A CLASS OF POWER SERIES q-DISTRIBUTIONS

Charalambos A. Charalambides
Department of Mathematics, University of Athens
GR-15784 Athens, Greece

E-mail address:ccharal@math.uoa.gr

Contents

Preliminary notions and remarks

Distributions on i.i.d Bernoulli trials

The classical binomial and negative binomial (or Pascal) distributions are defined on the stochastic model of a sequence of independent and **identically** distributed Bernoulli trials.

The Poisson distribution may be considered as a limiting case of the binomial distribution as the number of trials tends to infinity.

Also, the logarithmic distribution may be considered as a limiting case of the zero-truncated negative binomial distribution as the number of successes tends to zero.

Generalized binomial and negative binomial distributions

Poisson (1837), generalized the binomial distribution (and implicitly the negative binomial distribution) by assuming that the probability of success at a trial varies with the number of previous trials.

Woodbury (1949) generalized the negative binomial distribution (and implicitly the binomial distribution) by assuming that the probability of success at a trial varies with the number successes occurring in the previous trials.

The probability functions of these distributions were derived in terms of generalized Stirling numbers of the first and second kind; these numbers, which are explicitly given by **multiple sums**, constitute elementary and product sum **symmetric functions**.

The meaning of "implicitly"

Let X_n be the number of successes in n trials, and W_n be the number of failures until the occurrence of the nth success. Then,

$$P(W_n = w) = P(X_{n+w-1} = n-1)P(n\text{th success at the } (n+w)\text{th trial})$$

and, equivalently,

$$P(X_n = x) = P(W_{x+1} = n-x)/P((x+1)\text{st success at the } (n+1)\text{st trial}).$$

Note that instead of W_n, the number T_n of trials until the occurrence of the nth success may be considered. Clearly, $\{T_n = t\} = \{W_n = t - n\}$ and the preceding expressions are accordingly modified.

Usefulness of generalizations

It should be noticed that a stochastic model of a sequence of independent Bernoulli trials, in which the probability of success at a trial is assumed to vary with the number of trials and/or the number of successes, is advantageous in the sense that it permits incorporating the experience gained from previous trials and/or successes.

If the probability of success at a trial is a very general function of the number of trials and/or the number successes, very little can be inferred from it about the distributions of the various random variables that may be defined on this model.

Success probability varying geometrically

In order to be able to infer more about these distributions, the general expression of the probability of success in terms of the number of trials and/or the number successes should be restricted.

The assumption that the probability of success (or failure) at a trial varies geometrically, with rate (proportion) q, leads to the introduction of discrete q-distributions.

The study of these distributions is greatly facilitated by the wealth of existing q-sequences and q-functions, in q-combinatorics, and the theory of q-hypergeometric series.

Selected classical discrete q-distributions

q-Numbers and q-factorials

Let x and q be real numbers, with $q \neq 1$, and k be an integer. The number

$$[x]_q = \frac{1 - q^x}{1 - q}$$

is called *q-number* and in particular $[k]_q$ is called *q-integer*. Note that $\lim_{q \to 1} [x]_q = x$.

The kth order factorial of the q-number $[x]_q$, which is defined by

$$[x]_{k,q} = [x]_q [x-1]_q \cdots [x-k+1]_q, \quad k = 1, 2, \ldots,$$

is called *q-factorial* of x of order k. In particular,

$$[k]_q! = [1]_q [2]_q \cdots [k]_q, \quad k = 1, 2, \ldots,$$

is called *q-factorial* of k.

q-Binomial coefficients and q-Pascal's triangle

The *q-binomial coefficient* (or *Gaussian polynomial*) is defined by

$$\begin{bmatrix} x \\ k \end{bmatrix}_q = \frac{[x]_{k,q}}{[k]_q!}, \quad k = 0, 1, \ldots,$$

and so $\lim_{q \to 1} \begin{bmatrix} x \\ k \end{bmatrix}_q = \binom{x}{k}$, $k = 0, 1, \ldots$.

