ACTUARIAL RESEARCH CLEARING HOUSE
1993 VOL. 1

SIX BRIDGES TO Ψ'S

by

William A. Bailey

Kemper National Insurance Companies
1 Kemper Drive
Long Grove, Illinois 60049-0001

ABSTRACT

Six methods are included for calculating the probability of ultimate ruin, where the waiting times between claims are independent and identically distributed; *those distributions need not be exponential.*

The first three use Monte Carlo and the last three do not.

The first Monte Carlo method uses two-dimensional random sequences to generate a distribution of maximal aggregate losses and the connection between the maximal aggregate loss random variable and the probability of ruin. The first Convolution method is similar to the first Monte Carlo method except that the distribution of maximal aggregate loss is generated by performing two-dimensional generalized numerical "convolutions".

The second Monte Carlo method uses one-dimensional random sequences to obtain the probability of ruin, starting with a given initial surplus. The second Convolution method is similar to the second Monte Carlo method except that the probability of ruin is generated by one-dimensional numerical convolutions.

The third Monte Carlo method uses the second Monte Carlo method starting with zero initial surplus, together with the fact that the maximal aggregate loss random variable has a compound geometric distribution. The third Convolution method uses the second Convolution method starting with zero initial surplus, together with the fact that the maximal aggregate loss random variable has a compound geometric distribution.

Some preliminary one-dimensional convolutions are performed to establish two parameters $\hat{n}$ and $\hat{\Delta}$, which are then used in any of the six methods.

The appendixes introduce techniques to facilitate the performing of one- and two-dimensional regular and generalized numerical convolutions.

TABLE OF CONTENTS

There is an apparent paradox (ref[2]) in the idea that Monte Carlo trials (simulation by random numbers) can be used to evaluate probabilities of ultimate ruin. However, such an approach is possible, and quite practical except where the relative security margin is close to zero.

Six methods are included for calculating the probability of ultimate ruin, where the waiting times between claims are independent and identically distributed; *those distributions need not be exponential.*

The first three use Monte Carlo and the last three do not.

The first Monte Carlo method uses two-dimensional random sequences to generate a distribution of maximal aggregate losses and the connection between the maximal aggregate loss random variable and the probability of ruin. The first Convolution method is similar to the first Monte Carlo method except that the distribution of maximal aggregate loss is generated by performing two-dimensional generalized numerical "convolutions".

The second Monte Carlo method uses one-dimensional random

sequences to obtain the probability of ruin, starting with a given initial surplus. The second Convolution method is similar to the second Monte Carlo method except that the probability of ruin is generated by one-dimensional numerical convolutions.

The third Monte Carlo method uses the second Monte Carlo method starting with zero initial surplus, together with the fact that the maximal aggregate loss random variable has a compound geometric distribution. The third Convolution method uses the second Convolution method starting with zero initial surplus, together with the fact that the maximal aggregate loss random variable has a compound geometric distribution.

Some preliminary one-dimensional convolutions are performed to establish two parameters $\hat{n}$ and $\hat{\Delta}$, which are then used in any of the six methods.

The appendixes introduce techniques to facilitate the performing of one- and two-dimensional regular and generalized numerical convolutions.

We will use the following random variables among others:

W_i is the waiting time between claim number i-1 and claim number i; the W_i's are independent identically distributed random variables. The W_i's could be exponentially distributed, as is

true in the Compound Poisson Process; but they need not be so distributed.

X_i is the size of claim number i. The X_i's are independent identically distributed random variables. The X_i's are not restricted to taking on only nonnegative values; however, we are requiring that the $E[X_i]$ be nonnegative for each i.

We assume that W_1, W_2, ... ,X_1, X_2, ... are mutually independent.

$H_i = c \cdot W_i - X_i$ is the change in surplus from just after claim number i-1 to just after claim number i, where c is the rate at which premiums are collected. We further assume that the H_i's are mutually independent; they are identically distributed random variables.

$V_n(u) = u + H_1 + H_2 + \ldots + H_n$ is the surplus just after claim number n, having started with an initial surplus of u.

Other subscripted small letters such as w_i, x_i, h_i and u_i will be used to represent possible outcomes of the random variables W_i, X_i, H_i and U_i, respectively.

μ and σ^2 will denote the expected value and variance, respectively, of the random variable shown in the subscript position of those symbols.

u will denote the initial surplus.

$\psi(u;n)$ will denote the probability of ruin before n+1 claims, where the initial surplus was u.

$Pr(L \leq u;n)$ will denote the probability that the maximal loss random variable L will be less than u during the period of n claims.

$\psi(u)$ will denote the probability of ultimate ruin, where the initial surplus was u.

If we were interested in the probability of ruin occurring anytime before a certain number, say n_0+1, of claims have occurred, then we could use a random number generator and compute a number, say m, of trial sequences

$$w_1, x_1, w_2, x_2, \ldots, w_{n_0}, x_{n_0}$$

of outcomes for

$$W_1, X_1, W_2, X_2, \ldots, W_{n_0}, X_{n_0},$$

respectively, and record whether any v_1 in the sequence

$$v_0 = u \text{ (the initial surplus)}$$

$$v_1 = v_0 + c \cdot w_1 - x_1$$

$$v_2 = v_1 + c \cdot w_2 - x_2$$

$$\vdots$$

$$v_{n_0} = v_{n_0-1} + c \cdot w_{n_0} - x_{n_0}$$

is less than zero.

Then, the ratio r/m, where r is the number of trial sequences in which for some $i < n_0+1$ v_i is less than zero, and m is the total number of trial sequences, would be an estimate of the probability $\psi(u;n_0)$ of ruin occurring sometime before n_0+1 claims. The expected error in the estimate would, of course, tend to decrease as m increases.

We can use a modification of this procedure to establish intervals within which the probability $\psi(u)$ of ultimate ruin is likely to lie. We will describe six such modifications, referred to as "bridges" here.

DETERMINING $\hat{v}$, $\hat{n}$ AND $\hat{\Delta}$

In APPENDIX #1 we prove three assertions, from which we can conclude that given an initial surplus of, say u, and an $\varepsilon > 0$ there exists a negative real number $\hat{v}$ and a positive integer $\hat{n}$ of claims such that

(a) the probability $Pr(L>-\hat{v};\hat{n})$ of a maximal loss of more that $-\hat{v}$ before $\hat{n}+1$ claims is less than ε; and,

(b) the probability of a surplus of less than $-\hat{v}$ just after claims number $\hat{n}$ is less than ε.

The validity of these assertions leads us to the following heuristic technique to obtain $\hat{v}$ and $\hat{n}$.

Let $H = c \cdot W - X$ and generate the distribution of H by convoluting the distribution of $c \cdot W$ with the distribution of $-X$.[1]

---

[1]For details on how to perform one-dimensional convolutions see APPENDIX #2. Throughout the paper we assume that we are dealing with either discrete or discretized distributions. The section A NUMERICAL EXAMPLE illustrates how continuous distributions can be discretized.

Let H_1, H_2, ... be mutually independent and identically distributed as H.

Let $V_1 = H_1$ and let V_1 be distributed as H_1

Let $V_2 = V_1 + V_1$ and obtain the distribution of V_2 by convoluting the distribution of V_1 with itself.
Let $V_4 = V_2 + V_2$ and obtain the distribution of V_4 by convoluting the distribution of V_2 with itself.

Etc.

As we complete each of these convolutions, we observe the minimum outcome after discarding the "probability products".[2] The minimums will tend to decrease, reach a minimum, and then increase.[3] We take $\hat{v}$ to be the minimum of the minimums. See Comment below.

Generating the minimum of the minimums from the distributions of V_2, V_4, V_8, ... will often be good enough for the purpose of determining $\hat{v}$. Closer estimates can be obtained by generating

---

[2]"probability products" for univariate convolutions are defined at the (*) in APPENDIX #2.

[3]Although this tendency is real, this statement should not be taken too literally. Transient local minimums are not ruled out.

the minimum of the minimums from the distributions of V_2, V_3, V_4, ..., but the increase in computer time may outweigh the value of the improvement in accuracy.

Suppose $\bar{n}$ is a positive integer such that $\hat{v}$ turns out to lie in the distribution of $V_{\bar{n}}$. Generate the distributions of

$$V_{\bar{n}\cdot 2} = V_{\bar{n}} + V_{\bar{n}}$$
$$V_{\bar{n}\cdot 2^2} = V_{\bar{n}\cdot 2} + V_{\bar{n}\cdot 2}$$
$$V_{\bar{n}\cdot 2^3} = V_{\bar{n}\cdot 2^2} + V_{\bar{n}\cdot 2^2}$$
$$\vdots$$

again observing as we complete each of these convolutions the minimum outcome after discarding the "probability products." The minimums will tend to increase.[4] Continue the process until the minimum is $\geq$ the positive real number $-\hat{v}$. We record the smallest positive integer $\hat{n}$ for which the distribution of $V_{\hat{n}}$ has a minimum $\geq$ the positive real number $-\hat{v}$.

Comment: $\hat{n}$ can turn out to be inordinately large; and, this situation may become apparent as we are looking for the minimum of the minimums. This may happen if the relative security margin

[4]Although this tendency is real, this statement should not be taken too literally. Transient decreases in the minimums are not ruled out.

is very close to zero, or if one of the distributions (of the univariate random variables W or X) is sufficiently pathological. However, these situations are not likely to occur in practice.

Once again, closer estimates can be obtained by generating minimums for the distributions of $V_{\hat{n}}$, $V_{\hat{n}+1}$, $V_{\hat{n}+2}$, ..., but the increase in computer time may outweigh the value of the improved accuracy.

The final cumulative of the distribution of V_n will be less than unity, because of discarding "probability products" totalling, say, $\hat{\Delta}$. At this point we have established values for the positive real number $-\hat{v}$ and the positive integer $\hat{n}$ such that

$$\Pr\{ \mathcal{L} > -\hat{v} \; ; \; \hat{n} \} < \hat{\Delta} \text{ and}$$
$$\Pr\{ \mathcal{V}_{\hat{n}} < -\hat{v} \} < \hat{\Delta}.$$

Each of the Bridges to follow make use of the values of $\hat{n}$ and $\hat{\Delta}$. So, although "probability products" totalling $\hat{\Delta}$ were discarded in the convolutions, $\hat{\Delta}$ itself is not being discarded permanently.

If $\hat{\Delta}$ is larger than we wish, we can restart the calculations using a smaller value of ε in performing the numerical convolutions.[5]

---

[5]See (*) in APPENDIX #2 for definition of ε here.

This method involves setting the initial surplus u_0 equal to 0, using Monte Carlo to generate trial sequences of outcomes

$$W_1, X_1, W_2, X_2, \ldots, W_{\hat{n}}, X_{\hat{n}}$$

for

$$W_1, X_1, W_2, X_2, \ldots, W_{\hat{n}}, X_{\hat{n}}, \text{ respectively,}$$

and generating the number pairs (u_1, v_1) where

$$v_0 = 0$$

$$u_0 = 0$$

$$v_1 = 0 + c \cdot w_1 - x_1$$

$$u_1 = \min(v_1, u_0)$$

$$v_2 = v_1 + c \cdot w_2 - x_2$$

$$u_2 = \min(v_2, u_1)$$

$$\ldots$$

$$v_{\hat{n}} = v_{\hat{n}-1} + c \cdot w_{\hat{n}} - x_{\hat{n}}$$

$$u_{\hat{n}} = \min(v_{\hat{n}}, u_{\hat{n}-1}).$$

The value of $u_{\hat{n}}$ from each trial sequence would be recorded. Note that the $u_{\hat{n}}$'s are nonpositive.

Let n be the number of trial sequences generated. The distributuion implied by attaching 1/n to the negative of each $u_{\hat{n}}$

would be the approximate distribution of maximal losses, say $\hat{L}$, during the period of the first $\hat{n}$ claims. This distribution of $\hat{L}$ can be used to obtain the value of $\hat{\Psi}(u;\hat{n})$ using the relationship

$$\hat{\Psi}(u;\hat{n}) = 1 - Pr(\hat{L} \leq u; \hat{n}).$$

There is some degree of error due to using a finite number of Monte Carlo trial sequences, and we would like to know how close $\hat{\Psi}(u;\hat{n})$ is likely to be to the accurate probability $\Psi(u;\hat{n})$. We can obtain a k% confidence interval[6] for $\Psi(u;\hat{n})$ using the sample value $\hat{\Psi}(u;\hat{n})$, the number n of trial sequences, and the fact that we are dealing with Binomial random variables (i.e. for any given u we have $-u_{\hat{n}} \leq u$ or $-u_{\hat{n}} > u$). Let $\underline{p(u)}$ and $\overline{p(u)}$ be the positive real numbers such that ($\underline{p(u)}$, $\overline{p(u)}$) is the k% confidence interval for $\Psi(u;\hat{n})$.

Let $\underline{\Psi(u)} = \underline{p(u)}$.

