DOI: 10.1214/EJP.v8-137

Electronic Journal of Probability

Vol. 8 (2003) Paper no. 9, pages 1–26.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol8/paper9.abs.html

BERRY-ESSEEN BOUNDS FOR THE NUMBER OF MAXIMA IN PLANAR REGIONS

Zhi-Dong Bai
Department of Matematics, Northeast Normal University
130024, Changchun, Jilin, PRC
and
Department of Statistics and Applied Probability, National University of Singapore
119260, Singapore
Email: stabaizd@leonis.nus.edu.sg
URL: http://www.stat.nus.edu.sg/%7Estabaizd/

Hsien-Kuei Hwang
Institute of Statistical Science, Academia Sinica
Taipei 115, Taiwan
Email: hkhwang@stat.sinica.edu.tw
URL: http://algo.stat.sinica.edu.tw

Tsung-Hsi Tsai
Institute of Statistical Science, Academia Sinica
Taipei 115, Taiwan
Email: chonghi@stat.sinica.edu.tw
URL: http://www.stat.sinica.edu.tw/chonghi/stat.htm

Abstract We derive the optimal convergence rate $O(n^{-1/4})$ in the central limit theorem for the number of maxima in random samples chosen uniformly at random from the right triangle of the shape ◺. A local limit theorem with rate is also derived. The result is then applied to the number of maxima in general planar regions (upper-bounded by some smooth decreasing curves) for which a near-optimal convergence rate to the normal distribution is established.

Submitted to EJP on July 18, 2002. Final version accepted on June 3, 2003.

1 Introduction

Given a sample of points in the plane (or in higher dimensions), the *maxima* of the sample are those points whose first quadrants are free from other points. More precisely, we say that $p_1 = (x_1, y_1)$ *dominates* $p_2 = (x_2, y_2)$ if $x_1 > x_2$ and $y_1 > y_2$; the maxima of a point set $\{p_1, \dots, p_n\}$ are those p_j's that are dominated by no points. The main purpose of this paper is to derive convergence rates in the central limit theorems (or Berry-Esseen bounds) for the number of maxima in samples chosen uniformly at random from some planar regions. As far as the Berry-Esseen bounds are concerned, very few results are known in the literature for the number of maxima (and even in the whole geometric probability literature): precise approximation theorems are known only in the $\log$-class (regions for which the number of maxima has logarithmic mean and variance); see Bai et al. (2001) for more precise results. We propose new tools for dealing with the $\sqrt{n}$-class (see Bai et al. (2001)) in this paper.

Such a dominance relation among points is a natural ordering relation for multidimensional data and is very useful, both conceptually and practically, in diverse disciplines; see Bai et al. (2001) and the references cited there for more information. For example, it was used in analyzing the 329 cities in United States selected in the book *Places Rated Almanac* (see Becker et al., 1987). Naturally, city A is "better" than city B if factors pertaining to the quality of life of city A are all better than those of city B. The same idea is useful for educational data: a student is "better" than another if all scores of the former are better than those of the latter; also a student should not be classified as "bad" if (s)he is not dominated by any others. While traditional ranking models relying on average or weighted average may prove unfair for someone with outstanding performance in only one subject and with poor performance in all others, the dominance relation provides more auxiliary information for giving a less "prejudiced" ranking of students.

Some recent algorithmic problems in computational geometry involving quantitatively the number of maxima can be found in Chan (1996), Emiris et al. (1997), Ganley (1999), Zachariasen (1999), Datta and Soundaralakshmi (2000).

To further motivate our study on the number of maxima, we mention (in addition to applications in computational geometry) yet another application of dominance to knapsack problems, which consists in maximizing the weighted sum $\sum_{1 \le j \le n} p_j x_j$ by choosing an appropriate vector $(x_1, \dots, x_n)$ with $x_j \in \{0, 1\}$, subject to the restriction $\sum_{1 \le j \le n} w_j x_j \le W$, where p_j, w_j and W are nonnegative numbers. Roughly, item i dominates item j if $w_i \le w_j$ and $p_i \ge p_j$, so that a good heuristic is that if $\lfloor w_j / w_i \rfloor p_i \ge p_j$ then item j can be discarded from further consideration; see Martello and Toth (1990). A probabilistic study on the number of undominated variables can be found in Johnston and Khan (1995), Dyer and Walker (1997). Similar dominance relations are also widely used in other combinatorial search problems.

Interestingly, the problem of finding the maxima of a sample of points in a bounded planar region can also be stated as an optimization problem: given a set of points in a bounded region, we seek to minimize the area between the "staircase" formed by the selected points and the upper-right boundary; see Figure 1 for an illustration. Obviously, the minimum value is achieved by the set of maxima.

Let $\mathcal{D}$ be a given region in $\mathbb{R}^2$. Denote by $M_n(\mathcal{D})$ the number of maxima in a random sample of n points chosen uniformly and independently in $\mathcal{D}$.

It is known that the expected number of maxima in bounded planar regions $\mathcal{D}$ exhibits generally three different modes of behaviors: $\sqrt{n}$, $\log n$ and bounded (see Golin, 1993; Bai et al., 2001). Briefly, if the region $\mathcal{D}$ contains an upper-right corner (a point on the boundary that dominates all other points inside and on the boundary), then $E(M_n(\mathcal{D}))$ is roughly either of order $\log n$ or bounded; otherwise, $E(M_n(\mathcal{D}))$ is of order $\sqrt{n}$.

The $\log n$-class was studied in details in Bai et al. (2001), where the analysis relies heavily on the case when $\mathcal{D}$ is a rectangle $\mathcal{R}$ (or a square). Basically, $M_n(\mathcal{D})$ can be expressed (in case when $\mathcal{D}$ has an upper-right corner) as $1 + M_{I_n}(\mathcal{R})$, where the distribution of I_n depends on the shape of $\mathcal{D}$. While rectangle is

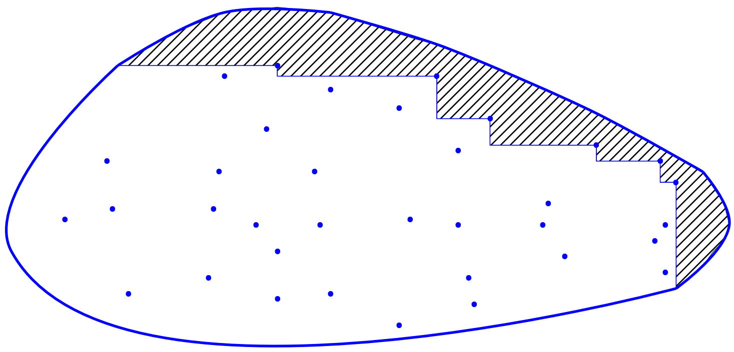

Figure 1: *The maxima-finding problem can be viewed as an optimization problem: minimizing the area in the shaded region (between the "staircase" formed by the selected points and the upper-right boundary).*

prototypical for $\log n$-class, we show in this paper that the right triangle $\mathcal{T}$ of the form ◺ plays an important role for the Berry-Esseen bound of $M_n(\mathcal{D})$ when the mean and the variance are of order $\sqrt{n}$. Thus we start by considering right triangles.

For simplicity, let $M_n = M_n(\mathcal{T})$, where $\mathcal{T}$ is the right triangle with corners $(0, 0)$, $(0, 1)$ and $(1, 0)$. It is known that (see Bai et al., 2001)

$$\frac{M_n - \sqrt{\pi n}}{\sigma n^{1/4}} \xrightarrow{d} N(0, 1), \tag{1}$$

where $\sigma^2 = (2 \log 2 - 1)\sqrt{\pi}$ and $N(0, 1)$ is a normal random variable with zero mean and unit variance. The mean and the variance of M_n satisfy

$$\mu_n := E(M_n) = \sqrt{\pi n} - 1 + O(n^{-1/2}), \tag{2}$$

$$\sigma_n^2 := \mathrm{Var}(M_n) = \sigma^2\sqrt{n} - \frac{\pi}{4} + O(n^{-1/2}). \tag{3}$$

See also Neininger and Rüschendorf (2002) for a different proof for (1) via contraction method.

We improve (1) by deriving an optimal (up to implied constant) Berry-Esseen bound and a local limit theorem for M_n. Let $\Phi(x)$ denote the standard normal distribution function.

Theorem 1 (Right triangle: Convergence rate of CLT).

$$\sup_{-\infty<x<\infty} \left| P\left(\frac{M_n - \mu_n}{\sigma_n} < x\right) - \Phi(x) \right| = O(n^{-1/4}). \tag{4}$$

Theorem 2 (Right triangle: Local limit theorem). *If $k = \lfloor \mu_n + x\sigma_n \rfloor$, then*

$$P(M_n = k) = \frac{e^{-x^2/2}}{\sqrt{2\pi}\,\sigma_n} \left(1 + O\left((1 + |x|^3)n^{-1/4}\right)\right), \tag{5}$$

uniformly for $x = o(n^{1/12})$.

