Canad. Math. Bull. Vol. **15** (3), 1972.

DISTRIBUTION OF THE SUM OF VARIATES FROM TRUNCATED DISCRETE POPULATIONS

BY

A. K. MD. EHSANES SALEH AND M. A. RAHIM

1. **Introduction.** Distributions often arise in practice where one or more values of the variate are unobserved. Various practical problems have been referred by Finney [1], David and Johnson [4], Bliss and Fisher [3]. It is of interest to know the exact distribution of the sum of variates from such truncated discrete population. In this paper, utilising the property of characteristic function certain general results are shown. The distribution of the sum of independent variables from a discrete population, truncated by any set of s distinct values, follows from them immediately. Using these results the exact distributions of the sum, from binomial, poisson, negative binomial and geometric population, truncated from anywhere, are derived.

2. **General results.** Let X be a discrete random variable with probability mass function

$$P(X = x) = f_X(x), \qquad X = 0, 1, 2, \ldots$$

which has characteristic function $\phi(t, X)$. Let any set of s distinct values of X, denoted by $r_1, r_2, \ldots, r_s$ be truncated. Then we have the new random variable X' with probability mass function

$$P(X' = x) = f_{X'}(x) = \frac{1}{1-Q} f_X(x) \tag{2.1}$$

where $Q = f_X(r_1)+f_X(r_2)+\cdots+f_X(r_s)$. We are interested to determine the exact distribution of $Y = X_1'+X_2'+\cdots+X_n'$ where each X_i' is distributed as in (2.1). The characteristic function of any X_i' is

$$\phi(t, X_i') = \sum \frac{1}{1-Q} e^{itx} f_X(x) = \frac{1}{1-Q} [\phi(t, X)-f_X(r_1)e^{itr_1}-\cdots-f_X(r_s)e^{itr_s}]$$

where the summation extends over all values of X except $r_1, r_2, \ldots, r_s$. Hence the characteristic function of Y is

$$\begin{aligned}\phi(t, Y) &= \frac{1}{(1-Q)^n} [\phi(t, X)-f_X(r_1)e^{itr_1}-\cdots-f_X(r_s)e^{itr_s}]^n \\ &= \frac{1}{(1-Q)^n} \sum_{k_0,k_1,\ldots k_s} (-1)^{n-k_0} \binom{n}{k_0, k_1, \ldots k_s} \\ &\quad \times [f_X(r_1)]^{k_1}[f_X(r_2)]^{k_2} \cdots [f_X(r_s)]^{k_s}[\phi(t, X)]^{k_0} e^{it\{r_1k_1+r_2k_2+\cdots+r_sk_s\}} \\ &= \sum_{k_0,k_1,\ldots k_s} B(k_0, k_1, \ldots k_s)[\phi(t, X)]^{k_0} e^{it\{r_1k_1+r_2k_2+\cdots+r_sk_s\}}\end{aligned} \tag{2.2}$$

Received by the editors February 5, 1970 and, in revised form, November 16, 1970.

https://doi.org/10.4153/CMB-1972-072-1 Published online by Cambridge University Press

where we write

$$(2.3)\quad B(k_0, k_1, \ldots k_s) = \frac{(-1)^{n-k_0}}{(1-Q)^n}\binom{n}{k_0, k_1, \ldots k_s}[f_X(r_1)]^{k_1}\cdots[f_X(r_s)]^{k_s}$$

and the summation extends over all combinations of $k_0, k_1, k_2, \ldots, k_s$ such that $k_0+k_1+k_2+\cdots+k_s = n$. Hence, the probability mass function of Y is given by

$$(2.4)\quad f_Y(y) = \sum_{k_0,k_1,\ldots k_s} B(k_0, k_1, \ldots k_s)\frac{1}{2\pi}\int_{-\pi}^{\pi}[\phi(t, X)]^{k_0}e^{-it\{y-(r_1k_1+\cdots+r_sk_s)\}}\,dt$$

In the case of infinite discrete population, if we want to truncate all values from the right up to a point, i.e., if we truncate the values of $X=r+1$, $r+2$, $R+3, \ldots \infty$; (2.2) reduces to

$$\phi(t, Y) = \frac{1}{Q'^n}[f_X(0)e^{0it}+f_X(1)e^{it}+f_X(2)e^{2it}+\cdots+f_X(r)e^{rit}]^n$$

$$= \frac{1}{Q'^n}\sum_{k_0,k_1,\ldots k_r}\binom{n}{k_0, k_1, \ldots k_r}[f_X(0)]^{k_0}[f_X(1)]^{k_1}\cdots[f_X(r)]^{k_r}e^{it\{k_1+2k_2+\cdots+rk_r\}}$$

where $Q' = 1-Q = f_X(0)+f_X(1)+\cdots+f_X(r)$. Hence the probability mass function of Y is given by

$$(2.5)\quad \begin{aligned} f_Y(y) &= \frac{1}{Q'^n}\sum_{k_0,k_1,\ldots k_r}\binom{n}{k_0, k_1, \ldots k_r}[f_X(0)]^{k_0}[f_X(1)]^{k_1}\cdots[f_X(r)]^{k_r} \\ &\quad\times\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-it\{y-(k_1+2k_2+\cdots+rk_r)\}}\,dt \\ &= \frac{1}{Q'^n}\sum_{k_0,\ldots k_r}\binom{n}{k_0, k_1, \ldots k_r}[f_X(0)]^{k_0}[f_X(1)]^{k_1}\cdots[f_X(r)]^{k_r} \\ &\quad\times\varepsilon[y-\{k_1+2k_2+\cdots+rk_r\}] \end{aligned}$$

where $\varepsilon[y-\{k_1+2k_2+\cdots+rk_r\}]$ is the one point distribution, such that

$$(2.6)\quad \begin{aligned} \varepsilon[y-\{k_1+2k_2+\cdots+rk_r\}] &= 1 \quad \text{when} \quad y = \{k_1+2k_2+\cdots+rk_r\} \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

