Pranab Kumar Sen

The Hájek asymptotics for finite population sampling and their ramifications

Kybernetika, Vol. 31 (1995), No. 3, 251--268

Persistent URL: `http://dml.cz/dmlcz/124724`

KYBERNETIKA — VOLUME 31 (1995), NUMBER 3, PAGES 251-268

THE HÁJEK ASYMPTOTICS FOR FINITE POPULATION SAMPLING AND THEIR RAMIFICATIONS

PRANAB KUMAR SEN

In finite population (equal as well as unequal probability) sampling late Jaroslav Hájek's contributions to the general asymptotics are fundamental. In the last two decades more research work has been accomplished in this area with the basic ideas germinating from Hájek's work. A systematic review of such developments with due emphasis on some martingale formulations is presented here.

1. INTRODUCTION

Finite population sampling (FPS) theory provides the most useful methodology for drawing statistical conclusions for population characteristics based on a representative sample from it. The simple random sampling with replacement (SRSWR) is the precursor of other relatively more complex (and yet useful) sampling designs which are adopted in FPS. Simple random sampling without replacement (SRSWOR) retains the equal probability sampling (EPS) structure but violates the independence of the sample observations. There are various unequal probability sampling (UPS) schemes which may have distinct advantages over SRSWR/SRSWOR. Although the underlying probability structure for such FPS schemes are well defined, their complexities increase with the increase in the size of the population and sample, so that asymptotics become indispensible for obtaining simplified (manageable) expressions for the sample inclusion probabilities and for studying various statistical properties of sample statistics. In particular, in order to set a confidence interval for a suitable population characteristic or to test for a plausible hypothesis on the same, we need to estimate from the sample the mean square error (MSE) of the estimator of this population characteristic. Then suitable large sample theory (generally leading to normal laws) can be incorporated to draw the desired statistical conclusions. The estimation of such MSE's usually involves estimation or simplification of inclusion probabilities which may require some nonstandard asymptotics. Moreover, lack of independence may preclude the use of standard central limit theorems (CLT) for establishing the asymptotic normality results. In fact, lack of independence, unequal probabilities for inclusion and complex sampling designs all add complexities in the treatment of general asymptotics. Hájek's ingenuity lies in the general formulation of FPS asymptotics in a rigorous probabilitistic framework which allows diverse modern tools to accomplish the needed task.

Hájek [12] contains a novel and basic probabilistic approach to FPS covering both EPS and UPS schemes in a common vein. While most of these works were accomplished in the sixties and published in contemporary journals, the publication of this impressive monograph was considerably delayed due to Hájek's serious health problems in early seventies culminating with his premature and unexpected death in 1974. Several colleagues and former students of Hájek voluntarily took up the pending task of putting the finishing touches to the material left behind by him. Needless to comment that as in nonparametrics, in FPS asymptotics too, under the pioneering guidance of Jaroslav Hájek, the Czech school of probability and statistics has made an outstanding contribution, and the flow of research is still on. Jitka Dupačová, Jan Ámos Víšek and Zuzana Prášková are all among the other disciples of Hájek whose contributions to this field are noteworthy, and their efforts have led to extensions of the basic results of Hájek [5], [8], [11] in various directions. Sampling theory (for Poisson sampling and rejective sampling) with varying probabilities (in FPS) developed by Hájek [8] immediately caught the attention of prominant statisticians from all over the world. Rosén [31] – [33], Sampford [34], [35] and Karlin [18] have significant contributions in this vein. In course of this study, certain generalized occupancy models cropped up in a natural way, and this led to some further interactive research in the area of capture-mark-recapture (CMR) methodology. Some martingale characterizations in this context have also evolved and led to further developments of asymptotics in successive (sub-) sampling with varying probability schemes (viz., Sen [38], [40]). Section 3 is devoted to the exposition of these works in FPS.

Invariance principles in FPS has been one of the areas of recent research interest. The main inspiration came from the basic work of Hájek in the mid-sixties, although most of these developments took place some ten years later on. Section 4 deals with these developments. The concluding section is devoted to some general remarks.

2. SRSWOR AND PCLT

Consider a finite population $\Pi_N = \{a_1, \ldots, a_N\}$ of size N (where the a_j need not be all distinct), and let $\boldsymbol{X}_n = \{x_1, \ldots, x_n\}$ be a sample of size n drawn from Π_N according to a probability law P_{Nn}. In SRSWR, this probability law is given by

$$P\{X_1 = a_{i_1}, \ldots, X_n = a_{i_n}\} = N^{-n}, \tag{2.1}$$

for every $i_j = 1, \ldots, N$, $j = 1, \ldots, n$, and $n \geq 1$. Therefore, the X_i are independent and identically distributed random variables (i.i.d.r.v.) with the probability distribution

$$P\{X_i = a_k\} = N^{-1}, \quad \text{for } 1 \leq k \leq N;\ i \geq 1. \tag{2.2}$$

In SRSWOR, we have

$$P\{X_1 = a_{i_1}, \ldots, X_n = a_{i_n}\} = N^{-[n]} = \{N \ldots (N - n + 1)\}^{-1}, \tag{2.3}$$

for every $1 \leq i_1 \neq \ldots \neq i_n \leq N$, $n \leq N$. Although, marginally each X_i has the same probability distribution in (2.2), they are no longer independent. The X_i are interchangeable (or exchangeable) r.v.'s, but their intra-class dependence pattern depends on $\{N, n\}$.

To introduce a permutational (conditional) probability measure ($\mathcal{P}_n$), let us consider a set $Y_1, \ldots, Y_n$ of i.i.d.r.v.'s with a continuous distribution function (d.f) F on $\mathbb{R}$. Let $\boldsymbol{Z}_n = \{Y_{n:1} < \cdots < Y_{n:n}\}$ be the collection of the ordered r.v.'s (order statistics) $Y_{n:1}, \ldots, Y_{n:n}$ corresponding to $Y_1, \ldots, Y_n$. Then given $\boldsymbol{Z}_n$, $\boldsymbol{Y}_n = (Y_1, \ldots, Y_n)$ takes on each permutation of the coordinates of $\boldsymbol{Z}_n$ with the common (conditional) probability $(n!)^{-1}$. This is the genesis of permutational probability laws. To deal with the case of possibility discrete distributions and to encompass the SRSWOR schemes as well, by reference to (2.3), we let

$$X_i = a_{R_i}, \quad i \geq 1, \tag{2.4}$$

so that $(R_1, \ldots, R_n)$ is a sub-vector of $(1, \ldots, N)$, such that

$$1 \leq R_1 \neq \ldots \neq R_n \leq N, \quad \forall\, 1 \leq n \leq N. \tag{2.5}$$

