Sankhyā : The Indian Journal of Statistics
San Antonio Conference: Selected articles
2002, Volume 64, Series A, Pt. 2, pp 293-305

CERTAIN q-ANALOGUES OF THE BINOMIAL DISTRIBUTION

By ADRIENNE W. KEMP
University of St Andrews, Scotland

SUMMARY. The paper studies the relationships between various q-analogues of the binomial distribution, both extant and new, using the now generally accepted Gasper-Rahman notation. Some properties and models are explored. Two fitted examples are given.

1. **Introduction**

The Poisson and binomial distributions are central to discrete distribution theory. The q-Poisson distribution appears in both the physics and the statistics literatures. Biedenhahn (1989), see also Macfarlane (1989), obtained it as the energy distribution in the q-coherent state in the theory of the q-analogue of the quantum harmonic oscillator. Benkherouf and Bather (1988) called it the Euler distribution and showed that this and an alternative q-analogue of the Poisson distribution (their Heine distribution) are feasible prior distributions for the number of undiscovered sources of oil; see also Kemp (1992a,b) and Benkherouf and Alzaid (1993).

Various q-analogues of the classical binomial distribution have also been derived. Dunkl (1981) saw that the distribution arising from direct absorption sampling is a binomial q-analogue. Its probability mass function (pmf) is

$$\Pr[X_{DA} = x] = q^{(n-x)(m-x)} \frac{(q^{m-x+1};q)_x (q^{n-x+1};q)_x}{(q;q)_x}, \qquad (1)$$

where $x = 0, 1, \ldots, \min(m,n)$, $(u;q)_0 = 1$, $(u;q)_x = (1-u)(1-uq)\cdots(1-uq^{x-1})$, m is the fixed (finite) size of the population of items, n is the fixed

Paper received October 2000; revised November 2001.
AMS (2000) subject classification. Primary 60E10, 62E15; secondary 33D15.
Keywords and phrases. q-binomial distribution, direct absorption sampling, log-linear odds, log-linear probabilities, q-deformed binomial state, Euler distribution, q-Poisson distribution, Rogers-Szegö polynomial, Stieltjes-Wigert polynomial, Jackson transformation, Wall polynomial, conditionality models, stationary state-dependant processes.

number of time periods, and q is the probability that a particular item fails to be sampled on a particular occasion, $0 < q < 1$. (The symmetry in m and n is surprising.) For properties, alternative models, and a characterization see Kemp (1998, 2001).

Kemp's (1987) treatment of a weapon defense problem yielded the distribution with probability generating function (pgf)

$$G_P(z) = \prod_{i=0}^{m-1} (1 - cq^i + cq^i z), \quad 0 < c < 1,\ 0 < q < 1. \tag{2}$$

This tends to a binomial as $q \to 1$ and to the Euler as $m \to \infty$. Gani (1974) had earlier encountered the case $c = 1$. The complicated probabilities make this distribution incompatible with the others in this paper.

Kemp and Kemp's (1991) q-analogue of the binomial distribution replaced the loglinear relationship for the Bernoulli probabilities in Poissonian binomial sampling in (2) with a loglinear odds relationship. Kemp and Kemp stated the pmf in the form

$$\Pr[X_{KB} = x] = \frac{(q;q)_m\ \theta^x q^{x(x-1)/2}}{(q;q)_x (q;q)_{m-x}} \Bigg/ \prod_{j=1}^{m} (1 + \theta q^{j-1})\ , \quad x = 0, 1, \ldots, m, \tag{3}$$

where $0 < \theta$, $0 < q < 1$. Kemp and Newton (1990) obtained the distribution via a state-dependent stochastic process.

In physics there is much interest in the q-deformed boson oscillator and its applications. This led Jing (1994) to extend the concept of the ordinary binomial state to the q-deformed binomial state by developing a q-analogue of the binomial distribution with pmf

$$\Pr[X_{QD} = x] = \frac{(q;q)_m (\tau;q)_{m-x} \tau^x}{(q;q)_x (q;q)_{m-x}}, \quad 0 < q < 1,\ 0 < \tau < 1. \tag{4}$$

This limits to a q-Poisson (Euler) distribution as $m \to \infty$. Jing and Fan (1994) have also interpreted the pmf (3) in terms of a q-deformed binomial state.