The q-binomial coefficient satisfies the triangular recurrence relation

$$\begin{bmatrix} x \\ k \end{bmatrix}_q = \begin{bmatrix} x-1 \\ k \end{bmatrix}_q + q^{x-k} \begin{bmatrix} x-1 \\ k-1 \end{bmatrix}_q, \quad k = 1, 2, \ldots, \quad \begin{bmatrix} x \\ 0 \end{bmatrix}_q = 1.$$

Alternatively

$$\begin{bmatrix} x \\ k \end{bmatrix}_q = q^k \begin{bmatrix} x-1 \\ k \end{bmatrix}_q + \begin{bmatrix} x-1 \\ k-1 \end{bmatrix}_q, \quad k = 1, 2, \ldots, \quad \begin{bmatrix} x \\ 0 \end{bmatrix}_q = 1.$$

The lack of uniqueness of q-analogues of expressions and formulae

The lack of uniqueness, due to the presence of powers of q in pseudo-isomorphisms as

$$[x+y]_q = [x]_q + q^x [y]_q \quad \text{and} \quad [x+y]_q = q^y [x]_q + [y]_q,$$

where $0 < q < 1$ or $1 < q < \infty$, should be remarked from the very beginning of the presentation of the basic q-sequences, q-functions and q-formulae. It should also be noticed that the two formulae may be considered as *equivalent* in the sense that any of these implies the other by replacing the base q by q^{-1}.

q-Binomial distribution of the first kind

Kemp and Kemp (1991), in their study of the Weldon's classical dice data, introduced a q-binomial distribution. Specifically, they observed that there is a spectrum of unfairness among the dice and in order to explain it made the a log-linear odds assumption

$$\log\theta_i = \log\theta + (i-1)\log q, \quad i = 1, 2, \ldots,$$

for $0 < \theta < \infty$ and $0 < q < 1$ or $1 < q < \infty$, which implies that the odds of success at the ith trial, $\theta_i = p_i/(1-p_i)$, varies geometrically with the number i of trials,

$$\theta_i = \theta q^{i-1}, \quad i = 1, 2, \ldots,$$

for $0 < \theta < \infty$ and $0 < q < 1$ or $1 < q < \infty$.

q-Binomial distribution of the first kind (cont)

Equivalently, the probability of success at the ith trial is given by

$$p_i = \theta q^{i-1}/(1+\theta q^{i-1}), \quad i = 1, 2, \ldots .$$

The distribution of the number of successes in a sequence of n independent Bernoulli trials, with the odds of success at a trial varying geometrically with the number of trials, is the q-binomial distribution of the first kind. Kemp and Newton (1990) further studied it as stationary distribution of a birth and death process. Its probability function is given by

$$p_x(\theta; q) = \begin{bmatrix} n \\ x \end{bmatrix}_q \frac{\theta^x q^{\binom{x}{2}}}{\prod_{i=1}^{n}(1+\theta q^{i-1})}, \quad x = 0, 1, \ldots, n,$$

where $0 < \theta < \infty$, and $0 < q < 1$ or $1 < q < \infty$.

Negative q-binomial distribution of the second kind

Charalambides (2010) in his study of the q-Bernstein polynomials as a q-binomial distribution of the second kind, introduced the negative q-binomial distribution of the second kind. Precisely, it is assumed that the probability of success at any Bernoulli trial, given that $j-1$ successes occur in the previous trials is given by

$$p_j = 1 - \theta q^{j-1}, \quad j = 1, 2, \ldots, \quad 0 < \theta < \infty, \quad 0 < q < 1.$$

The distribution of the number of failures until the occurrence of the nth success in a sequence of independent Bernoulli trials, with the probability of success at a trial varying geometrically with the number of successes is the negative q-binomial distribution of the second kind.

Negative q-binomial distribution of the second kind (cont)

Its probability function is given by

$$p_x(\theta; q) = \begin{bmatrix} n + x - 1 \\ x \end{bmatrix}_q \theta^x \prod_{i=1}^{n} (1 - \theta q^{i-1}), \quad x = 0, 1, \dots,$$

where $0 < \theta < 1$ and $0 < q < 1$.

q-Poisson distributions

Benkherouf and Bather (1998) derived the Heine and Euler distributions, which constitute q-analogs of the Poisson distribution, as feasible priors in a simple Bayesian model for oil exploration. The probability function of the q-Poisson distributions is given by (Charalambides (2016), p. 107)

$$p_x(\lambda; q) = E_q(-\lambda)\frac{\lambda^x}{[x]_q!}, \quad x = 0, 1, \ldots,$$

where $0 < \lambda < 1/(1-q)$ and $0 < q < 1$ (Euler distribution) or $0 < \lambda < \infty$ and $1 < q < \infty$ (Heine distribution).