---

[6]From ref[1] a $1-1/t^2$ confidence interval $[p_l, p_u]$ for a binomial parameter p such that

$$Pr(\, p_l \leq p \leq p_u \,) \geq 1-1/t^2$$

is given by

$$p_l = \bar{X}/(1+t^2/n) + (t^2/n - (4t^2/n \cdot \bar{X} \cdot (1-\bar{X}) + t^4/n^2)^{.5})/(2(1+t^2/n))$$

$$p_u = \bar{X}/(1+t^2/n) + (t^2/n + (4t^2/n \cdot \bar{X} \cdot (1-\bar{X}) + t^4/n^2)^{.5})/(2(1+t^2/n)),$$

where $\bar{X}$ is a sample mean for a sample of size n.

We must also consider the possibility that survivors through claim number $\hat{n}$ may eventually become ruined, even though the surplus just after claim number $\hat{n}$ is greater than $-\hat{v}$; and, finally, there are the "probability products" totalling $\hat{\Delta}$ which we discarded in determining $\hat{v}$ and $\hat{n}$. See section "Determining $\hat{v}$, $\hat{n}$ and $\hat{\Delta}$."

$$\begin{aligned}\text{Let } \overline{\Psi(u)} = \overline{p(u)} &+ \hat{\Delta} \\ &+ (1-\underline{\Psi(u)}) \cdot \hat{\Delta} \\ &+ (1-\underline{\Psi(u)})^2 \cdot \hat{\Delta} \\ &+ (1-\underline{\Psi(u)})^3 \cdot \hat{\Delta} \\ &+ \ldots \\ = \overline{p(u)} &+ \hat{\Delta}/(\underline{\Psi(u)}),\end{aligned}$$

collapsing a geometric series. Then, at least k% of the time we can expect $\Psi(u)$ to lie in the interval $(\underline{\Psi(u)} , \overline{\Psi(u)})$.

The fact that $\overline{\Psi(u)}$ is an upper bound (with at least k% confidence) for $\Psi(u)$ follows from a consideration of successive periods of $\hat{n}$ claims: for each such period we consider the product of (1) an upper bound (with at least k% confidence) for the probability of <u>not</u> going ruined before that period, and (2) an upper bound for the probability of going ruined during a period of $\hat{n}$ claims given that the surplus at the beginning of that period was at least $-\hat{v}$.

MONTE CARLO BRIDGE #2

This method involves setting the initial surplus u_0 equal to u, using Monte Carlo to generate trial sequences of outcomes

$$W_1, X_1, W_2, X_2, \ldots, W_{\hat{n}}, X_{\hat{n}}$$

for

$$W_1, X_1, W_2, X_2, \ldots, W_{\hat{n}}, X_{\hat{n}}, \text{ respectively,}$$

amd generating trial sequences of surplus values

$$v_0 = u$$

$$v_1 = v_0 + c \cdot w_1 - x_1$$

$$v_2 = v_1 + c \cdot w_2 - x_2$$

$$\ldots$$

$$v_{\hat{n}} = v_{\hat{n}-1} + c \cdot w_{\hat{n}} - x_{\hat{n}}.$$

Rather than each trial sequence consisting of exactly $\hat{n}$ terms, we truncate each trial sequence as soon as $v_i < 0$ for some i, recording the number,[7] say m, of trial sequences that terminate with $v_i < 0$ and the total number, say n, of trial sequences. m/n is a first estimate of $\Psi(u)$.

But how close in m/n likely to be to the exact value of $\Psi(u)$. First, there is the error due to using Monte Carlo.

---

[7]If we intend to use M.C.B.#3, we also record the absolute value of each v_i, calling it r_i. The empirical distribution of, say R, will then be available.

Second, there is the possibility that a survivor through claim number $\hat{n}$ will eventually go ruined.

And, thirdly, there are the "probability products" totalling, say, $\Delta_{\hat{n}}$, which we discarded in determining $\hat{v}$ and $\hat{n}$. We can obtain a k% confidence interval for $\Psi(u;\hat{n})$ using the sample value m/n, the number n of trial sequences, and the fact that we are dealing with a binomial distribution, i.e. for any given u a trial sequence either goes ruined or it does not. Let

$$(\underline{p(u)} , \overline{p(u)})$$

be that confidence interval for $\Psi(u;\hat{n})$. Letting $\underline{\Psi(u)} = \underline{p(u)}$ we are assured that at least k% of the time $\Psi(u)$ will be greater than $\underline{\Psi(u)}$.

To determine a value $\overline{\Psi(u)}$ such that $\Psi(u)$ can be expected to be less than $\overline{\Psi(u)}$ at least k% of the time, we let

$$\begin{aligned}\overline{\Psi(u)} &= \overline{p(u)} + \hat{\Delta} \\ &\quad + (1-\underline{\Psi(u)})\cdot\hat{\Delta} \\ &\quad + (1-\underline{\Psi(u)})^2\cdot\hat{\Delta} \\ &\quad + (1-\underline{\Psi(u)})^3\cdot\hat{\Delta} \\ &\quad + \ldots \\ &= \overline{p(u)} + \hat{\Delta}/\underline{\Psi(u)},\end{aligned}$$

by collapsing a geometric series.

$\Psi(u) \leq \overline{\Psi(u)}$ at least k% of the time, by the reasoning described in the last paragraph of MONTE CARLO BRIDGE #1.

Then at least k% of the time we can expect $\Psi(u)$ to lie in the interval $(\underline{\Psi(u)}\ ,\ \overline{\Psi(u)})$.

MONTE CARLO BRIDGE #3

This method involves setting the initial surplus u equal to zero, and using MONTE CARLO BRIDGE #2 to determine $\underline{\Psi(0)}$ and $\overline{\Psi(0)}$. We then construct compound geometric distributions using

for the parameter q: either $\underline{\Psi(0)}$ or $\overline{\Psi(0)}$ and

for the severity distribution:

the conditional distribution of R, recorded as we determined m/n in M.C.B.#2.

Each of these compound geometric distributions can be generated as follows:[8]

1-q is an estimate of the probability of never having a

[8]See Ref[4] for other ways of generating these and some other compound distributions.

surplus less than zero;[9] thus, this probability is associated with an ultimate maximal loss of 0.

$q \cdot (1-q)$ is an estimate of the probability[10] of having only one negative surplus, that is, a first negative surplus, followed by no additional negative surpluses thereafter; thus, this probability is multiplied into the probabilities in the distribution of R;

convolute the distribution of R with itself to obtain the distribution of R + R; $q^2 \cdot (1-q)$ is an estimate of the probability[11] of having a negative surplus, followed by an additional negative surplus, followed by no additional negative surpluses thereafter; thus, this probability is multiplied into the probabilities in the distribution of R + R;

recursively for each positive integer i (i = 2,3,...): convolute the distribution of R with the distribution of $\underbrace{R+R+\ldots+R}_{i \text{ terms}}$ to obtain the distribution of $\underbrace{R+R+\ldots+R}_{i+1 \text{ terms}}$; $q^{i+1} \cdot (1-q)$ is an estimate of the probability[12] of having a

---

[9]having started with an initial surplus of u=0.

[10]Ibid

[11]Ibid

[12]Ibid

negative surplus , followed by i additional negative surpluses, followed by no additional negative surpluses thereafter; thus this probability is multiplied into the probabilities in the distribution of $\underbrace{R+R+\ldots+R}_{i+1 \text{ terms}}$;

continue until the sum of all of the resulting probabilities is within say Δ,[13] of unity; the concatenation of the resulting partial distributions is, except for the discarded probabilities totalling Δ, the desired compound geometric distribution.

Letting $\underline{L}$ and $\overline{L}$ be the random variables associated with the compound geometric distributions generated using for q: $\underline{\Psi(0)}$ and $\overline{\Psi(0)}$, respectively, the probability of ultimate ruin $\Psi(u)$ can be expected at least k% of the time to lie in the interval

$$(1 - Pr(\overline{L} \leq u) - \Delta , 1 - Pr(\underline{L} \leq u + \Delta).$$

$\underline{L}$ and $\overline{L}$ are called maximal loss random variables.

---

[13] Δ is the total of the probabilities discarded in truncating the series $q^i \cdot (1-q)$ $(i=0,1,2,\ldots)$ and the "probability products" discarded in performing the convolutions to obtain the distributions of R, R+R, R+R+R, etc.

We cannot be as definite about the intervals containing $\Psi(u)$ for the CONVOLUTION BRIDGES as we were for the MONTE CARLO BRIDGES. This is because the performing of numerical convolutions (regular or generalized) involves selecting and using mesh intervals or mesh rectangles. The results are "almost exact" (in the sense that finer meshes will tend to produce more accurate results), but nevertheless approximate. This should be born in mind when reading about the CONVOLUTION BRIDGES because the description does not emphasize the less rigorous nature of the results.

CONVOLUTION BRIDGE #1

Let the initial surplus be u=0.

Let $H = c \cdot W - X$, and generate the distribution of H by convoluting the distribution of $c \cdot W$ with the distribution of $-X$.[14]

Let $H_1, H_2, \ldots, H_n$ be mutually independent and identically distributed as H.

[14]See APPENDIX #2 for an algorithm to perform univariate convolutions

Let $V_1 = H_1$ and $U_1 = \min(H_1 , 0)$. V_1 is the surplus just after the first claim. U_1 is the smaller of the initial surplus of 0 and the surplus just after the 1st claim.

Let $V_2 = V_1 + H_2$ and $U_2 = \min(U_1, V_1 + H_2)$. V_2 is the surplus just after the 2nd claim. U_2 is the lower of the surplus just after the 1st claim and the surplus just after the 2nd claim; that is, U_2 is the lowest surplus experienced at any time during the period of the first 2 claims. $-U_2$ is called the maximal loss during the period of the first 2 claims.

Etc.

Let $V_{\hat{n}} = V_{\hat{n}-1} + H_{\hat{n}}$ and $U_{\hat{n}} = \min(U_{\hat{n}-1}, V_{\hat{n}-1} + H_{\hat{n}})$. $V_{\hat{n}}$ is the surplus just after the $\hat{n}^{th}$ claim. $U_{\hat{n}}$ is the lowest of the surplus experienced at any time during the period of the first $\hat{n}$ claims. $-U_{\hat{n}}$ is called the maximal loss during the period of the first $\hat{n}$ claims.

Consider the bivariate random variables (U_1,V_1), (U_2,V_2), ... , $(U_{\hat{n}},V_{\hat{n}})$. The distribution of (U_1,V_1) is immediately available from the distribution of H_1, since

$$V_1 = H_1 \quad \text{and} \quad U_1 = \min(H_1 , 0).$$

For any n ($2 \le n \le \hat{n}$) the distribution of (U_n, V_n) can be obtained by "convoluting"[15] the distribution of (U_{n-1}, V_{n-1}) and the distribution of H_{n-1}, using the formulas

$$V_n = V_{n-1} + H_n$$

and

$$U_n = \min(U_{n-1}, V_{n-1} + H_n).$$

Thus we can generate the distribution of $(U_{\hat{n}}, V_{\hat{n}})$ recursively.

Alternatively, we could generate the distribution of $(U_{\hat{n}}, V_{\hat{n}})$ recursively as follows:

For any n ($2 \le n \le \hat{n}$) which is a power of 2 the distribution of (U_n, V_n) can be obtained by "convoluting"[16] the distribution of $(U_{n/2}, V_{n/2})$ and the distribution of $(U'_{n/2}, V'_{n/2})$, using the formulas

$$V_n = V_{n/2} + V'_{n/2}$$

and

---

[15]See APPENDIX #3 for an algorithm to perform these generalized bivariate convolutions. It may be helpful to think of H_i as a bivariate random variable $(H_i, 0)$. The word convoluting is shown in quotation marks to indicate that a generalized operation is involved and not merely a convolution for sums.

[16]The word convoluting is shown in quotation marks to indicate that a generalized operation is involved and not merely a convolution for sums. See APPENDIX #3 for an algorithm to perform these generalized bivariate convolutions.

$$U_n = \min\{ U_{n/2},\ V_{n/2} + U'_{n/2} \}$$

where $(U'_{n/2}, V'_{n/2})$ is a bivariate random variable independently and identically distributed as $(U_{n/2}, V_{n/2})$. If $\hat{n}$ is not a power of 2, the procedure is easily modified to "convolute" appropriate nonidentical distributions to generate the distribution of the desired bivariate random variable $(U_{\hat{n}}, V_{\hat{n}})$.

Summing the probabilities from the (partial) distribution of $(U_{\hat{n}}, V_{\hat{n}})$ we would have unity, except for the "probability products"[17], which have been discarded in performing the bivariate "convolutions". Let $\hat{\Delta}_{\hat{n}}$ be this deficiency in the probabilities.

