Mixing Collective Risk Models

Leigh J. Halliwell, FCAS, MAAA

Abstract: The form of the collective risk model is $S = X_1 + \ldots + X_M$, where X represents a common severity distribution and *M* is a claim-count random variable. The model for a second loss is $T = Y_1 + \ldots + Y_N$. Let Z represent the mixed severity, i.e., the distributions of X and Y weighted according to their expected claim counts. How does the mixed model $U = Z_1 + \ldots + Z_{M+N}$ compare with $S + T$? Although the mean is unaffected, we will show that if the claim-count distributions are negative binomial with a common contagion, $Var[U] \le Var[S+T]$. In other words, attendant to the reduction of homogeneity (upon mixing the severities) is a reduction of variance. An appendix reveals the conditions under which one may fully reduce a set of collective risk models to one mixed model. This should be of value to the task of modeling correlated exposures.

Keywords: collective risk model, mixed distribution, negative binomial, contagion, homogenous, moment generating function, multinomial

Most casualty actuaries know that the National Council on Compensation Insurance divides workers' compensation claims into five injury types: fatal, permanent total, permanent partial, temporary total, and medical only. It publishes excess loss factors (ELFs) for these injury types, to which one can fit underlying severity distributions. Moreover, by weighting these ELFs according to the expected losses of their injury types NCCI derives an overall ELF, which yields a mixed severity distribution. How does an overall collective risk model compare with the sum of five injury-type collective risk models? This was, for the author, the genesis of the problem of mixing collective risk models, which we will henceforth treat abstractly.

We will consider *k* loss models, $S_1 \ldots S_k$, and their mixture S_0:

$$\begin{array}{ll} S_1 = X_{11} + \ldots + X_{1N_1} & E[N_1] = \mu_1 \\ \vdots & \\ S_k = X_{k1} + \ldots + X_{kN_k} & E[N_k] = \mu_k \\ \hline S_0 = X_{01} + \ldots + X_{0N_0} & N_0 = \sum_{i=1}^{k} N_i \end{array}$$

Moreover, let $\mu_0 = E[N_0] = \sum_{i=1}^{k} \mu_i$. Since claim counts are non-negative integers, each μ is greater than or equal to zero, and a zero μ indicates a trivial model of no claims. We will assume that no model is trivial; hence, we can assign weights to each model, $\pi_i = \mu_i / \mu_0$, which together constitute a probability measure. The individual losses of the i^{th} model are independent, but identically distributed; so we will henceforth drop the column subscript and speak of a typical "severity" X_i, for $i \in \{0, 1, \ldots, k\}$. However, the distribution of X_0 is the mixture of the other severity distributions weighted according to their probabilities. All the X and N random variables are independent of each other, except for the obvious case that $N_0 = \sum_{i=1}^{k} N_i$.

Homogeneity arises from the models' having different severity distributions; we could and should combine models whose severity distributions are identical. Since X_0 is a mixture of the other X_i, the S_0 model is less homogenous than the summation of models $S_1, \ldots S_k$. But how does the reduction of homogeneity affect the distribution of S_0?

One might suppose that the reduction of homogeneity always affects S_0, and that $Var[S_0] < Var\left[\sum_{i=1}^{k} S_i\right] = \sum_{i=1}^{k} Var[S_i]$. However, at the 2008 Seminar on Reinsurance the author gave counterexamples to both suppositions.[1] First, he gave an example in which $Var[S_0] > Var\left[\sum_{i=1}^{k} S_i\right]$. Second and more surprising, he demonstrated that if the N_i are Poisson distributed, the moment generating function of S_0 is the same as that of $\sum_{i=1}^{k} S_i$, which implies the sameness of their probability distributions. So with Poisson claim counts, reducing homogeneity makes no difference to collective risk modeling. Although the author stated, "$Var[U] < Var[S + T]$ … might be the usual case in real insurance problems," he could produce no example. But subsequently he has found, and will show here, that if the N_i are negative-binomial distributed, $Var[S_0] \leq \sum_{i=1}^{k} Var[S_i]$, and realistically, that the inequality is strict.