The objective in this paper is to develop further q-analogues of the binomial distribution and to clarify their relationships to the above distributions via the now generally accepted Gasper-Rahman notation. Their shape, factorial moments, conditionality models, and stationary birth-death modes of genesis are explored. Two fitted examples are given.

2. **Notation**

The publication of the book on basic hypergeometric functions by Gasper and Rahman (G/R) (1990) has led to the universal adoption in the mathematical and physics literature of a new notation for q-series, with the advantage that limiting forms of G/R q-series are also G/R q-series. In the new notation,

$$\begin{aligned} &{}_A\phi_B(a_1,\dots,a_A;b_1,\dots,b_B;q,z)\\ &= \sum_{j=0}^{\infty} \frac{(a_1;q)_j\dots(a_A;q)_j z^j}{(b_1;q)_j\dots(b_B;q)_j(q;q)_j}\left[(-1)^j q^{\binom{j}{2}}\right]^{B-A+1}, \end{aligned}$$

where $0 \le q \le 1$, $(u;q)_0 = 1$, $(u;q)_x = (1-u)(1-uq)\cdots(1-uq^{x-1})$. Unfortunately this improved notation conflicts with the notation in the books by Bailey (1935) and Slater (1966) which was used in the early papers on q-Poisson and q-binomial analogues. The definition of the Gaussian binomial coefficient is unchanged,

$$\begin{bmatrix} m \\ 0 \end{bmatrix}_q = 1, \qquad \begin{bmatrix} m \\ x \end{bmatrix}_q = \frac{(q;q)_m}{(q;q)_x(q;q)_{m-x}}.$$

The pgf for the (underdispersed) Heine distribution becomes

$$G_H(z) = \frac{{}_0\phi_0(-;-;q,-\lambda z)}{{}_0\phi_0(-;-;q,-\lambda)}, \quad 0<q<1,\ 0<\lambda,$$

emphasizing its relationship to the Poisson distribution. The pgf for the (overdispersed) Euler distribution has an extra parameter equal to 0:

$$G_E(z) = \frac{{}_1\phi_0(0;-;q,\eta z)}{{}_1\phi_0(0;-;q,\eta)}, \quad 0<q<1,\ 0<\eta<1.$$

The pgf of Kemp and Kemp's (1991) distribution,

$$G_{KB}(z) = \frac{{}_1\phi_0(q^{-m};-;q,-q^m\theta z)}{{}_1\phi_0(q^{-m};-;q,-q^m\theta)}, \quad 0<q<1,\ 0<\theta, \tag{5}$$

is an analogue of the $G(z) = {}_1F_0(-m;-;-\theta z)/{}_1F_0(-m;-;-\theta)$, $0<\theta$, form of the binomial pgf, whereas (2) is analogous to $G(z) = {}_1F_0[-m;-;c(1-z)]$, $c=\theta/(1+\theta)$.

Kemp (1998) showed that Dunkl's pgf is

$$G_{DA}(z) = q^{mn}{}_2\phi_1(q^{-m},q^{-n};0;q,qz). \tag{6}$$

It is straightforward to show that Jing's q-deformed pgf can be reexpressed as

$$G_{QD}(z) = (\tau;q)_m \; {}_2\phi_1(q^{-m}, 0; q^{1-m}\tau^{-1}; q, qz). \tag{7}$$

3. The New q-binomial Distributions

Further q-analogues of the binomial distribution can be constructed by introducing extra parameters into the q-series formulation of the Kemp and Kemp q-binomial pgf and setting these equal to zero, e.g.

$$G_{RS}(z) = \frac{{}_2\phi_0(q^{-m}, 0; -; q, q^m\theta z)}{{}_2\phi_0(q^{-m}, 0; -; q, q^m\theta)}, \tag{8}$$

$$G_{SW}(z) = \frac{{}_1\phi_1(q^{-m}; 0; q, q^m\theta z)}{{}_1\phi_1(q^{-m}; 0; q, q^m\theta)}, \tag{9}$$

where $0 < q < 1,\ 0 < \theta$. When stated in terms of the Gaussian binomial coefficient their probabilities differ only by powers of q from the Kemp and Kemp q-binomial probabilities; for (8) and (9)

$$\Pr[X_{RS}=x] = C_{RS} \begin{bmatrix} m \\ x \end{bmatrix}_q \theta^x,$$

$$\Pr[X_{SW}=x] = C_{SW} \begin{bmatrix} m \\ x \end{bmatrix}_q q^{x(x-1)}\theta^x,$$

respectively, where C is a normalizing constant. For (5)

$$\Pr[X_{KB}=x] = C_{KB} \begin{bmatrix} m \\ x \end{bmatrix}_q q^{x(x-1)/2}\theta^x.$$

As $q \to 1$, all three tend to the ordinary binomial distribution.

