DEVELOPMENTS IN DISCRETE DISTRIBUTIONS 1969-1980. (U)
SEP 81 N L JOHNSON, S KOTZ
M/5-81-034
UNCLASSIFIED
DEVELOPMENTS IN DISCRETE DISTRIBUTIONS 1969-1980

Norman L. Johnson
University of North Carolina at Chapel Hill

Samuel Kotz
University of Maryland, College Park

Abstract

In this paper we survey and summarize developments in theory and methodology of discrete distributions during the period 1969-1980 since publication of our book Distributions in Statistics - Discrete Distributions (Wiley, 1969). A comprehensive (though not exhaustive) bibliography of some 600 items is included.

September 1980
techniques for closely related distributions.

The qualification "appropriate" in the previous paragraph is important.
We do not regard the mere number and variety of distributions included in a
class (or system) as being, in themselves measures of its importance - still
less of its practical value. (For example the class for which \( \nu \sim \text{Pr} \{ x \} = 1 \)
in very broad, but of no value for our purposes.) What is important, and useful,
is inclusion of as wide a variety of distributions as possible, within
an explicit formulation as possible. This removes the need for piecemeal
definitions and represents considerable economy of effort.

2. Notation

Unless otherwise indicated, we will always assume that the distributions
we discuss here are lattice distributions over non-negative values of the variable
\( n = 0, 1, 2, \ldots \).

We will use the abbreviations:

- \( F_n \) for \( \text{Pr} \{ x = n \} \); \( F_n \) for \( \text{Pr} \{ X = n \} \);
- cdf for "cumulative distribution function";
- pdf for "probability density function" (of a continuous variable);
- mgf for "moment generating function" \( (E[x^n]) = E[F_n x^n] \);
- pgf for "probability generating function" \( (E[x^n]) = E[F_n x^n] \);
- \( F_n \sim \text{distn} \) for "is distributed as".

The distribution obtained by describing a distribution \( F_n \) on parameter \( 
\alpha \) of a distribution \( F_\alpha \) will be denoted

\[ F_\alpha \sim F_n \quad \text{(1)} \]

This is called a "compound" \( F_\alpha \) distribution; \( F_n \) is the compounding distribution.

The distribution \( F_\alpha \) is often called a "\( F_\alpha \)-generalized" distribution (though sometimes,
and in our opinion preferably, the order is reversed).

Denoting the pgf corresponding to a distribution \( F_1 \) by \( G_1(s) \) the distri-
bution with pgf obtained by replacing \( s \) in \( G_1(s) \) by \( C_\alpha(s) \) (giving \( C_\alpha[G_1(s)] \))
is termed the \( F_\alpha \)-generalized \( F_1 \) distribution, and denoted

\[ F_1 = F_{\alpha} \quad \text{(2)} \]

The distribution \( F_\alpha \) is termed the generalizing distribution. (Note that while
\( F_\alpha \) in (1) may be continuous, \( F_\alpha \) in (2) must be a discrete lattice distribution.)

A simple interpretation of (2) is obtained by noting that \( G_{\alpha}(s) \) is the pgf of
the sum (convolution) of \( N \) independent \( F_\alpha \) variables, so that (2) is the distri-
bution of a "random sum" of \( N \) such variables, the number of variables in the sum
itself having a \( F_1 \) distribution. Kendall and Puri (1970) and Chesterfield and Theodoss
(1973) have suggested the same random sum distribution for such distributions, as
preference to "\( F_\alpha \)-generalised", as the latter term may be confused with other general-
izations of a different nature. Douglas (1970) refers to these distributions as
randomly stopped sums. Chesterfield and Theodoss (1973) also suggest the use of the term
"mixture distributions" in place of "compound distributions". Regarding \( \alpha \) as a para-
meter of the \( \alpha \)-th convolution, \( F_{\alpha \alpha} \) may, of \( F_1 \), (2) is also the compound distribution

\[ F_{\alpha \alpha} \sim F_{\alpha} \quad \text{(3)} \]

We will use some fairly standard notations for distributions such as

- Poisson \( (\theta) \quad \text{pdf} \quad \phi(x) = e^{-\theta} \theta^x / x! \quad (x = 0, 1, 2, \ldots) \)
- Binomial \( (n, p) \quad \text{pdf} \quad \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad (x = 0, 1, 2, \ldots, n) \)
- Negative Binomial \( (n, p) \quad \text{pdf} \quad \frac{\theta^n (1-\theta)^x}{x!} \quad (x = 0, 1, 2, \ldots) \)
- Logistic \( (\theta) \quad \text{pdf} \quad \phi(x) = \frac{1}{(1+e^{-\theta})} \frac{1}{\theta} \quad (x = 0, 1, 2, \ldots) \)
- Lognormal \( (a, b) \quad \text{pdf} \quad \phi(a, b) = \frac{1}{b \sqrt{2\pi}} e^{-(a-b)^2 / (2b^2)} \quad (a, 1, 2, \ldots) \)
Multinomial (n;p_1,...,p_k) \[ P_k = \binom{n}{x_k} \prod_{i=1}^{k} \frac{x_k!}{n!} \left( p_i^{x_k} \right) \] 

\[ \mathbb{E}(X) = \lambda \] 

\[ \text{Var}(X) = \text{Var}(\mathbb{E}(X)) + \text{Var}(\text{Var}(X)) \] 

\[ \text{Cov}(X,Y) = \text{Var}(X) \rho \] 

\[ \text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\text{SD}(X) \text{SD}(Y)} \] 

\[ \text{Prob}(X > Y) = \int_{-\infty}^{\infty} f_X(x) \left( 1 - F_Y(x) \right) dx \] 

\[ \text{Prob}(X < Y) = \int_{-\infty}^{\infty} f_Y(x) \left( 1 - F_X(x) \right) dx \] 

\[ \text{Prob}(X = Y) = \int_{-\infty}^{\infty} f_X(x) f_Y(x) dx \] 

\[ \text{Prob}(X \neq Y) = 1 - \text{Prob}(X = Y) \] 

