The Annals of Applied Probability
1998, Vol. 8, No. 3, 886–895

ON THE VARIANCE OF THE NUMBER OF MAXIMA IN RANDOM VECTORS AND ITS APPLICATIONS

BY ZHI-DONG BAI, CHERN-CHING CHAO,[1] HSIEN-KUEI HWANG AND WEN-QI LIANG[1]

National Sun Yat-Sen University, Academia Sinica, Academia Sinica and Academia Sinica

We derive a general asymptotic formula for the variance of the number of maxima in a set of independent and identically distributed random vectors in $\mathbb{R}^d$, where the components of each vector are independently and continuously distributed. Applications of the results to algorithmic analysis are also indicated.

1. Introduction. Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a set of independent and identically distributed (iid) random vectors in $\mathbb{R}^d$. A point $\mathbf{x}_i = (x_{i1}, \ldots, x_{id})$ is said to be *dominated* by $\mathbf{x}_j$ if $x_{ik} < x_{jk}$ for all $k = 1, \ldots, d$, and a point $\mathbf{x}_i$ is called a *maximum* of X if none of the other points dominates it. This paper is concerned with the number of maxima, denoted by $K_{n,d}$, of X.

The study of the number of maxima of a set of points was initiated by Barndorff-Nielsen and Sobel (1966) as an attempt to describe the boundary of a set of random points in $\mathbb{R}^d$. Due to its close relationships to convex hull, this problem has been developed to be one of the core problems in computational geometry, with many applications in diverse disciplines such as pattern classification, graphics, economics, data analysis, etc. The reader is referred to Preparata and Shamos (1985), Bentley, Kung, Schkolnick and Thompson (1978), Becker, Denby, McGill and Wilks (1987) and Bentley, Clarkson and Levine (1993), Golin (1993) for more information. This problem also arose in the multicriterial choice problem in operations research. Let x_{ij} represent a utility of variant (alternative, plan) i according to criterion j, $i = 1, \ldots, n$, $j = 1, \ldots, d$. If there is no relation of criteria according to importance, the choice is often made by relying on the partial order relation $\mathbf{x}_i \succ \mathbf{x}_j$ if $x_{ik} \geq x_{jk}$ for all k and $x_{il} > x_{jl}$ for some l. Then the optimal variants constitute the so-called *Pareto* set of X, that is, the set of all $\mathbf{x}_i$ which are not "$\prec$" by others. The Pareto set has been actively investigated since the seventies, notably in Russia; see the survey paper by Sholomov (1983). Under the assumptions that $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are iid and that the components of each vector are identically and continuously distributed, the Pareto set is identical to the set of maxima. In the sequel, all (with only one exception) results concerning the random variables $K_{n,d}$ mentioned in this paper are under the above assumptions.

Received September 1997; revised November 1997.

[1] Supported in part by NSC Grant NSC-86-2115-M001-022.

AMS 1991 *subject classifications*. Primary 60D05; secondary 68Q25, 65Y25.

Key words and phrases. Maximal points, multicriterial optimization, Eulerian sums.

Dominance is clearly one of the natural order relations in multivariate observations. Thus, the random variables $K_{n,d}$ play a fundamental role in diverse fields, and some of their probabilistic properties have been rediscovered in the literature. Barndorff-Nielsen and Sobel (1966) first showed, as a special case of their general results, that

$$\begin{aligned}\mu_{n,d} := \mathrm{E}(K_{n,d}) &= \sum_{1\le k\le n} \binom{n}{k}(-1)^{k-1}k^{1-d}\\ &= \frac{(\log n)^{d-1}}{(d-1)!}\big(1+O((\log n)^{-1})\big), \qquad n\to\infty,\end{aligned} \tag{1.1}$$

for $d \ge 2$. This problem is the object of many papers (some simplifying the proofs of the others) by Berezovskii and Travkin (1975), Ivanin (1975b), Bentley, Kung, Schkolnick and Thompson (1978), Devroye (1980), O'Neill (1980), Buchta (1989). The only exception mentioned above is that Ivanin (1975a) dropped the assumption of independence of components and derived an asymptotic formula for $\mathrm{E}(K_{n,d})$ for multivariate normal random variables $\mathbf{x}_i$.

