Smallest-fit selection of random sizes under a sum constraint: weak convergence and moment comparisons

Frans A. Boshuizen *
Econometric Institute, Erasmus Universiteit Rotterdam, The Netherlands

Robert P. Kertz †
School of Mathematics, Georgia Institute of Technology Atlanta, Georgia

April 28, 1995

Abstract

Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables with distribution function $F(x)$, and let $X_{1,n}, \ldots, X_{n,n}$ be the sequence of order statistics of $X_1, \ldots, X_n$. For a sequence $(c_n)_{n\geq 1}$ of positive constants the smallest fit off-line counting random variable is defined by $N^e(c_n) := \max\{j \leq n : X_{1,n} + \cdots + X_{j,n} \leq c_n\}$. In this paper, an asymptotic joint distributional comparison is given between the off-line count $N^e(c_n)$ and on-line counts N_n^τ for 'good' sequential (on-line) policies τ satisfying the sum constraint $\sum_{j\geq 1} X_{\tau_j} I(\tau_j \leq n) \leq c_n$. Specifically, for such policies τ, under appropriate conditions on the distribution function $F(x)$ and the constants $(c_n)_{n\geq 1}$, we find sequences of positive constants $(B_n)_{n\geq 1}$, $(\Delta_n)_{n\geq 1}$ and $(\Delta'_n)_{n\geq 1}$ such that

$$\left(\Delta_n \left(\frac{N^e(c_n)}{B_n} - 1\right), \Delta'_n \left(\frac{N_n^\tau}{B_n} - 1\right)\right) \Rightarrow (W, W') \text{ as } n \to \infty,$$

for some nondegenerate r.v.'s W and W'.
Abbreviated Title. Joint convergence of smallest-fit counts

1 Introduction

Let $X_1, X_2, \cdots$ be a sequence of positive independent, identically distributed random variables (i.i.d. r.v.'s) with distribution function $F(x)$ and denote the order statistics of $X_1, \ldots, X_n$ by $X_{1,n} \leq \ldots \leq X_{n,n}$.

*Research partly supported by a Fulbright Grant
†Corresponding author; Research partly supported by National Science Foundation Grant DMS 92-09586
AMS 1991 subject classification. Primary 60G40, 60G70; secondary 60F05, 60F15, 60G55, 90B50.
Key words and phrases. Joint convergence in distribution, extreme sums, order statistics, threshold strategies, counting r.v.'s, Poisson random measures, continuous mapping principle, strong approximations, Brownian bridge, on-line vs. off-line strategies

Let $(c_n)_{n\geq 1}$ be a sequence of positive numbers. The *smallest fit counting random variable* $N^e(c_n) = N_n^e(c_n)$ is defined by

$$N^e(c_n) := \max\{j : 1 \leq j \leq n \text{ and } X_{1,n} + \cdots + X_{j,n} \leq c_n\} \tag{1}$$

if this set is nonempty and $= 0$ otherwise. In a bin-packing context, the r.v. $N^e(c_n)$ is the maximal number of objects that can be packed into a bin of capacity c_n, chosen from n objects with random sizes sampled from the distribution F. This is the largest count obtainable by a 'prophet' or individual using an off-line strategy, that is, a strategy which uses knowledge of all the sizes in the sample, without any order restrictions on the sample values. This off-line, smallest-fit strategy is denoted by τ^e.

Define a *policy* τ to be any sequence of stopping times $(\tau_j)_{j\geq 1}$ with respect to $X_1, X_2, \cdots$ with $1 \leq \tau_1 < \tau_2 < \cdots$. For n objects, the counting r.v. associated with policy τ is defined by $N_n^\tau := \sum_{j\geq 1} I(\tau_j \leq n)$. For any policy $\tau = (\tau_j)_{j\geq 1}$ satisfying the sum constraint $\sum_{j\geq 1} X_{\tau_j} I(\tau_j \leq n) \leq c_n$, the r.v. N_n^τ has interpretation in a bin-packing context as the number of objects that are packed into a bin of capacity c_n, chosen sequentially under policy τ, without recall, as n objects with sizes sampled from distribution F appear in a given order. This N_n^τ is an 'on-line' count, a count under an on-line policy.

In this paper, a quantitative comparison is made between the off-line count $N^e(c_n)$ and on-line count N_n^τ for 'good' policies τ satisfying the sum constraint. There are several possible criteria for comparison of such counts. Coffman, Flatto and Weber [5] have used an expectation-based comparison. Under the hypotheses that $c_n \equiv c$ and $F(x)$ is continuous, strictly increasing on the support of F, and $F(x) \approx Ax^\alpha$ as $x \to 0$, for $A, \alpha > 0$, they used a 'Chernoff-estimates' based approach to show that $\lim_{n\to\infty} EN^e(c_n)/EN_n^*(c_n) = 1$. Here, for n objects and capacity c_n, an *optimal on-line counting r.v.* $N_n^*(c_n)$ is any r.v. satisfying

$$EN_n^*(c_n) = \sup\left\{EN_n^\tau : \tau \text{ is any policy such that } \sum_{j\geq 1} X_{\tau_j} I(\tau_j \leq n) \leq c_n\right\}. \tag{2}$$

As in Coffman et al. [5], one can observe that there exist optimal (on-line) policies τ^*, i.e. $N_n^*(c_n) = N_n^{\tau^*}$, by using a stochastic dynamic programming argument. Under weaker assumptions on $(c_n)_{n\geq 1}$ and F, Bruss and Robertson [2] used the same techniques to obtain asymptotic behavior of $EN^e(c_n)$, and Rhee and Talagrand [14] used a direct inequality based approach to obtain asymptotic behavior of $EN_n^*(c_n)$.

To fix ideas, consider the following comparison. A manufacturer anticipates production of an amount c_n of his product during the next business period. At the beginning of the period, the manufacturer knows there will be n customers (one order per customer) during the period. The manufacturer can not fill partial orders. The company requires the manufacturer to fill orders in such a way that the total customer base for the product is maximized. If the manufacturer has knowledge of the sizes of all customer orders at the beginning of the business period, then he will service smallest orders first and thereby fill $N^e(c_n)$ orders. If the manufacturer does not know the sizes of all customer orders at the beginning of the business period, but sees customer orders one-by-one as they arrive, and must either accept or reject each order immediately, without recall, then he must establish some criteria and choose a 'good' policy for handling orders under this criteria. Under an expectation based criteria, a best off-line policy would yield $EN^e(c_n)$ and a best on-line policy yields $EN_n^*(c_n)$, and the manufacturer wishes to compare $EN^e(c_n)$ and

$EN_n^*(c_n)$. As discussed in [5] and [14], an optimal policy τ^* for this expectation-based criteria can be described through the stochastic dynamic programming approach through certain functions $\kappa_j(x)$, $j \geq 1$, as the policy $\tau^* = (\tau_j^*)_{j\geq 1}$ with $\tau_1^* = \min\{1 \leq j \leq n : X_j \leq \kappa_{n-j}(c_n)\}$ if this set is nonempty and $= \infty$ otherwise, and for $k = 2, 3, \cdots$, $\tau_k^* = \min\{1 \leq j \leq n : \tau_{k-1}^* < j \leq n : X_j \leq \kappa_{n-j}(c_n - \sum_{i=1}^{k-1} X_{\tau_i^*})\}$ if $\tau_{k-1}^* < \infty$ and this set is nonempty and $= \infty$ otherwise (interpret $\kappa_0(x) \equiv 0$). However, identification of the functions $(\kappa_j(x))_{j\geq 1}$ and direct analysis using this policy τ^* is difficult, as mentioned in [5] and [14]; and one seeks simpler forms of policies under which analysis and implementation are possible.

One such policy, the *threshold policy t_n with horizon n*, used by Coffman, Flatto and Weber [5], is described by the following procedure. Let $(\varepsilon_n)_{n\geq 1}$ and $(c_n)_{n\geq 1}$ be sequences of positive constants. The ε_n's are used as thresholds and the c_n's are used as capacities. For n objects, accept only objects of size $\leq \varepsilon_n$ and select these only when the sum of the objects selected is $\leq c_n$. Specifically, t_n is the sequence of stopping times $(\tau_j)_{j\geq 1}$ defined by $\tau_1 = \min\{1 \leq j \leq n : X_j \leq \varepsilon_n \text{ and } X_j \leq c_n\}$ if this set is nonempty and $= \infty$ otherwise, and for $k = 2, 3, \cdots$, $\tau_k = \min\{\tau_{k-1} < j \leq n : X_j \leq \varepsilon_n \text{ and } X_{\tau_1} + \cdots + X_{\tau_{k-1}} + X_j \leq c_n\}$ if $\tau_{k-1} < \infty$ and this set is nonempty, and $= \infty$ otherwise. The counting r.v. associated with this policy, denoted $N^{t_n}(c_n) = N^{t_n}(\varepsilon_n, c_n) := N_n^{t_n}$, was shown in [5] to satisfy $\lim_{c'\uparrow c} \liminf_{n\to\infty} EN_n^{t_n}(c')/EN_n^e(c) = 1$ under the specific hypotheses mentioned above. Another such policy, the *stopped threshold policy s_n with horizon n*, used by Rhee and Talagrand [14], appears to be a simple, 'good' policy that is more accessible for analysis than t_n. This policy s_n uses the following procedure: accept objects only of size $\leq \varepsilon_n$ and select these only when the sum of the sizes of objects selected is $\leq c_n$; if the sum of the sizes $\leq \varepsilon_n$ up to the present exceeds c_n, then the present and all future objects are rejected. Specifically s_n is the sequence of stopping times $(\tau_j)_{j\geq 1}$ defined by $\tau_1 = \min\{1 \leq j \leq n : X_j \leq \varepsilon_n \text{ and } X_j \leq c_n\}$ if this set is nonempty and $= \infty$ otherwise, and for $k = 2, 3, \cdots$, $\tau_k = \min\{\tau_{k-1} < j \leq n : X_j \leq \varepsilon_n \text{ and } \sum_{i=1}^{j} X_i I(X_i \leq \varepsilon_n) \leq c_n\}$ if $\tau_{k-1} < \infty$ and this set is nonempty, and $= \infty$ otherwise. The counting r.v. associated with this policy is denoted $N^{s_n}(c_n) = N^{s_n}(\varepsilon_n, c_n) := N_n^{s_n}$. Note that $N^{s_n}(c_n)$ can be written as

$$N^{s_n}(c_n) = \sum_{i=1}^{\nu_n(c_n)} I(X_i \leq \varepsilon_n) \tag{3}$$

where $\nu_n(c_n) = \max\{j : 1 \leq j \leq n \text{ and } \sum_{i=1}^{j} X_i I(X_i \leq \varepsilon_n) \leq c_n\}$ if this set is nonempty and $= 0$ otherwise (interpret $\sum_{i=1}^{0}$ as 0). Under appropriate assumptions on the left tail of F and on the capacities $(c_n)_{n\geq 1}$, more general than those of [5], one can use results in [14] to show $\lim_{n\to\infty} EN^{s_n}(c_n)/EN^e(c_n) = 1$ and give some indication of the rate of this convergence (see Corollary 3.6 in this paper).

In this paper, the focus is on the asymptotic joint distributional comparisons of the off-line count $N^e(c_n)$ and on-line counts N_n^τ for 'good' policies τ satisfying the sum constraint. Specifically, a sequence of policies $(\tau^n)_{n\geq 1}$ satisfying the sum constraints $\sum_{j\geq 1} X_{\tau_j^n} I(\tau_j^n \leq n) \leq c_n$ for $n \geq 1$ is said to be a **consistent approximator** of the off-line smallest-fit strategy τ^e if there exists positive constants $(B_n)_{n\geq 1}$, $(\Delta_n)_{n\geq 1}$ and $(\Delta_n')_{n\geq 1}$ for which

$$\left(\Delta_n \left(\frac{N^e(c_n)}{B_n} - 1\right), \Delta_n' \left(\frac{N_n^{\tau^n}}{B_n} - 1\right)\right) \Rightarrow (W, W') \text{ as } n \to \infty, \tag{4}$$

for some nondegenerate r.v.'s W and W' (the convergence here, denoted by '$\Rightarrow$', is convergence in distribution). In Theorems 2.1, 2.5, 3.1 and 4.1, it is shown that the sequence of stopped threshold policies $(s_n)_{n\geq 1}$ is a consistent approximator of τ^e, for appropriate assumptions on capacities $(c_n)_{n\geq 1}$ and distribution F, and for appropriately chosen thresholds $(\varepsilon_n)_{n\geq 1}$. Consequences of these consistent approximator results include additional distributional comparisons and expectation-based comparisons of $N^e(c_n)$ and $N^{s_n}(c_n)$, in Corollaries 2.3, 2.6, 3.5, 3.6, 4.2 and 4.3.

