CUMULANTS OF CONVOLUTION—MIXED DISTRIBUTIONS

Alan Brown
Melbourne

1. Convolution—Mixed Distributions

Consider a risk process which is characterised by three stochastic variables

(1) the number of accidents, N,
(2) the number of claims per accident, C, and
(3) the amount of a claim, X.

Let Y be a random variable denoting the total loss in a given period. Suppose that

$$p_n = \text{Prob}\,(N = n) \qquad n = 0, 1, 2. \ldots$$

and

$$v_c = \text{Prob}\,(C = c \mid \text{an accident has occurred}) \qquad c = 1, 2, 3. \ldots$$

If P_r represents the probability that exactly r claims occur in the period, then Kupper [4] has shown on certain simplifying assumptions that

$$P_r = \sum_{n=0}^{\infty} p_n\, v_r^{*n} \tag{1}$$

where v_r^{*n}, the probability of exactly r claims in n accidents, is given by

$$v_r^{*n} = \sum_{c=n-1}^{r-1} v_c^{*(n-1)}\, v_{r-c} \qquad \text{for } r \geqslant n,\ n = 1, 2, 3. \ldots$$

and $v_r^{*n} = 0 \quad$ for $r < n$

Further

$$v_r^{*1} = v_r$$
$$v_r^{*0} = 1 \qquad \text{for } r = 0$$
and $v_r^{*0} = 0 \qquad$ for $r \neq 0$

Suppose that

$$F(x) = \text{Prob } (Y \leq x)$$

and $S(x) = \text{Prob } (X \leq x)$

The total loss can be expressed on certain simplifying assumptions by the well known formula

$$F(x) = \sum_{r=0}^{\infty} P_r \, S^{*r}(x) \tag{2}$$

where $S^{*r}(x)$, the r^{th} convolution of the distribution function $S(x)$, is given by

$$S^{*r}(x) = \int_0^x S^{*(r-1)}(x - z)dS(z) \qquad \text{for } r = 1, 2, 3 \ldots$$

$$S^{*1}(x) = S(x)$$

$$S^{*0}(x) = 1 \qquad \text{for } x \geq 0$$

$$S^{*0}(x) = 0 \qquad \text{for } x < 0$$

Combining equations (1) and (2) together we obtain

$$F(x) = \sum_{r=0}^{\infty} \sum_{n=0}^{\infty} p_n \, v_r^{*n} \, S^{*r}(x)$$

$$= \sum_{n=0}^{\infty} \sum_{r=0}^{\infty} p_n \, v_r^{*n} \, S^{*r}(x)$$

if we interchange the order of summation

Auxiliary Functions Associated with Probability Distributions

There are several useful auxiliary functions associated with a distribution function $F(x)$ of the random variable Y (see [3])

(1) Probability generating function

$$G_Y(z) = E_Y(z^x) = \int_{-\infty}^{\infty} z^x \, dF(x) \qquad (z \text{ real, positive})$$

(2) Moment generating function

$$M_Y(u) = E_Y(e^{ux}) = \int_{-\infty}^{\infty} e^{ux} \, dF(x) \quad (u \text{ real})$$

(3) Characteristic function

$$\phi_Y(t) = E_Y(e^{itx}) = \int_{-\infty}^{\infty} e^{itx} \, dF(x) \quad (t \text{ real})$$

(4) Cumulant generating function

$$K_Y(u) = \log M_Y(u)$$

Provided the various integrals exist we can change from one auxiliary function to another by the transformations

$$u = it = \log z$$

For instance $\quad G_Y(e^u) = M_Y(u)$

and $\quad K_Y(it) = \log M_Y(it)$

$$= \log \phi_Y(t)$$

The Application of Generating Functions to Convolution—Mixed Distributions

We depend heavily on the following well-known (see [3])

Lemma

If $X_1, X_2 \ldots . X_n$ are independent and identically distributed random variables

and $\quad Z = X_1 + X_2 + \ldots\ldots + X_n$

then $\quad G_Z(u) = [G_X(u)]^n$

Now from equation (3) we have

$$\begin{aligned} G_Y(z) &= \sum_{n=0}^{\infty} \sum_{r=0}^{\infty} p_n \, v_r^{*n} \, S^{*r}(x) \, z^x \\ &= \sum_{n=0}^{\infty} \sum_{r=0}^{\infty} p_n \, v_r^{*n} \, G_{X_1+X_2+\ldots.X_r}(z) \\ &= \sum_{n=0}^{\infty} \sum_{r=0}^{\infty} p_n \, v_r^{*n} \, [G_X(z)]^r \\ &= \sum_{n=0}^{\infty} p_n \, G_{C_1+C_2+\ldots.+C_n}(G_X(z)) \\ &= \sum_{n=0}^{\infty} p_n \, [G_C(G_X(z))] \\ &= G_N(G_C(G_X(z))) \end{aligned}$$

Thyrion [5] has introduced a very wide class of distributions, the distributions in a bunch ($m = 2$), and in a bunch of bunches ($m > 2$), defined by generating functions in the following general form

$$G_Y(z) = G_1(G_2(G_3, \ldots . G_{m-1}(G_m(z)) \ldots .)) \qquad m \geqslant 2$$

where $G_j(z)$ are probability generating functions of integer valued variables, $j = 1$ to $(m - 1)$, and $G_m(z)$ is any probability generating function.

