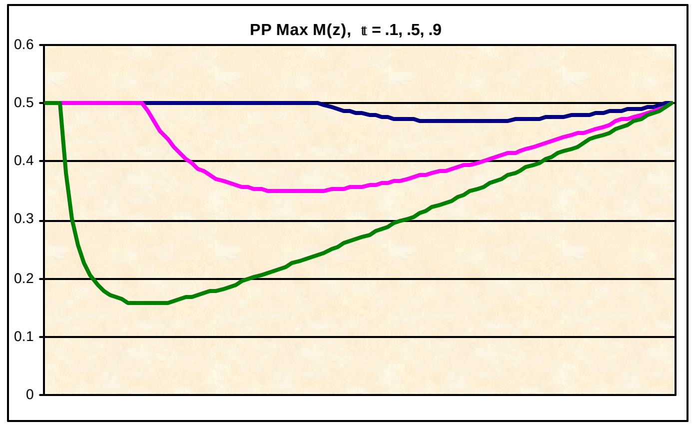
A function of interest is the conditional expected value of V|U=z. However this is often difficult to estimate from data, as there are usually not too many values of V for any given value of U. So a related function is chosen: the expected value of V given U<z. Let

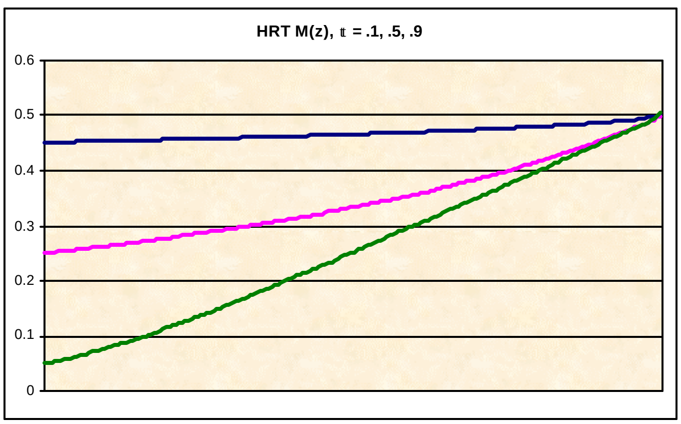
$$M(z) = E(V \mid U<z) = \int_0^z \int_0^1 v c(u,v)\, dv\, du \,/\, z$$

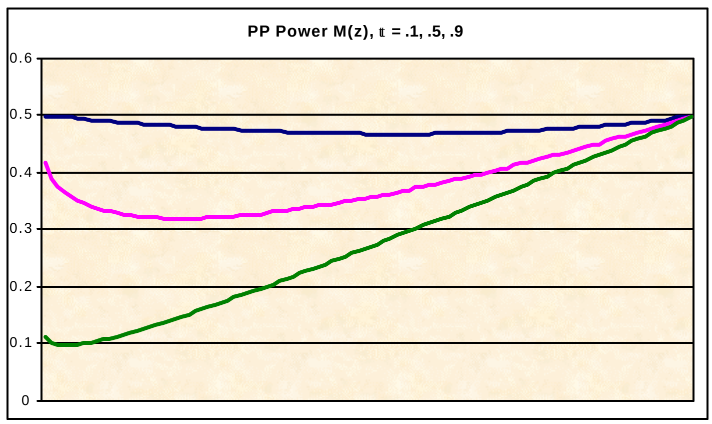
Since E(V) = ½, every copula will have M(1) = ½, so the differences in M among copulas will be for lower values of z and the shape of the curve approaching z = 1.