Also, $E_q(t) = \prod_{i=1}^{\infty}(1 + t(1-q)q^{i-1})$ is a q-exponential function. It should be noted that $e_q(t) = \prod_{i=1}^{\infty}(1 - t(1-q)q^{i-1})^{-1}$ is another q-exponential function and that these q-exponential functions are connected by $E_q(t)e_q(-t) = 1$ and $E_{q^{-1}}(t) = e_q(t)$.

q-Logarithmic distribution

A q-logarithmic distribution was studied by C. D. Kemp (1997) as a group size distribution. Its probability function is given by

$$p_x(\theta; q) = [-l_q(1-\theta)]^{-1} \frac{\theta^x}{[x]_q}, \quad x = 1, 2, \ldots,$$

where $0 < \theta < 1$, $0 < q < 1$, and

$$-l_q(1-\theta) = \lim_{x \to 0} \frac{1}{[x]_q} \left(\prod_{i=1}^{\infty} \frac{1 - \theta q^{x+i-1}}{1 - \theta q^{i-1}} - 1 \right) = \sum_{j=1}^{\infty} \frac{\theta^j}{[j]_q}$$

is a q-logarithmic function.

Power series q-distributions

The class of power series q-distributions, which is introduced in the sequel, provides a unified approach to the study of these distributions. Its probability generating function and q-factorial moments are derived.

Demonstrating this approach, the probability generating function and q-factorial moments of the q-Poisson (Heine and Euler), q-binomial, negative q-binomial, and q-logarithmic distributions are obtained.

q-Taylor expansion

Consider a positive function $g(\theta)$ of a positive parameter θ and assume that it is analytic with a q-Taylor expansion

$$g(\theta) = \sum_{x=0}^{\infty} a_{x,q}\theta^x, \quad 0 < \theta < \rho, \quad \rho > 0, \tag{1}$$

where the coefficient

$$a_{x,q} = \frac{1}{[x]_q!}[D_q^x g(t)]_{t=0} \geq 0,\ x = 0, 1, \ldots,\ 0 < q < 1, \text{ or } 1 < q < \infty,$$

with $D_q = d_q/d_q t$ the q-derivative operator,

$$D_q g(t) = \frac{d_q g(t)}{d_q t} = \frac{g(t) - g(qt)}{(1-q)t},$$

does not involve the parameter θ.

Definition of a family of power series q-distributions

Clearly, the function

$$p_x(\theta; q) = \frac{a_{x,q}\theta^x}{g(\theta)}, \quad x = 0, 1, \ldots, \tag{2}$$

with $0 < q < 1$ or $1 < q < \infty$, and $0 < \theta < \rho$, satisfies the properties of a probability (mass) function.

Definition

A family of discrete q-distributions $p_x(\theta; q)$, $\theta \in \Theta$, $q \in Q$, is said to be a class of power series q-distributions, with parameters θ, q and series function $g(\theta)$ if it has the representation (2), with series function satisfying condition (1).

Probability generating function and q-factorial moments

The probability generating function $P(t) = \sum_{x=0}^{\infty} p_x(\theta; q) t^x$, on using (1) and (3), is readily deduced as

$$P(t) = \frac{g(\theta t)}{g(\theta)}. \tag{3}$$

Taking its mth q-derivative, with respect to t, at $t = 1$, the mth q-factorial moment of the power series q-distribution is obtained as

$$E([X]_{m,q}) = \frac{1}{g(\theta)} \cdot \left[\frac{d_q^m g(\theta t)}{d_q t^m} \right]_{t=1} = \frac{\theta^m}{g(\theta)} \cdot \frac{d_q^m g(\theta)}{d_q \theta^m}, \; m = 1, 2, \ldots . \tag{4}$$

q-Mean and q-variance

In particular the q-mean is given by

$$E([X]_q) = \frac{\theta}{g(\theta)} \cdot \frac{d_q g(\theta)}{d_q \theta}. \tag{5}$$

Also, on using the expression

$$V([X]_q) = qE([X]_{2,q}) - E([X]_q)\big(E([X]_q) - 1\big),$$

the q-variance is obtained as

$$V([X]_q) = \frac{q\theta^2}{g(\theta)} \cdot \frac{d_q^2 g(\theta)}{d_q \theta^2} - \frac{\theta}{g(\theta)} \cdot \frac{d_q g(\theta)}{d_q \theta} \left(\frac{\theta}{g(\theta)} \cdot \frac{d_q g(\theta)}{d_q \theta} - 1 \right). \tag{6}$$

q-Poisson distributions

These are power series q-distributions, with series function

$$g(\lambda) = e_q(\lambda) = 1/E_q(-\lambda),$$

where $0 < \lambda < 1/(1-q)$ and $0 < q < 1$ or
$0 < \lambda < \infty$ and $1 < q < \infty$.
Since $D_q e_q(t) = e_q(t)$ and $e_q(0) = 1$, it follows from (2) that

$$a_{x,q} = \frac{1}{[x]_q!}[D_q^x e_q(t)]_{t=0} = \frac{1}{[x]_q!}, \quad x = 0, 1, \ldots,$$

