Ann. Inst. Statist. Math.
Vol. 41, No. 2, 305-321 (1989)

ON A NEW SYSTEM OF DISCRETE DISTRIBUTIONS AND CHARACTERIZATIONS OF SEVERAL DISCRETE DISTRIBUTIONS BY EQUALITY OF DISTRIBUTIONS

R. M. KORWAR

Department of Mathematics and Statistics, University of Massachusetts, Lederle GRC Towers, Amherst, MA 01003, U.S.A.

(Received April 14, 1988; revised October 31, 1988)

Abstract. This paper deals with a new system of discrete distributions. It also gives several characterizations of the Waring (and hence the Yule) distribution (and its truncated versions), the super-Poisson, the discrete uniform and other discrete distributions by using this system and other such systems existing in the literature, and linear regression. Continuous analogues of the above results are also briefly discussed.

Key words and phrases: Characterizing, linear regression, new system of distributions, Waring distribution.

1. Introduction

There are in the literature several systems of discrete distributions. They all start with a discrete distribution (call it the parent population) and generate a new one by the system. Such systems were considered, for example, by Katz (1945, 1965), Bissinger (1965) and Ord (1967*a*, 1967*b*). See also Johnson and Kotz ((1969), Chapter 2, Section 4). Rao (1965) introduced a system of weighted distributions (which are defined for starting discrete as well as continuous distributions). This paper introduces yet another system which is used, in conjunction with other systems, to obtain several characterizations of the Waring distribution (and hence the Yule), the discrete uniform and the super-Poisson. The method used here is the equality of distributions of two random variables.

The motivation for these characterizations comes mainly from the question: For what parent population (or populations) do two such systems lead to the same new distribution? It then turns out that only the Waring, Pareto and super-Poisson, among others, have this property. For further motivation and applications to income distribution see also Xekalaki (1983).

Section 2 introduces the new system and Section 3 contains the characterizations of the Waring and the super-Poisson. Section 4 deals with the characterization of the discrete uniform. In Section 5 we indicate the corresponding results for the continuous case.

2. The new system of discrete distributions

Let X be a positive integer-valued random variable (r.v.) with probability function

$$p_r = P(X = r), \quad r = 1, 2, \ldots . \tag{2.1}$$

Let Y be a new positive integral-valued r.v. with

$$q_y = P(Y = y) = \sum_{x=1}^{y} x p_x / y(y+1), \quad y = 1, 2, \ldots . \tag{2.2}$$

The new system (2.2) arises in at least two different ways. In one approach, suppose U is a uniform r.v. on (0, 1) independently distributed of X. Then

$$[X/U] \text{ has the same distribution as } Y, \tag{2.3}$$

where $[x]$ denotes the greatest integer in x. To see (2.3) we note that

$$\begin{aligned} P([X/U] = y) &= \sum_{x=1}^{y} P([X/U] = y, X = x) \\ &= \sum_{x=1}^{y} P\left(y \leq \frac{X}{U} < y+1, X = x \right) \\ &= \sum_{x=1}^{y} P\left(\frac{x}{y+1} < U \leq \frac{x}{y}, X = x \right) \\ &= \sum_{x=1}^{y} x p_x / y(y+1) . \end{aligned}$$

We can view $[X/U]$ as, for example, an over-report of a true insurance claim X.

Another approach to (2.2) is to consider Y as being increased from X by some creative process. For example, X particles of some kind give rise to Y particles by a process of splitting. In this approach, let the conditional distribution of Y given $X = x$ be the distribution on $x, x+1, \ldots$ given by

$$P(Y = y \mid X = x) = s(y \mid x) = x / y(y+1), \quad y = x, x+1, \ldots . \tag{2.4}$$

Then, clearly the unconditional distribution of Y is given by (2.2). Another example of a creative process by which Y results from X is the following.

Let X denote the number of accidents involving personal injuries on a certain highway during a given period of time. Then Y could be taken to be the number of persons injured in the X accidents.

Note that, since (2.4) has an infinite mean, so does the new system (2.2).

3. Characterizations of the Waring and Yule distributions

Starting from (2.1), Rao (1965) introduced a system of weighted distributions and the associated weighted r.v. Z by

$$q'_r = P(Z = r) = (r + a)p_r/(\mu + a), \qquad r = 1, 2,\dots, \tag{3.1}$$

where we assume $\mu = E(X)$ to be finite and $a > -1$ is a constant. Distribution (3.1) arises in what is called size-biased sampling.

A second system frequently used is biased on partial sums or the tail of the distribution (2.1). Let W be a nonnegative integer-valued r.v. with

$$q''_r = P(\mathrm{W} = r) = \sum_{j=r+1}^{\infty} p_j/\mu, \qquad r = 0, 1,\dots. \tag{3.2}$$

Distribution (3.2) has been found useful. See, for example, Johnson and Kotz ((1969), p. 261) and Patil and Rao (1977).