So first, let us denote as NB(μ, c) the negative binomial random variable with mean μ and variance $\mu(1 + c)$, where μ and c are greater than zero. The parameter c is called the contagion. The characteristic of the negative binomial random variable N is that $Var[N] = \mu(1 + c) > \mu = E[N]$; the Poisson distribution is the limiting case of the negative binomial as $c \to 0^+$. The moment generating function of NB(μ, c) is:[2]

$$M_N(t) = \left(1 - c\left[e^t - 1\right]\right)^{-\mu/c}$$

---

[1] www.casact.org/education/reinsure/2008/handouts/halliwell.pdf. The formulation there was less general than that used here: the mixed model U was compared with homogenous models S and T.

[2] The reader may verify that $E[N] = M'_N(0) = \mu$ and that $E[N^2] = M''_N(0) = \mu(\mu + 1 + c)$, whence it follows that $Var[N] = \mu(1 + c)$.

As a check, $\lim_{c \to 0^+} M(t) = e^{\mu[e^t - 1]}$, the moment generating function of the Poisson(μ) random variable. Furthermore, the moment generating function of the sum of two independent negative binomial random variables with common contagion *c* is:

$$\left(1 - c\left[e^t - 1\right]\right)^{-\mu_1/c}\left(1 - c\left[e^t - 1\right]\right)^{-\mu_2/c} = \left(1 - c\left[e^t - 1\right]\right)^{-(\mu_1+\mu_2)/c}$$

Hence, for independent random variables, $\sum_i \mathrm{NB}(\mu_i, c) \sim \mathrm{NB}\left(\sum_i \mu_i, c\right)$.[3]

Next, as is well known from the actuarial syllabus, the formulas for the mean and the variance of the collective risk model $S = X_1 + \ldots + X_N$ are:

$$E[S] = E[N]E[X] = \mu E[X]$$
$$Var[S] = E[N]Var[X] + Var[N]E[X]^2 = \mu Var[X] + Var[N]E[X]^2$$

For our purpose we need only to modify these formulas with our subscripting convention. Furthermore, if the N_i are negative binomial with common contagion *c*:

$$Var[S_i] = \mu_i Var[X_i] + \mu_i(1 + c)E[X_i]^2 = \mu_i E\left[X_i^2\right] + c\mu_i E[X_i]^2$$

Finally, in order to compare S_0 with $\sum_{i=1}^{k} S_i$, we require formulas for the two moments of X_0. Since X_0 is a mixture of the other X_i weighted according to p_i:

$$E\left[X_0^n\right] = E_{\Theta}\left[E\left[X_0^n \middle| \Theta\right]\right] = \sum_{i=1}^{n} \pi_i E\left[X_i^n\right]$$

Therefore, after all this preparation, we make a first-moment comparison:

$$\begin{aligned} E[S_0] &= \mu_0 E[X_0] \\ &= \mu_0 \sum_{i=1}^{k} \pi_i E[X_i] \\ &= \sum_{i=1}^{k} \mu_i E[X_i] \\ &= \sum_{i=1}^{k} E[S_i] \end{aligned}$$

The mixed model preserves the first moment of the sum. This holds true regardless of both the probability distributions of the N_i and their independence.

But implicit in the first line of the second-moment comparison are all the assumptions (viz., independent negative binomial distributions with common contagion *c*):

[3] The symbol '~' means 'is distributed as'.

$$
\begin{aligned}
Var[S_0] &= \mu_0 E[X_0^2] + c\mu_0 E[X_0]^2 \\
&= \sum_{i=1}^{k} \mu_i E[X_i^2] + c\mu_0 E[X_0]^2 \\
&= \sum_{i=1}^{k} \left(Var[S_i] - c\mu_i E[X_i^2]\right) + c\mu_0 E[X_0]^2 \\
&= \sum_{i=1}^{k} Var[S_i] - c\sum_{i=1}^{k} \mu_i E[X_i^2] + c\mu_0 E[X_0]^2 \\
&= \sum_{i=1}^{k} Var[S_i] - c\mu_0 \left\{\sum_{i=1}^{k} \pi_i E[X_i^2] - E[X_0]^2\right\} \\
&= \sum_{i=1}^{k} Var[S_i] - c\mu_0 \left\{\sum_{i=1}^{k} \pi_i E[X_i^2] - \left(\sum_{i=1}^{n} \pi_i E[X_i]\right)^2\right\} \\
&= \sum_{i=1}^{k} Var[S_i] - c\mu_0 \cdot VHM \\
&\leq \sum_{i=1}^{k} Var[S_i]
\end{aligned}
$$

