ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XII, 2 (1971)

ADRIENNE W. KEMP and C. D. KEMP (Bradford)

ON MIXING PROCESSES AND THE LOST-GAMES DISTRIBUTION

1. Introduction. The lost-games distribution arises from stochastic processes associated with the gambler's ruin problem, with McKendrick's epidemic in an infinite population and with the busy period for the $M/M/1$ queue. Kemp and Kemp [7] discussed the relationships between the three random walks corresponding to these processes. Properties of the distribution appear in the same paper.

More recently Kemp and Kemp [8] have examined the slightly more general distribution with probability generating function

$$K_{a,j}(s) = s^j p^a {}_2F_1[a/2, (a+1)/2; a+1; 4pqs], \tag{1}$$

where $p \geqslant q\ (=1-p)$ and $0 \leqslant j \leqslant a$, a and j integers ($j = a$ gives the lost-games distribution). Various branching and clustering processes giving rise to $K_{a,j}(s)$ distributions were investigated. In particular, it was shown that the $K_{a,0}(s)$ distribution arises from binomial, from negative binomial and from Poisson clustering processes. That is to say that when the lost-games distribution is shifted so as to start at the origin, then it becomes a generalized Poisson distribution in the sense of Gurland [4], and so has the important property of infinite divisibility.

The present paper is concerned with mixing processes leading to $K_{a,j}(s)$ distributions. Mixed negative binomial, mixed Poisson and mixed confluent hypergeometric models exist. In particular, the lost-games distribution, when shifted so as to start at the origin, is then a compound Poisson distribution in the sense of Gurland [4], and so provides yet another example of a distribution which is both a generalized Poisson and a compound Poisson distribution.

2. The special case $p = q$. The special case $p = q$, $j = 0$, yields the distribution with p.g.f. (probability generating function)

$$G(s) = {}_2F_1[a/2, (a+1)/2; a+1; s]/2^a,$$

which is in turn a special case of Kemp and Kemp's [6] generalized hypergeometric type IV distribution (known also as the negative binomial beta distribution, and also as the generalized Waring distribution, Irwin [5]).

A number of mixing models leading to the type IV distribution are known; models leading to our special case ($p = q$) distribution may be derived from these. Further models may be obtained in a similar manner.

The name negative binomial beta comes from a derivation which seems to have been given first by Kemp and Kemp [6], It comprises compounding the inverse sampling form of the negative binomial distribution with a beta distribution:

$$\int_0^1 \left(\frac{1-x}{1-xs}\right)^k \frac{x^{c-1}(1-x)^{d-1}dx}{B(c, d)}$$

$$= {}_2F_1[k, c; k+c+d; s]/{}_2F_1[k, c; k+c+d; 1].$$

A shifted form with p.g.f.

$$\int_0^1 s^k \left(\frac{1-x}{1-xs}\right)^k \frac{x^{c-1}(1-x)^{d-1}dx}{B(c, d)}$$

is Raiffa and Schlaifer's [9] beta-Pascal distribution.

Taking $k = a/2$, $c = (a+1)/2$, $d = \frac{1}{2}$ gives

$$G(s) = \int_0^1 \left(\frac{1-x}{1-xs}\right)^{a/2} \frac{x^{(a-1)/2}(1-x)^{-1/2}dx}{B\left((a+1)/2, \frac{1}{2}\right)};$$

alternatively, taking $k = (a+1)/2$, $c = a/2$, $d = \frac{1}{2}$ gives

$$G(s) = \int_0^1 \left(\frac{1-x}{1-xs}\right)^{(a+1)/2} \frac{x^{a/2-1}(1-x)^{-1/2}dx}{B(a/2, \frac{1}{2})}.$$

