The Annals of Statistics
1974, Vol. 2, No. 4, 819-821

A NOTE ON THE ASYMPTOTIC EQUIVALENCE OF SAMPLING WITH AND WITHOUT REPLACEMENT

BY OLAV KALLENBERG

University of Göteborg

The vague statement that "sampling with and without replacement from a finite population are approximately equivalent when the sampling fraction is small" is given a precise meaning in terms of limit theorems for distributions in R^∞ and $D[0, \infty)$.

It is a well-known and commonly used "fact" that sampling with and without replacement from a finite population are "approximately equivalent" provided the sampling fraction is "small". This vague statement has partially been made precise in papers by Hájek (1960) and Rosén (1964). The aim of the present note is to give a more complete justification in the form of two limit theorems which follow easily from results in [4].

For $k = 1, 2, \cdots$, let Π_k be a population with real (or vector-valued) elements $x_{k1}, \cdots, x_{kN_k}$, and let $\xi_{k1}, \xi_{k2}, \cdots$ and $\eta_{k1}, \cdots, \eta_{kN_k}$ be the values obtained by simple random sampling with and without replacement respectively from Π_k. For convenience of writing, let $\eta_{kj}, j > N_k$, be arbitrary random variables. We introduce the random sequences

$$\xi_k = (\xi_{k1}, \xi_{k2}, \cdots), \qquad \eta_k = (\eta_{k1}, \eta_{k2}, \cdots), \qquad k = 1, 2, \cdots,$$

and further, for arbitrary positive $n_1, n_2, \cdots$, the random processes X_k and Y_k, $k = 1, 2, \cdots$, defined by

$$X_k(t) = \sum_{j \leq n_k t} \xi_{kj}, \qquad Y_k(t) = \sum_{j \leq n_k t} \eta_{kj}, \qquad t \geqq 0.$$

The former will be considered as random elements in R^∞ endowed with the product topology (cf. [1], page 19), and the latter as random elements in $D[0, \infty)$ endowed with the Skorohod–Stone topology (cf. [5]). Write $=_d$ and $\to_d$ for equality and convergence in distribution respectively with respect to these topologies [1].

THEOREM 1. *Suppose that $N_k \to \infty$. Then $\xi_k \to_d$ some ξ if and only if $\eta_k \to_d$ some η, and in case of convergence, $\xi =_d \eta$.*

THEOREM 2. *Suppose that $n_k \to \infty$ and $n_k/N_k \to 0$. Then the following statements are equivalent.*

(i) $X_k \to_d$ *some* X,

(ii) $Y_k \to_d$ *some* Y,

Received March 1973; revised August 1973.

AMS 1970 *subject classifications.* 62D05, 60F05.

Key words and phrases. Sampling from finite populations, drawing with and without replacement, convergence in distribution.

(iii) $X_k(t) \to_d$ *some* α_t *for some (any)* $t > 0$,
(iv) $Y_k(t) \to_d$ *some* β_t *for some (any)* $t > 0$.

If the statements are true, then $X =_d Y$ *and* $X(t) =_d \alpha_t =_d \beta_t =_d Y(t)$, $t > 0$.

In Theorem 1, the "only if" part is due to Rosén (1964, Lemma 3.1). As for Theorem 2, the equivalence of (i) and (iii) is due to Prohorov (1956, page 197) and Skorohod (1957, Theorem 2.7), while the equivalence of (iii) and (iv) is due to Hájek (1960, Theorem 5.1) in the particular case when $\text{Var}(Y_k(t)) \to \text{Var}(\beta_t) < \infty$. (Note, however, that the statements of Theorem 2 may be true even without this assumption, cf. Theorem 4.1 in [4].)

To indicate how the above theorems may be deduced from [4], note that Theorem 1 follows from the fact that, by Theorems 1.2 and 1.3 in [4], the convergence in distribution of $\{\xi_k\}$ and $\{\eta_k\}$ are both equivalent to weak convergence of the empirical distributions

$$\pi_k = N_k^{-1} \sum_{j=1}^{N_k} \delta_{x_{kj}}, \qquad k = 1, 2, \cdots,$$

towards some probability distribution μ, and that, in case of convergence, the components of the limiting sequence are independent with common distribution μ. Similarly, statements (i) and (ii) of Theorem 2 are equivalent (with $X =_d Y$ in case of convergence) since, by Theorems 3.2 and 4.1 in [4], they are both equivalent to the conditions on $\{\pi_k\}$ occurring in the classical limit theorem for null-arrays (cf. [2], page 564). Next, by continuity, (ii) implies (iv) with $\beta_t =_d Y(t)$, so it remains to derive (ii) from (iv). Now (iv) implies tightness of $\{Y_k\}$ by Lemma 2.1 in [4], and any limit Y of $\{Y_k\}$ has stationary independent increments since this is true for $\{X_k\}$. Again $Y(t) =_d \beta_t$, determining the distribution of Y uniquely, so we may use Theorem 2.3 in [1] to complete the proof.

It may be worthwhile to point out that Theorem 1 is equivalent to the assertion (i) $\Leftrightarrow$ (ii) of Theorem 2 in the case $n_1 = n_2 = \cdots$. However, the assertions (iii) $\Rightarrow$ (i) and (iv) $\Rightarrow$ (ii) are false in this case, although the corresponding tightness implications remain true (cf. Lemma 2.1 in [4]). It is also interesting to note that, if the statements of Theorem 2 (or Theorem 1) are true for one particular sequence $\{n_k\}$, then nothing can be said in general about the asymptotic behavior when using an essentially different sequence $\{n_k'\}$ (and an appropriate renormalization of the X_k and Y_k). In particular, neither of the above theorems may be deduced from the other one.

Acknowledgment. I am grateful to the referee for some valuable suggestions.

REFERENCES

[1] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.
[2] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, **2** (2nd ed.). Wiley, New York.
[3] Hájek, J. (1960). Limiting distributions in simple random sampling from a finite population. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** 361-374.

[4] KALLENBERG, O. (1973). Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **27** 23–36.

[5] LINDVALL, T. (1973). Weak convergence of probability measures and random functions in the function space $D[0, \infty)$. *J. Appl. Probability* **10** 109–121.

[6] PROKHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1** 157–214.

[7] ROSÉN, B. (1964). Limit theorems for sampling from a finite population. *Ark. Mat.* **5** 383–424.

[8] SKOROHOD, A. V. (1957). Limit theorems for stochastic processes with independent increments. *Theor. Probability Appl.* **2** 138–171.

DEPARTMENT OF MATHEMATICS
CHALMERS INSTITUTE OF TECHNOLOGY
AND THE UNIVERSITY OF GÖTEBORG
FACK S-402 20 GÖTEBORG 5
SWEDEN