Suppose now that our interest lies in the estimation of the population mean $\bar{a}_N = N^{-1} \sum_{i \leq N} a_i$. The sample mean $\bar{X}_n = n^{-1} \sum_{i \leq n} X_i$ is a natural estimator of $\bar{a}_N$ having some optimal properties (viz., Nandi and Sen [22] covering the more general case of U-statistics in SRSWOR). We may then write equivalently

$$\bar{X}_n = \sum_{i \leq N} \delta_i a_{R_i}, \tag{2.6}$$

where

$$\delta_i = \begin{cases} 1/n, & 1 \leq i \leq n, \\ 0, & n < i < N; \end{cases} \tag{2.7}$$

$$\{R_{n+1}, \ldots, R_N\} = \{1, \ldots, N\} \backslash \{R_1, \ldots, R_n\}. \tag{2.8}$$

Recall that $\boldsymbol{R}_N = (R_1, \ldots, R_N)$ takes on each permutation of $\{1, \ldots, N\}$ with the common probability $(N!)^{-1}$ (independently of the a_j, $j \leq N$), so that we have a completely specified probability measure. We denote this permutation law by $\mathcal{P}_N$. Then, for every $n\colon 1 \leq n \leq N$, it follows that

$$\mathrm{E}_{\mathcal{P}_N}(\bar{X}_n) = \bar{a}_N \quad \text{and} \quad n\mathrm{V}_{\mathcal{P}_N}(\bar{X}_n) = [(N-n)/(N-1)]A_N^2, \tag{2.9}$$

where

$$A_N^2 = N^{-1} \sum_{i \leq N} \{a_i - \bar{a}_N\}^2. \tag{2.10}$$

Madow [19] used the representation in (2.6) and incorporated the classical Wald–Wolfowitz [50] PCLT to establish that

$$n^{1/2}(\bar{X}_n - \bar{a}_N)/\{A_N\left((N-n)/(N-1)\right)^{1/2}\} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{2.11}$$

whenever N and n are large. The regularity conditions underlying the Wald–Wolfowitz PCLT were rather stringent, and later on, these were relaxed by Noether [23], Hoeffding [16], Motoo [21], and others. The final say in this context is due to Hájek [6]. We present this evolutionary picture in a proper perspective.

Let $\Pi_N = \{a_{N1}, \ldots, a_{NN}\}$ and $\Gamma_N = \{b_{N1}, \ldots, b_{NN}\}$ be two sequences of real numbers, and define $\bar{a}_N$, $\bar{b}_N$, A_N^2 and B_N^2 as in (2.9) and earlier. Then a linear permutation statistic is defined as

$$L_N = \sum_{i \le N} b_{Ni} a_{NR_i}, \tag{2.12}$$

where $\boldsymbol{R}_N = (R_1, \ldots, R_N)$ takes on each permutation of $(1, \ldots, N)$ with the common probability $(N!)^{-1}$. Side by side, we may also introduce a bilinear permutation statistic as

$$Q_N = \sum_{i \le N} d_N(i, R_i), \tag{2.13}$$

where $D_N = \{d_N(i,j) : 1 \le i, j \le N\}$ is a double sequence of real numbers. A condition, known in the literature as the *Noether condition*, that arises invariably with PCLT states that for a sequence, say Π_N, as $N \to \infty$,

$$\max_{1 \le i \le N} (a_{Ni} - \bar{a}_N)^2 \Big/ \sum_{j \le N} (a_{Nj} - \bar{a}_N)^2 \to 0. \tag{2.14}$$

Noether [23] replaced one of the Wald–Wolfowitz conditions by (2.14) and showed that the PCLT holds. Hoeffding [16] extended this result for Q_N. Both the approaches are based on the so called "method of moments", and thereby relate to sufficient (but not necessary) regularity conditions.

Hájek [6] had a completely different approach, and he obtained a necessary and sufficient condition for the PCLT to hold for L_N. For this (without loss of generality), we let $a_{N1} \le \cdots \le a_{NN}$, and introduce a quantile function $a_N(\cdot) = \{a_N(\lambda) : 0 < \lambda \le 1\}$ by letting

$$a_N(\lambda) = a_{Ni} \quad \text{for } (i-1)/N < \lambda \le i/N, \quad 1 \le i \le N. \tag{2.15}$$

Also, let $U_1, \ldots, U_N$ be i.i.d.r.v.'s having the uniform (0,1) d.f., and let

$$L_N^0 = \sum_{i \le N} (b_{Ni} - \bar{b}_N) a_N(U_i) + N \bar{b}_N \bar{a}_N. \tag{2.16}$$

Then, Hájek [6] succeeded in showing that under minimal regularity conditions, as $N \to \infty$,

$$E(L_N - L_N^0)^2 / V(L_N) \to 0. \tag{2.17}$$

When Γ_N satisfy (2.14) and $a_N(\cdot)$ is square integrable, the CLT applies to L_N^0, so that (2.17) and the Slutzky theorem lead to the CLT for L_N as well. Some further simplifications can be made when Π_N(or Γ_N), $N \ge N_o$, are defined in a special way (as is usually done in nonparametrics). Let $U_{N:1} < \cdots < U_{N:N}$ be the order statistics corresponding to $U_1, \dots, U_N$, and define

$$a_{Ni}^* = E\phi(U_{N:i}), \quad 1 \le i \le N, \tag{2.18}$$

where $\phi\colon (0,1) \to \mathbb{R}$, is assumed to be square integrable and without loss of generality, we set $\bar{a} = \int_0^1 \phi(u)\,du = 0$ ($\Rightarrow \bar{a}_N^* = 0$, $\forall\ N \ge 1$). Also, let

$$A^2 = \int_0^1 \phi^2(u)\,du \quad (< \infty) \quad \text{and} \quad A_N^{*2} = \frac{1}{N-1}\sum_{i=1}^{N} a_{Ni}^{*2}. \tag{2.19}$$

In (2.12), replacing the a_{Ni} by a_{Ni}^*, we define L_N^* (in place of L_N). Further let $\mathcal{B}_N = \mathcal{B}(\boldsymbol{R}_N)$ be the sigma-field generated by the rank vector $\boldsymbol{R}_N$, $N \ge 1$. Then,

$$\mathrm{E}(L_N^0)|\mathcal{B}_N) = L_N^*, \quad \forall\ N \ge 1, \tag{2.20}$$

so that

$$\mathrm{E}(L_N^* - L_N^0)^2 = \mathrm{E}(L_N^{02}) - \mathrm{E}(L_N^{*2}) = NB_N^2(A^2 - A_N^{*2}), \tag{2.21}$$

and it is easy to verify that A_N^{*2} is $\uparrow$ in N with $\lim_{N\to\infty} A_N^{*2} = A^2$. Therefore, (2.17) follows readily from (2.21). Moreover, note that for an arbitrary Π_N, we have

$$\mathrm{E}_{\mathcal{P}_N}(L_N - L_N^*)^2/\mathrm{E}_{\mathcal{P}_N} L_N^{*2} \le \left(\sum_{i\le N}[a_{Ni} - a_{Ni}^*]^2\right) \Big/ \left(\sum_{i\le N} a_{Ni}^{*2}\right), \tag{2.22}$$

so that for an arbitrary L_N, whenever,

$$N^{-1}\sum_{i\le N}(a_{Ni} - a_{Ni}^*)^2 \to 0, \quad \text{as } N \to \infty, \tag{2.23}$$

(2.17) can be verified through (2.22) and (2.23). Hájek [6] has an elegant mathematical treatment of this quadratic mean equivalence results in PCLT. In fact, his treatment goes far beyond this basic result; the genesis of martingale characterizations in nonparametrics and FPS lies in this treatise.