It is natural to ask: For what distribution (2.1) is (3.1) and (3.2) essentially the same? Since Z takes values 1, 2,... and W = 0, 1,..., we need to truncate W on the left at zero. It then turns out that only the Waring distribution (introduced by Irwin (1965)) has this property. For our purpose, a discrete r.v. is said to have a Waring distribution with parameters c and λ and denoted $\mathrm{W}(c, \lambda)$ if

$$p_r = P(X = r) = (\lambda - c)c^{[r-1]}/\lambda^{[r]}, \qquad r = 1, 2,\dots, \tag{3.3}$$

where $\lambda - c > 0$, $c > 0$ and

$$c^{[r]} = c(c+1)\cdots(c+r-1), \qquad r = 1, 2,\dots, c^{[0]} = 1 .$$

A special case ($c = 1$) is the Yule distribution (see Johnson and Kotz (1969), p. 250).

We are now ready to characterize the Waring distribution.

THEOREM 3.1. *Let X be a positive integral-valued r.v. given by* (2.1) *with a finite mean μ. Define Z and* W *by* (3.1) *and* (3.2), *respectively, for* *some $a > -1$. Then* W *truncated (on the left) at* 0 *has the same distribution as Z if, and only if, X has the Waring distribution* (3.3).

PROOF. "Only if" part. Let q'_r and q''_r be given by (3.1) and (3.2), respectively, for some $a > -1$. Further let

$$q'_r = q''_r/(1 - q''_0), \qquad r = 1, 2, \dots , \tag{3.4}$$

where

$$q''_0 = \mu^{-1} .$$

Then it follows from (2.1), (3.1), (3.2) and (3.4) that

$$q''_r = x \sum_{k=r+1}^{\infty} q''_k/(k + a), \qquad r = 1, 2, \dots , \tag{3.5}$$

where

$$x = (\mu + a)/(\mu - 1) . \tag{3.6}$$

From (3.5) we obtain

$$q''_r - q''_{r+1} = xq''_{r+1}/(r + 1 + a) ,$$

yielding us the recurrent relation

$$q''_{r+1}/q''_r = (r + 1 + a)/(r + a + 1 + x), \qquad r = 1, 2, \dots . \tag{3.7}$$

Note that (3.7) holds even for $r = 0$. This follows from (3.1), (3.2), (3.4), (3.6) and $q_0 = \mu^{-1}$. From (3.7) we get

$$q''_r = (1 + a)^{[r]}q''_0/(1 + a + x)^{[r]}, \qquad r = 0, 1, \dots . \tag{3.8}$$

Equations (3.4)–(3.8) and (3.1) lead to

$$q'_r = (1 + a)^{[r]}q''_0/(1 + a + x)^{[r]}(1 - q''_0), \qquad r = 1, 2, \dots \tag{3.9}$$

and

$$p_r = x(1 + a)^{[r-1]}/(1 + a + x)^{[r]}, \qquad r = 1, 2, \dots . \tag{3.10}$$

But (3.10) is of the form (3.3) with $\lambda = 1 + a + x$ and $c = 1 + a$. Note that $x = (\mu + a)/(\mu - 1) > 1$ and $c > 0$ since we have assumed $a > -1$. This proves the "only if" part.

"If" part. Let X have the Waring distribution given by (3.3) with $\lambda - c > 1$ (which is required for X to have a finite mean μ). Then using the Waring expansion

$$(\lambda - c)^{-1} = \sum_{r=0}^{\infty} c^{[r]}/\lambda^{[r+1]},$$

which is valid for $\lambda > c > 0$, we find that

$$\sum_{j=r+1}^{\infty} p_j = (c + r - 1)p_r/(\lambda - c).$$

From this last equation and the fact that $\mu = c/(\lambda - c - 1) + 1$, we can easily verify that $q''_r = (1 - q''_0)q'_r$, $r = 1, 2, \dots$, where q'_r and q''_r are given by (3.1) and (3.2), respectively, with $a = c - 1 > -1$ and $q''_0 = \mu^{-1}$. This completes the proof of the "if" part and the proof of the theorem.

