

A Partially Comonotonic Algorithm for Loss Generation

Rodney E. Kreps
Guy Carpenter InStrat, Seattle, USA

1501 Fourth Avenue, Suite 1400, Seattle WA 98101, USA
Tel. 206.621.3088 – Fax. 206. 625.6926 – E-mail. Rodney.E.Kreps@GuyCarp.com

ABSTRACT

A simple multivariate algorithm and corresponding copula are introduced which allow varying dependency as a function of loss size. In contrast to the Gaussian copula, large losses from Pareto distributions can be correlated. For the bivariate problem, special cases give the upper Frechet bound, independence, and the lower bound.

1. INTRODUCTION

It is important in modeling dependencies to try to model from the physical cause, and not just try to stuff data into a convenient mathematical form. One source of correlations is the occurrence of catastrophes. In particular, one would expect from a large earthquake that there would be large losses to property, workers' compensation, and automobile lines. However, in general these losses are not particularly correlated except by inflation. What is desired is a method of generating correlations that will correlate large losses, but not small.

What follows is a copula method with a very simple form and simulation algorithm. Its advantage is that for arbitrary marginal distributions, and in particular for Paretos, the correlation remains even in the large loss limit. As pointed out by Embrechts, McNeil and Strauman in *Correlation and Dependency in Risk Management* in the XXX ASTIN Colloquium, asymptotically the normal copula loses the dependency between the large losses. The present method allows direct control of the tail dependencies. The present method also can be applied to an arbitrary number of simultaneously dependent losses.

The disadvantages are that the copula is very concentrated where there are dependencies, and that simultaneous claims are either all independent or all comonotonic.

A discussion of the bivariate case is given first with sections on the algorithm, the cumulative distribution function, the copula, and the density functions. Symmetry considerations, linear correlation, and examples follow this. The whole of the above is repeated in condensed form for the multivariate version, followed by a small conclusion and an appendix with sample realizations.

2. THE BIVARIATE ALGORITHM:

The basic idea is to have losses be, on a random basis, either independent or comonotonic. This method begins with a generating function $h(U_1, U_2)$ which is symmetric, integrable, has range between 0 and 1 in the unit square, but is otherwise arbitrary.

The algorithm is as follows: Draw uniform random deviates U_1, U_2, U_3 .

If $U_3 > h(U_1, U_2)$ (independent case)
then $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(U_2)$.

If $U_3 < h(U_1, U_2)$ (comonotonic case)

then $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(U_2)$.

where X_1 and X_2 are the desired random deviates.

3. CUMULATIVE DISTRIBUTIONS:

In order to get the cumulative joint distribution from this algorithm, consider that for given U_1, U_2 , and U_3 the conditional probabilities give

$$\begin{aligned} \Pr(X_1 < x_1, X_2 < x_2) &= \Pr(U_1 < F_1, U_2 < F_2 | U_3 > h(U_1, U_2)) \Pr(U_3 > h(U_1, U_2)) \\ &+ \Pr(U_1 < F_1, U_1 < F_2 | U_3 < h(U_1, U_2)) \Pr(U_3 < h(U_1, U_2)) \end{aligned}$$

Here and subsequently F_1 is an abbreviation of $F_1(x_1)$, and so on. These probabilities are all zero or 1, and can be written in terms of the index function $\Theta(x)$ [= 0 for x negative, 1 for x positive]:

$$\begin{aligned} \Pr(X_1 < x_1, X_2 < x_2) &= \Theta(F_1 - U_1) \Theta(F_2 - U_2) \Theta(U_3 - h(U_1, U_2)) \\ &+ \Theta(F_1 - U_1) \Theta(F_2 - U_1) \Theta(h(U_1, U_2) - U_3) \\ &= \Theta(F_1 - U_1) \Theta(F_2 - U_2) [1 - \Theta(h(U_1, U_2) - U_3)] \\ &+ \Theta(F_1 - U_1) \Theta(F_2 - U_1) \Theta(h(U_1, U_2) - U_3) \end{aligned}$$

using $\Theta(-x) = 1 - \Theta(x)$ in the last step.