In Sections 2, 3 and 4, joint weak convergence results are proved for $N^e(c_n)$, the smallest fit counting r.v., and $N^{s_n}(c_n)$, the counting r.v. associated with stopped threshold policy s_n, as the sample size grows to infinity. The analysis divides naturally into three subcases: (i) the counting r.v.'s have finite limits; (ii) the counting r.v.'s have infinite limits, and the numbers selected are small compared to the sample sizes; and (iii) the counting r.v.'s have infinite limits and the numbers selected are proportional to the sample sizes. The three subcases are treated respectively in Sections 2, 3 and 4. The respective techniques used in the three cases are (i) convergence of point processes to Poisson random measures and continuous mapping theorems; (ii) strong approximation results of Csörgő, Mason, Haeusler, Horváth, et al.; and (iii) results from renewal theory. In Sections 2 and 3, for the main results it is assumed that distribution function $F(x)$ is in the domain of attraction for minima for one of the extreme value d.f.'s (see the Appendix for properties of these classes of d.f.'s). In all cases it is assumed that $l_F \geq 0$ with $F(l_F) = 0$, for left-end point l_F of the support of $F(x)$. (So the case of $F(x)$ in Case II for minima does not arise here, since in this case $l_F = -\infty$.)

Throughout the paper the following notation is used. For a nondecreasing function $h(x)$ defined on a subset S of $\mathbb{R}$, the left-continuous inverse of h is defined by $h^{\leftarrow}(s) = \inf\{x \in S : h(x) \geq s\}$. For a distribution function F, $l_F = \inf\{x : F(x) > 0\}$ denotes the left end point of the support of F. The notations $o_P(1)$ and $O_P(1)$ are used to denote sequences of random variables which are respectively converging to zero in probability and bounded above and below by a finite constant uniformly for all n large. For two r.v.'s X and Y, $X \stackrel{\mathrm{d}}{=} Y$ if the distributions of X and Y are the same. For a real number x, $[x]$ denotes the largest integer $\leq x$, and $\lceil x \rceil$ denotes the smallest integer $\geq x$, and for two sequences $(m_n)_{n\geq 1}$ and $(l_n)_{n\geq 1}$ we write $m_n \approx l_n$ if $\lim_{n\to\infty} m_n/l_n = 1$.

2 Finite Limit of Numbers Selected

In this section, results are given on the weak convergence of the counting r.v.'s $N^e(c_n)$ and $N^{s_n}(c_n)$ to nondegenerate r.v.'s as the sample size $n \to \infty$, without normalizations. The setting is that of classical extreme value theory, and the i.i.d. r.v.'s $X_1, X_2, \cdots$ have d.f. $F(x)$ in a domain of attraction for minima, of Case I or III type. Pertinent definitions and properties of these r.v.'s are recalled in the Appendix. In particular, since joint convergence results for the counting r.v.'s are desired, techniques of proof emphasize use of the approach based on convergence of appropriate point processes to Poisson random measure and on continuous mapping theorems. Specifically, recall that for d.f.'s $F(x)$ in the domain of attraction for minima, there exist constants $a_n > 0$ and b_n, $n = 1, 2, \cdots$, for which

> point process $\xi_n = \sum_{k=1}^{n} \epsilon_{(k/n, a_n(X_k - b_n))}$ converges weakly to Poisson random measure ξ on $[0,1] \times E$, with intensity measure $dt \times d\nu$, where ξ may be assumed to have representation $\xi = \sum_{j\geq 1} \epsilon_{(\tau_j, W_j)}$ and $W_1 < W_2 < \cdots$ in E, with probability one, and $a_n(X_{j,n} - b_n) \Rightarrow W_j$ for all $j = 1, 2, \cdots$. (see, e.g., [13, Corollary 4.19]). (5)

For $F(x)$ in Case I or III for minima, respectively $E = [-\infty, \infty)$ or $E = [0, \infty)$ with $\nu[-\infty, x] = e^x$ for $x \in \mathbb{R}$, or $\nu[0, x] = x^\alpha$ for some $\alpha > 0$ for $0 < x < \infty$. The constants $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ are identified in the Appendix; in particular, in Case III, $b_n = 0$ for all $n \geq 1$.

It is straightforward to show that, for these distribution function settings, one can obtain the desired type of joint convergence results without normalizations only if the distributions and the capacities $(c_n)_{n\geq 1}$ satisfy either (a) $F(x)$ is in Case I for minima with $l_F = 0$ and $a_n(c_n - kb_n) \to \gamma$ for some positive integer $k = 1, 2, \cdots$ and some $\gamma \in \mathbb{R}$; or (b) $F(x)$ is in Case III for minima with $l_F = 0$ and $a_n c_n \to \gamma$ for some $0 < \gamma < \infty$. In particular, it is required that $c_n \to 0$ at the rates determined by the two conditions (a) and (b). In the cases not covered under these two conditions one can show that sequence $\{N^e(c_n)\}_{n\geq 1}$ without normalizations has a degenerate limit, if a limit exists.

For the statement of the Theorem giving the asymptotic behavior when d.f. $F(x)$ is in Case III for minima with $l_F = 0$ and index $\alpha > 0$, we introduce r.v.'s $\mathcal{N}^e = \mathcal{N}^e(\gamma)$ and $\mathcal{N}^s = \mathcal{N}^s(\delta, \gamma)$ for $0 < \delta, \gamma < \infty$ in terms of the Poisson random measure $\xi = \sum_{j\geq 1} \epsilon_{(\tau_j, W_j)}$ on $[0,1] \times [0, \infty)$ with intensity measure $dt \times d\nu$, where $\nu[0, x] = x^\alpha$ for $0 < x < \infty$, of statement (5). Define $\mathcal{N}^e$ and $\mathcal{N}^s$ by

$$\mathcal{N}^e = \mathcal{N}^e(\gamma) := \max\{j \geq 1 : W_1 + \cdots W_j \leq \gamma\} \text{ if this set is nonempty and } = 0 \text{ otherwise;} \quad (6)$$

and $\mathcal{N}^s = \mathcal{N}^s(\delta, \gamma) := \sum_{l\geq 1} I(\rho_l < 1)$ where $\rho_1 := \min\{0 < \tau_k < 1 : W_k \leq \delta \text{ and } W_k \leq \gamma\}$ if this set is nonempty and $= 1$ otherwise; and, having defined $\rho_1, \ldots, \rho_{l-1}$, define ρ_l by $\rho_l := \min\{\rho_{l-1} < \tau_k < 1 : W_k \leq \delta \text{ and } \sum_{j\,:\,\tau_j \leq \tau_k} W_j I(W_j \leq \delta) \leq \gamma\}$ if $\rho_{l-1} < 1$ and this set is nonempty and $= 1$ otherwise. We have the representation

$$\mathcal{N}^s = \sum_{j\,:\,\tau_j \leq \nu} I(W_j \leq \delta) \quad (7)$$

where $\nu = \nu(\delta, \gamma) = \max\{0 < \tau_k < 1 : \sum_{j\,:\,\tau_j \leq \tau_k} W_j I(W_j \leq \delta) \leq \gamma\}$ if this set is nonempty and $= 0$ otherwise.

Theorem 2.1 *Let distribution function $F(x)$ be in Case III for minima with $l_F = 0$ and with index $\alpha > 0$ and associated positive constants $(a_n)_{n\geq 1}$, and let $(c_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$ be sequences of positive constants and $0 < \delta, \gamma < \infty$ such that $a_n c_n \to \gamma$ and $a_n \varepsilon_n \to \delta$. Then*

$$(N^e(c_n), N^{s_n}(\varepsilon_n, c_n)) \Rightarrow (\mathcal{N}^e(\gamma), \mathcal{N}^s(\delta, \gamma)).$$

In particular, the sequence of threshold policies $(s_n)_{n\geq 1}$ is a consistent approximator of the off-line smallest fit strategy τ^e.

Proof. Let $\xi = \sum_{j\geq 1} \epsilon_{(\tau_j, W_j)}$ be the Poisson random measure on $[0,1] \times [0, \infty)$ with intensity measure $dt \times d\nu$, where $\nu[0, x] = x^\alpha$ for $x > 0$, as in statement (5); and let $\xi_\delta = \xi|_{[0,1]\times[0,\delta]}$ denote the restriction of ξ to $[0,1] \times [0, \delta]$ where $\xi_\delta = \sum_{j=1}^{L} \epsilon_{(\sigma_j, U_j)}$ with $L := \#\{(\tau_j, W_j) \in [0,1] \times [0, \delta]\}$. Without loss of generality, it is assumed that with probability one

$$\begin{aligned} &(i)\ \tau_j \neq 0, 1 \text{ for all } j \geq 1;\ 0 < W_1 < W_2 < \cdots;\ W_j \neq \delta \text{ for all } j \geq 1; \text{ and} \\ &\quad\ W_1 + \cdots + W_j \neq \gamma \text{ for all } j \geq 1; \text{ and} \\ &(ii)\ L < \infty, \text{ and for } L \geq 1,\ 0 < \sigma_1 < \cdots < \sigma_L < 1 \text{ and } U_1 + \cdots + U_j \neq \gamma \\ &\quad\ \text{for all } j = 1, \ldots, L. \end{aligned} \quad (8)$$

Thus, with probability one, $\mathcal{N}^s = \mathcal{N}^s(\delta,\gamma) = \#\{(\sigma_j, U_j) : \sum_{i=1}^{j} U_i \le \gamma\}$ if this set is nonempty, and $= 0$ otherwise.

Also, let ξ_n denote the point process $\xi_n = \sum_{k=1}^{n} \epsilon_{(k/n, a_n X_k)}$ of statement (5), so that $\xi_n \Rightarrow \xi$ in $M_p = M_p([0,1]\times[0,\infty))$. To prove the conclusion, set up an application of the Continuous Mapping Theorem, found e.g. in Billingsley [1, Theorem 5.5] or Whitt [15], as follows.

Let $\delta_n = a_n\varepsilon_n$ and $\gamma_n = a_n c_n$ for $n \ge 1$, and $\delta^* = \sup_n \delta_n$. Define functions $\Lambda^e : M_p \times (0,\infty) \to \{0,1,\cdots\}$, $\Lambda^1 : M_p \times (0,\delta^*] \to M_p([0,1]\times[0,\delta^*])$, and $\Lambda^2 : M_p([0,1]\times[0,\delta^*]) \times (0,\infty) \to \{0,1,\cdots\}$ by

$$\Lambda^e(m,\eta) = \#\{k \ge 1 : \sum_{j=1}^{k} y_j \le \eta\} \text{ if this set is nonempty, and } = 0 \text{ otherwise;}$$

$$\Lambda^1(m,\mu) = m_\mu = m|_{[0,1]\times[0,\mu]}, \text{ the restriction of } m \text{ to } [0,1]\times[0,\mu]$$

for $m \in M_p$ with representation $m = \sum_{j\ge1} \epsilon_{(t_j,y_j)}$, where $0 \le y_1 \le y_2 \le \cdots$, $\eta > 0$ and $\mu \in (0,\delta^*]$; and for $m \in M_p([0,1]\times[0,\delta^*])$ with $m = \sum_{j=1}^{l} \epsilon_{(s_j,u_j)}$ with $0 \le s_1 \le \cdots \le s_l$ for $l \ge 1$ (interpret m as the zero measure if $l = 0$) and $\eta > 0$

$$\Lambda^2(m,\eta) = \#\{k \ge 1 : \sum_{j=1}^{k} u_j \le \eta\} \text{ if this set is nonempty, and } = 0 \text{ otherwise.}$$

Let $n \ge 1$ and define Λ^e_n, $\Lambda^e : M_p \to \{0,1,\cdots\}$ by $\Lambda^e_n(m) = \Lambda^e(m,\gamma_n)$ and $\Lambda^e(m) = \Lambda^e(m,\gamma)$, define Λ^1_n, $\Lambda^1 : M_p \to M_p([0,1]\times[0,\delta^*])$ by $\Lambda^1_n(m) = m_{\delta_n}$ and $\Lambda^1(m) = m_\delta$, and define Λ^2_n, $\Lambda^2 : M_p([0,1]\times[0,\delta^*]) \to \{0,1,\cdots\}$ by $\Lambda^2_n(m) = \Lambda^2(m,\gamma_n)$ and $\Lambda^2(m) = \Lambda^2(m,\gamma)$.

