

## Partial Independence and Finite Distributions

J. PANARETOS<sup>1</sup>

The conditional distribution plays an important role in statistical inference mainly because it can transform a problem defined in a given sample space to another problem defined in a reduced sample space. This can happen, for example, when we deal with a bivariate distribution  $f(x, y)$  and we find it convenient to talk about the distribution of  $Y$  for a given value of  $X$ . A problem of potential interest related to the consideration of the conditional distribution is to recover information regarding the distribution of the conditioning random variable (r.v.)  $X$  once the form of the conditional distribution  $s(y | x)$  of another r.v., say  $Y$ , on  $X$  is known. This problem becomes even more interesting when one moves away from the assumption of independence between functions of the r.v.'s involved to situations involving only partial independence.

In the case where the r.v.  $X$  is bounded and  $s(y | x)$  is of a certain structural form PANARETOS (1981) proved the following theorem.

**Theorem 1.** *Let  $N$  and  $m$  be positive integers such that  $N > m$ , and let  $k_0$  be the integral part of  $(N-1)/m$ . (i.e.  $k_0 = [(N-1)/m]$ ). Let  $\{(a_n, b_n); n=0, 1, \dots\}$  be a sequence of non-negative real vectors such that  $a_n > 0, n=0, 1, \dots, m; b_j > 0, j=0, 1, m+1, 2m+1, \dots, (k-1)m+1$ , for some integer  $k, 0 < k < k_0$ . For  $n=0, 1, \dots, N$ , put  $c_n = \sum_{r=0}^n a_r b_{n-r}$ . Consider a random vector  $(X, Y)$  of non-negative, integer-valued components such that  $P(X=x) = g_x, x=0, 1, \dots, N$  with  $g_0 < 1$ .*

<sup>1</sup> Department of Statistics and Actuarial Science, University of Iowa, Iowa City, Iowa 52 242, U.S.A.

Suppose that

$$\mathbf{P}(Y=y | X=x) = \frac{a_y b_{x-y}}{c_x}, \quad \text{whenever } g_x > 0, \quad (1)$$

$$y=0, 1, \dots, x; \quad x=0, 1, \dots, N.$$

If  $0 < k \leq k_0$ , then the  $k$  relations

$$\mathbf{P}(Y=y | X=Y) = \mathbf{P}(Y=y | X=Y+(j-1)m+1), \quad j=1, 2, \dots, k \quad (2)$$

hold iff

$$g_x = g_0 \frac{c_x}{c_0} \vartheta^x, \quad \text{for some } \vartheta > 0, x=0, \dots, km+l. \quad (3)$$

This result was extended (PANARETOS, 1982) to the more general case where no assumption is made about the structural form of the distribution  $s(y|x)$ . In fact the following theorem was proved.

**Theorem 2.** (PANARETOS, 1982). Consider a random vector  $(X, Y)$  of non-negative, integer-valued components such that  $\mathbf{P}(X=x) = g_x$ ,  $x=0, 1, \dots, N$  with  $g_0 < 1$ . Denote by  $[a]$  the integral part of  $a$  and let  $k_0 = \left[ \frac{N-1}{m} \right]$ . Let  $\mathbf{P}(Y=y | X=x) = s(y|x)$ ,  $y=0, 1, \dots, x$  with  $\{s(y|x) : y=0, 1, \dots, x\}$  as distributions such that for  $y=0, 1, \dots, m$  ( $m \leq x$ ),  $s(y|y) > 0$ ,  $s(y|y+qm+1) > 0$ ;  $q=0, 1, \dots, k-1$ , where  $k$  is an integer,  $l \leq k \leq k_0$ . Let

$$\begin{aligned} \mathbf{P}(Y=y | X=Y) &= \mathbf{P}(Y=y | X=Y+1) = \mathbf{P}(Y=y | X=Y+m+1) \\ &= \dots = \mathbf{P}(Y=y | X=Y+(k-1)m+1). \end{aligned} \quad (4)$$

Then

$$g_x = g_0 A(x) \lambda_0^x, \quad x=0, 1, \dots, km+1$$

where  $\lambda_0$  is a positive constant and  $A(x)$  is a function depending only on  $s(\cdot|x)$ .

However, it is interesting to observe that a similar result is valid when  $X$  is defined on  $\{l, l+1, \dots, N\}$  ( $l \geq 0$ ). The result can formally be presented in the form of the following theorem.