Finding the distribution of $K_{n,d}$ for general $d \ge 3$ is definitely more difficult (see discussions in the Bulletin Board of the newly established WEB site: `http://www-rocq.inria.fr/algo/AofA/index.html`). However, for $d = 2$, if we arrange x_{i1}, $i = 1, \dots, n$, in decreasing order, then it is easily seen that $K_{n,2}$ is essentially identical to the number of record values in a set of n iid random variables with a common continuous distribution. Thus, the exact distribution is nothing but the Stirling numbers of the first kind given by [see Barndorff-Nielsen and Sobel (1966)],

$$\mathrm{E}\big(z^{K_{n,2}}\big) = \frac{z(z+1)\cdots(z+n-1)}{n!}, \qquad n \ge 1,$$

and its asymptotic normality is also implied. In addition, Barndorff-Nielsen and Sobel put forth methods for calculating the distribution of $K_{n,d}$ for (1) small d and general n, and (2) small n and general d, and carried out the computations for $n = 2, 3, 4, 5$.

For the variance, it is known that

$$\mathrm{Var}(K_{n,2}) = H_n - H_n^{(2)} = \log n + \gamma - \frac{\pi^2}{6} + O(n^{-1}), \qquad n\to\infty,$$

where $H_n^{(j)} = \sum_{1\le k\le n} k^{-j}$ denotes the harmonic numbers and γ is Euler's constant. Barndorff-Nielsen and Sobel (1966) also showed that

$$\begin{aligned}\mathrm{Var}(K_{n,3}) &= 6\sum_{1\le i\le j<k\le l\le n}\frac{1}{ijkl} + \sum_{1\le i\le j\le n}\frac{1}{ij} - \big(\mathrm{E}(K_{n,3})\big)^2\\ &= \left(\frac{1}{2}+\frac{\pi^2}{6}\right)(\log n)^2\big(1+O((\log n)^{-1})\big), \qquad n\to\infty.\end{aligned}$$

For general d, Devroye (1997) derived the general estimate

$$\mathrm{Var}(K_{n,d}) = O\big(\mathrm{E}(K_{n,d})\big).$$

This implies, by Chebyshev's inequality, that

$$\frac{K_{n,d}}{(\log n)^{d-1}/(d-1)!} \to 1 \quad \text{in probability}, \qquad n \to \infty.$$

On the other hand, Ivanin [(1976), page 99] derived an exact formula for the second moment of $K_{n,d}$:

$$\mathrm{E}\bigl(K_{n,d}^2\bigr) = \mu_{n,d} + \sum_{1\le t<d} \binom{d}{t} \sum_{l=1}^{n-1} \frac{1}{l} \sum\nolimits^{(1)} \frac{1}{i_1\cdots i_{d-2}j_1\cdots j_{d-1}}, \tag{1.2}$$

where the summation $\sum^{(1)}$ runs over all indices satisfying the inequalities

$$1 \le i_1 \le \cdots \le i_{t-1} \le l, \qquad 1 \le i_t \le \cdots \le i_{d-2} \le l$$

and

$$l < j_1 \le \cdots \le j_{d-1} \le n.$$

From this expression, the asymptotics of the variances for $d = 2, 3, 4$ were further simplified:

$$\mathrm{Var}(K_{n,4}) \sim \biggl(\sum_{1\le j\le n} j^{-3} + \sum_{2\le j\le n} j^{-2} \sum_{1\le i<j} j^{-1} + \tfrac{1}{6} \biggr) (\log n)^3, \qquad n \to \infty. \tag{1.3}$$

As the main result of this paper, we establish the following theorem.