For this application the functionals of interest are $(\Lambda^e_n)_{n\ge1}$ and Λ^e, and $(\Lambda^s_n)_{n\ge1}$ and Λ^s, where $\Lambda^s_n, \Lambda^s : M_p \to \{0,1,\cdots\}$ are defined by $\Lambda^s_n = \Lambda^2_n \circ \Lambda^1_n$ and $\Lambda^s = \Lambda^2 \circ \Lambda^1$ for $n \ge 1$. Indeed, observe from the definitions that with probability one

$$\begin{aligned} (\Lambda^e_n(\xi_n), \Lambda^s_n(\xi_n)) &= (N^e(c_n), N^{s_n}(\varepsilon_n, c_n)) \text{ for } n \ge 1, \text{ and} \\ (\Lambda^e(\xi), \Lambda^s(\xi)) &= (\mathcal{N}^e(\gamma), \mathcal{N}^s(\delta,\gamma)). \end{aligned}$$

For application of the Continuous Mapping Theorem, it suffices to exhibit a set $\mathcal{D} = \mathcal{D}(\delta,\gamma)$ in M_p which satisfies

$$\begin{aligned} &(i)\ \mathcal{D} \text{ is a measurable subset of } M_p \text{ and } P(\xi \in \mathcal{D}) = 1, \text{ and} \\ &(ii) \text{ for each } m \in \mathcal{D} \text{ and each sequence } (m_n)_{n\ge1} \subset M_p \text{ for which } m_n \xrightarrow{\mathrm{v}} m, \\ &\quad \text{it follows that } (\Lambda^e_n(m_n), \Lambda^s_n(m_n)) \to (\Lambda^e(m), \Lambda^s(m)). \end{aligned} \tag{9}$$

Define $\mathcal{D} = \mathcal{D}(\delta,\gamma)$ as the set of point measures m in M_p satisfying $m = \sum_{j\ge1} \epsilon_{(t_j,y_j)}$ for some $\{(t_j,y_j)\}_{j\ge1}$ with $0 < t_j < 1$ for all $j \ge 1$, $0 < y_1 < y_2 < \cdots$, $y_j \ne \delta$ and $y_1 + \cdots + y_j \ne \gamma$ for all $j \ge 1$, and either $m_\delta = 0$, or for some $l \ge 1$, $m_\delta = \sum_{i=1}^{l} \epsilon_{(r_i,u_i)}$ with $0 < r_1 < \cdots < r_l < 1$ and $u_1 + \cdots + u_j \ne \gamma$ for all $1 \le j \le l$. One can show that $\mathcal{D}$ is a measurable subset of M_p (a set in the σ-algebra of Borel sets generated by the vague topology on M_p), for example, by representing $\mathcal{D}$ as $\mathcal{D} = \bigcap_{L=[\delta]}^{\infty} \bigcup_{k=0}^{\infty} \bigcap_{K=k+1}^{\infty} \bigcup_{q\in Q\cap[L,\infty)} \bigcup_{h_1,h_2=0}^{k} \bigcup_{\mu=0}^{h_1} \mathcal{D}(L,k,K,q,h_1,h_2,\mu)$ where the sets $\mathcal{D}(L,k,K,q,h_1,h_2,\mu)$ have the form $\{m = \sum_{j\ge1} \epsilon_{(t_j,y_j)} \in M_p$: from $\{(t_j,y_j)\}_{j\ge1}$, the points $\{(t_j,y_j)\}_{1\le j\le k}$ are in $[0,1]\times[0,L)$ and only these points, $0 < t_j < 1$ for all $1 \le j \le k$,

$0 < y_1 < \cdots < y_k < L$, $y_{h_1} < \delta < y_{h_1+1}$, $y_0 + \cdots + y_{h_2} < \gamma < y_0 + \cdots + y_{h_2+1}$ and $m|_{[0,1]\times[0,\delta]} = \sum_{i=1}^{h_1} \epsilon_{(r_i,u_i)}$ with $0 < r_1 < \cdots < r_{h_1} < 1$ and $u_0 + \cdots + u_\mu < \gamma < u_0 + \cdots + u_{\mu+1}$; and there are K distinct points from $\{(t_j, y_j)\}_{j\geq 1}$ in $[0,1] \times [0,q)$, each with point mass one } (suitably modified in 'boundary index' cases), and then by using the standard neighborhoods of the vague topology in M_p to show that the sets $\mathcal{D}(L,k,K,q,h_1,h_2,\mu)$ are open in the vague topology (for definitions and applications of the vague topology, see Chapters 3 and 4 of Resnick [13]). From (8), it follows that $P(\xi \in \mathcal{D}) = 1$.

We sketch a representative part of the proof of the continuity result (9) (ii). For $m \in M_p$, denote $m_\mu = m|_{[0,1]\times[0,\mu]}$ as before and $l_\mu = m_\mu([0,1]\times[0,\mu])$. Let $m = \sum_{j\geq 1} \epsilon_{(t_j,y_j)} \in \mathcal{D}$ with $\Lambda^e(m) = B$ and $\Lambda^s(m) = C$ for some $B, C \in \{1,2,\cdots\}$ (the argument for B or $C = 0$ is analogous), so that

(a) $\sum_{j=0}^{B} y_j < \gamma < \sum_{j=0}^{B+1} y_j$; and
(b) $m_\delta = \sum_{j=1}^{l_\delta} \epsilon_{(s_j,u_j)}$ with $0 < s_1 < \cdots < s_{l_\delta}$ and $0 < u_j < \delta$ for $1 \leq j \leq l_\delta$ and either (i) $\sum_{j=0}^{C} u_j < \gamma < \sum_{j=0}^{C+1} u_j$, or (ii) $\sum_{j=0}^{C} u_j < \gamma$ and $l_\delta = C$.

Let $(m_n)_{n\geq 1} \subset M_p$ with $m_n \xrightarrow{\text{v}} m$, and use Proposition 3.13 of [13] to obtain that for all n large there is a representation $m_n = \sum_{j\geq 1} \epsilon_{(t_j^{(n)}, y_j^{(n)})}$ for which

(a) $0 < y_1^{(n)} < \cdots < y_{B+1}^{(n)} < y_k^{(n)}$ for all $k > B+1$, $\lim_{n\to\infty}(t_j^{(n)}, y_j^{(n)}) = (t_j, y_j)$ for $1 \leq j \leq B+1$, and $\sum_{j=0}^{B} y_j^{(n)} < \gamma_n < \sum_{j=0}^{B+1} y_j^{(n)}$; and
(b) $l_{\delta_n} = l_\delta$ and $(m_n)_{\delta_n} = \sum_{j=1}^{l_\delta} \epsilon_{(s_j^{(n)}, u_j^{(n)})}$ with $0 < s_1^{(n)} < \cdots < s_{l_\delta}^{(n)}$ and $0 < u_j^{(n)} < \delta$ for $1 \leq j \leq l_\delta$ and either (i) $\sum_{j=0}^{C} u_j^{(n)} < \gamma < \sum_{j=0}^{C+1} u_j^{(n)}$, or (ii) $\sum_{j=0}^{C} u_j^{(n)} < \gamma$ and $l_\delta = C$.

It follows that for all large n, $\Lambda_n^e(m_n) = B$ and $\Lambda_n^s(m_n) = C$ and thus $\lim_{n\to\infty}(\Lambda_n^e(m_n), \Lambda_n^e(m_n)) = (\Lambda^e(m), \Lambda^s(m))$; and one may apply the Continuous Mapping Theorem to obtain the desired conclusion. □

Theorem 2.1 can be used to obtain asymptotic bounds on the sequences of expectations $\{EN^e(c_n)\}_{n\geq 1}$ and $\{EN^{s_n}(\varepsilon_n, c_n)\}_{n\geq 1}$. These asymptotic bounds can then be used as a measure of comparison between the off-line strategy τ^e and the on-line strategies $\{s_n\}$. To illustrate this, we first give general bounds for $N^e(c_n)$ and $N^{s_n}(\varepsilon_n, c_n)$ (used again in Section 3) and then combine these with the weak convergence result of Theorem 2.1 to obtain asymptotic bounds.

Lemma 2.2 *For all nonnegative r.v.'s $X_1, X_2, \cdots$ and positive constants $(c_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$, r.v. $N^e(c_n)$ satisfies*

$$N^e(c_n) \leq c_n \varepsilon_n^{-1} + \varepsilon_n^{-1} \sum_{i=1}^{n} (\varepsilon_n - X_i) I(X_i \leq \varepsilon_n), \tag{10}$$

and r.v. $N^{s_n}(\varepsilon_n, c_n)$ satisfies

$$EN^{s_n}(\varepsilon_n, c_n) \geq nF(\varepsilon_n) - 1 - \frac{F(\varepsilon_n) E|\sum_{i=1}^{n} X_i I(X_i \leq \varepsilon_n) - c_n|}{EX_i I(X_i \leq \varepsilon_n)}. \tag{11}$$

Proof. Observe that inequality (10) can be rewritten as

$$N^e(c_n) \le N'_n + \varepsilon_n^{-1}\left(c_n - \sum_{i=1}^n X_i I(X_i \le \varepsilon_n)\right) \tag{12}$$

where $N'_n = \sum_{i=1}^n I(X_i \le \varepsilon_n)$; and inequality (12) is immediate, since, on the set $\{N'_n \le N^e(c_n)\}$,

$$c_n \ge \sum_{i=1}^{N^e(c_n)} X_{i,n} = \sum_{i=1}^{N'_n} X_{i,n} + \sum_{i=N'_n+1}^{N^e(c_n)} X_{i,n} \ge \sum_{i=1}^n X_i I(X_i \le \varepsilon_n) + \varepsilon_n(N^e(c_n) - N'_n)$$

and on the set $\{N'_n \ge N^e(c_n)\}$,

$$\sum_{i=1}^n X_i I(X_i \le \varepsilon_n) = \sum_{i=1}^{N'_n} X_{i,n} = \sum_{i=1}^{N^e(c_n)} X_{i,n} + \sum_{i=N^e(c_n)+1}^{N'_n} X_{i,n} \le c_n + \varepsilon_n(N'_n - N^e(c_n)).$$

To prove inequality (11), use an argument similar to that used by Rhee and Talagrand to prove inequality (3) in [14]. □

Corollary 2.3 *Under the hypotheses of Theorem* 2.1

$$\begin{aligned}
&((\alpha+1)\gamma/\alpha)^{\alpha/(\alpha+1)} \\
&\ge\ E\mathcal{N}^e(\gamma) = \lim_{n\to\infty} EN^e(c_n) \\
&\ge\ E\mathcal{N}^s(\delta,\gamma) = \lim_{n\to\infty} EN^{s_n}(\varepsilon_n, c_n) \\
&\ge\ \left(\left(\frac{\alpha+1}{\alpha}\right)\left(\frac{\gamma}{\delta}\right) - \frac{1}{2}\left(\frac{\alpha+1}{\alpha}\right)^2\left(\frac{\gamma}{\delta}\right)^2\delta^{-\alpha} - 1 - \frac{1}{2}\frac{(\alpha+1)^2}{\alpha(\alpha+2)}\right)_+
\end{aligned} \tag{13}$$

(with $a_+ = a \vee 0$*); and*

$$\lim_{\gamma\to\infty} \inf_{\delta} \frac{E\mathcal{N}^e(\gamma)}{E\mathcal{N}^s(\delta,\gamma)} \le g(\alpha) \tag{14}$$

where $g(\alpha) = ((\alpha+2)/2)^{1/(\alpha+1)}((\alpha+2)/(\alpha+1))$, *a function on* $[0,\infty)$ *decreasing from* $g(0) = 2$ *to* $g(\infty) = 1$.

Proof. First, observe that the r.v.'s $N'_n = \sum_{i=1}^n I(X_i \le \varepsilon_n)$, $n \ge 1$, in the proof of Lemma 2.2 satisfy $\lim_{n\to\infty} EN'_n = \delta^\alpha$ and $\lim_{n\to\infty} E(N'_n)^2 = \delta^\alpha + (\delta^\alpha)^2$. Since $0 \le N^{s_n}(\varepsilon_n, c_n) \le N^e(c_n) \le c_n\varepsilon_n^{-1} + N'_n$ (from inequality (10)), it follows that each of the sequences $\{N^{s_n}(\varepsilon_n, c_n)\}_{n\ge1}$, $\{N^e(c_n)\}_{n\ge1}$ and $\{N'_n\}_{n\ge1}$ is uniformly integrable. Furthermore, from Karamata's Theorem (as given in the Appendix), $E\left(\varepsilon_n^{-1}\sum_{i=1}^n X_i I(X_i \le \varepsilon_n)\right) = na_n\int_0^{F(\varepsilon_n)} F^{\leftarrow}(s)ds \approx nF(\varepsilon_n)a_n\varepsilon_n(\alpha/(\alpha+1))$; and so from Theorem 2.1 and inequality (10), one concludes that

$$0 \le E\mathcal{N}^s(\delta,\gamma) = \lim_{n\to\infty} EN^{s_n}(\varepsilon_n, c_n) \le E\mathcal{N}^e(\gamma) = \lim_{n\to\infty} EN^e(c_n) \le \gamma\delta^{-1} + (\alpha+1)^{-1}\delta^\alpha.$$

Minimization of the right-hand-side function over δ occurs at $\delta = ((\alpha+1)\gamma/\alpha)^{1/(\alpha+1)}$ and yields the first inequality of the conclusion.