With our primary focus on the FPS asymptotics, we discuss only briefly some relevant martingale characterizations for L_N^* and some related rank statistics. As in (2.6), often, the Γ_N arise from a single sequence $\{b_n;\ n \ge 1\}$ of real numbers (e.g., $b_{Ni} = (b_i - \bar{b}_N)/B_N$, $1 \le i \le N$). Also, the R_i depend on N, and hence, we write them as R_{Ni}, $1 \le i \le N$; $N \ge 1$. As such, we write

$$L_N^* = \sum_{i\le N}(b_i - \bar{b}_N)a_N^*(R_{Ri}), \quad N \ge 1, \tag{2.24}$$

where a_{Ni}^* is written as $a_N^*(i)$, $1 \le i \le N$. As in Hájek [12], we may relate the SRSWR to a superpopulation model, and assume that this superpopulation has a continuous distribution. Then, along the lines of Sen and Ghosh [46], it is easy to very that

$$\mathrm{E}\{L_{N+1}^*|\mathcal{B}_N\} = L_N^* \text{ a.e.}, \ \forall\ N \ge 1, \tag{2.25}$$

where $\mathcal{B}_N = \mathcal{B}(\boldsymbol{R}_N)$ is nondecreasing in N. Therefore,

$$\{L_N^*, \mathcal{B}_N;\ N \geq 1\} \quad \text{is a zero mean martingale.} \tag{2.26}$$

A similar martingale characterization holds for signed-rank statistics and some other related ones. Such martingale characterizations (and approximations for other Π_N) provide access to general asymptotics, and these are presented in a systematic manner in Sen [41]). In passing, we may remark that in a multivariate setup (i.e., when the elements of Π_N are themselves p-vectors, for some $p \geq 1$), the treatment of PCLT is a little bit more complex: Chatterjee and Sen [2] formulated the rank-permutation approach which extends (2.3) in a natural way (to column permutations of a $(p \times N)$ rank collection matrix). In this setup too, the Hájek [6] quadratic mean equivalence plays a basic role, and a martingale approach to such multivariate PCLT has been formulated in Sen [44].

Let us return to the SRSWOR asymptotics in a more general setup. The sample mean $\bar{X}_n$, in (2.6), is a special case of a U-statistic. Based on the sample $\boldsymbol{X}_n$ and a symmetric kernel $g(x_1, \ldots, x_k)$ of degree $k (\geq 1)$, we define

$$U_n = U(\boldsymbol{X}_n) = \binom{n}{k}^{-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} g(X_{i1}, \ldots, X_{ik}), \tag{2.27}$$

and by (2.3), we obtain that U_n is an unbiased estimator of

$$\begin{aligned} \Theta_N &= \Theta(\Pi_N) = U(\Pi_N) \\ &= N^{-[n]} \sum_{\{1 \leq i_1 \neq \cdots \neq i_k \leq N\}} g(a_{i1}, \ldots, a_{ik}). \end{aligned} \tag{2.28}$$

Nandi and Sen [22] established the minimum variance property of U_n (in a nonparametric setup), obtained a compact expression for the variance of U_n, and through detailed combinatorial arguments showed that as $N \to \infty$, $n \to \infty$ (but n/N need not go to a positive limit),

$$n^{1/2}[U_n - \Theta_N]/(1 - n/N)^{1/2} \quad \text{is asymptotically normal.} \tag{2.29}$$

If we define $g_1(x) = \mathrm{E}_{\mathcal{P}_n}\{g(X_1, \ldots, X_k) \mid X_1 = x\}$, and let

$$U_n^{(1)} = \frac{1}{n} \sum_{i=1}^{n} g_1(X_i), \quad n \geq 1, \tag{2.30}$$

then it follows from Nandi and Sen [22] that as $n \to \infty$,

$$n\mathrm{E}[U_n - \Theta_N - k\{U_n^{(1)} - \Theta_N\}]^2 \to 0. \tag{2.31}$$

On the other hand, $U_n^{(1)}$ is a linear statistic for which the Hájek [6] PCLT applies, so that (2.31) extends the PCLT for general U-statistics. For some related results, we may refer to Puri and Sen [30], Ch. 3. We may even extend the PCLT for SRSWOR in a more general setup as follows. Let $T_n = T(\boldsymbol{X}_n)$ be an arbitrary statistic, such that $\Theta_N = \mathrm{E}T_n$ exists and the second moment of T_n is also finite. Let us define for n $\leq N-1$,

$$\widehat{T}_n = \frac{N-1}{N-n}\sum_{n=1}^{n}\mathrm{E}\{T_n \mid X_i] - \frac{N(n-1)}{N-n}\Theta_N. \tag{2.32}$$

Then, it follows by some standard steps that

$$\begin{aligned}\mathrm{E}(T_n - \widehat{T}_n)^2 &= \mathrm{E}(T_n - \Theta_N)^2 + \mathrm{E}(\widehat{T}_n - \Theta_N)^2 - 2\mathrm{E}\{(T_n - \Theta_N)(\widehat{T}_n - \Theta_N)\} \quad (2.33)\\ &= \mathrm{E}(T_n - \Theta_N)^2 - \mathrm{E}(\widehat{T}_n - \Theta_N)^2,\end{aligned}$$

so that whenever

$$\lim_{n,N\to\infty} \mathrm{E}(\widehat{T}_n - \Theta_N)^2/\mathrm{E}(T_n - \Theta_N)^2 = 1, \tag{2.34}$$

the Hájek [10] projection applies to T_n, and as $\widehat{T}_n$ is a linear statistic, Hájek's [6] PCLT applies to $\widehat{T}_n$ under minimal regularity conditions, the PCLT holds for general T_n under (2.34) and the same regularity conditions (on T_n). In particular for U-statistics, (2.34) follows from Nandi and Sen [22]. It is clear from the above discussion that the impact of Hájek's [10] projection and quadratic mean approximation on PCLT in a SRSWOR setup goes far beyond linear statistics.