In order to get the probabilities not conditioned on the U_i , integrate over their (uniform) distributions:

$$\begin{aligned} F(x_1, x_2) &= \int_0^{F_1} dU_1 \int_0^1 dU_2 \int_0^1 dU_3 \left\{ \begin{aligned} &\Theta(F_1 - U_1) \Theta(F_2 - U_2) [1 - \Theta(h(U_1, U_2) - U_3)] \\ &+ \Theta(F_1 - U_1) \Theta(F_2 - U_1) \Theta(h(U_1, U_2) - U_3) \end{aligned} \right\} \\ &= F(x_1) F(x_2) - \int_0^{F_1} dU_1 \int_0^{F_2} dU_2 h(U_1, U_2) + \int_0^{F_<} dU_1 \int_0^1 dU_2 h(U_1, U_2) \end{aligned}$$

where $F_<$ is the smaller of $F_1(x_1)$ and $F_2(x_2)$.

4. THE COPULA:

A reminder on copula methods: the marginal distributions $F_i(x_i)$ are specified, and the dependency between variables is carried in the copula $C(U_1, U_2)$, which is a joint distribution function of two uniform deviates. The actual joint distribution is then

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$$

The obvious extension for N variables holds.

The copula here is

$$C(U, V) = UV - \int_0^U dU_1 \int_0^V dU_2 h(U_1, U_2) + \int_0^{\min(U, V)} dU_1 \int_0^1 dU_2 h(U_1, U_2)$$

It is simple to verify that $C(U, 1) = U$ and $C(1, V) = V$ as is necessary to have a copula. For $h = 1$ the copula is the upper Frechetⁱ bound, $\min(U, V)$. For $h = 0$ the distributions are independent.

There is an extension to the algorithm which will give the mathematically interesting but physically rare case of countermonotonic behavior. Allow h to take on negative values, and expand the algorithm to

If $U_3 > |h(U_1, U_1)|$ (independent)
 then $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(U_2)$.

If $U_3 < |h(U_1, U_1)|$
 then if $h(U_1, U_2) > 0$ (comonotonic)
 then $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(U_2)$.

Else $[h(U_1, U_2) < 0]$ (countermonotonic)
 $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(1 - U_1)$.

This leads in similar fashion to the above to the copula

$$C(U, V) = UV - \int_0^U dU_1 \int_0^V dU_2 |h(U_1, U_2)| + \int_0^{\min(U, V)} dU_1 \int_0^1 dU_2 (h(U_1, U_2))_+ \\ + \Theta(U + V - 1) \int_{1-V}^U dU_1 \int_0^1 dU_2 (-h(U_1, U_2))_+$$

where as usual $(g)_+ \equiv g \Theta(g) = \max(g, 0)$. However, in order to be a copula this formula needs an additional symmetry $h(U, V) = h(1 - U, V)$ as well as the earlier $h(U, V) = h(V, U)$.

This leads to reflection symmetry about the diagonals of the unit square and also about the lines $U = 0.5$ and $V = 0.5$. Thus, the values of h in the upper right corner of the square are replicated in all the other corners. If h is non-zero for joint large claims then to the same degree it is also non-zero for large and small claims, and for joint small claims.

Note that $h = 1$ is the upper Frechet bound $\min(U, V)$; $h = 0$ is independence; and now $h = -1$ is the lower Frechet bound $(U + V - 1)_+$.

5. PROBABILITY DENSITY FUNCTIONS:

The joint probability density function is

$$\frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} = f_1(x_1) f_2(x_2) \left\{ 1 - h(F_1, F_2) + \delta(F_1 - F_2) \int_0^1 h(U, F_c) dU \right\}$$

where the f_i are the marginal densities.

For readers unfamiliar with it, the δ function was introduced to the physics community by Diracⁱⁱ in 1926. It can conveniently be thought of as the density function of a normal distribution with mean zero and standard deviation arbitrarily small compared to everything else under consideration. It has the effect of being zero except where its argument is zero, and integrates to the index function. Conversely, it allows differentiation of the index function and simple representation and manipulation of otherwise troublesome objects.

In particular, $\frac{\partial}{\partial x} \min(x, y) = \Theta(y - x)$ and $\frac{\partial^2}{\partial x \partial y} \min(x, y) = \delta(y - x)$. In the present case the δ function ensures that the final expression is only non-zero when $F_1 = F_2 = F_{\leftarrow}$. It describes a mass line in the density.

For the copula the density is

$$\frac{\partial^2 C(U, V)}{\partial U \partial V} = 1 - h(U, V) + \delta(U - V) \int_0^1 h(U_1, V) dU_1$$

The second argument of h under the integral could equally well be U or $(U+V)/2$ because the δ function requires $U = V$.