Also, the probability generating function of the q-Poisson distributions, on using (3), is deduced as

$$P(t) = \frac{e_q(\lambda t)}{e_q(\lambda)} = E_q(-\lambda)e_q(\lambda t).$$

q-Poisson distributions (cont)

The q-factorial moments, since $D_q^m e_q(\lambda) = e_q(\lambda)$, are readily deduced as

$$E([X]_{m,q}) = \lambda^m, \quad m = 1, 2, \dots .$$

In particular, the q-mean is given by

$$E([X]_q) = \lambda.$$

Also, using (6), the q-variance is obtained as

$$V([X]_q) = q\lambda^2 - \lambda(\lambda - 1) = \lambda(1 + (q - 1)\lambda).$$

q-Binomial distribution of the first kind

The series function of this distribution is

$$g(\theta) = \prod_{i=1}^{n} (1 + \theta q^{i-1}),$$

where $0 < \theta < \infty$ and $0 < q < 1$ or $1 < q < \infty$. Since

$$D_q g(\theta) = [n]_q \prod_{i=1}^{n-1} (1 + (\theta q) q^{i-1}),$$

it follows successively that

$$D_q^x g(\theta) = [n]_{x,q} q^{\binom{x}{2}} \prod_{i=1}^{n-x} (1 + (\theta q^x) q^{i-1}),$$

for $x = 1, 2, \ldots, n$. Thus,

$$a_{x,q} = \frac{1}{[x]_q!} [D_q^x g(t)]_{t=0} = \begin{bmatrix} n \\ x \end{bmatrix}_q q^{\binom{x}{2}}, \quad x = 0, 1, \ldots, n,$$

q-Binomial distribution of the first kind (cont)

The probability generating function of the q-binomial distribution of the first kind is readily deduced as

$$P(t) = \frac{\prod_{i=1}^{n}(1+\theta t q^{i-1})}{\prod_{i=1}^{n}(1+\theta q^{i-1})}.$$

The q-factorial moments, since

$$D_q^m g(\theta) = [n]_{m,q} q^{\binom{m}{2}} \prod_{i=1}^{n-m}(1+(\theta q^m)q^{i-1}) = [n]_{m,q} q^{\binom{m}{2}} \prod_{i=m+1}^{n}(1+\theta q^{i-1}),$$

are obtained as

$$E([X]_{m,q}) = \frac{[n]_{m,q}\theta^m q^{\binom{m}{2}}}{\prod_{i=1}^{m}(1+\theta q^{i-1})}, \quad m = 1, 2, \ldots .$$

q-Binomial distribution of the first kind (cont)

In particular, the q-mean is

$$E([X]_q) = \frac{[n]_q\theta}{(1+\theta)}.$$

Also, using (6), the q-variance is obtained as

$$V([X]_q) = \frac{[n]_q\theta}{(1+\theta)(1+\theta q)}\left(1 + \frac{[n]_q\theta(q-1)}{1+\theta}\right).$$

Negative q-binomial distribution of the second kind

It is a power series q-distribution, with series function

$$g(\theta) = \prod_{i=1}^{n}(1-\theta q^{i-1})^{-1}, \quad 0<\theta<1, \quad 0<q<1.$$

Since

$$D_q g(\theta) = [n]_q \prod_{i=1}^{n+1}(1-\theta q^{i-1})$$

it follows successively that

$$D_q^x g(\theta) = [n+x-1]_{x,q} \prod_{i=1}^{n+x}(1-\theta q^{i-1}),$$

for $x = 1, 2, \ldots$. Thus,

$$a_{x,q} = \frac{1}{[x]_q!}[D_q^x g(t)]_{t=0} = \begin{bmatrix} n+x-1 \\ x \end{bmatrix}_q, \quad x = 0, 1, \ldots .$$

Negative q-binomial distribution of the second kind (cont)

The probability generating function of the negative q-binomial distribution of the second kind is deduced as

$$P(t) = \frac{\prod_{i=1}^{n}(1-\theta t q^{i-1})^{-1}}{\prod_{i=1}^{n}(1-\theta q^{i-1})^{-1}}.$$

The q-factorial moments, since

$$D_q^m g(\theta) = [n+m-1]_{m,q} \prod_{i=1}^{n+m}(1-\theta q^{i-1})^{-1},$$

are obtained as

$$E([X]_{m,q}) = [n+m-1]_{m,q}\theta^m \prod_{i=1}^{m}(1-\theta q^{n+i-1})^{-1}, \quad m = 1, 2, \ldots .$$