**Theorem 3.** Let  $(X, Y)$  be a random vector of non-negative, integer-valued components such that  $\mathbf{P}(X=x) = g_x$ ,  $x=l, l+1, \dots, N$  ( $N > l \geq 0$ ), with  $g_l < 1$ . Denote by  $[a]$  the integral part of  $a$  and let  $k_0 = \left[ \frac{N-1}{m} \right]$ . Let  $\mathbf{P}(Y=y | X=x) = s(y|x)$ ,  $y=0, 1, \dots, x$  with  $\{s(y|x) : y=0, 1, \dots, x\}$  as distributions such that for  $y=0, 1, \dots, m$  ( $m \leq x$ ),  $s(y|y+l) > 0$ ,  $s(y|y+l+qm+1) > 0$ ;  $q=0, 1, \dots, k-1$  where  $k$  is an integer  $l \leq k \leq k_0$ . Let

$$\mathbf{P}(Y=y | X=Y+l) = \mathbf{P}(Y=y | X=Y+l+(j-1)m+1) \quad (5)$$

$$j=1, 2, \dots, k; \quad y=0, 1, \dots, x.$$

Then

$$g_{l+y+em+1} = g_l \lambda_0^{y+em+1} A(l, y, q, m) \quad (6)$$

$y=0, 1, \dots, m; \varrho=0, 1, \dots, k-1; \lambda_0 = \frac{\mathbf{P}(X=Y+l+1)}{\mathbf{P}(X=Y+l)}$  a constant and  $A(l, y, \varrho, m)$  a function depending only on  $s(\cdot | x)$  and given by

$$A(l, y, \varrho, m) = \frac{s(y | l+y+1)}{s(y | l+y+\varrho m+1)} \left\{ \prod_{i=0}^y \frac{s(i | l+i)}{s(i | l+i+1)} \right\} \times \frac{s(m | l+1)}{s(0 | l+1)} \prod_{j=0}^{\varrho} \frac{s(0 | l+jm+1)}{s(m | l+jm+1)} \times \left\{ s(m | l+m+1) \prod_{i=0}^m \frac{s(i | l+i)}{s(i | l+i+1)} \right\}^{\varrho} \tag{7}$$

Clearly for  $l=0$  theorem 3 is reduced to theorem 2.

Proof. Following PANARETOS (1982) we first observe that the conditions imposed on  $s(y | x)$  imply that  $g_x > 0$  for all  $x=l, l+1, \dots, N$ . Moreover, (5) for  $j=1, 2, \dots, k$  implies

$$g_{l+y+1} = \lambda_0 g_{l+y} \frac{s(y | y+l)}{s(y | l+y+1)} \tag{8}$$

and

$$g_{l+y+\varrho m+1} = \lambda_{\varrho} g_{l+y+(\varrho-1)m+1} \frac{s(y | l+y+(\varrho-1)m+1)}{s(y | l+y+\varrho m+1)}$$

for all  $y=0, 1, \dots, m$  and all  $\varrho=1, 2, \dots, k-1$  where  $\lambda_i, i=0, 1, \dots, k-1$  are constants given by

$$\lambda_0 = \frac{\mathbf{P}(X=l+Y+1)}{\mathbf{P}(X=l+Y)}, \quad \lambda_i = \frac{\mathbf{P}(X=l+Y+im+1)}{\mathbf{P}(X=l+Y+(i-1)m+1)}; \quad i=1, 2, \dots, k-1.$$

From (8) we get

$$g_{l+y+\varrho m+1} = \lambda_{(\varrho)} g_{l+y} \frac{s(y | l+y)}{s(y | l+y+\varrho m+1)}, \tag{9}$$

$$\varrho=0, 1, \dots, k-1; \quad y=0, 1, \dots, m \quad \text{and} \quad \lambda_{(\varrho)} = \prod_{i=0}^{\varrho} \lambda_i.$$

Further, from (9) we have

$$g_{l+y+\varrho m+1} = g_l \lambda_0^y \lambda_{(\varrho)} \frac{s(y | l+y+1)}{s(y | l+y+\varrho m+1)} \prod_{i=0}^y \frac{s(i | l+i)}{s(i | l+i+1)}, \tag{10}$$

$$y=0, 1, \dots, m; \quad \varrho=0, 1, \dots, k-1.$$

It is easy, though, to verify that, for  $\varrho=1, 2, \dots, k-1$

$$\lambda_{(\varrho)} = \lambda_0^m \lambda_{(\varrho-1)} s(0 | l+\varrho m+1) \frac{s(m | l+m+1)}{s(m | l+\varrho m+1)} \prod_{i=0}^m \frac{s(i | l+i)}{s(i | l+i+1)}.$$

Therefore

$$\lambda_{(\varrho)} = \lambda_0^{\varrho m+1} \frac{s(m | l+1)}{s(0 | l+1)} \left\{ \prod_{j=0}^{\varrho} \frac{s(0 | l+jm+1)}{s(m | l+jm+1)} \right\} \times \left\{ s(m | l+m+1) \prod_{i=0}^m \frac{s(i | l+i)}{s(i | l+i+1)} \right\}^{\varrho} \quad \varrho=0, 1, \dots, k-1. \tag{11}$$

Substituting  $\lambda_{(\varrho)}$  in (10) from (11) concludes the proof of the theorem.