Therefore, according to our realistic assumptions, the variance of mixed model S_0 is less than or equal to that of the sum $\sum_{i=1}^{k} S_i$. Equality obtains only in the extremely trivial case that μ_0 equals zero, or in the Poisson limiting case that $c = 0$, or in the case that the mean severities of all the models are equal (or $VHM = 0$).[4] In the first and second cases the distributions of S_0 and $\sum_{i=1}^{k} S_i$ are identical. In the third case the distributions are the same if and only if *all* the moments of the X_i are equal, i.e., if and only if the severity distributions are identical. But as long as the mean severities of the S_i models are equal, the distributions of S_0 and $\sum_{i=1}^{k} S_i$ match to two moments. Aside from these three cases, the variance of the mixed model is strictly less.

A trick allows us to extend the inequality. The binomial distribution BN(n, p) has mean np and variance $np(1-p)$. Formally this is a negative binomial distribution with a negative contagion, i.e., $p = -c$. In fact, BN(n, p) is equivalent to NB(np, $-p$), as the moment generating function proves;

[4] In actuarial parlance *VHM* stands for 'variance of the hypothetical means'; here it can be interpreted as 'variance of the homogenous mean severities'. To the extent to which the severity distributions do not vary the mixed model suffers no reduction in homogeneity. The converse is not necessarily true, for a mixed model with Poisson claim counts fully preserves the sum of its parts.

$$M_{\mathrm{NB}(np,-p)}(t) = \left(1-(-p)\left[e^t-1\right]\right)^{-np/(-p)}$$
$$= \left(1+p\left[e^t-1\right]\right)^n = M_{\mathrm{BN}(n,p)}(t)$$

Hence, if the claim counts are BN(*n*, *p*) ~ NB(*np*, –*p*):

$$Var[S_0] = \sum_{i=1}^{k} Var[S_i] - c\mu_0 \cdot VHM$$
$$\sum_{i=1}^{k} Var[S_i] - (-p)\mu_0 \cdot VHM$$
$$= \sum_{i=1}^{k} Var[S_i] + p\mu_0 \cdot VHM$$
$$\geq \sum_{i=1}^{k} Var[S_i]$$

So, perhaps contrary to intuition, the variance of a collective risk model mixed from homogeneous collective risk models whose claim-count distributions are compatibly binomial is *greater* than the variance of the sum of the models.[5]

To conclude, the divergence of a mixed collective risk model from the sum of the models from which it is mixed is determined by the claim-count distributions, not by the severity distributions. The distribution of S_0 given that $N_0 = \sum_{i=1}^{k} N_i = 1$ is the distribution of X_0; the mixed severity is fully adequate to a one-claim outcome. The divergence arises from a plurality of claims. One may gain some insight from BN(n_i, *p*) claim counts: the claim count of the i^{th} model cannot exceed n_i.[6] However, the claim-count vector of the mixed model is in effect:

$$\begin{bmatrix} N_1^* \\ \vdots \\ N_k^* \end{bmatrix} \sim \text{MULTINOMIAL}\left[N_0 \sim \text{BN}[n_0, p], \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_n \end{bmatrix} \right]$$

[5] And, similarly to the negative binomial, one obtains the Poisson limiting case by keeping $E[N] = np$ constant as $p \to 0^+$, but discretely stepwise so as to maintain integer values for $n = E[N]/p$. Equivalently, let integer $n \to \infty$ and $p = E[N]/n$.

[6] Because the support of the binomial distribution is finite, it is theoretically unsuited for modeling most casualty insurance claims: the possible number of claims from a unit of exposure is unlimited. Binomial claim counts may suit life insurance, since there can be no more deaths than there are persons insured. Of course, for large *n* and small *p* the binomial distribution approximates the Poisson. Then again, a realistic claim-count distribution, though of infinite support, may not be negative binomial; e.g., it could be bimodal. How this affects our analysis is unknown; certainly it ruins the desirable property assumed herein of "divisibility," i.e., that sums of independent random variables of the same family remain in the family.

