TWO CLASSES OF NUMBERS APPEARING IN THE CONVOLUTION OF BINOMIAL-TRUNCATED POISSON AND POISSON-TRUNCATED BINOMIAL RANDOM VARIABLES

M. Koutras
Statistical Unit, University of Athens, Athens, Greece
(Submitted October 1988)

1. Introduction

The Stirling numbers of the second kind, known to mathematicians as the coefficients in the factorial expansion of powers, are of great importance in the calculus of finite differences, and have been extensively studied, especially with respect to their mathematical properties (see Jordan [14], Riordan [17] and references therein). Recently, several extensions and modifications were considered, which have proved useful in various combinatorial, probabilistic and statistical applications. Of the most interesting variations are the Lah numbers, Lah [16], and their generalization C-numbers, Charalambides [8], [9], appearing in the expansion of a factorial of t, scaled by a parameter s, in a sum of factorials of t.

The present paper was motivated by the problem of providing explicit expressions for the distribution of two-sample sums from Poisson and binomial distributions, one of which is left-truncated. Specifically, the distribution of the statistic $Z = X_1 + \cdots + X_v + X_{v+1} + \cdots + X_{v+n}$, where

(a) $X_1, \ldots, X_v$ is a random sample from a Poisson and $X_{v+1}, \ldots, X_{v+n}$ an independent random sample from a left-truncated binomial distribution and

(b) $X_1, \ldots, X_v$ is a random sample from a binomial and $X_{v+1}, \ldots, X_{v+n}$ an independent random sample from a left-truncated Poisson distribution,

led to the introduction of two double sequences of Stirling and C-related numbers, obtained from the expansion of certain classes of polynomials in a series of factorials.

In Section 2, we discuss some general results relating the expansion of polynomials in factorials and the corresponding exponential generating functions (egf's). In Section 3, we consider two specific families of polynomials (r-q polynomials) and introduce two double sequences of numbers (R-Q numbers). Notice that in Tauber's [19] terminology these numbers might be called generalized Lah numbers. Next, the egf's of the R-Q numbers are used to derive recurrence relations and initial conditions and the connection to well-known numbers is examined in Section 4. In Section 5, it is shown how R-Q numbers can be used for the introduction of two new families of truncated discrete probability functions including binomial and hypergeometric distributions as special cases; also for the solution of the above-mentioned problems (a) and (b). An application to occupancy problems is also provided. Finally, in Section 6, a further generalization of the R-Q numbers, through egf's, is also discussed, along with its properties and applications.

2. Preliminary General Results

Let $\{p_m(x),\ m = 0, 1, \ldots\}$ be a class of polynomials, and consider the double sequence $\{P(m, n),\ m = 0, 1, \ldots,\ n = 0, 1, \ldots, m\}$ obtained by expanding the polynomial $p_m(x)$ in a series of factorials, namely

$$(2.1) \qquad p_m(x) = \sum_{n=0}^{m} P(m, n)(x)_n .$$

Denote the egf of the numbers $P(m, n)$ with respect to the index m by $f_n(t)$, and the egf of the polynomials $p_m(x)$ by $p(x; t)$, that is

$$(2.2) \qquad f_n(t) = \sum_{m=n}^{\infty} P(m, n)\frac{t^m}{m!}, \quad p(x; t) = \sum_{m=0}^{\infty} p_m(x)\frac{t^m}{m!}.$$

On using (2.1), we may easily verify that

$$p(x, t) = \sum_{n=0}^{\infty} f_n(t)(x)_n ,$$

and the next theorem is an immediate consequence of Newton's formula (see Jordan [14]).

Theorem 2.1: Let $p(x; t)$ denote the egf of a class of polynomials $\{p_m(x), m = 0, 1, \ldots\}$ and $f_n(t)$ the egf of the corresponding numbers $P(m, n)$ as defined in (2.1). Then

$$(2.3) \qquad f_n(t) = \frac{1}{n!}[\Delta_x^n p(x, t)]_{x=0}.$$

We now state some general results referring to recurrence relations satisfied by the polynomials $p_m(x)$ and the numbers $P(m, n)$, when a certain partial differential equation holds true for the egf $p(x, t)$.

Theorem 2.2: If the egf $p(x, t)$ of the polynomials $p_m(x)$ satisfies the partial differential equation

$$(2.4) \qquad (1 + Bt + Ct^2)\frac{\partial p(x, t)}{\partial t} = (D + Et)p(x, t),$$

where B, C, D, and E may be functions of x, then there is a recurrence relation connecting three polynomials $p_m(x)$ with consecutive indices (degrees), namely,

$$(2.5) \qquad p_{m+1}(x) = (D - Bm)p_m(x) + ((E + C)m - Cm^2)p_{m-1}(x).$$

Proof: Differentiate $p(x, t)$ of (2.2) term by term, substitute in (2.4) and equate the coefficients of $t^m/m!$ in the right and left sides of the resulting identity.

Note that (2.5) is true for $m \geq 1$, while, for $m = 0$, it reduces to

$$(2.6) \qquad p_1(x) = Dp_0(x),$$

which suggests that $D = D(x)$ must be at least of order 1 with respect to x.