A special case where the G_j, $j = 1$ to m are all identical, occurs in the theory of branching processes, where Y is the size of the m^{th} generation. The principal result of this paper is contained in the following theorem, which is a generalisation of a known result in the theory of branching processes (see [2]).

Theorem

If $\qquad G_Y(z) = G_N(G_C(G_X(z)))$

then $\qquad K_Y(n) = K_N(K_C(K_X(u))) \qquad (4)$

Proof

$$
\begin{aligned}
\text{Let} \qquad u &= \log z \\
\text{then} \qquad M_Y(u) &= G_Y(z) \\
&= G_N(G_C(G_X(z))) \\
&= G_N(G_C(M_X(u))) \\
&= G_N(G_C(e^{\log M_X(u)})) \\
&= G_N(M_C(K_X(u))) \\
&= G_N(e^{\log M_c(K_X(u))}) \\
&= M_N(K_C(K_X(u))) \\
\text{so that} \qquad K_Y(u) &= K_N(K_C(K_X(u))) \qquad \text{as required}
\end{aligned}
$$

This theorem can obviously be extended to include the distributions, a bunch of bunches. By differentiating the cumulant generating function and setting u = 0 we can obtain the cumulants of a distribution. Using an obvious notation we can derive the following relationships between the cumulants of a low order from equation (4).

$$\kappa_{1Y} = \kappa_{1N}\,\kappa_{1C}\,\kappa_{1X} \tag{5}$$

$$\kappa_{2Y} = \kappa_{2N}\,\kappa_{1C}^2\,\kappa_{1X}^2 + \kappa_{1N}\,\kappa_{2C}\,\kappa_{1X}^2 + \kappa_{1N}\,\kappa_{1C}\,\kappa_{2X} \tag{6}$$

$$
\begin{aligned}
\kappa_{3Y} = {}& \kappa_{3N}\,\kappa_{1C}^3\,\kappa_{1X}^3 + 3\kappa_{2N}\,\kappa_{1C}\,\kappa_{2C}\,\kappa_{1X}^3 + 3\kappa_{2N}\,\kappa_{1C}\,\kappa_{2X}\,\kappa_{1X} \\
& + \kappa_{1N}\,\kappa_{3C}\,\kappa_{1X}^3 + 3\kappa_{1N}\,\kappa_{2C}\,\kappa_{2X}\,\kappa_{1X} + \kappa_{1N}\,\kappa_{1C}\,\kappa_{3X}
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
\kappa_{4Y} = {}& \kappa_{4N}\,\kappa_{1C}^4\,\kappa_{1X}^4 + 6\kappa_{3N}\,\kappa_{2C}\,\kappa_{1C}^2\,\kappa_{1X}^4 + 6\kappa_{3N}\,\kappa_{1C}^3\,\kappa_{2X}\,\kappa_{1X}^2 \\
& + 4\kappa_{2N}\,\kappa_{3C}\,\kappa_{1C}\,\kappa_{1X}^4 + 3\kappa_{2N}\,\kappa_{2C}^2\,\kappa_{1X}^4 + 18\kappa_{2N}\,\kappa_{2C}\,\kappa_{1C}\,\kappa_{2X}\,\kappa_{1X}^2 \\
& + 4\kappa_{2N}\,\kappa_{1C}^2\,\kappa_{3X}\,\kappa_{1X} + 3\kappa_{2N}\,\kappa_{1C}^2\,\kappa_{2X}^2 \\
& + \kappa_{1N}\,\kappa_{4C}\,\kappa_{1X}^4 + 6\kappa_{1N}\,\kappa_{3C}\,\kappa_{2X}\,\kappa_{1X}^2 + 4\kappa_{1N}\,\kappa_{2C}\,\kappa_{3X}\,\kappa_{1X} \\
& + 3\kappa_{1N}\,\kappa_{2C}\,\kappa_{2X}^2 + \kappa_{1N}\,\kappa_{1C}\,\kappa_{4X}
\end{aligned} \tag{8}
$$

These formulae, given in equations (5)-(8) can be used in the normal power expansion [1]

$$F(x) = \Phi(y)$$

where $\Phi(y)$ is the cumulative Normal distribution and

$$\frac{x - \varkappa_{1Y}}{(\varkappa_{2Y})^{1/2}} = y + \frac{\varkappa_{3Y}}{6(\varkappa_{2Y})^{3/2}}(y^2 - 1) + \frac{\varkappa_{4Y}}{24\varkappa_{2Y}^2}(y^3 - 3y) + \frac{\varkappa_{3Y}^2}{36\varkappa_{2Y}^3}(2y^3 - 5y) + \ldots \quad (9)$$

In particular if the number of accidents, N, has a Poisson distribution with expected value λt, where λ is a constant, then the cumulants

$$\varkappa_{jN} = \lambda t \qquad \text{for all } j > 0$$

It follows that

$$\varkappa_{jY} = o(t) \qquad \text{for all } j > 0$$

which is all that is required to establish the validity of the asymptotic expansion (9) for large values of t.

References

[1] Beard, R. E., Pentikainen T. & Pesonen, E., 1969. Risk Theory.
[2] Cox, D. R. & Miller, H. D., 1965. The Theory of Stochastic Processes.
[3] Kendall, M. G. & Stuart, A., 1963. The Advanced Theory of Statistics.
[4] Kupper, J., 1963. Some Aspects of the Cumulative Risk, Astin Bulletin, III, 1.
[5] Thyrion, P., 1960. Note sur les distributions 'par grappes', Bulletin de l'Association Royale des Actuaries Belges.