Remark 3.1. Our theorem covers the characterization of the Yule distribution. In the "only if" part the Yule distribution corresponds to the case $a = 0$ and in the "if" part to $c = 1$. If we take $a = 0$ in the "only if" part, then (3.9) becomes

$$q'_r = P(\mathrm{W} = r) = xr!/(1 + x)^{[r]}, \qquad r = 1, 2, \dots, \tag{3.11}$$

where we used (3.6) and the fact $q''_0 = \mu^{-1}$. Krishnaji (1970) calls (3.11) a Yule distribution and characterizes it (see below). In fact, it is not surprising that Krishnaji's characterization should come up in the context of our theorem. The connection to his characterization is as follows. When $a = 0$, (3.5) in our "only if" part becomes

$$(1 - q''_0)q'_r = \sum_{j=r+1}^{\infty} q'_j/j, \qquad j = 1, 2, \dots, \tag{3.12}$$

where

$$q''_0 = 1/\mu = \sum_{j=1}^{\infty} q'_j/j. \tag{3.13}$$

The right-hand side of (3.12) is $P([\mathrm{W}U] = r)$, as shown by Krishnaji (1970), where U is a uniformly distributed r.v. on $(0, 1)$ independently of W. Thus (3.12) can be interpreted as

(3.14) W and $[U\mathrm{W}]$ truncated at 0 have the same distribution.

Krishnaji (1970) has characterized W of (3.11) by the property (3.14). Note that the right-hand side of (3.12) for $r = 0, 1, \ldots$, defines the so-called Bissinger (1965) system of discrete distributions corresponding to the parent $\{q_r'\}_{r=1}^{\infty}$.

Remark 3.2. Suppose we let X take on nonnegative integral values and replace (3.1) by

$$q_r' = P(Z = r) = \omega(r)p_r/v, \qquad r = 0, 1, \ldots,$$

where $\omega(r) > 0$ is a general weight function and $v = E(\omega(X)) < \infty$. Then if Z and W have the same distribution, we get, corresponding to (3.7) that

$$q_{r+1}'/q_r' = \omega(r+1)/\{x + \omega(r+1)\}, \qquad r = 0, 1, \ldots,$$

where $x = v/\mu$. Thus, more distributions can be characterized by using a more general weight function ω. For example, if we take $\omega(r) = 1/(r+1)$, then

$$q_{r+1}'/q_r' = 1/x\{r + 2 + 1/x\}, \qquad r = 0, 1, \ldots,$$

yielding

$$q_r' = \Gamma(\lambda)\theta^r q_0'/\Gamma(\lambda + r), \qquad r = 0, 1, \ldots,$$

where $\theta = 1/x$ and $\lambda = 2 + \theta$. This is the probability mass function of a super-Poisson distribution (see Patil and Joshi (1968), p. 16). Thus we can characterize a super-Poisson by our methods: Let $\omega(r) = 1/(r+1)$. Then Z and W have the same distribution if, and only if, Z has a super-Poisson distribution with parameters θ and $\lambda = 2 + \theta$, $\theta > 0$.

Next we will characterize the Waring distribution (3.3) with $\lambda = c + 2$ by using the system of discrete distributions (2.2), (3.1) and (3.2). The following two theorems characterize, as expected in view of Theorem 3.1, the same Waring distribution.

THEOREM 3.2. *Let r.v.'s X, Y and Z have the distributions given by* (2.1), (2.2) *and* (3.1) *for some $a > -1$, respectively. Assume $E(X^2) < \infty$. Then Y and Z have the same distribution if, and only if, X has the Waring distribution* W$(\mu - 1, \mu + 1)$.

THEOREM 3.3. *Let X, Y and* W *be r.v.'s with distributions given by* (2.1), (2.2) *and* (3.2), *respectively. Then Y and* W *truncated at* 0 *have the same distribution if, and only if, X has the Waring distribution* W$(\mu - 1, \mu + 1)$.

PROOF OF THEOREM 3.2. "Only if" part. Let Y and Z have the same distribution for some $a > -1$. Then

$$\sum_{x=1}^{y} xp_x/y(y+1) = (r+a)p_r(\mu+a), \qquad r = 1, 2, \ldots .$$

From this it follows that

$$(\mu+a)p_{r+1} = (r+2)(r+1+a)p_{r+1} - r(r+a)p_r .$$

That is,

$$p_{r+1}\{(r+2)(r+1+a) - (\mu+a)\} = r(r+a)p_r, \qquad r = 1, 2, \ldots . \tag{3.15}$$

We first show that $a = \mu - 2$. Summing (3.15) over $r = 1, 2, \ldots$, we get

$$\sum_{r=2}^{\infty} r(r-1)p_r + (2+a)\sum_{r=2}^{\infty} rp_r - \mu\sum_{r=2}^{\infty} p_r = \sum_{r=1}^{\infty} r(r-1)p_r + (1+a)\sum_{r=1}^{\infty} rp_r .$$

That is,

$$\mu^{(2)} + (2+a)(\mu - p_1) - \mu(1-p_1) = \mu^{(2)} + (1+a) ,$$

where $\mu^{(2)}$ is the second factorial moment of X. If $\mu^{(2)}$ is finite, then the last equation yields $(2+a-\mu)p_1 = 0$. Since we have assumed $p_1 \neq 0$, we must have $a = \mu - 2$.