Kemp and Kemp [6] considered the effect of applying the transformation $y = x/(1-x)$ to their compound inverse sampling model. This leads to compounding the "compounded accident proneness" form of the negative binomial distribution, having p.g.f. $(1+y-ys)^{-k}$, with an F-distribution,

$$\int_0^\infty (1+y-ys)^{-k} \frac{y^{c-1}(1+y)^{-c-d}dy}{B(c, d)},$$

(where $F = yd/c$), and yields Irwin's [5] compounded liability, compounded proneness accident model. Again taking $k = a/2$, $c = (a+1)/2$, $d = \frac{1}{2}$ gives

$$G(s) = \int_0^\infty (1+y-ys)^{-a/2} \frac{y^{(a-1)/2}(1+y)^{-a/2-1}dy}{B\big((a+1)/2, \frac{1}{2}\big)},$$

and $k = (a+1)/2$, $c = a/2$, $d = \frac{1}{2}$ gives

$$G(s) = \int_0^\infty (1+y-ys)^{-(a+1)/2} \frac{y^{(a/2-1)}(1+y)^{-(a+1)/2}dy}{B\,(a/2, \frac{1}{2})}.$$

These negative binomial models for the type IV distribution imply that it also results from a mixed Poisson model, see Gurland [4]. Such a model has been given by Dacey [2],

$$\int_0^\infty e^{x(s-1)} \frac{\Gamma(c+d)\,\Gamma(k+d)}{\Gamma(c)\,\Gamma(d)\,\Gamma(k)}\; e^{x/2} x^{(c+k-3)/2} W_{(1-c-2d-k)/2,(c-k)/2}(x)\,dx,$$

where $W_{\varkappa,\mu}$ is a Whittaker function. As we would expect from the uniqueness of the Laplace transform, both substitutions, $k = a/2$, $c = (a+1)/2$ $d = \frac{1}{2}$, and $k = (a+1)/2$, $c = a/2$, $d = \frac{1}{2}$, lead to the same model for $G(s)$, namely

$$G(s) = \int_0^\infty e^{x(s-1)} \frac{ae^{x/2}x^{a/2-5/4}}{2\sqrt{\pi}}\; W_{-\frac{1}{2}a-\frac{1}{4},\frac{1}{4}}(x)\,dx$$

$$= \int_0^\infty e^{x(s-1)} \frac{ax^{a/2-1}}{2\sqrt{\pi}}\; \Psi\big((a+1)/2, \tfrac{1}{2}; x\big)\,dx,$$

where $\Psi(a, \beta; x)$ is a confluent hypergeometric function of the second kind.

The type IV distribution may also be obtained as a mixed confluent hypergeometric distribution. Bhattacharya [1] obtained the negative binomial in the following manner:

$$\int_0^\infty \frac{{}_1F_1(b; d; xs)}{{}_1F_1(b; d; x)} \frac{a^b(a+1)^{d-b}x^{d-1}e^{-(a+1)x}}{\Gamma(d)}\; {}_1F_1(b; d; x)\,dx$$

$$= \frac{{}_1F_0[b; \cdot; s/(a+1)]}{{}_1F_0[b; \cdot; 1/(a+1)]} = [1+(1-s)/a]^{-b}.$$

More generally, we can show that

$$\int_0^\infty \frac{{}_1F_1(b; d; xs)}{{}_1F_1(b; d; x)} \frac{(a+1)^c x^{c-1} e^{-(a+1)x} {}_1F_1(b; d; x)\,dx}{\Gamma(c)\, {}_2F_1[b, c; d; 1/(a+1)]}$$

$$= {}_2F_1[b, c; d; s/(a+1)]/{}_2F_1[b, c; d; 1/(a+1)],$$

and $a = 0$ yields the type IV distribution. The substitutions $b = a/2$, $c = (a+1)/2$, $d = a+1$, and $b = (a+1)/2$, $c = a/2$, $d = a+1$, then give the following models for our special case ($p = q$) distribution:

$$G(s) = \int_0^\infty \frac{{}_1F_1[a/2; a+1; xs]}{{}_1F_1[a/2; a+1; x]} \frac{e^{-x} x^{(a-1)/2}}{2^a \Gamma\big((a+1)/2\big)} {}_1F_1[a/2; a+1; x]\,dx$$

and

$$G(s) = \int_0^\infty \frac{{}_1F_1[(a+1)/2; a+1; xs]}{{}_1F_1[(a+1)/2; a+1; x]} \frac{e^{-x} x^{a/2-1}}{2^a \Gamma(a/2)} {}_1F_1[(a+1)/2; a+1; x]\,dx.$$

It is possible to show for all these models that the mixing distributions are indeed valid probability distributions.