6. ASYMMETRIC GENERATOR:

If $h(U_1, U_2)$ is not symmetric in its arguments, then the algorithm as given does not in general result in a copula. For example,

$$C(1, V) = V + \int_0^V dU_1 \int_0^1 dU_2 [h(U_1, U_2) - h(U_2, U_1)] \neq V$$

The algorithm can be modified to: Draw uniform random deviates U_1, U_2, U_3, U_4 .

If $U_3 > h(U_1, U_2)$ (independent case)

then if $U_4 < 1/2$ $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(U_2)$.

otherwise $X_1 = F_1^{-1}(U_2)$ and $X_2 = F_2^{-1}(U_1)$.

If $U_3 < h(U_1, U_2)$ (comonotonic case)

then if $U_4 < 1/2$ $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(U_1)$.

otherwise $X_1 = F_1^{-1}(U_2)$ and $X_2 = F_2^{-1}(U_2)$.

Where X_1 and X_2 are the desired deviates.

The copula becomes

$$C(U, V) = UV - \int_0^U dU_1 \int_0^V dU_2 \frac{h(U_1, U_2) + h(U_2, U_1)}{2} + \int_0^{\min(U, V)} dU_1 \int_0^1 dU_2 \frac{h(U_1, U_2) + h(U_2, U_1)}{2}$$

Clearly this only involves the symmetric part of h , so in this case nothing is gained by the asymmetric form. The author has not found a good algorithm which uses a general asymmetric h . There may be particular exceptions.

7. LINEAR CORRELATION:

The covariance between the variables is given by

$$Cov(X_1, X_2) = \int_0^1 dU F_1^{-1}(U) F_2^{-1}(U) \int_0^1 dV h(U, V) - \int_0^1 dU \int_0^1 dV F_1^{-1}(U) F_2^{-1}(V) h(U, V)$$

The maximum dependency is when $h(U_1, U_2) = 1$ everywhere. This implies that X_1 and X_2 are comonotonic and leads to the bound

$$\text{Cov}(X_1, X_2) \leq \int_0^1 F_1^{-1}(U) F_2^{-1}(U) dU - \text{Mean}_1 \text{Mean}_2$$

The implied correlation is 1 when the two distributions are identical, but is in general less than 1 otherwise.

Clearly, for any given value of the covariance and forms of the distributions, there are many (or perhaps no) forms for h which will give the same covariance and different joint distributions.

8. PARTICULAR CASES:

For the special case where h is multiplicatively separable [$h(U_1, U_2) = h(U_1)h(U_2)$] define

$$H(x) = \int_0^x h(U) dU. \text{ Then the copula takes on the particularly simple form}$$

$$C(U, V) = UV - H(U)H(V) + H(\min(U, V))H(1)$$

and
$$\frac{\partial^2 C(U, V)}{\partial U \partial V} = 1 - h(V)[h(U) - \delta(U - V)H(1)].$$

The case $h(U_1, U_2) = h\Theta(U_1 - K)\Theta(U_2 - K)$ can be used, for example with $K = 0.9$, to correlate only large claims, as motivated the development in the first place. This would seem the simplest example. Other possibilities that do not necessarily require both U_1 and U_2 to be large are forms such as $h(U_1, U_2) = h\Theta(U_1 + U_2 - K)$ and

$h(U_1, U_2) = h \min(1, \Theta(U_1 - K) + \Theta(U_2 - K))$, and the even simpler and analytic

$$h(U_1, U_2) = h \left[\frac{(U_1 + U_2)}{2} \right]^\alpha \text{ and } h(U_1, U_2) = (U_1 U_2)^\alpha h. \text{ The latter two have the advantage}$$

that there is increasing dependency as the losses get larger. Clearly, by choosing different forms of h many different joint distributions can be obtained.

Returning the first example, taking $h(U, V) = h\Theta(U - K)\Theta(V - K)$ with $0 \leq h, K \leq 1$ implies

$$F(x_1, x_2) = F(x_1)F(x_2) - h(F_1 - K)_+(F_2 - K)_+ + h(F_2 - K)_+(1 - K)$$

The joint density is

$$\frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} = f_1(x_1)f_2(x_2)\{1 - h\Theta(F_1 - K)[\Theta(F_2 - K) - \delta(F_1 - F_2)]\}$$

9. N VARIABLES:

Choose U_1, U_2, \dots, U_N uniform deviates. Define $2^N - N - 1$ functions h , one for each possible combination of more than one variables. Label them in some descriptive way, such as using subscripts for the appropriate variables. For example, the four functions for three variables are h_{12}, h_{13}, h_{23} , and h_{123} . Each h should be a symmetric function of its arguments and restricted to the range $[0, 1]$. Note that some or all of the functions may be identically zero, in which case that particular dependency is not physically required.

