THE INSTITUTE OF STATISTICS

THE CONSOLIDATED UNIVERSITY
OF NORTH CAROLINA

PRECISION CALCULATION OF DISTRIBUTIONS

FOR TRIMMED SUMS

by

Sándor Csörgő and Gordon Simons

March 1994

Mimeo Series #2317

PRECISION CALCULATION OF DISTRIBUTIONS FOR TRIMMED SUMS*

By SÁNDOR CSÖRGŐ and GORDON SIMONS

University of Michigan and University of North Carolina at Chapel Hill

Recursive methods are described for computing the frequency and distribution functions of trimmed sums of independent and identically distributed nonnegative integer-valued random variables. Surprisingly, for fixed arguments, these can be evaluated with just a *finite* number of arithmetic operations (and whatever else it takes to evaluate the common frequency function of the original summands). These methods give rise to very accurate computational algorithms which permit a delicate numerical investigation, herein described, of Feller's weak law of large numbers for repeated St. Petersburg games.

1. Introduction. Trimmed sums of iid (independent and identically distributed) random variables appear in many contexts. Applied statisticians use them to improve estimators when the parent distribution has a heavy tail. (See, for instance, David (1981), pp. 158–163; and for a list of relevant references, see Stigler (1973).) Probabilists, who have studied them extensively, have clearly documented the heavy influence, in some settings, of the largest observation(s). (Extensive reference lists can be found in a recent book edited by Hahn, Mason and Weiner (1991).)

We are concerned here with describing effective recursive methods for computing the frequency and distribution functions for trimmed sums of iid nonnegative integer-valued random variables. Not only do such methods exist, but, as we shall see, there exist methods that can be fully implemented with just a *finite* number of arithmetic operations. The model for this is provided by the convolution-based recursion for untrimmed sums

$$(1) \qquad P\{S_n = s\} = \sum_{k=0}^{s} P\{X = k\} \, P\{S_{n-1} = s - k\},$$

and by the simple sum

$$(2) \qquad P\{S_n \le s\} = \sum_{k=0}^{s} P\{S_n = k\}, \quad s = 0, 1, 2, \dots,$$

where $S_n := X_1 + \cdots + X_n$ is the sum of n iid random variables distributed as X. Both sums, (1) and (2), contain just a finite number of terms.

Turning to trimmed sums, let $S_n(m)$ denote the same sum but with the m largest summands excised, $m \le n$, i.e. let

$$S_n(m) := X_{n,1} + X_{n,2} + \cdots + X_{n,n-m}, \quad m = 0, 1, 2, \dots, n,$$

where $X_{n,1} \le X_{n,2} \le \dots \le X_{n,n}$ denote the order statistics for $X_1, X_2, \dots, X_n$, so that $S_n(0) = S_n$. (Throughout, $S_0 := 0$ and $S_n(n) := 0$.) While, clearly, the distribution function $P\{S_n(m) \le s\}$ can be obtained by a finite sum, as in (2), the frequency function $P\{S_n(m) = s\}$, for $m \ge 1$, can not be computed via a simple analogue of (1).

Consider the special case $m = 1$. A simple recursion in n, described in Theorem 3 below, *requiring nothing but finite sums*, links the functions $P\{S_n(1) = s, X_{n,n} = t\}$, $s, t = 0, 1, 2, \dots$, $n \ge 1$. But this approach leads to an *infinite* sum

$$(3) \qquad P\{S_n(1) = s\} = \sum_{t=0}^{\infty} P\{S_n(1) = s, X_{n,n} = t\}.$$

Fortunately, this shortcoming can be finessed with an application of the following.

THEOREM 1. *For integers $n \ge m \ge 0$ and $r \ge s \ge 0$,*

$$(4) \qquad P\{S_n(m) = s, X_{n,n} \le r\} = \sum_{k=0}^{m} (-1)^k \binom{n}{k} [1 - F(r)]^k P\{S_{n-k}(m-k) = s\},$$

where $F(r) := P\{X \le r\}$ is the distribution function of X.