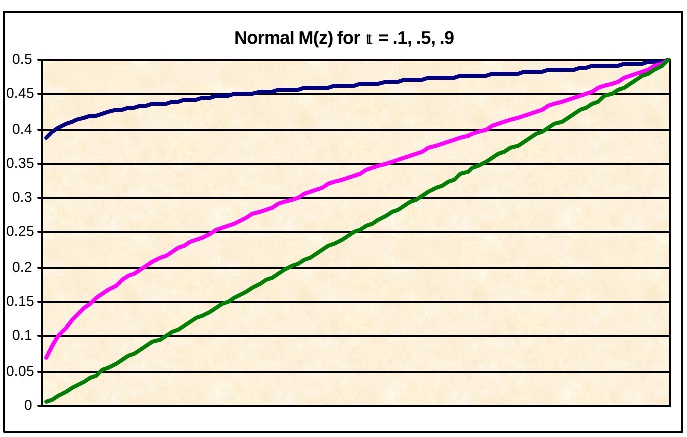
Often the integral has to be done numerically, but for a few copulas it is done explicitly in Appendix 2. Graphs of this function for several copulas are shown below. For this function, the lower τ is, the closer the values stay to ½.

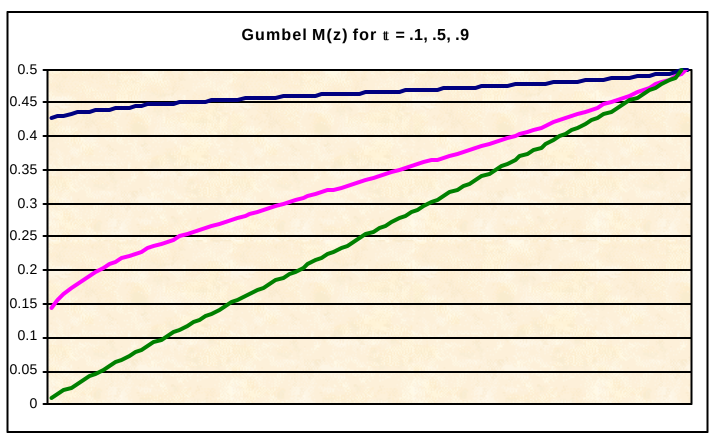












Copula Distribution Function

Genest and Rivest[4] define a function $K(z) = Pr(C(u,v)<z)$. 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.

For a sample of n pairs of observations (x_j,y_j), $K(z)$ can be estimated by first calculating the empirical $F(x,y)$ at each observation as $1/(n-1)$ times the number of other observations that are lower in both variables. Then for any z, the empirical $K(z)$ is the proportion of points with this empirical distribution function less than or equal to z.

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

4. Flipping a Copula

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),

---

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

but C_F is a copula. We will 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[5], 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:

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

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

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

---

[5] 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.

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

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

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[6] assume Pareto marginals for both distributions, but compare fits for several copulas. Klugman and Parsa[7], 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 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 to select among copulas. Plotting the empirical K(z) against the values from several copulas, they found the Gumbel looked best. It also gave the best value for the Akaike information criterion, which 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 was produced by the Gumbel copula with a = 1.453. This gives $\tau = 0.31$. Klugman and Parsa estimate the Frank a = 3.07438, which also gives $\tau = 0.31$.

---

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

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

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 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[8] for loss, a = 1.046 = c, b = 11,577.7, and for expense, a = 1.57658, b = 10,100.2, c = 0.573534. The table below converts these parameters to median and tail heaviness (= c for the inverse Burr). 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

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:

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

---

[8] The inverse Burr with a = c they call the inverse paralogistic, which is 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.