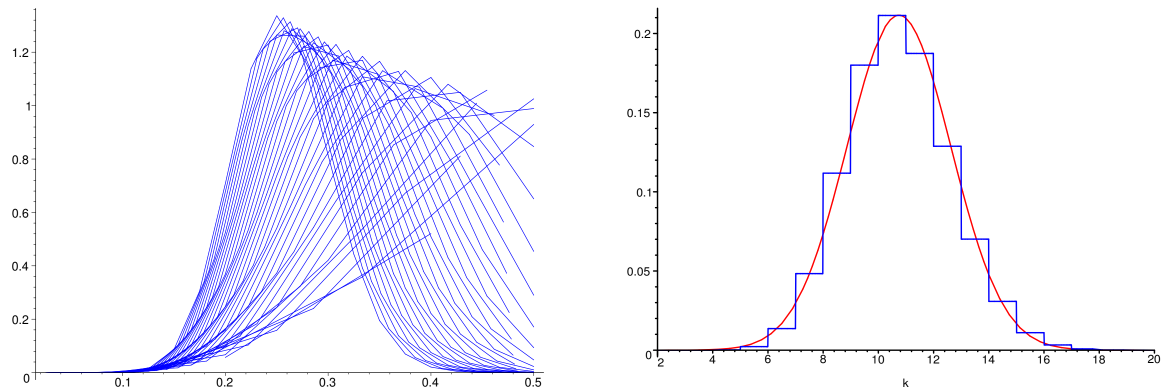

Figure 2: *Exact histograms of* $n^{1/2}P(M_n = \lfloor xn \rfloor)$ *for* n *from 5 to 40 and* $0 \le x \le 1/2$ *(left) and* $P(M_{40} = k)$ *for* $2 \le k \le 20$ *with the associated Gaussian density (right).*

Note that the same error terms in (4) and (5) hold if we replace μ_n or σ_n in the left-hand side by their asymptotic values $\sqrt{\pi n}$ and $\sigma n^{1/4}$, respectively.

The Berry-Esseen bound and the local limit theorem are derived by a refined method of moments introduced in Hwang (2003), coupling with some inductive arguments and Fourier analysis; the technicalities are quite different and more involved here. Roughly, We start the approach by considering the normalized moment generating functions $\phi_n(y) := E(e^{(M_n - N(\mu_n, \sigma_n^2))y})$. We next show that

$$|\phi_n^{(m)}(0)| = |E(M_n - N(\mu_n, \sigma_n^2))^m| \le m! A^m n^{m/6} \qquad (m \ge 0),$$

for a sufficiently large A. This is the hard part of the proof. Such a precise upper bound suffices for deriving the estimate

$$\left| E(e^{(M_n - \mu_n)iy/\sigma_n}) - e^{-y^2/2} \right| = O\left(n^{-1/4} |y|^3 e^{-y^2/2} \right),$$

uniformly for $|y| = o(n^{1/12})$. We then use another inductive argument to derive a uniform estimate for $E(e^{(M_n - \mu_n)iy/\sigma_n})$ for $|y| \le \pi \sigma_n$ and conclude (4) by applying the Berry-Esseen smoothing inequality (see Petrov, 1975) and (5) by the Fourier inversion formula.

Although the proof is not short, the results (4) and (5) are the first and the best, up to the implied constants, of their kind. Also the approach used (based on estimates of normalized moments) can be applied to other recursive random variables; it is therefore of some methodological interests *per se*. While the usual method of moments proves the convergence in distribution by establishing the stronger convergence of all moments, our method shows that in the case of a normal limit law the convergence of moments can sometimes be further refined and yields stronger quantitative results.

The result (4) will also be applied to planar regions upper-bounded by some smooth curves of the form

$$\mathcal{D} = \{(u, v) \,:\, 0 < u < 1, 0 < v < f(u)\},$$

where $f(u) > 0$ is a nonincreasing function on $(0, 1)$ with $f(0^+) < \infty$, $f(1^-) = 0$ and $\int_0^1 f(u)\,\mathrm{d}u = 1$.

Devroye (1993) showed that if f is either convex, or concave, or Lipschitz with order 1, then

$$E(M_n(\mathcal{D})) \sim \alpha \sqrt{\pi n}, \quad \text{where} \quad \alpha = \int_0^1 \sqrt{|f'(u)|/2}\,\mathrm{d}u.$$

Our result says that if f is smooth enough (roughly twice differentiable with $f' \neq 0$), then the number of maxima converges (properly normalized) in distribution to the standard normal distribution with a rate of order $n^{-1/4}\gamma_n \log^2 n$, where γ_n is some measure of "steepness" of f defined in (28); see Theorem 3 for a precise statement. While the order of γ_n can vary with f, it is bounded or at most logarithmic for most practical cases of interests. The method of proof of Theorem 3 is different from that for Theorem 1; it proceeds by splitting $\mathcal{D}$ into many smaller regions and then by transforming $\mathcal{D}$ in a way that $M_n(\mathcal{D})$ is very close to the number of maxima in some right triangle $\mathcal{T}_n$. Then we can apply (4). The hard part is that we need precise estimate for the difference between the number of maxima in $\mathcal{D}$ and those in the approximate triangle $\mathcal{T}_n$.

The proofs of Theorems 1 and 2 are given in the next section. We derive a Berry-Esseen bound for $M_n(\mathcal{D})$ for nondecreasing f in Section 3. We then conclude with some open questions.

Results related to ours have very recently been derived by Barbour and Xia (2001), where they study the bounded Wasserstein distance between the number of maxima in certain planar regions and the normal distribution. Since the bounded Wasserstein distance is in general larger than (roughly of the order $\mathrm{Var}(M_n(\mathcal{D}))^{-1/4}$) the Kolmogorov distance, our results are stronger as far as the Kolmogorov distance is concerned. On the other hand, our settings and approach are completely different. Their approach relies on the Stein method using arguments on point processes similar to those used in Chiu and Quine (1997). The latter paper studies the number of seeds in some stochastic growth model with inhomogeneous Poisson arrivals; in particular, since the point processes are assumed to be spatially homogeneous in Chiu and Quine (1997), Theorem 6.1 there can be translated into a Berry-Esseen bound for M_n (maxima in right triangle) with a rate of order $n^{-1/4}\log n$; see Barbour and Xia (2001) for some details. See also Baryshnikov (2000) for other limit theorems for maxima. While it is likely that Stein's method can be further improved to give an alternative proof of (4), it is unclear how such an approach can be used for proving our local limit theorem (5).

Notations. Throughout this paper, we use the generic symbols ε, c and B (without subscripts) to denote suitably small, absolute and large positive constants, respectively, whose values may vary from one occurrence to another. For convenience of reference, we also index these symbols by subscripts to denote constants with *fixed* values. The abbreviation "iid" stands for "independent and identically distributed". The convention $0^0 = 1$ is adopted.

2 Right triangle

Theorems 1 and 2 are proved by first establishing the following two estimates. Define

$$\varphi_n(y) := E\left(e^{iy(M_n-\mu_n)/\sigma_n}\right) = e^{-\mu_n iy/\sigma_n} P_n\left(\frac{iy}{\sigma_n}\right),$$

where $P_n(y) := E(e^{M_n y})$.

Proposition 1. *The estimate*

$$\left|\varphi_n(y) - e^{-y^2/2}\right| = O\left(n^{-1/4}|y|^3 e^{-y^2/2}\right) \tag{6}$$

holds uniformly for $|y| \le \varepsilon n^{1/12}$.

Proposition 2. *Uniformly for* $|y| \le \pi\sigma_n$ *and* $n \ge 2$,

$$|\varphi_n(y)| \le e^{-\varepsilon y^2}. \tag{7}$$

The hard part of the proof is the locally more precise bound (6).