THEOREM. *For* $d \ge 2$,

$$\mathrm{Var}(K_{n,d}) = \biggl(\frac{1}{(d-1)!} + c_d \biggr) (\log n)^{d-1}\bigl(1 + O\bigl((\log n)^{-1}\bigr)\bigr), \qquad n \to \infty, \tag{1.4}$$

where

$$\begin{aligned} c_d = \frac{1}{(d-1)!} \sum_{l\ge 1} \frac{1}{l^2} \sum_{1\le p,q\le l} \binom{l}{p}\binom{l}{q} (-1)^{p+q} pq \\ \times \bigl((p^{-1} + q^{-1})^{d-1} - p^{1-d} - q^{1-d} \bigr). \end{aligned} \tag{1.5}$$

Thus asymptotically $\mathrm{Var}(K_{n,d}) \ge \mathrm{E}(K_{n,d})$ as n becomes large.

The proof will be presented in Section 2. We first give an alternative derivation of (1.2) and then consider $\mathrm{E}(K_{n,d}^2) - \mu_{n,d}^2$. Comparing (1.1) and (1.4), we see that the major task in proving (1.4) is to cancel the first $d-1$ terms in the asymptotic expansion of $\mu_{n,d}^2$ and to identify the dth term.

For constants c_d, we have, in particular,

$$c_2 = 0 \quad \text{and} \quad c_3 = \zeta(2) = \sum_{l\ge 1} l^{-2} = \frac{\pi^2}{6},$$

where $\zeta(s)$ denotes Riemann's zeta function $\zeta(s) = \sum_{k\geq 1} k^{-s}$ for $\Re s > 1$. While a general reduction of c_d to Riemann's zeta function at integer arguments may seem impossible in view of current knowledge on multiple harmonic sums [cf. Bailey, Borwein and Girgensohn (1994), Flajolet and Salvy (1996), Hoffman (1992) and Zagier (1992)], we have, by a well-known formula due to Euler,

$$c_4 = \sum_{l\geq 1} \frac{H_l}{l^2} = 2\zeta(3)$$

[simplifying (1.3)], and, by reductions to suitable Euler sums,

$$c_5 = \frac{5}{12}\sum_{l\geq 1}\frac{H_l^2}{l^2} + \frac{1}{6}\sum_{l\geq 1}\frac{H_l^{(2)}}{l^2} = \frac{33}{16}\zeta(4),$$

$$c_6 = \frac{7}{72}\sum_{l\geq 1}\frac{H_l^3}{l^2} + \frac{1}{8}\sum_{l\geq 1}\frac{H_l H_l^{(2)}}{l^2} + \frac{1}{36}\sum_{l\geq 1}\frac{H_l^{(3)}}{l^2} = \frac{5}{4}\zeta(5) + \frac{1}{6}\zeta(2)\zeta(3).$$

These identities can be derived by the results in Flajolet and Salvy (1996); details are omitted here. (They are numerically easy to check by using symbolic computation packages like MAPLE or MATHEMATICA.) Note that

$$c_d = \sum_{l\geq 1}\frac{1}{l^2}\sum_{j=1}^{d-2}\frac{\mu_{l,j}\mu_{l,d-1-j}}{j!(d-1-j)!}.$$

Applications of our theorem to algorithmic analysis will be briefly discussed in Section 3. Based on numerical simulations, we predict that the asymptotic distribution of $K_{n,d}$ would be Gaussian. However, we have not found any proof for $d \geq 3$.

2. Proof of the theorem. Without loss of generality, we assume that n iid random vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are uniformly distributed over $(0,1)^d$. Denote by G_k the event (as well as the indicator of the event) that $\mathbf{x}_k$ is a maximum in $\mathbf{x}_1, \dots, \mathbf{x}_n$. Then

$$K_{n,d} = \sum_{k=1}^{n} G_k.$$

If there are exactly $r-1$ points dominating $\mathbf{x}_k$, then $\mathbf{x}_k$ is called an rth layer maximum. Denote this event by $G_k(r)$. Thus, the total number of rth layer maxima can be expressed by

$$K_{n,d}(r) = \sum_{k=1}^{n} G_k(r).$$

To prove the theorem, we first derive a lemma and the mean of $K_{n,d}(r)$.