Combining (3) and (4), for the case $m = 1$, leads to the finite sum: for integers $n \ge 1$ and $r \ge s \ge 0$,

$$(5) \qquad P\{S_n(1) = s\} = \sum_{t=0}^{r} P\{S_n(1) = s, X_{n,n} = t\} + n[1 - F(r)] P\{S_{n-1} = s\}.$$

2

The same trick works for a general $m \leq n$. For instance, for $n \geq m = 2$ and $r \geq s \geq 0$,

$$
P\{S_n(2) = s\} = \sum_{v=0}^{r}\sum_{u=0}^{v} P\{S_n(2) = s, X_{n,n-1} = u, X_{n,n} = v\}
$$

(6)

$$
+ n[1 - F(r)]\, P\{S_{n-1}(1) = s\} - \frac{n(n-1)}{2}[1 - F(r)]^2 P\{S_{n-2} = s\}.
$$

We readily concede the point to any critic who would argue, at this point, that it is possible, with proper care, to throw away an infinite number of small summands without introducing a substantial amount of error. This is true. But we would make three rejoinders:

1) Precise computations are more easily achieved when the issue of truncation does not arise (or is circumvented).

2) The exercise of "proper care" with a formula like (3) requires more memory, than with (5), to accomplish comparable accuracy. This can be a significant issue.

3) Recursive methods tend to propagate errors. Thus the "proper care" sufficient to handle the case $n = 10$, for instance, might not be adequate when the same calculations are extended to $n = 100$.

We have used the methods described herein in various settings. Based on considerable experience, we feel quite confident that they yield excellent results, when performed with double precision arithmetic, even when n assumes values in the low thousands. For example, we have obtained essentially identical results when using equation (5) with various values of r, the free choice of which is a potential means for checking accuracy. But a careful error analysis has not been made.

The presence of mixed signs on the right of (6) can be overcome, in order to avoid potential losses of computational accuracy, by applying (5) to the next to last term in (6) to obtain

$$
P\{S_n(2) = s\} = \sum_{v=0}^{r}\sum_{u=0}^{v} P\{S_n(2) = s, X_{n,n-1} = u, X_{n,n} = v\}
$$

(7)

$$
+ n[1 - F(r)] \sum_{t=0}^{r} P\{S_{n-1}(1) = s, X_{n-1,n-1} = t\}
$$

$$
+ \frac{n(n-1)}{2}[1 - F(r)]^2 P\{S_{n-2} = s\}.
$$

Everything on the right side of (7), other than the factor $1 - F(r)$, can be evaluated *without subtracting terms of positive sign.* (A check of the recursion described in Theorem

3 below is required to verify this assertion.) Formula (7) is also a simple consequence, for the case $m = 2$, of Theorem 2 below, which, in a sense, inverts equation (4).

THEOREM 2. *For integers $n \geq m \geq 0$ and $r \geq s \geq 0$,*

$$(8) \quad P\{S_n(m) = s\} = \sum_{k=0}^{m} \binom{n}{k} [1 - F(r)]^k P\{S_{n-k}(m-k) = s, X_{n-k,n-k} \leq r\} \,.$$

Section 2 gives a proof of Theorems 1 and 2, and describes in Theorem 3 a recursion for general $m \geq 1$ which yields intermediate probabilities such as $P\{S_n(1) = s, X_{n,n} = t\}$ and $P\{S_n(2) = s, X_{n,n-1} = u, X_{n,n} = v\}$ in (5) and (7). Section 3 discusses an illustrative application, which is of relevance to an ongoing study by the authors of the "St. Petersburg Paradox".

2. Theory.

Here, we assume the notation appearing in the introduction. We begin with the proofs of Theorems 1 and 2. Then we describe in Theorem 3 an essential recursion, for general m, which, with the recursion in either Theorem 1 or 2, leads to a general scheme, in the spirit of (6) or (7), for computing $P\{S(m) = s\}$, for any integer $s \geq 0$, with just a finite number of arithmetic operations.

PROOF OF THEOREM 1. There is really nothing to prove when $m = 0$; equation (4) reduces to $P\{S_n = s, X_{n,n} \leq r\} = P\{S_n = s\}$, which is obvious when $r \geq s$.