Moment generating function. Our starting point is the recurrence for the moment generating function (see Bai et al., 2001)

$$P_n(y) = e^y \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) P_j(y) P_k(y), \tag{8}$$

for $n \geq 1$ with the initial condition $P_0(y) = 1$, where the sum is extended over all nonnegative integer triples (j, k, ℓ) such that $j + k + \ell = n - 1$ and

$$\pi_{j,k,\ell}(n) := \frac{(n-1)!}{j!k!\ell!} 2^\ell \int_0^1 x^{2j+\ell}(1-x)^{2k+\ell}\,\mathrm{d}x = \frac{(n-1)!}{j!k!\ell!} \frac{(2j+\ell)!(2k+\ell)!2^\ell}{(2n-1)!}. \tag{9}$$

Recurrences. We first prove the estimate (6), starting by defining the normalized moment generating function

$$\phi_n(y) := E(e^{(M_n - N(\mu_n, \sigma_n^2))y}) = e^{-\mu_n y - \sigma_n^2 y^2/2} P_n(y),$$

which satisfies, by (8), the recurrence $\phi_0(y) = 1$ and for $n \geq 1$

$$\phi_n(y) = \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) \phi_j(y) \phi_k(y) e^{\Delta y + \delta y^2}, \tag{10}$$

where

$$\Delta := 1 + \mu_j + \mu_k - \mu_n, \quad \delta := \frac{1}{2}\left(\sigma_j^2 + \sigma_k^2 - \sigma_n^2\right).$$

Then we consider $\phi_{n,m} := \phi_n^{(m)}(0)$, which satisfies $\phi_{0,0} = 1$, $\phi_{0,m} = 0$ for $m \geq 1$, and by (10) (cf. Bai et al., 2001),

$$\phi_{n,m} = \frac{\Gamma(n)}{\Gamma(n+1/2)} \sum_{1 \leq j < n} \frac{\Gamma(j+1/2)}{\Gamma(j+1)} \phi_{j,m} + \psi_{n,m} \qquad (n \geq 1; m \geq 3),$$

with $\phi_{n,0} = 1$, $\phi_{n,1} = \phi_{n,2} = 0$ (by definition), where

$$\psi_{n,m} = \sum_{\substack{p+q+r+2s=m \\ p,q<m}} \frac{m!}{p!q!r!s!} \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) \phi_{j,p} \phi_{k,q} \Delta^r \delta^s. \tag{11}$$

We need tools for handling recurrences of the type

$$a_n = b_n + \frac{\Gamma(n)}{\Gamma(n+1/2)} \sum_{1 \leq j < n} \frac{\Gamma(j+1/2)}{\Gamma(j+1)} a_j \qquad (n \geq 1), \tag{12}$$

with $a_0 = b_0 := 0$, where $\{b_n\}_{n \geq 1}$ is a given sequence.

Asymptotic transfers for (12).

Proposition 3. (i) *The conditions* $b_n = o(n^{1/2})$ *and* $\sum_j b_j j^{-3/2} < \infty$ *are necessary and sufficient for*

$$a_n \sim c_0 n^{1/2}, \quad \text{where} \quad c_0 := \sum_{j \geq 1} \frac{\Gamma(j+1/2)}{\Gamma(j+2)} b_j; \tag{13}$$

(ii) *if* $|b_n| \leq c_1 n^\alpha$ *for all* $n \geq 1$, *where* $\alpha > 1/2$ *and* $c_1 > 0$, *then*

$$|a_n| \leq B c_1 \frac{2\alpha+1}{2\alpha-1} n^\alpha, \tag{14}$$

for all $n \geq 1$ *and* $\alpha > 1/2$, *where* c *is independent of* α.

Proof. The exact solution to (12) is given by (see Bai et al., 2001)

$$a_n = b_n + \frac{\Gamma(n+1)}{\Gamma(n+1/2)} \sum_{1\le j<n} \frac{\Gamma(j+1/2)}{\Gamma(j+2)}\, b_j \qquad (n \ge 1), \tag{15}$$

from which the sufficiency part of (i) follows (using Stirling's formula). On the other hand, if $a_n \sim cn^{1/2}$ for some c, then by (12)

$$\begin{aligned} b_n &= a_n - \frac{\Gamma(n)}{\Gamma(n+1/2)} \sum_{1\le j<n} \frac{\Gamma(j+1/2)}{\Gamma(j+1)}\, a_j \\ &\sim cn^{1/2} - cn^{-1/2} \sum_{1\le j\le n} 1 \\ &= o(n^{1/2}). \end{aligned}$$

Then by (15), we deduce that $c = c_0$ and that the series $\sum_j b_j j^{-3/2} < \infty$.

For case (ii), we have, by assumption and by (15),

$$\begin{aligned} |a_n| &\le c_1 n^\alpha + c_1 \frac{\Gamma(n+1)}{\Gamma(n+1/2)} \sum_{1\le j<n} \frac{\Gamma(j+1/2)}{\Gamma(j+2)}\, j^\alpha \\ &\sim c_1 n^\alpha + c_1 n^{1/2} \sum_{1\le j<n} j^{\alpha-3/2} \\ &\sim c_1 \frac{2\alpha+1}{2\alpha-1}\, n^\alpha. \end{aligned}$$

Thus (14) holds for, say $1/2 < \alpha \le 1$. For $\alpha \ge 1$, again by (15),

$$\begin{aligned} |a_n| &\le c_1 n^\alpha + Bc_1 n^{1/2} \sum_{1\le j<n} j^{\alpha-3/2} \\ &\le Bc_1 \frac{2\alpha+1}{2\alpha-1}\, n^\alpha, \end{aligned}$$

by a proper choice of B (independent of α). This proves (14). ∎

Estimates. We derive some estimates that will be used later.

Lemma 1. *Let J be a binomial random variable* $\mathrm{Binom}(n-1; U^2)$, *where U has a uniform prior over* $(0,1)$. *Then, for any $r \ge 0$*

$$E\left|\sqrt{J} - \sqrt{n}U\right|^{2r} = \int_0^1 E_x \left|\sqrt{J} - \sqrt{n}x\right|^{2r} \mathrm{d}x \le \frac{(2r)!e^{2r}}{(\log(2r+1))^{2r}}. \tag{16}$$

Proof. We consider two cases. If $x \le 1/\sqrt{n}$, then

$$E_x \left|\sqrt{J} - \sqrt{n}x\right|^{2r} \le 1 + E_x J^r.$$

Note that $\sum_{r\ge 0} E_x(J^r) z^r/r! = (1 + x^2(e^z-1))^{n-1}$. This implies that for any $\beta > 0$

$$E_x(J^r)\beta^r/r! \le (1 + x^2(e^\beta - 1))^{n-1} \le e^{nx^2(e^\beta-1)} \le e^{e^\beta - 1}.$$

Thus

$$E_x \left|\sqrt{J} - \sqrt{n}x\right|^{2r} \leq 1 + r!\beta^{-r} e^{e^{\beta}-1}.$$

Taking $\beta = \log r$, we obtain

$$E_x \left|\sqrt{J} - \sqrt{n}x\right|^{2r} \leq r! e^{r-1} (\log r)^{-r} + 1,$$

for $x \leq 1/\sqrt{n}$ and $r \geq 2$.

Similarly, when $x \geq 1/\sqrt{n}$, we have

$$\begin{aligned}
E_x \left|\sqrt{J} - \sqrt{n}x\right|^{2r} &\leq E_x \Big((J - nx^2)/|\sqrt{J} + \sqrt{n}x| \Big)^{2r} \\
&\leq (\sqrt{n}x)^{-2r} E_x (J - nx^2)^{2r} \\
&\leq (\sqrt{n}x)^{-2r} (2r)! \beta^{-2r} e^{-nx^2\beta} \Big(1 + x^2(e^{\beta} - 1)\Big)^{n-1} \\
&\leq (\sqrt{n}x)^{-2r} (2r)! \beta^{-2r} e^{nx^2(e^{\beta}-1-\beta)}.
\end{aligned}$$

Taking $\beta = (\sqrt{n}x)^{-1} \log(2r+1)$, we obtain

$$\begin{aligned}
E_x \left|\sqrt{J} - \sqrt{n}x\right|^{2r} &\leq (\log(2r+1))^{-2r} (2r)! \exp\left(nx^2 \sum_{j \geq 2} \frac{(\log(2r+1))^j}{j!(\sqrt{n}x)^j} \right) \\
&\leq \frac{(2r)! e^{2r}}{(\log(2r+1))^{2r}} \qquad (r \geq 0).
\end{aligned}$$

The lemma follows from the inequality

$$\frac{r! e^{r-1}}{(\log r)^r} + 1 \leq \frac{(2r)! e^{2r}}{(\log(2r+1))^{2r}},$$

for $r \geq 2$. Note that (16) also holds for $r = 1$. Finally, if $r = 0$ then (16) becomes an identity. ∎

Proposition 4. *If (J, K, L) is trinomial vector $T(n-1; U^2, (1-U)^2, 2U(1-U))$, where U has a uniform prior over $(0,1)$. Then, for any integers $r, p, q \geq 0$,*

$$S_r(n) := E\left|\sqrt{J} + \sqrt{K} - \sqrt{n}\right|^{2r} \leq \frac{(2r)! 4^r e^{2r}}{(\log(2r+1))^{2r}} \leq \frac{r!^2 (4e)^{2r}}{(\log(2r+1))^{2r}}, \tag{17}$$

$$S_{p,q}(n) := E(J^{p/3} K^{q/3}) \leq B \frac{\Gamma(p/3 + 1/2)\Gamma(q/3 + 1)}{\Gamma((p+q)/3 + 3/2)}, \tag{18}$$

where B is independent of p and q.