Next, let $n \to \infty$ in inequality (11) to obtain

$$E\mathcal{N}^s(\delta,\gamma) \geq \delta^\alpha - 1 - ((\alpha+1)/(\delta\alpha))E\left|\sum_i U_i - \gamma\right|, \tag{15}$$

since $EN^{s_n}(\varepsilon_n, c_n) \to E\mathcal{N}^s(\delta,\gamma)$, $nF(\varepsilon_n) \to \delta^\alpha$, $na_nEX_iI(X_i \leq \varepsilon_n) \to (\alpha/(\alpha+1))\delta^{\alpha+1}$, and $E\left|\sum_{i=1}^n a_nX_iI(a_nX_i \leq a_n\varepsilon_n) - a_nc_n\right| \to E\left|\sum_i U_i - \gamma\right|$ (using the Continuous Mapping Theorem and uniform integrability). Now observe that $\delta^{-1}U_1, \ldots, \delta^{-1}U_l$ are i.i.d. r.v.'s given $N = l$, with common d.f. $F(x) = x^\alpha$, $0 \leq x \leq 1$, where $N = \xi([0,1]\times[0,\delta])$ is Poisson (δ^α)-distributed. Use this fact and (15) to obtain

$$\begin{aligned} E\mathcal{N}^s(\delta,\gamma) &\geq \delta^\alpha - 1 - ((\alpha+1)/(\delta\alpha))\left(E\left(\sum_i U_i - \gamma\right)^2\right)^{1/2} \\ &\geq \left(\frac{\alpha+1}{\alpha}\right)\left(\frac{\gamma}{\delta}\right) - \frac{1}{2}\left(\frac{\alpha+1}{\alpha}\right)^2\left(\frac{\gamma}{\delta}\right)^2\delta^{-\alpha} - 1 - \frac{1}{2}\frac{(\alpha+1)^2}{\alpha(\alpha+2)}, \end{aligned}$$

and so the last inequality of (13) is proved. A simple calculus argument, based on the bounds in (13), yields (14). □

Example 2.4 Let $F(x)$ be the Uniform $(0,1)$-distribution function (so that $F(x)$ is in Case III for minima with $\alpha = 1$). By maximizing the right-hand expression in (13), one obtains that for $\gamma > 0$ and $\delta = \delta(\gamma) = \sqrt{3\gamma}$

$$\begin{aligned} \sqrt{2\gamma} &\geq E\mathcal{N}^e(\gamma) = \lim_{n\to\infty} EN^e(c_n) \\ &\geq E\mathcal{N}^s(\delta(\gamma),\gamma) = \lim_{n\to\infty} EN^{s_n}(\varepsilon_n, c_n) \geq \left((4\sqrt{3}/9)\sqrt{\gamma} - (5/3)\right)_+ \end{aligned} \tag{16}$$

where the limits use sequences of positive numbers $(\varepsilon_n)_{n\geq1}$ and $(c_n)_{n\geq1}$ satisfying $a_nc_n \to \gamma$ and $a_n\varepsilon_n \to \delta(\gamma) = \sqrt{3\gamma}$. One can then use these bounds to obtain inequality (14) in this case, which implies that there are constants γ and δ sufficiently large, so that for these sequences $(\varepsilon_n)_{n\geq1}$ and $(c_n)_{n\geq1}$, and for all n sufficiently large, $1 \leq EN^e(c_n)/EN^{s_n}(\varepsilon_n, c_n) < 1.838$. For this special d.f. $F(x)$, one can also use techniques of Section 5 of [5] to obtain the bound $E\mathcal{N}^e(\gamma) \leq (6/5)\gamma$, an improvement on inequality (16) for $0 < \gamma < 25/18$.

Next, let distribution function $F(x)$ be in Case I for minima with $l_F = 0$. In this case, the convergence result of the non-normalized counting r.v.'s differs markedly from Theorem 2.1 for $F(x)$ in Case III, due to the necessity here of location scaling in the underlying point process convergence. In this case, the limiting pair $(\mathcal{N}^e, \mathcal{N}^s)$ is defined as follows. Let $\xi = \sum_{j\geq1}\epsilon_{(\tau_j, W_j)}$ denote the Poisson random measure on $[0,1]\times[-\infty,\infty)$ with intensity measure $dt \times d\nu$, where $\nu[-\infty, x] = e^x$ for $x \in \mathbb{R}$, as in (5). For integers $K = 1, 2, \cdots$ and $\delta, \gamma \in \mathbb{R}$, define $\mathcal{N}^e = \mathcal{N}^e(\gamma; K)$ by

$$\mathcal{N}^e = \begin{cases} K-1 & \text{if} \quad \gamma < \sum_{i=1}^K W_i \\ K & \text{if} \quad \sum_{i=1}^K W_i \leq \gamma \end{cases} \tag{17}$$

and define $\mathcal{N}^s = \mathcal{N}^s(\delta,\gamma;K)$ by

$$\mathcal{N}^s = \begin{cases} j & \text{if} \quad W_j \leq \delta < W_{j+1} \text{ for } j = 0, \ldots, K-2 \\ K-1 & \text{if} \quad W_{K-1} \leq \delta < W_K, \text{ or if } W_K \leq \delta \text{ and } \gamma < \sum_{i=1}^K U_i \\ K & \text{if} \quad W_K \leq \delta \text{ and } \sum_{i=1}^K U_i \leq \gamma \end{cases} \tag{18}$$

where $W_0 \equiv 0$, and we have assumed, without loss of generality, that $-\infty < W_1 < W_2 < \cdots$ and $\xi_\delta = \xi|_{[0,1]\times[-\infty,\delta]} = \sum_{j=1}^{L} \epsilon_{(\sigma_j,U_j)}$ with finite $L = \#\{(\tau_j, W_j) \in [0,1]\times[-\infty,\delta]\}$, and for $L \geq 1$, $0 < \sigma_1 < \cdots < \sigma_L < 1$, with probability one.

Theorem 2.5 *Let distribution function $F(x)$ be in Case I for minima with $l_F = 0$ and associated positive constants $a_n > 0$ and $b_n \in \mathbb{R}$, $n \geq 1$. Let $(c_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$ be sequences of positive constants, K be a positive integer and δ and γ be constants in $\mathbb{R}$ such that $a_n(c_n - Kb_n) \to \gamma$ and $a_n(\varepsilon_n - b_n) \to \delta$. Then*

$$(N^e(c_n), N^{s_n}(\varepsilon_n, c_n)) \Rightarrow (\mathcal{N}^e, \mathcal{N}^s).$$

Proof. The proof follows along the same lines of that for Theorem 2.1. □

For d.f. $F(x)$ satisfying the conditions of Theorem 2.5, we can obtain from this result that for appropriately chosen $(c_n)_{n\geq 1}$, $(\varepsilon_n)_{n\geq 1}$ and K, $EN^{s_n}(\varepsilon_n, c_n)$ is close to $EN^e(c_n)$. This is illustrated in the following corollary.

Corollary 2.6 *Let $F(x)$ be in Case I for minima with $l_F = 0$. Then for each $K = 1, 2, \cdots$,*

$$\lim_{\gamma\to\infty} \lim_{\delta\to\infty} E\mathcal{N}^s(\delta,\gamma)/E\mathcal{N}^e(\gamma) = 1;$$

and under the hypotheses of Theorem 2.5,

$$\lim_{\gamma\to\infty} \lim_{\delta\to\infty} \lim_{n\to\infty} EN^{s_n}(\varepsilon_n, c_n)/EN^e(c_n) = 1$$

where the inside limit is taken over $n \to \infty$, with $a_n(c_n - Kb_n) \to \gamma$ and $a_n(\varepsilon_n - b_n) \to \delta$, and

$$\lim_{\gamma\to\infty} \lim_{n\to\infty} EN_n^*(c_n)/EN^e(c_n) = 1$$

where $N_n^*(c_n)$ was defined through (2) and the inside limit is taken over $n \to \infty$ with $a_n(c_n - Kb_n) \to \gamma$.

Proof. We have in this case that

$$\begin{aligned} E\mathcal{N}^e(\gamma) &= K - 1 + P(\sum_{i=1}^{K} W_i > \gamma) \text{ and} \\ E\mathcal{N}^s(\delta,\gamma) &= \sum_{j=0}^{K-1} j e^{\delta j} \exp(-e^\delta)/j! + (K-1)P(W_k \leq \delta) + P(\sum_{i=1}^{K} U_i \leq \gamma) \end{aligned}$$

where

$$\begin{aligned} P\left(\sum_{i=1}^{K} U_i \leq \gamma\right) &= P(\text{ there are } K \text{ or more points in } [0,1]\times[-\infty,\delta]) \\ &\times \quad P(E_1 + \cdots + E_K \geq (K\delta - \gamma)^+) \\ &= \left(\sum_{j=K}^{\infty} e^{\delta j} \exp(-e^\delta)/j!\right) P(Y \geq (K\delta - \gamma)^+) \end{aligned}$$

and $E_1, \ldots, E_K$ are i.i.d. Exponential (1)-distributed r.v.'s and Y is a Gamma $(K,1)$-distributed r.v. The first limit result is now immediate (by choosing γ sufficiently large, then δ sufficiently large); and the other limit results follow from Theorem 2.5 (use the convergence-in-distribution together with uniform integrability of $\{N^{s_n}(\varepsilon_n, c_n)\}_{n\geq 1}$ and $\{N^e(c_n)\}_{n\geq 1}$ to obtain $\lim_{n\to\infty} EN^{s_n}(\varepsilon_n, c_n) = E\mathcal{N}^s$ and $\lim_{n\to\infty} EN^e(c_n) = E\mathcal{N}^e$). □

3 Numbers Selected Grow to infinity, but are Small Compared to Increasing Sample Sizes

In this section, settings are considered in which counting r.v.'s $N^e(c_n)$ and $N^{s_n}(c_n) \nearrow \infty$, but $N^e(c_n)/n$ and $N^{s_n}(c_n)/n \to 0$, as the sample size $n \to \infty$. Results are proved on the weak convergence of appropriate normalizations of $N^e(c_n)$ and $N^{s_n}(c_n)$ to nondegenerate random variables as $n \to \infty$.

Specifically, for the settings of this section, the distribution functions $F(x)$ are in Case I or Case III for minima with $l_F \geq 0$. For $F(x)$ in Case I for minima the function $c(s)$ is defined by $c(s) = s^{-1} \int_0^s u dF^{\leftarrow}(u)$; and for $F(x)$ in Case III for minima recall the representation $F^{\leftarrow}(s) = l_F + s^{-a}L(s)$ where $L(s)$ is a function slowly varying at zero and $a = -1/\alpha$, for $\alpha > 0$. See the Appendix for background results on extreme value theory used in this section, and in particular, for properties of the functions $F^{\leftarrow}(s)$ and $c(s)$. Observe that for these types of distribution functions $F(l_F) = 0$.

Given a sequence of positive constants $(c_n)_{n\geq 1}$, define constants $(B_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$ by

$$c_n = n \int_0^{B_n/n} F^{\leftarrow}(s)ds \text{ and } \varepsilon_n = F^{\leftarrow}(B_n/n). \tag{19}$$

Why is the threshold ε_n chosen this way? Intuitively, an on-line estimate of $X_{B_n,n} = F^{\leftarrow}(U_{B_n,n})$ or $U_{B_n,n}$ is desired; so we use $E(U_{B_n,n}) = B_n/n$ and compare X_j's to the threshold $F^{\leftarrow}(B_n/n)$. It would be useful to have a 'better' on-line estimator of $X_{B_n,n}$ that would still be simple enough to carry through an analysis of the problem.