If we go back to the superpopulation model introduced just after (2.24) then Θ_N in (2.28) can be regarded as a U-statistic (say, U_N) based on a sample of size N from this superpopulation. This enables one to incorporate the reverse martingale property of U_n, $n \geq k$, for i.i.d. sampling, to conclude that for every $N(\geq n \geq k)$,

$$\{U_n - \Theta_N;\ k \leq n \leq N\} \quad \text{is a reverse martingale.} \tag{2.35}$$

Actually, in FFS, SRSWOR schemes, this reverse margingale characterization of U-statistics (due to Sen [36]) follows directly by using the permutation probability law $\mathcal{P}_N$, without necessarily appealing to any super-population structure. The past two decades have witnessed the phenomenal growth of research literature on asymptotics based on martingales and reverse martingales, and (2.35) provides the access to incorporating such asymptotics in SRSWOR schemes for a broad class of statistics. In particular, using the Chow-extension of the celebrated Hájek–Rényi [13] inequality, we obtain that for a nondecreasing sequence $\{c_n\}$ of positive numbers and for every $(k \leq)\ n \leq N,\ t > 0$,

$$\begin{aligned}&P\left\{\max_{n\leq m\leq N} c_m|U_m - \Theta_N| \geq t\right\} \quad (2.36)\\ &\leq t^{-2}\left\{c_n^2\mathrm{E}(U_n - \Theta_N)^2 + \sum_{m=n+1}^{N}(c_m^2 - c_{m-1}^2)\mathrm{E}(U_m - \Theta_N)^2\right\},\end{aligned}$$

where we know that

$$\mathrm{E}(U_m - \Theta_N)^2 = O\left((N-m)m^{-1}(N-1)^{-1}\right), \quad \forall\, m \geq k. \tag{2.37}$$

In particular, letting $c_m = c = 1$, $\forall\, m \geq k$ and $t = \epsilon > 0$, we obtain from (2.36) and (2.37) that for $n \in [k, N]$,

$$P\left\{\max_{n\leq m\leq N} |U_m - \Theta_N| \geq \epsilon\right\} \leq \epsilon^{-2}\mathrm{E}[U_n - \Theta_N]^2 = O\left(\epsilon^{-2}n^{-1}(1-n/N)\right). \quad (2.38)$$

This Kolmogorov-type maximal inequality (in SRSWOR) implies that whenever n $(< N)$ increases (without necessarily assuming that n/N converges to a positive limit),

$$\max_{n\leq m\leq N} |U_m - \theta_N| \to 0, \quad \text{in probability.} \quad (2.39)$$

This mode of convergence is stronger than the usual stochastic convergence result that $U_n - \Theta_N \to 0$, in probability, as n increases, and this result is referred to in the literature as the *strong convergence* in FPS (SRSSWOR). In sequential analysis relating to SRSWOR schemes, (2.39) is quite useful (see, for example, Sen [41]). We shall consider some allied results in the last section.

Hájek [9] considered an interesting Kolmogorov-type inequality for dependent summands (relating to FPS), and that is allied to (2.36). This inequality, discussed in detail in Chapter 5 of Hájek and Šidák [14], plays a vital role in nonparametrics asymptotics. Let $d_1, \ldots, d_N$ be a set of real numbers. $R_1, \ldots, R_n$ be defined as in (2.5) and let $\bar{d}_N = N^{-1}\sum_{i=1}^N d_i$, $D_N^2 = \sum_{i=1}^N (d_i - \bar{d}_N)^2$ and $D_N^* = \sum_{i=1}^N (d_i - \bar{d}_N)^4$. Then, for every $n(< N)$,

$$\mathrm{E}\left[\sum_{i=1}^n d_{Ri} - n\bar{d}_N\right]^4 = \frac{n(N-n)}{N^{[4]}}\{3(N-n-1)(n-1)D_N^4 + (N^2 - 6nN + 6n^2 + N)D_N^*\} \quad (2.40)$$

Using (2.40), Hájek [9] showed that for every $\epsilon > 0$, $n < N$,

$$\begin{aligned} &P\left\{\max_{1\leq k\leq n}\left|\sum_{i=1}^k d_{Ri} - k\bar{d}_N\right| > \epsilon D_N\right\} \quad (2.41)\\ &\leq \frac{n}{N}\left\{\max_{1\leq i\leq N}(d_i - \bar{d}_N)^2/D_N^2 + \frac{3n}{N}\right\}\epsilon^{-4}(1-\frac{n}{N})^{-3}(1+\eta_N), \end{aligned}$$

where $\eta_N \to 0$ as $N \to \infty$.

The main utility of (2.41) relates to the case where n/N is small, and for this the 4th moment in (2.40) is utilized. However, the 4th moment condition may not be necessary. Recall that $n^{-1}\sum_{i=1}^n d_{Ri}$ is a U-statistic (of degree 1), so that letting $c_n = n$, $n \geq 1$, we obtain from (2.36) that for every $n < N$, $\epsilon > 0$,

$$\begin{aligned} &P\left\{\max_{1\leq k\leq n}\left|\sum_{i=1}^k d_{Ri} - k\bar{d}_N\right| > \epsilon D_N\right\} \quad (2.42)\\ &\leq \left[\sum_{k=1}^{n-1}\frac{k}{k+1} + \frac{n(N-n)}{N}\right]\Big/\{\epsilon^2(N-1)\} \end{aligned}$$

$$< \left[\frac{n-1}{N-1} + \frac{n(N-n)}{N(N-1)} \right] \epsilon^{-2}.$$

This inequality (Sen [36]) has further been strengthened by incorporating higher order moments or moment generating function of U-statistics in a FPS setup. With due emphasis on FPS asymptotics, we may allow N to be large, and, as in Sen [37], consider the following.

Let us define $Y_N = \{Y_N(t),\ 0 \le t \le 1\}$ by letting

$$Y_N(t) = D_N^{-1} \sum_{i \le [Nt]} (d_{Ri} - \bar{d}_N), \quad 0 < t < 1, \tag{2.43}$$

where conventionally, $Y_N(t) = 0$, $\forall\ 0 \le t < 1/N$. Then,

$$Y_N \xrightarrow{D} W^o, \quad \text{in the } J_1\text{-topology on D[0,1]}, \tag{2.44}$$

where $W^o = \{W^o(t),\ 0 \le t \le 1\}$ is a standard Brownian bridge on [0,1]. A direct consequence of (2.44) is that for every α $(0 \le \alpha \le 1)$: $\lim_{N\to\infty} n/N = \alpha$, and $\epsilon > 0$,

$$\lim_{N\to\infty} P\left\{ \sup_{0\le t\le \alpha} |Y_N(t)| \ge \epsilon \right\} = P\left\{ \sup_{0\le t\le \alpha} |W^o(t)| \ge \epsilon \right\}. \tag{2.45}$$

Recall that by the Doob [3] construction, $\{(t+1)W^o(t/(t+1)) : t \ge 0\} \stackrel{\mathcal{D}}{=} \{W(t),\ t \ge 0\}$, where $W(t)$ is a standard Wiener process on $[0, \infty)$. As such, for every $\epsilon > 0$, $0 \le \alpha \le 1$, we have

$$\begin{aligned} &P\left\{ \sup_{0\le t\le \alpha} |W^o(t)| > \epsilon \right\} \\ &= P\left\{ \sup_{0\le u\le \alpha/(1-\alpha)} (1+u)^{-1}|W(u)| > \epsilon \right\}. \end{aligned} \tag{2.46}$$