Proof. Applying Lemma 1 and the inequality

$$E\left|\sqrt{J} + \sqrt{K} - \sqrt{n}\right|^{2r} \leq 2^{2r-1} \left(E\left|\sqrt{J} - \sqrt{n}U\right|^{2r} + E\left|\sqrt{K} - \sqrt{n}(1-U)\right|^{2r} \right),$$

we obtain the first inequality in (17). The second inequality in (17) follows from the inequality

$$\frac{(2r)!}{r!^2} \leq 4^r \qquad (r \geq 0).$$

The estimate (18) is obtained as follows.

$$\begin{aligned}
S_{p,q}(n) &\le E\left(J^{p/3}(n-1-J)^{q/3}\right)\\
&= \sum_{1\le j<n} \binom{n-1}{j} j^{p/3}(n-1-j)^{q/3} \int_0^1 x^{2j}(1-x^2)^{n-1-j}\,\mathrm{d}x\\
&= \sum_{1\le j<n} \frac{\Gamma(n)\Gamma(j+1/2)}{2\Gamma(n+1/2)\Gamma(j+1)}\, j^{p/3}(n-1-j)^{q/3}\\
&\le c\,\frac{\Gamma(n)}{\Gamma(n+1/2)} \sum_{1\le j<n} j^{p/3-1/2}(n-1-j)^{q/3}\\
&\le c\frac{\Gamma(p/3+1/2)\Gamma(q/3+1)}{\Gamma((p+q)/3+3/2)} n^{(p+q)/3},
\end{aligned}$$

where c is independent of p and q. ▮

Estimate for $\phi_{n,3}$. We first determine the order of $\phi_{n,3}$.

Observe that, by (2) and (3),

$$|\Delta|, |\delta| \le B_1(1 + |\sqrt{j} - \sqrt{k} + \sqrt{n}|),$$

for all $0 \le j, k \le n-1$, so that

$$\begin{aligned}
|\Delta|^r|\delta|^s &\le B_1^{r+s}\left(1 + |\sqrt{j} + \sqrt{k} - \sqrt{n}|\right)^{r+s}\\
&\le B_2^{r+s}\left(1 + |\sqrt{j} + \sqrt{k} - \sqrt{n}|^{r+s}|\right) \qquad (r, s \ge 0).
\end{aligned} \tag{19}$$

Since $\phi_{n,1} = \phi_{n,2} = 0$, we have, by (11) with $m = 3$ and (17),

$$\begin{aligned}
\psi_{n,3} &= O\left(\sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n)\left(|\Delta|^3 + |\Delta||\delta|\right)\right)\\
&= O\left(\sum_{0\le p\le 3} E\left|\sqrt{J} + \sqrt{K} - \sqrt{n}\right|^{2p}\right)\\
&= O(1).
\end{aligned}$$

Thus by (13), we obtain

$$\phi_{n,3} \le B_3 n^{1/2}.$$

An upper estimate for $\phi_{n,m}$. We now show by induction that

$$|\phi_{n,m}| \le m!A^m n^{m/6} \qquad (m \ge 0; n \ge 0), \tag{20}$$

where A is a sufficiently large positive constant. The cases $0 \le m \le 3$ hold with $A \ge (B_3/6)^{1/3}$.

Estimate for $\psi_{n,m}$. We consider $m \geq 4$. By definition (11) and induction, we have (using (19))

$$|\psi_{n,m}| \leq m! \sum_{\substack{p+q+r+2s=m \\ p,q<m}} \frac{B_2^{r+s}}{r!s!} A^{p+q} \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) j^{p/6} k^{q/6} \left(1 + |\sqrt{j} + \sqrt{k} - \sqrt{n}|^{r+s}\right),$$

which, by Cauchy-Schwarz inequality, is bounded above by

$$|\psi_{n,m}| \leq m! \sum_{\substack{p+q+r+2s=m \\ p,q<m}} \frac{B_2^{r+s}}{r!s!} A^{p+q} \left(\sqrt{S_{p,q}(n)} + \sqrt{S_{p,q}(n) S_{r+s}(n)}\right), \tag{21}$$

where $S_{p,q}(n)$ and $S_r(n)$ are defined in Proposition 4.

Substituting the estimates (17) and (18) derived above into (21) yields

$$\begin{aligned} |\psi_{n,m}| &\leq Bm! \sum_{\substack{p+q\leq m \\ p,q<m}} A^{p+q} n^{(p+q)/6} \sqrt{\frac{\Gamma(p/3+1/2)\Gamma(q/3+1)}{\Gamma((p+q)/3+3/2)}} \\ &\quad \times \sum_{r+2s=m-p-q} \frac{B_2^{r+s}}{r!s!} \left(1 + \frac{(r+s)!(4e)^{r+s}}{(\log(2r+2s+1))^{r+s}}\right) \\ &\leq Bm! \sum_{0\leq\ell\leq m} A^{\ell} n^{\ell/6} \sum_{\substack{p+q=\ell \\ p,q<m}} \sqrt{\frac{\Gamma(p/3+1/2)\Gamma((\ell-p)/3+1)}{\Gamma(\ell/3+3/2)}}, \end{aligned} \tag{22}$$

since, by Stirling's formula, the sum

$$\sum_{r+2s=k} \frac{(r+s)!(4e)^{r+s}}{r!s!(\log(2r+2s+1))^{r+s}}$$

is bounded for all $k \geq 0$.

The inner sum in (22) is estimated as follows. For $0 \leq \ell \leq m$

$$\begin{aligned} &\sum_{0\leq p\leq\ell} \sqrt{\frac{\Gamma(p/3+1/2)\Gamma((\ell-p)/3+1)}{\Gamma(\ell/3+3/2)}} \\ &\quad \leq c(\ell+1)^{-1/4} + c(\ell+1)^{-5/12} + c(\ell+1)^{-1} \\ &\qquad + \sum_{2\leq p<\ell} \sqrt{\frac{\Gamma(p/3+1/2)\Gamma((\ell-p)/3+1)}{\Gamma(\ell/3+3/2)}} \\ &\quad \leq c(\ell+1)^{-1/4} + \sqrt{\ell+1} \sqrt{\sum_{2\leq p<\ell} \frac{\Gamma(p/3+1/2)\Gamma((\ell-p)/3+1)}{\Gamma(\ell/3+3/2)}}. \end{aligned}$$

Now

$$\sum_{2\le p<\ell} \frac{\Gamma(p/3+1/2)\Gamma((\ell-p)/3+1)}{\Gamma(\ell/3+3/2)}$$
$$= \int_0^1 x^{-1/2}\frac{x^{2/3}(1-x)^{(\ell-1)/3} - x^{\ell/3}(1-x)^{1/3}}{(1-x)^{1/3}-x^{1/3}}\,\mathrm{d}x$$
$$\sim \int_0^{1/2} x^{1/6}e^{-\ell x/3}\,\mathrm{d}x + \int_0^{1/2} x^{1/3}e^{-\ell x/3}\,\mathrm{d}x$$
$$= O((\ell+1)^{-7/6}).$$

Thus

$$\sum_{0\le p\le\ell}\sqrt{\frac{\Gamma(p/3+1/2)\Gamma((\ell-p)/3+1)}{\Gamma(\ell/3+3/2)}} \le c(\ell+1)^{-1/12},$$

and it follows that

$$|\psi_{n,m}| \le Bm!\sum_{0\le\ell\le m} A^\ell n^{\ell/6}(\ell+1)^{-1/12}$$
$$\le B_6 m! A^m n^{m/6} m^{-1/12} \qquad (m\ge 4).$$

Applying now the inequality (14), we obtain

$$|\phi_{n,m}| \le B_6\frac{m+3}{m-3}\, m^{-1/12} m! A^m n^{m/6}. \tag{23}$$

Choose now $m_0 > 3$ so large that $B_6 m^{-1/12}(m+3)/(m-3) \le 1$ for all $m > m_0$. Then (20) holds for $m > m_0$. It remains to tune the value of A such that $|\phi_{n,m}| \le m! A^m n^{m/6}$ for all $m \le m_0$. This is possible since the factor A in (23) depends only on $|\phi_{n,j}|$, $0\le j<m$. Thus if we take

$$A_r = \left(\frac{B_3}{6}\right)^{1/3}\prod_{4\le j\le r}\left(B_6\,\frac{j+3}{j^{1/12}(j-3)}\right)^{1/j} \qquad (3\le r\le m_0),$$

then

$$|\phi_{n,r}| \le B_6\frac{r+3}{r-3}\, r^{-1/12} r! A_{r-1}^r n^{r/6} = r! A_r^r n^{r/6} \qquad (4\le r\le m_0).$$

The proof of (20) is complete by taking $A = A_{m_0}$.