Throughout this section, *the positive constants* $(c_n)_{n\geq 1}$ *and* $(B_n)_{n\geq 1}$ *are assumed to satisfy* $c_n/n \to 0$ *and* $B_n \to \infty$ *as* $n \to \infty$. For the norming constants in Theorem 3.1, this implies that the location parameters $(B_n)_{n\geq 1}$ satisfy $B_n \to \infty$ but $B_n << n$, and the scaling parameters $(A_n)_{n\geq 1}$, $(B_n^{1/2})_{n\geq 1}$ satisfy $A_n, B_n^{1/2} << n^{1/2}$, B_n, as $n \to \infty$. So the total number of sizes selected goes to infinity, but remains small in comparison to the sample size n, as $n \to \infty$, both for the off-line smallest fit strategy τ^e and for the on-line stopped threshold policies $(s_n)_{n\geq 1}$. However, as noted in the proof of Theorem 3.1 below, the 'over-capacity' r.v.'s $(\nu_n(c_n))_{n\geq 1}$ in the representation (3) satisfy $\nu_n(c_n)/n \to 1$ in probability as $n \to \infty$; so the policies $(s_n)_{n\geq 1}$ view larger and larger portions of the samples before reaching capacity, as $n \to \infty$. Observe also that under this assumption, the following 'capacities vs. thresholds' comparisons holds: $c_n \approx B_n\varepsilon_n$ for $F(x)$ in Case I, and in Case III for $l_F > 0$; and $c_n \approx B_n\varepsilon_n/(1+\alpha^{-1})$ for $F(x)$ in Case III with $l_F = 0$ and parameter $\alpha > 0$. Thus, the policies $(s_n)_{n\geq 1}$ are feasible, for all n large.

In the main results of this section, it is also assumed that

$$B_n^{1/2}\left(\frac{F(\varepsilon_n)}{B_n/n} - 1\right) \to 0 \text{ as } n \to \infty \tag{20}$$

and, as $n \to \infty$,

$$\begin{aligned} B_n^{1/2}c(B_n/n)/F^{\leftarrow}(B_n/n) &\to \infty \quad \text{for} \quad F(x) \text{ in Case I, and} \\ B_n^{1/2}(F^{\leftarrow}(B_n/n) - l_F) &\to \infty \quad \text{for} \quad F(x) \text{ in Case III with } l_F > 0. \end{aligned} \tag{21}$$

Condition (20) ensures that the location parameters coincide for the convergence of the on-line and off-line counting r.v.'s in Theorem 3.1. If d.f. $F(x)$ is a continuous d.f., then $F(\varepsilon_n) =$

$F(F^{\leftarrow}(B_n/n)) = B_n/n$ for all $n \geq 1$, and (20) holds. Also, for all d.f.'s in Cases I or III for minima with $l_F \geq 0$, it follows that $1 \leq F(\varepsilon_n)/(B_n/n) \leq F(\varepsilon_n)/F(\varepsilon_n-) \to 1$ as $n \to \infty$, for example, by using Karamata's Representation for F (see the Appendix). Thus, condition (20) is a condition on how fast $F(\varepsilon_n)/(B_n/n) \to 1$. For the types of distribution functions of interest in this section, those in Case I or III for minima with $l_F \geq 0$, there are distribution functions that do not satisfy condition (20) for some sequences $(B_n)_{n\geq 1}$ satisfying the hypotheses of this section and for which Theorem 3.1 does not hold (see Example 3.7). Condition (21) controls the effect of variations in the off-line counting r.v.'s in the convergence results of Theorem 3.1; no such additional condition is required when d.f. $F(x)$ is in Case III for minima with $l_F = 0$. Condition (21) can be thought of as a restriction on the speed of convergence of $c(s)/F^{\leftarrow}(s) \to 0$ (or equivalently, $\int_0^s F^{\leftarrow}(u)du/(sF^{\leftarrow}(s)) \to 1$) for $F(x)$ in Case I, and on the speed of convergence of $F^{\leftarrow}(s) - l_F = s^{-a}L(s) \to 0$ in Case III with $l_F > 0$ (see Example 3.8).

The next theorem is the main result of this section. Observe that in this theorem, the case $l_F > 0$ is allowed; however, in this case the assumptions on the constants $(c_n)_{n\geq 1}$ and $(B_n)_{n\geq 1}$ of this section force $c_n \to \infty$. Contrast this to the setting and results of Section 2, in which $c_n \to 0$ and only $l_F = 0$ gives a nondegenerate result. In the case of $l_F = 0$, it is possible in the next theorem that $c_n \to c$, for any $0 \leq c \leq \infty$ (see Example 3.4).

Theorem 3.1 *Let $F(x)$ be in Case I or III for minima with $l_F \geq 0$ and let the positive constants $(c_n)_{n\geq 1}$, $(B_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$ satisfy (19), (20) and (21). Then there exist positive constants $(A_n)_{n\geq 1}$ for which*

$$\left(A_n^{-1}\left(N^e(c_n) - B_n\right), B_n^{-1/2}\left(N^{s_n}(c_n) - B_n\right)\right) \Rightarrow (\mathcal{N}^e, \mathcal{N}^s)$$

where $(\mathcal{N}^e, \mathcal{N}^s) = (W_1, W_2 + (W_3 \wedge 0))$ and $\mathbf{W} = (W_1, W_2, W_3)$ is $N(\mathbf{0}, \Sigma_{\mathbf{W}})$-distributed.

For $F(x)$ in Case I, the constants $(A_n)_{n\geq 1}$ satisfy

$$A_n \approx B_n^{1/2} c(B_n/n)/F^{\leftarrow}(B_n/n), \text{ and } \Sigma_{\mathbf{W}} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}.$$

For $F(x)$ in Case III, the constants $(A_n)_{n\geq 1}$ satisfy

$$A_n \approx \begin{cases} B_n^{1/2} & \text{for} \quad l_F = 0 \\ B_n^{1/2}\left(F^{\leftarrow}(B_n/n) - l_F\right)/l_F & \text{for} \quad 0 < l_F < \infty \end{cases}$$

$$\text{and } \Sigma_{\mathbf{W}} = \begin{pmatrix} K_a^2 & \frac{-a}{1-a} & \frac{a}{1-2a} \\ \frac{-a}{1-a} & 1 & -1 \\ \frac{a}{1-2a} & -1 & \frac{(1-a)^2}{1-2a} \end{pmatrix} \text{ if } l_F = 0, \text{ and } = \begin{pmatrix} K_a^2 & \frac{-a}{1-a} & \frac{a}{1-a} \\ \frac{-a}{1-a} & 1 & -1 \\ \frac{a}{1-a} & -1 & 1 \end{pmatrix} \text{ if } l_F > 0,$$

and $K_a^2 = 2a^2/((1-2a)(1-a))$.

In fact, $W_3 = -W_2$ in both Case I, and Case III with $l_F > 0$; and $W_3 = (1-a)(W_1 - W_2)$ in Case III with $l_F = 0$.

Note that the inequality $N^{s_n}(c_n) \leq N^e(c_n)$ (a consequence of the definitions) carries over (with probability one) to the limit r.v.'s as $\mathcal{N}^s \leq \mathcal{N}^e$ if $F(x)$ is in Case III with $l_F = 0$, and as $\mathcal{N}^s \leq 0$ if $F(x)$ is in Case I, or in Case III with $l_F > 0$. The latter inequality is clear since, in these cases, $B_n^{-1/2}(N^{s_n}(c_n) - B_n) \leq \delta_n A_n^{-1}(N^e(c_n) - B_n)$ where $\delta_n = c(B_n/n)/F^{\leftarrow}(B_n/n)$ for $F(x)$ in Case I, and $= (F^{\leftarrow}(B_n/n) - l_F)/l_F$ for $F(x)$ in Case III with $l_F > 0$, and $\delta_n \to 0$.

The proof of Theorem 3.1 is based on a Brownian bridge approximation to the uniform empirical process. Before Theorem 3.1 is proved, some lemmas concerning the Brownian bridge approximations are given. In the sequel we work on a probability space $(\Omega, \mathcal{A}, P)$ constructed by Csörgő et al. [6] carrying an infinite sequence $U_1, U_2, \cdots$ of i.i.d. r.v.'s uniformly distributed on $(0,1)$ and a sequence of Brownian bridges $U_n(s)$, $0 \leq s \leq 1$, $n = 1, 2, \cdots$ such that for the uniform empirical process $\alpha_n(s) = n^{1/2}(G_n(s) - s)$, $0 \leq s \leq 1$, where $G_n(s) = n^{-1}\sum_{i=1}^n I(U_i \leq s)$

$$\sup_{1/n \leq s \leq 1-1/n} n^{\nu} \frac{|\alpha_n(s) - U_n(s)|}{(s(1-s))^{(1/2)-\nu}} = O_P(1) \tag{22}$$

as $n \to \infty$, where ν is any fixed number such that $0 \leq \nu < 1/4$. This can be assumed without loss in generality. Note that for an F-distributed r.v. X_i we have $X_i \overset{\mathrm{d}}{=} F^{\leftarrow}(U_i)$, so on this space we use $X_i = F^{\leftarrow}(U_i)$ for $i = 1, 2, \cdots$. Recall that the constants $(B_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$ are defined in equation (19). The proofs of the following two lemmas use standard arguments found, for example, in the papers by Csörgő and Mason [8], Csörgő, Haeusler and Mason [10] and Lo [12].

Lemma 3.2 *Let $F(x)$ be in Case I or III for minima with $l_F \geq 0$ and let $(m_n)_{n\geq 1}$ be any sequence of positive integers such that $m_n \approx B_n$. Then*

$$\frac{\sum_{i=1}^{m_n} X_{i,n} - n\int_0^{m_n/n} F^{\leftarrow}(s)ds}{n^{1/2}A^e(m_n/n)} = -\frac{\int_{1/n}^{B_n/n} U_n(s)dF^{\leftarrow}(s)}{A^e(B_n/n)} + o_P(1)$$

where for $0 < s < 1$, $A^e(s) = \begin{cases} s^{1/2}c(s) & \text{in Case I} \\ s^{(1/2)-a}L(s) & \text{in Case III} \end{cases}$.

Lemma 3.3 *Let $F(x)$ be in Case I or III for minima with $l_F \geq 0$, and let $\tau_n := F(\varepsilon_n)$ for $n \geq 1$. Then*

$$(n\tau_n)^{-1/2}\left(\sum_{i=1}^n I(X_i \leq \varepsilon_n) - n\tau_n\right) = \tau_n^{-1/2}U_n(\tau_n) + o_P(1) = (B_n/n)^{-1/2}U_n(B_n/n) + o_P(1)$$

and

$$\begin{aligned}
&\frac{\sum_{i=1}^n X_i I(X_i \leq \varepsilon_n) - n\int_0^{\tau_n} F^{\leftarrow}(s)ds}{(n\tau_n)^{1/2}F^{\leftarrow}(\tau_n)} \\
&= \begin{cases} \tau_n^{-1/2}U_n(\tau_n) - \dfrac{\int_{1/n}^{\tau_n} U_n(s)dF^{\leftarrow}(s)}{\tau_n^{1/2}F^{\leftarrow}(\tau_n)} + o_P(1) & \text{in Case III with } l_F = 0 \\ \tau_n^{-1/2}U_n(\tau_n) + o_P(1) & \text{in Case I, and Case III with } l_F > 0 \end{cases}.
\end{aligned}$$

If, in addition, condition (20) holds as a hypotheses in Lemma 3.3, then the conclusions of Lemma 3.3 hold with τ_n replaced by B_n/n.

