Mixing Collective Risk Models

Leigh J. Halliwell, FCAS, MAAA

CARe Seminar Research Corner
Boston, MA
May 19, 2008

1. A setting for the problem:

Within a state and hazard group, NCCI has five injury types (Fatal, PT, PP, TT, Med). One can simulate or derive the ST×HG aggregate distributions from summing five IT collective-risk models of the form: $S = X_1 + \ldots + X_M$. Can we mix the five severity distributions so as to form one equivalent collective-risk model?

2. Let's simplify to two groups and two collective-risk models:

$$\begin{array}{lr} S = X_1 + \ldots + X_M & E[M] = \mu > 0 \\ T = Y_1 + \ldots + Y_N & E[N] = \nu > 0 \\ \hline U = Z_1 + \ldots + Z_{M+N} & E[M+N] = \mu + \nu \end{array}$$

Here Z is the mixture of X and Y in proportion to μ and ν.

3. Since $E[Z] = \dfrac{\mu E[X] + \nu E[Y]}{\mu + \nu}$,

$$E[U] = E[M+N]E[Z] = (\mu + \nu)\frac{\mu E[X] + \nu E[Y]}{\mu + \nu} = \mu E[X] + \nu E[Y] = E[S+T]$$

Hence, the mixed model, in which the severity distributions are mixed according to expected claim frequencies, preserves the total expected loss. So this simplification is good to the first moment. Under what conditions is it completely equivalent?

4. If X and Y are identically distributed (viz., $X \sim Y$), then $Z = mix(X, Y) \sim X \sim Y$, and the simplified model is identically distributed to the sum of the group models, i.e., $U \sim S + T$. But this is a trivial case; normally the severity distributions are different. Another trivial case is for either μ or ν to be zero.

5. In some examples $Var[U] < Var[S] + Var[T] = Var[S+T]$. This might be the usual case in real insurance problems. After all, we might feel that a loss of information goes with a loss of variance. However, here's an example in which $Var[U] > Var[S+T]$:

$Prob[X = 0] = Prob[Y = 2] = Prob[M = \mu] = Prob[N = \nu] = 1$, i.e., all the variables are constants (again, both μ and ν are positive) $Var[S + T] = 0$. However, $Z = mix(X, Y)$ has a positive variance. Then the variance of the simplified model is greater than the sum of the group variances.

6. Variables whose moment generating functions are equal are identically distributed. The same holds true for cumulant generating functions, $\psi_X(t) \equiv \ln M_X(t) = \ln E[e^{tX}]$.

<u>First</u>, as to severity:

$$\begin{aligned}
M_z(t) &= E[e^{tZ}] \\
&= E[e^{tZ} | Z = X] \cdot Prob[Z = X] + E[e^{tZ} | Z = Y] \cdot Prob[Z = Y] \\
&= E[e^{tX}] \cdot Prob[Z = X] + E[e^{tY}] \cdot Prob[Z = Y] \\
&= \frac{\mu M_X(t) + \nu M_Y(t)}{\mu + \nu}
\end{aligned}$$

In other symbols, $M_Z(t) = mix(M_X(t), M_Y(t))$ in proportion to μ and ν.

<u>Second</u>, as to the collective risk model:

$$\begin{aligned}
M_S(t) &= E[e^{tS}] = E[e^{t(X_1 + \ldots X_M)}] \\
&= \underset{M}{E}[E[e^{t(X_1 + \ldots X_M)} | M = m]] \\
&= E[M_X(t)^M] \\
&= E[e^{\ln M_X(t) \cdot M}] = M_M(\ln M_X(t))
\end{aligned}$$

$$\psi_S(t) = \psi_M(\ln M_X(t))$$

(Please pardon the confusion here between M as a random variable and M as the moment generating function.) For the other models:

$$\begin{aligned}
&\psi_T(t) = \psi_N(\ln M_Y(t)) \\
&\psi_U(t) = \psi_U(\ln M_Z(t)) = \psi_S(\ln M_Z(t)) + \psi_T(\ln M_Z(t))
\end{aligned}$$

<u>Third</u>, the moment and cumulant generating functions a Poisson random variable with mean θ are:

$$M_{Poisson[\theta]}(t) = E\left[e^{tP}\right] = \sum_{k=0}^{\infty} e^{tk} \frac{\theta^k e^{-\theta}}{k!} = e^{-\theta} \sum \frac{\left(\theta e^t\right)^k}{k!} = e^{-\theta} e^{\theta e^t} = e^{\theta\left(e^t - 1\right)}$$

$$\psi_{Poisson[\theta]}(t) = \ln M_{Poisson[\theta]}(t) = \theta\left(e^t - 1\right)$$

7. Therefore, if *M* and *N* are Poisson-distributed (with respective means μ and ν):

$$\psi_S(t) = \mu\left(e^{\ln M_x(t)} - 1\right) = \mu\left(M_x(t) - 1\right) \qquad \psi_T(t) = \nu\left(M_Y(t) - 1\right)$$

And hence,

$$\begin{aligned}
\psi_U(t) &= (\mu + \nu)\left(M_Z(t) - 1\right) \\
&= (\mu + \nu)\left(\frac{\mu M_X(t) + \nu M_Y(t)}{\mu + \nu} - 1\right) \\
&= \mu M_X(t) + \nu M_Y(t) - \mu - \nu \\
&= \mu\left(M_X(t) - 1\right) + \nu\left(M_Y(t) - 1\right) \\
&= \psi_S(t) + \psi_T(t) \\
&= \psi_{S+T}(t)
\end{aligned}$$

And thus, $U \sim S + T$, i.e., the distribution of *U* is that of the sum of *S* and *T*. In words, if the frequency distributions are Poisson, the mixed model contains all the information of the sum of the group models.

In non-trivial cases the converse seems to be true, although I cannot prove it. However, the key seems to be that $\psi_M(t)$ and $\psi_N(t)$ are linear in $e^t - 1$, and only the Poisson distribution has such a cumulant generating function. Hence, the likelihood of the converse.