For fixed integers $n \geq m \geq 1$ and $r \geq s \geq 0$, let $A_i = A_i(n, m)$ denote the event $A_i(n, m) := \{S_n(m) = s, X_i > r\}$, $i = 1, \ldots, n$. Then, by inclusion and exclusion,

$$P\{S_n(m) = s, X_{n,n} > r\} = P\{\cup_{i=1}^n A_i\} = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} P\{\cap_{j=1}^k A_{i_j}\}$$

$$= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} P\{\cap_{i=1}^k A_i\}$$

$$= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} P\{X_i > r, 1 \leq i \leq k, \text{ and the sum of the}$$

$$n - m \text{ smallest among } X_{k+1}, \ldots, X_n \text{ equals } s\}$$

$$= \sum_{k=1}^{m} (-1)^{k-1} \binom{n}{k} [1 - F(r)]^k P\{S_{n-k}(m-k) = s\} \,,$$

4

where the assumption that $r \geq s$ is essential for the fourth equality. Since the probability $P\{S_n(m) = s, X_{n,n} > r\} = P\{S_n(m) = s\} - P\{S_n(m) = s, X_{n,n} \leq r\}$, we see that the desired probability $P\{S_n(m) = s, X_{n,n} \leq r\}$ is

$$P\{S_n(m) = s\} + \sum_{k=1}^{m} (-1)^k \binom{n}{k} [1 - F(r)]^k P\{S_{n-k}(m - k) = s\}$$

$$= \sum_{k=0}^{m} (-1)^k \binom{n}{k} [1 - F(r)]^k P\{S_{n-k}(m - k) = s\},$$

proving the theorem. □

PROOF OF THEOREM 2. While this theorem can be viewed as a corollary of Theorem 1, with (8) following from (4) by a direct combinatorial calculation, it is more instructive to present a straightforward probabilistic argument. The statement is trivial if $n = m$. Fix $n > m \geq 0$ and $r \geq s \geq 0$, and let $K_n(r)$ denote the number of $X_i > r$, $1 \leq i \leq n$. Then,

$$P\{S_n(m) = s\} = \sum_{k=0}^{m} P\{S_n(m) = s, K_n(r) = k\}.$$

Now, for $0 \leq k \leq m$, introduce the events $B_k := \{X_i > r \text{ for } n - k < i \leq n\}$ and $C_k := \{\text{the } n - m \text{ smallest among } X_1, \ldots, X_{n-k} \text{ sums to } s, X_i \leq r \text{ for } 1 \leq i \leq n-k\}$. Since there are $\binom{n}{k}$ ways of choosing exactly k of the X_i's to exceed r, $1 \leq i \leq n$, and thereby making $K_n(r) = k$, we have

$$P\{S_n(m) = s, K_n(r) = k\} = \binom{n}{k} P\{B_k \cap C_k\} = \binom{n}{k} P\{B_k\} P\{C_k\}$$

$$= \binom{n}{k} [1 - F(r)]^k P\{S_{n-k}(m - k) = s, X_{n-k,n-k} \leq r\}.$$

The two equations together complete the proof of (8). □

To proceed, we need some additional notation. For fixed $n \geq m \geq 1$, let

$$X_n(m) = (X_{n,n-m+1}, \ldots, X_{n,n}).$$

When $X_n(m) = t$, then $t = (t_1, \ldots, t_m)$, where $0 \leq t_1 \leq \cdots \leq t_m$ are some integers. Given such a t, let $\underline{t}$ denote the smallest (the first) component of t, let $\{t\}$ denote the

set of integer values appearing in t, without repetitions, and, for integers j and k, with $0 \leq j \leq \underline{t}$ and $k \in \{t\}$, let $t[j, k]$ denote the vector formed by augmenting t from the left with the integer j and deleting one of the k's appearing in t. Thus, $t[j, k]$ remains an m-dimensional vector with the same properties as t. Finally, for a given $t = (t_1, \ldots, t_m)$ as above, set $\{t\}_* := \{0, 1, 2, \ldots, \underline{t} - 1\}$, and let us agree that $\{t\}_* = \emptyset$ if $\underline{t} = 0$.