The name Rogers-Szegö is appropriate for (8) since the pgf has the form of a Rogers-Szegö polynomial, Andrews (1976), Fine (1988); here we will set $\theta = \rho,\ 0 < \rho$, in order to distinguish between formulas. Similarly, (9) is a Stieltjes-Wigert polynomial, Szegö (1975); we will call this a Stieltjes-Wigert distribution and put $\theta = q^{1-m}\beta$, $0 < \beta$. When (8), (5), and (9) are reversed, the distribution of $m-X$ has the same nature and properties as the distribution of X. These distributions resemble the binomial distribution in this respect.

Kemp (1998) found that this reversal property does not hold for Dunkl's q-binomial distribution. When reversed the pgf becomes

$$G_{DAR}(z) = (q^{n+1-m};q)_m \, {}_1\phi_1(q^{-m};q^{n+1-m};q,q^{n+1}z). \tag{10}$$

Also the reversed form of Jing's q-deformed binomial has the pgf

$$G_{QDR}(z) = \tau^m \, {}_2\phi_0(q^{-m},\tau;-;q,q^m z/\tau). \tag{11}$$

In both cases the expressions for $\Pr[X=0]$ and $\Pr[X=m]$ have different forms.

Two further distributions are constructed in this paper. The first of these has the pgf

$$G_J(z) = \frac{{}_2\phi_1(q^{-m},\gamma;0;q,-qz)}{{}_2\phi_1(q^{-m},\gamma;0;q,-q)}, \quad 0<q<1,\ 0<\gamma<1; \tag{12}$$

its reversed form is

$$G_{JR}(z) = \frac{{}_1\phi_1(q^{-m};q^{1-m}/\gamma;q,-qz/\gamma)}{{}_1\phi_1(q^{-m};q^{1-m}/\gamma;q,-q/\gamma)}. \tag{13}$$

It resembles Dunkl's distribution, but with q^{-n} replaced by γ and a change in the sign of the argument; this alters its properties. Application of the Jackson (1927) transformation formula

$${}_2\phi_1(q^{-m},b;c;q,u] = \frac{(c/b;q)_m b^m}{(c;q)_m} \, {}_3\phi_1(q^{-m},b,q/u;bq^{1-m}/c;q,u/c], \tag{14}$$

with $b=\gamma$, $c=0$, and $u=-qz$, to (12) gives

$$\begin{aligned} G_J(z) &= \frac{\gamma^m \, {}_3\phi_0(q^{-m},\gamma,-1/z;-;q,-q^m z/\gamma)}{{}_2\phi_1(q^{-m},\gamma;0;q,-q)} \\ &= \frac{\gamma^m + \sum_{j=1}^{m} \begin{bmatrix} m \\ j \end{bmatrix}_q (\gamma;q)_j \, \gamma^{m-j}(1+z)(1+q^{-1}z)\cdots(1+q^{1-j}z)}{{}_2\phi_1(q^{-m},\gamma;0;q,-q)}, \end{aligned} \tag{15}$$

showing that the distribution is a mixture of Kemp and Kemp q-binomial distributions. The name 'Jackson distribution' will be used; the theory of q-series owes much to the long series of papers, 1904-1954, by F. H. Jackson.

The other new distribution resembles the q-deformed binomial distribution, but with q^{-n} instead of τ and a change in the sign of the argument. The pgf is

$$G_W(z) = \frac{{}_2\phi_1(q^{-m},0;q^{1-m+n};q,-qz)}{{}_2\phi_1(q^{-m},0;q^{1-m+n};q,-q)}, \quad 0<q<1,\ n\in\mathbb{Z}^+, \tag{16}$$

which is a Wall polynomial in z (a Wall polynomial is a special case of a little q-Jacobi polynomial, see G/R (1990)). The reversed Wall distribution has the pgf

$$G_{WR}(z) = \frac{{}_2\phi_0(q^{-m}, q^{-n}; -; q, -q^{m+n}z)}{{}_2\phi_0(q^{-m}, q^{-n}; q, -q^{m+n})}. \tag{17}$$