Proof of Proposition 1. With the precise estimate (20) available, we now have

$$\left|\varphi_n(y) - e^{-y^2/2}\right| = e^{-y^2/2}\left|\phi_n(iy/\sigma_n) - 1\right|$$
$$\le e^{-y^2/2}\sum_{m\ge 3}\frac{|\phi_{n,m}|}{m!\sigma_n^m}|y|^m$$
$$\le e^{-y^2/2}\sum_{m\ge 3}\left(A\sigma_n^{-1}n^{1/6}|y|\right)^m$$
$$\le ce^{-y^2/2}|y|^3 n^{-1/4},$$

if $|y| \le \varepsilon n^{1/12}$. This proves (6).

We also need an estimate for $|\varphi_n(y)|$ for larger values of $|y|$.

Note that from (6) we have

$$\begin{aligned}|\varphi_n(y)| &\le e^{-y^2/2}\left(1 + c|y|^3 n^{-1/4}\right)\\ &\le e^{-y^2/2 + c|y|^3/n^{1/4}},\end{aligned}$$

for $|y| \le \varepsilon n^{1/12}$. Also, by definition, $|P_n(iy)| = 1$ for $n = 0, 1$ and $P_2(iy) = e^{iy}/3 + 2e^{2iy}/3$. Thus, by (3),

$$\begin{aligned}|P_n(iy)| &\le e^{-\sigma_n^2 y^2/2 + c_2\sqrt{n}|y|^3}\\ &\le e^{-(\tau_1\sqrt{n}+\tau_2)y^2} \qquad (n \ge 2),\end{aligned} \tag{24}$$

for some $c_2 > 0$ and $|y| \le \varepsilon_1 n^{-1/6}$, where $\tau_1, \tau_2 > 0$ are constants satisfying

$$\tau_1\sqrt{n} + \tau_2 \le \frac{\sigma_n^2}{2} - c_2\varepsilon_1 n^{1/3} \qquad (n \ge 2). \tag{25}$$

Here ε_1 is chosen so small that the right-hand side is positive for all $n \ge 2$ (we may take $\varepsilon_1 = 1/(12c_2)$).

Proof of Proposition 2. We now show, again by induction, that the same estimate (24) holds for $|y| \le \pi$, provided the constants τ_1 and τ_2 are suitably tuned.

Note that since the span of M_n is 1 (by induction, $P(M_n = k) > 0$ for $1 \le k \le n$),

$$|P_n(iy)| \le e^{-c_3 y^2} \qquad (|y| \le \pi),$$

for $2 \le n \le n_0$, where n_0 is a sufficiently large number (see (27)). [Numerically, $c_3 = 1/9$ suffices.] This gives another condition for τ_1 and τ_2:

$$\tau_1\sqrt{n} + \tau_2 \le \tau_1\sqrt{n_0} + \tau_2 \le c_3 \qquad (2 \le n \le n_0). \tag{26}$$

By induction using (8) and (24),

$$\begin{aligned}|P_n(iy)| &\le \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n)|P_j(iy)||P_k(iy)|\\ &\le \sum_{j+k+\ell=n-1} \pi_{j,k,\ell}(n) e^{-(\tau_1\sqrt{j}+\tau_1\sqrt{k}+2\tau_2)y^2}\\ &\quad + 2\sum_{0\le j\le 1}\sum_{0\le k<n} \pi_{j,k,n-1-k-j}(n) e^{-(\tau_1\sqrt{k}+\tau_2)y^2}\left(1 - e^{-(\tau_1\sqrt{j}+\tau_2)y^2}\right)\\ &\quad + \sum_{0\le j,k\le 1} \pi_{j,k,n-1-k-j}(n)\left(1 - e^{-(\tau_1\sqrt{j}+\tau_1\sqrt{k}+2\tau_2)y^2}\right)\\ &=: G + 2\sum_{0\le j\le 1} G_j + \sum_{0\le j,k\le 1} G_{jk},\end{aligned}$$

say. By (17), we obtain (with the notation of Lemma 4)

$$
\begin{aligned}
e^{(\tau_1\sqrt{n}+\tau_2)y^2}G &= e^{-\tau_2 y^2} E\left(e^{-\tau_1(\sqrt{J}+\sqrt{K}-\sqrt{n})y^2}\right) \\
&= e^{-\tau_2 y^2} \sum_{m\ge 0} \frac{(-\tau_1 y^2)^m}{m!} E\left(\sqrt{J}+\sqrt{K}-\sqrt{n}\right)^m \\
&\le e^{-\tau_2 y^2} \sum_{m\ge 0} \frac{(\tau_1 y^2)^m}{m!} \left(E(\sqrt{J}+\sqrt{K}-\sqrt{n})^{2m}\right)^{1/2} \\
&\le e^{-\tau_2 y^2} \left(1 + \sum_{m\ge 1} \frac{(\tau_1 y^2)^m}{m!} \cdot \frac{m!(4e)^m}{(\log(2m+1))^m}\right) \\
&\le e^{-\tau_2 y^2}\left(1 + c_4\tau_1 y^2\right) \\
&\le e^{-\tau_2 y^2 + c_4\tau_1 y^2}.
\end{aligned}
$$

Choose $\tau_2 = 2c_4\tau_1$. Then

$$
G \le e^{-(\tau_1\sqrt{n}+\tau_2)y^2 - c_4\tau_1 y^2}.
$$

We now estimate G_0.

$$
\begin{aligned}
\frac{e^{(\tau_1\sqrt{n}+\tau_2)y^2}}{1-e^{-\tau_2 y^2}} G_0 &= \sum_{0\le k<n} \pi_{0,k,n-1-k}(n) e^{\tau_1(\sqrt{n}-\sqrt{k})y^2} \\
&= \sum_{0\le k<n} \binom{n-1}{k} 2^{n-1-k} e^{\tau_1(\sqrt{n}-\sqrt{k})y^2} \int_0^1 x^{n-1-k}(1-x)^{n-1+k}\,\mathrm{d}x \\
&\le \sum_{0\le k<n} \binom{n-1}{k} 2^{n-1-k} e^{\tau_1(n-k)y^2/\sqrt{n}} \int_0^1 x^{n-1-k}(1-x)^{n-1+k}\,\mathrm{d}x \\
&= \sum_{0\le k<n} \binom{n-1}{k} 2^{k} e^{\tau_1(k+1)y^2/\sqrt{n}} \int_0^1 x^{k}(1-x)^{2n-2+k}\,\mathrm{d}x \\
&= e^{\tau_1 y^2/\sqrt{n}} \int_0^1 \left(1 + 2\left(e^{\tau_1 y^2/\sqrt{n}}-1\right)x - \left(2e^{\tau_1 y^2/\sqrt{n}}-1\right)x^2\right)^{n-1} \mathrm{d}x \\
&\le \frac{c}{n}\int_0^n \exp\left(\frac{2\tau_1 y^2}{\sqrt{n}}x - \frac{x^2}{n} + O\left(\frac{x}{n}+\frac{x^2}{n^{3/2}}\right)\right)\mathrm{d}x \\
&\le \frac{c}{n}\int_0^{4\tau_1 y^2\sqrt{n}} 1\,\mathrm{d}x + \frac{c}{n}\int_{4\tau_1 y^2\sqrt{n}}^{\infty} e^{-x^2/(4n)}\,\mathrm{d}x \\
&\le cn^{-1/2}.
\end{aligned}
$$

It follows that

$$
e^{(\tau_1\sqrt{n}+\tau_2)y^2} G_0 \le c\tau_2 y^2 n^{-1/2} \le c_5\tau_1 y^2 n^{-1/2},
$$

where c_5 is independent of τ_1.

