Publishers' page

Publishers' page

Publishers' page

Publishers' page

CONTENTS

An application of Stein's method to maxima in hypercubes

Z. D. Bai
Department of Statistics and Applied Probability, National University of Singapore
2 Science Drive 2, Singapore 117543, Republic of Singapore
E-mail: stabaizd@nus.edu.sg

Luc Devroye
School of Computer Science
McGill University
Montreal, H3A 2K6
Canada
E-mail: luc@cs.mcgill.ca

and
Tsung-Hsi Tsay
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan
E-mail: chonghi@stat.sinica.edu.tw

We show that the number of maxima in random samples taken from $[0,1]^d$ is asymptotically normally distributed. The method of proof relies on Stein's method and gives a convergence rate.

Contents

1. Introduction

A point $\mathbf{y}$ in $\mathbb{R}^d$ is said to be *dominated* by another point $\mathbf{x}$ if the (vector) difference $\mathbf{x} - \mathbf{y}$ has only nonnegative coordinates. We write $\mathbf{y} \prec \mathbf{x}$. The points in a given sample that are not dominated by any other points are called *maxima*. We derive in this short note a central limit theorem (CLT) for the number of maxima in random samples independently and identically distributed (iid) in the hypercube $[0,1]^d$. A proof with the same rate was given previously in our paper Bai *et al.* (2004). We provide an alternative proof here using more original ideas introduced by Stein, which, in addition to methodological interest, also sheds more light on the complexity of the problem.

For concrete motivations and information regarding dominance and maxima, we refer the reader to our paper Bai *et al.* (2004).

Maxima in hypercubes. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sequence of iid points chosen uniformly at random from $[0,1]^d$, $d \geq 2$. Denote by $K_n = K_{n,d}$ the number of maxima in $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

The mean of K_n is known to be

$$\mathbb{E}[K_{n,d}] = \frac{(\log n)^{d-1}}{(d-1)!} \left(1 + O((\log n)^{-1})\right), \tag{1}$$

for bounded d; see Bai *et al.* (2004) and the references therein for more information.

The variance satisfies (see Bai *et al.*, 1998)

$$\frac{\mathbb{V}[K_n]}{(\log n)^{d-1}} = \left(\frac{1}{(d-1)!} + \kappa_d\right) \left(1 + O\left((\log n)^{-1}\right)\right), \tag{2}$$

where

$$\kappa_d = \sum_{1 \leq k \leq d-2} \frac{1}{k!(d-1-k)!(k-1)!(d-2-k)!} \int_0^1 \int_0^1 \frac{(-\log x)^{k-1}(-\log z)^{d-2-k}}{x+z-xz} \, dx \, dz,$$

is a bounded constant for $d \geq 2$. An asymptotic expansion for $\mathbb{V}[K_n]$ was derived in Bai *et al.* (2004).

A Berry-Esseen bound for $K_{n,d}$. Suppose that $Y_1, Y_2, \ldots$ is a sequence of random variables. Write $\{Y_n\} \in CLT(r_n)$, if

$$\sup_x \left| P\left(\frac{Y_n - \mathbb{E}[Y_n]}{\sqrt{\mathbb{V}[Y_n]}} < x\right) - \Phi(x) \right| = O\left(r_n\right),$$

where $r_n \to 0$ and $\Phi(x)$ is the standard normal distribution function. A sequence r_n will be referred to as a *convergent sequence* if $r_n \to 0$.

We will construct a sequence of random variables $K_{n,w}$ satisfying the following two theorems.

Theorem 1.1: For a convergent sequence $r_n \geq \Omega((\ln n)^{-\frac{d-1}{2}})$,

$$\{K_n\} \in CLT\,(r_n) \text{ iff } \{K_{n,w}\} \in CLT\,(r_n)\,.$$

Theorem 1.2: The normalized random variables $K^*_{n,w} := (K_{n,w} - \mathbb{E}[K_{n,w}])/\sqrt{\mathbb{V}[K_{n,w}]}$ converge to the standard normal distribution with a rate

$$d_1\left(K^*_{W_n}, \mathcal{X}\right) = O\left((\log\log n)^{2d} (\log n)^{-\frac{d-1}{2}}\right),$$

where $\mathcal{X}$ denotes the standard normal distribution and

$$d_1(X,Y) := \sup\left\{|\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]| : \sup_x |h(x)| + \sup_x |h'(x)| \leq 1\right\}.$$

From Theorem 1.2, it is easy to derive a rate for the Kolmogorov distance between the distribution of $(K_n - \mathbb{E}[K_n])/\sqrt{\mathbb{V}[K_n]}$ and that of a standard normal.