In the example case when a distribution (P_0) is modified by increasing P_0 and compensating for this by multiplying the remaining P_j's by an appropriate constant, the resultant distribution (P'_0) is called a deflected (P') distribution. We have

\[ P'_0 = (1-a)P_0 \]

\[ P'_j = \frac{aP_j}{(1-a)} \] 

with 0 < a < 1. The pdf is C^r(a) = (1-a)^{r-1} + aC(a). If a > 1, the result is called a deflected distribution. The greatest possible value of a is \((1-P_0)^{-1}\).

If a takes this value, \(P'_0 = 0\).

If P_0 = 0 (the case most commonly considered) and P'_0 = 0, then we have a zero-truncated - or, more prolixly, "decremented" - \(P'_0\) distribution. Yet another name is "positive \(P'_0\)" distribution.

If \(P_0 = 0\) and \(0 < a < 1\) the phrase "with added success" is sometimes used.

General relations among the moments of \(P_0\) and \(P'_0\) include:

\[ E[X^r] = \frac{E[X^r]}{E[X]} \] 

\[ E[X^r] = \frac{E[X^{r-1}]}{E[X]} \] 

If \(P_0 = 0\)

\[ E[X^r] = E[X^{r-1}] \] 

If there is a recurrence relation \(g(x) \rightarrow \mathfrak{R}^2 = 0\) for \(P_0\), then

\[ E[X^r] = \frac{E[X^{r-1}]}{E[X]} \] 

If \(P_0 = 0\) we have
Another type of imperfection arises when observation is not always accurate. An example is the "faulty inspection" situation (Johnson et al. (1960)) wherein the chance of detecting that an item is defective is not 1, even when that item is selected in the sample. If the probability of detection is p, and the distribution of the actual number of defectives in the sample is \( F \), then the observed distribution is the compound

\[
\text{Binomial}(V,p) \approx F,
\]

(For sampling without replacement from a lot of size \( N \) containing \( D \) defectives, \( F \) would be Hypergeometric (\( N,D,N \) where \( N \) is the sample size.)

These models belong to the general class of "damage models", in which the observed value \( \bar{X} \) is not, in general, the original value \( X \), but that is left after some damage, resulting in reduction of \( Y \) to \( \bar{X} \), has been experienced, so that

\[
\bar{X} \times \bar{Y} = X \times Y,
\]

where \( X \) is the amount removed by damage. For example, in faulty inspection,

\( Y \) is the actual number of defectives in a sample and \( X \) the number detected. The "damage" is incurred by the imperfection of the inspection.

Damage models have been studied extensively, mainly with a view to obtaining conditions characterising the distribution of \( Y \) from that of \( X \). This work originated with Rao and Rubin (1956, page 94) and has been progressively extended, and in some cases, proofs simplified by Govindarajulu and Leslie (1970), Valuvalker (1970, 1980), Scodel (1975), Shambhag (1974, 1977), Bishnupriya (1974), Aced (1972, 1977), Consul (1974), Srivastava and Singh (1974), Puri and Ratanaperthi (1975), Panaretos (1980), and Shambhag and Tallie (1980).

The earlier papers were mostly concerned with results of the form: assuming \( X \) and \( Y \) to be mutually independent and the conditional distribution of \( X \), given \( (X+Y) \), is

(i) Binomial, then \( X \) and \( Y \) are each Poisson
(ii) Hypergeometric, then \( X \) and \( Y \) are each negative binomial
(iii) \( q \)-Poisson, then \( X \) and \( Y \) are each hypergeometric.

Consul (1974b) and Janardan (1974) have obtained results of similar type for the (Lagrangean) double Poisson, quasi-binomial, quasi-hypergeometric and quasi-Poisson (see Section 14) distributions. Multivariate extensions have been obtained by Shambhag (1974a) and Panaretos (1980).

4. Negative Moments

Chao and Straussman (1972) showed that, if \( r \) is a positive integer

\[
R[(X+a)^{-r}] = \sum_{k=0}^{\infty} a_r(s) s^k,
\]

where \( a_r(s) \) can be calculated from

\[
a_r(s) = E[(X+a)^{-r}] = s^{-1} G(s),
\]

and

\[
a_r(s) = s^{-1} \sum_{k=0}^{\infty} a_r(s) s^k (k=1,2,...r-1), \quad a \neq 0.
\]

Kabe (1976), generalising Stone's (1968) approach, obtained the formula

\[
R[(X+a)^{-r}] = \sum_{j=0}^{\infty} \frac{b_r-j}{(j+1)^{r+1}} S(j), \quad a \neq 0.
\]

In particular cases, special techniques may be used. For example, Gupta (1979) uses a recurrence relation to obtain the expected value of the reciprocal of a zero-truncated generalized Lagrangean double Poisson variable. See Baru and Consul (1979) on modified power series with applications to generalized...
(lognormal) double Poisson and binomial distributions. See also, Law (1974) for some general bounds on inverse moments.