The partial sum G_1 is estimated similarly, giving

$$\begin{aligned}\frac{e^{(\tau_1\sqrt{n}+\tau_2)y^2}}{1-e^{-(\tau_1+\tau_2)y^2}}G_1 &\le \sum_{0\le k\le n-2}\pi_{1,k,n-2-k}(n)e^{\tau_1(\sqrt{n}-\sqrt{k})y^2}\\ &= \sum_{0\le k\le n-2}\binom{n-2}{k}2^{n-2-k}e^{\tau_1(\sqrt{n}-\sqrt{k})y^2}\int_0^1 x^{n-2-k}(1-x)^{n-2+k}\,\mathrm{d}x\\ &\le e^{\tau_1 y^2/\sqrt{n}}\int_0^1(1-x)^{n-2}\left(1+2e^{\tau_1 y^2/\sqrt{n}}x-x\right)^{n-2}\mathrm{d}x\\ &\le cn^{-1/2},\end{aligned}$$

and, consequently,

$$e^{(\tau_1\sqrt{n}+\tau_2)y^2}G_1 \le c_6\tau_1 y^2 n^{-1/2},$$

where c_6 is independent of τ_1.

The remaining terms G_{jk} are easy.

$$\begin{aligned}\pi_{0,0,n-1}(n) &= \frac{2^{n-1}(n-1)!(n-1)!}{(2n-1)!}\sim\frac{\sqrt{\pi}}{2^n\sqrt{n}},\\ \pi_{0,1,n-2}(n) = \pi_{1,0,n-2}(n) &= \frac{2^{n-2}(n-1)!n!}{(2n-1)!}\sim\frac{\sqrt{\pi n}}{2^{n+1}},\\ \pi_{1,1,n-3} &= \frac{2^{n-3}(n-1)(n-2)(n-1)!(n-1)!}{(2n-1)!}\sim\frac{\sqrt{\pi}\,n^{3/2}}{2^{n+2}}.\end{aligned}$$

Thus there exists a constant c_7 such that for $n\ge 2$

$$\pi_{j,k,n-1-j-k}(n)\le e^{-c_7 n}\qquad(0\le j,k\le 1).$$

[Numerically, $c_7 = 1/3$ suffices.] Therefore,

$$\begin{aligned}\sum_{0\le j,k\le 1}G_{jk} &\le 4e^{-c_7 n}\left(1-e^{-2(\tau_1+\tau_2)y^2}\right)\\ &\le c_8\tau_1 y^2 e^{-c_7 n},\end{aligned}$$

for $n\ge 2$, where c_8 is independent of τ_1.

Collecting these estimates, we have

$$|P_n(iy)|\le \Pi_n(y)e^{-(\tau_1\sqrt{n}+\tau_2)y^2},$$

where

$$\Pi_n(y) := e^{-c_4\tau_1 y^2}+c_9\tau_1 y^2 n^{-1/2}+c_8\tau_1 y^2 e^{-c_7 n+(\tau_1\sqrt{n}+\tau_2)y^2},$$

where $c_9 = c_4 + c_5$. Now choose n_0 so large that

$$\max_{\varepsilon_1 n^{-1/2}\le|y|\le\pi}\Pi_n(y)\le 1\qquad(\varepsilon_1 n^{-1/6}\le|y|\le\pi),\tag{27}$$

for $n \geq n_0$. This is possible since $|y| \geq \varepsilon_1 n^{-1/6}$. Also n_0 is independent of τ_1 (since c_8 and c_9 are independent of τ_1). Once n_0 is specified, the two conditions (25) and (26) (together with $\tau_2 = 2c_4\tau_1$) become

$$\tau_1 \leq \begin{cases} \dfrac{\sigma_n^2/2 - c_2\varepsilon_1 n^{1/3}}{\sqrt{n} + 2c_4}, \\ \dfrac{c_3}{\sqrt{n_0} + 2c_4}. \end{cases}$$

Thus we can take

$$\tau_1 = \min\left\{\frac{c_3}{\sqrt{n_0} + 2c_4}, \min_{n\geq 2}\left\{\frac{\sigma_n^2/2 - c_2\varepsilon_1 n^{1/3}}{\sqrt{n} + 2c_4}\right\}\right\}.$$

We thus proved that

$$|P_n(iy)| \leq e^{-\tau_1(\sqrt{n}+2c_4)y^2} \qquad (|y| \leq \pi; n \geq 2),$$

which implies (7) by a proper choice of ε.

Berry-Esseen smoothing inequality. We now apply the Berry-Esseen smoothing inequality (see Petrov, 1975), which states for our problem that

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{M_n - \mu_n}{\sigma_n} < x\right) - \Phi(x)\right| = O\left(T^{-1} + \int_{-T}^{T}\left|\frac{\varphi_n(y) - e^{-y^2/2}}{y}\right| \mathrm{d}y\right),$$

where T is taken to be $\varepsilon n^{1/4}$. By the two estimates (6) and (7), we obtain

$$\begin{aligned}\int_{-T}^{T}\left|\frac{\varphi_n(y) - e^{-y^2/2}}{y}\right| \mathrm{d}y &= O\left(n^{-1/4}\int_{-\varepsilon n^{1/12}}^{\varepsilon n^{1/12}} e^{-y^2/2}y^2\,\mathrm{d}y + \int_{\varepsilon n^{1/12}}^{n^{1/4}} \frac{e^{-\varepsilon y^2} + e^{-y^2/2}}{y}\,\mathrm{d}y\right) \\ &= O\left(n^{-1/4}\right) + O\left(n^{-1/12}e^{-\varepsilon n^{1/6}}\right).\end{aligned}$$

Accordingly,

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{M_n - \mu_n}{\sigma_n} < x\right) - \Phi(x)\right| = O(n^{-1/4}). \quad \blacksquare$$

Local limit theorem. By the inversion formula

$$\begin{aligned}P(M_n = k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iky}P_n(iy)\,\mathrm{d}y \\ &= \frac{1}{2\pi\sigma_n}\int_{-\pi\sigma_n}^{\pi\sigma_n} e^{-ixy}\varphi_n(y)\left(1 + O\left((1+|y|)n^{-1/4}\right)\right)\,\mathrm{d}y,\end{aligned}$$

where $k = \lfloor \mu_n + x\sigma_n \rfloor$, we deduce, by splitting the integral similarly as above, the local limit theorem (5).

3 Planar regions upper-bounded by decreasing curves

Notations. Recall that $f(u) > 0$ is a decreasing function on $(0,1)$ with $f(0^+) < \infty$, $f(1^-) = 0$ and $\int_0^1 f(u)\,\mathrm{d}u = 1$. Define $\mathcal{D} = \{(u,v) : 0 < u < 1, 0 < v < f(u)\}$. Then $|\mathcal{D}| = 1$.

Let C^2 denote the class of twice differentiable functions f in the unit interval with $f'(u) < 0$ for $0 < u < 1$.

A measure of "steepness". Assume $f \in C^2$. Construct $\lambda - 1$ points u_j's in the interval $0 = u_0 < u_1 < \cdots < u_{\lambda-1} < u_\lambda = 1$ such that

$$\nabla u_k \nabla f_k = \frac{4 \log n}{n},$$

for $1 \le k < \lambda$ and $\nabla u_\lambda \nabla f(u_\lambda) < 4 \log n/n$, where $\nabla u_k := u_k - u_{k-1}$ and $\nabla f_k := f(u_{k-1}) - f(u_k)$. We assume implicitly that n is sufficiently large so that the u_k's are well-defined.

Define a measure of "steepness" or "flatness":

$$\gamma_n = \max \left\{ 1, \sum_{2 \le k \le \lambda - 3} \sup_{\substack{u_{k-1} \le \zeta_k \le u_{k+1} \\ u_{k-1} \le \xi_k \le u_k}} \left| \frac{f''(\zeta_k)}{f'(\xi_k)} \right| (u_k - u_{k-1}) \right\}. \tag{28}$$

Note that the ranges $u_{k-1} \le \zeta_k \le u_k$ alone are not sufficient for our proof. In particular, in (36) and (38) we use the ranges $u_{k-1} \le \zeta_k \le u_k$, but $u_{k-1} \le \zeta_k \le u_{k+1}$ are needed in (44). Also the index range $2 \le k \le \lambda - 3$ is chosen for technical convenience.

Theorem 3. *Assume $f \in C^2$. Then*

$$\sup_{-\infty < x < \infty} \left| P\left(\frac{M_n(\mathcal{D}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x \right) - \Phi(x) \right| = O\left(n^{-\frac{1}{4}} \gamma_n \log^2 n \right), \tag{29}$$

where $\sigma^2 = 2\log 2 - 2$ and

$$\alpha = \int_0^1 \sqrt{|f'(u)|/2}\, \mathrm{d}u.$$

Although γ_n may diverge with n even under the stronger assumption that $f \in C^\infty$, it is small compared to $n^{1/4}$ in most cases.

Corollary 1. *Assume $f \in C^2$. If both $f'(0^+)$ and $f'(1^-)$ exist and*

$$-\infty < f'(0^+), f'(1^-) < 0,$$

then (29) holds with

$$\gamma_n = O(1).$$

For example, if $f(u) = f_0(u)/\int_0^1 f_0(t)\,\mathrm{d}t$, where $f_0(u) := 2 - u - u^b$, $b > 1$, then $\gamma_n = O(1)$. This example indicates that the error term in (29) may not in general be optimal in view of (4).