Theorem 1.3:

$$\{K_n\} \in CLT\left((\log\log n)^d (\log n)^{-\frac{d-1}{4}}\right).$$

Indeed, Theorem 1.3 follows from Theorem 1.2 and the fact that

$$\begin{aligned}\mathbb{E}\left[h\left(\frac{K_n - \mathbb{E}[K_n]}{\sqrt{\mathbb{V}[K_n]}} + \sqrt{r_n}\right)\right] &\leq \sqrt{r_n}\,\mathbb{P}\left(\frac{K_n - \mathbb{E}[K_n]}{\sqrt{\mathbb{V}[K_n]}} < x\right) \\ &\leq \mathbb{E}\left[h\left(\frac{K_n - \mathbb{E}[K_n]}{\sqrt{\mathbb{V}[K_n]}}\right)\right],\end{aligned}$$

where

$$h(u) = \begin{cases} \sqrt{r_n}, & \text{if } u \leq x, \\ 0, & \text{if } u > x + \sqrt{r_n}, \\ \text{linear, otherwise,} \end{cases}$$

and $\sqrt{r_n} = (\log\log n)^d (\log n)^{-\frac{d-1}{4}}$.

The proof given in this short note is similar to that given in Bai *et al.* (2004) since they both relies on the log-transformation first introduced by Baryshnikov (2000) and Stein's method. The main difference is that we give here a more self-contained proof based on Stein's original procedures (instead of just applying a theorem formulated in the book Janson *et al.* 2000). Also the conditioning arguments used in Bai *et al.* (2004) are replaced by more direct calculations.

2. CLT for K_n

As in Bai *et al.* (2004), the proof of Theorem 1.1 is divided into several steps.

The log-transformation. Assume now that $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are iid points uniformly distributed in the cube $(-1,0)^d$. The transformation $\mathbf{x} = (x_1, \ldots, x_d) \to \mathbf{y} = (y_1, \ldots, y_d)$, where (see Baryshinikov, 2000)

$$y_i = -\log(-x_i), \quad i = 1, \ldots, d,$$

from $(-1,0)^d$ to $\mathbb{R}_+^d = \{\mathbf{x} : x_i > 0 \text{ for all } i = 1, \ldots, d\}$, preserves the dominance relation and thus transforms exactly maximal point to maximal point. Denote by $\mathbf{y}_1, \ldots, \mathbf{y}_n$ the images of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ under such a transformation. Then the components of $\mathbf{y}_1$ are i.i.d. with exponential distribution ($\lambda = 1$). We define $\|\mathbf{x}\| = x_1 + \cdots + x_d$ for $\mathbf{x} \in \mathbb{R}_+^d$. Then $\|\mathbf{y}_1\|$ has a gamma distribution with parameter $(d, 1)$, i.e., $\|\mathbf{y}_1\|$ has the density function $\frac{x^{d-1}}{(d-1)!}e^{-x}$.

Approximation to K_n by the number of maxima in a strip. Let $B_\alpha = \{\mathbf{x} : \|\mathbf{x}\| > \alpha\} \cap \mathbb{R}_+^d$ and $B_\alpha^c = \{\mathbf{x} : \|\mathbf{x}\| \le \alpha\} \cap \mathbb{R}_+^d$. Take

$$\alpha = \ln n - \ln\left(4(d-1)\ln\ln n\right),$$
$$\beta = \ln n + 4(d-1)\ln\ln n.$$

Let $\widetilde{K}_n$ be the number of maxima of the points falling in the strip $T := B_\alpha \cap B_\beta^c$. We prove that for a convergent sequence $r_n \ge \Omega((\ln n)^{-\frac{d-1}{2}})$,

$$\{K_n\} \in CLT\,(r_n) \text{ iff } \{\widetilde{K}_n\} \in CLT\,(r_n). \tag{3}$$

To prove (3), we use the following Lemma whose proof is omitted.

