Ann. Inst. Statist. Math.
35 (1983), Part A, 95-103

CHARACTERIZATIONS OF DISCRETE DISTRIBUTIONS BY A CONDITIONAL DISTRIBUTION AND A REGRESSION FUNCTION

T. CACOULLOS AND H. PAPAGEORGIOU

(Received Nov. 2, 1981; revised Apr. 9, 1982)

Summary

The bivariate distribution of (X, Y), where X and Y are non-negative integer-valued random variables, is characterized by the conditional distribution of Y given $X=x$ and a consistent regression function of X on Y. This is achieved when the conditional distribution is one of the distributions: a) binomial, Poisson, Pascal or b) a right translation of these. In a) the conditional distribution of Y is an x-fold convolution of another random variable independent of X so that Y is a generalized distribution. A main feature of these characterizations is that their proof does not depend on the specific form of the regression function. It is also indicated how these results can be used for goodness-of-fit purposes.

1. Introduction

Here we are concerned with characterizing the distribution of non-negative integer-valued random variables (r.v.) X and Y in terms of the conditional distribution of Y given X and the regression function $E(X|Y)$ of X on Y. Several papers have appeared in this direction. Korwar [7] considered a conditional binomial or Pascal distribution combined with linear regression and characterized the Poisson, binomial and negative binomial distributions in the former case and the geometric in the latter case. Dahiya and Korwar [4] extended these characterizations to bivariate X and Y under conditional distributions which are independent binomials or Pascal and linear regression. Khatri [5], [6], using a slightly more general approach gave similar results for the multivariate case. A case of non-linear regression was treated by Xekalaki [11] in characterizing the bivariate Poisson distribution.

In this paper more general and unifying results are obtained by by-passing the unnecessary details involved in obtaining a specific characterization under a specific regression function. This is achieved by appealing to the unicity of a solution of a first-order difference equation. Specifically, it is shown that certain conditional distributions along with the regression functions determine uniquely the distributions of X and Y, hence also of (X, Y). This fact can be used to generate a wide spectrum of distributions characterized under these conditions. For example (Theorem 3.1), given a conditional binomial distribution (of Y on X), we may choose an arbitrary given distribution for X, which in turn gives a specific regression function; thus (see Section 3), in addition to the distributions mentioned earlier, characterizations were obtained for the logarithmic distribution and several generalized (compound) distributions, e.g., Neyman, Poisson binomial, binomial Poisson, logarithmic binomial etc. Multivariate analogues are given in Papageorgiou [8].

2. Some preliminary results

It will be shown that certain conditional probability functions (p.f.) $p(y|x)$ of Y given $X=x$ together with the regression function

$$m(y)=\mathrm{E}\,[X|Y=y] \qquad y=0, 1, 2, \cdots$$

determine the bivariate distribution of (X, Y). We consider two categories of conditional distributions $p(y|x)$.

In the first category of conditional distributions, $p(y|x)$ defines an x-fold convolution of a non-negative integer-valued r.v. Z_i, say, with probability generating function h_0. Thus Y admits the representation

$$Y=Z_1+Z_2+\cdots+Z_X\,, \tag{2.1}$$

where the Z_i are i.i.d. r.v.'s independent of X, so that the p.g.f. $h_x(\cdot)$ of $p(y|x)$, $g(\cdot)$ of X, $h(\cdot)$ of Y and $G(\cdot, \cdot)$ of (X, Y) satisfy

$$h_x(v)=[h_0(v)]^x\,, \qquad h(v)=g(h_0(v))\,, \qquad G(u, v)=g(uh_0(v)) \tag{2.2}$$

(c.f. Cacoullos and Papageorgiou, [2]). Specifically, we examine the following cases:

1. *Binomial*. For a $p=1-q$, $0<p<1$, and an integer $n>0$

$$p(y|x)=\binom{nx}{y}p^y q^{nx-y} \quad y=0, 1, \cdots, nx, \quad x=0, 1, \cdots, \quad (q=1-p)\,. \tag{2.3}$$