Theorem 2.3: If $p(x, t)$ satisfies the partial differential equation

$$(2.7) \qquad (1 + bt)\frac{\partial p(x, t)}{\partial t} = (c_0 + c_1 x + c_2 t + c_{12} xt)p(x, t)$$

with b, c_0, c_1, c_2, and c_{12} being constants, then the polynomials $p_m(x)$ and the numbers $P(m, n)$ satisfy the recurrences

$$(2.8a) \qquad p_{m+1}(x) = (c_0 + c_1 x - bm)p_m(x) + (c_2 + c_{12}x)mp_{m-1}(x), \quad m \geq 0,$$

$$(2.8b) \qquad p_1(x) = (c_0 + c_1 x)p_0(x),$$

$$(2.9a) \qquad P(m + 1, n) = (c_0 - bm + c_1 n)P(m, n) + c_1 P(m, n - 1) + m(c_2 + nc_{12})P(m - 1, n) + c_{12}mP(m - 1, n - 1), \quad 1 \leq n \leq m - 1,$$

$$(2.9b) \qquad P(m + 1, m + 1) = c_1 P(m, m),$$

$$(2.9c) \qquad P(m + 1, m) = (c_0 + (c_1 - b)m)P(m, m) + c_1 P(m, m - 1) + c_{12}mP(m - 1, m - 1).$$

Proof: For (2.8), apply Theorem 2.2 in the special case

$$B(x) = b,\ c(x) = 0,\ D(x) = c_0 + c_1 x,\ E(x) = c_2 + c_{12}x.$$

For (2.9) observe that, after expanding $p_{m+1}(x)$, $p_m(x)$, and $p_{m-1}(x)$ by (2.1), one obtains

$$\sum_{n=0}^{m+1} P(m+1, n)(x)_n = (c_0 - bm)\sum_{n=0}^{m} P(m, n)(x)_n + c_1 \sum_{n=0}^{m} nP(m, n)(x)_n + c_1 \sum_{n=1}^{m+1} P(m, n-1)(x)_n + m(c_2 + nc_{12})\sum_{n=0}^{m-1} P(m-1, n)(x)_n + mc_{12}\sum_{n=1}^{m} P(m-1, n-1)(x)_n,$$

which establishes the proof.

It is worth noticing that many classes of well-known numbers with special interest in statistical and combinatorial applications, have an egf $p(x, t)$ obeying the partial differential equation (2.7). For example,

α. If $p_m(x) = x^m$, we obtain the Stirling numbers of the second kind (see Jordan [14]) and

$$p(x, t) = \sum_{m=0}^{\infty} x^m \frac{t^m}{m!} = e^{tx}, \quad \frac{\partial p(x, t)}{\partial t} = xp(x, t).$$

β. If $p_m(x) = (x - a)^m$, we are led to the noncentral Stirling numbers of the second kind, Koutras [15], or weighted Stirling numbers, Carlitz [4], [5], or r-Stirling numbers, Broder [1], with

$$p(x, t) = \sum_{m=0}^{\infty} (x - a)^m \frac{t^m}{m!} = e^{t(x-a)}, \quad \frac{\partial p(x, t)}{\partial t} = (x - a)p(x, t).$$

γ. If $p_m(x) = (-x)_m$, or more generally, $p_m(x) = (sx)_m$, we obtain the Lah or C-numbers, respectively, Lah [16], Charalambides [8], [9], with

$$p(x, t) = \sum_{m=0}^{\infty} (sx)_m \frac{t^m}{m!} = (1 + t)^{sx}, \quad (1 + t)\frac{\partial p(x, t)}{\partial t} = sxp(x, t).$$

δ. If $p_m(x) = (sx + r)_m$, the resulting numbers are the Gould and Hopper numbers studied by Charalambides & Koutras [10]. In this case, we have

$$p(x, t) = \sum_{m=0}^{\infty} (sx + r)_m \frac{t^m}{m!} = (1 + t)^{sx+r}, \quad (1 + t)\frac{\partial p(x, t)}{\partial t} = (sx + r)p(x, t).$$

Notice how simple it is to compute the egf for any of the above-mentioned special cases. The egf $f_n(t)$ and the recurrences for the corresponding numbers are then easily obtained as a direct application of Theorems 2.1 and 2.3.

3. The r-q Polynomials and Numbers—Generating Functions and Recurrence Relations

Let us define two classes of polynomials by the formulas

$$r_m(x) = r_m(x; s, a) = e^{-a}\frac{d^m}{dt^m}[t^{sx}e^{at}]_{t=1}, \tag{3.1}$$

$$q_m(x) = q_m(x; a) = e^{-x}\frac{d^m}{dt^m}[t^a e^{xt}]_{t=1}. \tag{3.2}$$

Thus, the first few r-q polynomials are

$$r_0(x) = 1, \quad r_1(x) = sx + a, \quad r_2(x) = s^2x^2 + (2a - 1)sx + a^2,$$

$$q_0(x) = 1, \quad q_1(x) = x + a, \quad q_2(x) = x^2 + 2ax + a(a - 1).$$

Considering the Newton expansion of r-q polynomials in a series of factorials, we may define the R-Q numbers by

$$(3.3)\qquad r_m(x;\, a,\, s) = \sum_{n=0}^{m} R(m,\, n;\, s,\, a)(x)_n = \sum_{n=0}^{m} R(m,\, n)(x)_n,$$

$$(3.4)\qquad q_m(x;\, a) = \sum_{n=0}^{m} Q(m,\, n;\, a)(x)_n = \sum_{n=0}^{m} Q(m,\, n)(x)_n.$$

Since for $a = 0$ the r-q polynomials reduce to

$$r_m(x) = (sx)_m,\quad q_m(x) = x^m,$$

it follows that

$$R(m,\, n;\, s,\, 0) = C(m,\, n,\, s)$$

the C-numbers,

$$R(m,\, n;\, -1,\, 0) = L(m,\, n)$$

the Lah numbers, and

$$Q(m,\, n;\, 0) = S(m,\, n)$$

the Stirling numbers of the second kind.