The sum of the h must be ≤ 1 . Let h_0 (all variables independent) = $1 -$ the sum of the h . Thus the functions h together with the h_0 form a partition of the unit interval.

With U another uniform deviate, the algorithm for fixed $U_1..U_N$ is:

Let U pick out the corresponding function by a random choice on the partition; make those variables comonotonic (for example on the variable with the smallest index).

The corresponding CDF is obtained similarly to the two variable development, again converting the index function on h_0 by using $\Theta(-x) = 1 - \Theta(x)$. The result is most easily expressed in terms of auxiliary functions D :

Let n_1, n_2, \dots, n_K be an arbitrary non-empty combination of the integers 1 to N , denoted collectively by the K -vector \bar{n} . Represent the corresponding K -subset of U as $U_{\bar{n}}$ and define

$$D_{\bar{n}}(U_{\bar{n}}) \equiv \int_0^{\min(U_{\bar{n}})} dV_1 \int_0^1 dV_2 \dots \int_0^1 dV_K h_{\bar{n}}(\bar{V}) - \int_0^{U_{n_1}} dV_1 \int_0^1 dV_2 \dots \int_0^1 dV_K h_{\bar{n}}(\bar{V})$$

Note that $D_i(U_i) = 0$ identically, since there is only one variable. Also, as any particular U_i in the combination goes to 1, the function D goes to the function on the combination without U_i but with a revised h .

Then the general form of the copula for N variables under this algorithm may be expressed as

$$C(\bar{U}) = \prod_{i=1}^N U_i + \sum_{\text{all combinations}} D_{\bar{n}}(U_{\bar{n}}) \prod_{i \neq \bar{n}} U_i$$

As usual $F(x_1, x_2, \dots, x_N) = C(F_1(x_1), F_2(x_2), \dots, F_N(x_N))$.

The Frechet bound is obtained for $h_{12..N} = 1$ and all other $h = 0$. All $h = 0$ gives uncorrelated deviates.

Further, as $N-2$ variables go to 1 the distribution reduces to the two-variable distribution above. In fact, as $N-K$ variables go to 1, the N -variable distribution reduces to the appropriate K -variable distribution. This property is not only desirable, but also probably even a good requirement for multivariate forms.

It is also interesting that there are dependencies possible here that are not possible in the multivariate normal form, in which only the pairwise dependencies are addressed.

The density for D is

$$d_{\bar{n}}(U_{\bar{n}}) = \left\{ \int_0^1 \prod_{i=1}^K \delta(U_{n_i} - V_i) dV_1 \int_0^1 dV_2 \dots \int_0^1 dV_K h_{\bar{n}}(\bar{V}) \right\} - h_{\bar{n}}(U_{\bar{n}})$$

Although the integration over V_1 can be done immediately using any one of the K δ functions, it is left in this form to show explicitly the symmetry among the variables. Note also that because of the remaining $K-1$ δ functions the first term is actually restricted to an $N-K+1$ dimensional subspace.

The copula density is $c(\bar{U}) = 1 + \sum_{\text{all combinations}} d_{\bar{n}}(U_{\bar{n}})$.

The corresponding general multivariate density function is as usual

$$f(x_1, x_2, \dots, x_N) = c(F_1, F_2, \dots, F_N) \prod_{i=1}^N f_i(x_i)$$

The covariance of any two variables can be easily obtained as

$$\text{Cov}(X_i, X_j) = \int_0^1 d\bar{U} F_i^{-1}(U_i) F_j^{-1}(U_j) c(\bar{U}) - \text{mean}_i \text{mean}_j$$

It is helpful in evaluation to note that integrating $d_{\bar{n}}(U_{\bar{n}})$ over any $K-1$ of its variables results in zero, so that the only combinations which survive are those that involve both i and j . This is exactly what one expects intuitively, of course.