Lemma 2.1: *Let X_n, Y_n be two sequences of random variables and r_n be a convergent sequence. Suppose that (i) the total variation distance $d(X_n, Y_n)$ between X_n and Y_n is bounded above by $O(r_n)$, (ii)*

$$|\mathbb{E}[X_n] - \mathbb{E}[Y_n]| = O(r_n\sqrt{\mathbb{V}[X_n]}),$$

and (iii)

$$|\mathbb{V}[X_n] - \mathbb{V}[Y_n]| = O(r_n \sqrt{\mathbb{V}[X_n]}).$$

Then $\{X_n\} \in CLT(r_n)$ *iff* $\{Y_n\} \in CLT(r_n)$.

Let $N_n(A)$ denote the number of points of $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ falling in A. Denote by $K_n(A)$ the number of maxima of $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ falling in A and by V_n the event that no points of $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ fall in B_β. Clearly, $K_n(A) \leq N_n(A)$. Note that maximal points contributing to $\widetilde{K}_n$ may not be maximal points contributing to K_n when $N_n(B_\beta) > 0$. However, we have $K_n(B_\alpha)1_{V_n} = \widetilde{K}_n 1_{V_n}$, which implies that

$$K_n = \widetilde{K}_n 1_{V_n} + K_n(B_\alpha^c)1_{V_n} + K_n 1_{V_n^c}. \tag{4}$$

To apply Lemma 2.1, we need to estimate the following quantities:

$$d(K_n, \widetilde{K}_n) \leq \mathbb{P}(K_n 1_{V_n^c} \geq 1) + \mathbb{P}(K_n(B_\alpha^c)1_{V_n} \geq 1) + \mathbb{P}(\widetilde{K}_n 1_{V_n^c} \geq 1),$$
$$\left|\mathbb{E}[K_n] - \mathbb{E}[\widetilde{K}_n]\right| \leq \mathbb{E}[\widetilde{K}_n 1_{V_n^c}] + \mathbb{E}[K_n(B_\alpha^c)] + \mathbb{E}[K_n 1_{V_n^c}],$$
$$\left|\mathbb{V}[K_n] - \mathbb{V}[\widetilde{K}_n]\right| \leq \left|\mathbb{E}[K_n^2] - \mathbb{E}[\widetilde{K}_n^2]\right| + \left|\mathbb{E}[K_n] - \mathbb{E}[\widetilde{K}_n]\right| \left(\mathbb{E}[K_n] + \mathbb{E}[\widetilde{K}_n]\right),$$

where

$$\left|\mathbb{E}[K_n^2] - \mathbb{E}[\widetilde{K}_n^2]\right| = \left|\mathbb{E}[K_n^2 1_{V_n^c}] + \mathbb{E}[K_n^2(B_\alpha^c)1_{V_n}] - \mathbb{E}[\widetilde{K}_n^2 1_{V_n^c}]\right|.$$

Observe that $1_{V_n^c} \leq N_n(B_\beta)$,

$$\begin{aligned}
\mathbb{E}[N_n(B_\beta)] &= n\mathbb{P}\left(\|\mathbf{y}_1\| \geq \beta\right) \\
&= n \int_\beta^\infty \frac{x^{d-1}}{(d-1)!} e^{-x} dx \\
&= O\left(n\beta^{d-1}e^{-\beta}\right) \\
&= O\left((\ln n)^{-3(d-1)}\right),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[N_n(T)] &= n\mathbb{P}\left(\alpha \leq \|\mathbf{y}_1\| \leq \beta\right) \\
&= n \int_\alpha^\beta \frac{x^{d-1}}{(d-1)!} e^{-x} dx \\
&= O\left(n\alpha^{d-1}e^{-\alpha}\right) \\
&= O\left((\ln n)^{d-1} \ln \ln n\right).
\end{aligned}$$

Recall that $\mathbb{E}[K_n] \asymp (\ln n)^{d-1}$ and $\mathbb{E}[K_n^2] \asymp (\ln n)^{2(d-1)}$; see (1) and (2). We show that $\mathbb{E}[\widetilde{K}_n]$ and $\mathbb{E}[\widetilde{K}_n^2]$ have the similar asymptotic order. By (4),