As a starting point, let us derive the egf of the r-q polynomials and numbers, namely

$$(3.5)\qquad r(x,\, t;\, s,\, a) = \sum_{m=0}^{\infty} r_m(x;\, s,\, a)\frac{t^m}{m!},$$

$$f_n(t;\, s,\, a) = \sum_{m=n}^{\infty} R(m,\, n;\, s,\, a)\frac{t^m}{m!},$$

$$q(x,\, t;\, a) = \sum_{m=0}^{\infty} q_m(x;\, a)\frac{t^m}{m!},\quad g_n(t;\, a) = \sum_{m=n}^{\infty} Q(m,\, n;\, a)\frac{t^m}{m!}.$$

Regarding $t^{sx}e^{at}$ and $t^a e^{xt}$ as functions of t and expanding in a Taylor series around $t = 1$, we obtain

$$t^{sx}e^{at} = \sum_{m=0}^{\infty} \frac{d^m}{dt^m}[t^{sx}e^{at}]_{t=1}\frac{(t-1)^m}{m!},$$

$$t^a e^{xt} = \sum_{m=0}^{\infty} \frac{d^m}{dt^m}[t^a e^{xt}]_{t=1}\frac{(t-1)^m}{m!},$$

and, using definitions (3.1) and (3.2), we get

$$(3.6)\qquad r(x,\, t) = r(x,\, t;\, s,\, a) = (1+t)^{sx}e^{at},$$

$$(3.7)\qquad q(x,\, t) = q(x,\, t;\, a) = (1+t)^a e^{xt}.$$

As regards the egf's of R-Q numbers, they may be obtained easily from Theorem 2.1, which, in view of (3.6) and (3.7), gives

$$f_n(t;\, s,\, a) = \frac{1}{n!}\, e^{at}\, [\underset{x}{\Delta}^n (1+t)^{sx}]_{x=0} = \frac{1}{n!}\, e^{at}\, [(1+t)^{sx}\{(1+t)^s - 1\}^n]_{x=0}.$$

$$g_n(t;\, a) = \frac{1}{n!}(1+t)^a\, [\underset{x}{\Delta}^n e^{tx}]_{x=0} = \frac{1}{n!}(1+t)^a\, [e^{tx}\{e^t - 1\}^n]_{x=0}.$$

Therefore,

$$(3.8)\qquad f_n(t) = f_n(t;\, s,\, a) = \frac{1}{n!}\, e^{at}\{(1+t)^s - 1\}^n,$$

(3.9) $$g_n(t) = g_n(t;\ a) = \frac{1}{n!}(1 + t)^a\{e^t - 1\}^n.$$

Differentiating (3.6) and (3.7) with respect to t, we obtain the partial differential equations

$$(1 + t)\frac{\partial r(x,\ t)}{\partial t} = (sx + at + a)r(x,\ t)$$

and

$$(1 + t)\frac{\partial q(x,\ t)}{\partial t} = (a + xt + x)r(x,\ t),$$

and using Theorem 2.3, we may establish the following recurrence relations for the r-q polynomials and R-Q numbers

(3.10) $$\begin{aligned} r_{m+1}(x) &= (a + sx - m)r_m(x) + amr_{m-1}(x),\ m \ge 1,\\ r_1(x) &= (a + sx)r_0(x); \end{aligned}$$

(3.11) $$\begin{aligned} q_{m+1}(x) &= (a + x - m)q_m(x) + mxq_{m-1}(x),\ m \ge 1,\\ q_1(x) &= (a + x)q_0(x); \end{aligned}$$

(3.12) $$\begin{aligned} R(m + 1,\ n) = {}& (a + sn - m)R(m,\ n) + amR(m - 1,\ n)\\ & + sR(m,\ m - 1),\ m \ge n + 1; \end{aligned}$$

(3.13) $$R(m + 1,\ m) = sR(m,\ m - 1) + (a + sm - m)R(m,\ m);$$

(3.14) $$R(m,\ m) = sR(m - 1,\ m - 1);$$

(3.15) $$\begin{aligned} Q(m + 1,\ n) = {}& (a + n - m)Q(m,\ n) + nmQ(m - 1,\ n)\\ & + Q(m,\ n - 1) + mQ(m - 1,\ n - 1),\ m \ge n + 1; \end{aligned}$$

(3.16) $$Q(m + 1,\ m) = aQ(m,\ m) + Q(m,\ m - 1) + mQ(m - 1,\ m - 1);$$

(3.17) $$Q(m,\ m) = Q(m - 1,\ m - 1).$$

Notice that both relations (3.12) and (3.15) are not "triangular array recurrences" since, for the computation of the $(m + 1,\ n)$ term, they require the value of the $(m - 1,\ n)$ term. It is also obvious that, in order to compute all the terms of the double sequences $R(m,\ n)$ and $Q(m,\ n)$, $m \ge n$ via recurrences (3.12) and (3.15), respectively, one should at least know the following "initial" (boundary) conditions

a. m-axis values $R(m,\ 0)$, $Q(m,\ 0)$, $m = 0,\ 1,\ \ldots,$
b. first-diagonal values $R(m,\ m)$, $Q(m,\ m)$, $m = 1,\ 2,\ \ldots,$
c. second-diagonal values $R(m,\ m - 1)$, $Q(m,\ m - 1)$, $m = 1,\ 2,\ \ldots\ .$

For (a), consider the egf's (3.8) and (3.9) which, in the special case $n = 0$, give

$$f_0(t) = \sum_{m=0}^{\infty} R(m,\ 0)\frac{t^m}{m!} = e^{at} = \sum_{m=0}^{\infty} a^m\frac{t^m}{m!},$$

$$g_0(t) = \sum_{m=0}^{\infty} Q(m,\ 0)\frac{t^m}{m!} = (1 + t)^a = \sum_{m=0}^{\infty}\binom{a}{m}t^m.$$