The reason why this result is of interest is because of its possible use in practical applications. Before giving an example, we present two special cases that are of interest as corollaries.

**Corollary 1.** (PANARETOS, 1981) *Suppose that*

$$P(Y=y | X=x) = \frac{\binom{m}{y} \binom{N-m}{x-y}}{\binom{N}{x}}, \quad y \leq x, \quad m < N, \quad m, n, N > 0 \quad (12)$$

*i.e.  $s(y | x) \sim \text{hypergeometric}(m, x, N)$ .*

*Then*

$$P(Y=y | X=Y) = P(Y=y | X=Y+(j-1)m+1), \quad j=1, 2, \dots, k_0 \quad (13)$$

*with*

$$k_0 = \left\lfloor \frac{N-1}{m} \right\rfloor$$

*iff*

$$g_x = \binom{N}{x} p^x q^{N-x}; \quad 0 < p < 1, \quad q = 1-p, \quad N > 0, \quad x = 0, 1, \dots, N \quad (14)$$

**Corollary 2.** *Suppose that a random vector  $(X, Y)$  with nonnegative, integer-valued components is such that*

$$P(X=x, Y=y) = g_x(a_y b_{x-y}/c_x), \quad y=0, 1, \dots, x; \quad x=0, 1, \dots, N \quad (15)$$

*where  $g_x$  is binomial  $(x; N, p)$ . Then*

$$P(Y=y | X=Y) = P(Y=y | X=Y+(j-1)m+1); \quad (16)$$

$$j=0, 1, \dots, k_0; \quad k_0 = \left\lfloor \frac{N-1}{m} \right\rfloor$$

*iff*

$$s(y | x) \text{ is hypergeometric}(m, x, N).$$

**Proof.** The "if" part of the proof is a side result of corollary 1. To show the "only if" part we observe from (15) that  $s(y | x)$  is of the form  $a_y b_{x-y}/c_x$ . Therefore, from theorem 1, we have that (16) is equivalent to

$$c_x = \frac{g_x}{g_0} c_0 \lambda^{-x} \quad x = 0, 1, 2, \dots, N.$$

This combined with the fact that  $g_x$  is binomial  $(x; N, p)$  implies that

$$\frac{c_x}{\sum c_n} = \binom{N}{x} \pi^x (1-\pi)^{N-x}$$

where  $\pi = \frac{p}{q\lambda + p}$ .

So, the distribution  $c_x / \sum_{n=0}^N c_n, x=0, 1, \dots, N$  is of a binomial form. But it is well-known (see e.g. RAMACHANDRAN (1967)) that the binomial distribution is uniquely

decomposable into two binomials with the same probability of success. This implies that

$$a_r = c_1 \binom{m}{r} \pi^r (1-\pi)^{m-r} \quad \text{and} \quad b_n = c_2 \binom{N-m}{n} \pi^n (1-\pi)^{N-m-n}$$

$$r = 0, 1, \dots, m; \quad n = 0, 1, \dots, N-m$$

where  $c_1$  and  $c_2$  are constants such that  $c_1 c_2 = \sum_{n=0}^N c_n$ . The result now follows easily if we replace  $a_r, b_{n-r}, c_n$  to the given form  $a_y b_{x-y} / c_x$  of  $s(y | x)$ .

**Remark.** It is perhaps worth pointing out that using the results of theorems 1, 2 and 3 characterizations of tail-truncated distributions as well as doubly truncated distributions can be obtained. (For more information concerning such distributions the reader is referred to JOHNSON and KOTZ (1969) pp. 26–27). To mention a few interesting examples in this direction one can easily verify the following corollaries.