Example: $h_{12..N} = h \prod_{i=1}^N \Theta(U_i - K)$ with $0 \leq h, K \leq 1$ and all other $h = 0$. Then

$$F(x_1, x_2, \dots, x_N) = \prod_{i=1}^N F(x_i) - h \prod_{i=1}^N (F_i - K)_+ + h(1-K)^{N-1} (F_{\zeta} - K)_+$$

The density and covariances are

$$f(x_1, x_2, \dots, x_N) = \prod_{i=1}^N f_i(x_i) \left\{ 1 - h\Theta(F_{\zeta} - K) + h \left[\prod_{i=2}^N (1-K) \delta(F_i - F_1) \right] \Theta(F_1 - K) \right\}$$

and

$$\text{Cov}_{i,j} = h(1-K)^{N-2} \left\{ (1-K) \int_K^1 dU F_i^{-1}(U) F_j^{-1}(U) - \int_K^1 dU F_i^{-1}(U) \int_K^1 dV F_j^{-1}(V) \right\}$$

10. CONCLUSION:

The algorithm clearly is simple to implement. Doing so in a spreadsheet is both straightforward and instructive, as repeated simulation allows one to develop intuition about the effects of different forms of the generating function and of parameters within forms.

The algorithm allows control over the degree of dependency for any range of values, and can obviously be tailored to keep tail dependencies.

The biggest drawback is the concentration of values along the comonotonic curve. That is, there is no spread away from it except for what is provided by the random background. That being said, the generated data sets of the author's experience look fairly good, and certainly useable for simulation. See the examples on the next few pages.

The four examples are all from the choice $h(U, V) = \Theta(UV - K^2) H(UV)^\alpha$ with marginal distributions being different Paretos. The Pareto parameters are characteristic of property and casualty lines. The examples are repeated twice. The first time is just a scatter plot, the second has guidelines for (1) the means of the distributions, (2) the curve above which the partial comonotonicity comes into play, and (3) the curve along which the comonotonic points lie. Note that with the choice of parameters in the appendix, it is increasingly likely that losses will be correlated as they get large. The parameters were chosen to give reasonable values for the fraction of claims comonotonic and their dollar value. Linear

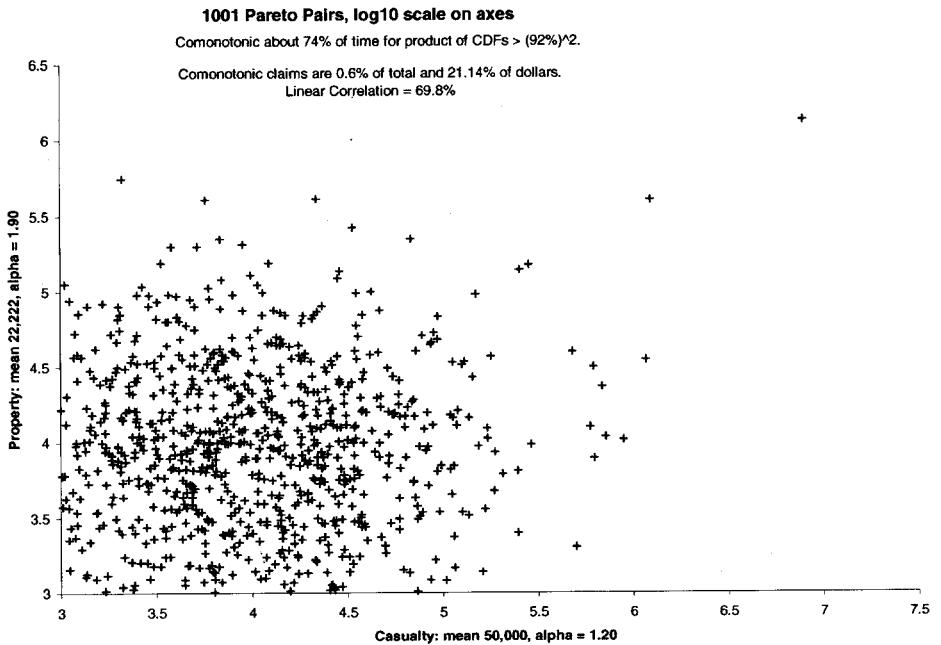
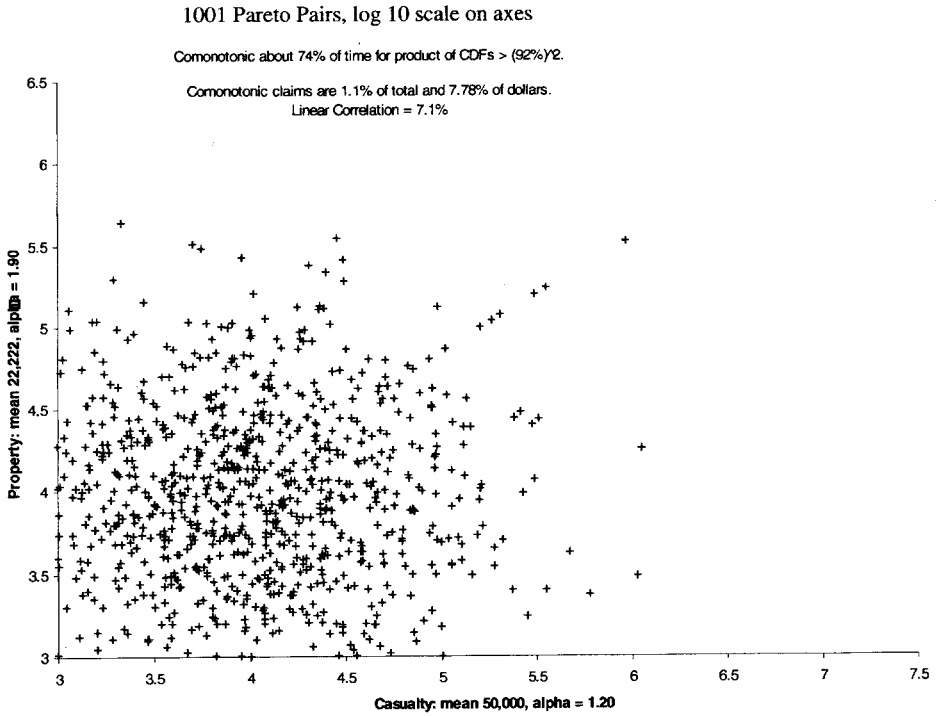
correlation varies considerably. Specifically, on samples of 1001 claim pairs the mean of the fraction of claims comonotonic is 0.90% and its standard deviation is 0.30%; the mean of the dollar fraction of claims comonotonic is 6.9% and its standard deviation is 7.3%; and the mean of the linear correlation is 7.2% and its standard deviation is 16.2%. The latter distribution is extremely skewed, usually by the presence of one or two extremely large pairs. On 10,000 simulations of 1001 pairs the minimum observed was -5.0% and the maximum was 99.9%.ⁱⁱⁱ