Hence,

(3.18) $$R(m,\ 0) = a^m,\quad Q(m,\ 0) = (a)_m.$$

The initial condition (b) is readily obtained through (3.14), (3.17), (3.18), as

(3.19) $$R(m,\ m) = s^m,\quad Q(m,\ m) = 1.$$

As regards condition (c), we proceed as follows: relations (3.13) and (3.16), in view of (3.18), may be written in the form

$$\underset{m}{\Delta} s^{-m+1}R(m,\ m-1) = a + (s-1)m, \quad \underset{m}{\Delta} Q(m,\ m-1) = a + m,$$

and inverting the difference operator Δ_m, we obtain

$$s^{-m+1}R(m,\ m-1) = am + (s-1)\binom{m}{2} + k_1, \quad Q(m,\ m-1) = am + \binom{m}{2} + k_2.$$

Since

$$R(2,\ 1) = 2R(1,\ 0) + (a + s - 1)R(1,\ 1) = (2a + s - 1)s$$

and

$$Q(2,\ 1) = Q(1,\ 0) + aQ(1,\ 1) + Q(0,\ 0) = 2a + 1,$$

both constants k_1 and k_2 should vanish, and we finally deduce that

$$\begin{aligned} R(m,\ m-1) &= ams^{m-1} + \binom{m}{2}(s)_2 s^{m-2} \\ Q(m,\ m-1) &= am + \binom{m}{2} \end{aligned} \qquad m = 2,\ 3,\ \ldots\ . \tag{3.20}$$

It is obvious that the recurrences (3.12) and (3.15), along with initial conditions (3.18), (3.19) and (3.20) determine the double sequences $R(m,\ n)$, $Q(m,\ n)$, $m \geq n$.

4. Connection with Other Numbers

Let us denote by

$$s(m,\ n;\ a) = \frac{1}{n!}\left[\frac{d}{dx^n}(x)_m\right]_{x=a}, \quad S(m,\ n;\ a) = \frac{1}{n!}[\Delta^n x^m]_{x=a}$$

the noncentral Stirling numbers of the first and second kind, respectively, and

$$C(m,\ n;\ s,\ a) = \frac{1}{n!}[\Delta^n (sx + a)_m]_{x=0}$$

the noncentral C or Gould and Hopper numbers.

The first class of numbers has been recently studied by Carlitz [4], [5] as weighted Sterling numbers, by Koutras [15], as noncentral Stirling numbers, by Broder [1] as r-Stirling numbers, and by Shanmugan [18]. The second class, which was introduced by Chak [6] and Gould & Hopper [12], and subsequently investigated by Charalambides & Koutras [10], is closely related to Howard's [13] degenerate weighted Stirling numbers $S_1(m,\ n,\ \lambda|\theta)$ and $S(m,\ n,\ \lambda|\theta)$ by

$$S_1(m,\ n,\ \lambda|\theta) = (-1)^{m-n}C(m,\ n;\ \theta - \lambda,\ \theta)/\theta^n,$$

$$S(m,\ n,\ \lambda|\theta) = \theta^m C(m,\ n;\ \lambda\theta^{-1},\ \theta^{-1}).$$

In order to establish the connection between the R-Q numbers and the abovementioned classes, let us denote by

$$H_n(t) = H_n(t;\ a) = \sum_{m=n}^{\infty} S(m,\ n;\ a)\frac{t^m}{m!} = \frac{1}{n!}e^{at}[e^t - 1]^n$$

and

$$C_n(t) = C_n(t;\ s,\ a) = \sum_{m=n}^{\infty} C(m,\ n;\ s,\ a)\frac{t^m}{m!} = \frac{1}{n!}(1 + t)^a[(1 + t)^s - 1]^n$$

the egf's of noncentral Stirling and C-numbers, respectively. Comparing with formulas (3.8) and (3.9), we obtain

$$f_n(t;\ a,\ s) = e^{at}f_n(t;\ 0,\ s) = e^{at}C_n(t;\ s,\ 0),$$

$$g_n(t;\ a) = (1 + t)^a g_n(t;\ 0) = (1 + t)^a H_n(t;\ 0),$$

$$(1 + t)^a f_n(t;\ s,\ a) = e^{at}C_n(t;\ s,\ a),$$

$$e^{at}g_n(t;\ a) = (1 + t)^a H_n(t;\ a),$$

which imply the corresponding relations

$$R(m,\ n;\ s,\ a) = \sum_{k=n}^{m}\binom{m}{k}a^{m-k}R(k,\ n;\ s,\ 0) = \sum_{k=n}^{m}\binom{m}{k}a^{m-k}C(k,\ n,\ s) \tag{4.1}$$

and

$$Q(m, n; a) = \sum_{k=n}^{m} \binom{m}{k} (a)_{m-k} Q(k, n; 0) = \sum_{k=n}^{m} \binom{m}{k} (a)_{m-k} S(k, n);$$

$$(4.2) \qquad \sum_{k=n}^{m} \binom{m}{k} (a)_{m-k} R(k, n; s, a) = \sum_{k=n}^{m} \binom{m}{k} a^{m-k} C(k, n; s, a),$$

and

$$\sum_{k=n}^{m} \binom{m}{k} a^{m-k} Q(k, n; a) = \sum_{k=n}^{m} \binom{m}{k} (a)_{m-k} S(k, n; a).$$

Note also that (4.1) leads to the inverse relations

$$(4.3) \qquad C(m, n; s) = \sum_{k=n}^{m} \binom{m}{k} (-a)^{m-k} R(k, n; s, a),$$

and

$$S(m, n) = \sum_{k=n}^{m} \binom{m}{k} (-a)_{m-k} Q(k, n; a),$$

which imply that the RHS sums are independent of the parameter a.