LEMMA 1. *Let $0 \le t < d$. Then,*

$$\int_{(0,\,1)^d}\biggl(1-\prod_{i=1}^{d} x_i\biggr)^n\biggl(\prod_{i=t+1}^{d} x_i\biggr)^l\,d\mathbf{x} = \sum_{1\le i_1\le\cdots\le i_{d-1}\le n+1}\frac{n!(i_t+l-1)!}{(n+l+1)!(i_1\cdots i_{t-1})i_t!(i_{t+1}+l)\cdots(i_{d-1}+l)}. \tag{2.1}$$

PROOF. Rewrite

$$1-\prod_{i=1}^{d} x_i = (1-x_1)x_2\cdots x_d + (1-x_2)x_3\cdots x_d + \cdots + (1-x_d).$$

Then

$$\biggl(1-\prod_{i=1}^{d} x_i\biggr)^n\biggl(\prod_{i=t+1}^{d} x_i\biggr)^l = \sum_{\substack{i_1+\cdots+i_d=n\\ i_1,\dots,i_d\ge 0}}\frac{n!}{i_1!\cdots i_d!}(1-x_1)^{i_1}x_2^{i_1}(1-x_2)^{i_2}\cdots$$
$$\times\, x_t^{i_1+\cdots+i_{t-1}}(1-x_t)^{i_t}x_{t+1}^{i_1+\cdots+i_t+l}(1-x_{t+1})^{i_{t+1}}\cdots$$
$$\times\, x_d^{i_1+\cdots+i_{d-1}+l}(1-x_d)^{i_d},$$

and (2.1) follows. □

In particular, for $t=0$, we have

$$\int_{(0,\,1)^d}\biggl(1-\prod_{i=1}^{d} x_i\biggr)^n\biggl(\prod_{i=1}^{d} x_i\biggr)^l\,d\mathbf{x} = \frac{n!l!}{(n+l+1)!}\sum_{l+1\le i_1\le\cdots\le i_{d-1}\le n+l+1}\frac{1}{i_1\cdots i_{d-1}}. \tag{2.2}$$

COROLLARY 1 [Barndorff-Nielsen and Sobel (1966)]. *The mean number of rth layer maxima is given by*

$$\mu_{n,\,d}(r) := \mathrm{E}\bigl(K_{n,\,d}(r)\bigr) = \sum_{r\le i_1\le\cdots\le i_{d-1}\le n}\frac{1}{i_1\cdots i_{d-1}}. \tag{2.3}$$

PROOF. The result follows from

$$\mu_{n,\,d}(r) = n\,\mathrm{E}(G_1(r)) = n\binom{n-1}{r-1}\int_{(0,\,1)^d}\biggl(1-\prod_{i=1}^{d} x_i\biggr)^{n-r}\biggl(\prod_{i=1}^{d} x_i\biggr)^{r-1}d\mathbf{x}$$

and (2.2). □

REMARK. It is interesting to note that the probability that a point, say $\mathbf{x}_i$, is a maximal point satisfies

$$\frac{\mu_{n,\,d}}{n} = \mathrm{P}(Y_2+\cdots+Y_n < d),$$

where each Y_j is geometrically distributed:

$$\mathrm{E}(z^{Y_j}) = \frac{1-1/j}{1-z/j}, \qquad 2 \le j \le n.$$

This follows from the fact that $\mu_{n,d}/n$ equals the coefficient of z^{d-1} in

$$\frac{1}{n}\prod_{1\le j\le n}\frac{1}{1-z/j} = \frac{1}{1-z}\prod_{2\le j\le n}\frac{1-1/j}{1-z/j}.$$

Also from a computational point of view, it is useful to use the recurrence $\mu_{n,1} = 1$ for $n \ge 1$ and for $d \ge 2$,

$$\mu_{n,d} = \frac{1}{d-1}\sum_{1\le j\le d-1} H_n^{(d-j)}\mu_{n,j},$$

by taking derivatives on both sides of $\sum_{d=1}^{\infty}\mu_{n,d}z^{d-1} = \prod_{j=1}^{n} 1/(1-z/j)$ and by equating coefficients of the same powers.