5. Minimum Variance Unbiased Estimators (MVUE)

A considerable amount of work has been done on the construction of MVUE's, but little attention has been paid to possible usefulness of these estimators in particular cases. In many cases the sum (n-fold convolution)

\[ S_n = \sum_{j=1}^{n} X_j \]

of n independent variables each having the distribution \( F \) is a sufficient statistic for the parameter \( \theta \) being estimated, and so the MVUE is a function of \( S_n \). As a consequence the distribution of \( S_n \) is of importance. Particular attention has been devoted to determining convolutions of decapitated, and other truncated distributions (Ahoja 1971b; Ahoja and Emekking 1974; Bache and Robin 1972).

If \( P_x \) is of the general linear exponential form, then

\[ P_x = g_1(0)g_2(x) \exp(g_2(0)) \]

\[ Pr[S_n = x] = (g_1(0))^{n} \frac{g_2(x)}{g_2(0)} \exp(g_2(0)) \]

where \( g_2(x) = \sum_{j=1}^{n} g_2(x_j) \). From (13) it can be seen that \( S_n \) is a sufficient statistic for \( \theta \).

If \( n \) now \( (P_n) \) be truncated by changing \( g_2(x) \) to zero for some \( x \)'s and changing

\[ g_2(x) = \left( \sum_{x=0}^{n} g_2(x) \exp(g_2(0)) \right)^{-1} \]

the \( g_2(x) \), appropriately then \( S_n \) is still sufficient for \( \theta \). If we have zero-truncation so that

\[ Pr[S_n = x] = (g_1(0))^{n} \frac{g_2(x)}{g_2(0)} \exp(g_2(0)) \]

(we, n=1, \ldots)

with \( S_n(x) \) obtained from \( S_n(x) \) by omitting all terms in the summation in which any \( x_j \) is zero, the essential technical problem in determining the convolution of the decapitated variable is the evaluation of this sum. If \( S_n(x) \) (for the untruncated distribution) is known we can use the recurrence formula

\[ S_n(x) = S_n(x) - n g_2(0) S_{n-1}(x) - \left( \frac{n}{2} \right) g_2(0)^2 S_{n-2}(x) \ldots \]

(15)

to determine \( S_n(x) \)'s progressively, starting from \( S_1(x) = S_1(x) \).

6. Infinite Divisibility

Marde and Katti (1971) show that the discrete distribution \( P_n \) on \( x = 0, 1, \ldots \), with \( P_0 \neq 0 \) and \( P_1 \neq 0 \) is infinitely divisible if the sequence \( (P_{n+1}/P_n) \) increases with \( n \). This condition is, in fact, quite a strong one; it requires that \( (P_{n+1}/P_n) \) be a decreasing sequence. However, it is satisfied by geometric and logseries distributions, later a.lso. Marde and Katti also give a method for constructing a new infinitely divisible distribution from a known one.

James (1974) shows that if \( P_0 \neq 0 \) then a necessary condition for infinite divisibility is

\[ P_0 \text{ var}(x) = P_1 \]

If this is an equality, the distribution must be Poisson. (See also Schwaner (1970).)

Bendelsen (1972, 1979) has introduced the idea of a generalized negative binomial convolution (g.n.b.c.). "Generalized" applies to "convolution" and not to "negative binomial". A g.n.b.c. is any distribution which can be obtained as a "weak limit of finite convolutions of negative binomial distributions". He demonstrates the parallelism of g.n.b.c.'s with "generalized gamma convolutions" defined by Therri (1978). Among distributions identified as being g.n.b.c.'s, and as infinitely divisible, is
\[ P_n = \sum_{k=0}^{n} b(k) k^{n-1} \]

where \( a_j, b_j > 0 \) and \( 1 \leq n \leq \frac{1}{2} y_j \).

7. A General Result

Hoensmann and Wormalth (1979) have shown that

\[ P_n = \left( (n+1) \frac{a_j}{b_j} \right)^{-1} (n+1)^{-1} \]

where \( p \) is the integer part of \( n+1 \), for any discrete non-negative unit lattice distribution.


Among discrete distributions, the class of power series distributions (Patil 1962, 1963) is one of the most useful, in the sense described in Section 1.1. They are defined by

\[ P_n = \sum_{k=0}^{n} b(k) k^{n-1} \]

where \( b(k) \) and \( b(k) \) are called the generating function and the parameter, respectively.

of the distribution. Since \( \sum_{n=0}^{\infty} P_n = 1 \)

\[ g(0) = \frac{1}{2} \sum_{n=0}^{\infty} b(n) \]

and so, subject to some regularity conditions

\[ b(0) = \frac{\partial^2 g(0)}{\partial x^2} \mid_{x=0} \]

If \( X \) has distribution (19), we write

\[ X \sim \text{PSD}(n; g(0)) \]

Modified power series distributions (Gupta 1974b) are distributions with \( u \) in (19) replaced by \( (u(n))^{1/2} \). Given \( g(0) \), values of \( b(n) \) can be obtained, if a Lagrange expansion (see (571)) can be used, from the formula

\[ b(n) = \frac{\partial^n}{\partial x^n} \left( e^{x} \right) \left( \frac{g(x)}{g(0)} \right) \mid_{x=0} \]

See also H. C. Gupta (1974b) and Kumar and Consul (1979) (Section 4) for further developments and applications. Janardan (1960b) discusses a "discrete exponential" family:

\[ P_n = h(n) \exp \left( \lambda X \right) \lambda^{b(0)} \]

with \( h(0), c(0) \) positive and differentiable, and \( c(0), C(0) \) strictly increasing; \( h(x) > 0 \).