3. The general case. The models given in the last section for the special case $p = q$ $(= \frac{1}{2})$ may all be augmented so as to yield models for the general case. We list these models below and restrict ourselves to brief comments.

(i)

$$K_{a,j}(s) = \int_0^1 s^j \left(\frac{1-4pqx}{1-4pqxs}\right)^{a/2} \frac{(2p)^a x^{(a-1)/2} (1-x)^{(a-1)/2}}{B\big((a+1)/2, \frac{1}{2}\big)} (1-4pqx)^{-a/2}\,dx.$$

This model will be valuable in inverse sampling for a fixed number of successes where the probability of success varies and has an upper limit $4pq$, above which it may be regarded as negligible. The mixing distribution can be obtained from the Leipnik distribution

$$dF = \frac{(1-r^2)^{(a-1)/2} (1+\varrho^2-2\varrho r)^{-a/2}}{B\big((a+1)/2, \frac{1}{2}\big)}\,dr$$

by the simple linear transformation $x = (1+r)/2$, also $p = 1/(1+\varrho)$. The mean probability of success is therefore $4pq[\frac{1}{2} + aq/2p(a+2)]$, i.e. $2q[1-2q/(a+2)]$.

(ii)

$$K_{a,j}(s) = \int_0^1 s^j \left(\frac{1-4pqx}{1-4pqxs}\right)^{(a+1)/2} \frac{(2p)^a x^{a/2-1} (1-x)^{a/2} (1-4pqx)^{-(a+1)/2}}{B(a/2, \frac{1}{2})}\,dx.$$

This model will be useful in situations similar to those for model (i). The mixing distribution is closely allied to the Leipnik distribution. The maximum probability of success is again $4pq$, and the mean probability of success is now $2qa/(a+1)$.

(iii) $K_{a,j}(s)$

$$=\int_0^\infty s^j\left(\frac{1+y-4qpy}{1+y-4pqys}\right)^{a/2}\frac{(2p)^a y^{(a-1)/2}(1+y)^{-a/2-1}(1+y-4pqy)^{-a/2}}{B\big((a+1)/2,\frac{1}{2}\big)}\,dy.$$

(iv) $K_{a,j}(s)$

$$=\int_0^\infty s^j\left(\frac{1+y-4pqy}{1+y-4pqys}\right)^{(a+1)/2}\frac{(2p)^a y^{a/2-1}(1-y)^{-(a+1)/2}(1+y-4pqy)^{-(a+1)/2}}{B(a/2,\frac{1}{2})}\,dy.$$

Applications for models (iii) and (iv) seem unlikely.

(v) $$K_{a,j}(s)=\int_0^\infty s^j e^{4pqx(s-1)}\frac{(2p)^a a x^{a/2-1}e^{(4pq-1)x}}{2\sqrt{\pi}}\,\Psi\big((a+1)/2,\tfrac{1}{2};x\big)\,dx;$$

it is possible to show that the probability density function of the mixing distribution is equivalent to

$$e^{(4pq-1)x}ap^a 8^{a/2}x^{a/2-1}I_a(\sqrt{2x})\,dx,$$

where $I_a(u)$ is Fisher's [3] repeated integral of the normal probability integral, and hence is a valid p.d.f. This model is theoretically very important as it enables us to conclude that the lost-games distribution, when shifted so as to start at the origin, is both a generalized and a compound Poisson distribution. A number of other distributions (e.g. negative binomial, Neyman type A, Poisson-Pascal) have this dual property; the resultant flexibility of such distributions leads to their frequent use in summarising observed data.