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, $b = (\text{median})1.15^{1/p}$. The medians from Klugman and Parsa with p = 1 and q = 1.5 give (rounded):

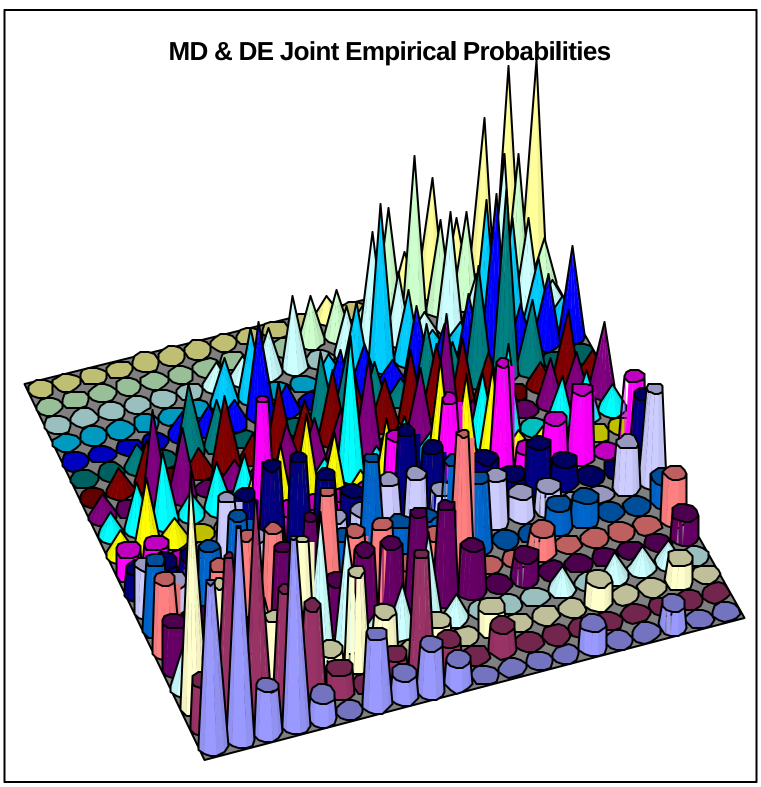
$$F(x,y) = 1-[1+x/14150]^{-1.11}-[1+(y/6450)^{1.5}]^{-1.11}+[1+x/14150+(y/6450)^{1.5}]^{-1.11}$$

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

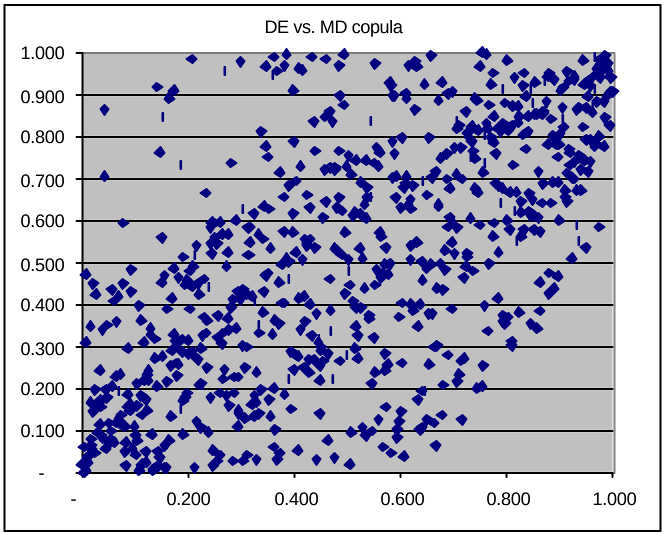
$$F_{Y|X}(y \mid x) = 1-[1+(y/d_x)^{1.5}]^{-2.11} \text{ with } d_x = 6450 + 11x^{2/3}$$

Simulated Hurricane Losses

A simulation of n=727 losses from a hurricane loss generator for a sample data set of Maryland and Delaware exposures will be used as an example of 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 above. The graph 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 3. A scatterplot of the empirical probabilities is shown below.