ⁱ See the cited paper by Embrechts *et al.* for the Frechet bounds and other discussion of copula methods.

ⁱⁱ P.A.M. Dirac, "The Physical Interpretation of Quantum Dynamics" Dec. 1926 Proceedings of the Royal Society of London, Series A, Vol. 113 pp. 621-641

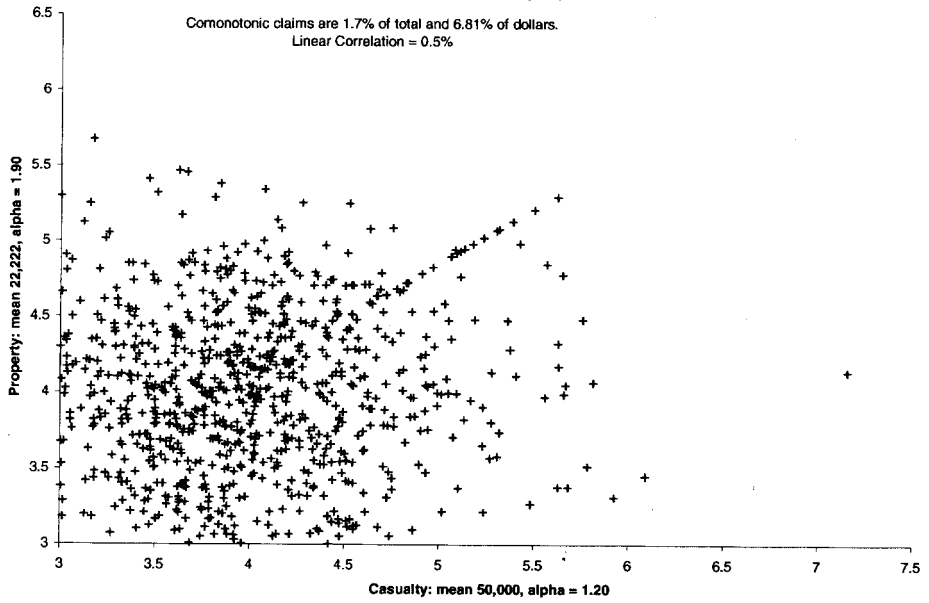
ⁱⁱⁱ

A Appendix: Some sample simulations.