Next, we derive the second moment of $K_{n,d}$. Let $(t) = \{(\mathbf{x},\mathbf{y});\, x_1 > y_1, \dots, x_t > y_t,\, x_{t+1} < y_{t+1}, \dots, x_d < y_d\}$. We have

$$\begin{aligned}
\mathrm{E}(K_{n,d}^2) &= \mu_{n,d} + n(n-1)\mathrm{P}(G_1G_2)\\
&= \mu_{n,d} + n(n-1)\sum_{t=1}^{d-1}\binom{d}{t}\int_{(t)}\Bigg(1-\prod_{i=1}^{d}(1-x_i)-\prod_{i=1}^{d}(1-y_i)\\
&\qquad\qquad + \prod_{i=1}^{t}(1-x_i)\prod_{i=t+1}^{d}(1-y_i)\Bigg)^{n-2} d\mathbf{x}\,d\mathbf{y}\\
&= \mu_{n,d} + n(n-1)\sum_{t=1}^{d-1}\binom{d}{t}\\
&\quad \times \int_{(0,1)^d}\int_{(0,1)^d}\left(1-\prod_{i=1}^{d}x_i\prod_{i=1}^{t}y_i - \prod_{i=t+1}^{d}x_i\prod_{i=1}^{d}y_i + \prod_{i=1}^{d}x_i\prod_{i=1}^{d}y_i\right)^{n-2}\\
&\qquad \times \prod_{i=t+1}^{d}x_i\prod_{i=1}^{t}y_i\,d\mathbf{x}\,d\mathbf{y}\\
&= \mu_{n,d} + n(n-1)\sum_{t=1}^{d-1}\binom{d}{t}\\
&\quad \times \int_{(0,1)^d}\int_{(0,1)^d}\left(1-\prod_{i=1}^{t}x_i\prod_{i=1}^{d}y_i - \prod_{i=t+1}^{d}x_i\prod_{i=1}^{d}y_i + \prod_{i=1}^{d}x_i\prod_{i=1}^{d}y_i\right)^{n-2}\\
&\qquad \times \prod_{i=1}^{d}y_i\,d\mathbf{x}\,d\mathbf{y}\\
&:= \mu_{n,d} + \sum_{t=1}^{d-1}\binom{d}{t}J_t.
\end{aligned}$$

Noting that

$$\begin{aligned}1-\prod_{i=1}^{t}x_i\prod_{i=1}^{d}y_i-\prod_{i=t+1}^{d}x_i\prod_{i=1}^{d}y_i+\prod_{i=1}^{d}x_i\prod_{i=1}^{d}y_i\\
=\left(1-\prod_{i=1}^{d}y_i\right)+\prod_{i=1}^{d}y_i\left(1-\prod_{i=1}^{t}x_i\right)\left(1-\prod_{i=t+1}^{d}x_i\right),\end{aligned}$$

we have, by Lemma 1,

$$\begin{aligned}J_t&=\sum_{l=0}^{n-2}\frac{n!}{l!(n-2-l)!}\int_{(0,1)^d}\int_{(0,1)^d}\left(1-\prod_{i=1}^{d}y_i\right)^{n-2-l}\left(\prod_{i=1}^{d}y_i\right)^{l+1}\\
&\qquad\times\left(1-\prod_{i=1}^{t}x_i\right)^{l}\left(1-\prod_{i=t+1}^{d}x_i\right)^{l}d\mathbf{x}\,d\mathbf{y}\\
&=\sum_{l=0}^{n-2}\frac{n!}{l!(n-2-l)!}\frac{(l+1)!(n-2-l)!\mu_{n,d}(l+2)}{n!}\frac{\mu_{l+1,t}}{l+1}\frac{\mu_{l+1,d-t}}{l+1}\\
&=\sum_{l=1}^{n-1}\frac{1}{l}\mu_{n,d}(l+1)\mu_{l,t}\mu_{l,d-t}.\end{aligned}$$

Therefore,

$$\mathrm{E}(K_{n,d}^2)=\mu_{n,d}+\sum_{t=1}^{d-1}\binom{d}{t}\sum_{l=1}^{n-1}\frac{1}{l}\mu_{n,d}(l+1)\mu_{l,t}\mu_{l,d-t},$$

and we finally obtain (1.2).