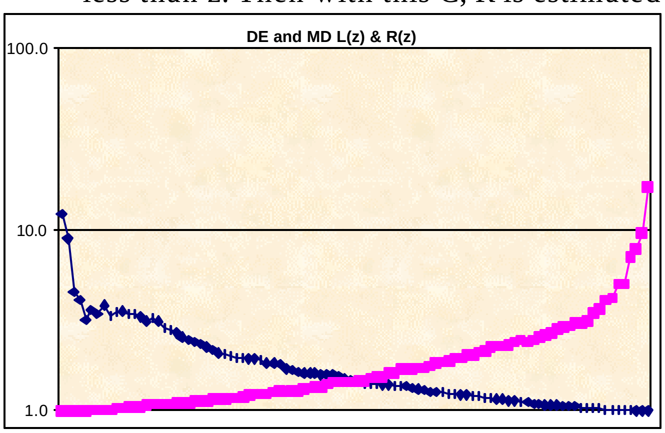


The usual estimate for the Kendall tau is to compute the average value over all pairs of observations (u_i,v_i), (u_j,v_j), $i<j$ of $\text{sign}[(u_i - u_j)(v_i - v_j)]$. In this case the estimate is $\tau = .4545$.

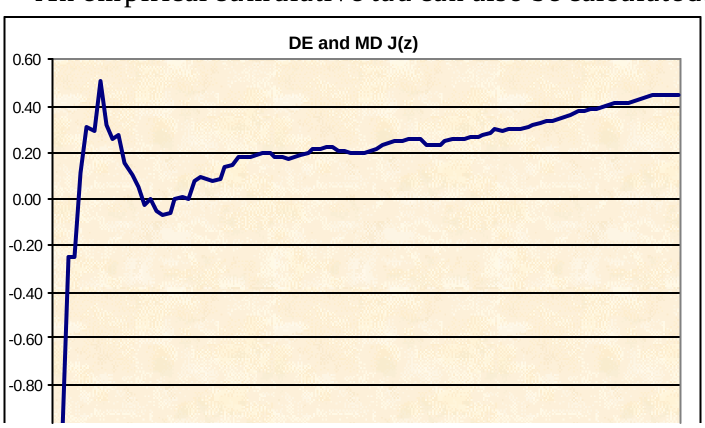


An empirical copula can also be built at each point by counting the other points that are less in both states. As there are n–1 other pairs, the count divided by n–1 can be taken as an estimate of the copula at that point. For this data, the maximum empirical copula value is 0.9821 and the average is 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 $C(z,z)/z^2$ 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 $R(z) = [1 - 2z + C(z,z)]/(1 - z)^2$. These functions are graphed at the left. Since both tails are fairly concentrated, single-tailed copulas like the HRT, PP Max, and Clayton are not indicated. The right tail looks a little light for the PP Power and Gumbel as well, so the Frank and normal copulas are likely to fit the best.

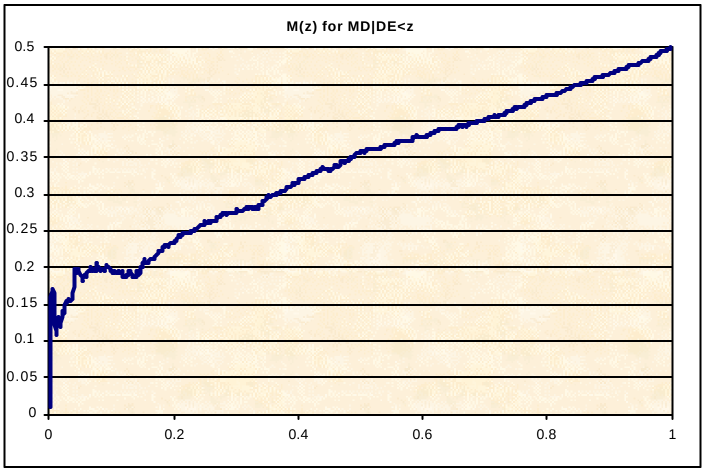


An empirical cumulative tau can also be calculated. For each z, the empirical C(u,v) can be computed for each (u,v) pair with both u and v less than z. Then the average of these values estimates the average copula in the square from



(0,0) to (z,z). This divided by C(z,z), times four less one, is the estimate of J(z).

The graph of this at right 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 J for the Frank copula.