Proof of Theorem 3.1. Consider first $F(x)$ in Case III with $l_F = 0$. For the convergence analysis of $N^{s_n}(c_n)$, use condition (20) to observe that

$$\frac{N^{s_n}(c_n) - B_n}{B_n^{1/2}} = \frac{\sum_{i=1}^{\nu_n(c_n)} (I(X_i \leq \varepsilon_n) - F(\varepsilon_n))}{B_n^{1/2}} + B_n^{1/2}\left(\frac{\nu_n(c_n)}{n} - 1\right) + o_P(1)$$

where $\nu_n(c_n) = \max\{j : 1 \leq j \leq n \text{ and } \sum_{i=1}^{j} X_i I(X_i \leq \varepsilon_n) \leq c_n\}$ if this set is nonempty and $= 0$ otherwise; and obtain the following representations associated with this sum by using the argument in the proof of the Doeblin-Ascombe Central Limit Theorem as given in the book by Chow and Teicher [3, Theorem 1, page 317], the result that $\nu_n(c_n)/n \to 1$ in probability, and Lemma 3.3:

$$B_n^{-1/2} \sum_{i=1}^{\nu_n(c_n)} (I(X_i \leq \varepsilon_n) - F(\varepsilon_n)) = (B_n/n)^{-1/2} U_n(B_n/n) + o_P(1)$$

and

$$\frac{\sum_{i=1}^{l_n} X_i I(X_i \leq \varepsilon_n) - l_n c_n/n}{n^{1/2}(B_n/n)^{1/2} F^{\leftarrow}(B_n/n)} = (B_n/n)^{-1/2} U_n(B_n/n) - \frac{\int_{1/n}^{B_n/n} U_n(s)dF^{\leftarrow}(s)}{(B_n/n)^{1/2}F^{\leftarrow}(B_n/n)} + o_P(1)$$

where $l_n := \lceil n + (nz_2/B_n^{1/2})\rceil$ for $n \geq 1$, with $-\infty < z_2 < 0$. Also define constants $m_n := \lceil B_n + A_n\mu\rceil$ for $n \geq 1$, with $-\infty < \mu < \infty$, where the constants $(A_n)_{n\geq 1}$ are defined in the statement of the Theorem, and obtain from condition (21) and results on slowly varying functions in the Appendix that

$$\frac{c_n - n\int_0^{m_n/n} F^{\leftarrow}(s)ds}{n^{1/2}(B_n/n)^{(1/2)-a}L(B_n/n)} = -\mu + o(1) \text{ and } \frac{c_n - (l_n c_n/n)}{n^{1/2}(B_n/n)^{1/2}F^{\leftarrow}(B_n/n)} = -z_2/(1-a) + o(1).$$

Now use these representations and convergence results to obtain the weak convergence

$$\begin{aligned}
&P\left(\frac{N^e(c_n) - B_n}{A_n} < \mu, \frac{\sum_{i=1}^{\nu_n(c_n)} (I(X_i \leq \varepsilon_n) - F(\varepsilon_n))}{B_n^{1/2}} < z_1, B_n^{1/2}\left(\frac{\nu_n(c_n)}{n} - 1\right) < z_2\right) \\
&= P\left(\sum_{i=1}^{m_n} X_{i,n} > c_n, \frac{\sum_{i=1}^{\nu_n(c_n)} (I(X_i \leq \varepsilon_n) - F(\varepsilon_n))}{B_n^{1/2}} < z_1, \sum_{i=1}^{l_n} X_i I(X_i \leq \varepsilon_n) > c_n\right) \\
&= P\left(\frac{\int_{1/n}^{B_n/n} U_n(s)dF^{\leftarrow}(s)}{(B_n/n)^{(1/2)-a}L(B_n/n)} < \mu, (B_n/n)^{-1/2}U_n(B_n/n) < z_1,\right. \\
&\qquad \left.(1-a)\left(\frac{\int_{1/n}^{B_n/n} U_n(s)dF^{\leftarrow}(s)}{(B_n/n)^{(1/2)-a}L(B_n/n)} - (B_n/n)^{-1/2}U_n(B_n/n)\right) < z_2\right) + o(1) \\
&= P(W_1 < \mu, W_2 < z_1, W_3 < z_2) + o(1)
\end{aligned}$$

where $\mathbf{W} = (W_1, W_2, W_3) \stackrel{\text{d}}{=} N(\mathbf{0}, \Sigma_{\mathbf{W}})$ and the covariance matrix $\Sigma_{\mathbf{W}}$ is defined in the assertion of the Theorem.

Hence

$$P\left(\frac{N^e(c_n) - B_n}{A_n} \leq \mu, \frac{N^{s_n}(c_n) - B_n}{B_n^{1/2}} \leq \eta\right) = P\left(W_1 \leq \mu, W_2 + (W_3 \wedge 0) \leq \eta\right) + o(1)$$

and the Theorem is proved for $F(x)$ in Case III with $l_F = 0$. For the proof of the Theorem in the other cases, use straightforward modifications of this argument, together with the Lemmas 3.2 and 3.3 and background convergence results from the Appendix. $\square$

Example 3.4 For distribution function $F(x) = x^{-1/a}$, $0 \leq x \leq 1$, with $-\infty < a < 0$, and for positive constants $(c_n)_{n\geq 1}$ with $c_n/n \to 0$ and $c_n n^{-a} \to \infty$, it follows that $B_n = ((1-a)c_n n^{-a})^{1/(1-a)}$ and Theorem 3.1 gives

$$\left(B_n^{1/2}\left(\frac{N^e(c_n)}{B_n} - 1\right), B_n^{1/2}\left(\frac{N^{s_n}(c_n)}{B_n} - 1\right)\right) \Rightarrow (W_1, W_2 + ((1-a)(W_1 - W_2) \wedge 0))$$

where (W_1, W_2) is $N(\mathbf{0}, \Sigma_{\mathbf{W}})$-distributed with $\Sigma_{\mathbf{W}} = \begin{pmatrix} \frac{2a^2}{(1-a)(1-2a)} & \frac{-a}{1-a} \\ \frac{-a}{1-a} & 1 \end{pmatrix}$.

In particular, if $F(x)$ is the d.f. of a random variable uniformly distributed on $(0,1)$, so $F(x) = x$ for $0 \leq x \leq 1$, then $F(x)$ is in Case III for minima with $a = -1$. Moreover the sequences $(A_n)_{n\geq 1}$ and $(B_n)_{n\geq 1}$ are given by $B_n = (2nc_n)^{1/2}$ and $A_n \approx B_n^{1/2} = (2nc_n)^{1/4}$, and $(2nc_n)^{-1/4}\left(N^e(c_n) - (2nc_n)^{1/2}\right) \Rightarrow N(0, 1/3)$ and $(2nc_n)^{-1/4}\left(N^{s_n}(c_n) - (2nc_n)^{1/2}\right) \Rightarrow W_2 + (W_3 \wedge 0)$ where $(W_2, W_3) \stackrel{\text{d}}{=} N(\mathbf{0}, \begin{pmatrix} 1 & -1 \\ -1 & 4/3 \end{pmatrix})$. As a comparison, note that the mean and variance of the limit r.v. $\mathcal{N}^s$ is given by $E(W_2 + (W_3 \wedge 0)) = -(2/(3\pi))^{1/2}$ and $\mathrm{Var}(W_2 + (W_3 \wedge 0)) = 2(1 - 1/\pi)/3$, respectively.

Additional asymptotic distributional comparisons between $N^e(c_n)$ and $N^{s_n}(c_n)$ can be obtained from Theorem 3.1, as stated in the following corollaries. Analogous comparisons can be made between $(s_n)_{n\geq 1}$ and τ^e based on Theorem 4.1 (see, e.g. Corollary 4.2). The proof of the following corollary is immediate.

Corollary 3.5 *Let $F(x)$ be in Case I or III for minima with $l_F \geq 0$ and constants $(c_n)_{n\geq 1}$, $(B_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$ as in Theorem* 3.1. *Then*
(i) the sequence of stopped threshold policies $(s_n)_{n\geq 1}$ is a consistent approximator of the off-line smallest fit strategy τ^e;
(ii) $N^e(c_n)/N^{s_n}(c_n) \to 1$ in probability as $n \to \infty$; and
(iii) for $F(x)$ in Case III with $l_F = 0$,

$$B_n^{-1/2}(N^e(c_n) - N^{s_n}(c_n)) \Rightarrow \mathcal{N}^e - \mathcal{N}^s$$

and

$$(N^{s_n}(c_n) - B_n)/(N^e(c_n) - B_n) \Rightarrow \mathcal{N}^s/\mathcal{N}^e$$

where $\mathcal{N}^e$ and $\mathcal{N}^s$ are given in Theorem 3.1.

Corollary 3.6 *Let $F(x)$ be in Case I or Case III for minima with $l_F \geq 0$ and let the constants $(B_n)_{n\geq 1}$ and $(\varepsilon_n)_{n\geq 1}$ be given by* (19) *and satisfy condition* (20). *Then*

$$\begin{aligned}\kappa(a) &\leq \liminf_{n\to\infty} B_n^{-1/2}\left(EN^{s_n}(c_n) - B_n\right) \\ &\leq \liminf_{n\to\infty} B_n^{-1/2}\left(EN_n^*(c_n) - B_n\right) \leq \limsup_{n\to\infty} B_n^{-1/2}\left(EN_n^*(c_n) - B_n\right) \\ &\leq \limsup_{n\to\infty} B_n^{-1/2}\left(EN^e(c_n) - B_n\right) \leq 0\end{aligned} \tag{23}$$

where $\kappa(a) = -(1-a)^{1/2}$ in Case III with $l_F = 0$, and $= -1$ in Case III with $l_F > 0$ and Case I.

Proof. Only the first inequality and the last inequality in (23) need some explanation since $EN^{s_n}(c_n) \leq EN_n^*(c_n) \leq EN^e(c_n)$. For the last inequality in (23) use (10) to obtain

$$B_n^{-1/2}\left(N^e(c_n) - B_n\right) \leq B_n^{-1/2}\left(N_n' - B_n\right) + \frac{c_n - \sum_{i=1}^n X_i I(X_i \leq \varepsilon_n)}{B_n^{1/2}\varepsilon_n} \tag{24}$$

where $N_n' = \sum_{i=1}^n I(X_i \leq \varepsilon_n)$. Using an argument similar to that in the proof of Theorem 3.3 of Bruss and Robertson [2], and the convergence results from Lemma 3.3, it is easy to see that $B_n^{-1/2}\left(N_n' - B_n\right) \Rightarrow N(0,1)$, $B_n^{-1/2}\varepsilon_n^{-1}\left(c_n - \sum_{i=1}^n X_i I(X_i \leq \varepsilon_n)\right) \Rightarrow N(0,\sigma_a^2)$ where $\sigma_a^2 = (1-2a)^{-1}$ in Case III with $l_F = 0$, and $= 1$ in the other cases, and that the sequences $\{B_n^{-1/2}\left(N_n' - B_n\right)\}$ and $\{B_n^{-1/2}\varepsilon_n^{-1}\left(c_n - \sum_{i=1}^n X_i I(X_i \leq \varepsilon_n)\right)\}$ are uniformly integrable. This settles the last inequality in (23).

To see the first inequality in (23) use inequality (3) in Rhee and Talagrand [14] to obtain $B_n^{-1/2}\left(EN^{s_n}(c_n) - B_n\right) \geq -B_n^{-1/2} - c_n^{-1/2}\left(B_n F^{\leftarrow}(B_n/n)\right)^{1/2} + o(1)$. By Karamata's Theorem (see the Appendix), the right hand side of the latter inequality converges to $-(1-a)^{1/2}$ in Case III with $l_F = 0$. The other cases follow similarly by the convergence results from the Appendix. □

Example 3.7 This example shows the necessity of condition (20) for Theorem 3.1. Let $F(x)$ be given by $F(x) = k^{-3}$ if $(k+1)^{-1} \leq x < k^{-1}$, for $k = 1, 2, \cdots$; so d.f. $F(x)$ is in Case III for minima with $l_F = 0$, for $\alpha = 3$ and $a = -1/3$. Its inverse function is given by $F^{\leftarrow}(w) = (k+1)^{-1}$ if $(k+1)^{-3} < w \leq k^{-3}$, $k = 1, 2\cdots$. Let $k_n := [\log n]$ for $n \geq 1$ and $(\gamma_n)_{n\geq 1}$ be constants $0 < \gamma_n < 1$ with $\gamma_n \uparrow 1$, and let $(c_n)_{n\geq 1}$ be the positive constants satisfying

$$\begin{aligned}c_n/n &= \int_0^{(k_n+\gamma_n)^{-3}} F^{\leftarrow}(s)ds \\ &= \sum_{l=k_n+1}^{\infty} l^{-1}((l-1)^{-3} - l^{-3}) - (k_n+1)^{-1}(k_n^{-3} - (k_n+\gamma_n)^{-3}) \\ &= (k_n+\gamma_n)^{-3}(k_n+1)^{-1} - \sum_{l=k_n+1}^{\infty} l^{-4}(l+1)^{-1} \text{ for } n \geq 1.\end{aligned}$$

Thus, $c_n/n \approx (3/4)(\log n)^{-4}$ as $n \to \infty$. These constants $(c_n)_{n\geq 1}$ were chosen so that the constants $(B_n)_{n\geq 1}$ of (19) are given by $B_n = n(k_n+\gamma_n)^{-3}$ for $n \geq 1$. Thus $B_n \approx n/(\log n)^3$ as

$n \to \infty$. Also, the constants $(\varepsilon_n)_{n\geq 1}$ of (19) are given by $\varepsilon_n = (k_n+1)^{-1}$ for $n \geq 1$. It follows that for this d.f. $F(x)$ and this sequence $(c_n)_{n\geq 1}$,

$$B_n^{1/2}\left(\frac{F(\varepsilon_n)}{B_n/n}-1\right) = n^{1/2}(k_n+\gamma_n)^{-3/2}\left(\left(\frac{k_n+\gamma_n}{k_n}\right)^3-1\right) \approx 3n^{1/2}k_n^{-5/2} \to \infty \text{ as } n \to \infty.$$

Using the method of proof of Theorem 3.1 and direct calculation, one can show that for this example both $\nu_n(c_n)/n \to 1$ and $(F(\varepsilon_n)\nu_n(c_n) - B_n)/B_n^{1/2} \to -\infty$ in probability as $n \to \infty$, and so $B_n^{-1/2}(N^{s_n}(c_n) - B_n) \to -\infty$ in probability as $n \to \infty$. In contrast, as in Theorem 3.1, $B_n^{-1/2}(N^e(c_n) - B_n) \Rightarrow \mathcal{N}^e$, where $\mathcal{N}^e$ is given in the statement of Theorem 3.1, for $F(x)$ in Case III for minima with $l_F = 0$.