Noting that in (1.2) the sum of those terms with at least two identical j indices in $\sum^{(1)}$ is at most $O((\log n)^{d-3})$, we further have

$$\begin{aligned}\mathrm{E}(K_{n,d}^2)=\mu_{n,d}+\sum_{t=1}^{d-1}\binom{d}{t}\sum_{l=1}^{n-1}\frac{1}{l}\sum{}^{(*)}\frac{1}{i_1\cdots i_{d-2}j_1\cdots j_{d-1}}\\
+O\big((\log n)^{d-3}\big),\end{aligned}\tag{2.4}$$

where the last summation $\sum^{(*)}$ is extended over all indices satisfying the inequalities

$$1\le i_1\le\cdots\le i_{t-1}\le l,\qquad 1\le i_t\le\cdots\le i_{d-2}\le l$$

and

$$l<j_1<\cdots<j_{d-1}\le n.$$

Now, let us compare the second term in the above expression (2.4) with $\mu_{n,d}^2$. By (2.3),

$$\mu_{n,d}^2=\sum{}^{(2)}\frac{1}{i_1\cdots i_{d-1}j_1\cdots j_{d-1}},$$

where the summation $\sum^{(2)}$ runs over all combinations

$$1 \le i_1 \le \cdots \le i_{d-1} \le n \quad \text{and} \quad 1 \le j_1 \le \cdots \le j_{d-1} \le n.$$

Write the $(d-1)$st largest index among $\{i_1, \ldots, i_{d-1}, j_1, \ldots, j_{d-1}\}$ as l and the $d-1$ indices greater than or equal to l as $k_1, \ldots, k_{d-1}$. If the k_j's are not all distinct or if there is one $k_h = l$ $(1 \le h \le d-1)$, then the sum of all those terms is $O((\log n)^{d-2})$. Now, consider the sum of those terms for which the $d-1$ k-indices are distinct and not equal to l. Rearrange the k-indices as $l < k_1 < \cdots < k_{d-1}$. Suppose that there are t $(1 \le t \le d-1)$ indices (among l, $k_1, \ldots, k_{d-1}$) from the j-indices. There are $\binom{d}{t}$ such possibilities. Write the other $t-1$ i-indices as $1 \le i_1 \le \cdots \le i_{t-1} \le l$ and the remaining $d-1-t$ j-indices as $1 \le j_1 \le \cdots \le j_{d-1-t} \le l$. Note that the reindexing is unique if $i_{t-1} < l$ and $j_{d-1-t} < l$. However, ambiguity arises when $i_{t-1} = l$ or $j_{d-1-t} = l$. For, if there is a term with $i_{t-1} = l$ and $j_{d-t-1} = l$, then this term is counted once in the case when l is an i-index and the number of j-indices in $\{l, k_1, \ldots, k_{d-1}\}$ is t as well as once in the case when l is a j-index and the number of j-indices in $\{l, k_1, \ldots, k_{d-1}\}$ is $t+1$. Thus,

$$\begin{aligned}
\mu_{n,d}^2 = {} & \sum_{t=1}^{d-1} \binom{d}{t} \sum_{l=1}^{n-1} \frac{1}{l} \sum^{(*)} \frac{1}{i_1 \cdots i_{d-2} j_1 \cdots j_{d-1}} \\
& - \sum_{t=1}^{d-2} \binom{d-1}{t} \sum_{l=1}^{n-1} \frac{1}{l^2} \sum^{(**)} \frac{1}{i_1 \cdots i_{t-1} j_1 \cdots j_{d-2-t} k_1 \cdots k_{d-1}} \\
& + O\big((\log n)^{d-2}\big),
\end{aligned}$$