Finally, we mention that, in view of (4.1), formulas (3.3) and (3.4) lead to the following explicit expressions for the r-q polynomials

$$(4.4) \qquad r_m(x; s, a) = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} (sx)_k,$$

and

$$q_m(x; a) = \sum_{k=0}^{m} \binom{m}{k} (a)_{m-k} x^k.$$

Remark 1: The proof of (4.4) could also be obtained through the egf's $r(x, t)$, $q(x, t)$, by expanding the RHS of (3.6) and (3.7) in a power series with respect to t.

Remark 2: Comparing (4.3) with the binomial and Vandermonde formulas,

$$(a + x)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} x^k, \quad (a + sx)_m = \sum_{k=0}^{m} \binom{m}{k} (a)_{m-k} (sx)_k,$$

one might say that the r-q polynomials are the "intermediate connectors" in the transition from powers to factorials and vice versa.

Another important formula for the R-Q numbers may be obtained as follows: Multiplying (4.1) by $C(n, \lambda, s^{-1})$ and summing up for $n = \lambda, \lambda + 1, \ldots, m$, we have

$$\sum_{n=\lambda}^{m} R(m, n; s, a) C(n, \lambda, s^{-1}) = \sum_{n=\lambda}^{m} \sum_{k=n}^{m} \binom{m}{k} a^{m-k} C(k, n, s) C(n, \lambda, s^{-1})$$
$$= \sum_{k=\lambda}^{m} \binom{m}{k} a^{m-k} \sum_{n=\lambda}^{k} C(k, n, s) C(n, \lambda, s^{-1})$$

and on using the orthogonality property of C-numbers, we deduce that

$$(4.5) \qquad \sum_{n=\lambda}^{m} R(m, n; s, a) C(n, \lambda, s^{-1}) = \binom{m}{\lambda} a^{m-\lambda}.$$

Similarly, the orthogonality property of Stirling numbers implies that

$$(4.6) \qquad \sum_{n=\lambda}^{m} Q(m, n; a) s(n, \lambda) = \binom{m}{\lambda} (a)_{m-\lambda}.$$

In matrix notation, formulas (4.5) and (4.6) could be stated as follows: If $R = (R_{mn})$, $Q = (Q_{mn})$, $C = (C_{mn})$, and $s = (s_{mn})$ are the infinite matrices with

$$R_{mn} = R(m, n; s, a), \quad Q_{mn} = Q(m, n; a), \quad m, n = 0, 1, \ldots,$$
$$C_{mn} = C(m, n, s^{-1}), \quad s_{mn} = s(m, n), \qquad n, \lambda = 0, 1, \ldots,$$

respectively, then

$$RC = \left(\binom{m}{\lambda} a^{m-\lambda}\right), \qquad QS = \left(\binom{m}{\lambda}(a)_{m-\lambda}\right)$$

5. Applications

a. Two new families of discrete truncated distributions

It is obvious that the egf's (3.8) and (3.9) satisfy the relations

$$f_{n+\lambda}(t;\ s,\ a+b) = \frac{n!\lambda!}{(n+\lambda)!} f_n(t;\ a,\ s) f_\lambda(t;\ b,\ s)$$

and

$$g_{n+\lambda}(t;\ a+b) = \frac{n!\lambda!}{(n+\lambda)!} g_n(t;\ a) g_\lambda(t;\ b),$$

which imply the following addition theorems:

$$R(m,\ n+\lambda;\ s,\ a+b) = \binom{n+\lambda}{n}^{-1} \sum_{k=0}^{m} \binom{m}{k} R(k,\ n;\ s,\ a) R(m-k,\ \lambda;\ s,\ b);$$

$$Q(m,\ n+\lambda;\ a) = \binom{n+\lambda}{n}^{-1} \sum_{k=0}^{m} \binom{m}{k} Q(k,\ n;\ a) Q(m-k,\ \lambda;\ b).$$

For $\lambda = 0$ one obtains, by virtue of (3.18),

$$R(m,\ n;\ s,\ a+b) = \sum_{k=0}^{m} \binom{m}{k} b^{m-k} R(k,\ n;\ s,\ a),$$

$$Q(m,\ n;\ a+b) = \sum_{k=0}^{m} \binom{m}{k} (b)_{m-k} Q(k,\ n;\ a),$$

and, therefore, we are led to the conclusion that

$$(5.1) \qquad f(x;\ m,\ n;\ a,\ b) = \binom{m}{x} \frac{b^{m-x} R(x,\ n;\ s,\ a)}{R(m,\ n;\ s,\ a+b)}, \quad x = n,\ n+1,\ \ldots,\ m,$$

and

$$g(x;\ m,\ n;\ a,\ b) = \binom{m}{x} \frac{(b)_{m-x} Q(x,\ n;\ a)}{Q(m,\ n;\ a+b)}, \quad x = n,\ n+1,\ \ldots,\ m,$$

define families of multiparameter discrete distributions with range

$$R_x = \{n,\ n+1,\ \ldots,\ m\}.$$

Note that probability functions (5.1) could be regarded as generalizations of binomial and hypergeometric laws, respectively, since

$$f(x;\ m,\ 0;\ a,\ b) = \binom{m}{x} \left(\frac{a}{a+b}\right)^x \left(\frac{b}{a+b}\right)^{m-x}$$

$$g(x;\ m,\ 0;\ a,\ b) = \binom{a}{x} \binom{b}{m-x} \Big/ \binom{a+b}{m}.$$

b. Convolution of binomial and Poisson distributions with truncation away from zero