4. The Shapes of the Distributions

In this paper only the Heine and the infinitely divisible Euler distribution have infinite support. All the q-binomial analogues terminate (they cannot therefore be infinitely divisible). Their logconcavity/logconvexity properties determine their shapes. A terminating discrete distribution is logconcave or logconvex according as

$$D(x) = \frac{\Pr[X = x+1]}{\Pr[X = x]} - \frac{\Pr[X = x+2]}{\Pr[X = x+1]} \gtrless 0, \quad x = 0, 1, \ldots, m-2.$$

Most, but not all, of the q-binomial analogues in this paper are logconcave like the binomial distribution. For Dunkl's q-binomial, the Kemp and Kemp q-binomial, Rogers-Szegö, Stieltjes-Wigert, and Wall distributions we have

$$\begin{aligned}
D_{DA}(x) &= \frac{q^{2x-m-n+1}(1-q)f_{QA}(x)}{(1-q^{x+1})(1-q^{x+2})}, \\
D_{KB}(x) &= \frac{\theta q^x f_{KB}(x)}{(1-q^{x+1})(1-q^{x+2})}, \\
D_{RS}(x) &= \frac{\rho(1-q)f_{RS}(x)}{(1-q^{x+1})(1-q^{x+2})}, \\
D_{SW}(x) &= \frac{\beta q^{2x-m+1}(1-q)f_{SW}(x)}{(1-q^{x+1})(1-q^{x+2})}, \\
D_W(x) &= \frac{q^{x+1-m}(1-q)f_W(x)}{(1-q^{x+1})(1-q^{x+2})(1-q^{n-m+x+1})(1-q^{n-m+x+2})},
\end{aligned}$$

where
$$\begin{aligned}
f_{DA}(x) &= (1-q^{m-x})(1-q^{n-x})+q(1-q^{x+1})(1-q^{m+n-2x-1})>0, \\
f_{KB}(x) &= (1-q)(1-q^{m+1}) > 0, \\
f_{RS}(x) &= q^{x+1}(1-q^{m-x-1}) + q^{m-x-1}(1-q^{x+2}) > 0, \\
f_{SW}(x) &= (1-q^{m-x}) + q(1-q^{x+1}) > 0, \\
f_W(x) &= (1-q^{m-x-1})(1-q^{n-m+2x+3}) \\
&\quad + q^{m-x-1}(1-q^{x+2})(1-q^{n-m+x+2}) > 0.
\end{aligned}$$

These distributions, and their reversed forms, are therefore all logconcave. This implies that they are all IFR and strongly unimodal.

This is not true for the q-deformed, reversed q-deformed, Jackson, and reversed Jackson distributions. Here

$$D_{QD}(x) = \frac{\tau q^{-x-1}(1-q)f_{QD}(x)}{(1-q^{x+1})(1-q^{x+2})(1-\tau q^{m-x-1})(1-\tau q^{m-x-2})}$$

and
$$D_J(x) = \frac{q^{x+1-m}(1-q)f_J(x)}{(1-q^{x+1})(1-q^{x+2})},$$

where
$$f_{QD}(x) = (q^{x+1}-q^m)(q^{x+1}-q^{m+1})+q^{m-1}(q-\tau)(1-q^{m+1})$$

and
$$f_J(x) = (1-q^{x+1})(1-\gamma q^{x+1}) + q^x(q-\gamma)(1-q^{m-x}).$$

The q-deformed and reversed q-deformed binomial distributions are necessarily logconcave when $\tau \le q$ but may not be logconcave when $\tau > q$. For example, if $m = 6$, $q = 0.2$, $\tau = 0.22$, then $D_{QD}(x) > 0$ for $x = 0, 1, 2$ but $D_{QD}(x) < 0$ for $x = 3, 4$. The equation $f_{QD}(x) = 0$ is a quadratic in q^x and hence there are at most two values of q^x in $0 \le x \le (m-2)$ for which $f_{QD}(x) = 0$. Numerical exploration suggests that the q-deformed binomial distribution can have a mode followed by an antimode and a sesquimode at $x = m$. The reversed q-deformed binomial has the reverse shape.