Corollary 2. *Assume $f \in C^2$. If $f''(u) \ne 0$ and there exists a sequence $\kappa_n > 0$ such that*

$$\frac{1}{\kappa_n} \le \left| \frac{f''(\zeta_k)}{f''(\xi_k)} \right| \le \kappa_n,$$

for all $\zeta_k, \xi_k \in (u_{k-1}, u_{k+1})$, $k = 2, 3, ..., \lambda - 3$, then (29) holds with

$$\gamma_n = O\left(\kappa_n \int_{u_1}^{u_{\lambda-3}} \left| \frac{f''(u)}{f'(u)} \right| \mathrm{d}u \right).$$

For example, if $f(u) = f_1(u)/\int_0^1 f_1(t)\,\mathrm{d}t$, where $f_1(u) := (1-u^{b_1})^{b_2}$, $b_1, b_2 > 0$, then there is a $c > 0$ such that

$$\frac{1}{c} \le \left|\frac{f''(\zeta_k)}{f''(\xi_k)}\right| \le c,$$

for all $\zeta_k, \xi_k \in (u_{k-1}, u_{k+1})$, $k = 2, 3, ..., \lambda - 3$, and

$$\int_{u_1}^{u_{\lambda-3}} \left|\frac{f''(u)}{f'(u)}\right| \mathrm{d}u = O\left(\int_{1/n}^{1-1/n} u^{-1} + (1-u)^{-1}\,\mathrm{d}u\right) = O\left(\log n\right).$$

Thus

$$\gamma_n = O\left(\log n\right).$$

If f behaves very flatly or steeply near the origin or the unit, say

$$f(u) = \frac{1 - e^{1-1/u}}{\int_0^1 (1 - e^{1-1/u})\,\mathrm{d}u},$$

then tedious calculations yield $\gamma_n = O(\log^2 n)$.

Main steps of the proof of Theorem 3. Let $\nu = \sqrt{\frac{n}{4\log n}}$. Choose $0 = u_0 < u_1 < \cdots < u_{\lambda-1} < u_\lambda = 1$ as above. Let $\mathcal{T}$ be the right triangle region formed by $(0,0)$, $(\lambda/\nu, 0)$ and $(0, \lambda/\nu)$. Let $\mathcal{X}$ be a Poison Process on the plane with density n. Denote by $\mathcal{M}_n(\mathcal{A})$ the number of maxima of $\mathcal{X} \cap \mathcal{A}$.

(i) We first show that

$$|\mathcal{T}| = \alpha^2 + O\left(\gamma_n \sqrt{\log n/n}\right). \tag{30}$$

(ii) From (30), we next deduce that the number of maxima of the Poisson process satisfies

$$\sup_{-\infty<x<\infty} \left| P\left(\frac{\mathcal{M}_n(\mathcal{T}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\right) - \Phi(x)\right| = O\left(n^{-\frac{1}{4}}\gamma_n \log n\right). \tag{31}$$

The introduction of the Poisson process has the advantage of simplifying the analysis.

(iii) The next step is to show (quantitatively) that most maxima appear near the boundary:

$$P\left(|\mathcal{M}_n(\mathcal{T}) - \mathcal{M}_n(\mathcal{Z})| > c\gamma_n \log^2 n\right) = O\left(n^{-1/2}\right), \tag{32}$$

where $\mathcal{Z}$ is region close to the boundary defined below (see (40)).

(iv) We then introduce a mapping on $\mathcal{D}$ that basically transforms $\mathcal{D}$ into $\mathcal{T}$ and show that the numbers of maxima under the Poisson process model and the iid uniform distribution model are very close; more precisely,

$$\sup_x \left| P\left(\frac{\mathcal{M}_n(\mathcal{Z}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\right) - P\left(\frac{M_h(\mathcal{Z}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\right)\right| = O\left(\sqrt{\log n/n}\right), \tag{33}$$

where $M_h(\mathcal{A})$ is the number maximal points of $\{h(X_1), ..., h(X_n)\}$ that also lie in $\mathcal{A}$. Here h is a mapping on $\mathcal{D}$; see (39).

(v) The final step is to construct mappings h_1 and h_2 such that $M_{h_1}(\mathcal{Z}) \stackrel{d}{=} M_{h_2}(\mathcal{Z}) \stackrel{d}{=} M_h(\mathcal{Z})$, and

$$P\left(M_{h_1}(\mathcal{Z}) - M_n(\mathcal{D}) > c\gamma_n \log n\right) = O\left(n^{-1}\right), \tag{34}$$

$$P\left(M_n(\mathcal{D}) - M_{h_2}(\mathcal{Z}) > c\gamma_n \log n\right) = O\left(n^{-1}\right). \tag{35}$$

The reason for introducing these mappings is that the dominance relations for some parts of $\mathcal{D}$ may be changed by the mapping h. So we need to "fine-tune" the number of maxima.

We then conclude our Berry-Esseen bound (29) for $M_n(\mathcal{D})$ from **(ii)**, **(iii)**, **(iv)** and **(v)** since $\Phi'(x)$ is bounded.

Proof of (30). It suffices to prove that the total number of sections of the unit interval satisfies

$$\lambda = \sqrt{2}\nu\alpha + O(\gamma_n).$$

Note that

$$\nu \int_{u_{k-1}}^{u_k} \sqrt{\frac{\nabla f_k}{\nabla u_k}}\, \mathrm{d}u = \int_{u_{k-1}}^{u_k} \frac{1}{\nabla u_k}\, \mathrm{d}u = 1 \qquad (1 \le k < \lambda),$$

and by Cauchy-Schwarz inequality

$$\int_0^{u_1} \sqrt{|f'(u)|}\, \mathrm{d}u \le \sqrt{\nabla f_1}\sqrt{\nabla u_1} = O\left(1/\nu\right).$$

It follows that

$$\begin{aligned}
\left|\lambda - \sqrt{2}\nu\alpha\right| &= \left|\sum_{2\le k\le \lambda-3} \nu \int_{u_{k-1}}^{u_k} \sqrt{\frac{\nabla f_k}{\nabla u_k}}\, \mathrm{d}u - \nu \sum_{2\le k\le \lambda-3} \int_{u_{k-1}}^{u_k} \sqrt{|f'(u)|}\, \mathrm{d}u\right| + O(1) \qquad (36)\\
&\le \nu \sum_{2\le k\le \lambda-3} \int_{u_{k-1}}^{u_k} \left|\sqrt{|f'(\xi_k)|} - \sqrt{|f'(u)|}\right| \mathrm{d}u + O(1)\\
&\qquad \left(f'(\xi_k) = -\frac{\nabla f_k}{\nabla u_k}, \xi_k \in [u_{k-1}, u_k]\right)\\
&\le \nu \sum_{2\le k\le \lambda-3} \int_{u_{k-1}}^{u_k} \frac{|f'(\xi_k) - f'(u)|}{\sqrt{|f'(\xi_k)|} + \sqrt{|f'(u)|}}\, \mathrm{d}u + O(1)\\
&\le \nu \sum_{2\le k\le \lambda-3} \int_{u_{k-1}}^{u_k} |\xi_k - u| \frac{|f''(\zeta_k)|}{\sqrt{|f'(\xi_k)|} + \sqrt{|f'(u)|}}\, \mathrm{d}u + O(1) \quad \left(\zeta_k \in [u_{k-1}, u_k]\right)\\
&\le \sum_{2\le k\le \lambda-3} \int_{u_{k-1}}^{u_k} \frac{\nabla u_k}{\sqrt{\nabla u_k \nabla f_k}} \cdot \frac{|f''(\zeta_k)|}{\sqrt{|f'(\xi_k)|} + \sqrt{|f'(u)|}}\, \mathrm{d}u + O(1) \qquad (37)\\
&\le \sum_{2\le k\le \lambda-3} \int_{u_{k-1}}^{u_k} \frac{|f''(\zeta_k)|}{|f'(\xi_k)| + \sqrt{|f'(\xi_k) f'(u)|}}\, \mathrm{d}u + O(1)\\
&= O\left(\gamma_n\right).
\end{aligned}$$

Proof of (31). From (30), there is $c > 0$ such that

$$P\left(\left|\mathcal{N}_n(\mathcal{T}) - \alpha^2 n\right| > c\gamma_n \sqrt{n}\log n\right) = O(n^{-1}),$$

where $\mathcal{N}_n(\mathcal{A})$ is the number of points in $\mathcal{X} \cap \mathcal{A}$. Applying Theorem 1 (conditioning on fixed number of points in $\mathcal{T}$), we have

$$\begin{aligned}
&\sup_x \left| P\left(\frac{\mathcal{M}_n(\mathcal{T}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\right) - \Phi(x)\right| \\
&= \sum_{|k-\alpha^2 n| \le c\gamma_n\sqrt{n}\log n} \sup_x \left| P\left(\frac{M_k - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x \,|\mathcal{N}_n(\mathcal{T}) = k\right) - \Phi(x)\right| P\left(\mathcal{N}_n(\mathcal{T}) = k\right) \\
&\quad + O(n^{-1}) \\
&= \sum_{|k-\alpha^2 n| \le c\gamma_n\sqrt{n}\log n} O\left(\frac{\left|\sqrt{\pi k} - \alpha\sqrt{\pi n}\right|}{n^{1/4}}\right) P\left(\mathcal{N}_n(\mathcal{T}) = k\right) + O(n^{-1}) \\
&= O\left(n^{-\frac{1}{4}}\gamma_n \log n\right).
\end{aligned}$$

To prove (32)–(35), we define first a dissection and then a transformation on $\mathcal{D}$.