$$\begin{aligned}\widetilde{K}_n &= \widetilde{K}_n 1_{V_n^c} + K_n - K_n(B_\alpha^c)1_{V_n} - K_n 1_{V_n^c}\\ &\le \widetilde{K}_n 1_{V_n^c} + K_n\\ &\le N_n(T)1_{V_n^c} + K_n.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}\widetilde{K}_n &\le \mathbb{E}[N_n(T)N_n(B_\beta)] + \mathbb{E}[K_n]\\ &= \mathbb{E}[K_n] + n(n-1)\mathbb{P}(\mathbf{y}_1 \in T)\mathbb{P}(\mathbf{y}_2 \in B_\beta)\\ &= \mathbb{E}[K_n] + \mathbb{E}[N_n(B_\beta)]\mathbb{E}[N_{n-1}(T)]\\ &= O((\ln n)^{d-1}),\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[\widetilde{K}_n^2] &\le 2\mathbb{E}[N_n^2(T)N_n(B_\beta)] + 2\mathbb{E}[K_n^2]\\ &= 2\mathbb{E}[K_n^2] + 2n(n-1)\mathbb{P}(\mathbf{y}_1 \in T)\mathbb{P}(\mathbf{y}_2 \in B_\beta) + 4n(n-1)(n-2)\mathbb{P}^2(\mathbf{y}_1 \in T)\mathbb{P}(\mathbf{y}_2 \in B_\beta)\\ &= 2\mathbb{E}[K_n^2] + 2\mathbb{E}[N_n(B_\beta)]\mathbb{E}[N_{n-1}(T)] + 4\mathbb{E}[N_n(B_\beta)]\mathbb{E}[N_{n-1}(T)]\mathbb{E}[N_{n-2}(T)]\\ &= O((\ln n)^{2(d-1)}).\end{aligned}$$

Estimates needed. We now claim that

(i) $\mathbb{E}[K_n(B_\alpha^c)] = O((\ln n)^{-3(d-1)})$,
(ii) $\mathbb{E}[K_n 1_{V_n^c}] = O((\ln n)^{-2(d-1)})$,
(ii′) $\mathbb{E}[\widetilde{K}_n 1_{V_n^c}] = O((\ln n)^{-2(d-1)})$,
(iii) $\mathbb{E}[K_n^2 1_{V_n^c}] = O((\ln n)^{-(d-1)})$,
(iii′) $\mathbb{E}[\widetilde{K}_n^2 1_{V_n^c}] = O((\ln n)^{-(d-1)})$ and
(iv) $\mathbb{E}[K_n^2(B_\alpha^c)] = O((\ln n)^{-2(d-1)})$.

From these it follows that

$$d(K_n, \widetilde{K}_n) = O((\ln n)^{-2(d-1)}),$$

$$\left|\mathbb{E}[K_n] - \mathbb{E}[\widetilde{K}_n]\right| = O((\ln n)^{-2(d-1)}),$$

and

$$\left|\mathbb{V}[K_n] - \mathbb{V}[\widetilde{K}_n]\right| = O((\ln n)^{-(d-1)}).$$

Proof of (i)**.** If $\mathbf{y}$ is a maximal point, then there are no points in the region $C_{\mathbf{y}} = \{\mathbf{z} : \ z_i > y_i, i \le d\}$. The probability that $\mathbf{y}_1$ falls in $C_{\mathbf{y}}$ is

$$\int_{\|\mathbf{y}\|}^{\infty} \frac{(u - \|\mathbf{y}\|)^{d-1}}{(d-1)!} e^{-u} du = e^{-\|\mathbf{y}\|}.$$

Therefore, for large n

$$\begin{aligned}
\mathbb{E}[K_n(B_\alpha^c)] &= n \int_0^\alpha (1 - e^{-y})^{n-1} \frac{y^{d-1}}{(d-1)!} e^{-y} dy \\
&\le n \int_0^\alpha \frac{\alpha^{d-1}}{(d-1)!} e^{-y-(n-1)e^{-y}} dy \\
&= O\left(n(n-1)^{-1} \alpha^{d-1} e^{-(n-1)e^{-\alpha}} \right) \\
&= O\left((\ln n)^{-3(d-1)} \right).
\end{aligned}$$

Proof of (ii)**.** Note that

$$\begin{aligned}
G_{n:n} &:= \{\mathbf{y}_1 \text{ is a maximum in } \{\mathbf{y}_1, \cdots, \mathbf{y}_n\}\} \\
&\subset G_{n:n-1} := \{\mathbf{y}_1 \text{ is a maximum in } \{\mathbf{y}_1, \cdots, \mathbf{y}_{n-1}\}\}.
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}[K_n N_n(B_\beta)] &\le n\mathbb{P}\left(\|\mathbf{y}_1\| \ge \beta\right) + n(n-1)\mathbb{P}(G_{n:n} \cap \{\mathbf{y}_n \in B_\beta\}) \\
&\le n\mathbb{P}\left(\|\mathbf{y}_1\| \ge \beta\right) + n(n-1)\mathbb{P}(G_{n:n-1} \cap \{\mathbf{y}_n \in B_\beta\}) \\
&= \mathbb{E}[N_n(B_\beta)] + \mathbb{E}[K_{n-1}]\mathbb{E}[N_n(B_\beta)] \\
&= O\left((\ln n)^{-3(d-1)} \right) + O\left((\ln n)^{(d-1)} \right) O\left((\ln n)^{-3(d-1)} \right) \\
&= O\left((\ln n)^{-2(d-1)} \right).
\end{aligned}$$