The Jackson distribution is logconcave when $\gamma \le q$, but may not be logconcave when $\gamma > q$. If $m = 10$, $q = 0.9$, $\gamma = 0.95$, then $D_J(x) > 0$ for $x = 1, 2, \ldots, 8$, but $D_J(0) < 0$. The equation $f_J(x) = 0$ is again a quadratic in q^x with at most two values of q^x in $0 \le x \le (m-2)$ for which $f_J(x) = 0$; numerical exploration suggests that this distribution is able to have a sesquimode at the origin followed by an antimode and another mode. The reversed Jackson distribution has the reverse shape.

For $\tau \le q$ and $\gamma \le q$ these distributions, like the other q-binomials in this paper, are IFR and strongly unimodal as a consequence of their logconcavity.

5. Moment Properties

Only a few of these distributions have tractable moments. The pgf (5) can be restated as

$$G_{KB}(z) = \prod_{j=1}^{m} \frac{(1+\alpha q^j z)}{(1+\alpha q^j)}, \quad \text{where} \quad \alpha = \theta/q.$$

This is the convolution of m independent Bernoulli pgfs and therefore, as Kemp and Kemp (1991) realized, the cumulants are the sums of the m Bernoulli cumulants, giving

$$\mu = \sum_{j=1}^{m} \alpha q^j/(1+\alpha q^j), \qquad \sigma^2 = \sum_{j=1}^{m} \alpha q^j/(1+\alpha q^j)^2. \tag{18}$$

The mean and variance of Dunkl's absorption distribution were found by Kemp (1998) using the Jackson transformation (14). This gave

$$\begin{aligned} G_{DA}(z) &= {}_3\phi_0(q^{-m}, q^{-n}, z^{-1}; -; q, q^{m+n}z) \\ &= 1 + \sum_{j=1}^{m} \begin{bmatrix} m \\ j \end{bmatrix}_q (1-q^n)\cdots(1-q^{n-j+1})(z-1)\cdots(z-q^{j-1}), \end{aligned}$$

whence

$$\mu = \sum_{j=1}^{m} (1-q^m)\cdots(1-q^{m-j+1})(1-q^n)\cdots(1-q^{n-j+1})/(1-q^j), \tag{19}$$

$$\mu'_{[2]} = 2\sum_{j=2}^{m} \frac{(1-q^m)\cdots(1-q^{m-j+1})(1-q^n)\cdots(1-q^{n-j+1})}{(1-q^j)} \sum_{k=1}^{j-1} \frac{1}{(1-q^k)} \tag{20}$$

The moments of the Jackson distribution can be found from (15):

$$\mu = C\sum_{j=1}^{m} \begin{bmatrix} m \\ j \end{bmatrix}_q (\gamma;q)_j\, \gamma^{m-j}(1+q^0)\cdots(1+q^{1-j}) \sum_{\ell=0}^{j-1} \frac{1}{(1+q^\ell)}, \tag{21}$$

$$\mu'_{[2]} = 2C\sum_{j=2}^{m} \begin{bmatrix} m \\ j \end{bmatrix}_q (\gamma;q)_j\, \gamma^{m-j}(1+q^0)\cdots(1+q^{1-j}) \sum_{\ell_1,\ell_2}^{j-1} \frac{1}{(1+q^{\ell_1})(1+q^{\ell_2})}, \tag{22}$$

where $\ell_1 = 0,\ldots,j-1$, $\ell_2 = 0,\ldots,j-1$, $\ell_1 < \ell_2$, and $C = [{}_2\phi_1(q^{-m}, \gamma; 0; q, -q)]^{-1}$.

The Jackson transformation (14) applied to the q-deformed pgf and the Wall pgf gives

$$\begin{aligned} G_{QD}(z) &= {}_2\phi_0(q^{-m}, z^{-1}; -; q, q^m \tau z) \\ G_W(z) &= C^*\, {}_2\phi_0(q^{-m}, -z^{-1}; -; q, -q^{m-n}z) \end{aligned}$$

where $C^* = [{}_2\phi_0(q^{-m}, -1; -; q, -q^{m-n})]^{-1}$, enabling their moments to be got similarly.

No such transformation appears to be available for the Stieltjes-Wigert and Rogers-Szegö distributions; there it seems necessary to use

$$\mu'_{[r]} = \sum_{j=0}^{m} x(x-1)\cdots(x-r+1)\Pr[X=x].$$

The disadvantage of using this formula to calculate a factorial moment of a discrete distribution is that the use of only the first few terms in the series does not give a good approximation unless the distribution is strongly reverse J-shaped.