Clearly, $X_n \in \{t\}_* \cup \{t\}$ and $X_{n,n-m} \in \{0, 1, 2, \ldots, \underline{t}\} = \{t\}_* \cup \{\underline{t}\}$ when $X_n(m) = t$. Furthermore, to handle an incoming new observation X_n at time n, we claim that for $n - 1 \geq m$, if $X_n(m) = t$, then

$$(9) \quad \left(S_{n-1}(m), X_{n-1}(m)\right) = \begin{cases} \left(S_n(m) - X_n, t\right), & \text{if } X_n \in \{t\}_*, \\ \left(S_n(m) - X_{n,n-m}, t[X_{n,n-m}, X_n]\right), & \text{if } X_n \in \{t\}. \end{cases}$$

The first of these is obvious because a value of $X_n \in \{t\}_* = \{0, 1, 2, \ldots, \underline{t} - 1\}$ can not be in $\{t\}$, and hence $X_{n-1}(m) = X_n(m) = t$, and, for the same reason, the difference in the trimmed sums $S_n(m) - S_{n-1}(m)$ must be X_n. But when $X_n \in \{t\}$, the new observation X_n is trimmed at time n, or, optionally, can be trimmed if it and some previous random variable are both equal to $\underline{t}$. If X_n is trimmed, the smallest member of $X_{n-1}(m)$, namely $X_{n-1,(n-1)-m+1} = X_{n,n-m}$ must be deleted from the trimmed set at time n, because X_n is trimmed instead of it, and so $X_{n,n-m}$ is a term in $S_n(m)$, and hence $X_{n-1}(m) = t[X_{n,n-m}, X_n]$, and the difference $S_n(m) - S_{n-1}(m) = X_{n,n-m}$. Notice, for the optional case, that $S_n(m) - S_{n-1}(m)$ is $X_{n,n-m} = \underline{t} = X_n$ and $t[X_{n,n-m}, X_n] = t[\underline{t}, \underline{t}] = \{t\}$, so that either form of (9) can be used.

Understanding an empty sum as zero, using the indices j and k to represent the values of $X_{n,n-m}$ and X_n, respectively, and noting the independence of the vector $\left(S_{n-1}(m), X_{n-1}(m)\right)$ and X_n, we are lead to the following recursion.

THEOREM 3. *For integers* $n > m \geq 1$ *and* $t_m \geq \cdots \geq t_1 = \underline{t}, s \geq 0$, *with* $t = (t_1, \ldots, t_m)$,

$$(10) \quad \begin{aligned} P\{S_n(m) = s, X_n(m) = t\} &= \sum_{k=0}^{\underline{t}-1} P\{X = k\} P\{S_{n-1}(m) = s - k, X_{n-1}(m) = t\} \\ &+ \sum_{k \in \{t\}} P\{X = k\} \sum_{j=0}^{\underline{t}} P\{S_{n-1}(m) = s - j, X_{n-1}(m) = t[j, k]\}. \end{aligned}$$

6

PROOF. If $X_n(m) = t$, the observation X_n is confined to the set $\{t\}_* \cup \{t\} = \{0, 1, 2, \ldots, \underline{t} - 1\} \cup \{t\}$. The values $X_n = k \in \{t\}_* = \{0, 1, 2, \ldots, \underline{t} - 1\}$ give rise to the first sum on the right side of (10):

$$P\{S_n(m) = s, X_n(m) = t, X_n = k\} = P\{X = k\}\, P\{S_{n-1}(m) = s - k, X_{n-1}(m) = t\},$$

in accordance with the first case in (9). The values $X_n = k \in \{t\}$ give rise to the double sum on the right side of (10) as follows:

$$P\{S_n(m) = s, X_n(m) = t, X_n = k\}$$

$$= \sum_{j=0}^{\underline{t}} P\{S_n(m) = s, X_n(m) = t, X_n = k, X_{n,n-m} = j\}$$

$$= \sum_{j=0}^{\underline{t}} P\{S_n(m) = s, X_{n-1}(m) = t[j,k], X_n = k\}$$

$$= \sum_{j=0}^{\underline{t}} P\{S_{n-1}(m) = s - j, X_{n-1}(m) = t[j,k], X_n = k\}$$

$$= P\{X = k\} \sum_{j=0}^{\underline{t}} P\{S_{n-1}(m) = s - j, X_{n-1}(m) = t[j,k]\},$$

in accordance with the second case in (9). This establishes (10). $\qquad\square$