Proof of (iii)**.**

$$\begin{aligned}
\mathbb{E}[K_n^2 N_n(B_\beta)] &\le n\mathbb{P}\left(\|\mathbf{y}_1\| \ge \beta\right) + 3n(n-1)\mathbb{P}(G_{n:n} \cap \{\mathbf{y}_n \in B_\beta\}) \\
&\quad + n(n-1)(n-2)\mathbb{P}(F_{n:n} \cap \{\mathbf{y}_n \in B_\beta\}) \\
&\le \mathbb{E}[N_n(B_\beta)] + 3\mathbb{E}[K_{n-1}]\mathbb{E}[N_n(B_\beta)] + n(n-1)(n-2)\mathbb{P}(F_{n:n-1} \cap \{\mathbf{y}_n \in B_\beta\}) \\
&\le \mathbb{E}[N_n(B_\beta)] + 3\mathbb{E}[K_{n-1}]\mathbb{E}[N_n(B_\beta)] + \mathbb{E}[K_{n-1}^2]\mathbb{E}[N_n(B_\beta)] \\
&= O\left((\ln n)^{-(d-1)} \right),
\end{aligned}$$

where

$$\begin{aligned}
F_{n:n} &:= \{\mathbf{y}_1, \mathbf{y}_2 \text{ are two maxima in } \{\mathbf{y}_1, \cdots, \mathbf{y}_n\}\} \\
&\subset F_{n:n-1} := \{\mathbf{y}_1, \mathbf{y}_2 \text{ are two maxima in } \{\mathbf{y}_1, \cdots, \mathbf{y}_{n-1}\}\}.
\end{aligned}$$

Proof of (ii')**.** Similarly as above, we have

$$\begin{aligned}\mathbb{E}[\widetilde{K}_n N_n(B_\beta)] &\le n(n-1)\mathbb{P}(\mathbf{y}_1 \in T \text{ not dominated by points in } T \text{ and } \mathbf{y}_2 \in B_\beta)\\ &\le \mathbb{E}[\widetilde{K}_{n-1}]\mathbb{E}[N_n(B_\beta)]\\ &= O\left((\ln n)^{-2(d-1)}\right).\end{aligned}$$

Proof of (iii')**.**

$$\begin{aligned}\mathbb{E}[\widetilde{K}_n^2 N_n(B_\beta)] &\le 2n(n-1)\mathbb{P}(\mathbf{y}_1 \in T \text{ not dominated by points in } T \text{ and } \mathbf{y}_2 \in B_\beta)\\ &\quad + n(n-1)(n-2)\mathbb{P}(\mathbf{y}_1, \mathbf{y}_2 \in T \text{ not dominated by points in } T \text{ and } \mathbf{y}_3 \in B_\beta)\\ &\le 2\mathbb{E}[\widetilde{K}_{n-1}]\mathbb{E}[N_n(B_\beta)] + \mathbb{E}[\widetilde{K}_{n-1}^2]\mathbb{E}[N_n(B_\beta)]\\ &= O\left((\ln n)^{-(d-1)}\right).\end{aligned}$$

Proof of (iv)**.** Given $\mathbf{y}_1, \mathbf{y}_2$, the conditional probability that $\mathbf{y}_3$ falls in $C_{\mathbf{y}_1} \cup C_{\mathbf{y}_2}$ is

$$\mathbb{P}(C_{\mathbf{y}_1}) + \mathbb{P}(C_{\mathbf{y}_2}) - \mathbb{P}(C_{\mathbf{y}_1} \cap C_{\mathbf{y}_2}) \ge \frac{1}{2}\left(e^{-\|\mathbf{y}_1\|} + e^{-\|\mathbf{y}_2\|}\right);$$

the conditional probability that both $\mathbf{y}_1$ and $\mathbf{y}_2$ are maximal is less than

$$\left(1 - \frac{1}{2}\left(e^{-\|\mathbf{y}_1\|} + e^{-\|\mathbf{y}_2\|}\right)\right)^{n-2} \le e^{-\frac{1}{2}(n-2)\left(e^{-\|\mathbf{y}_1\|}+e^{-\|\mathbf{y}_2\|}\right)}.$$