There is positive probability that $N_i^* > n_i$; the claim counts implicit in the mixed model can range more than those of the homogenous models. Thus a binomial distribution can splay more in the mixed model. On the other hand, the support of the negative binomial distribution is infinite. It seems to splay more in the separate models, especially when one considers that the Poisson distribution is the equilibrium.

Appendix

When Mixing preserves the Distribution

In our conclusion we remarked that the mixed severity is fully adequate to a one-claim outcome. So why can't a convolution argument be invoked for *n*-claim outcomes? Just where will a convolution argument take us? We will answer these questions in this technical appendix, which requires an excursion into random vectors.

First, let **X** be a *k*×1 random vector with probability distribution $f_{\mathbf{X}}(x_1,\ldots,x_k)$. The multivariate moment generating function, though still yielding a scalar, is the function of the *k*×1 vector t:[7]

$$M_{\mathbf{X}}(\mathrm{t}) = E\left[e^{\mathrm{t}'\mathbf{X}}\right] = E\left[e^{t_1X_1+\ldots+t_kX_k}\right]$$

Its partial derivatives evaluated at zero generate the joint moments of **X**:

$$\left.\frac{\partial^{n_1+\ldots n_k} M_{\mathbf{X}}(\mathrm{t})}{\partial t_1^{n_1}\ldots\partial t_k^{n_k}}\right|_{\mathrm{t}=0} = E\left[X_1^{n_1}\ldots X_k^{n_k}\right]$$

Furthermore, the moment generating function of a random vector whose elements are independent is the product of the moment generating functions:

$$M_{\mathbf{X}}(\mathrm{t}) = E\left[e^{t_1X_1+\ldots+t_kX_k}\right] = E\left[\prod_{i=1}^{k} e^{t_iX_i}\right] = \prod_{i=1}^{k} E\left[e^{t_iX_i}\right] = \prod_{i=1}^{k} M_{X_i}(t_i)$$

As with scalars, identity of multivariate moment generating functions implies identity of probability distributions.

Second, we introduce the notation $\mathbf{N} \oplus \mathbf{X}$ for the *k* collective risk models, considered as a *k*×1 random vector:

$$\begin{bmatrix} X_{11} + \ldots + X_{1N_1} \\ \vdots \\ X_{k1} + \ldots + X_{kN_k} \end{bmatrix}$$

[7] The roman and italic typefaces distinguish 't' as a vector from '*t*' as a scalar. One must be careful also to distinguish $M_{\mathbf{X}}(\mathrm{t}) = E\left[e^{t_1X_1+\ldots+t_kX_k}\right]$ from $M_{\mathbf{X}}(t \cdot 1_{\mathrm{k}}) = E\left[e^{tX_1+\ldots+tX_k}\right] = M_{X_1+\ldots+X_k}(t)$, where 1_k is the *k*×1 vector whose elements are ones.

As in the body of this paper, **X** is a $k\times1$ "severity" vector; the X_{ij} terms are independent instances of the severities. The $k\times1$ random claim-count vector **N** specifies the number of instances; however, we will henceforth relax the assumption that its elements N_i are independent of each other. The moment generating function of $\mathbf{N}\oplus\mathbf{X}$ is:

$$\begin{aligned}
M_{\mathbf{N}\oplus\mathbf{X}}(\mathrm{t}) &= E\left[e^{\mathrm{t}'(\mathbf{N}\oplus\mathbf{X})}\right] \\
&= E\left[e^{t_1\left(X_{11}+\ldots+X_{1N_1}\right)+\ldots+t_k\left(X_{k1}+\ldots+X_{kN_k}\right)}\right] \\
&= \underset{\mathbf{N}}{E}\left[E\left[e^{t_1\left(X_{11}+\ldots+X_{1N_1}\right)+\ldots+t_k\left(X_{k1}+\ldots+X_{kN_k}\right)}\middle|\mathbf{N}\right]\right] \\
&= \underset{\mathbf{N}}{E}\left[\prod_{i=1}^{k} E\left[e^{t_i\left(X_{i1}+\ldots+X_{iN_i}\right)}\middle|N_i\right]\right] \\
&= E\left[\prod_{i=1}^{k} M_{X_i}\left(t_i\right)^{N_i}\right] \\
&= E\left[e^{N_1 \ln M_{X_1}(t_1)+\ldots N_k \ln M_{X_k}(t_k)}\right] \\
&= M_{\mathbf{N}}\left(\begin{bmatrix}\ln M_{X_1}\left(t_1\right) \\ \vdots \\ \ln M_{X_k}\left(t_k\right)\end{bmatrix}\right)
\end{aligned}$$