where the summation $\sum^{(**)}$ is extended over all indices satisfying

$$1 \le i_1 \le \cdots \le i_{t-1} \le l, \qquad 1 \le j_1 \le \cdots \le j_{d-2-t} \le l$$

and

$$l < k_1 < \cdots < k_{d-1} \le n.$$

Therefore, we finally obtain

$$\begin{aligned}
\operatorname{Var}(K_{n,d}) = {} & \mu_{n,d} + \sum_{t=1}^{d-2} \binom{d-1}{t} \sum_{l=1}^{n-1} \frac{1}{l^2} \sum^{(**)} \frac{1}{i_1 \cdots i_{t-1} j_1 \cdots j_{d-2-t} k_1 \cdots k_{d-1}} \\
& + O\big((\log n)^{d-2}\big) \\
= {} & \left(\frac{1}{(d-1)!} + c_d \right) (\log n)^{d-1} \big(1 + O\big((\log n)^{-1}\big)\big),
\end{aligned}$$

where

$$c_d = \sum_{t=1}^{d-2} \frac{1}{t!(d-1-t)!} \sum_{l \ge 1} \frac{1}{l^2} \sum_{\substack{1 \le i_1 \le \cdots \le i_{t-1} \le l \\ 1 \le j_1 \le \cdots \le j_{d-2-t} \le l}} \frac{1}{i_1 \cdots i_{t-1} j_1 \cdots j_{d-2-t}}.$$

Using the finite difference formula

$$\mu_{n,h} = \sum_{1\le i_1\le i_2\le\cdots\le i_h\le n} \frac{1}{i_1 i_2\cdots i_h} = \sum_{1\le j\le n}\binom{n}{j}\frac{(-1)^{j-1}}{j^h}, \qquad h=1,2,\dots,$$

we obtain (1.5) and complete the proof of the theorem. □

3. Algorithmic applications. In this section we briefly discuss an implication of our main result: the asymptotic linearity of the variance of the cost of maxima-finding algorithms using divide-and-conquer paradigm.

There exists a large number of algorithms for finding the maxima in a given set of points [cf. Preparata and Shamos (1985), Bentley, Clarkson and Levine (1993), Devroye (1997)]. A naive divide-and-conquer algorithm runs as follows [cf. Devroye (1983)]. Divide the points $\{\mathbf{x}_1,\dots,\mathbf{x}_n\}$ into two groups $\{\mathbf{x}_1,\dots,\mathbf{x}_{\lfloor n/2\rfloor}\}$ and $\{\mathbf{x}_{\lfloor n/2\rfloor+1},\dots,\mathbf{x}_n\}$, where $\lfloor y\rfloor$ denotes the largest integer less than or equal to y. Find (recursively) the (set of) maxima of each group, denoted by M_1 and M_2, respectively. Then find, by pairwise comparisons, the maxima of M_1 and M_2. Note that the randomness is preserved in the process. The worst case behavior of this algorithm is obviously quadratic in n. But the expected number of comparisons as well as the variance are both linear under our uniform distribution assumption. This is seen by noting that both quantities satisfy recurrences of the form

$$f_n = f_{\lfloor n/2\rfloor} + f_{\lfloor (n+1)/2\rfloor} + g_n,$$

for $n\ge n_0\ge 1$ with suitable initial conditions and that $g_n = O((\log n)^{2d-2})$ for mean and $g_n = O((\log n)^{3d-3})$ for variance. It follows that $f_n = O_d(n)$, where the implied constant depends on d. The linear terms are oscillating in nature; see Flajolet and Golin (1993).

Other divide-and-conquer algorithms, such as that in Bentley and Shamos (1978), can also be shown to have linear variance for its cost.

Acknowledgments. The authors thank Luc Devroye for making his preprint available to them. They also thank the referee for pointing out an error in an earlier version of the paper.