We thus have

$$\begin{aligned}\mathbb{E}[K_n^2(B_\alpha^c)] &= \mathbb{E}\left[\sum_{i=1}^n 1_{\mathbf{y}_i \text{ is maxima and} \|\mathbf{y}_i\| \le \alpha}\right]^2\\ &= \mathbb{E}[K_n(B_\alpha^c)] + n(n-1)\mathbb{P}(\text{both } \mathbf{y}_1 \text{ and } \mathbf{y}_2 \text{ are maxima falling in } B_\alpha^c)\\ &\le \mathbb{E}[K_n(B_\alpha^c)] + \frac{n^2}{[(d-1)!]^2}\int_0^\alpha\int_0^\alpha (xy)^{d-1} e^{-\frac{1}{2}(n-2)[e^{-x}+e^{-y}]-x-y}dx\\ &\le \mathbb{E}[K_n(B_\alpha^c)] + \frac{n^2(\ln n)^{2(d-1)}}{[(n-2)(d-1)!]^2}e^{-(n-2)e^{-\alpha}}\\ &= O\left((\ln n)^{-2(d-1)}\right).\end{aligned}$$

Approximation by Poisson process. Construct a Poisson process $\{\mathbf{W}_n\}$ on T with intensity function $\lambda_n = ne^{-\|\mathbf{w}\|}$. Denote by N_w the number of points of the Poisson process falling in T. Also, let $K_{n,w}$ denote the number of maxima of the Poisson process and $\widetilde{N}_n$ be the number of points of $\{\mathbf{y}_1, \cdots, \mathbf{y}_n\}$ that falls in T. It is easy to see that the conditional distribution of $\widetilde{K}_n$ given $\widetilde{N}_n = m$ is identical to the conditional distribution of

$K_{n,w}$ given $N_w = m$. Thus, the total variation distance between $\widetilde{K}_n$ and $K_{n,w}$ satisfies

$$\sup_A \left|\mathbb{P}(\widetilde{K}_n \in A) - \mathbb{P}(K_{n,w} \in A)\right|$$
$$= \sup_A \left| \sum_{0\le m\le n} \mathbb{P}(\widetilde{N}_n = m)\mathbb{P}(\widetilde{K}_n \in A|\widetilde{N}_n = m) - \sum_{0\le m<\infty} \mathbb{P}(N_w = m)\mathbb{P}(K_{n,w} \in A|N_w = m)\right|$$
$$\le \sum_{0\le m\le n} \left|\mathbb{P}(\widetilde{N}_n = m) - \mathbb{P}(N_w = m)\right| + \sum_{n<m<\infty} \mathbb{P}(N_w = m)$$
$$\le O(p_n),$$

(see Prohorov, 1953) where

$$p_n := P(\mathbf{y}_1 \in T) = \int_\alpha^\beta \frac{x^{d-1}}{(d-1)!} e^{-x} dx = O\left(\frac{(\ln n)^{d-1} \ln\ln n}{n}\right).$$

Similarly, we have

$$\left|\mathbb{E}[\widetilde{K}_n] - \mathbb{E}[K_{n,w}]\right| \le np_n^2,$$

and

$$\left|\mathbb{E}[\widetilde{K}_n(\widetilde{K}_n - 1)] - \mathbb{E}[K_{n,w}(K_{n,w} - 1)]\right| \le n(n-1)p_n^3.$$

The above three estimates imply that for a convergent sequence $r_n \ge \Omega((\ln n)^{-\frac{d-1}{2}})$,

$$\{\widetilde{K}_n\} \in CLT\,(r_n) \text{ iff } \{K_{n,w}\} \in CLT\,(r_n)\,. \tag{5}$$

3. A central limit theorem for $K_{n,w}$

We prove in this section Theorem 1.2. We first give a lemma on Stein's method.

Let $h(x)$ be a function such that

$$\sup_x |h(x)| + \sup_x |h'(x)| \le 1. \tag{6}$$

Let f be the solution of the differential equation

$$xf(x) - f'(x) = h(x) - Eh,$$

where

$$Eh = \mathbb{E}[h(\mathcal{X})] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(x)e^{-x^2/2}dx,$$

where $\mathcal{X}$ is the standard normal variable.