**Corollary 3.** *Suppose that  $X$  is a random variable defined on  $\{0, 1, \dots, N\}$ . Then condition (13) is necessary and sufficient for  $g_x, x = 0, 1, \dots, N$  to be*

- i) *tail-truncated POISSON (given that  $s(y | x)$  is binomial)*
- ii) *tail-truncated negative binomial (given that  $s(y | x)$  is negative hypergeometric)*
- iii) *tail-truncated binomial (given that  $s(y | x)$  is hypergeometric)*

**Corollary 4.** *Suppose that  $X$  is a random variable defined on  $\{l, l+1, \dots, N\}$  ( $l, N \geq 0, N > l$ ). Then condition (5) with  $k = k_0$  is necessary and sufficient for  $g_x, x = l, l+1, \dots, N$  to be*

- i) *doubly-truncated POISSON (given that  $s(y | x)$  is binomial)*
- ii) *doubly-truncated negative binomial (given that  $s(y | x)$  is hypergeometric)*
- iii) *doubly-truncated binomial (given that  $s(y | x)$  is hypergeometric)*

### A Possible Application

A potential application of the results discussed in this paper is illustrated by the following example. Consider a population consisting of  $N$  individuals inoculated against a disease. Denote by  $X$  the number of individuals attacked by the disease even though inoculated. Assume that  $m$  individuals (out of  $N$ ) have been attacked in the past by the disease, and let  $Y$  be the number of individuals out of  $X$  attacked for a second time. (Clearly  $Y \leq X$ ). The distribution of  $Y$  and its relation to the distribution of  $X$  is obviously of interest. The consideration of the conditional distribution of  $Y$  given  $X$  might then help. Under the circumstances, it does not seem unreasonable to assume that for a given sample of  $n$  people attacked (out of  $N$ ) the distribution of  $Y$  given  $X = x$  is hypergeometric ( $m, x, N$ ). Since the binomial distribution is reproducible with respect to hypergeometric sampling (see SKIBINSKY (1970)), testing for a binomial form for  $Y$  is equivalent to testing

for a binomial distribution for  $X$ . Further, because of our findings this is equivalent to testing the validity of condition (13). Moreover, the number of relations one has to check depends on the magnitude of  $m$  in relation to  $N$ . If for example we have a situation with  $m > \frac{N-1}{2}$  it is evident from the theorem that we have to test only one condition (namely  $P(Y=y | X=Y) = P(Y=y | X=Y+1)$ ) to be able to say whether  $X$  follows a binomial distribution or not. On the other hand, if we have reasons to believe that the tail of the distribution of  $X$  can be ignored (and this may well be the case since the probabilities of the values of  $X$  will tend to be negligible as  $n$  approaches  $N$ ) we can appeal to the results involving tail-truncated distributions.

#### Acknowledgements

I would like to thank an Associate Editor and two referees for a careful reading of an earlier version of this paper and their constructive comments and suggestions.

#### References

- JOHNSON, N. L. and KOTZ, S. (1969). *Discrete distributions*. Boston, U.S.A.
- PANARETOS, JOHN (1981). On the joint distribution of two discrete random variables. *Ann. Inst. Statist. Math.*, A, **33** (2), 191–198.
- PANARETOS, JOHN (1982). On a structural property of finite distributions. *J. Roy. Statist. Soc.*, B, **44** (2), 209–211.
- RAMACHANDRAN, B. (1967). *Advanced theory of characteristic functions*. Statistical Publishing Society, Calcutta.
- SKIBINSKY, M. (1970). A characterization of hypergeometric distributions. *J. Amer. Stat. Assoc.* **65**, 926–929.

#### Zusammenfassung

Es ist bekannt, daß die bloße Kenntnis der bedingten Verteilung zweier zufälliger Variablen nicht ausreicht, um die Randverteilungen eindeutig zu spezifizieren. Dazu sind gewisse Zusatzinformationen notwendig. Diese werden üblicher Weise vorausgesetzt in Gestalt der Unabhängigkeit von Funktionen der beiden betrachteten zufälligen Variablen. PANARETOS (1981) führte eine Methode zur Herleitung der Randverteilungen basierend auf der Kenntnis der bedingten Verteilung und einer Annahme über die partielle Unabhängigkeit ein. Es wird eine Erweiterung dieses Ergebnisses dargestellt, das sich auf gestutzte Verteilungen bezieht. Basierend auf der eindeutigen Zerlegung des Binomialgesetzes wird eine interessante Eigenschaft der hypergeometrischen Verteilung enthüllt.

**Резюме**

Известно, что знание только условных распределений двух случайных величин не хватает чтобы специфицировать функции маргинального распределения. Некоторая дополнительная информация необходима. Обычно эта информация находится в некотором форме независимости функций двух случайных величин.

Панаретос (1981) вводил метод вычисления маргинального распределения, который ползуется только знанием условных распределений и некотором предположением о частичной независимости. В данной работе дается расширение этого результата на классе усеченных распределении. Открывается интересное свойство гипергеометрического распределения с помощью разложения биномиального закона.

Received December 1981; revised December 1982.