REFERENCES

BAILEY, D. D., BORWEIN, J. M. and GIRGENSOHN, R. (1994). Experimental evaluation of Euler sums. *Experiment. Math.* **3** 17–30.

BARNDORFF-NIELSEN, O. and SOBEL, M. (1966). On the distribution of the number of admissible points in a vector random sample. *Theory Probab. Appl.* **11** 249–269.

BECKER, R. A., DENBY, L., MCGILL, R. and WILKS, A. R. (1987). Analysis of data from Places Rated Almanac. *Amer. Statist.* **41** 169–186.

BENTLEY, J. L., CLARKSON, K. L. and LEVINE, D. B. (1993). Fast linear expected-time algorithms for computing maxima and convex hulls. *Algorithmica* **9** 168–183.

BENTLEY, J. L., KUNG, H. T., SCHKOLNICK, M. and THOMPSON, C. D. (1978). On the average number of maxima in a set of vectors and applications. *J. Assoc. Comput. Mach.* **25** 536–543.

BENTLEY, J. L. and SHAMOS, M. I. (1978). Divide and conquer for linear expected time. *Inform. Process. Lett.* **7** 87–91.

BEREZOVSKII, B. A. and TRAVKIN, S. I. (1975). Supervision of queues of requests in computer systems. *Automat. Remote Control* **36** 1719–1725.

BUCHTA, C. (1989). On the average number of maxima in a set of vectors. *Inform. Process. Lett.* **33** 63–65.

DEVROYE, L. (1980). A note on finding convex hulls via maximal vectors. *Inform. Process. Lett.* **11** 53–56.

DEVROYE, L. (1983). Moment inequalities for random variables in computational geometry. *Computing* **30** 111–119.

DEVROYE, L. (1997). A note on the expected time for finding maxima by list algorithms. *Algorithmica*. To appear.

FLAJOLET, P. and GOLIN, M. (1993). Exact asymptotics of divide-and-conquer recurrences. *Lecture Notes in Comput. Sci.* **700** 137–149. Springer, Berlin.

FLAJOLET, P. and SALVY, B. (1996). Euler sums and contour integral representations. *Experiment. Math.* To appear.

GOLIN, M. J. (1993). How many maxima can there be? *Comput. Geom.* **2** 335–353.

HOFFMAN, M. E. (1992). Multiple harmonic sums. *Pacific J. Math.* **152** 275–290.

IVANIN, V. M. (1975a). Asymptotic estimate for the mathematical expectation of the number of elements in the Pareto set. *Cybernetics* **11** 108–113.

IVANIN, V. M. (1975b). Estimate of the mathematical expectation of the number of elements in a Pareto set. *Cybernetics* **11** 506–507.

IVANIN, V. M. (1976). Calculation of the dispersion of the number of elements of the Pareto set for the choice of independent vectors with independent components. In *Theory of Optimal Decisions* 90–100. Akad. Nauk. Ukrain. SSR Inst. Kibernet., Kiev. (In Russian.)

O'NEILL, B. (1980). The number of outcomes in the Pareto-optimal set of discrete bargaining games. *Math. Oper. Res.* **6** 571–578.

PREPARATA, F. P. and SHAMOS, M. I. (1985). *Computational Geometry: An Introduction*. Springer, New York.

SHOLOMOV, L. A. (1983). Survey of estimational results in choice problems. *Engrg. Cybernetics* **21** 51–75.

ZAGIER, D. (1992). Values of zeta functions and their applications. In *First European Congress of Mathematics* **2** 497–512. Birkhäuser, Berlin.

Z.-D. BAI
DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
KENT RIDGE, SINGAPORE 0511
SINGAPORE

C.-C. CHAO
H.-K. HWANG
W.-Q. LIANG
INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI, 115
TAIWAN
E-MAIL: ccchao@stat.sinica.edu.tw
hkwang@stat.sinica.edu.tw
wqliang@stat.sinica.edu.tw