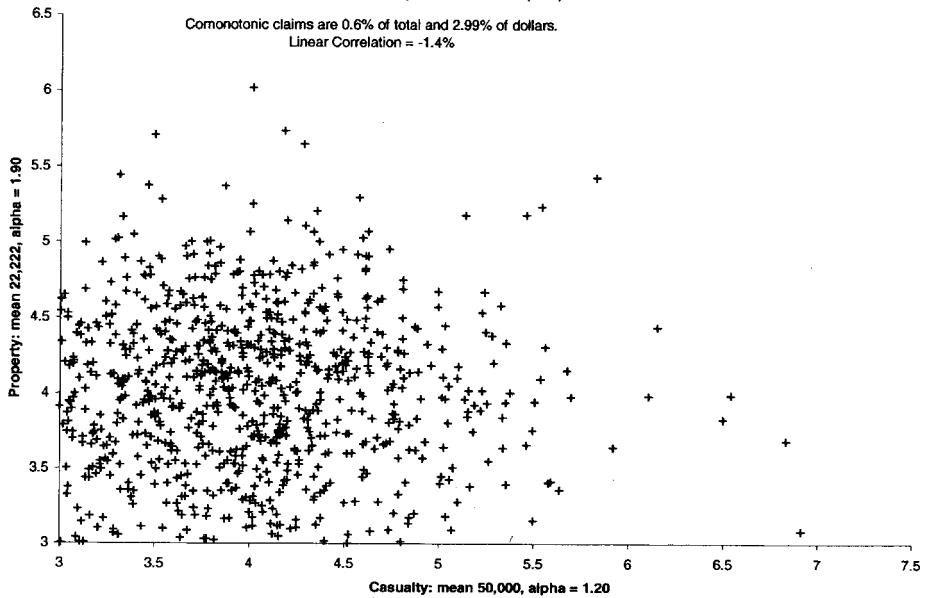
1001 Pareto Pairs, log10 scale on axes

Comonotonic about 75% of time for product of CDFs $> (92\%)^2$.
Comonotonic claims are 1.7% of total and 6.81% of dollars.
Linear Correlation = 0.5%



1001 Pareto Pairs, log10 scale on axes

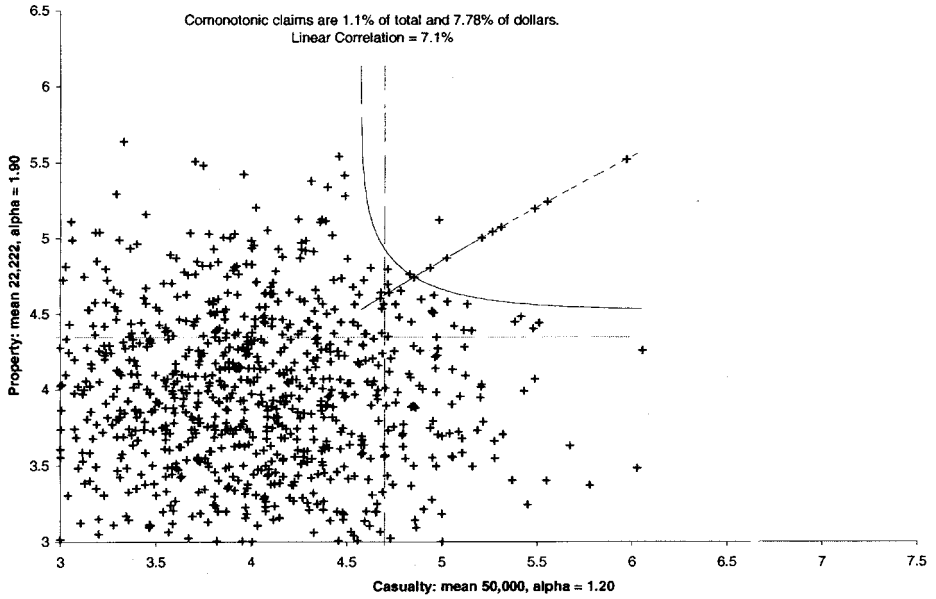
Comonotonic about 68% of time for product of CDFs $> (92\%)^2$.
Comonotonic claims are 0.6% of total and 2.99% of dollars.
Linear Correlation = -1.4%



1001 Pareto Pairs, log10 scale on axes

Comonotonic about 74% of time for product of CDFs > (92%)².