The marginal distribution of the absolute values $|U_{\hat{n}}|$ is the estimated distribution of maximal losses, say $\tilde{L}$, during the period of the first $\hat{n}$ claims, ignoring the set of outcomes with probability less than $\hat{\Delta}_{\hat{n}}$. The distribution of $\tilde{L}$ can be used to determine the distribution of $\Psi(u;\hat{n})$, using the formula

$$\Psi(u;\hat{n}) = 1 - \Pr\{ \tilde{L} \le u;\ \hat{n} \}.$$

Let $\underline{\Psi(u)} = \hat{\Psi}(u;\hat{n})$.

---

[17]"probability products" for bivariate "convolutions" are defined at the (**) in APPENDIX #3.

Let $\overline{\Psi(u)} = \hat{\Psi}(u;\hat{n}) + \hat{\Delta}_{\hat{n}}$

$$+\left(1 - \hat{\Psi}(u;\hat{n})\right)\cdot\hat{\Delta}$$

$$+\left(1 - \hat{\Psi}(u;\hat{n})\right)^2\cdot\hat{\Delta}$$

$$+\left(1 - \hat{\Psi}(u;\hat{n})\right)^3\cdot\hat{\Delta}$$

$$+\ldots$$

$$= \hat{\Psi}(u;\hat{n}) + \hat{\Delta}_{\hat{n}} + \hat{\Delta}\cdot\left(1/\hat{\Psi}(u;\hat{n}) - 1\right)$$

$\underline{\Psi(u)}$ and $\overline{\Psi(u)}$ are lower and higher estimates, respectively, for $\Psi(u)$. We can use the nonrigorous intervals $\left(\underline{\Psi(u)}\ ,\ \overline{\Psi(u)}\right)$ to decide how far to round our estimate of $\Psi(u)$.

Defining $\overline{\Psi(u)}$ as we have is motivated by a consideration of successive periods of $\hat{n}$ claims. For each such period we consider the product of (1) an approximate upper bound for the probability of not going ruined by the end of that period, and (2) an approximate upper bound for the probability of going ruined during a period of $\hat{n}$ claims given that the surplus at the beginning of that period was at least $-\hat{v}$.

This method is analogous to Monte Carlo Bridge #2, except that here we use univariate numerical convolutions.

Let $H = c \cdot W - X$ and generate the distribution of H by convoluting the distribution of $c \cdot W$ with the distribution of $-X$. Let H_1, H_2, ... , $H_{\hat{n}}$ be mutually independent and identically distributed as H. Let $V_1 = u + H_1$ and obtain the distribution of V_1 by adding the initial surplus u to the amounts (not the probabilities) in the distribution of H_1.

For each positive integer i ($2 \leq i \leq \hat{n}$):

let S_i be the random variable which assumes the zero and positive values of V_i with the associated probabilities. The values of S_i can be considered to be outcomes for survivors just after the i^{th} claim. The (partial) distribution of S_i is thus part of the distributon of V_i. Let R_i be the random variable which assumes the absolute value of the negative values of V_i with the associated probabilites. The values of R_i can be considered to be outcomes for ruineds which occurred with the i^{th} claim. The

(partial) distribution of R_i is thus part of the distribution of V_i. Let $V_{i+1} = S_i + H_{i+1}$ and obtain the (partial) distribution of V_{i+1} by convoluting the (partial) distribution of S_i with the distribution of H_i.[18]

Summing the probabilities from the (partial) distributions of $R_1, R_2, \ldots, R_{\hat{n}}$ and $S_{\hat{n}}$ we would have unity, except for the "probability products"[19] which have been discarded in performing the numerical convolutions. Let $\Delta_{\hat{n}}(u)$ be this deficiency in the probabilities, where the initial surplus was u.

The probability of an outcome being a survivor just after $\hat{n}$ claims is the sum of the probabilities in the distribution of $S_{\hat{n}}$. The probability of an outcome being ruined anytime prior to $\hat{n}+1$ claims is the sum of the probabilities in the (partial) distributions

$$R_1, R_2, \ldots, R_{\hat{n}}.$$

And there are the probabilities (totalling $\hat{\Delta}$) which were discarded in determining $\hat{n}$ and $\hat{v}$.

Let $\underline{\Psi(u)} = \sum_{i=1}^{\hat{n}} Pr(R_i)$. We must consider the possibility that the survivors after $\hat{n}$ claims eventually become ruineds.

---

[18]See APPENDIX #2 for an algorithm to perform univariate convolutions

[19]See (*) in APPENDIX #2 for definition of "probability products" which are being discarded in performing univariate convolutions.

Let $\overline{\Psi(u)} = \underline{\Psi(u)} + \Delta_{\hat{n}}(u)$

$$+ (1-\underline{\Psi(u)}) \cdot \hat{\Delta}$$

$$+ (1-\underline{\Psi(u)})^2 \cdot \hat{\Delta}$$

$$+ (1-\underline{\Psi(u)})^3 \cdot \hat{\Delta}$$

$$+ \ldots$$

$$= \underline{\Psi(u)} + \Delta_{\hat{n}}(u) + \hat{\Delta} \cdot (1/ \underline{\Psi(u)} - 1).$$

$\underline{\Psi(u)}$ and $\overline{\Psi(u)}$ are lower and higher estimates, respectively, for $\Psi(u)$. We can use the nonrigorous intervals ($\underline{\Psi(u)}$, $\overline{\Psi(u)}$) to decide how far to round our estimate of $\Psi(u)$. The motivation for defining $\overline{\Psi(u)}$ as we have here is analogous to that described in the last paragraph of CONVOLUTION BRIDGE #1.

CONVOLUTION BRIDGE #3

This method involves using CONVOLUTION BRIDGE #2 to generate a lower estimate $\underline{\Psi(0)}$ and a higher estimate $\overline{\Psi(0)}$ of $\Psi(0)$. We then construct a compound geometric distribtuion using

for the parameter q: either $\underline{\Psi(0)}$ or $\overline{\Psi(0)}$, and

for the severity distribution: the conditional distribution of the amounts in the severity distribution obtained by concatenating the C.B.#2 (partial) distributions of $R_1, R_2, \ldots, R_n$ and normalizing the probabilities to sum to unity.

For details on how to generate these compound geometric distributions using 1-dimensional numerical convolutions, see the relevant part of MONTE CARLO BRIDGE #3.

Letting $\underline{L}$ and $\overline{L}$ be the random variables associated with the compound geometric distributions generated using for q: $\underline{\Psi(0)}$ and $\overline{\Psi(0)}$, respectively,

$$1 - \Pr(\overline{L} \leq u) - \Delta$$

and

$$1 - \Pr(\underline{L} \leq u) + \Delta$$

would be higher and lower estimates for $\Psi(u)$, respectively, where Δ is as defined in a footnote to MONTE CARLO BRIDGE #3. We can use the nonrigorous intervals ($\underline{\Psi(u)}$, $\overline{\Psi(u)}$) to decide how far to round our estimate of $\Psi(u)$.

COMPARISON OF BRIDGES

Whether you will ultimately prefer to use a MONTE CARLO BRIDGE or a CONVOLUTION BRIDGE may be difficult to determine in advance.

The MONTE CARLO BRIDGES use random numbers and have the advantage of producing rigorous confidence intervals for the probability $\Psi(u)$ of ultimate ruin, but have the disadvantage in certain situations of requiring the generation of more pseudo-random numbers than can be done in a reasonable time.

The CONVOLUTION BRIDGES use regular or generalized univariate or bivariate "convolutions" and do not use random numbers, but have the disadvantage of not producing rigorous confidence intervals for the probabiliy $\Psi(u)$ of ultimate ruin. The bivariate generalized "convolutions" are not quick even on the fastest of personal computers, and are not likely to produce as fine results as univariate convolutions.

MONTE CARLO BRIDGES #1 and #3, and CONVOLUTION BRIDGES #1 and #3, have the advantage of producing distributions from which $\Psi(u)$ is available for all values of u at once.

A NUMERICAL EXAMPLE

Assume that the size X of individual claims is distributed according to probability density function

$$p(x) = 12 \cdot (e^{-3x} - e^{-4x}), \quad x>0;$$

and assume that the waiting time W between claims is distributed according to probability density function

$$q(w) = e^{-w}, \quad w>0;$$

and let the rate c (at which the premiums are received) be 1.

Find the probability $\psi(u)$ of ultimate ruin, given an initial surplus of u.

Solution

In practice the random variable X and the random variable W are likely to be chosen as discrete empirical distributions based on relevant experience. However, an advantage of solving this particular numerical example (which happens to involve continuous distributions) is that Reference [1] provides exact (and various approximate) values of $\psi(u)$ for u=.0, .5, 1.0, 1.5, ... , 10.0. So we can compare the results of our six methods with those of the three methods used in Reference [1].

Using the definition of p(x) and integrating, we have for any fixed positive integer n:

$$(\#) \qquad \int_{1/n}^{(i+1)/n} p(x)dx = \left(-4 \cdot e^{-3 \cdot (i+1)/n} + 3 \cdot e^{-4 \cdot (i+1)/n} \right) - \left(-4 \cdot e^{-3 \cdot 1/n} + 3 \cdot e^{-4 \cdot 1/n} \right), (i=0,1,2,\ldots);$$

and

$$(\#\#) \qquad \int_{1/n}^{(i+1)/n} x \cdot p(x)dx =$$

$$12\cdot\Big(\ -(i+1)/n\cdot 1/3\cdot e^{-3\cdot(i+1)/n}\ +\ i/n\cdot 1/3\cdot e^{-3\cdot i/n}\ +$$
$$1/9\cdot e^{-3\cdot(i+1)/n}\ +\ 1/9\cdot e^{-3\cdot i/n}\ +$$
$$+(i+1)/n\cdot 1/4\cdot e^{-4\cdot(i+1)/n}\ -\ i/n\cdot 1/4\cdot e^{-4\cdot i/n}\ +$$
$$+1/16\cdot e^{-4\cdot(i+1)/n}\ -\ 1/16\cdot e^{-4\cdot i/n}\ \Big)$$

$$,(i=0,1,2,\ldots).$$

Select a small positive real number, say ε. By simple trial and error we determine a positive integer , say n_0, such that the inequality

$$1\ -\ \sum_{i=0}^{r_0-1}\ \int_{1/n_0}^{(i+1)/n_0} p(x)\,dx\ <\ \varepsilon$$

is satisfied, where r_0 is a reasonable number of rows to use in the discretized version of the distribution of X.

In our solution here we decided to use an r_0 of about 4000 and an $\varepsilon=10^{-15}$. This led to our choosing $n_0=360$, for which r_0 turned out to be 4311. So, we used the following matrix as the discretized version of the distribution of X:

$$\left[\ \frac{\displaystyle\int_{1/n_0}^{(i+1)/n_0} x\cdot p(x)\,dx}{\displaystyle\int_{1/n_0}^{(i+1)/n_0} p(x)\,dx}\ ,\ \int_{1/n_0}^{(i+1)/n_0} p(x)\,dx\ \right]_{i=0,4310;\,n_0=360}$$

where the values of the integrals are immediately available from formulas (#) and (##) above.

A similar procedure led to our using the following matrix as the discretized version of the distribution of W:

$$\left[\frac{\int_{i/n_0}^{(i+1)/n_0} w\cdot q(w)\,dw}{\int_{i/n_0}^{(i+1)/n_0} q(w)\,dw}\ ,\ \int_{i/n_0}^{(i+1)/n_0} q(w)\,dw\right]_{i=0,4144;n_0=120}$$

where the values of the integrals are immediately available from the following formulas:

$$\int_{i/n}^{(i+1)/n} q(w)\,dw =$$

$$-e^{-(i+1)/n} + e^{-i/n}\ ,(i=0,1,2,\ldots);$$

and

$$\int_{i/n}^{(i+1)/n} w\cdot q(w)\,dw =$$

$$-(i+1)/n\cdot e^{-(i+1)/n} + i/n\cdot e^{-i/n} +$$
$$-e^{-(i+1)/n} + e^{-i/n}\ ,(i=0,1,2,\ldots);$$

MONTE CARLO BRIDGE #1

See below!

MONTE CARLO BRIDGE #2

Choosing parameters nax=4000, mesh=1 and epsilon=10^{-15}, $\hat{n}$ and $\hat{\Delta}$ turned out to be 387 and $.0^{5}128746$, respectively. Then, choosing $n=10^{6}$, $\underline{\psi(0)}$ and $\overline{\Psi(0)}$ turned out to be .582671 and .585213, respectively. The exact answer given in ref[1] is .583333.

MONTE CARLO BRIDGE #3

To generate the appropriate compound geometric distribution we used

for the severity distribution:
the conditional distribution of the absolute value of the losses at time of ruin captured as a byproduct of the MCB#2 run;

for the geometric distributions:

$q=\underline{\psi(0)}=.582671$ and

$q=\overline{\psi(0)}=.585213$;

nax=4000, mesh=1 and epsilon=10^{-15};

The geometric distribution involved enough rounding that the probabilities totaled just slightly different than unity; we normalized the probabilities to sum to unity;

The compound geometric distribution discarded probabilities totaling on the order of 10^{-10}; and we simply ignored this fact.