A related class (see Irwin (1975), Berg (1974, 1975a-b, 1978, 1980)) is that of factorial series distributions. The factorial series distribution with generating function \( h(0) \) and parameter \( n \) (an integer) is defined by

\[ P_n = \sum_{k=0}^{n} c(k) \]

with

\[ h(n) = \sum_{k=0}^{n} c(k) \]

(24)

(25)
so that
\[ c(n) = \delta^n h(n) |_{n=0} = \delta^n h(0). \]  
(26)

We write
\[ X \sim \text{PSD}(h,h(n)). \]

There are close similarities between properties of PSD's and FSD's. For example, the rth factorial moment is
\[ \mu_r = \frac{(\delta^r g(0))/g(0)}{\delta^r h(0)}/. \]
(27)

for the FSD (19), and
\[ \mu_r = \frac{(\delta^r g(0))/g(0)}{\delta^r h(0)}/. \]
(28)

for the FSD (24).

The distribution of the sum of a independent random variables \[ X_1, X_2, \ldots, X_n \]
each having the PSD (19) is also a PSD, with generating function \( g(0)^n \). Berg (1979) obtains an analogous result for FSD's which, however, requires the \( X \)'s to have a certain kind of dependence. He introduces a "parameter translated" modification of PSD's, defined by
\[ Pr[X_n = (n-1)(n) \delta^n h(n-n)/h(n) (n=0,1,\ldots,N-a) \]
(29)

written
\[ X \sim \text{PSD}(h(n);h(n)) \]

where a is a positive integer. Then if
\[ X_1 \sim \text{PSD}(h(n)); X_2 \sim \text{PSD}(h(n-a)); \ldots \]

and generally
\[ X_{j+1} \sim \text{PSD}(h(n-a)); \ldots, \]

it follows that
\[ X_1 \sim \text{PSD}(h(n);h(n))^n. \]

Berg (1979) also studies compound distributions of form
\[ \text{PSD}(h(n);h(n))^n \]
and (Berg (1976, 1980)) applies them in a generalized form of "snowball sampling" (Goodman (1961)).

Multivariate power series distributions (Patil (1965), J. K. page 33) are defined by
\[ P_X = \prod_{i=1}^{\infty} \frac{\sum_{k=0}^{\infty} b(k)}{b(k)} g(k) \]
(31)

with
\[ b(k) = \prod_{i=1}^{\infty} \frac{d_i(k)}{b(k)} g(k) \]
(32)

where \( d_i \) denotes differentiation with respect to \( \theta_i \). We write
\[ X \sim \text{MPS}(g(0), \theta). \]

Joshi and Patil (1971, 1974) have studied a special class of \( h \) FSD, introduced by Patil (1968), called same-symmetric ("SSS", PSD), for which \( g(0) \) is a function only of \( \theta = \sum_{i=1}^{\infty} d_i \).

Multivariate factorial series distributions are defined by Berg (1977) analogously to \( h \) PSD, with
\[ P_X = \prod_{i=1}^{\infty} \frac{\sum_{k=0}^{\infty} c(k)}{c(k)} g(k) \]
(33)

where
\[ c(k) = \prod_{i=1}^{\infty} d_i(k) g(k) \]
(34).
can be used to construct a broad class of discrete distributions with

\[ P_n = \binom{h}{n} \left( \frac{x^{h-1}}{x^n} \right) \left( \frac{1}{1-x} \right)^n \]

(37)

provided no \( P_n \)'s are negative and the series converges (Kemp, 1968). Note that \( a[0] = 1 \) so that

\[ P_0 = \binom{h}{n} (x^0) \phi \]

The pgf is

\[ n^F_r (x; b; \theta) / n^F_r (x; \theta) \]

We may write

\[ X \sim \text{Hyper}(x; b; \theta) \]

(40)

These generalized hypergeometric (series) distributions form a very extensive class and have some specially attractive features (such as simple recurrence relationships for probabilities, factorial moments, etc.). Decay (1972, Table 2) gives a list of some 50 pgf's expressed in terms of \( F_r \) functions, with \( p, q \leq 3 \).