Now, with $a = \mu - 2$, (3.15) becomes

$$p_{r+1}(r+3+a) = p_r(r+a) .$$

That is,

$$p_{r+1}/p_r = (r+a)/(r+3+a), \qquad r = 1, 2, \ldots , \tag{3.16}$$

yielding

$$p_r = (1+a)^{[r-1]}p_1(3+a)/(3+a)^{[r]}, \qquad r = 1, 2, \ldots .$$

Summing both sides over $r = 1, 2, \ldots$ and using the Waring expansion (referred to earlier) we get $p_1 = 2/(3+a)$. Hence

$$p_r = 2(1+a)^{[r-1]}/(3+a)^{[r]}, \qquad r = 1, 2, \ldots , \tag{3.17}$$

which is $\mathrm{W}(1+a,\ 3+a) = \mathrm{W}(\mu-1,\ \mu+1)$ (since, $a = \mu - 2$). Note that

$\mu + 1 > \mu - 1 > 0$. This completes the proof of the "only if" part.

The "if" part follows in a straightforward fashion by the use of the Waring expansion. The details are omitted here. This completes the proof.

PROOF OF THEOREM 3.3. "Only if" part. Let Y and W have the same distribution. Then

$$\sum_{r=1}^{y} rp_r/y(y+1) = \sum_{r=y+1}^{\infty} p_r/\mu(1-q_0''), \qquad y = 1, 2, \ldots ,$$

where $q_0'' = 1/\mu$. From this, one easily obtains the recurrent relation

$$p_{y+1}/p_y = (y + \mu - 2)/(y + \mu - 2 + 3), \qquad y = 1, 2, \ldots ,$$

which is (3.16) since $a = \mu - 2$. Hence p_r is given by (3.17). That is, X is Waring $(\mu - 1, \mu + 1)$. This completes the proof of the "only if" part.

The "if" part follows in a straightforward fashion by the use of the Waring expansion. The details are omitted here. This completes the proof.

4. A characterization of the discrete uniform

In introducing the new system of discrete distributions in Section 2, we assumed that the parent population X took all the positive integral values. If, however, X is a finite r.v. taking the values $1, \ldots, m$ for some positive integer m

$$p_r = P(X = r), \qquad r = 1, \ldots, m , \tag{4.1}$$

then the new r.v. Y has the distribution

$$q_y = P(Y = y) = \begin{cases} \sum_{r=1}^{y} rp_r/y(y+1), & y = 1, \ldots, m , \\ \sum_{r=1}^{m} rp_r/y(y+1), & y = m+1, m+2, \ldots . \end{cases} \tag{4.2}$$

We can write (4.2) as

$$q_r = p\left\{ \sum_{x=1}^{y} xp_x/py(y+1) \right\} I(y) \\ + (1-p)\{(m+1)/y(y+1)\}\{1 - I(y)\} , \tag{4.3}$$

where

$$1-p=\sum_{r=1}^{m} rp_r/(m+1)\,, \tag{4.4}$$

and $I(y)=1$ if $y=1,\ldots,m$ and 0 otherwise. Thus Y has a finite mixture distribution of (2.2) truncated on the right at m and the Yule distribution

$$f(y)=1/y(y+1), \qquad y=1,2,\ldots\,,$$

truncated on the left at m. This representation makes it clear that: (1) Y does not have finite positive order moments, and (2) we can characterize a finite discrete distribution by using only the first distribution in the mixture (4.3). The system (4.2) arises in the same two ways as (2.2).

Let Y^* be the r.v. Y truncated on the right at m. Then the r-th ascending factorial moment $\mu^*_{[r]}=E(Y^{*[r]})$ of the r.v. Y^* can be expressed in terms of $\mu=E(X)$ and the r-th ascending factorial moment $\mu_{[r]}$ of X as follows:

$$\mu^*_{[r]}=\{\mu(m+2)^{[r-1]}-\mu_{[r]}\}/(r-1)p, \qquad r\geq 2\,,$$

where p is given by (4.4). No such simple relation exists for $\mu^*_{[1]}=E(Y^*)$, the mean of Y^*.

Turning to the characterizations referred to in (2) above, we state the following counterparts of Theorems 3.4 and 3.5 characterizing the truncated $\mathrm{W}(\mu-1,\mu+1)$ on the right at m.