6. Models

Two types of model are considered in this paper, conditionality models and equilibrium birth-death processes. The former have been considered in depth for various q-series distributions in Kemp (2002), a paper that uses the Rao-Shanbhag-Kapoor Theorem, Rao and Shanbhag (1994), as the starting-point for generalizations of the Moran (1952) characterization of the Poisson distribution; see also Shanbhag and Kapoor (1993).

The Rao-Shanbhag-Kapoor Theorem characterizes the Kemp q-binomial (5) as the pgf for $U|(U+V=m)$ where U has a Heine and V has an independent Euler distribution. Theorem 2 in Kemp (2002) characterizes the Stieltjes-Wigert pgf (9) as the pgf for $U|(U+V=m)$ where U and V are independent Heine variables. Theorem 5 characterizes the Rogers-Szegö pgf (8) as the pgf for $U|(U+V=m)$ where U and V are independent Euler variables.

Kemp (2001) characterized Dunkl's absorption pgf (6) as the pgf for $U|(U+V=m)$ where U has a Kemp and Kemp q-binomial and V has an independent Heine distribution, given that both have the same argument parameter.

We can characterize the other distributions in this paper similarly. For the q-deformed binomial (7), U is an Euler $(q,\eta\tau)$ variable and V is an independent Dunkl (1981) q-negative binomial variable with parameters (q,η,τ). If U has a Dunkl q-negative binomial (q,λ,τ) distribution and V has an independent Heine (q,λ) distribution, then $U|(U+V=m)$ is a Jackson (12) variable. And if U has an Euler (q,η) distribution while V has an independent Kemp and Kemp q-binomial (q,η,n) distribution, then $U|(U+V=m)$ is a Wall (16) variable. The reversed distributions arise when the rôles of U and V are interchanged.

In a discussion of stationary birth-death models for wild animal populations, Kemp (1998) commented that for low population sizes it may be hard to find a mate; also death rates may be very high when population sizes are high. For the following non-linear birth and death rates,

$$\alpha_x = c(1-q^{m-x}), \qquad \beta_x = cq^{m+n+1-2x}(1-q^x)/(1-q^{n-x+1}),$$

the equilibrium outcome is Dunkl's absorption distribution.

She also considered a random walk on $0, 1, \ldots, m$ that is both timid and hesitant. If $(1-q^r)/(1-q^m)$ and $(q^r-q^m)/(1-q^m)$, $0<q<1$, are the probabilities that a particle at r steps from the origin attempts to move down or up, respectively, also aq^r, $0<a<1$, is the probability that an attempt to move up is successful and bq^r, $0<b<1$, is the probability that an attempt to move down is a failure, then the equilibrium pgf is

$$H(z) = {}_1\phi_1(q^{-m}; bq; q, aq^m z)/{}_1\phi_1(q^{-m}; bq; q, aq^m). \qquad (23)$$

As the particle gets further away from the origin any attempt to move up becomes less likely to succeed and any attempt to move down becomes more likely to succeed. The pgf for the reversed absorption distribution is the outcome when $a = q^{n-m+1}$, $b = q^{n-m}$.

Taking $\alpha_x = c\alpha q^{x+1}(1-q^{m-x})$ and $\beta_x = c(1-q^x)$ instead gives the Kemp and Kemp q-binomial, taking $\alpha_x = c\beta q^{2x}(1-q^{m-x})$ and $\beta_x = cq^{m-1}(1-q^x)$ gives the Stieltjes-Wigert distribution, and setting $\alpha_x = c\rho(1-q^{m-x})$ and $\beta_x = c(1-q^x)$ yields the Rogers-Szegö distribution.

The death rate $\beta_x = c(1-q^x)$ implies that the overall death rate increases with x but the death rate per individual β_x/x decreases. If $\alpha_x = cq$ for $0 \le x \le m-1$ and falls catastrophically to zero at $x=m$, then the equilibrium pmf is $\Pr[X=x] = q^x(q;q)_m/(q;q)_x$ and we have the q-deformed binomial distribution (7) with $\tau = q$. Suppose now that $\alpha_x = c\tau(1-q^{m-x})/(1-\tau q^{m-x-1})$, $x = 1, 2, \ldots, m$, instead. If $\tau < q$, then the birth rate decreases steadily to zero when $x = m$; increasing population size has a deleterious effect on both the overall birth rate and the death rate per individual. If $\tau > q$, then the birth rate rises steadily from $x=0$ to $x=m-1$ but drops catastrophically to zero at $x=m$. In both cases the outcome equilibrium distribution is the q-deformed binomial distribution.