The formula for the ratio of successive probabilities

\[ \frac{P_{n+1}}{P_n} = \frac{n+1}{x+1} \cdot \frac{\binom{n+1}{x+1}}{n+1} \]

(41)

may be regarded as a generalization of the formula

\[ \frac{P_{n+1}}{P_n} = \frac{n+1}{x+1} \]

(42)

used by Katz (1965) to generate a family of distributions. (In the symbol of (40) above, (42) corresponds to

\[ X \sim \text{Hyper}(x; b; \theta) \]

(38)
The subclass \( \mathcal{G} \) \( \mathcal{G}(a_1, a_2; b_1; 0) \) has been used much more than any other numbers of the class. In Duncan's (1972) list, about 30, among 50 entries use \( \mathcal{G} \). Kemp and Kemp (1966, 1969, 1971) discuss a number of interesting special cases, including a "lost-games" distribution, that of the number of games lost by a player starting with \( a \) units of money before losing all his money. When in each game the probability of losing a unit is \( p \) (and of gaining a unit, \( q \)), and \( (1 + p) \) distributions having applications in epidemic theory and queueing theory.

They also show (1971) that these distributions can be obtained in a variety of ways as mixtures of negative binomial or Poisson distributions.

More recently Kemp and Kemp (1974) introduced a class of distributions with \( \text{pgf} F_Z(z; a, b; \lambda) \) which they called generalized hypergeometric factorial-moment distributions. Besides some well-known distributions like Poisson, negative binomial and Pólya, this class contains some matching and occupancy distributions (see Johnson and Kotz (1977) and Kemp and Kemp (1978)) and other compound distributions. Again, recurrence relationships for probabilities, moments, etc., are easily obtainable.

Gurland and Tripathi (1974, 1977) consider the more general class - termed by them "extended Katz" (EK) for which

\[
P_{X} = \frac{e^{z \lambda}}{\sum_{n=0}^{\infty} \frac{\lambda^n}{n!}} \quad (z > 0; \lambda > 0; a > 0; b > 0)
\]

which has \( n = 2, k = 1, a_0 = a, b_1 = \lambda, b_0 = b \). The factorial moments satisfy the recurrence relation

\[
(1-\lambda)u_{p+2} + \left( (\lambda + r - 1) - r \right)u_{p+1} + (r+1)(\lambda+\epsilon^2)u_{p} = 0.
\]

Also

\[
E[X] = \frac{\lambda}{\lambda - n(1-p)}.
\]

The class is extended to negative values of \( \theta \) by using the function

\[
F_{Z}(z; a, b; \lambda) = \sum_{n=0}^{\infty} \frac{a^n}{n!} \cdot \frac{\lambda^n}{(\lambda - n)^n}
\]

with \( n \) equal to the integer part of \(-a/b\). As \( b > 0 \) (either through positive or negative values) the class (CB) described by Crow and Hardwell (1963) - and termed hyper-Poisson by these authors - is approached. The pgf for this class is

\[
\frac{F_{Z}(z; a, b; \lambda)}{F_{1}(1; \lambda; a)} = \frac{1}{F_{Z}(z; a, b; \lambda)}.
\]

Gurland and Tripathi (1974) extend the CB class to

(1) \( E_1 \) CB, with \( p_{x+1}/p_x = \gamma x (1+x) \)

and

\[
\text{pgf} F_{Z}(z; a, b; \lambda) = \frac{F_{Z}(z; a, b; \lambda)}{F_{1}(1; \lambda; a)}
\]

(with either \( \alpha, \gamma, \lambda \) > 0 or \( \alpha > 0 \) and \( \gamma, \lambda < 0 \) with \( \gamma \) and \( \lambda \) having the same integer parts) and

(2) \( E_2 \) CB, with pgf

\[
\frac{F_{Z}(z; a, b; \lambda)}{F_{1}(1; \lambda; a)} = \frac{F_{Z}(z; a, b; \lambda)}{F_{2}(1; \lambda; a)}.
\]

This class (ii) is not included in the general form (27). It includes

(1) as a special case with \( a_0 = 0 \). Conditions on the parameters are

\[
a_0, a_1, \alpha, \gamma, \lambda > 0,
\]

\[
f \alpha > 0, \lambda > 0
\]

\[
\gamma < 0, \lambda < 0, \gamma \text{ and } \lambda \text{ having the same integer parts, } a_1 > 0,
\]

\[
a_0a_1 \neq 0.
\]

Ord (1972), following K. Pearson's (1805) original approach, obtains a system of hypergeometric (series) distributions as solutions of the difference equation

\[
\frac{x+1}{x} - \frac{(a+1)x}{b_0x+b_0x^2} = 0,
\]

Ord (1972, Table 5.1) gives a summary of the members of this system. Note that for all cases except Type IX, \( b_0 \neq 0 \). As a way of distinguishing among those different members, for which \( b_0 = 0 \), Ord notes that (51) can then be rewritten

\[
xP_{Z}(z; a, b; \lambda) = \frac{a+1}{b_1b_2(x-1)}
\]

If \( b_2 = 0 \), \( xP_{Z}(z; a, b; \lambda) \) is a linear function of \( x \).
It is suggested that \( u_{k} = x_{k}/f_{k-1} \) be plotted against \( x \), and if the plot appears to be linear, the following table (based on Table 5.4 of Ord (1972)) may be used to choose the appropriate distribution.

<table>
<thead>
<tr>
<th>Intercept (&quot;( u_{0} ))</th>
<th>Slope</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi &gt; 0 )</td>
<td>( \phi )</td>
<td>Poisson (( \phi ))</td>
</tr>
<tr>
<td>(( n-1 ))p/q &gt; 0</td>
<td>-p/q &lt; 0</td>
<td>Binomial (( n, p ))</td>
</tr>
<tr>
<td>(( n-1 ))p/q &gt; 0</td>
<td>p/q &gt; 0</td>
<td>Negative binomial (N, ( p ))</td>
</tr>
<tr>
<td>( \phi &lt; 0 )</td>
<td>( \phi &gt; 0 )</td>
<td>Log series (( \phi )) (note: intercept = -(slope))</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>Discrete rectangular</td>
</tr>
</tbody>
</table>

When \( \phi = 1 \) we have generalized hypergeometric distributions ("series" is dropped).