3. Application to the St. Petersburg game. The context of the St. Petersburg paradox is a game, based on a sequence of fair coin tosses, in which Peter agrees to pay Paul $X = 2^k$ ducats, where k is the number of tosses required to produce the first head, so that $P\{X = 2^k\} = 2^{-k}$, $k = 1, 2, \ldots$. The simple fact that Paul's expected winnings, $E(X)$, is infinite provides the basis for the paradox. For as Nicolaus Bernoulli, who posed the problem in 1713, wrote in 1728 to his younger cousin Daniel, "... there ought not to exist any even halfway sensible person who would not sell the right of this gain for forty ducats." (The original numbers of ducats are doubled here and everywhere in our discussion to conform with a more convenient payoff scheme used by many subsequent writers. The translation from the Latin is taken from Martin-Löf (1985); we like it better

7

than the standard form in the English translation of Bernoulli (1738), where Daniel cites Nicolaus' letter. See Jorland (1987) and Dutka (1988) for recent historical accounts.)

Despite the attention of many well known mathematicians, stretching over a quarter of a millennium, a significant mathematical treatment of the subject did not occur until Feller (1945) addressed the topic, arguing that the question of "Paul's fair price" only makes sense when one considers a sequence of independent St. Petersburg games, with payoffs $X_1, X_2, \ldots$ distributed as X, and asks what a "fair price" would be for playing n such games. Addressing this issue, he showed, see also Feller (1950), that

$$\frac{S_n}{n \operatorname{Log} n} \to 1 \quad \text{in probability as} \quad n \to \infty,$$

where $S_n := X_1 + \cdots + X_n$, and $\operatorname{Log} n$ denotes the base 2 logarithm of n. Subsequently, Chow and Robbins (1961) showed that the convergence in Feller's law can not be upgraded to almost sure convergence. Indeed, it can easily be shown that $P\{X_n > c n \operatorname{Log} n \text{ infinitely often}\} = 1$ for every $c > 0$. On the other hand, the authors (1994) have shown for every $m \geq 1$ that

$$\frac{S_n(m)}{n \operatorname{Log} n} \to 1 \quad \text{almost surely as} \quad n \to \infty.$$

These facts suggest that $p_n(\varepsilon) := P\{S_n > (1 + \varepsilon) n \operatorname{Log} n\}$ might go to zero quite slowly with n, and its trimmed analogue, $p_n(m, \varepsilon) := P\{S_n(m) > (1 + \varepsilon) n \operatorname{Log} n\}$ might converge to zero more rapidly when $m \geq 1$. This conjecture is investigated numerically for $m = 1$ in Figures 1 and 2 below, which contain overlaying plots of $p_n(\varepsilon)$ and $p_n(1, \varepsilon)$ for $\varepsilon = 0.25$ and $\varepsilon = 1$, respectively.

The plots in the untruncated case are based on the simple recursion appearing in (1) and (2), made somewhat easier to compute by the fact that the index k in (1) is restricted to integer powers of 2. Frankly, we were initially surprised that this simple recursion could be run out to n values well into the thousands, with high accuracy maintained, and without major difficulties.

The calculations for $m = 1$ and $n = 1, \ldots, N$ proceed as follows: Beginning with

$$P\{S_1(1) = s, X_{1,1} = 2^t\} = \frac{I_{\{0\}}(s)}{2^t}, \quad t \geq 1, \quad \text{and} \quad P\{S_1(1) = s\} = I_{\{0\}}(s),$$