Let $X_1,\ X_2,\ \ldots,\ X_v$ be a random sample from the binomial distribution

$$(5.2) \qquad P[X = x] = (1+\theta)^{-N} \binom{N}{x} \theta^x, \quad x = 0,\ 1,\ 2,\ \ldots,\ N,$$

where $\theta = p/(1-p) > 0$ and N is a positive integer. It is well known that the sum $Z_1 = X_1 + \cdots + X_v$ is again a binomial variable $b(vN,\ p)$ with probability function

$$(5.3) \qquad P[Z_1 = z] = (1-\theta)^{-a} \binom{a}{z} \theta^z, \quad z = 0,\ 1,\ \ldots,\ a, \quad a = vN.$$

Assume further that another independent sample $X_{v+1}, \ldots, X_{v+n}$ coming from the zero-truncated Poisson distribution with parameter θ is available. For statistical inference purposes, it would be interesting to establish explicit formulas for the distribution of the two-sample sum $Z = X_1 + \cdots + X_{v+n}$. To this end, we proceed as follows: The probability function of $Z_2 = X_{v+1} + \cdots + X_{v+n}$ was obtained by Cacoullos [2] in the form

$$P[Z_2 = z] = \frac{n!S(z, n)\theta^z}{[e^\theta - 1]^n z!}, \quad z = n, n + 1, \ldots .$$

Therefore,

$$P[Z = z] = \sum_{x=n}^{z} P[Z_1 = z - x]P[Z_2 = x]$$
$$= \frac{n!}{(1 + \theta)^a (e^\theta - 1)^n} \frac{\theta^z}{z!} \sum_{x=n}^{z} \binom{z}{x} (a)_{z-x} S(x, n)$$

which, on using (4.1), gives

$$(5.4) \qquad P[Z = z] = \frac{n!Q(z, n; a)}{(1 + \theta)^a (e^\theta - 1)^n} \frac{\theta^z}{z!}, \quad z = n, n + 1, \ldots .$$

Expression (5.4) may be used to obtain an explicit formula for the (unique) unbiased estimator of the parametric function θ^k (k a positive integer) that is based on the two-sample sum Z. Thus, from the condition of unbiasedness

$$E[h_k(Z)] = \theta^k \text{ for every } \theta > 0,$$

we obtain, by virtue of (5.4), (3.5), and (3.9),

$$\sum_{z=n}^{\infty} h_k(z) Q(z, n; a) \frac{\theta^z}{z!} = \sum_{z=n+k}^{\infty} (z)_k Q(z - k, n; a) \frac{\theta^z}{z!},$$

which implies that

$$h_k(z) = \begin{cases} \dfrac{(z)_k Q(z - k, n; a)}{Q(z, n; a)} & \text{if } z \ge n + k, \\ 0 & \text{if } z < n + k. \end{cases}$$

Hence,

$$h_1(Z) = ZQ(Z - 1, n; a)/Q(Z, n; a), \quad Z \ge n + 1$$

is an unbiased estimator of θ, and since

$$\mathrm{Var}[h_1(Z)] = E[(h_1(Z))^2] - \theta^2 = E[(h_1(Z))^2] - E[h_2(Z)]$$

the statistic

$$h^*(Z) = [h_1(Z)]^2 - h_2(Z)$$

will be an unbiased estimator of the variance of the unbiased estimator of θ.

Consider the case where $X_1, X_2, \ldots, X_v$ is a random sample from the Poisson distribution with parameter 0 and $X_{v+1}, \ldots, X_{v+n}$ is an independent sample from the zero-truncated binomial law with probability function

$$P[X = x] = [(1 + \theta)^N - 1]^{-1} \binom{N}{x} \theta^x, \quad x = 1, 2, \ldots .$$

The distribution of $Z_1 = X_1 + \cdots + X_v$ is, of course, Poisson with parameter $v\theta$, while the probability function of $Z_2 = X_{v+1} + \cdots + X_{v+n}$ is given by (see Cacoullos & Charalambides [3])

$$P[Z_2 = x] = \frac{n!C(x, n, N)}{[(1 + \theta)^N - 1]^n} \frac{\theta^x}{x!}, \quad x = n, n + 1, \ldots .$$

Therefore, the probability function of the two-sample sum $Z = X_1 + \cdots + X_{v+n}$ is

$$P[Z = z] = \sum_{x=n}^{m} P[Z_1 = z - x]P[Z_2 = x]$$
$$= \frac{n!}{e^{v\theta}[(1+\theta)^N - 1]^n} \frac{\theta^z}{z!} \sum_{x=n}^{z} \binom{z}{x} v^{z-x} C(x, n, N)$$

which, on using (4.1), reduces to

$$(5.5) \quad P[Z = z] = \frac{n! R(z, n; N, v)}{e^{v\theta}[(1+\theta)^N - 1]^n} \frac{\theta^z}{z!}, \quad z = n, n+1, \ldots .$$

Following similar arguments with the binomial-zero truncated Poisson problem, one could easily verify that $h_k(Z)$ with

$$h_k(Z) = \begin{cases} \dfrac{(z)_k R(z-k, n; N, v)}{R(z, n; N, v)} & \text{if } z \geq n+k, \\ 0 & \text{if } z < n+k, \end{cases}$$

is an unbiased estimator of the parametric function θ^k, while

$$h^*(z) = [h_1(Z)]^2 - h_2(Z)$$

is an unbiased estimator of the variance of the unbiased estimator of θ.

c. Occupancy problems

Formula (4.1) implies the following combinatorial interpretation of the numbers $Q(m, n; a)$: Consider n identical cells with no capacity restrictions and a control cell of $a \in Z^+$ different (distinguishable) compartments, each of capacity 1. If $a + m \geq n$, then $Q(m, n; a)$ is equal to the number of ways of distributing m distinct balls into the cells so that none of the n identical cells is empty.