A considerable part of the increased attention devoted to multivariate discrete distributions in recent years has been directed to generalizations of multivariate hypergeometric distributions. (Note that "generalized" is not used in the limited "random sum" sense.)

Although this work has been done fairly recently, most of the distributions discussed are included in a very broad class described by Steyn (1951) some 50 years ago. For this class

\[
P_{x} = \frac{(c - \sum_{j=1}^{b_{j}} x_{j})^{[a]}}{c^{[a]}} \cdot \prod_{j=1}^{b_{j}} \frac{x_{j}^{[a]} b_{j}}{j!^{[a]}}
\]

(53)

the values of \( a, b_{j} \) and \( c \) being such that \( P_{x} \geq 0 \).

Janardan and Patil (1972) show that this class includes many established distributions, among them the multivariate hypergeometric (JE, page 300), inverse hypergeometric, negative hypergeometric, negative inverse hypergeometric, Polya, and inverse Polya distributions. They also summarize some possible generalizations of these distributions.

10. Digamma and Trigamma Distributions.

Sibuya (1979a), by taking limiting cases of the zero-truncated generalized hypergeometric (Inverse Polya-Eggenberger) \((\text{Negative Binomial } (a, p) \times \text{Beta } (b, y))\) distribution

\[
P_{x} = \frac{\sum_{t=0}^{\infty} x^{t}[t]}{\gamma (a+b+y)} - \frac{\sum_{t=0}^{\infty} x^{t}[t]}{\gamma (a+y)} - 1 (x = 1, 2, \ldots) \quad (54)
\]

(interpreting \( a^{(b)} \) as \( t(a+b)/t(a) \)) constructed two new distributions as limiting cases.

(i) As \( b = 0 \) (with \( a \neq 0 \))

\[
P_{x} = (\phi(a+y) - \psi(y))^{-1} \frac{a^{[x]}}{(a+y)^{[x]}}, \quad (x = 1, 2, \ldots; y > 0; \quad a > 1; \quad a+y > 0)
\]

(55)

where \( \psi(y) = d \log \Gamma(y)/dy \) is the digamma function. This is termed a digamma distribution.

(ii) As \( a = 0 \) and \( b = 0 \)

\[
P_{x} = \frac{1}{\psi^{(1)}(\gamma)} \cdot \frac{x[x]}{\gamma[x]}, \quad (x = 1, 2, \ldots; y > 0).
\]

(56)

The function \( \psi^{(1)}(y) = d^{2} \log \Gamma(y)/dy^{2} \) is the trigamma function, and (56) is termed a trigamma distribution.

Clearly the trigamma distributions are limit distributions (as \( a = 0 \)) of the digamma distributions. If \( a \) and \( y \) tend to infinity with \( a(a+y)^{-1} \rightarrow 0 \) the distribution tends to a logseries distribution with parameter \( \phi \).
The trigamma distribution (56) is very similar to the beta distribution (JE page 240). In fact, when \( r = 1 \), the trigamma distribution is a beta distribution with parameter 1. Silvey (1970a) suggests that the digamma distribution may be used in place of logseries distributions when the tails of the latter are not long enough to fit the data. Multivariate digamma distributions can be derived from a truncated multivariate inverse Polya-Eggenberger, but when a multivariate trigamma is sought, it turns out to be degenerate.

11. Lagrange Distributions.

Exploitation of the Lagrange series expansion

\[
f(s) = f(0) + \sum_{j=1}^{\infty} \frac{u^j}{j!} \left( \frac{\delta (s)}{s} \right)^j f(t) t^0 \tag{57}
\]

where \( u = s/g(s) \) has produced some useful distributions. There is a good general introduction in Consul and Skenton (1972b), and useful information on construction of Lagrange distributions in Jain (1974a).

The functions \( f(\cdot) \) and \( g(\cdot) \) are each taken to be pgf's. Since \( u = 1 \) when \( u = 1 \), \( s = s(u) \) can also be a pgf, and will be so if the coefficients of powers of \( u \) in the expansion of \( s(u) \) are all positive. Then (57) will represent the pgf with argument \( u \) of a “generalization” (in the sense of Durand - see Chapter 8 of JK) of the distribution with pgf \( f(\cdot) \) by the distribution with pgf \( g(\cdot) \). For these distributions

\[
\begin{align*}
P_n &= f(0) \\
P_s &= u^{n-1} (\delta_n) g(\cdot) f(\cdot) 
\end{align*}
\tag{58}
\]

may form the class of Lagrange distributions. We write

\[
\mathcal{L}(g(\cdot); f(\cdot)) .
\tag{59}
\]

Specific subclasses are obtained by choosing specific forms for \( g(\cdot) \) and \( f(\cdot) \) termed, respectively, the transformation and generating pgf's for \( L(g(\cdot); f(\cdot)) \).