One may then use the basic results of Anderson [1] to obtain an algebraic expression for (2.46) (albeit in an infinite series form). Alternatively, as in Sen [37], we may bound (2.46) from above by

$$4\left[1 - \Phi\left(\epsilon(\alpha^{-1}(1-\alpha))^{1/2}\right)\right], \tag{2.47}$$

where $\Phi(x)$ is the standard normal d. f. For small α, (2.47) provides a good approximation for (2.46). Whenever α is not so small, one may even use the Kolmogorov–Smirnov bound:

$$P\left\{ \sup_{0\le t\le 1} |W^o(t)| > \epsilon \right\} = 2\sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2\epsilon^2}, \tag{2.48}$$

so that the left hand side of (2.45) is bounded from above by (2.48). Since by the Mills ratio, $1 - \Phi(x) \le x^{-1}\phi(x)$, where $\phi(x) = \Phi'(x)$, comparing (2.47) and (2.48), we gather that for small values of α (i.e., $\epsilon/\sqrt{\alpha}$ large), (2.47) provides a better bound that (2.48).

The Hájek-type inequalities in FPS discussed above have found their ways to mingle with other applications as well. Again, martingale characterizations play a basic role in this context. Let us define the $\{d_n\}$ and the R_i as in before. Also, let $q = \{q(t): 0 \leq t \leq 1\}$ be a continuous, nonnegative, U-shaped and square integrable function inside (0,1). Then, for every N (≥ 1), $\lambda > 0$,

$$P\left\{\max_{1\leq k\leq N} q(k/N)\left|\sum_{i=1}^{k}(d_{Ri} - \bar{d}_N)\right| \geq D_N\lambda\right\} \leq \lambda^{-2}\int_0^1 q^2(t)\,dt. \tag{2.49}$$

The proof (Sen [39]) exploits martingale characterizations and the Hájek–Rényi [13] inequality. Some applications of these probability inequalities will be considered in the concluding section.

3. UPS ASYMPTOTICS

Hájek ([12], Ch.3) contains an elegant formulation of probability sampling covering EPS as well as UPS schemes. We define a sample s as an arbitrary subset of the population $S = \{1, \ldots, N\}$, so that there are 2^N subsets (including the empty set $\emptyset$ and the whole set S). A *Sampling design* is specified by a probability law

$$P = \{P(s);\ s \in S\}, \tag{3.1}$$

which characteristingly defines the *inclusion probabilities*:

$$\pi_{i_1\ldots i_l} = \sum_{\{s:i_1,\ldots,i_l\in s\}} P(s), \tag{3.2}$$

for $1 \leq i_1 < \cdots < i_l \leq N$; $l \geq 1$. The most important entities relate to the case of $l = 1$ and 2.

For SRSWOR, we have $P(s) = \binom{N}{n}^{-1}$ if size of $s = n$, and 0 otherwise. We define the *inclusion indicators* by

$$I_i(s) = 1 \text{ or } 0 \text{ according as } i \in s \text{ or not, } 1 \leq i \leq N. \tag{3.3}$$

In *Poisson Sampling* the indicators I_i are taken as independent r.v.'s. For any sequence of positive numbers $p_1, \ldots, p_N$: $\sum_{i=1}^N p_i = 1$, the corresponding Poisson sampling is defined by

$$P(s) = \prod_{i\in s} p_i \prod_{i\in S\setminus s}(1 - p_i). \tag{3.4}$$

The sample size of s, denoted by ($\#s$) is therefore a r.v., and

$$\mathrm{E}(\#s) = \sum_{i=1}^N p_i,\ \mathrm{Var}(\#s) = \sum_{i=1}^N p_i(1 - p_i); \tag{3.5}$$

$$\pi_i = p_i; \ \pi_{ij} = p_i p_j, \ i \neq j = 1, \ldots, N. \tag{3.6}$$

The Poisson Sampling leads to a unified way of presenting some other sampling methods including the rejective sampling and successive sampling schemes. Professor Hájek made outstanding contributions in this field too (see Hájek [8]).

Rejective Sampling (Hájek [8]) may be defined either as conditional Poisson Sampling or as conditional sampling with replacement. We let here

$$P(s) = \begin{cases} c^* \prod_{i \in s} \alpha_i, & \text{if } \#S = n; \\ 0, & \text{otherwise,} \end{cases} \tag{3.7}$$

where $\alpha_1, \ldots, \alpha_N$ are positive numbers and $\sum_{i=1}^N \alpha_i = 1$, $c^* > 0$. The probability α_i of selecting the unit i in individual draws has been termed by Hájek ([12], p. 67) the *drawing probability*. For some unified treatment of rejective sampling, we refer to Hájek ([12], Ch. 7). In the *Sampford–Durbin modification of rejective sampling* there is a two-phase scheme: The first unit is drawn with the drawing probabilities

$$\alpha_{i(1)} = n^{-1}\pi_i, \quad 1 \leq i \leq N, \tag{3.8}$$

while the remaining $n-1$ units are drawn with drawing probabilities

$$\alpha_{i(*)} = \lambda\pi_i(1-\pi_i)^{-1}, \quad 1 \leq i \leq N, \ \lambda > 0. \tag{3.9}$$

In this scheme, the inclusion probabilities are exactly equal to π_i if λ is so chosen.

Successive Sampling consists of a sequence of independent draws of one unit with some constant probabilities $\alpha_1, \ldots, \alpha_N$: $\sum_{i \leq N} \alpha_i = 1$. If a draw yields an unit already selected in an earlier draw it is ignored, and the sequence stops as soon as there are n distinct units in the sample s. The advantage of successive sampling is that the average number of draws may only moderately exceed the sample size n, and the disadvantage is the methodological complications that may generally arise when the sample size is not small. Rosén [33] has made a fundamental contribution to general asymptotics, and Hájek ([12]; Ch. 9) has a unified treatise of the same. Some martingale characterizations (viz., Sen [38], [40]) add more flexibilities to such asymptotic methods. We intend to present a broader review of these asymptotics. For a population $\boldsymbol{\Pi}_N = (a_1, \ldots, a_N)$ with drawing probabilities $\boldsymbol{P}_N = (p_1, \ldots, p_N)$, in a SSVPWOR (successive sampling with varying probabilities (without replacement)), let $\Delta(r,n) = P(r \in s)$, the probability that unit r is included in a sample of size n. Then the well known *Horvitz-Thompson* estimator of the population total ($t = \sum_{i \leq N} a_i$) is

$$\begin{aligned} HT_n &= \sum_{r=1}^N w_{nr} a_r / \Delta(r,n) \\ &= \sum_{r \in s} a_r / \Delta(r,n), \end{aligned} \tag{3.10}$$

where

$$w_{nr} = \begin{cases} 1, & r \in s \\ 0, & \text{otherwise, for } 1 \leq r \leq N. \end{cases} \tag{3.11}$$