6. The Generalized R-Q Numbers

Following the technique used by Charalambides [8] and Charalambides & Koutras [10], we may define the generalized R-Q numbers

$$R_r(m, n; a, s) = R_r(m, n) \quad \text{and} \quad Q_r(m, n; a) = Q_r(m, n)$$

by their egf's as follows [cf. (3.8) and (3.9)],

$$(6.1) \quad f_{n,r}(t) = f_{n,r}(t; s, a) = \sum_{m=rn}^{\infty} R_r(m, n)\frac{t^m}{m!}$$
$$= \frac{1}{n!} e^{at}\left\{(1+t)^s - \sum_{k=0}^{r-1}\binom{s}{k}t^k\right\}^n,$$

$$(6.2) \quad g_{n,r}(t) = g_{n,r}(t; a) = \sum_{m=rn}^{\infty} Q_r(m, n)\frac{t^m}{m!}$$
$$= \frac{1}{n!}(1+t)^a\left\{e^t - \sum_{k=0}^{r-1}\frac{t^k}{k!}\right\}^n.$$

The generalized R-Q numbers retain many of the properties of the R-Q numbers and may be studied in a similar way.

Thus, differentiating (6.1) and (6.2) with respect to t, we obtain the difference-differential equations

$$(1+t)\frac{d}{dt} f_{n,r}(t) = (a + sn + at)f_{n,r}(t) + (s)_r \frac{t^{r-1}}{(r-1)!} f_{n,r-1}(t),$$

and

$$(1+t)\frac{d}{dt} g_{n,r}(t) = (a + n + nt)g_{n,r}(t) + (1+t)\frac{t^{r-1}}{(r-1)!} g_{n,r-1}(t),$$

which imply the following recurrence relations:

$$R_r(m+1, n) = (a + sn - m)R_r(m, n) + amR_r(m-1, n) + \binom{m}{r-1}(s)_r R_r(m-r+1, n-1), \quad m \ge rn+1;$$

$$R_r(rn+1, n) = (a + sn - rn)R_r(rn, n) + \binom{rn}{r-1}(s)_r R_r(rn-r+1, n-1);$$

$$R_r(rn, n) = \binom{rn-1}{r-1}(s)_r R_r(rn-r, n-1);$$

$$Q_r(m+1, n) = (a + n - m)Q_r(m, n) + nmQ_r(m-1, n) + \binom{m}{r-1}Q_r(m-r+1, n-1) + r\binom{m}{r}Q(m-r, n-1), \quad m \ge rn+1;$$

$$Q_r(rn+1, n) = (a + n - rn)Q_r(rn, n) + \binom{rn}{r-1}Q_r(rn-r+1, n-1) + r\binom{rn}{r}Q_r(rn-r, n-1);$$

$$Q_r(rn, n) = \binom{rn}{r}rQ_r(rn-r, n-1).$$

Notice also that the m-axis values for $R_r(m, n)$, $Q_r(m, n)$ are

$$R_r(m, 0) = a^m, \quad Q_r(m, 0) = (a)_m,$$

as may be readily verified from (6.1) and (6.2).

Another set of recurrences (with respect to r) useful for tabulation purposes is the following:

$$R_{r+1}(m, n) = \sum_{k=0}^{n}(-1)^k \frac{(m)_{rk}}{k!}\binom{s}{r}^k R_r(m-rk, n-k);$$

$$R_r(m, n) = \sum_{k=0}^{n}\frac{(m)_{rk}}{k!}\binom{s}{r}^k R_{r+1}(m-rk, n-k);$$

$$Q_{r+1}(m, n) = \sum_{k=0}^{n}(-1)^k \frac{(m)_{rk}}{k!(r!)^k} Q_r(m-rk, n-k);$$

$$Q_r(m, n) = \sum_{k=0}^{n}\frac{(m)_{rk}}{k!(r!)} Q_{r+1}(m-rk, n-k).$$

This set of recurrences results from the formulas:

$$f_{n,r+1}(t) = \sum_{k=0}^{n}(-1)^k\binom{s}{r}^k \frac{t^{rk}}{k!} f_{n-k,r}(t);$$

$$f_{n,r}(t) = \sum_{k=0}^{n}\binom{s}{r}^k \frac{t^{rk}}{k!} f_{n-k,r+1}(t);$$

$$g_{n,r+1}(t) = \sum_{k=0}^{n}(-1)^k \frac{t^{rk}}{k!(r!)^k} f_{n-k,r}(t);$$

$$g_{n,r}(t) = \sum_{k=0}^{n}\frac{t^{rk}}{k!(r!)^k} f_{n-k,r+1}(t).$$

It is also worth noticing that:

a. The generalized R-Q numbers are connected to the generalized C and Stirling numbers (see [8]) by relations analogous to those of Section 4 for the "non-generalized" quantities.

b. The form of the egf's (6.1) and (6.2) imply "proper" addition theorems for the generalized R-Q numbers, which lead to the definition of two multi-parameter discrete distributions with probability functions

$$f(x;\ m,\ n;\ a,\ b,\ r) = \binom{m}{x}\frac{b^{m-x}R_r(x,\ n;\ s,\ a)}{R_r(m,\ n;\ s,\ a+b)},\quad x = rn,\ rn+1,\ \ldots,\ m,$$

and

$$g(x;\ m,\ n;\ a,\ b) = \binom{m}{x}\frac{(b)_{m-x}Q_r(x,\ n;\ a)}{Q_r(m,\ n;\ a+b)},\quad x = rn,\ rn+1,\ \ldots,\ m.$$

c. The generalized R-Q numbers appear in the convolution of two samples coming from a binomial and a Poisson law, when one of the distribution laws is truncated on the left away from a given nonnegative integer r. More precisely, we have:

(i) If $X_1,\ X_2,\ \ldots,\ X_v$ is a random sample from the binomial distribution $b(N,\ p)$ and $X_{v+1},\ \ldots,\ X_{v+n}$ another independent sample from the Poisson distribution with parameter $\theta = p/(1-p)$, truncated away from r, i.e.,

$$P[X_i = x] = \left[e^{\theta} - \sum_{k=0}^{r-1}\frac{\theta^k}{k!}\right]^{-1}\frac{\theta^x}{x!},\quad \begin{array}{l} x = r,\ r+1,\ \ldots, \\ i = v+1,\ \ldots,\ v+n, \end{array}$$

then the distribution of the statistic $Z = X_1 + \cdots + X_{v+n}$ is given by

$$P[Z = z] = \frac{Q_r(z,\ n;\ a)}{g_{n,r}(\theta;\ a)}\frac{\theta^z}{z!},\quad z = rn,\ rn+1,\ \ldots\ .$$

(ii) If $X_1,\ X_2,\ \ldots,\ X_v$ is a random sample from the Poisson distribution with parameter θ, and $X_{v+1},\ \ldots,\ X_{v+n}$ another independent sample from the binomial law with probability function

$$P[X_i = x] = \left[(1+\theta)^N - \sum_{k=0}^{r-1}\binom{N}{x}\theta^k\right]^{-1}\binom{N}{x}\theta^x,\quad \begin{array}{l} x = r,\ r+1,\ \ldots, \\ i = v+1,\ \ldots,\ v+n, \end{array}$$

then the distribution of the statistic $Z = X_1 + \cdots + X_{v+n}$ is given by

$$P[Z = z] = \frac{R_r(z,\ n;\ N,\ n)}{f_{n,r}(\theta;\ N,\ v)}\frac{\theta^z}{z!},\quad z = rn,\ rn+1,\ \ldots\ .$$

d. The numbers $Q_r(m,\ n;\ a)$ admit a combinatorial interpretation similar to the one given for $Q(m,\ n;\ a)$ in Section 5c. In the expression "none of the n identical cells is empty," simply replace "is empty" by "contains less rhan r balls."

Acknowledgments

The author wishes to thank Professors T. Cacoullos and Ch. Charalambides for many helpful suggestions and critical comments during the preparation of the present work.

References

1. A. Broder. "The r-Stirling Numbers." *Discrete Math.* *49* (1984):241-59.
2. T. Cacoullos. "A Combinatorial Derivation of the Distribution of the Truncated Poisson Sufficient Statistic." *Ann. Math. Statist.* *32* (1961):904-05.
3. T. Cacoullos & Ch. A. Charalambides. "On Minimum Variance Unbiased Estimation for Truncated Binomial and Negative Binomial Distributions." *Ann. Inst. Statist. Math.* *27* (1975):235-44.
4. L. Carlitz. "Weighted Stirling Numbers of the First and Second Kind—I." *Fibonacci Quarterly* *18* (1980):147-62.

5. L. Carlitz. "Weighted Stirling Numbers of the First and Second Kind—II." *Fibonacci Quarterly* *18* (1980):242-57.
6. M. Chak. "A Class of Polynomials and a Generalization of Stirling Numbers." *Duke Math. J.* *23* (1956):45-55.
7. Ch. A. Charalambides. "Minimum Variance Unbiased Estimation for a Class of Left-Truncated Discrete Distributions." *Sankhya*, Ser. A, *36* (1974):397-418.
8. Ch. A. Charalambides. "The Generalized Stirling and C Numbers." *Sankhya*, Ser. A, *36* (1974):419-36.
9. Ch. A. Charalambides. "A New Kind of Numbers Appearing in the n-Fold Convolution of Truncated Binomial and Negative Binomial Distributions." *SIAM J. Appl. Math.* *33* (1977):297-98.
10. Ch. Charalambides & M. Koutras. "On the Differences of the Generalized Factorials at an Arbitrary Point and Their Combinatorial Applications." *Discrete Math.* *47* (1983):183-201.
11. Ch. Charalambides & J. Singh. "A Review of the Stirling Numbers, Their Generalizations and Statistical Applications." To appear in *Communications in Statistics*.
12. H. W. Gould & A. T. Hopper. "Operational Formulas Connected with Two Generalizations of Hermite Polynomials." *Duke Math. J.* *29* (1962):51-63.
13. F. Howard. "Degenerate Weighted Stirling Numbers." *Discrete Math.* *57* (1985):45-88.
14. C. Jordan. *Calculus of Finite Differences*. New York: Chelsea, 1960.
15. M. Koutras. "Non-Central Stirling Numbers and Some Applications." *Discrete Math.* *42* (1982):73-89.
16. I. Lah. "Eine neue Art von Zahlen, ihre Eigenschaften und Anwendung in der Mathematichen Statistik." *Mitteilungsbl. Math. Statist.* *7* (1955):203-12.
17. J. Riordan. *An Introduction to Combinatorial Analysis*. New York: Wiley, 1958.
18. R. Shanmugan. "On Central versus Factorial Moments." *South African Statist. J.* *18* (1984):97-110.
19. S. Tauber. "Lah Numbers for R-Polynomials." *Fibonacci Quarterly* *6* (1968): 100-07.

*****