For example, the Poisson-Poisson (Lagrange “double Poisson” distribution, or generalized Borel-Tanner distribution (JK, page 254)) is obtained by taking

\[
f(s) = \exp(\lambda_1 (s-1)); \ g(s) = \exp(\lambda_2 (s-1)) - \text{the pgf's of Poisson distributions with parameters } \lambda_1, \lambda_2 \text{ respectively. In order to satisfy the condition on } s(u) \text{ (see above) we must have } \lambda_2 < 1. \text{ For this distribution}
\]

\[
P_n = \lambda_1 (\lambda_2 + \lambda_2 + \lambda_2) s^{n-1} e^{-\lambda_1 s} / s ! \quad (n = 0, 1, \ldots)
\tag{60}
\]

(Consul and Jain (1970)).

Table 6.1 of Consul and Skenton (1972) lists thirteen Lagrange distributions obtained by combining binomial, Poisson, negative binomial and “delta” pgf’s (the latter has \( P_n = 1 \)). The same authors also describe Lagrange Poisson-rectangular and Poisson-logseries distributions.

Consul and Skenton (1972) show that

\[
\begin{align*}
\mathbb{E}(x) &\left( g(\cdot); f(\cdot) \right) = f_1 (1 - g_1)^{-1} \\
\text{var}(x) &\left( g(\cdot); f(\cdot) \right) = f_2 (1 - g_2)^{-2} + f_2 g_2 (1 - g_1)^{-3}
\end{align*}
\tag{61}
\]

(where \( f_1, f_2, g_1, g_2 \) are the first and second cumulants of the distributions with pgf's \( f(s) \), \( g(s) \) respectively and also give formulas for \( u_2 (x) \) and \( v_2 (x) \). If \( g(\cdot) \) is such that \( |g_1| < 1 \), then all moments of \( L(g(\cdot); f(\cdot)) \) exist.

Consul and Skenton (1974) give a useful summary of properties of Lagrange distributions. They also identify several Lagrange distributions arising in queueing theory.

Consul and Skenton (1973) have studied certain limit cases of Lagrange distributions. They show that

\begin{enumerate}
\item if \( g_1 \) is kept fixed and \( f_1 \to \) the standardized \( L(g(\cdot); f(\cdot)) \) distribution tends to normality, but
\end{enumerate}
(ii) If \( f_1 = \cdots = f_n = 1 \) in such a way that \( f_1(1-x) x^{-1} = c^2 \), then
the standardized distribution tends to the inverse Gaussian distribution with pdf
\[
\frac{c^{1/2}}{2\pi(y-c)^{3/2}} \exp\left(-\frac{1}{2} cy^{-1}(y-c)^{-1}\right). \tag{62}
\]

Consul and Shenton (1973a), utilizing a multivariate extension of Lagrange's expansion due to Good (1959), described a general approach to construction of multivariate distributions with (univariate) Lagrange marginals. They showed that such distributions appeared naturally in queuing theory, gave general formulas for variances and covariances and constructed a multivariate Nessel-Tanner ("double Poisson") distribution with
\[
P_{\lambda} = \sum_{j=1}^{\infty} \frac{\left(\frac{\lambda}{j!}\right)^j}{(j^\lambda j!)^2} \exp\left(-\frac{\lambda}{j^\lambda}\right) \left[1 - \lambda (x_j - 1)^{-1}\right] \tag{63}
\]
where the \((p,h)\)-th element of the \(\lambda\) matrix \(\lambda(x)\) is \(\frac{\lambda}{j^\lambda}\) and the \(\lambda\)'s are positive constants for \(x_j > 0\). A queuing theory interpretation is given.

The same authors (Shenton and Consul (1973)) consider the bivariate case in more detail.

Jain (1974b) extended his treatment of univariate Lagrange power series distributions to multivariate distributions. In particular Jain and Singh (1975) obtained the Lagrange bivariate negative binomial
\[
P_{\lambda} = \sum_{j=1}^{\infty} \frac{\left(\frac{\lambda}{j!}\right)^j}{(j^\lambda j!)^2} \exp\left(-\frac{\lambda}{j^\lambda}\right) \left[1 - \lambda (x_j - 1)^{-1}\right] \tag{63}
\]
where \(0 < \lambda, \lambda_1, \lambda_2 < \infty\) and \(\lambda_1, \lambda_2 > 1\).

They also obtain a bivariate Nessel-Tanner distribution
\[
P_{\lambda} = \frac{\lambda^{1/2}}{2\pi(y-c)^{3/2}} \exp\left(-\frac{1}{2} cy^{-1}(y-c)^{-1}\right) \left[1 - \lambda (x_j - 1)^{-1}\right] \tag{63}
\]
where \(\lambda, \lambda_1, \lambda_2, \gamma_1, \gamma_2 > 0\), which differs from (63).

12. Discrete (Linear) Exponential Distributions.
These are defined by
\[
P_{\lambda} = \sum_{j=1}^{\infty} \frac{\lambda^{1/2}}{2\pi(y-c)^{3/2}} \exp\left(-\frac{1}{2} cy^{-1}(y-c)^{-1}\right) \left[1 - \lambda (x_j - 1)^{-1}\right] \tag{63}
\]
where \(\lambda, \lambda_1, \lambda_2, \gamma_1, \gamma_2 > 0\), which differs from (63).