The varying probability structure introduces additional complications in the study of asymptotics for such estimators. Rosén [32], [33] considered an alternative approach (via the *coupon collector's problem*) and presented deeper results. Let $\{J_k;\ k \geq 1\}$ be a sequence of i.i.d.r.v.'s where

$$P\{J_k = r\} = p_r, \quad 1 \leq r \leq N,\ \forall\, k \geq 1. \tag{3.12}$$

Let then

$$Y_{nk} = \begin{cases} a_{J_k}/\Delta(J_k, n), & \text{if } J_k \notin \{J_1, \ldots, J_{k-1}\} \\ 0, & \text{otherwise, for } k \geq 1; \end{cases} \tag{3.13}$$

$$\nu_k = \inf\{m(\geq k) : \text{number of distinct } J_1, \ldots, J_m = k\},\ k \geq 1. \tag{3.14}$$

Then Rosén showed that for every $n \geq 1$,

$$HT_n \stackrel{D}{=} \sum_{k=1}^{n} Y_{n\nu_k} = Bn\nu_n, \tag{3.15}$$

where for each $m \geq 1$, B_{nm} is the *bonus sum* at the mth stage in a coupon collector's problem with the set $\{a^*_{nr},\ p_r;\ 1 \leq r \leq n\}$, with $a^*_{nr} = a_r/\Delta(r,n)$, for $r = 1, \ldots, N$. Thus, the asymptotic behavior of *randomly stopped* bonus sums provides the access to the general asymptotics for HT_n and other related estimators. Rosen's formulation rests on sophisticated nonstandard mathematical analysis, and some simplifications and generalizations based on martingale approximations are due to Sen [38]. By reference to a coupon collector's model, we consider a sequence $\{\Omega_N\}$ where for each N, $\Omega_N = \{(a_N(1), p_1(1)), \ldots, (a_N(N), p_N(N))\}$ and the nonnegative $p_N(s)$ add upto 1. Define the J_{Nk} as in (3.12) and the Y_{Nk} as in (3.13) (with the $a_{J\,k}/\Delta(J_k, n)$ being replaced by $a_N(J_{Nk})$). Let then

$$Z_{Nn} = \sum_{k \leq n} Y_{nk}, \quad n \geq 1, \quad Z_{N0} = 0. \tag{3.16}$$

Then Z_{Nn} is the *bonus sum* after n coupons in the collector's situation Ω_N. If the $a_N(k)$ are all nonnegative then Z_{Nn} is $\uparrow$ in n, so that if we let $U_n = \{U_N(t),\ t \in \mathbb{R}^+\}$, where

$$U_N(t) = \min\{k \colon Z_{Nk} \geq t\},\ t \geq 0, \tag{3.17}$$

then $U_N(t)$ is the *waiting time* to obtain the bonus sum t in the coupon collector's situation Ω_N. Let then

$$Q_{Nn} = \sum_{s \leq N} a_N(s)\,[1 - \exp(-np_N(s))], \tag{3.18}$$

$$\begin{aligned} d^2_{Nn} &= \sum_{s \leq N} a^2_N(s) e^{-np_N(s)} \left(1 - e^{-np_N(s)}\right) \\ &\quad - \left(\sum_{s \leq N} a_N(s) p_N(s) e^{-np_N(s)}\right)^2, \ n \geq 1 \end{aligned} \tag{3.19}$$

Rosén [31], [32] and Holst [17], among others, showed that under certain regularity conditions, as $n \to \infty$,

$$(Z_{Nn} - Q_{Nn})/d_{Nn} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{3.20}$$

Further, using the identity that for every $x, t > 0$, $P\{U_N(t) > x\} = P\{Z_{N[x]} < t\}$, one can derive the asymptotic normality of the normalized version of $U_N(t)$, as $t \to \infty$. Note that the $\{Z_{Nn}, n \geq 0\}$, $N \geq N_o$ may not generally be a martingale array, and hence, the proof of (3.20) rests on some sophisticated analysis. Martingale approximations provide simpler solutions. Let $Q_{Nk} = p_N(J_{Nk})$, $k \geq 1$, and let

$$X_{Nk}^{(n)} = Y_{Nk}(1+Q_{Nk})^{k-1}e^{-nQ_{NR}}, \ k \geq 1, \ X_{N0}^{(n)} = 0; \tag{3.21}$$

$$\begin{aligned} \widetilde{X}_{Nk}^{(n)} &= X_{Nk}^{(n)} - \mathrm{E}(X_{Nk}^{(n)}|\mathcal{B}_{Nk-1}) \\ &= X_{Nk}^{(n)} - \zeta_{Nk}^{(n)} + \sum_{s=0}^{k-1} X_{Ns}^{(n)} Q_{Ns}(1+Q_{Ns})^{k-s}, \end{aligned} \tag{3.22}$$

where $\mathcal{B}_{Nk} = \mathcal{B}(J_{Nj},\ j \leq k)$, $k \geq 0$, and

$$\zeta_{Nk}^{(n)} = \sum_{s \leq N} a_N(s)p_N(s)\exp\left(-np_N(s)\right)[1+p_N(s)]^{k-1}, \ k \geq 1; \tag{3.23}$$

and conventionally, we let $\widetilde{X}_{N0}^{(n)} = 0 = \zeta_{N0}^{(n)}$, $\forall\ n \geq 1$. Also, let

$$\widetilde{S}_{Nk}^{(n)} = \sum_{j \leq k} \widetilde{X}_{Nn}^{(n)} \quad \text{and} \quad \widetilde{\zeta}_{Nk}^{(n)} = \sum_{j \leq k} \zeta_{Nk}^{(n)}, \ k \geq 0. \tag{3.24}$$

Then, it can be shown that for every $\epsilon > 0$, as $N \to \infty$,

$$P\{d_{Nn}^{-1}|\widetilde{S}_{Nn}^{(n)} - Z_{Nn} + Q_{Nn}| > \epsilon\} = 0, \tag{3.25}$$

while by construction, for every (N, n), $\{\widetilde{S}_{Nk}^{(n)}, B_{Nk};\ k \geq 1\}$ is a martingale array, so that by martingale CLT,

$$d_{Nn}^{-1}\widetilde{S}_{Nn}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{3.26}$$

and (3.20) follows directly from (3.25) and (3.26). As a matter of fact, the asymptotic normality in (3.20) or (3.26) extends directly to suitable weak invariance principles for bonus sums and waiting times; these are presented in Sen [38]. The coupon collector's problem has important application in the generalized occupancy problems with access to the classical mark-capture-release-recapture methodology. The number of catches needed to obtain exactly n distinct units relates essentially to the waiting time in the coupon collector's problem. In UPS asymptotics these play a fundamental role. A renewal theorem in this context is due to Sen [42].