And third, let **C** be what is known as the categorical random variable: C_i is Bernoulli with probability π_i, and $\sum_{i=1}^{k}\pi_i = 1$. This is a "switching" or "indicator" random variable; one of the elements of **C** must be one and the others zero. Its moment generating function is:

$$M_{\mathbf{C}}(\mathrm{t}) = E\left[e^{\mathrm{t}'\mathbf{C}}\right] = \sum_{i=1}^{k}\pi_i e^{t_i \cdot 1} = \sum_{i=1}^{k}\pi_i e^{t_i}$$

Obviously, the C_i are not independent; they must sum to one. As one can show from differentiation, $E\left[C_i\right] = \pi_i$, $Var\left[C_i\right] = \pi_i\left(1-\pi_i\right)$, and $Cov\left[C_i, C_{j\neq i}\right] = -\pi_i\pi_j$. Actually, the categorical distribution is a multinomial distribution of one trial. Therefore, the mixture of the severities of **X** is $\mathbf{C}'\mathbf{X}$, and its moment generating function is:

$$\begin{aligned} M_{\mathbf{C}'\mathbf{X}}(t) &= E\left[e^{t\cdot\mathbf{C}'\mathbf{X}}\right] \\ &= \underset{\Theta}{E}\left[E\left[e^{t\cdot\mathbf{C}'\mathbf{X}}\middle|\Theta\right]\right] \\ &= \sum_{i=1}^{k}\pi_i M_{X_i}(t) \\ &= M_{\mathbf{C}}\left(\begin{bmatrix}\ln M_{X_1}(t) \\ \vdots \\ \ln M_{X_k}(t)\end{bmatrix}\right) \\ &= M_{\mathbf{C}\oplus\mathbf{X}}(t\cdot 1_k) \end{aligned}$$

This amounts to a proof of our assertion that the mixed severity is fully adequate to a one-claim outcome; in symbols, $\mathbf{C}'\mathbf{X} \sim 1_k'(\mathbf{C}\oplus\mathbf{X})$

Finally, let the $k\times 1$ random vector $\mathbf{N}$ represent the sum of N categorical random variables:

$$\mathbf{N} = \mathbf{C}_1 + \ldots + \mathbf{C}_N = \begin{bmatrix}C_{11} \\ \vdots \\ C_{k1}\end{bmatrix} + \ldots + \begin{bmatrix}C_{1N} \\ \vdots \\ C_{kN}\end{bmatrix} = \begin{bmatrix}C_{11} + \ldots + C_{1N} \\ \vdots \\ C_{k1} + \ldots + C_{kN}\end{bmatrix} = (N\cdot 1_k)\oplus\mathbf{C}$$

This is like a group of k collective risk models whose claim counts are equal.[8] Because of the binary nature of the $\mathbf{C}_i$, $1_k'\mathbf{N} = N$. $\mathbf{N}$ is any claim-count random variable derivable from convolutions of $C(\pi_1, \ldots, \pi_k)$, and its moment generating function is:

$$M_{\mathbf{N}}(\mathrm{t}) = \underset{N}{E}\left[E\left[M_{\mathbf{N}}(\mathrm{t})\middle|N\right]\right] = \underset{N}{E}\left[M_{\mathbf{C}}(\mathrm{t})^N\right] = M_N(\ln M_{\mathbf{C}}(\mathrm{t}))$$

Consequently, the moment generating function of the k collective risk models $\mathbf{N}\oplus\mathbf{X}$ is:

$$M_{\mathbf{N}\oplus\mathbf{X}}(\mathrm{t}) = M_N\left(\ln M_{\mathbf{C}}\left(\begin{bmatrix}\ln M_{X_1}(t_1) \\ \vdots \\ \ln M_{X_k}(t_k)\end{bmatrix}\right)\right)$$