8

where $I_{\{0\}}(s) = 0$ or 1 as $s \geq 1$ or $s = 0$, one computes $P\{S_n(1) = s, X_{n,n} = 2^t\}$ and $P\{S_n(1) = s\}$ successively for $n = 2, \ldots, N$ with the recursions

$$P\{S_n(1) = s, X_{n,n} = 2^t\} = \sum_{k=0}^{t-1} \frac{1}{2^k} P\{S_{n-1}(1) = s - 2^k, X_{n-1,n-1} = 2^t\}$$

$$+ \frac{1}{2^t} \sum_{j=0}^{t} P\{S_{n-1}(1) = s - 2^j, X_{n-1,n-1} = 2^j\}$$

for the integers $0 \leq s \leq r$, $0 \leq t \leq \lfloor \text{Log } r \rfloor$, and

$$P\{S_n(1) = s\} = \sum_{t=0}^{\lfloor \text{Log } r \rfloor} P\{S_n(1) = s, X_{n,n} = 2^t\} + \frac{n}{2^{\lfloor \text{Log } r \rfloor}} P\{S_{n-1} = s\}$$

for $0 \leq s \leq r$, where $\lfloor x \rfloor := \max\{k = 0, 1, 2, \ldots : k \leq x\}$ is the usual integer part and, below, $\lceil x \rceil := \min\{k = 0, 1, 2, \ldots : k \geq x\}$ is the upper integer part of $x \geq 0$. Then

$$p_n(1, \varepsilon) = 1 - P\{S_n(1) \leq (1 + \varepsilon) n \text{ Log } n\} = 1 - \sum_{s=0}^{r(n)} P\{S_n(1) = s\},$$

where $r(n) := \lfloor (1 + \varepsilon) n \text{ Log } n \rfloor$.

In order to do all the required calculations for n up to N, *with just one set of recursions*, one must work with a single $r \geq r(N) = \lfloor (1 + \varepsilon) N \text{ Log } N \rfloor$. Here, $N = 4096 = 2^{12}$ for Figures 1 and 2, resulting in $r(N) = 71\,680$ and $r(N) = 114\,688$ for $\varepsilon = 0.25$ and $\varepsilon = 1$, respectively.

While these calculations were carried out with good accuracy, a substantial memory burden was encountered that required the storage of approximately $\lfloor (1 + \varepsilon) N (\text{Log } N)^2 \rfloor$ double precision numbers: about 4 million when $\varepsilon = 0.25$ and about 6.5 million when $\varepsilon = 1$.

Place Figures 1 and 2 about here

The horizontal axes in Figures 1 and 2 are expressed in units of $\text{Log } n$, rather than n, in order to draw attention to empirical evidence indicating a link between the values of $p_n(\varepsilon)$ and $p_n(1, \varepsilon)$ and the location of n between consecutive integer powers of 2. Theoretical support for this link is provided by the fact that the distribution functions of

9

$(S_n - n \operatorname{Log} n)/n$ and $(S_n(1) - n \operatorname{Log} n)/n$ are asymptotically approximated, as $n \to \infty$, by the distribution functions of certain infinitely divisible random variables and their "trimmed" analogues, respectively, chosen on the basis of the value $\gamma_n := n/2^{\lceil \operatorname{Log} n \rceil}$, $1/2 < \gamma_n \le 1$. This is described in a forthcoming book by the authors.

Evidence of slow convergence to zero of both $p_n(\varepsilon)$ and $p_n(1, \varepsilon)$, and especially of $p_n(\varepsilon)$, is apparent in Figures 1 and 2. By methods outside the scope of the present paper, we can prove for every fixed $\varepsilon > 0$ that

$$(11) \qquad 1 \le \liminf_{n \to \infty} \left[\varepsilon \operatorname{Log} n \right] p_n(\varepsilon) \le \limsup_{n \to \infty} \left[\varepsilon \operatorname{Log} n \right] p_n(\varepsilon) \le 2 .$$

Also, we have sufficient grounds to conjecture, but for the time being can not prove that for every fixed $\varepsilon > 0$,

$$(12) \qquad \frac{1}{(m+1)!} \le \liminf_{n \to \infty} [\varepsilon \operatorname{Log} n]^{m+1} p_n(m, \varepsilon) \le \limsup_{n \to \infty} [\varepsilon \operatorname{Log} n]^{m+1} p_n(m, \varepsilon) \le \frac{2^{m+1}}{(m+1)!}.$$

Notice that (12) reduces to (11) when $m = 0$.