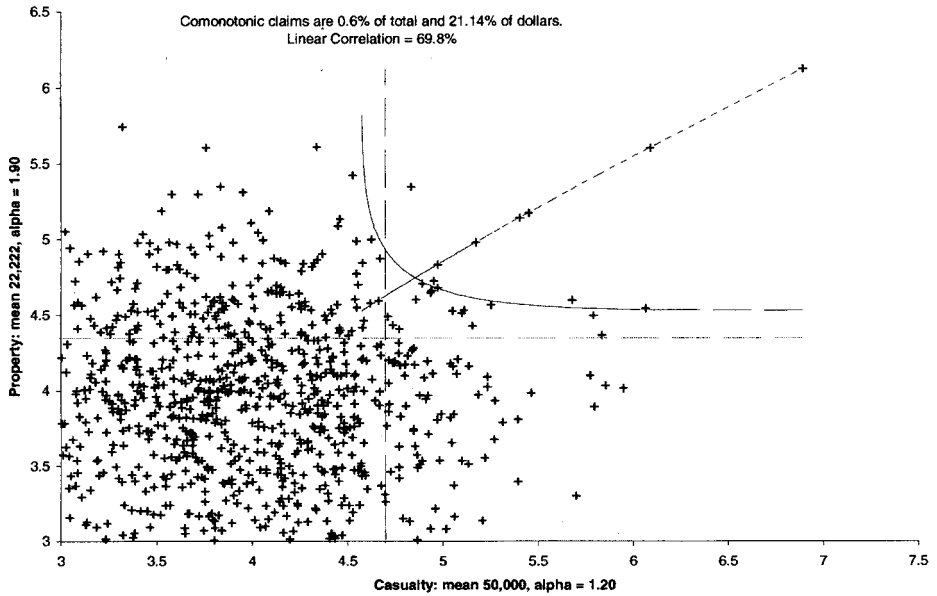
Comonotonic claims are 1.1% of total and 7.78% of dollars.
Linear Correlation = 7.1%



1001 Pareto Pairs, log10 scale on axes

Comonotonic about 74% of time for product of CDFs > (92%)².

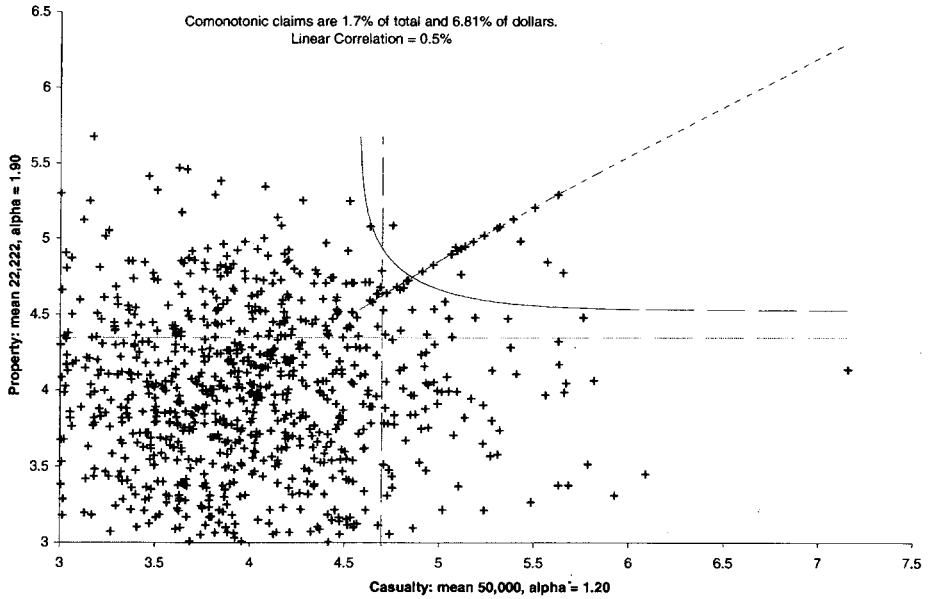
Comonotonic claims are 0.6% of total and 21.14% of dollars.
Linear Correlation = 69.8%



1001 Pareto Pairs, log10 scale on axes

Comonotonic about 75% of time for product of CDFs > (92%)².

Comonotonic claims are 1.7% of total and 6.81% of dollars.
Linear Correlation = 0.5%



1001 Pareto Pairs, log10 scale on axes

Comonotonic about 68% of time for product of CDFs > (92%)².

Comonotonic claims are 0.6% of total and 2.99% of dollars.
Linear Correlation = -1.4%