THEOREM 4.1. *Let X, Y and Z be r.v.'s with distributions* (4.1), (4.2) *and* (3.1) *for some $a>-1$, respectively. Then Y truncated on the right at m and Z have the same distribution if, and only if, X has the Waring $(\mu-1, \mu+1)$ distribution truncated on the right at m.*

THEOREM 4.2. *Let X, Y and* W *be r.v.'s with distributions* (4.1), (4.2) *and* (3.2), *respectively. Then Y truncated on the right at m and* W *truncated on the left at* 0 *have the same distribution if, and only if, X has the* $\mathrm{W}(\mu-1,\mu+1)$ *distribution truncated on the right at m.*

A different kind and new characterization of the discrete uniform is obtained below by requiring that either X and Y truncated on the right at m have the same distribution, or X has a linear regression on Y with slope $2/3$ and intercept $1/3$.

THEOREM 4.3. *Let X and Y be r.v.'s with distributions* (4.1) *and* (4.2), *respectively, for $m>1$. Then X and Y truncated on the right at m have the same distribution if, and only if, X has the uniform distribution*

(4.5) $$p_r = P(X = r) = 1/m, \qquad r = 1,\ldots,m\ .$$

PROOF. "Only if" part. Let X and Y truncated on the right at m have the same distribution. Then

(4.6) $$\sum_{r=1}^{y} rp_r/py(y+1) = p_y, \qquad y = 1,\ldots,m\ .$$

Setting $y = 1$ in (4.6), we obtain $p = 1/2$. This and (4.6) give

$$2(y+1)p_{y+1} = (y+1)(y+2)p_{y+1} - y(y+1)p_y\ .$$

That is,

$$p_{y+1} = p_y, \qquad y = 1,\ldots,m-1\ ,$$

giving us (4.5). This completes the proof of the "only if" part.

The proof of the "if" part is straightforward and will be omitted here. This completes the proof.

We next characterize the discrete uniform (and other distributions) by linear regression of X on Y. But first we deal with some preliminaries.

Let, then, X and Y be given by (4.1) and (4.2), respectively, where m now is a positive integer greater than 1 or ∞. The joint distribution of X and $Y = [X/U]$, where as before U is uniformly distributed on $(0,1)$ independently of X, is given by

(4.7) $$\begin{aligned} P(X = x,\ Y = y) &= P(X = x, [X/U] = y) \\ &= P(X = x,\ x/(y+1) < U \leq x/y) \\ &= \begin{cases} xp_x/y(y+1), & y \geq x\ , \\ 0, & y < x\ . \end{cases} \end{aligned}$$

Hence, the conditional distribution of X given $Y = y$ is given by

(4.8) $$P(X = x \mid Y = y) = \begin{cases} xp_x\Big/ \sum_{x=1}^{y} xp_x, & x \leq y \leq m\ , \\ xp_x\Big/ \sum_{x=1}^{m} xp_x, & x < m < y\ . \end{cases}$$

Assume $E(X)$ is finite. It then follows that

$$E(X|Y=y) = \begin{cases} \sum_{x=1}^{y} x^2 p_x \Big/ \sum_{x=1}^{y} x p_x, & y \le m\,, \\ \sum_{x=1}^{m} x^2 p_x \Big/ \sum_{x=1}^{m} x p_x, & y > m\,. \end{cases} \tag{4.9}$$

Note (from (4.8) and(4.9)) that: (a) the conditional distribution of X given Y is a weighted distribution; and (b) the regression of X on Y is constant for $Y \ge m$.

Lemma 4.1 shows that if the regression of X on Y is linear, then m must be a positive integer (and not ∞).

LEMMA 4.1. *Let X and $Y = [X/U]$ be given by* (4.1) *and* (4.2) *where m is a positive integer or ∞. Let $E(X)$ be finite and $E(X|Y)$ be linear. Then m must be a positive integer and not ∞.*

PROOF. Let

$$E(X|Y=y) = \alpha + \beta y\,, \tag{4.10}$$

for some constants α and β. Then from (4.9) with $y = 1$ and (4.10), we get

$$\alpha + \beta = 1\,.$$

It can be shown that for $y \le m$ the right-hand side of (4.9) is an increasing function of y. Hence $\beta > 0$. From (4.10) and $\alpha + \beta = 1$, it also follows that $\beta \le (y-1)/y$. Thus we have

$$0 < \alpha < 1, \quad 0 < \beta < 1 \quad \text{and} \quad \alpha + \beta = 1\,.$$

Now equating the right-hand sides of (4.9) and (4.10) we have, for $y \le m$,

$$\sum_{x=1}^{y} x^2 p_x = (\alpha + \beta y) \sum_{x=1}^{y} x p_x\,. \tag{4.11}$$

From $\alpha + \beta = 1$ and this last equation, we get

$$p_y / p_{y-1} = \{(y-2) + (1-\alpha)/\alpha\}/y, \quad y = 2, \ldots, m\,, \tag{4.12}$$

yielding, finally,

$$p_y = p_1 \{(1-\alpha)/\alpha\}^{[y-1]}/y!, \quad y = 1, \ldots, m\,. \tag{4.13}$$