2. *Pascal*.

$$\begin{aligned} &(\mathrm{i})\quad p(y|x)=\binom{y-1}{x-1}p^x q^y \qquad y=x, x+1, \cdots\,, \\ &(\mathrm{ii})\quad p(y|x)=\binom{x+y-1}{y}p^x q^y \qquad y=0, 1, \cdots, \quad x=1, 2, \cdots\,. \end{aligned} \tag{2.4}$$

3. *Poisson.* For a $\lambda>0$

$$p(y|x)=e^{-\lambda x}\frac{(\lambda x)^y}{y!} \qquad y=0,1,\cdots, \quad x=0,1,\cdots . \tag{2.5}$$

Under (2.3)–(2.5), the bivariate distribution of (X, Y) is determined by the distribution of Y. This follows from (2.2) directly by taking into account the relations between h_0 and g for each of (2.3)–(2.5), respectively:

$$h_0(v)=(pv+q)^n \qquad g(v)=h\left(\frac{v^{1/n}-q}{p}\right), \tag{2.3a}$$

$$\text{(2.4a)}\quad \begin{aligned} &\text{(i)} \quad h_0(v)=\frac{pv}{1-qv}, \qquad g(v)=h\left(\frac{v}{p+qv}\right), \\ &\text{(ii)} \quad h_0(v)=\frac{p}{1-qv}, \qquad g(v)=h\left(\frac{v-p}{qv}\right), \end{aligned}$$

$$h_0(v)=e^{\lambda(v-1)} \qquad g(v)=h\left(\frac{\log v}{\lambda}+1\right). \tag{2.5a}$$

Equivalently, the mixtures defined by $p(y|x)$ are identifiable, in the sense that for given $p(y|x)$ there is a one-to-one correspondence between the p.f. of X, as shown by Teicher [10] (see also Seshadri and Patil [9]) in view of the closure under convolution of the families (2.3)–(2.5) with parameter the mixing variable x.

The second category of $p(y|x)$ defines a shift to the right by x of a non-negative integer-valued r.v. with p.g.f. $h^*(v)$, so that now the corresponding p.g.f.'s are given by

$$h_x(v)=v^x h^*(v), \qquad h(v)=g(v)h^*(v). \tag{2.6}$$

Thus, given $h^*(v)$, h clearly determines g and vice-versa.

Summarizing the preceding discussion, we have

PROPOSITION 2.1. For each of the conditional distributions (mixtures) defined by (2.3)–(2.6), there is a one-to-one correspondence between $g(\cdot)$ and $h(\cdot)$ (the mixtures are identifiable).

COROLLARY 2.1. *Under* (2.3)–(2.6), *the p.f.* $p(y)$ *determines uniquely the distribution of* X.

There remains, therefore, the problem of finding $p(y)$. As already stated, this is achieved by specifying the regression function $m(y)$, as shown in Section 3.

3. The main results

In determining $p(y)$, we first show

LEMMA 3.1. *Let $p(y|x)$ satisfy*

$$xp(y|x)=c_{-1}(y)p(y-1|x)+c_0(y)p(y|x)+c_1(y)p(y+1|x) \tag{3.1}$$

with one of the $c_{-1}(y)$, $c_1(y)$ equal to zero, i.e.,

$$c_1(y)c_{-1}(y)\equiv 0 \ . \tag{3.2}$$

Then $m(y)$ determines $p(y)$.

PROOF. We have

$$m(y)=\sum_{x=0}^{\infty} x\,\mathrm{P}\,[X=x|Y=y]=\sum_{x=0}^{\infty} x\frac{p(y|x)}{p(y)}\,\mathrm{P}\,[X=x]$$

which by (3.1) can be written as

$$m(y)p(y)=c_0(y)p(y)+c_1(y)p(y+1)+c_{-1}(y)p(y-1) \ .$$

This, in virtue of (3.2), reduces to a first-order linear difference equation, which is well known to have a unique solution with a given arbitrary $p(y)$ for one y-value. This combined with the (initial) condition $\sum_{y=0}^{\infty} p(y)=1$ determines $p(y)$ uniquely for all y.