Let Z_v be a set of random variables and

$$\begin{aligned}
U_v &= \{j : Z_j \text{ is dependent of } Z_v\}, \\
V_v &= \sum_{j \in U_v} Z_j, \\
U_{v,j} &= \{k : Z_k \text{ is dependent of } Z_v \text{ or } Z_j\}, \\
V_{v,j} &= \sum_{k \in U_{v,j}} Z_k, \\
S &= \sum_v Z_v, \\
S_v &= S - V_v, \\
S_{v,j} &= S - V_{v,j}.
\end{aligned}$$

Lemma 3.1: *Use the above notation and assume that* $\mathbb{E}[Z_v] = 0$ *and* $\mathbb{E}[S^2] = \sum_v \mathbb{E}[Z_v V_v] = 1$.

$$d_1(S, \mathcal{X}) \le C \sum_v \sum_{j \in U_v} \sum_{k \in U_{v,j} \cup U_v} \left(\mathbb{E}|Z_v Z_j Z_k| + \mathbb{E}|Z_v Z_j| \mathbb{E}|Z_k|\right).$$

The lemma is essentially the same as Theorem 6. 31 of Janson *et al* (2000, Page 158).

Split $\mathbb{R}_+^d$ into cubes of edge-length δ_n where δ_n is a small positive number to be specified later. Let Z_v denote the number of maxima of the Poisson process falling in the cell T_v (only cubes intersecting with T are counted). Set

$$K_{n,w} = \sum_v Z_v.$$

and

$$K_{n,w}^* = (K_{n,w} - \mathbb{E}[K_{n,w}])/\sqrt{\mathbb{V}[K_{n,w}]} = \sum_v (Z_v - \mathbb{E}[Z_v])/\sqrt{\mathbb{V}[K_{n,w}]}.$$

Replacing Z_v in Lemma 3.1 by $(Z_v - \mathbb{E}[Z_v])/\sqrt{\mathbb{V}[K_{n,w}]}$, we obtain

$$\begin{aligned}
&d_1(K_{n,w}^*, \mathcal{X}) \\
&\le C \mathbb{V}[K_{n,w}]^{-\frac{3}{2}} \sum_v \sum_{j \in U_v} \sum_{k \in U_{v,j} \cup U_v} \left(\mathbb{E}[Z_v Z_j Z_k] + \mathbb{E}[Z_v Z_j]\mathbb{E}[Z_k] + \mathbb{E}[Z_v]\mathbb{E}[Z_j]\mathbb{E}[Z_k]\right).
\end{aligned} \tag{7}$$

We now show that

(i) If $v \neq j$, then

$$\mathbb{E}[Z_v^{\ell_1} Z_j^{\ell_2}] \le \mathbb{E}[Z_v^{\ell_1} N_j^{\ell_2}] \le \mathbb{E}[Z_v^{\ell_1}]\mathbb{E}[N_j^{\ell_2}] \qquad (\ell_1, \ell_2 = 1, 2),$$

where N_j is the number of Poisson process points falling in the region T_j.

This follows from the fact that $\mathbb{E}[Z_v^{\ell_1} | N_j = m]$ is decreasing in m.

(ii) If v, j, k are pairwise distinct, then

$$\mathbb{E}[Z_v Z_j Z_k] \le \mathbb{E}[Z_v Z_j N_k] \le \mathbb{E}[Z_v]\mathbb{E}[N_j]\mathbb{E}[N_k].$$

Similar to the proof for (i), $\mathbb{E}[Z_v Z_j | N_k = m]$ is a decreasing function of m. Thus, $\mathbb{E}[Z_v Z_j N_k] \le \mathbb{E}[Z_v Z_j]\mathbb{E}[N_k]$. Then (ii) follows from (i).