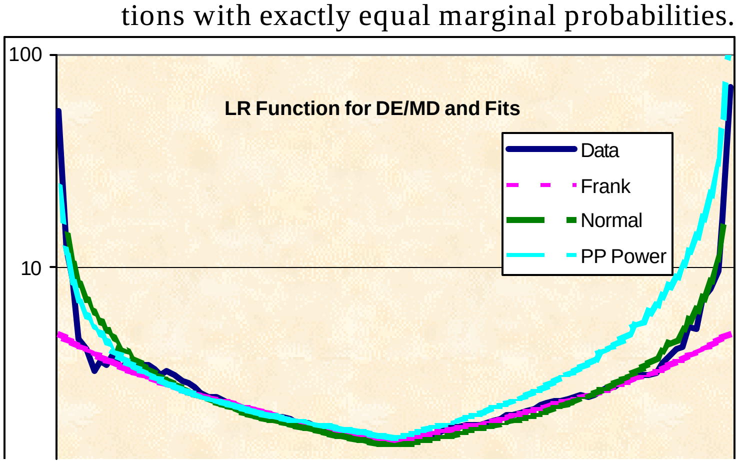
The M(z) function can be calculated either for DE | MD or MD | DE. The graph at right shows MD | DE. It is most like the M function for the normal copula.

The descriptive functions thus suggest that the normal and Frank copulas are the most likely to fit this data.

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 parameter, the ordering of the likelihood function is the same as those from the various information criteria such as AIC, etc.

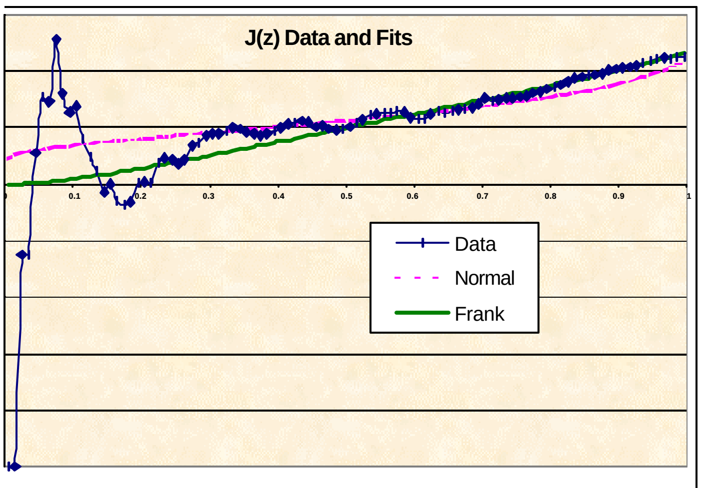
	HRT	Gumbel	Frank	Normal	Flipped Gumbel
Parameter	0.968	1.67	4.92	0.624	1.68
Ln Likelihood	124	157	183	176	161
Tau	0.34	0.40	0.45	0.43	0.40

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



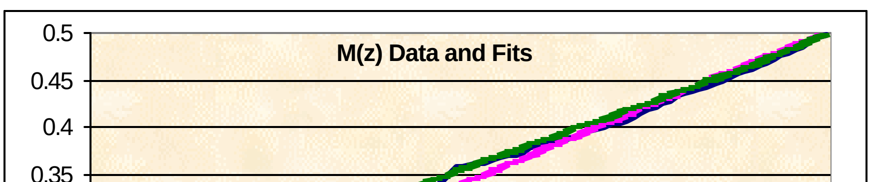
this data some of the descriptive functions seem to make this copula unlikely. The likelihood function favors the Frank copula in this case. Some of the functions are graphed for the fit and the data for this and in some cases some other copulas at left and below.