Negative q-binomial distribution of the second kind (cont)

In particular, the q-expected value is

$$E([X]_q) = \frac{[n]_q\theta}{1-\theta q^n}.$$

Also, the q-variance obtained as

$$V([X]_q) = \frac{[n]_q\theta}{(1-\theta q^n)(1-\theta q^{n+1})}\left(1 + \frac{[n]_q\theta(q-1)}{1-\theta q^n}\right).$$

q-Logarithmic distribution

The series function of this distribution is

$$g(\theta) = -l_q(1-\theta) = \sum_{j=1}^{\infty} \frac{\theta^j}{[j]_q}, \quad 0 < \theta < 1, \quad 0 < q < 1.$$

Taking successively its q-derivatives,

$$D_q^x g(\theta) = \sum_{j=x}^{\infty} [j-1]_{x-1,q} \theta^{j-x} = [x-1]_q! \sum_{j=x}^{\infty} \begin{bmatrix} j-1 \\ j-x \end{bmatrix}_q \theta^{j-x},$$

and using the negative q-binomial formula, we find

$$D_q^x g(\theta) = [x-1]_q! \prod_{i=1}^{x} (1-\theta q^{i-1})^{-1}.$$

q-Logarithmic distribution (cont)

Thus,

$$a_{x,q} = \frac{1}{[x]_q!}[D_q^x g(t)]_{t=0} = \frac{1}{[x]_q}, \quad x = 1, 2, \ldots .$$

Also, the probability generating function of the q-logarithmic distribution is deduced as

$$P(t) = \frac{-l_q(1-\theta t)}{-l_q(1-\theta)}.$$

The q-factorial moments, are obtained as

$$E([X]_{m,q}) = \frac{[-l_q(1-\theta)]^{-1}[m-1]_q!\theta^m}{\prod_{i=1}^{m}(1-\theta q^{i-1})}, \quad m = 1, 2, \ldots .$$

q-Logarithmic distribution (cont)

In particular, the q-mean value is

$$E([X]_q) = \frac{[-l_q(1-\theta)]^{-1}\theta}{1-\theta}.$$

Also, the q-variance is obtained as

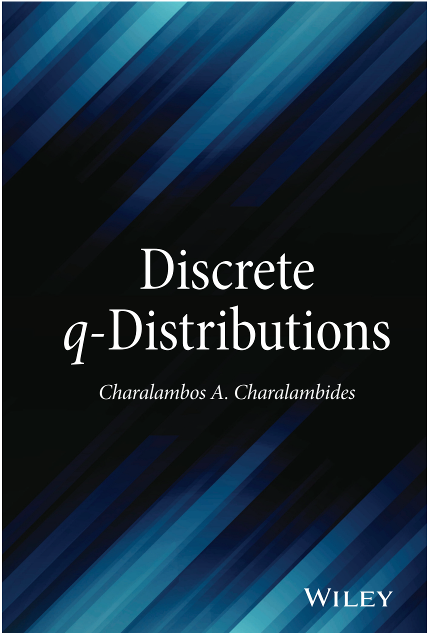
$$V([X]_q) = \frac{[-l_q(1-\theta)]^{-1}\theta}{1-\theta}\left(\frac{1}{1-\theta q} - \frac{[-l_q(1-\theta)]^{-1}\theta}{1-\theta}\right).$$

Bibliography

- Benkherouf, L. and Bather, J. A. (1988) Oil exploration: sequential decisions in the face of uncertainty. *J. Appl. Probab.* **25**, 529-543.
- Charalambides, Ch. A. (2010) The q-Bernstein basis as a q-binomial distribution, *J. Statist. Plann.Inference* **140**, 2184-2190.
- Charalambides, Ch. A. (2016) *Discrete q-Distributions*, John Wiley & Sons, Hoboken, New Jersey.
- Kemp, C. D. (1997) A q-logarithmic distribution. In Balakrishnan, N. (Ed.) *Advances in Combinatorial Methods and Applications to Probability and Statistics*, Birkhäuser, Boston, MA, pp. 465-470.

Bibliography (cont)

- Kemp, A. and Kemp, C. D. (1991) Weldon's dice data revisited. *Amer. Statist.* **45**, 216-222.
- Kemp, A. and Newton, J. (1990) Certain state-dependent processes for dichotomized parasite population. *J. Appl. Probab.* **27**, 251-258.


Discrete
q-Distributions
Charalambos A. Charalambides
WILEY