Substituting these into (7), we obtain

$$d_1(K^*_{n,w}, \mathcal{X}) \le C\mathbb{V}[K_{n,w}]^{-\frac{3}{2}} \left(\sum_v \mathbb{E}[Z_v^3] + \sum_v \mathbb{E}[Z_v]^2 \sum_{j \in U_v} \mathbb{E}[N_j] + \sum_v \mathbb{E}[Z_v] \sum_{j \in U_v} \mathbb{E}[N_j] \sum_{k \in U_{v,j} \cup U_v} \mathbb{E}[N_k] \right). \tag{8}$$

Recall that $\mathbb{V}[K_{n,w}] \asymp (\ln n)^{d-1}$. Define

$$M_n = \sum_v \mathbb{E}[Z_v] \asymp (\ln n)^{d-1},$$
$$p_v = \int_{T_v} n e^{-\|\mathbf{y}\|} d\mathbf{y},$$
$$P_n = \int_T n e^{-\|\mathbf{y}\|} d\mathbf{y} \sim \frac{(\ln n)^{d-1}}{(d-1)!} n e^{-\alpha} \sim \frac{a(\ln n)^{d-1} \ln \ln n}{(d-1)!}.$$

Then we have

$$\begin{aligned}
\sum_v \mathbb{E}[Z_v^3] &= \sum_v \sum_{m \ge 1} \mathbb{E}[Z_v^3 | N_v = m] \frac{p_v^m}{m!} e^{-p_v} \\
&\le \sum_v \sum_{m \ge 1} \mathbb{E}[Z_v | N_v = m] m^2 \frac{p_v^m}{m!} e^{-p_v} \\
&\le 9 \sum_v \sum_{m \ge 1} \mathbb{E}[Z_v | N_v = m] \frac{p_v^m}{m!} e^{-p_v} + \sum_v \sum_{m \ge 4} m^3 \frac{p_v^m}{m!} e^{-p_v} \\
&\le 9M_n + 5 \sum_v p_v^4 \\
&\le 9M_n + 5 \max_v p_v^3 P_n.
\end{aligned}$$

(Recall that $\alpha = \ln n - \ln(4(d-1)\ln\ln n)$). If we choose T_v (i.e. δ_n) so small that

$$\max_v p_v^3 P_n \le 1/5,$$

then

$$\sum_v \mathbb{E}[Z_v^3] \le 9M_n + 1.$$

Similarly, we can prove that

$$\sum_v \mathbb{E}[Z_v^2] \le 3M_n + 1.$$

Combining the above estimates, we have

$$d_1(K_{n,w}^*, \mathcal{X}) \le C\mathbb{V}[K_{n,w})]^{-\frac{3}{2}} \left(M_n(1 + Q_1 + Q_2^2) + 1\right),$$

where

$$Q_1 = \max_v \sum_{j \in U_v} \mathbb{E}[N_j]$$
$$Q_2 = \max_{v,j} \sum_{k \in U_{v,j} \cup U_v} \mathbb{E}[N_k].$$

On the other hand, $Q_1 \le Q_2$ and

$$Q_2 = O\left((\ln\ln n)^{d-1} \int_\alpha^\beta n e^{-x} dx\right) = O((\ln\ln n)^d).$$

Therefore, we conclude that

$$d_1(K_{n,w}^*, \mathcal{X}) = O\left((\ln\ln n)^{2d} (\ln n)^{-\frac{d-1}{2}}\right).$$

References

1. Z.-D. Bai, C.-C. Chao, H.-K. Hwang and W.-Q. Liang (1998). On the variance of the number of maxima in random vectors and its applications. *Annals of Applied Probability*, **8** 886–895.
2. Z.-D. Bai, L. Devroye, H.-K. Hwang and T.-H. Tsai (2004). Maxima in hypercubes. Preprint submitted for publication.
3. Z.-D. Bai, H.-K. Hwang, W.-Q. Liang and T.-H. Tsai (2001). Limit theorems for the number of maxima in random samples from planar regions. *Electronic Journal of Probability*, **6** Article 3, 41 pp.
4. A. D. Barbour and A. Xia (2001). The number of two dimensional maxima. *Advances in Applied Probability*, **33** 727–750.

5. O. Barndorff-Nielsen and M. Sobel (1966). On the distribution of the number of admissible points in a vector random sample. *Theory of Probability and its Applications*, **11** 249–269.
6. Y. Baryshnikov (2000). Supporting-points processes and some of their applications. *Probability Theory and Related Fields* , **117** 163–182.
7. S. Janson, T. Luczak and A. Ruciński (2000). *Random Graphs*. John Wiley & Sons, Inc.
8. Yu. V. Prohorov (1953). Asymptotic behavior of the binomial distribution. in *Selected Translations in Mathematical Statistics and Probability*, Vol. 1, pp. 87–95, ISM and AMS, Providence, R.I. (1961); translation from Russian: *Uspehi Matematičeskih Nauk*, **8** (1953), no. 3 (35), 135–142.