THEOREM 3.1. *Let the conditional distribution $p(y|x)$ be one of the distributions* (2.3)–(2.5). *Let $m(y)=\mathrm{E}\,[X|Y=y]$ be an arbitrary function of y consistent with $p(y|x)$. Then $p(y|x)$ and $m(y)$ together characterize the distributions of X, Y and (X, Y).*

PROOF. By Corollary 2.1 and Lemma 3.1, it is enough to verify that $p(y|x)$ satisfies (3.1) and (3.2). Thus, using the conbinatorial identity

$$x\binom{x}{y}=(y+1)\binom{x}{y+1}+y\binom{x}{y}$$

for (2.3), and the identities

$$x\binom{y-1}{x-1}=y\binom{y-1}{x-1}-(y-1)\binom{y-2}{x-1} ,$$

$$x\binom{x+y-1}{y}=(y+1)\binom{x+y}{y+1}-y\binom{x+y-1}{y}$$

for (2.4), we obtain the respective first-order difference equations (recurrences):

$$m(y)=\frac{q}{np}(y+1)\frac{p(y+1)}{p(y)}+\frac{y}{n}\,, \qquad y=0,1,\cdots, \tag{3.3}$$

(3.4)

$$\text{(i)}\quad m(y)=y-q(y-1)\frac{p(y-1)}{p(y)}\,, \qquad y=1,2,\cdots,$$

$$\text{(ii)}\quad m(y)=(y+1)\frac{p(y+1)}{qp(y)}-y\,, \qquad y=0,1,\cdots,$$

$$m(y)=\frac{y+1}{\lambda}\frac{p(y+1)}{p(y)}\,, \qquad y=0,1,\cdots. \tag{3.5}$$

Hence the proof of the theorem is complete.

In point of fact, the explicit solutions of (3.3)–(3.5) are, respectively,

$$p(y)=p(0)\left(\frac{p}{q}\right)^{y}\prod_{k=0}^{y-1}\frac{1}{k+1}(nm(k)-k)\,, \qquad y=1,2,\cdots, \tag{3.3a}$$

(3.4a)

$$\text{(i)}\quad p(y)=p(1)q^{y-1}\prod_{k=1}^{y}\frac{k-1}{k-m(k)}\,, \qquad y=2,3,\cdots,$$

$$\text{(ii)}\quad p(y)=p(0)q^{y}\prod_{k=0}^{y-1}\frac{1}{k+1}(m(k)+k)\,, \qquad y=1,2,\cdots,$$

$$p(y)=p(0)\lambda^{y}\prod_{k=0}^{y-1}\frac{m(k)}{k+1}\,, \qquad y=1,2,\cdots, \tag{3.5a}$$

with $p(0)$ or $p(1)$ determined from $\sum_{y=0}^{\infty}p(y)=1$.

Remark 3.1. In view of the representation (2.1), Theorem 3.1 can be used to characterize an infinite variety of discrete distributions: If X is an arbitrary non-negative integer-valued r.v., then Y has the generalized (compound) distribution of X by another r.v. Z (denoted by $X\vee Z$), the distribution of Z being determined by an h_0 of (2.3a)–(2.5a). This can be formulated into the constructive result:

COROLLARY 3.1. *Take $Y=X\vee Z$, where Z is a Bernoulli, a geometric or a Poisson r.v. and X arbitrary ($X=0,1,2,\cdots$). Find $m(y)$ from $m(y)=g_y'(1)$ where the conditional p.g.f. $g_y(\cdot)$ of X given $Y=y$ is given by*

$$g_y(u)=G_v^{(y)}(u,0)/G_v^{(y)}(1,0)\,, \qquad G_v^{(y)}(u_0,v_0)=\left.\frac{\partial^y G(u,v)}{\partial v^y}\right|_{v=v_0}^{u=u_0}$$

(see related results in Cacoullos and Papageorgiou [2]*). Then, each such*

$m(y)$ characterizes the distribution of X and the corresponding compound distribution of $Y=X\vee Z$. More easily given X we may determine $p(y)$ from $Y=X\vee Z$ and then use (3.3)-(3.5).