Consul and Nettel (1975) contain a thorough discussion of this family.

Estimation of the \(\gamma_j\)'s from \(x_1, x_2, \ldots, x_n\) when the \(x_i\)'s are independent and
\[
Pr[x_j = x] = \frac{g^{(j,x)}(x)}{g^{(j)}(x)} \tag{67}
\]
is discussed by Tsui (1979b) who obtains results on admissibility (with squared error loss function) similar to those of Stein for multinormal distributions.

Methods of shrinking the MVUE's to improve estimation are also discussed.

Recent papers by Lauritzen (1975) and Janardan (1980a) discuss some further generalizations.

13. 'Abel' Distributions.

Consul (1974a) and Consul and Nettel (1975) have derived a number of distributions which are related to Abel's generalizations (e.g., Riordan (1976)) of the binomial expansion (see also Janardan (1974)). These include

Quasi-binomial distribution I,
\[
P_{\lambda} = \frac{\lambda^{1/2}}{2\pi(y-c)^{3/2}} \exp\left(-\frac{1}{2} cy^{-1}(y-c)^{-1}\right) \left[1 - \lambda (x_j - 1)^{-1}\right] \tag{63}
\]
where \(\lambda, \lambda_1, \lambda_2, \gamma_1, \gamma_2 > 0\), which differs from (63).
Quasi-binomial distribution II.
\[
P_n = \frac{a^x b^{n-x}}{(a+n-1)(b+n-x-1)} \frac{(a+n-1)x}{(n-x)(n-x)} \quad (x=0,1,\ldots,n). \quad (68.2)
\]
Quasi-hypergeometric distribution.
\[
P_x = \frac{a^x b^{n-x}}{(a+b)(a+n-x)(b+n-x)} \frac{x}{n-x} \quad (x=1,\ldots,n). \quad (69)
\]
Quasi-Polya distribution.
\[
P_n = \frac{a^n b^n}{(a+b)^n} (n-1) \frac{(a+n-x)(b+x)}{(a+b)} \quad (n=0,1,\ldots,n). \quad (70)
\]
In all formulae (68)-(70) the parameters \( a, b, x \) can be any positive numbers.

Proper distributions are also obtained with some other sets of values of the parameters.

Multivariate extensions ("quasi-multinomial" distributions), based on Harviti's extensions of Abel's identities are described in Consul and Mittal (1977).


The preceding text is based on a considerably more extensive (though still not exhaustive) original version, which the authors plan to publish, if opportunity arises.

15. Acknowledgments.

Dr. R. Kott's research related to this paper and bibliography were supported in part (until July 1, 1980) by US Air Force Grant No. 78-567; after July 1, 1980, by NSF Grant MSC-79-12-130. Revisions and amendments of this paper (summer 1981) were supported by ONR Contract NO0014-81-K-0301. Dr. R. L. Johnson's work was supported by the National Science Foundation under Grant MCS 8021704.

BIBLIOGRAPHY

The first short list contains useful collections of results. In some cases abbreviations (JX, RCSW, etc) are given which are used in the text, and in other parts of the bibliography.

The second, much longer list includes many papers not referred to explicitly in the text. As we have noted, some severe selection was forced upon us by reasons of space. In most cases, the content of papers can be inferred from the titles.

AS


NATO

ISSE

JX


RCSW

SDSW


Boedova, T. J. and Bendell, B. J. (1970) Tables of exact and approximate values of certain ratios and reciprocals of multinomial random variables useful in life testing, Technometrics, 12, 901-908.


Cacoullos, T. (1972) On the convolution of Poisson distributions with the zero class missing, Amer. Statist., 26 (2), 44.


Bibliography


Gerrard, D. J. (1972) Inverse binomial sampling as a basis for estimating negative binomial densities, Biometrics, 28, 971-980.


Mintz (1933) The movement of people at the ticket office of Shinjuku station. Keisha (Sci) No. 3.


Huang, J. T. (1966a) Improving upon standard estimators in discrete exponential families with applications to Poisson and negative binomial cases. Cornell Univ.

Huang, J. T. (1967b) Some tail upper bounds on the class of admissible estimators with applications to Poisson and negative binomial families. Cornell Univ.


Kende, B. (1978) Sufficiency and the number of level crossings by a stationary process, Biometrika, 65, 97-100.


Sarkar, B. P. and Adhikari, A. K. (1968) Distribution of most significant digit in certain functions whose arguments are random variables, Sankhya, 29, 47-56.


Shahbagh, D. N. (1979) Some refinements in distribution theory, Sankhyā, 42A.


Toms, L. K., Habina, K. and Isakov, H. (1979) Confidence limits for the number of defective items in a lot, J. Qual. Tech., 11, 1979, 190-204.


Developments in Discrete Distributions 1969-1980

Technical

Herman E. Johnson and Samuel Kets

E. R. Shreve, Office of Naval Research
Statistics and Probability Program
Arlington, Va. 22217

September 1981

75

Approved for Public Release - Distribution Unlimited

Discrete distributions; systems of distributions; mixtures; damage models; estimation procedures; multivariate distributions.

In this paper we survey and summarize developments in the theory and methodology of discrete distributions during the period 1969-1980 since publication of our book 'Distributions in Statistics - Discrete Distributions' (Wiley 1969). A comprehensive (though not exhaustive) bibliography of some 400 items is included.