Now, suppose $(1-\alpha)/\alpha > 1$. Then (4.13) yields $p_y > p_1/y$, $y = 1, \ldots, m$, and $\sum_{y=1}^{m} p_y$ will diverge if $m = \infty$. Thus, in this case, m must be a positive integer. Next, suppose $(1-\alpha)/\alpha \leq 1$. Then, if $m=\infty$, (4.13) is the Waring distribution (3.3) with $\lambda = 1$ and $c=(1-\alpha)/\alpha \leq 1$. Such a Waring distribution has an infinite mean (see, for example, Patil and Joshi (1968), p. 50). This contradicts the assumption that $E(X) < \infty$, completing the proof.

Theorem 4.4 below characterizes distributions (4.13) (with m a positive integer) which include the discrete uniform, by linear regression of X on Y.

THEOREM 4.4. *Let X and Y have the distributions* (4.1) ($m > 1$) *and* (4.2), *respectively. Let $E(X)$ be finite. Then the regression of X on Y*

$$E(X|\,Y=y)=\begin{cases}\alpha+\beta y, & y\leq m\,,\\ const., & y>m\,,\end{cases} \tag{4.14}$$

for some constants α and β if, and only if, X has the distribution (4.13). *In particular,* (4.14) *holds with $\alpha = 1/3$ and $\beta = 2/3$ if, and only if, X has the discrete uniform distribution* (4.5).

PROOF. "Only if" part. Let (4.14) hold for some constants α and β. Then from the proof of Lemma 4.1, we get (4.13), where m now is a positive integer.

To prove the "if" part we assume (4.13) and make use of the identity

$$\sum_{x=1}^{y} a^{[x]}/x! = (a+1)^{[y]}/y! - 1, \qquad a\neq 0\,, \tag{4.15}$$

repeatedly to obtain

$$\sum_{x=1}^{y} xa^{[x-1]}/x! = (a+1)^{[y-1]}/(y-1)!, \qquad y=1,2,\ldots \tag{4.16}$$

and

$$\sum_{x=1}^{y} x^2a^{[x-1]}/x! = (ay+1)(a+1)^{[y-1]}/(a+1)(y-1)!, \quad y=1,2,\ldots\,. \tag{4.17}$$

Finally, (4.9), (4.16) and (4.17) (with $a=(1-\alpha)/\alpha$) give us (4.14), completing the proof of the "if" part.

The "only if" part of the second assertion of the theorem is obvious from (4.13). The "if" part follows from (4.16) and (4.17) with $a=(1-\alpha)/\alpha = 2$. This completes the proof.

The following theorem shows that if a positive r.v. T has a linear regression on X, then it will continue to have linear regression on $Y=$

$[X/U]$ if X has the distribution (4.13) provided U is independently distributed of (X, T).

THEOREM 4.5. *Let X, Y have distributions given by* (4.1) *and* (4.2), *respectively, for a positive integer $m > 1$. Assume that the r.v. T has a finite mean and has a linear regression on X*

$$E(T|X = x) = a + bx, \quad b \neq 0\,, \tag{4.18}$$

for some constants a and b. Further assume that $E(X) < \infty$ and U is independent of (X, T). Then, T has a linear regression on Y

$$E(T|\,Y = y) = \begin{cases} \gamma + \delta y, & y \leq m\,, \\ const., & y > m\,, \end{cases} \tag{4.19}$$

for some constants γ and δ if, and only if, X has the distribution (4.13). *In particular,* (4.19) *holds with $\gamma = a + b/3$ and $\delta = 2b/3$ if, and only if, X has the discrete uniform distribution* (4.5).

PROOF. "Only if" part. Suppose (4.19) holds for some constants γ and δ. Then

$$\begin{aligned} E(T|\,Y) &= E(E(T|X, U)|\,Y), \quad \text{a.s.} \\ &= E(a + bx|\,Y) \\ &= a + bE(X|\,Y)\,, \end{aligned} \tag{4.20}$$

the middle equality following from (4.18). Hence, from (4.10) and (4.19), we have

$$\gamma + \delta y = a + b \sum_{x=1}^{y} x^2 p_x \Big/ \sum_{x=1}^{y} x^2 p_x, \quad y \leq m\,. \tag{4.21}$$

Now, (4.21) is (4.11) of Lemma 4.1 with $\alpha = (\gamma - a)/b$ and $\beta = \delta/b$. Thus, from Lemma 4.1, (4.13) holds with $(1-\alpha)/\alpha = (b - \gamma + a)/(\gamma - a) = \delta/(b - \delta)$ (the last equality following from (4.21) with $y = 1$).

The "if" part follows from the "if" part of Theorem 4.5 and (4.20).