The following table shows the resulting $1-Pr(\bar{L} \le u; \hat{n})$ and $1-Pr(\underline{L} \le u; \hat{n})$ for each of the indicated values of u; and, for comparison, shows the corresponding exact answers to 6 decimal places from ref[1].

MCB#3

u	$Pr(\bar{L} \le u)$	$1-Pr(\bar{L} \le u)$	exact	$1-Pr(\underline{L} \le u)$	$Pr(\underline{L} \le u)$
0.0	.417329	.582671	.583333	.585213	.414787
0.5	.625097	.374903	.375661	.377637	.622363
1.0	.771621	.228379	.229644	.230781	.769219
1.5	.861310	.138690	.139433	.140571	.859429
2.0	.915977	.084023	.084583	.085475	.914525
2.5	.949118	.050882	.051303	.051914	.948086
3.0	.969159	.030841	.031117	.031573	.968427
3.5	.981316	.018684	.018873	.019183	.980817
4.0	.988688	.011312	.011447	.011658	.988342
4.5	.993146	.006854	.006943	.007086	.992914
5.0	.995847	.004153	.004211	.004307	.995693
5.5	.997485	.002515	.002554	.002617	.997383
6.0	.998477	.001523	.001549	.001590	.998410
6.5	.999077	.000923	.000940	.000967	.999033
7.0	.999441	.000559	.000570	.000587	.999413
7.5	.999661	.000339	.000346	.000357	.999643
8.0	.999795	.000205	.000210	.000217	.999783
8.5	.999876	.000124	.000127	.000132	.999868
9.0	.999925	.000075	.000077	.000080	.999920
9.5	.999954	.000046	.000047	.000049	.999951
10.0	.999972	.000028	.000028	.000030	.999970

CONVOLUTION BRIDGE #1

See below!

CONVOLUTION BRIDGE #2

Choosing parameters nax=4000, mesh=1 and epsilon=10^{-15}, $\hat{n}$ turned out to be 387. Again using nax=4000, mesh=1 and epsilon=10^{-15} $\underline{\psi(0)}$ turned out to be .583239; and,

$\overline{\Psi(0)}$=.583239+.$0^5$128746/.583239=.583241. That is, there was almost no difference between $\underline{\psi(0)}$ and $\overline{\Psi(0)}$. The exact answer given in ref[1] is .583333.

CONVOLUTION BRIDGE #3

To generate the appropriate compound geometric distribution we used

for the severity distribution:
the conditional distribution of the absolute value of the losses at time of ruin captured as a byproduct of the CB#5 run;

for the geometric distribution:
q=.583239;
nax=4000, mesh=1 and epsilon=10^{-15};

The geometric distribution involved enough rounding that the probabilities totaled just slightly different than unity; we normalized the probabilities to sum to unity.

The resulting compound geometric distribution discarded probabilities totaling on the order of 10^{-10}; and we simply

ignored this fact.

The following table shows the resulting $\psi(u)$ for each of the indicated values of u; and, for comparison, shows the corresponding exact answers to 6 decimal places from ref[1].

CB#3

u	$\Psi(u)$†	exact
0.0	.583239	.583333
0.5	.375688	.375661
1.0	.229678	.229644
1.5	.139398	.139433
2.0	.084590	.084583
2.5	.051308	.051303
3.0	.031116	.031117
3.5	.018874	.018873
4.0	.011449	.011447
4.5	.006943	.006943
5.0	.004210	.004211
5.5	.002554	.002554
6.0	.001549	.001549
6.5	.000939	.000940
7.0	.000570	.000570
7.5	.000346	.000346
8.0	.000210	.000210
8.5	.000127	.000127
9.0	.000077	.000077
9.5	.000047	.000047
10.0	.000028	.000028

†Rounding both $\Psi(u)$ and $\overline{\Psi(u)}$ to the number of decimal places shown in the column labelled $\Psi(u)$, we exhibit a single value.

Looking ahead to having generated the 2-dimensional random sequences dictated by this method, consider the formula for $\overline{\Psi(10)}$ for example. If $\hat{\Delta}=1.28\cdot 10^{-6}$ and $\underline{\Psi(10)} \doteq .000028$, then since the ratio of these two figures is about .046, $\overline{\Psi(10)}$ would turn out to be much too large for practical purposes. Therefore, if we are interested in determining a practical interval within which $\Psi(10)$ will fall with at least 95% confidence, we will have to recompute $-\hat{v}$ and $\hat{n}$ starting with a smaller epsilon. By so doing we can generate a smaller value of $\hat{\Delta}$.

Doing the recomputation starting with epsilon = 10^{-21}, we find that $\hat{v}=-41.92$, $\hat{n}=768$ and $\hat{\Delta}=.0^{9}16678$; in fact, $\hat{n}$ was smaller than 768, but we did not need to take the time to determine it more precisely. Once again we choose $n=10^{6}$.

The following table shows the resulting $\underline{\Psi(u)}$ and $\overline{\Psi(u)}$ for each of the indicated values of u; and, for comparison, shows the corresponding exact answers to 6 decimal places from ref[1].

MCB#1

u	Pr($\hat{L}\leq u$)	1-Pr($\hat{L}\leq u$)	$\underline{\Psi(u)}$	exact	$\overline{\Psi(u)}$
0.0	.416741	.583259	.576271	.583333	.590214
0.5	.625398	.374602	.367783	.375661	.381471
1.0	.772625	.227375	.221503	.229644	.233356
1.5	.862786	.137214	.132421	.139433	.142152
2.0	.917755	.082245	.078443	.084583	.086214
2.5	.950694	.049306	.046333	.051303	.052459
3.0	.970730	.029270	.026979	.031117	.031750
3.5	.982486	.017514	.015753	.018873	.019468
4.0	.989544	.010456	.009112	.011447	.011996
4.5	.993864	.006136	.005126	.006943	.007343
5.0	.996359	.003641	.002883	.004211	.004598
5.5	.997850	.002150	.001587	.002554	.002912
6.0	.998737	.001263	.000851	.001549	.001875
6.5	.999265	.000735	.000439	.000940	.001231
7.0	.999556	.000444	.000230	.000570	.000859
7.5	.999730	.000270	.000117	.000346	.000624

CONVOLUTION BRIDGE #1

Using the same reasoning as described under MONTE CARLO BRIDGE #1 immediately above, we again use $\hat{n}$=768 and $\hat{\Delta}$=$.0^{9}16678$. For the bivariate generalized numerical convolutions we use nax=64, nay=320 and epsilon=10^{15}. $\hat{\Delta}_{768}$ turned out to be $.0^{5}19475$.

C.B.#1 takes the most computer time relative to the accuracy obtained in the results; for that reason we obtained the numerical results for C.B.#1 by running an IBM RS 6000/530 at

PolySystems, Inc.; the numerical results for each of the other Bridges were obtained by running a Gateway 2000 (486/25mh).

CB#1

u	$\Psi(u)$†	exact
0.0	.58473	.58521
0.5	.37771	.37764
1.0	.23212	.22964
1.5	.14319	.13943
2.0	.08590	.08458
2.5	.05292	.05130
3.0	.03153	.03112
3.5	.01951	.01887
4.0	.01154	.01145
4.5	.00720	.00694
5.0	.00426	.00421
5.5	.00265	.00255
6.0	.00158	.00155
6.5	.00098	.00094
7.0	.00058	.00057
7.5	.00036	.00035
8.0	.00022	.00021
8.5	.00013	.00013
9.0	.00008	.00008
9.5	.00005	.00005
10.0	.00003	.00003

---

†Rounding both $\Psi(u)$ and $\overline{\Psi(u)}$ to the number of decimal places shown in the column labelled $\Psi(u)$, we exhibit a single value.

APPENDIX #1-ASSERTIONS

We will prove three assertions, from which we can conclude that given an initial surplus of, say u, and an $\varepsilon > 0$ there exist a negative real number $\hat{v}$ and a (positive) number $\hat{n}$ of claims such that

(a) the probability of a maximal loss of more than $\hat{v}$ during the period of the first $\hat{n}$ claims is less than ε;

and

(b) the probability of a surplus of less than $-\hat{v}$ just after claim number $\hat{n}$ is less than ε.

Assertion #1:

Given an $\varepsilon_1 > 0$ there exists a positive integer $\tilde{n}$ such that

$$\Pr\left\{ V_n(0) \leq 0 \right\} < \varepsilon_1 \text{ for all } n \geq \tilde{n}.$$

Assertion #2

Given an $\varepsilon_2 > 0$ there exists a negative real number $\hat{v}$ such that

$$\Pr\left\{ V_n(0) \le \hat{v} \right\} < \varepsilon_2 \text{ for all } n \text{ such that } 0 < n < \tilde{n}.$$

Assertion #3

Given an $\varepsilon_3 > 0$ there exists a positive integer $\hat{n}$ such that

$$\Pr\left\{ V_n(0) \le -\hat{v} \right\} < \varepsilon_3 \text{ for all } n > \hat{n}.$$

PROOF OF ASSERTION #1

From probability theory we know that

$$E[V_n(0)] = 0 + n \cdot E[H]$$

and

$$Var[V_n(0)]] = n \cdot Var[H]$$

and Chebyshev's Inequality† assures that

$$\Pr\left\{ \left| V_n(0) - E[V_n(0)] \right| \ge E[V_n(0)] \right\} \le \frac{Var[V_n(0)]}{E[V_n(0)]^2}$$

---

†See Ref[3].

by choosing t in the Chebyshev Inequality to be $E[V_n(0)]$.†

Substituting for $E[V_n(0)]$ and $Var[V_n(0)]$ in this statement we have

$$Pr\left\{ \mid V_n(0) - (0 + n\cdot E[H]) \mid \geq 0 + n\cdot E[H] \right\} \leq \frac{n\cdot Var[H]}{(0 + n\cdot E[H])^2}$$

Now, by L'Hopital's Rule

$$\lim_{n\to\infty} \frac{n\cdot Var[H]}{(0+n\cdot E[H])^2} = \lim_{n\to\infty} \frac{Var[H]}{2\cdot(0+n\cdot E[H])\cdot E[H]} = 0$$

So, given $\varepsilon_1 > 0$ there exists a positive integer $\tilde{n}$ such that

$$\frac{n\cdot Var[H]}{(0+n\cdot E[H])^2} < \varepsilon_1 \quad \text{for all } n > \tilde{n}.$$

Thus

$$Pr\left\{ \mid V_n(0) - (0 + n\cdot E[H]) \mid \geq 0 + n\cdot E[H] \right\} \leq \varepsilon_1 \text{ for all } n>\tilde{n}.$$

If $V_n(0) \leq 0+ n\cdot E[H]$, then

$$\mid V_n(0) - (0 + n\cdot E[H]) \mid = 0 + n\cdot E[H] - V_n(0)$$

---

†We are assuming that the relative security loading ϑ (i.e. $E[W]/E[X] - 1$) is positive, so t is positive as required by Chebyshev's Inequality:

$$Pr\left\{ \mid X - E[X] \mid \geq t \right\} \leq \frac{Var[X]}{t^2}.$$

and

$$\Pr\left\{ 0 + n\cdot E[H] - V_n(u) \geq 0 + n\cdot E[H] \right\} \leq \varepsilon_1 \text{ for all } n>\tilde{n}.$$

Therefore,

$$\Pr\left\{ V_n(0) \leq 0 \right\} \leq \varepsilon_1 \text{ for all } n>\tilde{n}.$$

If $V_n(0) > 0 + n\cdot E[H]$ then $V_n(0) > 0$, since $0 + n\cdot E[H] > 0$; i.e. $V_n(0)$ is <u>not</u> less than or equal to zero.

<u>qed</u>

PROOF OF ASSERTION #2

If H assumes only positive values,
$\Psi(u) = 0$ for all $u > 0$.
If H assumes only negative values, then
$\Psi(u) = 1$ for all $u > 0$.
So we assume that H takes on some positive and some negative values.

If H is bounded from below by a negative quantity, say $\hat{n}$, then

$$H_1,\ H_1 + H_2,\ \ldots,\ H_1 + H_2 + \ldots + H_{\tilde{n}}$$

are bounded from below by

$$\hat{n},\ 2\cdot\hat{n},\ 3\cdot\hat{n},\ \ldots,\ \tilde{n}\cdot\hat{n}, \text{ respectively;}$$

so, $\tilde{n}\cdot\hat{n}$ is a lower bound of

$$V_1(0),\ V_2(0),\ \ldots\ ,\ V_{\tilde{n}}(0);$$

and, letting $\hat{v} = \tilde{n}\cdot\hat{h}$, we have

$$\Pr\left\{\ V_n(0) \leq \hat{v}\ \right\} = 0 < \varepsilon \text{ for any } \varepsilon > 0.$$

This would prove the assertion.