And the moment generating function for the sum of the models $1_k'(\mathbf{N}\oplus\mathbf{X})$ is:

[8] It fails to be a collective risk model in that the C_{ij} are not independent; hence, its moment generating function must be derived anew.

$$\begin{aligned}
M_{1_k'(\mathbf{N}\oplus\mathbf{X})}(t) &= M_{\mathbf{N}\oplus\mathbf{X}}(t\cdot 1_k) \\
&= M_N\left(\ln M_{\mathbf{C}}\left(\begin{bmatrix}\ln M_{X_1}(t)\\ \vdots \\ \ln M_{X_k}(t)\end{bmatrix}\right)\right) \\
&= M_N\left(\ln M_{\mathbf{C}'\mathbf{X}}(t)\right) \\
&= M_{N\oplus(\mathbf{C}'\mathbf{X})}(t)
\end{aligned}$$

From this follows the distributional identity $1_k'(\mathbf{N}\oplus\mathbf{X}) \sim N\oplus(\mathbf{C}'\mathbf{X}) = (1_k'\mathbf{N})\oplus(\mathbf{C}'\mathbf{X})$. In words, if **N** is a convolution of the categorical distribution (i.e., a multinomial distribution whose parameter for the number of trials is the claim-count scalar *N*), then the *k* homogenous models of $\mathbf{N}\oplus\mathbf{X}$ are reducible to one model of *N* claims and mixed severity $\mathbf{C}'\mathbf{X}$. It is truly unexpected that the severity distributions **X** have nothing to do with this reducibility; it all depends on the claim-count random vector **N**.

A broad family of suitable random vectors consists of negative binomial convolutions of $C(\pi_1, \ldots, \pi_k)$, i.e., $\mathbf{N} \sim (\text{NB}(\mu, c)\cdot 1_k)\oplus\mathbf{C}$. The moment generating function of **N** is:

$$\begin{aligned}
M_{\mathbf{N}}(\mathrm{t}) &= M_N\left(\ln M_{\mathbf{C}}(\mathrm{t})\right) \\
&= \left(1 - c\left[e^{\ln M_{\mathbf{C}}(\mathrm{t})} - 1\right]\right)^{-\mu/c} \\
&= \left(1 - c\left[M_{\mathbf{C}}(\mathrm{t}) - 1\right]\right)^{-\mu/c} \\
&= \left(1 - c\left[\sum_{i=1}^{k}\pi_i e^{t_i} - 1\right]\right)^{-\mu/c} \\
&= \left(1 - c\left[\sum_{i=1}^{k}\pi_i e^{t_i} - \sum_{i=1}^{k}\pi_i\right]\right)^{-\mu/c} \\
&= \left(1 - c\left[\sum_{i=1}^{k}\pi_i\left(e^{t_i} - 1\right)\right]\right)^{-\mu/c}
\end{aligned}$$

The marginal distribution of N_i, the i^{th} element of **N**, is:

$$\begin{aligned}
M_{N_i}(t) &= E\left[e^{tN_i}\right] \\
&= E\left[e^{0\cdot N_1 + \ldots + tN_i + \ldots 0\cdot N_k}\right] \\
&= \left(1 - c\left[\pi_i\left(e^t - 1\right)\right]\right)^{-\mu/c} \\
&= \left(1 - c\pi_i\left[\left(e^t - 1\right)\right]\right)^{-\frac{\mu\pi_i}{c\pi_i}} \\
&= M_{\text{NB}(\mu\pi_i,\, c\pi_i)}(t)
\end{aligned}$$

Hence, $N_i \sim \text{NB}(\mu\pi_i, c\pi_i)$. We know its mean and variance to be $\mu\pi_i$ and $\mu\pi_i(1+c\pi_i)$.