The second assertion follows from (4.13) (with $(1-\alpha)/\alpha = \delta(b-\delta) = 2$) and (4.20). This completes the proof.

Remark 4.1. In Theorem 4.5 we assumed that $b \neq 0$. However, if $b = 0$, then (4.19) will hold with $\gamma = a$ and $\delta = 0$ for any r.v. X. This follows from (4.20).

5. Continuous analogues

In this section we state, for most of the time, continuous analogues of the results obtained in earlier sections.

Let X be a positive r.v. on (a, ∞), $a \geq 0$ with density f. Then we define a new r.v. Y by the density

$$g(y) = \int_a^y xf(x)dx/y^2, \quad y > a. \tag{5.1}$$

It turns out that Y has the same density as X/U, where U is uniformly distributed on $(0, 1)$ independently of X. As before, Y has no positive order moments. This is clear from the representation that the conditional density of $Y = X/U$ given $X = x$ is the Pareto distribution

$$h(y|x) = x/y^2, \quad y \geq x, \tag{5.2}$$

and obviously, (5.2) has an infinite mean.

If, however, we restrict X to a finite range say, (a, b), $0 \leq a < b < \infty$, then taking $Y = X/U$, the density of Y will be

$$g(y) = p\left\{\int_a^y xf(x)dx/py^2\right\} I(y) + (1-p)\{b/y^2\}(1 - I(y)), \tag{5.3}$$

where $I(y)$ is the indicator of (a, b) and

$$1 - p = \int_a^b xf(x)dx/b, \tag{5.4}$$

we can, then compute the r-th moment of Y^*, the truncated version of Y on the right at b, as follows

$$E(Y^{*r}) = \begin{cases} \mu_r^* = (b^{r-1}\mu - \mu_r)/p(r-1), & r \geq 2, \\ \left(\int_a^b xf(x) \log x dx - \mu \log b\right)\Big/ p, & r = 1, \end{cases} \tag{5.5}$$

where $E(X^r) = \mu_r$.

Before turning to characterizations, let us note the analogues of (3.1) and (3.2) for the continuous case: A r.v. Z is said to have the weighted distribution corresponding to X with a weight function $\omega(x) > 0$ if its density is given by

$$k(z) = \omega(z)f(z)/E(\omega(X)), \quad z > a, \tag{5.6}$$

where $E(\omega(X)) < \infty$. Similarly let W be the r.v. with density

$$(5.7) \qquad l(w) = \{1 - F(w)\}/\mu, \qquad w > 0\,,$$

where $E(X) = \mu < \infty$. See Patil and Rao (1977) for an example in which (5.7) arises naturally. We are now ready to state characterizations of some well-known continuous distributions.

THEOREM 5.1. *Let X be a positive r.v. on (a, ∞), $a \geq 0$ with the density $f(x)$. Let Z and W be given by* (5.6) *and* (5.7), *respectively, for some weight function ω. Assume $\int_a^\infty dv/\omega(v) = \infty$. Then Z and W have the same distribution if, and only if, X is given by*

$$(5.8) \qquad 1 - F(x) = \exp\left(- d\int_a^x dv/\omega(v) \right),$$

where $d = E(\omega(X))/\mu$.

Thus we can characterize different distributions by taking different function ω. For example, if we take $\omega(x) = 1/x$ (and $a = 0$) we get the Rayleigh distribution, and if we take $\omega(x) = x$, we get the Pareto distribution.

The next result characterizes a Beta distribution of the second kind. A r.v. X is said to have a Beta distribution of the second kind with parameters (p, q), p, $q > 0$ if its density is given by

$$(5.9) \qquad f(x; p, q) = x^{p-1}/\{\beta(p, q)(1 + z)^{p+q}\}, \qquad x > 0\,,$$

where $\beta(p, q)$ is the beta function.

THEOREM 5.2. *Let X be a positive r.v. on $(0, \infty)$ with the density $f(x)$. Let Y and Z be given by* (5.1) *and* (5.6), *respectively, with $\omega(x) = x + c$ for some constant $c > 0$. Then Y and Z have the same distribution if, and only if, X/c has a Beta distribution of the second kind with parameters $(\mu/c, 2)$.*

The following result characterizes the Pareto distribution of the second kind (see, for a definition, Johnson and Kotz (1970), p. 234).

THEOREM 5.3. *Let X be a positive r.v. on $(0, \infty)$ with the density $f(x)$ and a finite mean μ. Let Y and W be given by* (5.1) *and* (5.7), *respectively. Then Y and W have the same distribution if, and only if, X has the Pareto distribution of the second kind with the distribution* *function F given by*

$$1 - F(x) = \{\mu/(\mu + x)\}^2, \quad x > 0 . \tag{5.10}$$

Now, turning to characterizations by linear regression, we start with a lemma.