If H is unbounded from below, then we can prove the assertion as follows. Since for each positive integer n ($1 \leq n \leq \tilde{n}$) we have

$$V_n(0) = H_1 + H_2 + \ldots + H_{\tilde{n}}.$$

We know from probability theory that

$$\mu_{V_n(0)} = n\cdot\mu_H \text{ and } \sigma^2_{V_n(0)} = n\cdot\sigma^2_H$$

where

μ_H is the expected value of H,

σ^2_H is the variance of H,

$\mu_{V_n(0)}$ is the expected value of $V_n(0)$

and

$\sigma^2_{V_n(0)}$ is the variance of $V_n(0)$.

For each n ($1 \le n \le \tilde{n}$) Chebyshev's Inequality guarantees that

$$\Pr\left\{ \left| V_n(0) - n\cdot\mu_H \right| \ge t_n \right\} \le \frac{n\cdot\sigma_H^2}{t_n^2} \text{ for any } t_n > 0.$$

Given $\varepsilon_2 > 0$, for each n ($1 \le n \le \tilde{n}$) choose $t_n = \left\{ \frac{n\cdot\sigma_H^2}{\varepsilon_2} \right\}^{1/2}$, which is > 0. Then

$$\varepsilon_2 = \frac{n\cdot\sigma_H^2}{t_n^2} \text{ and } \Pr\left\{ \left| V_n(0) - n\cdot\mu_H \right| \ge t_n \right\} \le \varepsilon_2;$$

$$\text{but, } \Pr\left\{ \left| V_n(0) - n\cdot\mu_H \right| \ge t_n \right\} =$$

$$\Pr\left\{ V_n(0) - n\cdot\mu_H \ge 0 \ \&\ V_n(0) - n\cdot\mu_H \ge t_n \text{ or } V_n(0) - n\cdot\mu_H < 0 \ \&\ n\cdot\mu_H - V_n(0) \ge t_n \right\}$$
$$\le \varepsilon_2$$

and

$$\Pr\left\{ V_n(0) < n\cdot\mu_H \text{ and } V_n(0) \le n\cdot\mu_H - t_n \right\} \le$$

$$\Pr\left\{ V_n(0) - n\cdot\mu_H \ge 0 \ \&\ V_n(0) - n\cdot\mu_H \ge t_n \text{ or } V_n(0) - n\cdot\mu_H < 0 \ \&\ n\cdot\mu_H - V_n(0) \ge t_n \right\};$$

hence

$$\Pr\left\{ V_n(0) < n\cdot\mu_H \text{ and } V_n(0) \le n\cdot\mu_H - t_n \right\} \le \varepsilon_2.$$

Let $\hat{v} = \min\left\{ 1\cdot\mu_H - t_1,\ 2\cdot\mu_H - t_2,\ \ldots,\ \tilde{n}\cdot\mu_H - t_{\tilde{n}} \right\}$,

noting that

$\hat{v} < 0$, since t_n was chosen $\geq n\cdot\mu_H$ for each n $(1\leq n\leq\tilde{n})$.

Then, for each n $(1\leq n\leq\tilde{n})$ we have

$$\Pr\left\{ V_n(0) < n\cdot\mu_H \text{ and } V_n(0) \leq n\cdot\mu_H - t_n \right\} \leq \varepsilon_2.$$

But

$$\Pr\left\{ V_n(0) < n\cdot\mu_H \text{ and } V_n(0) \leq \hat{v} \right\} \leq$$

$$\Pr\left\{ V_n(0) < n\cdot\mu_H \text{ and } V_n(0) \leq n\cdot\mu_H - t_n \right\}$$

since $\hat{v} \leq n\cdot\mu_H - t_n$;

so, $\Pr\left\{ V_n(0) < n\cdot\mu_H \text{ and } V_n(0) \leq \hat{v} \right\} \leq \varepsilon_2$ for each n $(1\leq n\leq\tilde{n})$.

ged

PROOF OF ASSERTION #3

Once again by Chebyshev's Inequality we have

$$\Pr\left\{ \left| V_n(0) - (0 + n\cdot E[H]) \right| \geq 0 + n\cdot E[H] + \hat{v} \right\}$$

$$\leq \frac{n\cdot Var[H]}{(0 + n\cdot E[H] + \hat{v})^2}$$

substituting $(0 + n\cdot E[H] + \hat{v})$ for t.

Now, by L'Hopital's Rule

$$\lim_{n\to\infty} \frac{n\cdot Var[H]}{(0 + n\cdot E[H] + \hat{v})^2} =$$

$$\lim_{n\to\infty} \frac{Var[H]}{2\cdot(0 + n\cdot E[H] + \hat{v})\cdot E[H]}.$$

So, given ε_3 there exists a positive integer $\hat{n}$ such that

$$\frac{n\cdot Var[H]}{(0 + n\cdot E[H] + \hat{v})^2} < \varepsilon_3 \text{ for all } n>\hat{n}.$$

Thus,

$$\Pr\left\{ \left| V_n(0) - (0 + n\cdot E[H]) \right| \geq 0 + n\cdot E[H] + \hat{v} \right\} \leq \varepsilon_3 \text{ for all } n>\hat{n}.$$

If $V_n(0) > (0 + n\cdot E[H])$, then

$$\Pr\left\{ V_n(0) - (0+n\cdot e[h]) \geq \hat{v} \right\} \leq \varepsilon_3 \text{ for all } n>\hat{n}.$$

If $V_n(0) < (0 + n\cdot E[H] + \hat{v})$, then

$$\Pr\left\{ 0 + n\cdot E[H] - V_n(0) \geq 0 + n\cdot E[H] + \hat{v} \right\} \leq c_3 \text{ for all } n>\hat{n}.$$

qed

APPENDIX #2 - UNIVARIATE GENERALIZED NUMERICAL CONVOLUTIONS

If f_X and f_Y are independent distributions of the discrete finite univariate random variables X and Y, respectively, then the distribution $f_Z = f_{X+Y}$ of the sum Z=X+Y is the convolution $f_X + f_Y$ of f_X and f_Y for sums.†

Let f_X be expressed in element notation as

$$\begin{bmatrix} x_1 & p_1 \\ & \vdots \\ x_m & p_m \end{bmatrix}$$

which we will also express as

$$\left(x_i \quad p_i \right)_{i=1,m}$$

Similarly, let f_Y be $\left(y_j \quad q_j \right)_{j=1,n}$

Then $f_Z = f_{X+Y} = f_X + f_Y =$

---

† We are using the operation + instead of * between two distributions to indicate convolution for sums; that is, $f_X + f_Y$ instead of $f_X * f_Y$. We use the notation f_X/f_Y for the convolution of f_X and f_Y for quotients X/Y.

$$\begin{bmatrix} x_1+y_1 & p_1 \cdot q_1 \\ x_1+y_2 & p_1 \cdot q_2 \\ \vdots & \\ x_1+y_n & p_1 \cdot q_n \\ x_2+y_1 & p_2 \cdot q_1 \\ x_2+y_2 & p_2 \cdot q_2 \\ \vdots & \\ x_2+y_n & p_2 \cdot q_n \\ \vdots & \\ x_m+y_1 & p_m \cdot q_1 \\ x_m+y_2 & p_m \cdot q_2 \\ \vdots & \\ x_m+y_n & p_m \cdot q_n \end{bmatrix}$$

Matrix (1)

which we might also express as $\left(x_i+y_j \quad p_i \cdot q_j \right)_{i=1,m;\, j=1,n}$ †

† For a generalized convolution of f_X and f_Y to generate the distribution $f_{X/Y}$ of the random variable X/Y this expression would be replaced by

$$\left(x_i/y_j \quad p_i \cdot q_j \right)_{i=1,m;\, j=1,n}$$

If m and n are (say) 1000, then generating this matrix would involve 10^6 lines.‡ This would be practical if we do not intend to use f_Z in further convolutions. But, if for example we want to convolute f_Z with itself, then we would be dealing with 10^{12} lines.‡ And, this would be impractical, because of both the amount of computer storage and the amount of computing time required. The following algorithm has been designed to overcome these problems.

The Univariate Generalized Numerical Convolution Algorithm

Choose $\varepsilon>0$. Typically ε is chosen to be 10^{-10} or 10^{-15}.

Loop #1:

Perform the calculations indicated in Matrix (1) above, discarding any lines for which the resulting probability is less than ε; that is discard lines for which

$$p_i \cdot q_j < \varepsilon. \qquad (*)$$

---

‡ There may be some collapsing due to identical amounts on different lines. The number of lines produced is reduced by representing on a single line all lines with identical amounts; on that line is the amount and the sum of the original probabilties.

The purpose of this is to avoid underflow problems and to increase the fineness of the partitions (meshes) to be imposed.

Calculate

$$lowz = \min_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ x_i + y_j \neq 0 \mid p_i \cdot q_j > \varepsilon \}^{\dagger\ \ddagger}$$

and

$$highz = \max_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ x_i + y_j \neq 0 \mid p_i \cdot q_j > \varepsilon \}^{\dagger\ddagger}$$

Let *naz* ≥4 be a positive even integer selected for the purpose of

---

† In many applications we replace x_i+y_j by $\log(x_i+y_j)$, which will allow finer subintervals at the low end of the range. Of course, to be able to use logs the range of X+Y should not include values less than one (to avoid theoretical and numerical problems).

‡ For a generalized convolution of f_X and f_Y to generate the distribution $f_{X/Y}$ of the random variable X/Y these expressions would be replaced by

$$lowu = \min_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ x_i/y_j \neq 0 \mid p_i \cdot q_j > \varepsilon \}$$

and

$$highu = \max_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ x_i/y_j \neq 0 \mid p_i \cdot q_j > \varepsilon \}$$

creating the following partition:

let $\Delta = \frac{highz-lowz}{naz/2-2}$;

partition the interval (*lowz*-Δ,*highz*+Δ) into *naz*/2 subintervals $I_{(r)}$:

r	Subinterval $I_{(r)}$
1	[0,0]
2	(*lowz*-Δ, *lowz*)
3	[*lowz*, *lowz*+1·Δ)
4	[*lowz*+1·Δ, *lowz*+2·Δ)
.	.
.	.
.	.
naz/2	[*highz*,*highz*+Δ)

Subinterval $I_{(1)}$ is the degenerate interval consisting of 0 alone. If for some $r_0>1$ $0 \in I_{(r_0)}$, then 0 is deleted from $I_{(r_0)}$; that is, that particular subinterval has a hole at 0.

Loop #2:

For each r (r=1,2,...,naz/2+1) set to zero the initial value of each of the accumulators

$$m_0(r)$$

$$m_1(r)$$

$$m_2(r) \text{ and}$$

$$m_3(r).$$

For each i (1=1,2,...,m) and j (j=1,2,...,n) for which

$$p_i \cdot q_j > \epsilon,$$

determine the positive integer r for which

$$x_i + y_j \in I(r)$$

and perform the accumulations

$$m_0(r) = m_0(r) + p_i \cdot q_j$$
$$m_1(r) = m_1(r) + (x_i+y_j)^1 \cdot p_i \cdot q_j$$
$$m_2(r) = m_2(r) + (x_i+y_j)^2 \cdot p_i \cdot q_j$$
$$m_3(r) = m_3(r) + (x_i+y_j)^3 \cdot p_i \cdot q_j$$

That is, we generate the probability and the 0^{th} through 3^{rd} moments for each mesh interval $I(r)$ (r=1,2,...,naz/2).

Loop #3:

VonMises Theorem and algorithm (ref[5]) guarantee that for each r (r=1,2,...,naz/2) there exist and we can find two pairs of real numbers

$$(\underline{z}(r), Pr\{\underline{z}(r)\} \text{ and } (\overline{z}(r), Pr_Z(\overline{z}(r)) \ ^{\dagger}$$

such that

$$\underline{z}(r) \in I(r) \text{ and } \overline{z}(r) \in I(r)$$

and such that the following local moment relationships hold:

$$\underline{z}^i \cdot Pr\{\underline{z}(r)\} + \overline{z}^i \cdot Pr\{\overline{z}(r)\} = m_i(r) \text{ for } i=0,1,2,3.$$

Using the 0th through 3rd moments we can produce two points† and

---

† In some cases $x_1=x_2$ and what would otherwise be two pairs ($\underline{z}$, $Pr_Z\{\underline{z}(r)\}$) and ($\overline{z}$, $Pr_Z(\overline{z}(r))$) collapse into one pair ($\underline{z}$, $Pr_Z\{\underline{z}(r)\}$). This would happen, for example, where the values of $x_i + y_j$ that fall into I(r) are all identical.

† Ibid

associated probabilities, with the feature that these moments are accurately retained.

Having kept accurately the 0^{th} through 3^{rd} moments of X+Y within each mesh interval, we have automatically kept accurately the corresponding global moments.