For illustration, suppose Z is a Bernoulli r.v. Then a linear regression function,

$$m(y)=a+by\ , \qquad y=0, 1, \cdots ,$$

characterizes both X and Y as (i) binomials if $0<b<1$ (ii) Poissons if $b=1$ and (iii) negative binomials if $b>1$. This is the result of Korwar [7]. In view of his limiting himself to linear $m(y)$, he missed, among others, the case of logarithmic X and Y, when

$$m(0)=a\ , \qquad m(y)=by\ , \qquad y=1, 2, \cdots , \qquad b>1\ .$$

As pointed out one may characterize a variety of interesting generalized distributions which have been given some attention in several contexts in the statistical literature. For the sake of brevity, we state only the X-distribution. Such are the Neyman (Poisson$\vee$Poisson), Poisson$\vee$binomial, Poisson$\vee$negative binomial, negative binomial$\vee$binomial, logarithmic$\vee$Poisson, etc. A unified treatment of generalized distributions is given in Charalambides [3].

Finally, let us consider the case of (2.6). More specifically, we consider shifts of a binomial, a Poisson and a negative binomial. We show

THEOREM 3.2. *Let $p(y|x)$ be one of the distributions* (3.6). *Let $m(y)$ be consistent with $p(y|x)$. Then $p(y|x)$ and $m(y)$ together characterize the distributions of X, Y and (X, Y).*

$$\text{(3.6)}\quad \begin{aligned} &\text{(i)}\quad p(y|x)=\binom{n}{y-x}p^{y-x}q^{n-y+x} && 0\leqq x\leqq y\leqq x+n\ ,\\ &\text{(ii)}\quad p(y|x)=e^{-\lambda}\frac{\lambda^{y-x}}{(y-x)!} && y\geqq x\geqq 0\ ,\\ &\text{(iii)}\quad p(y|x)=\binom{r+y-x-1}{r-1}p^r q^{y-x} && y\geqq x\geqq 0,\ (r>0)\ . \end{aligned}$$

PROOF. Under (3.6), h^* of (2.6) is given, respectively, by

$$\text{(3.7)}\quad \begin{aligned} &\text{(i)}\quad h^*(v)=(pv+q)^n\ , \qquad \text{(ii)}\quad h^*(v)=\exp\left[\lambda(v-1)\right]\ ,\\ &\text{(iii)}\quad h^*(v)=\left(\frac{p}{1-qv}\right)^r\ . \end{aligned}$$

Using the identity

$$x\binom{n}{y-x}=y\binom{n}{y-x}-(n-y+1)\binom{n}{y-x-1}-x\binom{n}{y-x-1},$$

we can write $m(y)$ for (3.6)-(i) in the form

$$m(y)=y-(n-y+1)\frac{p}{q}\,\frac{p(y-1)}{p(y)}-\frac{p}{q}\,\frac{p(y-1)}{p(y)}m(y-1).$$

For (3.6)-(ii) we easily find

$$m(y)=y-\lambda\frac{p(y-1)}{p(y)},$$

whereas for (3.6)-(iii) making use of the identity

$$x\binom{r+y-x-1}{y-x}=y\binom{r+y-x-1}{y-x}$$
$$-(r+y-1)\binom{r-y-x-2}{y-x}+x\binom{r+y-x-2}{y-x},$$

we obtain

$$m(y)=y-(r+y-1)q\frac{p(y-1)}{p(y)}+q\frac{p(y-1)}{p(y)}m(y-1).$$

Here, as for (2.3)-(2.5), $p(y)$ is found as the solution of the corresponding difference equation. Actually, we have

$$p(y)=p(0)\left(\frac{p}{q}\right)^y\prod_{k=1}^{y}\frac{n-k+1+m(k-1)}{k-m(k)},$$

(3.6a)
$$p(y)=p(0)\lambda^y\prod_{k=1}^{y}\frac{1}{k-m(k)},$$

$$p(y)=p(0)q^y\prod_{k=1}^{y}\frac{m(k-1)-(r+k-1)}{m(k)-k},$$

which completes the proof of the theorem.