When α_x is modified still further to $\alpha_x = c(1-q^{m-x})(1-\gamma q^x)$, with $\beta_x = cq^{m-x}(1-q^x)$, the equilibrium distribution is a Jackson distribution. Letting $\alpha_x = cq^x(1-q^{m-x})(1-q^{n-x})$ and $\beta_x = c(1-q^x)$ gives a reversed Wall distribution. Other equilibrium birth-death models can be constructed for the reversed distributions.

7. Applications

The following data sets have been fitted with some of the new distributions. Numerical search of the likelihood surface was used in order to obtain maximum likelihood fits because q-series are not amenable to ordinary integration and differentiation (the q-analogues of integration and differentiation use special differencing processes).

All the q-analogue binomial distributions that have been discussed here tend to the ordinary binomial distribution as $q \to 1$. When one of them gives a fit that is no closer than the binomial fit, the fit that it gives is the binomial fit.

(i) *Underdispersed word counts* (Bailey, 1990).

Samples of size 10 of words were drawn from the opening words on the printed lines of a copy of Macauley's "Essays on Milton"; the words were then classified as indefinite articles ('a', 'an') or otherwise. The following data were obtained:
(0: 27), (1: 44), (2: 26), (3: 3), ($\geq$ 4: 0). The poor binomial fit, (0: 33.0), (1: 38.7), (2: 20.4), (3: 6.4), ($\geq$ 4: 1.5), LL=–121.3, $\chi^2_{[2]}$=6.4, does not allow for the considerable underdispersion.

The Rogers-Szegö and Steiltjes-Wigert distributions have been constructed as weighted forms of the Kemps' q-binomial distribution, with weights $q^{x(1-x)/2}$ and $q^{x(x-1)/2}$, respectively. Use of the weights produces lesser or greater underdispersion compared with a binomial model, making the distributions suitable for the empirical fitting of data. When the Rogers-Szegö and Kemp and Kemp q-binomial distributions were fitted to the word count data, q limited to 1 and the fitted distributions limited to the fitted binomial distribution. The much improved Steiltjes-Wigert fit was
(0: 25.6), (1: 47.8), (2: 22.8), (3: 3.6), ($\geq$ 4: 0.2), LL=–117.7, $\chi^2_{[1]}$=1.0.

(ii) *Overdispersed Student absence data* (Ishii and Hawakawa, 1960).

The numbers of absences at lectures by 113 students attending a particular course of 9 lectures were as follows:
(0: 35), (1: 32), (2: 23), (3: 14), (4: 4), (5: 2), (6: 2), (7: 0), (8: 0), (9: 1); here there is marked overdispersion. The binomial fit, (0: 22.5), (1: 39.8), (2: 31.3), (3: 14.3), (4: 4.2), ($\geq$ 5: 0.9), LL=–190.0, $\chi^2_{[3]}$=13.6, was poor.

The Steiltjes-Wigert distribution arises as the distribution of $U|(U+V = m)$ when U and V are independent Heine random variables. The Rogers-Szegö is the distribution of $U|(U+V=m)$ when U and V are independent Euler random variables. Suppose that, overall, the students presences and

absences at lectures at their educational institution have independent Euler distributions, not necessarily with the same argument parameter; then the number of absences for a course of 9 lectures has a Rogers-Szegö distribution. Accordingly, a Rogers-Szegö distribution was fitted. It gave the much better fit:
(0: 29.1), (1: 35.6), (2: 25.2), (3: 13.6), (4: 6.1), ($\geq$ 5: 3.4), LL=–187.9, $\chi^2_{[2]}$=1.8.

A Rogers-Szegö distribution was chosen in preference to a Steiltjes-Wigert one because an Euler distribution is an overdispersed q-analogue of the Poisson distribution whereas a Heine distribution is an underdispersed Poisson q-analogue. It was thought that unconditional attendances and absences at lectures would both exhibit overdispersion. When a Steiltjes-Wigert distribution was fitted it gave its limiting form as $q \to 1$, i.e. the binomial fit.

References

ANDREWS, G.E. (1976). *The Theory of Partitions.* Addison-Wesley, Reading, Mass.

BAILEY, W.N. (1935). *Generalized Hypergeometric Series.* Cambridge University Press, London.