Similarly, the distribution of the sum of $N_i + N_j,\ i \neq j,$ is:

$$M_{N_i+N_j}(t) = E\left[e^{tN_i+tN_j}\right] = \left(1 - c\left[(\pi_i + \pi_j)(e^t - 1)\right]\right)^{-\mu/c} = M_{\text{NB}(\mu\pi_i+\mu\pi_j,\, c\pi_i+c\pi_j)}(t)$$

This does not contradict the earlier formula $\sum_i \text{NB}(\mu_i, c) \sim \text{NB}\left(\sum_i \mu_i, c\right)$, which formula assumed the independence of the negative binomial distributions. For here the elements of **N** are not necessarily independent. Although one can derive the covariances of the elements by twice differentiating $M_{\mathbf{N}}(\text{t})$, the following way is elegant and simpler:

$$\begin{aligned}
Cov\left[N_i, N_{j\neq i}\right] &= 2\cdot Cov\left[N_i, N_j\right]/2 \\
&= \left(Var\left[N_i + N_j\right] - Var\left[N_i\right] - Var\left[N_j\right]\right)/2 \\
&= \frac{\mu\{\pi_i + \pi_j\}(1 + c\{\pi_i + \pi_j\}) - \mu\pi_i(1 + c\pi_i) - \mu\pi_j(1 + c\pi_j)}{2} \\
&= \frac{\mu\{\pi_i + \pi_j\}c\{\pi_i + \pi_j\} - \mu\pi_i c\pi_i - \mu\pi_j c\pi_j}{2} \\
&= c\mu\pi_i\pi_j
\end{aligned}$$

In general, the ij^{th} element of $Var[\mathbf{N}]$ is $Cov\left[N_i, N_j\right] = \mu\left(\pi_i\delta_{ij} + c\pi_i\pi_j\right)$, where δ is the Kronecker delta. The off-diagonal correlation coefficient is $c\sqrt{\dfrac{\pi_i\pi_j}{(1+c\pi_i)(1+c\pi_j)}}$.

We saw that the binomial distribution can be treated as a negative binomial distribution with a negative contagion.[9] In this case, $c < 0$ and the claim counts of the k models are negatively correlated. In the case that $c = 0$, we know that the counts are Poisson($\mu_i = \mu\pi_i$) distributed. But furthermore,

$$M_{\mathbf{N}}(\text{t}) = \lim_{c\to 0}\left(1 - c\left[\sum_{i=1}^{k}\pi_i\left(e^{t_i} - 1\right)\right]\right)^{-\mu/c} = e^{\mu\sum_{i=1}^{k}\pi_i\left(e^{t_i}-1\right)} = \prod_{i=1}^{k} e^{\mu\pi_i\left(e^{t_i}-1\right)} = M_{\begin{bmatrix}\text{Poisson}(\mu_1)\\ \vdots \\ \text{Poisson}(\mu_k)\end{bmatrix}}(\text{t})$$

The product form indicates the elements of **N** not only to be Poisson, but also to be independent. If the N_i are independent, then only Poisson convolutions of the categorical distribution will allow for mixed-model equivalency. Now we can understand our earlier findings. The assumption of claim-count independence ignores the correlations of the N_i

[9] Of course, for the variance to be meaningful, $c \geq -1$. And $M_{N(\mu,-1)}(t) = \left(1 + \left[e^t - 1\right]\right)^{\mu} = 1\cdot e^{t\mu}$, which correctly represents the constant μ.

variables that provide for mixed-model equivalency. Binomial convolutions ($c < 0$) are thus given too much variance in the mixed model, and hence its variance is overstated. Likewise, negative binomial convolutions ($c > 0$) are given too little variance in the mixed model, which results in its understating the total variance. Lately, actuaries have sought to model correlated exposures. The windfall of mixed-model equivalence should invite them to test the top-down[10] modeling of claims as a convolution of the categorical distribution.[11]

---

[10] By "top-down" we mean that the sum N is logically prior to the summands N_i. A top-down simulation first simulates n total claims, which it then parcels into the n_i claims by simulating from a multinomial(n; $\pi_1, \dots, \pi_k$) distribution. The assumption that the N_i are independent comports with the reverse, or bottom-up, approach.

[11] We proved that claim-count distributions that are built from convoluting the categorical distribution allow for mixing; we did not prove that only such distributions allow for it. In other words, we proved sufficiency, but not necessity. However, necessity ("only if") follows from the fact that if a mixing scheme is not adequate to a one-claim outcome, it will not work for multi-claim outcomes. But if it does work for the one-claim outcome, we're back to the categorical distribution and its convolutions. There is no other starting point for the induction.