Sub-sampling (or *multi-stage sampling*) schemes are quite popular in practice. Here the primary units (say, N) of a population are composed of a number of smaller (sub-) units. Thus, it is customary to select first a sample of n primary units (out of N), and then, for each of the selected primary units, to draw a sample of subunits. At each stage, one may use EPS or UPS, and, as such, rejective sampling and successive sampling schemes are all relevant. The asymptotic distribution theory of estimators in successive sub-sampling with varying probabilities without replacement has been studied by Sen [40]. This was accomplished through an invariance principle for an extended coupon collector's problem wherein the basic martingale approach in Sen [38] has been exploited fully. These results provide a good theoretical justification for general asymptotics which have occasionally been adopted in survey sampling without proper motivation or analytical considerations. To conclude this Section, I may remark that the Hájek asymptotics in UPS opened the doors for rigorous theoretical treatise (often, in contrast to other heuristics in FPS), and his basic ideas also paved the way for martingale characterizations which, in turn, provided simpler proofs of many useful asymptotic results.

4. INVARIANCE PRINCIPLES IN FPS

In FPS, traditionally, the (asymptotic) normality of (the standardized form of) estimators is taken for granted and the prime emphasis is laid down on the estimation of its sampling variance, so that large sample confidence intervals and/or hypothesis testing can be validly worked out. Nevertheless, this asymptotic normality itself remains as a vital issue of serious study. The situation may become much more complex when the sample size may itself be a random variable. For example, in a stratified sampling scheme (EPS), if the Neyman allocation is based on the sample estimates of the within stratum variances, the resulting strata sample sizes are all random variables. Thus, there may be a need for extending the CLT's in FPS (EPS/UPS) to the cases where the sample size may not be prefixed. In fact, in inverse sampling schemes the sample size is typically random and is governed by a well defined stopping rule. In two (or multi-) stage sampling procedures, and more generally, in sequential ones, stochastic sample sizes are quite commonly encountered. In the literature on standard (parametric) asymptotic theory, the classical Anscombe theorem extends CLT's to random sample sizes. Another way of dealing with this problem is to formulating suitable invariance principles (weak as well as stronger ones) which yield the Anscombe-condition as a by-product, and, in addition, provide deeper asymptotic results.

Hájek [9] (and later reported in Hájek and Šidák [14]) provided an excellent introduction to such weak (and almost sure) convergence results for appropriate rank-processes, and opened the doors for a new approach to the deeper asymptotics in nonparametrics. Dealing with rank statistics, the situation is more complex than the case of sums of independent random variables, and Hájek [9] had to import some finer probability inequalities to verify the "compactness" (or tightness) part of the related weak convergence results. In particular, the inequality in (2.41) plays a key role in this context. Ranks are not stochastically independent, and hence, in PCLT's, the "independent increment" clause may not be generally true. The weak convergence result in (2.44) is based on a reversed martingale characterization of

U-statistics in FPS (Sen [36], [37]), and for such (reversed) martingale sequences, convergence of finite dimensional distributions implies the tightness condition (Sen [41], Ch.2)). Similar martingale characterizations have been worked out for various rank statistics, leading to appropriate invariance principles for then, and these are presented in a unified manner in Sen [41]. Because of the intrinsic connection between PCLT's and FPS asymptotics, explained in Section 2, it is quite intuitive to note that such invariance principles pertain to FPS as well.

Resampling plans (viz., *jackknife* and *bootstrap* methods) have gained a lot of scope for practical applications during the past fifteen years. Although most of these developments are related to SRSWR plans, there are some interesting developments relating to FPS as well. (The genesis of jackknifing lies in FPS). Invariance principles for jackknifing U-statistics for finite population sampling were developed by Majumdar and Sen [20], and applications to FPS schemes were also discussed. In this context too, reversed martingale characterizations in FPS play the key role. For bootstrap procedures, in FPS, the weak convergence of the normalized form of the bootstrap empirical distribution provides the desired key.

Invariance principles have also been developed for UPS schemes. For example, the asymptotic normality result in (3.25) has been strengthened to an invariance principle (Sen [38]) through some related developments for the coupon collector's bonus sum and waiting time problems. Again, in this context, martingale approximations provide useful tools. Sen [40] contains an invariance principle pertaining successive subsampling schemes discussed in the preceding section.

5. CONCLUDING REMARKS

Asymptotics in FPS, whether be in EPS or UPS schemes, follow a somewhat different track than in SRSWR. Complications may arise due to lack of independence and, probably, unequal drawing probabilities, and also there may be other constraints in the sampling design contributing more towards this complexity. Hájek's ingenuity in providing a sampling design in a deterministic probability modeling has indeed led to subsequent developments. In this respect, he did not hesitate to borrow tools from pure probability theory and stochastic processes to nonparametrics and general asymptotics, and the endproduct in a solid foundation of general asymptotics in FPS. In this respect, not only he was instrumental in providing the basic research work but also successfully developed the Prague School which has made genuine contributions in this field. During the last five years of his life, I had a good opportunity to know him not only professionally but also as a friend, colleague and mentor as well. With my dual interest in nonparametrics and FPS, I found in him an ideal person to follow the footsteps. It's difficult, but I learnt a lot. In particular, the role of martingale theory in nonparametrics and FPS, we have tried to explore fully during the past two decades, might not have come out in the present form without the foresight of late Professor Jaroslav Hájek. I therefore take this opportunity to pay my humble tribute and homage to this most pioneering researcher in asymptotic methods in our time.

(Received October 26, 1994.)

REFERENCES

[1] T. W. Anderson: A modification of the sequential probability ratio test to reduce the sample size. Ann. Math. Statist. *31* (1960), 165–197.

[2] S. K. Chatterjee and P. K. Sen: Nonparametric tests for the bivariate two-sample location problem. Calcutta Statist. Assoc. Bull. *13* (1964), 18–58.

[3] J. L. Doob: Heuristic approach to the Kolmogorov–Smirnov theorems. Ann. Math. Statist. *20* (1949), 393–403.

[4] J. Dupačová: A note on rejective sampling. In: Contributions to Statistics (J. Hájek Memorial Volume), (J. Jurečková, ed.), Academia, Prague and Reidel, Dordrecht 1979, pp. 72–78.

[5] J. Hájek: Optimum strategy and other problems in probability sampling. Časopis pro pěstování mat. *84* (1959), 387–423.

[6] J. Hájek: Some extensions of the Wald–Wolfowitz–Noether theorem. Ann. Math. Statist. *32* (1961), 506–523.

[7] J. Hájek: Asymptotically most powerful rank order tests. Ann. Math. Statist. *33* (1962), 1124–1147.

[8] J. Hájek: Asymptotic theory of rejective sampling with varying probabilities from a finite population. Ann. Math. Statist. *35* (1964), 1491–1523.