BAILEY, B.J.R. (1990). A model for function word counts. *Appl. Statist.* **39**, 107-114.

BENKHEROUF, L. and ALZAID, A.A. (1993). On the generalized Euler distribution. *Statist. Probab. Lett.* **18**, 323-326.

BENKHEROUF, L. and BATHER, J.A. (1988). Oil exploration: sequential decisions in the face of uncertainty. *J. Appl. Probab.* **25**, 529-543.

BIEDENHAHN, L.C. (1989). The quantum group $SU_q(2)$ and a q analogue of the boson operators. *J. Phys. A: Math. Gen.* **22**, L 873-878.

DUNKL, C.F. (1981). The absorption distribution and the q-binomial theorem. *Comm. Statist. Theor. Meth.* **10**, 1915-1920.

FINE, N.J. (1988). *Basic Hypergeometric Series and Applications.* American Mathematical Society, Providence, R.I.

GANI, J. (1974). Stochastic models for type counts in a literary text. In: Gani, J. (Ed.), *Perspectives in Probability and Statistics.* Academic Press, London, 313-323.

GASPER G. and RAHMAN M. (1990). *Basic Hypergeometric Series.* Cambridge University Press, Cambridge.

ISHII, G. and HAYAKAWA, R. (1960). On the compound binomial distribution. *Ann. Inst. Statistical Math.* **12**, 69-80, (Errata, **12**, 208).

JACKSON, F.H. (1927). A new transformation of Heinean Series. *Quart. J. Pure and Appl. Math.* **50**, 377-384.

JING, S.C. (1994). The q-deformed binomial distribution and its asymptotic behaviour. *J. Phys. A: Math. Gen.* **27**, 493-499.

JING, S.C. and FAN, H.Y. (1994). q-deformed binomial state. *Phys. Rev.* A. **49**, 2277-2279.

KEMP, A.W. (1987). A Poissonian binomial model with constrained parameters. *Naval Research Logistics Quarterly.* **34**, 853-858.

KEMP, A.W. (1992a). Heine-Euler extensions of the Poisson distribution. *Comm. Statist. Theor. Meth.* **21**, 571-588.

KEMP, A.W. (1992b). Steady State Markov chain models for the Heine and Euler distributions. *J. Appl. Probab.* **29**, 869-876.

KEMP, A.W. (1998). Absorption sampling and the absorption distribution. *J. Appl. Probab.* **35**, 1-6.

KEMP, A.W. (2001) A characterization of a distribution arising from absorption sampling. *Probability and Statistical Models with Applications,* (eds. Ch. A. Charalambides, M.V. Koutras and N. Balakrishnan) Chapman and Hall/CRC, Boca Raton, 239-246.

KEMP, A.W. (2002) Characterizations involving $U|(U+V=m)$ for certain discrete distributions. (To appear in a special issue of *J. Statist. Plann. Inference* in honor of Prof. C.R. Rao).

KEMP, A.W. and KEMP, C.D. (1991). Weldon's dice data revisited. *Amer. Statistician* **45**, 216–222.

KEMP, A.W. and NEWTON, J. (1990). Certain state-dependent processes for dichotomised parasite populations. *J. Appl. Probab.* **27**, 251-258.

MACFARLANE, A.J. (1989). On q-analogues of the quantum harmonic oscillator and the quantum group $SU(2)_q$. *J. Phys. A: Math. Gen.* **22**, 4581-4588.

MORAN, P.A.P. (1952). A characteristic property of the Poisson distribution. *Proc. Camb. Phil. Soc.* **48**, 206-207.

RAO, C.R. and SHANBHAG, D.N. (1994). *Choquet-Deny Type Functional Theorems with Applications to Stochastic Models.* Wiley, New York.

SHANBHAG, D.N. and KAPOOR, S. (1993). Some questions in characterization theory. *Math. Scientist* **18**, 127-133.

SLATER L.J. (1966). *Generalized Hypergeometric Functions.* Cambridge University Press, Cambridge.

SZEGÖ, G. (1975). *Orthogonal Polynomials.* American Mathematical Society, Providence, R.I.

FREDA (ADRIENNE) KEMP
MATHEMATICAL INSTITUTE
UNIVERSITY OF ST ANDREWS
NORTH HAUGH
ST ANDREWS KY16 9SS
SCOTLAND, UK
E-mail: awk@st-and.ac.uk