Example 3.8 Consider d.f. $F(x) = x - 1$ for $1 \leq x \leq 2$, so that $F(x)$ is in Case III for minima with $l_F = 1$, and has inverse function $F^{\leftarrow}(s) = 1 + s$ for $0 \leq s \leq 1$. Let $(B_n)_{n\geq 1}$ and $(c_n)_{n\geq 1}$ be positive constants for which $B_n \to \infty$, $B_n/n \to 0$ and $c_n = B_n(1 + (B_n/(2n)))$ for $n \geq 1$. If $B_n^{3/2}/n \to \infty$, then Theorem 3.1 gives that $(N^e(c_n) - B_n)/(B_n^{3/2}/n) \Rightarrow N(0, 1/3)$-distributed r.v. However, if $B_n^{3/2}/n \not\to \infty$, no such normal convergence holds, and asymptotic behavior of $\{N^e(c_n)\}$ may vary from one subsequence of n's to another. For example, consider the specific d.f. $F(x)$ of this example, with $B_n = n^{1/2}$ for $n \geq 1$ (so $A_n = B_n^{3/2}/n = n^{-1/4}$), and let $\mu < 0$. If $n = m^2$, for all positive integers m sufficiently large, then $P(n^{1/4}(N^e(c_n) - n^{1/2}) < \mu) \to 1/2$; but for $0 < \delta < 1$ small, and n's satisfying $(m+\delta)^2 < n < (m+1-\delta)^2$, for all positive integers m sufficiently large, then $P(n^{1/4}(N^e(c_n) - n^{1/2}) < \mu) \to 1$. To see this holds, use an argument analogous to part of the proof of Theorem 3.1, together with Lemma 3.2, and careful analysis of the sequence $\left\{(c_n - n\int_0^{m_n/n} F^{\leftarrow}(s)ds)/(n^{1/2}(B_n/n)^{(1/2)-a})\right\}$.

4 Numbers Selected Proportional to Increasing Sample Sizes

In this section, settings are considered in which $N^e(c_n)/n \to \tau$ and $N^{s_n}(c_n)/n \to \tau$, for $0 < \tau < \infty$, as the sample size $n \to \infty$. Results are proved on the weak convergence of appropriate normalizations of $N^e(c_n)$ and $N^{s_n}(c_n)$ to nondegenerate r.v.'s as $n \to \infty$.

In this section, *the sequence of capacity sizes $(c_n)_{n\geq 1}$ are assumed to satisfy the large capacity property $n^{-1/2}(c_n - n\theta) \to 0$, for $0 < \theta < EX_1$*, where $X_1, X_2, \cdots$ is the sequence of object sizes (nonnegative i.i.d. r.v.'s with d.f. $F(x)$). Throughout the section, the constants τ and ε satisfy $\theta = \int_0^\tau F^{\leftarrow}(s)ds$ and $\varepsilon = F^{\leftarrow}(\tau)$. The constant ε is used as a threshold for the stopped threshold policy s_n. For this theorem, the inequality $N^{s_n}(c_n) \leq N^e(c_n)$ carries over to inequality $\mathcal{N}^s \leq \mathcal{N}^e$ (this is easily checked from the conclusion, by using $0 < \theta \leq \tau\varepsilon$).

Theorem 4.1 *If $F(x)$ is continuous and strictly increasing on its support, $l_F \geq 0$ and $EX_1^2 < \infty$, then*

$$\left(n^{-1/2}\left(N^e(c_n) - n\tau\right), n^{-1/2}\left(N^{s_n}(c_n) - n\tau\right)\right) \Rightarrow (\mathcal{N}^e, \mathcal{N}^s)$$

where $(\mathcal{N}^e, \mathcal{N}^s) = (W_1, W_2 + (W_3 \wedge 0))$ and (W_1, W_2, W_3) is $N(\mathbf{0}, \Sigma(\theta))$-distributed. In fact, $W_3 = ((\tau\varepsilon)/\theta)(W_1 - W_2)$. The covariance matrix $\Sigma(\theta)$ is a symmetric matrix $\Sigma(\theta) = (\Sigma(\theta)_{i,j})_{i=1,2,3;\, j=1,2,3}$

given by

$$
\begin{aligned}
(\Sigma(\theta))_{i=1,\,j=1,2,3} &= \Big(\sigma^2(\tau)/\varepsilon^2, (1-\tau)(\tau-(\theta/\varepsilon))\,, \\
&\qquad\qquad \theta^{-1}\tau(\sigma^2(\tau)/\varepsilon-(1-\tau)(\tau\varepsilon-\theta))\Big)\,; \\
(\Sigma(\theta))_{i=2,\,j=2,3} &= (\tau(1-\tau), -\tau(1-\tau))\,; \\
(\Sigma(\theta))_{i=3,\,j=3} &= \theta^{-2}\tau^2\Big(\sigma^2(\tau)+2\varepsilon(1-\tau)\theta-\varepsilon^2\tau(1-\tau)\Big)\,;\ \textit{and} \\
\sigma^2(\tau) &= \int_0^\tau\int_0^\tau (s\wedge t-st)dF^{\leftarrow}(s)dF^{\leftarrow}(t).
\end{aligned}
$$

Proof. First, observe that

$$
n^{-1/2}\left(N^{s_n}(c_n)-n\tau\right)=n^{-1/2}\sum_{i=1}^{\nu_n(c_n)}\left(I(X_i\le\varepsilon)-\tau\right)+\tau n^{1/2}\left(\frac{\nu_n(c_n)}{n}-1\right).
$$

Follow the same lines as the proof of Theorem 3.1 and use standard approximation arguments (e.g. given by Csörgő, Csörgő and Horváth [7, Chapter 10]) to obtain

$$
\begin{aligned}
&P\left(n^{-1/2}\left(N^{e}(c_n)-n\tau\right)<\mu; n^{-1/2}\sum_{i=1}^{\nu_n(c_n)}\left(I(X_i\le\varepsilon)-\tau\right)<z_1; \tau n^{1/2}\left(\frac{\nu_n(c_n)}{n}-1\right)<z_2\right) \\
&= P\left(\frac{\int_{1/n}^{\tau}U_n(s)dF^{\leftarrow}(s)}{F^{\leftarrow}(\tau)}<\mu; U_n(\tau)<z_1\,;\right. \\
&\qquad\qquad \left.\frac{\tau}{\theta}\left(\int_{1/n}^{\tau}U_n(s)dF^{\leftarrow}(s)-F^{\leftarrow}(\tau)U_n(\tau)\right)<z_2\right)+o(1).
\end{aligned}
$$

In order to compute the covariance matrix $\Sigma(\theta)$, use $\int_0^\tau s dF^{\leftarrow}(s)=\tau F^{\leftarrow}(\tau)-\theta$. This completes the proof of the theorem. □

Corollary 4.2 *If $F(x)$ is continuous and strictly increasing on its support, $l_F\ge 0$, and $EX_1^2<\infty$, then*
(i) the sequence of stopped threshold policies $(s_n)_{n\ge1}$ is a consistent approximator of τ^e;
(ii) $(N^{s_n}(c_n)-n\tau)/(N^e(c_n)-n\tau)\Rightarrow \mathcal{N}^s/\mathcal{N}^e$; and
(iii) $n^{-1/2}(N^e(c_n)-N^{s_n}(c_n))\Rightarrow\mathcal{N}^e-\mathcal{N}^s$
where $(\mathcal{N}^e,\mathcal{N}^s)$ is given in Theorem 4.1.

Corollary 4.3 *If $c_n=n\theta$ for some $\theta>0$, and $F(x)$ is continuous and strictly increasing on its support, $l_F\ge0$, and $EX_1^2<\infty$, then*

$$
\begin{aligned}
-(2\pi)^{-1/2}\left(\Sigma(\theta)_{3,3}\right)^{1/2} &= \lim_{n\to\infty} n^{-1/2}\left(EN^{s_n}(c_n)-n\tau\right) \\
&\le \liminf_{n\to\infty} n^{-1/2}\left(EN_n^*(c_n)-n\tau\right)\le\limsup_{n\to\infty}n^{-1/2}\left(EN_n^*(c_n)-n\tau\right) \\
&\le \lim_{n\to\infty} n^{-1/2}\left(EN^{e}(c_n)-n\tau\right)=0
\end{aligned}
$$

and

$$
\lim_{n\to\infty} n^{-1/2}E\left(N^e(c_n)-N^{s_n}(c_n)\right)=(2\pi)^{-1/2}\left(\Sigma(\theta)_{3,3}\right)^{1/2}
$$

where $\Sigma(\theta)_{3,3}$ is defined in Theorem 4.1.

Proof. The proof of the corollary relies on the inequalities

$$\begin{aligned} n^{-1/2}\left(N^{s_n}(c_n)-n\tau\right) &\le n^{-1/2}\left(N^{e}(c_n)-n\tau\right) \\ &\le n^{-1/2}\left(N'_n-n\tau\right)+n^{-1/2}\varepsilon^{-1}\left(c_n-\sum_{i=1}^{n}X_iI(X_i\le\varepsilon)\right) \end{aligned} \tag{25}$$

where $N'_n=\sum_{i=1}^{n}I(X_i\le\varepsilon)$. The last inequality in (25) is proved in exactly the same manner as inequality (24). Since $n^{-1/2}\left(N^{s_n}(c_n)-n\tau\right)\Rightarrow\mathcal{N}^s$, $n^{-1/2}\left(N^{e}(c_n)-n\tau\right)\Rightarrow\mathcal{N}^e$ with $(\mathcal{N}^e,\mathcal{N}^s)$ from the assertion of Theorem 4.1, $E\mathcal{N}^s=-(2\pi)^{-1/2}\left(\Sigma(\theta)_{3,3}\right)^{1/2}$ and $E\mathcal{N}^e=0$, the proof of the theorem is completed if it is shown that the sequences $\left\{n^{-1/2}\left(N^{s_n}(c_n)-n\tau\right)\right\}$, $\left\{n^{-1/2}\left(N'_n-n\tau\right)\right\}$ and $\left\{n^{-1/2}\varepsilon^{-1}\left(c_n-\sum_{i=1}^{n}X_iI(X_i\le\varepsilon)\right)\right\}$ are uniformly integrable. Following an argument of Bruss and Robertson (see [2, Theorem 3.3]) it is enough to show that $P(|V_n|>v)\le Cv^{-2}$ for some constant C not depending on n, where $\{V_n\}$ represents each of the sequences above. We only consider the case $V_n=n^{-1/2}\left(N^{s_n}(c_n)-n\tau\right)$ since the other cases are straightforward.