LEMMA 5.1. *Let X be a positive r.v. on* $(0, b)$ *with the density* $f(x)$ *where b is a positive constant on* ∞. *Let the density of* $Y = X/U$ *be given by* (5.3). *Let* $E(X) < \infty$ *and assume that X has a linear regression on Y*

$$E(X|\, Y = y) = \begin{cases} \alpha + \beta y, & 0 < y < b \, , \\ const., & y > b \, , \end{cases} \tag{5.11}$$

for some constants α *and* β. *Then,* (i) $\alpha = 0$, $1/2 < \beta < 1$, (ii) $b < \infty$ *and* (iii) *X has the power distribution on* $(0, b)$

$$f(x) = \{\mu\beta/(1 - \beta)b^2\}(x/b)^{\beta/(1-\beta)-2}, \quad 0 < x < b \, . \tag{5.12}$$

In particular, X has a uniform distribution on $(0, b)$ *if* $\beta = 2/3$.

THEOREM 5.4. *Let X be a positive r.v. on* $(0, b)$, $b < \infty$, *with the density* $f(x)$. *Let* $Y = X/U$ *be given by* (5.3). *Then,* (5.11) *with* $\alpha = 0$, $1/2 < \beta < 1$ *holds if, and only if, X has the power distribution* (5.12). *In particular,* (5.11) *holds with* $\alpha = 0$ *and* $\beta = 2/3$ *if, and only if, X has a uniform distribution* $(0, b)$.

Our final result asserts that if a positive r.v. has a linear regression on X, then it will still have a linear regression on $Y = X/U$ if, and only if, X has the power distribution (5.12).

THEOREM 5.5. *Let X be a positive r.v. on* $(0, b)$, $b < \infty$, *with the density* $f(x)$. *Let* $Y = X/U$ *be given by* (5.3). *Let, further, T be a positive r.v. with a linear regression on X*

$$E(T|X = x) = \alpha' + \beta' x \, , \tag{5.13}$$

for some constants α' *and* β'. *If T has a linear regression also on Y,*

$$E(T|\, Y = y) = \begin{cases} \gamma + \delta y, & 0 < y < b \, , \\ const., & y > b \, , \end{cases} \tag{5.14}$$

for some constants γ *and* δ, *then,* (i) $\gamma = \alpha'$, (ii) $1/2 < \delta/\beta' < 1$ *and* (iii) *X* *has the distribution* (5.12) *with* $\beta = \delta/\beta'$. *Conversely, if* X *has the distribution* (5.12), *then,* (5.14) *holds with* $\gamma = \alpha'$ *and* $\delta = \beta\beta'$.

Acknowledgement

The author thanks the referee for helpful comments that improved the manuscript.

REFERENCES

Bissinger, B. H. (1965). A type-resisting distribution generated from considerations of an inventory decision model, *Classical and Contagious Discrete Distributions*, (ed. G. P. Patil), 15–17, Statistical Publishing Society, Calcutta.

Irwin, J. O. (1965). Inverse factorial series as frequency distributions, *Classical and Contagious Discrete Distributions*, (ed. G. P. Patil), 159–174, Statistical Publishing Society, Calcutta.

Johnson, N. L. and Kotz, S. (1969). *Discrete Distributions*, Houghton-Mifflin, Boston.

Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distributions*—1, Wiley, New York.

Katz, L. (1945). Characteristics of frequency functions obtained by first order difference equations, Dissertation, University of Michigan.

Katz, L. (1965). Unified treatment of a broad class of discrete probability distributions, *Classical and Contagious Discrete Distributions*, (ed. G. P. Patil), 175–182, Statistical Publishing Society, Calcutta.

Krishnaji, N. (1970). A characteristic property of the Yule distribution, *Sankhyā Ser. A*, **32**, 343–346.

Ord, J. K. (1967*a*). Graphical methods for a class of discrete distributions, *J. Roy. Statist. Soc. Ser. A*, **130**, 232–238.

Ord, J. K. (1967*b*). On a system of discrete distributions, *Biometrika*, **54**, 649–656.

Patil, G. P. and Joshi, S. W. (1968). *A Dictionary and Bibliography of Discrete Distributions*, Oliver & Boyd, Edinburgh.

Patil, G. P. and Rao, C. R. (1977). The weighted distributions: A survey of their applications, *Applications of Statistics*, (ed. P. R. Krishnaiah), 383–405, North-Holland, Amsterdam.

Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment, *Classical and Contagious Discrete Distributions*, (ed. G. P. Patil), 320–332, Statistical Publishing Society, Calcutta.

Xekalaki, E. (1983). A property of the Yule distribution and its applications, *Comm. Statist. A—Theory Methods*, **12**, 1181–1188.