We can then express the full distribution $f_Z = f_{X+Y}$ of the univariate random variable Z = X+Y as

$$\left[\begin{matrix} \underline{z}^{(r)} & Pr_Z(\underline{z}^{(r)}) \\ \overline{z}^{(r)} & Pr_Z(\overline{z}^{(r)}) \end{matrix}\right]_{r=1,naz/2}$$

We will now describe how we actually obtain the number pairs

$$(\underline{z}^{(r)}, Pr_Z(\underline{z}^{(r)})) \quad \text{and} \quad (\overline{z}^{(r)}, Pr_Z(\overline{z}^{(r)}))$$

for any given value of r (r=1,2,...,naz/2). To simplify the notation in this description we will replace the symbols

$$m_i^{(r)} \text{ by } m_i \text{ for } i=0,1,2,3$$
$$\text{and}$$
$$\underline{z}^{(r)} \text{ by } \underline{z} \text{ and } \overline{z}^{(r)} \text{ by } \overline{z}$$

If $m_0=0$ and $m_1 \neq 0$, then we let

$$\underline{z} = 0 \qquad Pr_Z(\underline{z}) = m_0$$

$$\overline{z} = 0 \qquad Pr_Z(\overline{z}) = 0;$$

otherwise,

if $m_0 \cdot m_2 - m_1 \cdot m_1 < 10^{-10} \cdot |m_1|$, we let

$$\underline{z} = m_1/m_0 \qquad Pr_Z(\underline{z}) = m_0$$

$$\overline{z} = 0 \qquad Pr_Z(\overline{z}) = 0;$$

that is, in effect, use a single number pair rather than two number pairs if the variance in $I(r)$ is close to zero.†

otherwise, perform the following calculations:

$$c_0 = \frac{m_1 \cdot m_3 - m_2 \cdot m_2}{m_0 \cdot m_2 - m_1 \cdot m_1}$$

$$c_1 = \frac{m_1 \cdot m_2 - m_0 \cdot m_3}{m_0 \cdot m_2 - m_1 \cdot m_1}$$

$$a_1 = \frac{1}{2} \cdot \left(-c_1 - |c_1 \cdot c_1 - 4 \cdot c_0|^{.5} \right)$$

$$a_2 = \frac{1}{2} \cdot \left(-c_1 + |c_1 \cdot c_1 - 4 \cdot c_0|^{.5} \right)$$

$$s_1 = \frac{m_0 \cdot a_2 - m_1}{a_2 - a_1}$$

$$s_2 = \frac{m_1 - m_0 \cdot a_1}{a_2 - a_1}$$

$$\underline{z} = a_1 \qquad Pr_Z(\underline{z}) = s_1$$

---

† We treat this situation differently in order to avoid exceeding the limits of precision of the numbers being held by the computer.

$$\overline{z} = a_2 \qquad Pr_Z(\overline{z}) = s_2;$$

We check that $\underline{z}$ and $\overline{z}$ both lie in $I(r)$; and, if not, then

if $I(r)$ is a degenerate interval (i.e. consists of a single point), then we let

$$\underline{z} = m_1/m_0 \qquad Pr_Z(\underline{z}) = m_0$$

$$\overline{z} = 0 \qquad Pr_Z(\overline{z}) = 0;$$

otherwise,† we let

$$\sigma = |(-m_1/m_0)\cdot(m_1/m_0)+(m_2/m_0)|^{.5}$$

$$k = \left| \frac{m_1/m_0 - \text{left endpoint of } I(r)}{\text{right endpoint of } I(r) - m_1/m_0} \right|$$

$$\underline{z} = -\sigma\cdot|k|^{.5} + m_1/m_0 \qquad Pr(\underline{z}) = m_0/(1+k)$$

$$\overline{z} = \sigma/|k|^{.5} + m_1/m_0 \qquad Pr_Z(\overline{z}) = Pr(\underline{z})\cdot k;$$

Thus, for each r (r=1,2,...,naz/2) we have generated

$$\begin{bmatrix} \underline{z}(r) & Pr_Z(\underline{z}(r)) \\ \overline{z}(r) & Pr_Z(\overline{z}(r)) \end{bmatrix}.$$

---

† This situation will occur only when the accuracy of the numbers being held by the computer is being impaired by the fact that the computer can hold numbers to only a limited degree of precision; since this situation occurs only where the associated probability is extremely small, the fact that not all of the first three moments are being retained in this situation is not of practical significance.

And the full distribution $f_Z = f_{X+Y}$ can be expressed as

$$\left[\begin{array}{ll} \underline{z}(r) & Pr_Z\{\underline{z}(r)\} \\ \overline{z}(r) & Pr_Z\{\overline{z}(r)\} \end{array} \right]_{r=1,naz/2} .$$

It is desirable to use double precision floating point numbers in performing these calculations; otherwise, you may run into numerical difficulties.

APPENDIX #3 - BIVARIATE NUMERICAL CONVOLUTIONS

If $f_{X,Y}$ and $f_{Z,T}$ are independent distributions of the bivariate discrete finite random variables (X,Y) and (Z,T), respectively, then the distribution $f_{U,V} = f_{X+Z,Y+T}$ of the bivariate random variable (U,V)=(X+Z,Y+T) is the bivariate convolution

$$f_{X,Y} + f_{Z,T}^{\dagger} \text{ of } f_{X,Y} \text{ and } f_{Z,T} \text{ for sums.}$$

Let $f_{X,Y}$ be expressed in element notation as[‡]

$$\begin{bmatrix} (x_1,y_1) & p_1 \\ (x_2,y_2) & p_2 \\ \vdots & \\ (x_m,y_m) & p_m \end{bmatrix}$$

which we will also express as

$$f_{X,Y} = \left((x_i,y_i) \quad p_i \right)_{i=1,m} .$$

Similarly, let $f_{Z,T}$ be

---

† We are using + instead of * to indicate convolution for sums; that is, $f_{X,Y} + f_{Z,T}$.

‡ We represent a line in the distribution $f_{X,Y}$ as a number pair (x_i,y_i) and its associated probability p_i. We could just as well have used a vector $[x_i \; y_i \; p_i]$.

$$f_{Z,T} = \left((z_j, t_j) \quad q_j \right)_{j=1,n} .$$

Then

$$f_{U,V} = f_{X+Z,Y+T} = f_{X,Y} + f_{Z,T} =$$

$$\begin{bmatrix}
(x_1+z_1, y_1+t_1) & p_1 \cdot q_1 \\
(x_1+z_2, y_1+t_2) & p_1 \cdot q_2 \\
 & \\
(x_1+z_n, y_1+t_n) & p_1 \cdot q_n \\
(x_2+z_1, y_2+t_1) & p_2 \cdot q_1 \\
(x_2+z_2, y_2+t_2) & p_2 \cdot q_2 \\
 & \\
(x_2+z_n, y_2+t_n) & p_2 \cdot q_n \\
 & \\
(x_m+z_1, y_m+t_1) & p_m \cdot q_1 \\
(x_m+z_2, y_m+t_2) & p_m \cdot q_2 \\
 & \\
(x_m+z_n, y_m+t_n) & p_m \cdot q_n
\end{bmatrix}$$

Matrix #1

which we might also express as

$$\left[(x_i+z_j, y_i+t_j) \quad p_i \cdot q_j \right]_{\substack{i=1,m \\ j=1,n}} \dagger$$

If m and n are (say) 10^3, then generating this matrix would involve as 10^6 lines.‡ This would be practical if we do not intend to use $f_{U,V}$ for further convolutions. On the other hand, if we want to perform further convolutions such as convoluting $f_{U,V}$ with itself, then we might be having to generate $10^{12} = 10^6 \cdot 10^6$ lines;†‡ and this would be impractical, because of the amount of both computer storage and computing time required. The following algorithm has been designed to overcome these problems.

---

† For a generalized convolution of $f_{X,Y}$ and $f_{Z,T}$ to generate for example the distribution $f_{\min(X,Y+Z),Y+T}$ of the bivariate random variable (min(X,Y+Z), Y+T) this expression would be replaced by

$$\left[(\min(x_i, y_i+z_j), \; y_i+t_j) \quad p_i \cdot q_j \right]_{\substack{i=1,m \\ j=1,n}} .$$

‡ There may be some collapsing due to identical amount pairs on different lines,; that is, the number of lines produced is reduced by representing on a single line all lines with identical amount pairs; on that single line is the amount pair and the sum of the original probabilities.

†‡ Ibid

The Bivariate Convolution Algorithm

Choose $\varepsilon>0$. Typically ε is chosen to be 10^{-10} or 10^{-15}.

Loop #1:

Perform the calculations indicated in Matrix #1 above, discarding any lines for which the resulting probability is less than ε; that is, discard lines for which

$$p_i \cdot q_j < \varepsilon.$$

The purpose of this is to avoid underflow problems and to increase the fineness of the partitions (meshes) to be imposed.

Calculate

$$lowu^{\dagger} = \min_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ x_i+z_j \neq 0 \mid p_i \cdot q_j > \varepsilon \}^{\ddagger}$$

and

---

†"*lowu*" stands for "lowest amount on the u-axis."

‡ For a generalized convolution of $f_{X,Y}$ and $f_{Z,T}$ to generate the distribution $f_{\min(X,Y+Z),Y+T}$ of the bivariate random variable $(\min(X,Y+Z),\ Y+T)$ these expressions would be replaced by

$$lowu = \min_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ \min(x_i,y_i+z_j) \neq 0 \mid p_i \cdot q_j > \varepsilon \})$$

$$highu = \max_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ \min(x_i,y_i+z_j) \neq 0 \mid p_i \cdot q_j > \varepsilon \}$$

$$highu^{\dagger\dagger} = \max_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{\, x_i+z_j \neq 0 \mid p_i \cdot q_j > \varepsilon \,\}^{\ddagger}$$

Let $nau^{\dagger} \geq 4$ be a positive even integer selected for the purpose of creating a partition.

$$\text{let } \Delta = \frac{highu-lowu}{nau/2 - 2};$$

partition the open interval $(lowu-\Delta, highu+\Delta)$ into $nau/2$ subintervals $I_{(r)}$:

r	Subinterval $I_{(r)}$
1	$[0,0]$
2	$(lowu-\Delta, lowu]$
3	$(lowu, lowu+\Delta)$
4	$[lowu+\Delta, lowu+2\cdot\Delta)$
.	.
.	.
.	.
$nau/2$	$[highu, highu+\Delta)$

Subinterval $I_{(1)}$ is the degenerate interval consisting of 0 alone.

If for some $r_0>1$ $0 \in I_{(r_0)}$, then 0 is deleted from $I_{(r_0)}$; that is, that particular subinterval has a hole at 0.

---

††"*highu*" stands for "highest amount on the u-axis."

†"*nau*" stands for "number of amounts along the u-axis."

Loop #2:

For each r (r=1,2,...,*nau*/2) calculate

$$low v_r = \min_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ y_i+t_j \neq 0 \mid p_i \cdot q_j > \varepsilon \text{ and } x_i+z_j \in I(r) \}^{\dagger}$$

and

$$high v_r = \max_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ y_i+t_j \neq 0 \mid p_i \cdot q_j > \varepsilon \text{ and } x_i+z_j \in I(r) \}^{\dagger}$$

let *nav*‡ be a positive integer selected for the purpose of

---

†For a generalized convolution of $f_{X,Y}$ and $f_{Z,T}$ to generate the distribution $f_{\min(X,Y+Z),Y+T}$ of the bivariate random variable (min{X,Y+Z}, Y+T) these expressions would be replaced by

$$low u = \min_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ y_i+t_j \neq 0 \mid p_i \cdot q_j > \varepsilon \text{ and } \min(x_i, y_i+z_j) \in I(r) \}$$

$$high u = \max_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} \{ y_i+t_j \neq 0 \mid p_i \cdot q_j > \varepsilon \text{ and } \min(x_i, y_i+z_j) \in I(r) \}$$

‡ "*nav*" stands for "number of amounts along the v axis."

creating the following partition:

calculate $\delta_r = \dfrac{highv_r - lowv_r}{nav/2 - 2}$;

partition each of the open intervals

$$(lowv_r - \delta_r,\ highv_r + \delta_r)$$

into $nav/2$ subintervals $J_{r(s)}$:

s	Subinterval $J_{r(s)}$
1	[0,0]
2	$(lowv_r - \delta_r, lowv_r]$
3	$(lowv_r, lowv_r + \delta_r)$
4	$[lowv_r + \delta_r, lowv_r + 2\cdot\delta_r)$
.	.
.	.
.	.
$nav/2$	$[highv_r, highv_r + \delta_r)$

Loop #3:

Let $I_{(r)} \times J_{r(s)}$ (called a mesh rectangle) denote the Cartesian product of $I_{(r)}$ and $J_{r(s)}$.