[9] J. Hájek: Extensions of the Kolmogorov–Smirnov tests to regression alternatives. In: Proc. Bernoulli–Bayes–Laplace Seminar, Berkeley (L. LeCam, ed.), Univ. of Calif. Press 1965, pp. 45–60.

[10] J. Hájek: Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Statist. *39* (1968), 325–346.

[11] J. Hájek: Asymptotic theory of sampling with varying probabilities without replacement. In: Proc. Prague Symp. Asymp. Statist. 1973, pp. 127–138.

[12] J. Hájek: Sampling from a Finite Population. Marcel Dekker, New York 1981.

[13] J. Hájek and A. Rényi: Generalization of an inequality of Kolmogorov. Acta. Math. Acad. Sci. Hung. *6* (1955), 281.

[14] J. Hájek and Z. Šidák: Theory of Rank Tests. Academia, Prague 1967.

[15] W. Hoeffding: A class of statistics with asymptotically normal distribution. Ann. Math. Statist. *19* (1948), 293–325.

[16] W. Hoeffding: A combinatorial central limit theorem. Ann Math. Statist. *22* (1951), 558–566.

[17] L. Holst: Asymptotic normality in a generalized occupancy problem. Z. Wahrsch. verw. Geb. *21* (1972), 109–120.

[18] S. Karlin: Inequalities for symmetric sampling plans I. Ann. Statist. *2* (1974), 1065–1094.

[19] W. G. Madow: On the limiting distributions of estimates based on samples from finite universes. Ann. Math. Statist. *19* (1948), 535–545.

[20] H. Majumdar and P. K. Sen: Invariance principles for jackknifing U-statistics for finite population sampling and some applications. Comm. Statist. *A7* (1978), 1007–1025.

[21] M. Motoo: On the Hoeffding's combinatorial central limit theorem. Ann. Inst. Statist. Math. *8* (1957), 145–154.

[22] H. K. Nandi and P. K. Sen: On the properties of U-statistics when the observations are not independent. Part II. Unbiased estimation of the parameters of a finite population. Calcutta Statist. Assoc. Bull. *12* (1963), 125–143.

[23] G. E. Noether: On a theorem by Wald and Wolfowitz. Ann. Math. Statist. *20* (1949), 455–458.

[24] Z. Prášková: Rate of convergence for simple estimate in the rejective sampling. In: Probability and Statistical Inference (W. Grossmann et al., eds.), Reidel, Dordrecht 1982, pp. 307–317.

[25] Z. Prášková: On the probability of very large deviations for the rejective sampling. In: Asymptotic Statistics 2, Proc. 3rd Prague Symp. on Asymptotic Statistics (P. Mandl and M. Hušková, eds.), North-Holland, Amsterdam 1984, pp. 387–389.

[26] Z. Prášková: On the rate of convergence in Sampford-Durbin sampling from a finite population. Statist. Decisions *2* (1984), 339–350.

[27] Z. Prášková: On the convergence to the Poisson distribution in rejective sampling from a finite population. In: Probability and Mathematical Statistics with Applications (Visegrád 1985, W. Grossmann et al. eds.), Reidel, Dordrecht 1985, pp. 285–294.

[28] Z. Prášková: Sampling from a finite set of random variables: The Berry–Eseén bound for the studentized mean. In: Proc. 4th Prague Symp. on Asymptotic Statistics (P. Mandl and M. Hušková, eds.), Charles University, Prague 1988, pp. 67–81.

[29] Z. Prášková: A note on Sampford–Durbin sampling. Comment. Math. Univ. Carolinae *31* (1990), 367–372.

[30] M. L. Puri and P. K. Sen: Nonparametric Methods in Multivariate Analysis. Wiley, New York 1971.

[31] B. Rosén: Asymptotic normality in a coupon collector's problem. Zeit. Wahrsch. verw. Geb. *13* (1969), 256–279.

[32] B. Rosén: On the coupon collectors waiting time. Ann. Math. Statist. *41* (1970), 1952–1969.

[33] B. Rosén: Asymptotic theory for successive sampling with varying probabilities without replacement, I, II. Ann. Math. Statist. *43* (1972), 373–397, 748–776.

[34] M. R. Sampford: On sampling without replacement with unequal probabilities of selection. Biometrika *54* (1967), 499–513.

[35] M. R. Sampford: A comparison of some possible methods of sampling from smallish populations, with units of unequal size. In: New Developments in Survey Sampling (N. L. Johnson and H. Smith, eds.), Wiley, New York 1969.

[36] P. K. Sen: The Hájek–Rényi inequality for sampling from a finite population. Sankhyā *A 32* (1970), 181–188.

[37] P. K. Sen: Finite population sampling and weak convergence to a Brownian bridge. Sankyā *A 34* (1972), 85–90.

[38] P. K. Sen: Invariance principles for the coupon collector's problem: A martingale approach. Ann. Statist. *7* (1979), 372–380.

[39] P. K. Sen: Weak convergence of some quantile processes arising in progressively censored tests. Ann. Statist. *7* (1979), 414–431.

[40] P. K. Sen: Limit theorems for an extended coupon collector's problem and for successive sampling with varying probabilities. Calcutta Statist. Assoc. Bull. *29* (1980), 113–132.

[41] P. K. Sen: Sequential Nonparametrics: Invariance Principles and Statistical Inference. Wiley, New York 1981.

[42] P. K. Sen: A renewal theorem for an urn model. Ann. Probab. *10* (1982), 838–843.

[43] P. K. Sen: On the asymptotic normality in sequential sampling tagging. Sankhyā *A 44* (1982), 352–363.

[44] P. K. Sen: On permutational central limit theorems for general multivariate linear rank statistics. Sankhyā *A 45* (1983), 141–149.

[45] P. K. Sen: Asymptotics in finite population sampling. In: Handbook of Statistics, Vol. 6: Sampling (P. R. Krishnaiah and C. R. Rao, eds.), North Holland, Amsterdam 1988, pp. 291–331.

[46] P. K. Sen and M. Ghosh: On bounded length sequential confidence intervals based on one-sample rank order statistics. Ann. Math. Statist. *42* (1971), 189–203.

[47] P. K. Sen and M. Ghosh: On strong convergence of regression rank statistics. Sankhyā *A 34* (1972), 335–348.

[48] R. J. Serfling: Probability inequalities for the sum sampling without replacement. Ann. Statist. *2* (1974), 39–48.
[49] J. Á. Víšek: Asymptotic distribution of sample estimate for rejective, Sampford and successive sampling. In: Contributions to Statistics: Jaroslav Hájek Memorial Volume (J. Jurečková, ed.), Academia, Prague and Reidel, Dordrecht 1979, pp. 363–376.
[50] A. Wald and J. Wolfowitz: Statistical tests based on permutations of the observations. Ann Math. Statist. *15* (1974), 358–372.

Prof. Dr. Pranab K. Sen, Departments of Biostatistics and Statistics, University of North Carolina, Chapel Hill, NC 27599–7400. U. S. A.