(vi) $K_{a,j}(s)$

$$=\int_0^\infty s^j\frac{{}_1F_1(a/2;a+1;4pqxs)}{{}_1F_1(a/2;a+1;4pqx)}\,p^a e^{-x}x^{(a-1)/2}\frac{{}_1F_1(a/2;a+1;4pqx)}{\Gamma(a+1)/2)}\,dx.$$

(vii) $K_{a,j}(s)$

$$= \int_0^\infty s^j \frac{{}_1F_1\big((a+1)/2; a+1; 4pqxs\big)}{{}_1F_1\big((a+1)/2; a+1; 4pqx\big)} \frac{p^a e^{-x} x^{a/2-1}}{\Gamma(a/2)} {}_1F_1\big((a+1)/2; a+1; 4pqx\big)\,dx.$$

Here the p.d.f. of the mixing distribution is equivalent to

$$a(p/q)^{a/2} \frac{e^{(2pq-1)x}}{2x} I_{a/2}(2pqx),$$

where $I_{a/2}(v)$ no longer represents Fisher's repeated integral, but instead is the modified Bessel function of the first kind; this is the busy period time length distribution for the $M/M/1$ queue with service and arrival rates equal to p^2 and q^2, respectively. When $a = j = 1$, the distribution that is being mixed is the zero truncated Poisson distribution.

References

[1] S. K. Bhattacharya, *Confluent hypergeometric distributions of discrete and continuous type with applications to accident proneness*, Calcutta Statist. Assn. Bull. 15 (1966), p. 20-31.

[2] M. F. Dacey, *A hypergeometric family of discrete probability distributions: properties and applications to location models*, Geographical Analysis 1 (1969), p. 283-317.

[3] R. A. Fisher, *Properties of the functions* (Part of Introduction) British Association Mathematical Tables 1 (1931), p. XXX.

[4] J. Gurland, *Some interrelations among compound and generalized distributions*, Biometrika 44 (1957), p. 265-268.

[5] J. O. Irwin, *The generalized Waring distribution applied to accident theory*, J. Roy. Stat. Soc. A 131 (1968), p. 205-225.

[6] C. D. Kemp and A. W. Kemp, *Generalized hypergeometric distributions*, J. Roy. Stat. Soc. B 18 (1956), p. 202-211.

[7] A. W. Kemp and C. D. Kemp, *On a distribution associated with certain stochastic processes*, J. Roy. Stat. Soc. B 30 (1968), p. 160-163.

[8] A. W. Kemp and C. D. Kemp, *Branching and clustering models associated with the 'lost-games distribution'*, J. Appl. Prob. 6 (1969), p. 700-703.

[9] H. Raiffa and R. Schlaifer, *Applied statistical decision theory*, Graduate School of Business Administration, Harvard University, Boston 1961.

SCHOOL OF MATHEMATICS
UNIVERSITY OF BRADFORD
BRADFORD, ENGLAND

Received on 12. 5. 1970

ADRIENNE W. KEMP i C. D. KEMP (Bradford)

O PROCESACH STOCHASTYCZNYCH PROWADZĄCYCH DO ROZKŁADU TYPU „LOST-GAMES"

STRESZCZENIE

Autorzy pokazali w poprzednich pracach, że trzy typy znanych procesów stochastycznych, a także pewne procesy kaskadowe i zbierające (*clustering*), prowadzą do rozkładów o funkcji tworzącej prawdopodobieństwa postaci (1). Obecnie pokazano, że istnieją mieszane modele typu ujemnego binomialnego, Poissona i hipergeometrycznego, które prowadzą do tych samych rozkładów.

Przypadek specjalny $j = 0$ daje rozkład typu *lost-games* (tj. rozkład długości gry w zadaniu o ruinie gracza), przesunięty do początku układu współrzędnych. Wysnuwa się z tego wniosek, że jest to jeszcze jeden przykład rozkładu, który jest zarówno uogólnionym, jak i złożonym rozkładem Poissona.