By Chebychev's inequality

$$\begin{aligned} &P(|V_n|>v)\le v^{-2}E|V_n|^2 \\ &\le\ 2v^{-2}\left(n^{-1}E\left(\sum_{i=1}^{\nu_n(c_n)}\left(I(X_i\le\varepsilon)-F(\varepsilon)\right)\right)^2+nE\left((\nu_n(c_n)/n)-1\right)^2\right)=:2v^{-2}(A+B) \end{aligned}$$

where $\nu_n(c_n)=\max\{j:1\le j\le n \text{ and } \sum_{i=1}^{j}X_iI(X_i\le\varepsilon)\le c_n\}$. Now

$$A\le 2n^{-1}\left(E\left(\sum_{i=1}^{\tilde\nu_n(c_n)}\chi_i\right)^2+E\left(\sum_{i=1}^{\tilde\nu_n(c_n)}\chi_i-\sum_{i=1}^{\nu_n(c_n)}\chi_i\right)^2\right)=:2n^{-1}(A_1+A_2)$$

where $\chi_i=I(X_i\le\varepsilon)-F(\varepsilon)$ and $\tilde\nu_n(c_n)=(\nu_n(c_n)+1)\wedge n=\min\{j:1\le j\le n \text{ and } \sum_{i=1}^{j}X_iI(X_i\le\varepsilon)>c_n\}$. Note that $\tilde\nu_n(c_n)$ is a stopping time (for $(X_j)_{j\ge1}$) and hence by Wald's identity (see [3, Theorem 3, page 139]), $A_1=E(\tilde\nu_n(c_n))\mathrm{Var}\left(I(X_1\le\varepsilon)\right)\le n\tau(1-\tau)$. It is easy to see that $A_2\le n\tau(1-\tau)$ and thus $A\le4\tau(1-\tau)$.

Consider B. Define $\bar\nu_n=\sup\{j\ge1:\sum_{i=1}^{j}Z_i\le n\}$ where $Z_i=(X_i/\theta)I(X_i\le\varepsilon)$ and observe that $B\le n^{-1}E(\bar\nu_n-n)^2$ since $c_n=n\theta$. From renewal theory (see Gut [11, Theorem II.5.2 (ii)]) it is known that $E\bar\nu_n=\frac{n}{\mu}+\frac{\sigma^2-\mu^2}{2\mu^2}+o(1)$ and $\mathrm{Var}\,\bar\nu_n=n\frac{\sigma^2}{\mu^3}+o(n)$, as $n\to\infty$, with $\sigma^2=\mathrm{Var}Z_1$ and $\mu=EZ_1=1$. Hence $B\le n^{-1}E(\bar\nu_n-n)^2=\sigma^2+o(1)\le\tau\varepsilon^2/\theta^2+o(1)$. This completes the proof of the theorem. □

A Appendix

Recall the following definitions, relations, and results concerned with domains of attraction for minima.

Let $X_1,X_2,\cdots$ be i.i.d. r.v.'s with d.f. F. $F(x)$ is said to be in the domain of attraction of d.f. $H(x)$ for minima if there exists constants $(a_n)_{n\ge1}$, $a_n>0$ and $(b_n)_{n\ge1}$, $b_n\in\mathbb{R}$, such that

$$P\left(a_n\left(\min_{1\le i\le n}X_i-b_n\right)\le x\right)\to H(x) \text{ for all continuity points } x \text{ of } H. \tag{26}$$

In this setting we say F is in Case I, in Case II with parameter $\alpha > 0$ or in Case III with parameter $\alpha > 0$, if the limit d.f. H is respectively given by $H_{\mathrm{I}}(x) = 1 - \exp(-e^x)$ for $x \in \mathbb{R}$; by $H_{\mathrm{II}}^\alpha(x) = 1 - \exp(-(-x)^{-\alpha})$ for $x < 0$; or by $H_{\mathrm{III}}^\alpha(x) = 1 - \exp(-x^\alpha)$ for $x > 0$.

Let $Y_i = -X_i$ for $i \geq 1$, so that $Y_1, Y_2, \cdots$ are i.i.d. r.v.'s with d.f. $G(x) = 1 - F((-x)-)$, and $r_G := \sup\{x : G(x) < 1\} = -l_F$. F is in the domain of attraction for minima in Case I, II or III iff respectively G is in the domain of attraction for maxima associated with d.f.'s Λ, Φ_α, or Ψ_α. See Resnick [13] for results on domains of attraction for maxima and definitions of these d.f.'s. In the following paragraphs we list some properties in each of the Cases I, II, and III which are used in this paper.

- If F is in Case I for minima, then constants $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ are given by

$$b_n = F^{\leftarrow}(1/n) \text{ and } a_n^{-1} = g(b_n) \text{ for } n \geq 1, \tag{27}$$

  for example with $g(t) = \int_{l_F}^t F(x)dx/F(t)$. The function $F^{\leftarrow}(s)$ is slowly varying at zero. An auxiliary function useful in analysis is the function $c(s)$ defined on $[0,1]$ by $c(s) := s^{-1}\int_0^s u dF^{\leftarrow}(u) = F^{\leftarrow}(s) - s^{-1}\int_0^s F^{\leftarrow}(u)du$. As proven by Lo [12], the function $c(s)$ satisfies the following properties. There exists a finite constant k such that for $0 < s \leq 1/2$, $F^{\leftarrow}(s) = k + c(s) - \int_s^1 u^{-1}c(u)du$. The function $\phi(s) := \int_0^s F^{\leftarrow}(u)du$ can be written as $\phi(s) = s\left(k - \int_s^1 u^{-1}c(u)du\right)$. Also, $c(s) > 0$; $c(s)$ is slowly varying at zero; $\lim_{s\downarrow 0} c(s)/F^{\leftarrow}(s) = 0$; and if l_F is finite, then $\lim_{s\downarrow 0} c(s) = 0$.

- If F is in Case II for minima, with $\alpha > 0$, and $a = 1/\alpha$, then

$$a_n^{-1} = -F^{\leftarrow}(1/n) \text{ and } b_n = 0 \text{ for } n \geq 1, \tag{28}$$

  $l_F = -\infty$ and $\lim_{n\to\infty} nF(a_n^{-1}x) = (-x)^{-\alpha}$ for $x < 0$. The function $F(-x)$ is regularly varying as $x \to \infty$ with index $-\alpha = -1/a$; and $F^{\leftarrow}(s) = -s^{-a}L(s)$ where L is slowly varying at zero.

- If F is in Case III for minima, with $\alpha > 0$, and $a = -1/\alpha$, then

$$a_n^{-1} = F^{\leftarrow}(1/n) - l_F \text{ and } b_n = l_F \text{ for } n \geq 1, \tag{29}$$

  l_F is finite and $\lim_{n\to\infty} nF(a_n^{-1}x + l_F) = x^\alpha$ for $x > 0$. The function $F(l_F + x^{-1})$ is regularly varying as $x \to \infty$ with index $-\alpha = 1/a$; and $F^{\leftarrow}(s) = l_F + s^{-a}L(s)$ where L is slowly varying at zero.

For definitions and results on Poisson random measures and related material, an excellent reference is Chapters 3 and 4 of Resnick [13]. In particular, see [13] for definitions and results on random measures associated with partial maxima. In the following, for use in Section 2 and for comparison purposes, notation and results associated with partial minima from this theory are given. For $i = \mathrm{I}, \mathrm{II}, \mathrm{III}$, let $N = N_i^\alpha$ denote the Poisson random measure with mean measure $d\mu_i^\alpha = dt \times d\nu_i^\alpha$ on $[0,\infty) \times E_i$ for $i = \mathrm{I}, \mathrm{II}$ and III where

$$\begin{aligned} E_{\mathrm{I}} = [-\infty,\infty) &\quad \text{and} \quad \nu_{\mathrm{I}}[-\infty,x] = e^x \text{ for } x \in \mathbb{R}; \\ E_{\mathrm{II}} = [-\infty,0) &\quad \text{and} \quad \nu_{\mathrm{II}}^\alpha[-\infty,x] = (-x)^{-\alpha} \text{ for } x < 0; \text{ and} \\ E_{\mathrm{III}} = [0,\infty) &\quad \text{and} \quad \nu_{\mathrm{III}}^\alpha[0,x] = x^\alpha \text{ for } x > 0, \end{aligned}$$

for $\alpha > 0$. The random measure N takes values in $M_p = M_p([0,\infty) \times E_i)$, the space of point measures on $[0,\infty) \times E_i$, and has representation $N = \sum_{j=1}^{\infty} \epsilon_{(T_j, Z_j)}$, where ϵ_x denotes point mass at x, and T_j and Z_j are r.v.'s taking values in $[0,\infty)$ and E_i respectively. The space M_p is given the usual measurable structure and the vague topology (see, e.g., Chapter 3 of [13]); convergence in the vague topology is denoted $\xrightarrow{v}$. The analogue of Corollary 4.19 of [13] for partial minima, used in Section 2, has the following form:

> Let d.f. $F(x)$ be in the domain of attraction for minima of $H(x)$, and with norming constants of (26) chosen as in (27), (28) and (29). Then
> for $H = H_{\mathrm{I}}$, (26) is equivalent to $N_n := \sum_{j=1}^{\infty} \epsilon_{(j/n, a_n(X_j - b_n))} \Rightarrow N_{\mathrm{I}}$;
> for $H = H_{\mathrm{II}}^{\alpha}$, assume $F(0) = 1$, then (26) is equivalent to
> $N_n := \sum_{j=1}^{\infty} \epsilon_{(j/n, a_n X_j)} \Rightarrow N_{\mathrm{II}}^{\alpha}$; and
> for $H = H_{\mathrm{III}}^{\alpha}$, (26) is equivalent to $N_n := \sum_{j=1}^{\infty} \epsilon_{(j/n, a_n(X_j - b_n)} \Rightarrow N_{\mathrm{III}}^{\alpha}$
> where the weak convergence takes place in $M_p([0,\infty) \times E_i)$
> for $i =$ I, II or III respectively. (30)

Finally we state a result concerning functions which are slowly varying at zero, which we frequently use in this paper.

- (Karamata's Theorem (see [8, Lemma 1]).) Let $L(x)$ be slowly varying at zero. If $\beta < 1$, then
$$\lim_{s \downarrow 0} \int_0^s u^{-\beta} L(u) du \Big/ \left(s^{1-\beta} L(s) \right) = \frac{1}{1-\beta}.$$
- (Karamata Representation (see e.g. [13, Section 0.4]).) For $0 < t < t_0$, $L(t)$ is slowly varying at zero if and only if $L(t) = c(t) \exp\left(\int_t^{t_0} \varepsilon(u)/u \, du \right)$ for some measurable functions $c : (0, t_0) \to \mathbb{R}_+$ and $\varepsilon : (0, t_0) \to \mathbb{R}$ satisfying $\lim_{t \downarrow 0} c(t) = c_0$ for some constant $c_0 > 0$ and $\lim_{t \to 0} \varepsilon(t) = 0$.

Acknowledgment. The authors are grateful to the School of Mathematics of the Georgia Institute of Technology in Atlanta, the Econometric Institute of the Erasmus University Rotterdam and the Stieltjes Institute in The Netherlands for reciprocal invitations to pursue this research, and for their hospitality.

References

[1] Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

[2] Bruss, F.T. and Robertson, J.B. (1991) 'Wald's lemma' for sums of order statistics of i.i.d. random variables. *Adv. Appl. Prob.* **23**, 612-623.

[3] Chow, Y.S. and Teicher, H. (1978) *Probability Theory: Independence, Interchangeability, Martingales.* Springer-Verlag, New York.

[4] Chung, K.L. (1974) *A Course in Probability Theory, 2nd edition.* Academic Press, New York.

[5] Coffman, E.G., Flatto, L. and Weber, R.R. (1987) Optimal selection of stochastic intervals under a sums constraint. *Adv. Appl. Prob.* **19**, 454-473.

[6] Csörgő, M., Csörgő, S., Horváth, L. and Mason, D.M. (1986) Weighted empirical and quantile processes. *Ann. Probab.* **14**, 31-85.

[7] Csörgő, M., Csörgő, S. and Horváth, L. (1986) *An Asymptotic Theory for Empirical Reliability and Concentration Processes.* Lecture Notes in Statistics **33**. Springer-Verlag, Berlin.

[8] Csörgő, S. and Mason, D.M. (1986) The asymptotic distribution of sums of extreme values from a regularly varying distribution. *Ann. Probab.* **14**, 974-983.

[9] Csörgő, S., Horváth, L. and Mason, D.M. (1986) What portion of the sample makes a partial sum asymptotically stable or normal? *Probab. Th. Rel. Fields* **72**, 1-16.

[10] Csörgő, S., Haeusler, E. and Mason, D. (1991) The asymptotic distribution of extreme sums. *Ann. Probab.* **19**, 783-811.

[11] Gut, A. (1988) *Stopped Random Walks: Limit Theorems and Applications.* Springer-Verlag, New York.

[12] Lo, G.S. (1989) A note on the asymptotic normality of extreme values. *J. Statist. Planning and Inference* **22**, 127-136.

[13] Resnick, S. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag, New York.

[14] Rhee, W. and Talagrand, M. (1991) A note on the selection of random variables under a sum constraint. *J. Appl. Prob.* **28**, 919-923.

[15] Whitt, W. (1980) Some useful functions for functional limit theorems. *Math. Operat. Res.* **5**, 67-85.