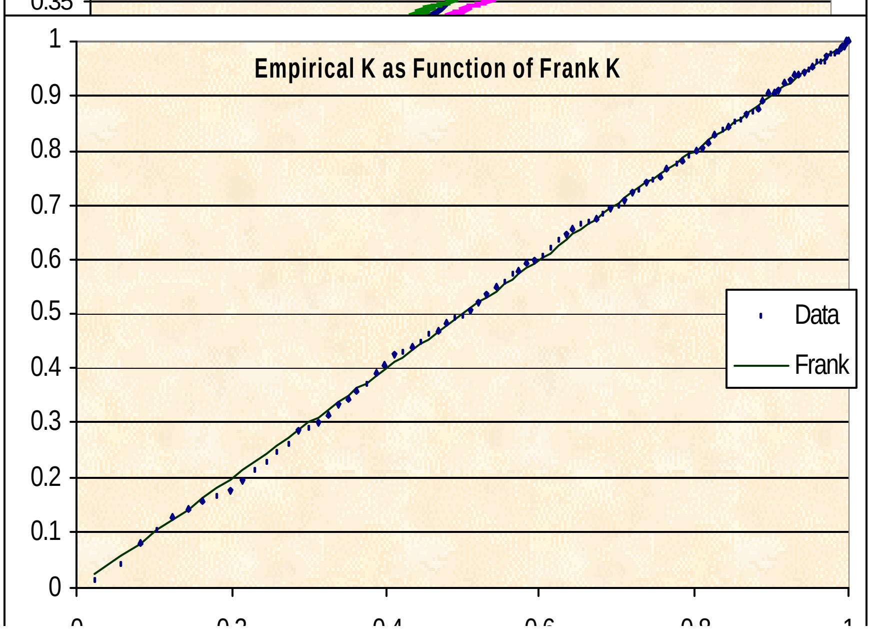
The L and R functions are combined in the graph at the left. R(z) is shown for z>0.5, and L(z) for z<0.5. The normal copula looks like a closer fit than the Frank in the right tail and in some of the left tail. The PP Power appears to be too heavy in the right tail for this data set.



The graph at right 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.

M(z) for the data and the Frank and normal copulas is graphed below. The normal copula seems to give the better fit for small events, and the Frank looks better in the middle of the range.





The K(z) function also has an empirical version. For any z this can be calculated as the proportion of empirical values of C(u,v) that are less than z. A scatterplot of the empirical K as a function of the Frank K is shown at left, along with the line x=y. The values are very close. This supports the fit, but the fit problems in the tails are difficult to discern with this function.

Even though the Frank copula provides the best fit according to the likelihood function, the normal copula might be a more useful fit for reinsurance applications, given its better match in the right tail.

6. Conclusion

Copulas provide a convenient way to model and simulate correlated variates. A number of copulas with varying shapes are available for modeling various types of relationships. Shape differences among copulas can be discerned using the descriptive functions. This can be helpful both in fitting copulas to data and in using informed judgment to select a copula for a given application.

Statisticians have identified a fair number of copulas[9]. The use of the descriptive functions provides an avenue for researching their properties. The hurricane model fit shows that a copula intermediate between the Frank and normal copulas could be useful, and some of the copulas in the literature may in fact be like that.

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. A version of J dividing by C(z,z) instead of $C(z,z)^2$ could be useful also. Even though that would increase from zero for every copula, there could be shape differences among copulas that would show up in QQ plots. 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.

---

[9] E.g., see R Nelson, *An Introduction to Copulas*, Springer Lecture Notes in Statistics, 1999

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[10].

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.

---

[10] See H. Joe, *Multivariate Models and Dependence Concepts*, Chapman and Hall, 1997

Appendix 1 – 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 $y=z^{a+1}$.

Appendix 2 – M(z)

Partial Perfect Maximum

$M(z) = ½ – ½ (z > a)(1 – a)(1 – z)(z – a) / z$

Partial Perfect Power

$M(z) = ½ + (z^{a+1} – z^{a}) / [(a+1)(a+2)]$

Appendix 3 – 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

This page intentionally left blank.