Figure 3 below provides numerical evidence, for four different values of ε, supporting the truth of (11), and it strongly indicates that the influence of the asymptotics arises quickly when ε is relatively large, and more slowly when it is small. Moreover these graphs suggest that neither bound in (11) is tight. Working with the two values of ε for which we have data, we see evidence in Figure 4 that supports our conjecture in (12) for the particular case $m = 1$.

```
Place Figures 3 and 4 about here
```

REFERENCES

BERNOULLI, D. (1738). Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, Tomus V, Ad annos 1730 et 1731, pp. 175–193; published in 1738. [Reprinted in: *Die Werke von Daniel Bernoulli*, Vol. II, pp. 223–234. Birkhäuser, Basel, 1982. English translation by L. Sommer (1954): Exposition of a new theory of the measurement of risk. *Econometrica* 22 23–36; reprinted in *Precursors in Mathematical Economics: An Anthology* (W. J. Baumol and S. M. Goldfeld, eds.), Series of Reprints of Scarce Works on Political Economy, No. 19, pp. 15–26, London School of Economics and Political Science, London, 1968; and in: *Utility Theory: A Book of Readings* (A. N. Page, ed.), pp. 199–214, Wiley, New York, 1968.]

DAVID, H. A. (1981). *Order Statistics*, Second edition. Wiley, New York.

10

CHOW, Y. S. and ROBBINS, H. (1961). On sums of independent random variables with infinite means and "fair" games. *Proc. Nat. Acad. Sci. U.S.A.* **47** 330–335.

CSÖRGŐ, S. and SIMONS, G. (1994). On the strong law of large numbers for trimmed sums. Submitted for publication.

DUTKA, J. (1988). On the St. Petersburg paradox. *Archive for History of Exact Sciences* **39** 13–39.

FELLER, W. (1945). Note on the law of large numbers and "fair" games. *Ann. Math. Statist.* **16** 301–304.

FELLER, W. (1950). *An Introduction to Probability Theory and its Applications*, Vol. I. Wiley, New York. [Second and third editions: 1957, 1968; Chapter X, Section 4, in all editions.]

HAHN, M. G., MASON, D. M. and WEINER D.C., editors (1991). *Sums, Trimmed Sums and Extremes* (Progress in Probability, Volume **23**). Birkhäuser, Boston.

JORLAND, G. (1987). The Saint Petersburg paradox 1713–1937. In: *The Probabilistic Revolution, Vol. I: Ideas in History* (L. Krüger, L. D. Daston and M. Heidelberger, eds.), pp. 157–190. MIT Press, Cambridge, Massachusetts.

MARTIN-LÖF, A. (1985). A limit theorem which clarifies the 'Petersburg paradox'. *J. Appl. Probab.* **22** 634–643.

STIGLER S. M. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1** 472–477.