For any specified discrete population, from which any set of finite distinct values are truncated away, the distribution of the sum Y is directly obtainable from (2.4). If however, the population is infinite and all values are truncated from the right up to certain point, the distribution of the sum Y is obtainable from (2.5). We now derive the distributions in a few specific cases.

3. **Case of binomial population.** In this case we have

$$f_X(x) = \binom{N}{x}q^{N-x}p^x; \qquad x = 0, 1, 2, \cdots N; \qquad \phi(t, X) = (q+pe^{it})^N.$$

https://doi.org/10.4153/CMB-1972-072-1 Published online by Cambridge University Press

The distribution of the sum Y follows from (2.4) as

$$f_Y(y) = \sum_{k_0,k_1,\ldots k_s} B(k_0, k_1, \ldots k_s) \frac{1}{2\pi} \int_{-\pi}^{\pi} (q+pe^{it})^{Nk_0} e^{-it\{y-(r_1k_1+\cdots+r_sk_s)\}}\, dt$$

(3.1)
$$= \sum_{k_0,k_1,\ldots k_s} B(k_0, k_1, \ldots k_s) \binom{Nk_0}{y-(r_1k_1+\cdots+r_sk_s)} \times q^{Nk_0-y+(r_1k_1+\cdots+r_sk_s)} p^{y-(r_1k_1+\cdots+r_sk_s)}$$

where $\binom{Nk_0}{y-(r_1k_1+\cdots+r_sk_s)}$ is defined to be zero when $y-(r_1k_1+\cdots+r_sk_s) > Nk_0$. The distribution obtained by Malik [2], when a single value $r_1 = 0$ is truncated away, follows from (3.1) as a special case.

4. **Case of poisson population.** In this case we have

$$f_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots; \qquad \phi(t, X) = e^{\lambda(e^{it}-1)}.$$

The distribution of the sum Y follows from (2.4) as

$$f_Y(y) = \sum_{k_0,k_1,\ldots k_s} B(k_0, k_1, \ldots k_s) e^{-\lambda k_0} \frac{(\lambda k_0)^{y-(r_1k_1+\cdots+r_sk_s)}}{[y-(r_1k_1+\cdots+r_sk_s)]!} \tag{4.1}$$

In the special circumstances when all the values $X = r+1, r+2, r+3, \ldots, \infty$ are truncated away from the right, the distribution of the sum Y follows from (2.5) as

(4.2)
$$f_Y(y) = \frac{e^{-n\lambda}}{Q'^n} \sum_{k_0,k_1,\ldots k_r} \binom{n}{k_0, k_1, \ldots k_r} \frac{\lambda^{(k_1+2k_2+\cdots+rk_r)}}{(2!)^{k_2}(3!)^{k_3}\cdots(r!)^{k_r}} \times \varepsilon[y-(k_1+2k_2+\cdots+rk_r)]$$

5. **Case of negative binomial population.** In this case we have

$$f_X(x) = \binom{-\theta}{x} p^\theta(-q)^x; \qquad \begin{matrix} 0 \le p \le 1,\ \theta > 0 \\ x = 0, 1, 2, \ldots \end{matrix}; \qquad \phi(t, X) = \left(\frac{p}{1-qe^{it}}\right)^\theta.$$

The distribution of the sum Y follows from (2.4) as

(5.1)
$$f_Y(y) = \sum_{k_0,k_1,\ldots k_s} B(k_0, k_1, \ldots k_s) \binom{-k_0\theta}{y-(r_1k_1+\cdots+r_sk_s)} p^{k_0\theta}(-q)^{y-(r_1k_1+\cdots+r_sk_s)}$$

It may be noted that putting $\theta = 1$, the distribution of the sum Y from a truncated geometric distribution follows from (5.1).

https://doi.org/10.4153/CMB-1972-072-1 Published online by Cambridge University Press

References

1. D. J. Finney, *The truncated binomial distribution*, Ann. of Eugenics, **14** (1949), 319–328.

2. H. J. Malik, *Distribution of the sum of truncated binomial variates*, Canad. Math. Bull. (3) **12** (1969), p. 344.

3. C. I. Bliss, and R. A. Fisher, *Fitting the binomial distribution to biological data and note on the efficient fitting of the negative binomial*, Biometrics, **9** (1953), 176–200.

4. F. N. David, and N. L. Johnson, *The truncated poisson*, Biometrics, **8** (1952), 275–285.

5. P. R. Rider, *Truncated binomial and negative binomial distributions*, Amer. Statist. Assoc. **50** (1955), 877–883.

6. A. K. Md. Ehsanes Saleh, *Distribution of the sum of independent truncated poisson variates*, Dacca Univ., Bangladesh, J. Statist. Res. (1) **4** (1970), 32–36.

Carleton University,
Ottawa, Ontario