For each ordered pair (r,s)

$$(r=1,2,\ldots,nau/2 \quad ; \quad s=1,2,\ldots,nav/2)$$

set to zero the initial value of each of the moment accumulators

$$m_{00}(r,s)$$
$$m_{10}(r,s)$$
$$m_{20}(r,s)$$
$$m_{30}(r,s)$$
$$m_{01}(r,s)$$
$$m_{02}(r,s)$$
$$m_{03}(r,s)$$
$$m_{11}(r,s)$$

and

for each i (i=1,2,...,m) and j (j=1,2,...,n) for which

$$p_i \cdot q_j > \varepsilon,$$

determine the positive integers r and s for which both

$$x_i+z_j \in I(r) \text{ and } y_i+t_j \in J_r(s)$$

and perform the accumulations

$$m_{00}(r,s) = m_{00}(r,s) + p_i \cdot q_j$$

$$m_{10}(r,s) = m_{10}(r,s) + (x_i+z_j)^1 \cdot p_i \cdot q_j$$

$$m_{20}(r,s) = m_{20}(r,s) + (x_i+z_j)^2 \cdot p_i \cdot q_j$$

$$m_{30}(r,s) = m_{30}(r,s) + (x_i+z_j)^3 \cdot p_i \cdot q_j$$

$$m_{01}(r,s) = m_{01}(r,s) + (y_i+t_j)^1 \cdot p_i \cdot q_j$$

$$m_{02}(r,s) = m_{02}(r,s) + (y_i+t_j)^2 \cdot p_i \cdot q_j$$

$$m_{03(r,s)} = m_{03(r,s)} + (y_i+t_j)^3 \cdot p_i \cdot q_j$$

$$m_{11(r,s)} = m_{11(r,s)} + (x_i+z_j) \cdot (y_i+t_j) \cdot p_i \cdot q_j;$$

that is, we generate the probability, the 0^{th} through 3^{rd} x-moments, the 0^{th} through 3^{rd} y-moments, and the joint first moment for each mesh rectangle

$$I_{(r)} \times J_{r(s)} \quad (r=1,2,\ldots,nau/2 \quad ; \; s=1,2,\ldots,nav/2 \quad).$$

Loop #4:

For each pair (r,s) (r=1,2,...,nau/2 ;s=1,2,...,nav/2) there exist and by the end of this Appendix we will have described how to calculate four triplets of real numbers

$$\begin{bmatrix} \underline{u}_{(r,s)} & \underline{v}_{(r,s)} & Pr_{U,V}\{(\underline{u}_{(r,s)}, \underline{v}_{(r,s)})\} \\ \underline{u}_{(r,s)} & \overline{v}_{(r,s)} & Pr_{U,V}\{(\underline{u}_{(r,s)}, \overline{v}_{(r,s)})\} \\ \overline{u}_{(r,s)} & \underline{v}_{(r,s)} & Pr_{U,V}\{(\overline{u}_{(r,s)}, \underline{v}_{(r,s)})\} \\ \overline{u}_{(r,s)} & \overline{v}_{(r,s)} & Pr_{U,V}\{(\overline{u}_{(r,s)}, \overline{v}_{(r,s)})\} \end{bmatrix}$$

such that $Pr_U\{(\underline{u}_{(r,s)}, \overline{v}_{(r,s)})\}$;

$$\underline{u}_{(r,s)} \in I_{(r)}$$

$$\overline{u}_{(r,s)} \in I_{(r)}$$

$$\underline{v}_{(r,s)} \in J_r(s)$$

$$\overline{v}_{(r,s)} \in J_r(s)$$

and such that the 0^{th} through the 3^{rd} (local) moments agree:

$$\underline{u}_{(r,s)}^{\,i} \cdot Pr_U(\underline{u}_{(r,s)}) + \overline{u}_{(r,s)}^{\,i} \cdot Pr_U(\overline{u}_{(r,s)}) = m_{i0}(r,s), \text{ where } i=0,1,2,3;$$

$$\underline{v}_{(r,s)}^{\,j} \cdot Pr_V(\underline{v}_{(r,s)}) + \overline{v}_{(r,s)}^{\,j} \cdot Pr_V(\overline{v}_{(r,s)}) = m_{0j}(r,s), \text{ where } j=0,1,2,3;$$

We note that each of the points

$$(\underline{u}_{(r,s)}, \underline{v}_{(r,s)})$$
$$(\underline{u}_{(r,s)}, \overline{v}_{(r,s)})$$
$$(\overline{u}_{(r,s)}, \underline{v}_{(r,s)})$$
and
$$(\overline{u}_{(r,s)}, \overline{v}_{(r,s)})$$

$$\in I_r \times J_r(s) .$$

VonMises To The Rescue In The U-Direction

To simplify the notation in this description we will abbreviate

$\underline{u}_{(r,s)}$ by $\underline{u}$

$\overline{u}_{(r,s)}$ by $\overline{u}$

and

$m_{i0}(r,s)$ by m_{i0} for i = 0,1,2,3.

For any particular (r,s):

if $m_{10}=0$ and $m_{00} \neq 0$, then we let

$$\underline{u} = 0 \qquad Pr_U(\underline{u}) = m_{10}$$

$$\overline{u} = 0 \qquad Pr_U(\overline{u}) = 0 \ ;$$

otherwise,

if $m_{00} \cdot m_{20} - m_{10} \cdot m_{10} < 10^{-10} \cdot |m_{10}|$, we let

$$\underline{u} = m_{10}/m_{00}$$

$$Pr_U(\underline{u}) = m_{00}$$

$$\overline{u} = 0$$

$$Pr_U(\overline{u}) = 0;$$

that is, in effect, use a single number pair rather than two number pairs if the variance in $I_{(r)}$ is close to zero.†

---

† We treat this situation differently in order to avoid exceeding the limits of precision of the numbers being held by the computer.

Otherwise, perform the following calculations:

$$c_0 = \frac{m_{10} \cdot m_{30} - m_{20} \cdot m_{20}}{m_{00} \cdot m_{20} - m_{10} \cdot m_{10}}$$

$$c_0 = \frac{m_{10} \cdot m_{20} - m_{00} \cdot m_{30}}{m_{00} \cdot m_{20} - m_{10} \cdot m_{10}}$$

$$a_1 = \frac{1}{2} \cdot (-c_1 - |c_1 \cdot c_1 - 4 \cdot c_0|^{.5})$$

$$a_2 = \frac{1}{2} \cdot (-c_1 + |c_1 \cdot c_1 - 4 \cdot c_0|^{.5})$$

$$s_1 = \frac{m_{00} \cdot a_2 - m_{10}}{a_2 - a_1}$$

$$s_2 = \frac{m_{10} - m_{00} \cdot a_1}{a_2 - a_1}$$

$$\underline{u} = a_1$$

$$Pr_U(\underline{u}) = s_1$$

$$\overline{u} = a_2$$

$$Pr_U(\overline{u}) = s_2;$$

we check that $\underline{u}$ and $\bar{u}$ both lie in $I_{(r)}$; and, if not, then

if $I_{(r)}$ is a degenerate interval (i.e. consists of a single point), then we let

$$\underline{u} = m_{10}/m_{00}$$

$$Pr_U(\underline{u}) = m_{00}$$

$$\bar{u} = 0$$

$$Pr_U(\bar{u}) = 0 \ ;$$

otherwise†, we let

$$\sigma = \left| \ (-m_{10}/m_{00}) \cdot (m_{10}/m_{00}) + (m_{20}/m_{00}) \ \right|^{.5}$$

$$k = \left| \ \frac{m_{10}/m_{00} - \text{left endpoint of } I_{(r)}}{\text{right endpoint of } I_{(r)} - m_{10}/m_{00}} \ \right|$$

$$\underline{u} = -\sigma \cdot |k|^{.5} + m_{10}/m_{00}$$

---

† This situation will occur only when the accuracy of the numbers being held by the computer is being impaired by the fact that the computer can hold numbers to ony a limited degree of precision; since this situation occurs only where the associated probability is extermely small, the fact that not all of the first three moments are being retained in this situation is not of practical significance.

$$Pr_U(\underline{u}) = m_{00}/(1+k)$$

$$\bar{u} = +\sigma/|k|^{.5} + m_{10}/m_{00}$$

$$Pr_U(\bar{u}) = Pr_U(\underline{u})\cdot k \ ;$$

It is desirable to use double precision floating point numbers† in performing these calculations; otherwise, you may run into numerical difficulties.

In this way we calculate $\underline{u}_{(r,s)}$ and $\bar{u}_{(r,s)}$ for each pair (r,s) (r=1,2,...nau/2 ;s=1,2,...,nav/2)

VonMises To The Rescue In The V-Direction

To simplify the notation in this description we will abbreviate

$\underline{v}_{(r,s)}$ by $\underline{v}$

$\bar{v}_{(r,s)}$ by $\bar{v}$

$m_{oj}{}_{(r,s)}$ by m_{oj} for j = 0,1,2,3.

---

†In the "C" programming language we use "long double" floating point numbers.

For any particular (r,s):

if $m_{01} = 0$ and $m_{\infty} \neq 0$, then we let

$$\underline{v} = 0$$

$$Pr_V(\underline{v}) = m_{01}$$

$$\overline{v} = 0$$

$$Pr_V(\overline{v}) = 0 \ ;$$

otherwise,

if $m_{\infty} \cdot m_{02} - m_{01} \cdot m_{01} < 10^{-10} \cdot |m_{01}|$, we let

$$\underline{v} = m_{01}/m_{\infty}$$

$$Pr_V(\underline{v}) = m_{\infty}$$

$$\overline{v} = 0$$

$$Pr_V(\overline{v}) = 0;$$

that is, in effect, use a single number pair rather than two number pairs if the variance in $J_{r(s)}$ is close tozero.†

---

† We treat this situation differently in order to avoid exceeding the limits of precision of the numbers being held by the computer.

otherwise, perform the following calculations:

$$c_0 = \frac{m_{01} \cdot m_{03} - m_{02} \cdot m_{02}}{m_{00} \cdot m_{02} - m_{01} \cdot m_{01}}$$

$$c_1 = \frac{m_{01} \cdot m_{02} - m_{00} \cdot m_{03}}{m_{00} \cdot m_{02} - m_{01} \cdot m_{01}}$$

$$a_1 = \frac{1}{2} \cdot \left(-c_1 - \left| c_1 \cdot c_1 - 4 \cdot c_0 \right|^{.5} \right)$$

$$a_2 = \frac{1}{2} \cdot \left(-c_1 + \left| c_1 \cdot c_1 - 4 \cdot c_0 \right|^{.5} \right)$$

$$s_1 = \frac{m_{00} \cdot a_2 - m_{01}}{a_2 - a_1}$$

$$s_2 = \frac{m_{01} - m_{00} \cdot a_1}{a_2 - a_1}$$

$$\underline{v} = a_1$$

$$Pr_V(\underline{v}) = s_1$$

$$\overline{v} = a_2$$

$$Pr_V(\overline{v}) = s_2 \ ;$$

check that $\underline{v}$ and $\overline{v}$ both lie in $J_{r(s)}$; and, if not, then if $J_{r(s)}$ is a degenerate interval (i.e. consists of a single point), then we let

$$\underline{v} = m_{01}/m_{00}$$

$$Pr_V(\underline{v}) = m_{00}$$

$$\overline{v} = 0$$

$$Pr_V(\overline{v}) = 0 \ ;$$

otherwise†, we let

$$\sigma = \left| \ (-m_{01}/m_{00}) \cdot (m_{01}/m_{00}) + (m_{02}/m_{00}) \ \right|^{.5}$$

$$k = \left| \ \frac{m_{01}/m_{00} - \text{left endpoint of } J_{r(s)}}{\text{right endpoint of } J_{r(s)} - m_{01}/m_{00}} \ \right|$$

---

† This situation will occur only when the accuracy of the numbers being held by the computer is being impaired by the fact that the computer can hold numbers to ony a limited degree of precision; since this situation occurs only where the associated probability is extremely small, the fact that not all of the first three moments are being retained in this situation is not of practical significance.

$$\underline{v} = -\sigma \cdot |k|^{-s} + m_{01}/m_{00} \qquad Pr_V(\underline{v}) = m_{00}/(1+k)$$

$$\overline{v} = +\sigma \cdot |k|^{-s} + m_{01}/m_{00}$$

$$Pr_V(\overline{v}) = (Pr_V(\underline{v})) \cdot k \ ;$$

In this way we calculate for each pair (r,s) (r=1,2,...,nau/2 ;s=1,2,...,nav/2)

$\underline{u}$ $\quad Pr_U(\underline{u})$

$\overline{u}$ $\quad Pr_U(\overline{u})$

$\underline{v}$ $\quad Pr_V(\underline{v})$

$\overline{v}$ $\quad Pr_V(\overline{v})$

with the feature that the first three marginal u-moments and the first three marginal v-moments are accurately retained.