DEPARTMENT OF STATISTICS
THE UNIVERSITY OF MICHIGAN
1444 MASON HALL
ANN ARBOR, MICHIGAN 48109–1027

DEPARTMENT OF STATISTICS
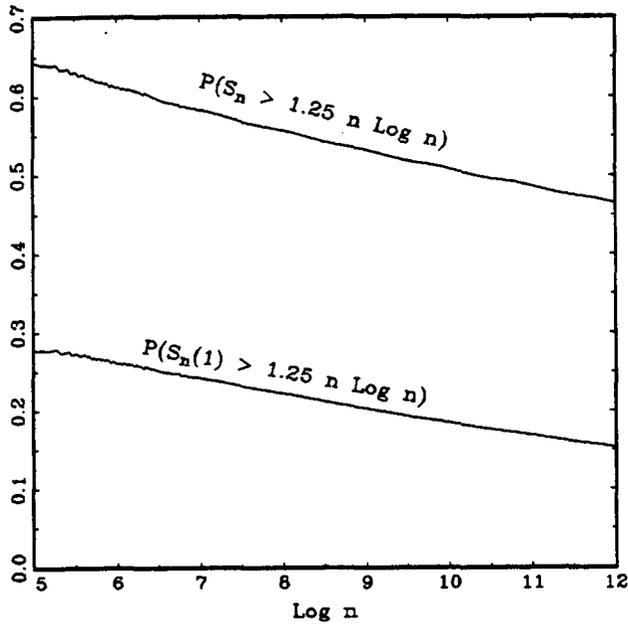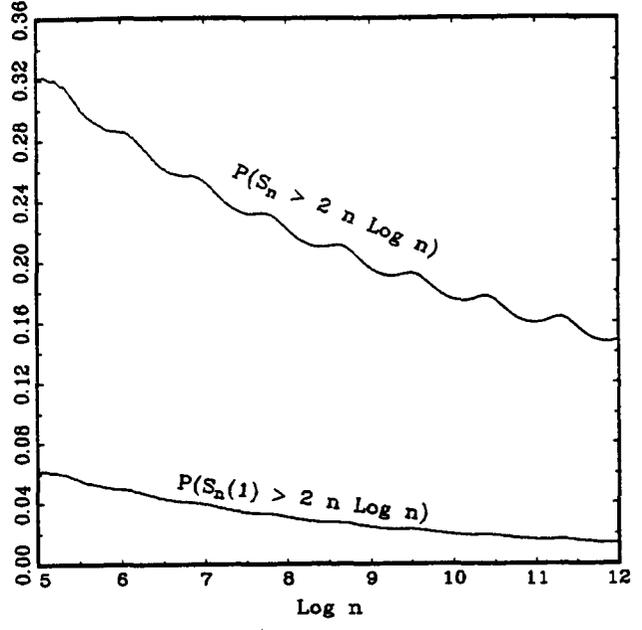THE UNIVERSITY OF NORTH CAROLINA
322 PHILLIPS HALL
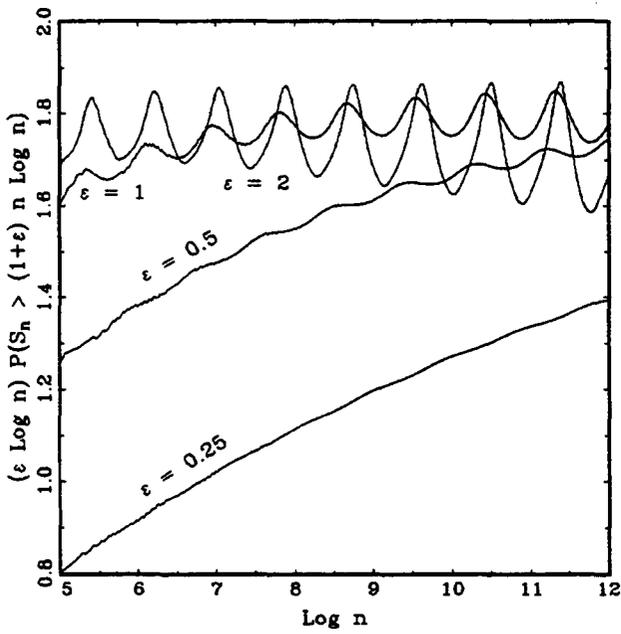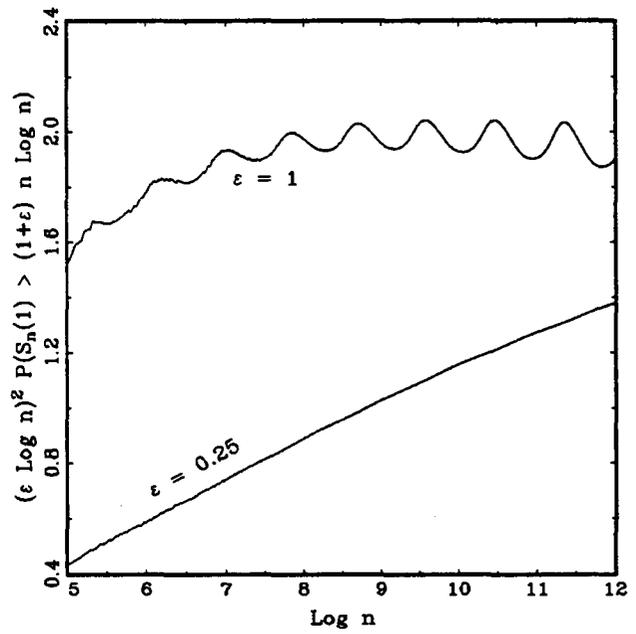CHAPEL HILL, NORTH CAROLINA 27599–3260

Figure 1

Figure 2

Figure 3

Figure 4