As a simple application of Theorem 3.2-(ii), consider the case of a Poisson r.v. X with parameter θ (cf. Remark 3.1), giving

(3.8)
$$m(y)=\frac{\theta}{\lambda+\theta}y \qquad y=0,1,2,\cdots.$$

Under (3.6)-(ii), the regression (3.8) characterizes X as a Poisson (λ) and Y as a Poisson ($\theta+\lambda$).

In general, using Theorem 3.2, one may obtain several characterizations, of (X, Y) regarding Y as a convolution,

$$Y=X+Z\ ,$$

where X is arbitrary and Z is independently distributed as a binomial, Poisson or negative binomial.

4. Some statistical applications of the characterizations

The preceding results, in addition to their probabilistic interest, can be used in goodness-of-fit tests in a variety of situations.

For illustration, consider the case in which records (X) of accidents and corresponding fatal accidents (Y) are available for a series of periods. Then we may be faced with identifying the distribution of X and Y under the natural assumption that Y given X is binomial. This is the situation described by (2.3). A possible test, within the framework of these characterizations, is to look at the regression function $m(y)$ of X on Y. Thus, if $m(y)$ is linear (see discussion after Theorem 3.1)

$$m(y)=a+by\ , \qquad y=0, 1, \cdots\ ,$$

then a regression line with slope $b=1$, shows that X (hence also Y) is a Poisson, a $b<1$ indicates that X (hence also Y) is a binomial and a $b>1$ suggests a negative binomial for X and Y. On the other hand, a line $m(y)=by$ with $b>1$ for $y\geqq 1$ and an isolated point at $y=0$ indicates a logarithmic X.

Similar remarks can be made concerning the cases of more complicated regression functions, which as a rule, take us away from the simple classical discrete distributions. This, however, is beyond the scope of the present investigation and we shall not pursue it here any further.

UNIVERSITY OF ATHENS

REFERENCES

[1] Cacoullos, T. and Papageorgiou, H. (1980). On some bivariate probability models applicable to traffic accidents and fatalities, *Int. Statist. Rev.*, 48, 345-356.

[2] Cacoullos, T. and Papageorgiou, H. (1981). On bivariate discrete distributions generated by compounding, *Statistical Distributions in Scientific Work*, Vol. 4, D. Reidel Publishing Company, Holland, 197-212.

[3] Charalambides, Ch. H. (1977). On the generalized discrete distributions and the Bell polynomials, *Sankhyā*, B, **39**, 36-44.

[4] Dahiya, R. C. and Korwar, R. M. (1977). On characterizing some bivariate discrete distributions by linear regression, *Sankhyā*, A, **39**, 124-129.

[5] Khatri, C. G. (1978a). Characterization of some discrete distributions by linear regression, *J. Indian Statist. Ass.*, **16**, 49-58.

[6] Khatri, C. G. (1978b). Characterization of some multivariate distributions by condi-

tional distributions and linear regression, *J. Indian Statist. Ass.*, **16**, 59-70.

[7] Korwar, R. M. (1975). On characterizing some discrete distributions by linear regression, *Commun. Statist.*, **4**, 1133-1147.

[8] Papageorgiou, H. (1983). On characterizing some discrete distributions, *Austral. J. Statist.*, **25**, in press.

[9] Seshadri, V. and Patil, G. P. (1964). A characterization of a bivariate distribution by the marginal and the conditional distributions of the same component, *Ann. Inst. Statist. Math.*, **15**, 215-221.

[10] Teicher, H. (1961). Identifiability of mixtures, *Ann. Math. Statist.*, **32**, 244-248.

[11] Xekalaki, E. (1980). On characterizing the bivariate Poisson, binomial and negative binomial distributions, *Colloquia Mathematica Societatis János Bolyai*, **21**, North-Holland, 369-379.