Loop #5

Let the partial distribution of (X + Z, Y + T) restricted to $I_{(r)} \times J_{(s)}$ be represented by the four points and associated probabilities

$$\begin{bmatrix} (\underline{u},\ \underline{v}) & Pr_{U,V}((\underline{u},\underline{v})) \\ (\underline{u},\ \overline{v}) & Pr_{U,V}((\underline{u},\overline{v})) \\ (\overline{u},\ \underline{v}) & Pr_{U,V}((\overline{u},\underline{v})) \\ (\overline{u},\ \overline{v}) & Pr_{U,V}((\overline{u},\overline{v})) \end{bmatrix}$$

where the probabilities are determined as follows:

we have $Pr_U(\underline{u})$, $Pr_U(\overline{u})$, $Pr_V(\underline{v})$ and $Pr_V(\overline{v})$; we want to determine

$$Pr_{U,V}((\underline{u},\underline{v}))$$
$$Pr_{U,V}((\underline{u},\overline{v}))$$
$$Pr_{U,V}((\overline{u},\underline{v}))$$
and
$$Pr_{U,V}((\overline{u},\overline{v})),$$

each in the interval [0,1], so that

$$Pr_{U,V}((\underline{u},\underline{v})) + Pr_{U,V}((\overline{u},\underline{v})) = Pr_V(\underline{v})$$

$$Pr_{U,V}((\underline{u},\overline{v})) + Pr_{U,V}((\overline{u},\overline{v})) = Pr_V(\overline{v})$$

$$Pr_{U,V}((\underline{u},\underline{v})) + Pr_{U,V}((\underline{u},\overline{v})) = Pr_U(\underline{u})$$

$$Pr_{U,V}((\overline{u},\underline{v})) + Pr_{U,V}((\overline{u},\overline{v})) = Pr_U(\overline{u})$$

and so as to minimize the absolute value of the error

$$\left| m_{11} - \left(\underline{u}\cdot\underline{v}\cdot Pr_{U,V}((\underline{u},\underline{v})) + \underline{u}\cdot\overline{v}\cdot Pr_{U,V}((\underline{u},\overline{v})) + \overline{u}\cdot\underline{v}\cdot Pr_{U,V}((\overline{u},\underline{v})) + \overline{u}\cdot\overline{v}\cdot Pr_{U,V}((\overline{u},\overline{v})) \right) \right|$$

in the joint moment.

We note that choosing a value for $Pr_{U,V}((\underline{u},\underline{v}))$ will determine the values for $Pr_{U,V}((\underline{u},\overline{v}))$, $Pr_{U,V}((\overline{u},\underline{v}))$ and $Pr_{U,V}((\overline{u},\overline{v}))$ in the four equations. Our choice of $Pr_{U,V}((\underline{u},\underline{v}))$ is not completely arbitrary, since $Pr_{U,V}((\underline{u},\overline{v}))$, $Pr_{U,V}((\overline{u},\underline{v}))$ and $Pr_{U,V}((\overline{u},\overline{v}))$ must each $\in [0,1]$. But, restricting $Pr_{U,V}((\underline{u},\underline{v}))$ to this extent, we can proceed to choose $Pr_{U,V}((\underline{u},\underline{v}))$ so as to minimize the absolute error in the joint moment. Choosing $Pr_{U,V}((\underline{u},\underline{v}))$ to do this is straightforward, because the error in the joint moment can be expressed as a linear function of $Pr_{U,V}((\underline{u},\underline{v}))$.† The next section describes the details of these calculations.

---

† The absolute value of the error in the joint moment is linear in two pieces, because of the absolute value being taken.

The Attempt To Minimize The Absolute Value Of The Error In The Joint Moment

Calculate

$t = \underline{u}\cdot\underline{v} - \underline{u}\cdot\overline{v} - \overline{u}\cdot\underline{v} + \overline{u}\cdot\overline{v}$

if $t < 10^{-10}$ and $| \underline{u} - \overline{u} | < 10^{-10}$ and $| \underline{v} - \overline{v} | < 10^{-10}$,
then let

$$Pr_{U,V}((\underline{u},\underline{v})) = Pr_U(\underline{u}) + Pr_U(\overline{u})$$

$$Pr_{U,V}((\underline{u},\overline{v})) = 0$$

$$Pr_{U,V}((\overline{u},\underline{v})) = 0$$

$$Pr_{U,V}((\overline{u},\overline{v})) = 0 \ ;$$

if $t < 10^{-10}$ and $| \underline{u} - \overline{u} | \geq 10^{-10}$ and $| \underline{v} - \overline{v} | < 10^{-10}$,

then let

$$Pr_{U,V}((\underline{u},\underline{v})) = Pr_U(\underline{u})$$

$$Pr_{U,V}((\overline{u},\underline{v})) = Pr_U(\overline{u})$$

$$Pr_{U,V}((\underline{u},\overline{v})) = 0$$

$$Pr_{U,V}((\bar{u},\bar{v})) = 0 \ ;$$

if $t < 10^{-10}$ and $| \underline{u} - \bar{u} | < 10^{-10}$ and $| \underline{v} - \bar{v} | \geq 10^{-10}$, then let

$$Pr_{U,V}((\underline{u},\underline{v})) = Pr_V(\underline{v})$$

$$Pr_{U,V}((\underline{u},\bar{v})) = Pr_V(\bar{v})$$

$$Pr_{U,V}((\bar{u},\underline{v})) = 0$$

$$Pr_{U,V}((\bar{u},\bar{v})) = 0 \ .$$

if $t \geq 10^{-10}$ and $Pr_V(\underline{v}) - Pr_U(\bar{u}) > 0$, then

let *lower* = $Pr_V(\underline{v}) - Pr_U(\bar{u})$;

let *higher* = minimum($Pr_V(\underline{v})$, $Pr_U(\bar{u})$);

let $Pr_{U,V}((\underline{u},\underline{v})) =$

$$\begin{array}{l} (\ m_{11} \\ \quad -\underline{u}\cdot\bar{v}\cdot Pr_U(\underline{u}) \\ \quad -\bar{u}\cdot\underline{v}\cdot Pr_V(\underline{v}) \\ \quad -\bar{u}\cdot\bar{v}\cdot (Pr_V(\bar{v}) - Pr_U(\underline{u}) \ ) \\ / \\ (\ \underline{u}\cdot\underline{v} - \underline{u}\cdot\bar{v} - \bar{u}\cdot\underline{v} \ ) \end{array}$$

if $Pr_{U,V}((\underline{u},\underline{v})) < lower$, then let $Pr_{U,V}((\underline{u},\underline{v})) = lower$;

if $Pr_{U,V}((\underline{u},\underline{v})) > higher$, then let $Pr_{U,V}((\underline{u},\underline{v})) = higher$;

if $Pr_{U,V}((\underline{u},\underline{v})) < 0$, then $Pr_{U,V}((\underline{u},\underline{v})) = 0$;

let $Pr_{U,V}((\underline{u},\overline{v})) = Pr_{U}(\underline{u}) - Pr_{U,V}((\underline{u},\underline{v}))$;

if $Pr_{U,V}((\underline{u},\overline{v})) < 0$, let $Pr_{U,V}((\underline{u},\overline{v})) = 0$;

let $Pr_{U,V}((\overline{u},\overline{v})) = Pr_{V}(\overline{v}) - Pr_{U}((\underline{u},\overline{v}))$;

if $Pr_{U,V}((\overline{u},\overline{v})) < 0$, let $Pr_{U,V}((\overline{u},\overline{v})) = 0$;

let $Pr_{U,V}((\overline{u},\underline{v})) = Pr_{U}(\overline{u}) - Pr_{U,V}((\overline{u},\overline{v}))$;

if $Pr_{U,V}((\overline{u},\underline{v})) < 0$, let $Pr_{U,V}((\overline{u},\underline{v})) = 0$;

If $t \geq 10^{-10}$ and $Pr_{V}(\underline{v}) - Pr_{U}(\overline{u}) \leq 0$, then

let $lower = Pr_{U}(\overline{u}) - Pr_{V}(\underline{v})$;

let $higher = minimum(\ Pr_{U}(\overline{u})\ ,\ Pr_{V}(\overline{v})\ )$;

$$\text{let } Pr_{U,V}((\bar{u},\bar{v})) =$$

$$\left(m_{11} - \underline{u}\cdot\bar{v}\cdot Pr_V(\bar{v}) - \bar{u}\cdot\underline{v}\cdot Pr_U(\bar{u}) - \underline{u}\cdot\underline{v}\cdot(Pr_U(\underline{u}) - Pr_V(\bar{v})) \right) / \left(\underline{u}\cdot\underline{v} - \underline{u}\cdot\bar{v} - \bar{u}\cdot\underline{v} \right)$$

if $Pr_{U,V}((\bar{u},\bar{v})) < lower$, then let $Pr_{U,V}((\bar{u},\bar{v})) = lower$;

if $Pr_{U,V}((\bar{u},\bar{v})) > higher$, then let $Pr_{U,V}((\bar{u},\bar{v})) = higher$;

if $Pr_{U,V}((\bar{u},\bar{v})) < 0$, then $Pr_{U,V}((\bar{u},\bar{v})) = 0$;

let $Pr_{U,V}((\bar{u},\underline{v})) = Pr_U(\bar{u}) - Pr_{U,V}((\bar{u},\bar{v}))$;

if $Pr_{U,V}((\bar{u},\underline{v})) < 0$, let $Pr_{U,V}((\bar{u},\underline{v})) = 0$;

let $Pr_{U,V}((\underline{u},\bar{v})) = Pr_V(\bar{v}) - Pr_{U,V}((\bar{u},\bar{v}))$;

if $Pr_{U,V}((\underline{u},\bar{v})) < 0$, let $Pr_{U,V}((\underline{u},\bar{v})) = 0$;

let $Pr_{U,V}((\underline{u},\underline{v})) = Pr_U(\underline{u}) - Pr_{U,V}((\underline{u},\bar{v}))$;

if $Pr_{U,V}((\underline{u},\underline{v})) < 0$, let $Pr_{U,V}((\underline{u},\underline{v})) = 0$;

Conclusion

Having kept accurately the first three moments of each of X + Z and Y + T, respectively, within each mesh rectangle, we have automatically kept accurately the corresponding global moments.

For each mesh rectangle

$$I(r)\times J_r(s) \qquad (r=1,nau/2 \;\; ; \;\; s=1,nav/2 \;\;)$$

there corresponds

$$\begin{bmatrix} (\underline{u}(r,s),\ \underline{v}(r,s)) & Pr_{U,V}\{(\underline{u}(r,s),\underline{v}(r,s))\} \\ (\underline{u}(r,s),\ \overline{v}(r,s)) & Pr_{U,V}\{(\underline{u}(r,s),\overline{v}(r,s))\} \\ (\overline{u}(r,s),\ \underline{v}(r,s)) & Pr_{U,V}\{(\overline{u}(r,s),\underline{v}(r,s))\} \\ (\overline{u}(r,s),\ \overline{v}(r,s)) & Pr_{U,V}\{(\overline{u}(r,s),\overline{v}(r,s))\} \end{bmatrix}$$

We can then express the full distribution $f_{U,V} = f_{X+Z,Y+T}$ of the bivariate random variable (U,V) = (X+Z,Y+T) as

$$\left.\begin{bmatrix} (\underline{u}(r,s),\ \underline{v}(r,s)) & Pr_{U,V}\{(\underline{u}(r,s),\underline{v}(r,s))\} \\ (\underline{u}(r,s),\ \overline{v}(r,s)) & Pr_{U,V}\{(\underline{u}(r,s),\overline{v}(r,s))\} \\ (\overline{u}(r,s),\ \underline{v}(r,s)) & Pr_{U,V}\{(\overline{u}(r,s),\underline{v}(r,s))\} \\ (\overline{u}(r,s),\ \overline{v}(r,s)) & Pr_{U,V}\{(\overline{u}(r,s),\overline{v}(r,s))\} \end{bmatrix}\right]_{\substack{r=1,nax/2 \\ s=1,nay/2}}$$

REFERENCES

1. Blum, J. R. and Rosenblatt, J. I., Probability and Statistics, W. B. Saunders Company, Philadelphia, 1972, pp 271-2

2. Dufresne, F. and Gerber, H. U., "Three Methods to Calculate the Probability of Ultimate Ruin", Astin Bulletin, Vol. 19, No. 1, April 1989

3. Feller, W., An Introduction to Probability Theory and Its Applications, Vol. I, John Wiley & Sons, Inc., New York, 1968, pp 233-4

4. Panjer, H., "Recursive Evaluation of a Family of Compound Distributions", Astin Bulletin, Vol. 12, 1981

5. Springer, M. D., The Algebra of Random Variables, John Wiley & Sons, Inc., New York, 1979, pp 269-270

ACKNOWLEDGEMENTS

I want to thank Larry Hickey of PolySystems, Inc., in Chicago for focusing my attention on Von Mises Theorem and for patiently tutoring me in the "C" programming language. He has been helful to me over the years as programmer, mathematicican and codesigner. Josh Zirin, formerly of Kemper Group, assisted in implementing the VonMises algorithm in a "C" language program. Charles Fuhrer, F.S.A., of Washington National provided editorial assistance. Thanks also to Anne Hoban of Kemper National and the CHIWRITER scientific/multifont word processor.