Dissection of $\mathcal{D}$. We split the region $\mathcal{D}$ into several smaller regions as follows. Let

$$\begin{aligned}
Q_k &= \left\{(u,v) \;:\; u_{k-1} \le u < u_k,\; f(u_k) \le v < f(u)\right\} && (1 \le k \le \lambda), \\
R_k &= \left\{(u,v) \;:\; u_{k-1} \le u < u_k,\; f(u_{k+1}) \le v < f(u_k)\right\} && (1 \le k \le \lambda - 1), \\
W_k &= \left\{(u,v) \;:\; u_{k-1} \le u < u_k,\; 0 < v < f(u_{k+1})\right\} && (1 \le k \le \lambda - 2).
\end{aligned}$$

Let T_k be the triangle formed by the vertices $(u_{k-1}, f(u_{k-1}))$, $(u_k, f(u_k))$ and $(u_{k-1}, f(u_k))$ and $F_k = (Q_k - T_k) \cup (T_k - Q_k)$. Denote by

$$T := \bigcup_{1 \le k \le \lambda} T_k, \quad Q := \bigcup_{1 \le k \le \lambda} Q_k, \quad F := \bigcup_{1 \le k \le \lambda} F_k, \quad R := \bigcup_{1 \le k \le \lambda-1} R_k, \quad W := \bigcup_{1 \le k \le \lambda-2} W_k.$$

Then $\mathcal{D}$ is the disjoint union of Q, R, and W, and F is the difference of $\mathcal{D}$ and the polygon $T \cup R \cup W$. The area of F satisfies

$$\begin{aligned}
|F| &\le \sum_{2 \le k \le \lambda-3} \frac{|f''(\zeta_k)|\,(\nabla u_k)^3}{12} + O\left(\frac{\log n}{n}\right) \\
&= \sum_{2 \le k \le \lambda-3} \frac{|f''(\zeta_k)|}{12\nu^2\,|f'(\xi_k)|} \nabla u_k + O\left(\frac{\log n}{n}\right) \\
&= O\left(\frac{\gamma_n \log n}{n}\right),
\end{aligned} \tag{38}$$

where $f'(\xi_k) = -\nabla f_k / \nabla u_k$ and $|f''(\zeta_k)| = \sup_{u_{k-1} \le u \le u_k} |f''(u)|$.

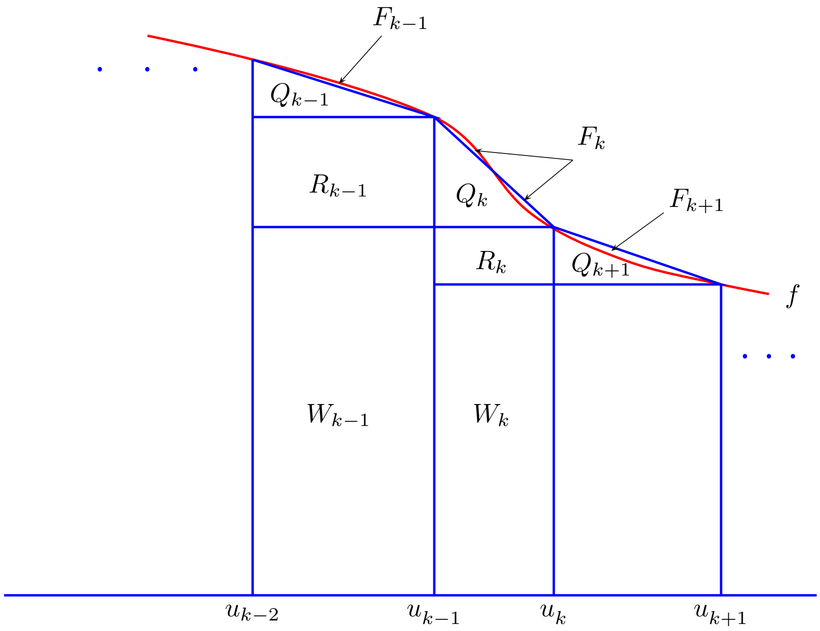

Figure 3: *The dissection of $\mathcal{D}$.*

A transformation. We define a transformation h on $0 \le u < 1$ as follows. For $u_{k-1} \le u < u_k$, $1 \le k \le \lambda$,

$$h(u,v) = \left(\frac{k-1}{\nu} + \left(u - u_{k-1}\right) \sqrt{\frac{\nabla f_k}{\nabla u_k}}, \ \frac{\lambda - k + 1}{\nu} - \left(f(u_{k-1}) - v\right) \sqrt{\frac{\nabla u_k}{\nabla f_k}} \right). \tag{39}$$

Note that $h(T_k)$ is the right triangle region formed by $(\frac{k-1}{\nu}, \frac{\lambda-k}{\nu})$, $(\frac{k}{\nu}, \frac{\lambda-k}{\nu})$ and $(\frac{k-1}{\nu}, \frac{\lambda-k+1}{\nu})$, and

$$h(R_k) = \left\{ (u,v) : \frac{k-1}{\nu} \le u < \frac{k}{\nu}, \ \frac{\lambda - k - \frac{\nabla u_k}{\nabla u_{k+1}}}{\nu} \le v < \frac{\lambda - k}{\nu} \right\}.$$

Basically the main effect of h is to transform $\mathcal{D}$ into $\mathcal{T}$.

By construction, the mapping h is piecewise linear and preserves measure. Thus $h(X_1), \ldots, h(X_n)$ are n iid random variables uniformly distributed in $h(\mathcal{D})$.

Proof of (32). Define

$$\mathcal{Z} := h(Q \cup R). \tag{40}$$

The proof for (32) consists of two parts: we first show that $\mathcal{M}_n(\mathcal{T} - \mathcal{Z} - h(F))$ is negligible; and then we estimate the difference between the number of maxima in $\mathcal{Z}$ and that in $h(T \cup R)$.

To show that $\mathcal{M}_n(\mathcal{T} - \mathcal{Z} - h(F))$ is negligible, we start from the following definitions. For $p = (p_1, p_2)$, let

$$\triangle(p, \mathcal{A}) := \{(u,v) : u > p_1, v > p_2, (u,v) \in \mathcal{A}\},$$

represents the first quadrant of p that also lies in $\mathcal{A}$, and define

$$\mathcal{W} := \{p : |\triangle(p, \mathcal{T} - \mathcal{Z} - h(F))| \le 2 \log n / n\}.$$

Then

$$P\left(\mathcal{M}_n(\mathcal{T}-\mathcal{Z}-h(F)-\mathcal{W})\geq 1\right)\leq 2\lambda\exp\left(-n\cdot\frac{2\log n}{n}\right)=O(n^{-1}).$$

We will show that

$$|\mathcal{W}|=O(\gamma_n\log n/n). \tag{41}$$

Then the number of maxima in $\mathcal{W}$ satisfies

$$P\left(\mathcal{M}_n(\mathcal{W})>c\gamma_n\log n\right)\leq P\left(\mathcal{N}_n(\mathcal{W})>c\gamma_n\log n\right)=O(n^{-1}).$$

And the contribution of $\mathcal{M}_n(\mathcal{T}-\mathcal{Z}-h(F))$ is absorbed in the error term $\gamma_n(\log n)^2/n^{1/4}$.

To show (41), we need to estimate $|h(W_k)\cap\mathcal{W}|$. Define (see Figure 4)

$$l_k=\max\left\{0,\frac{1}{\nu}\left(1-\frac{\nabla u_k}{\nabla u_{k+1}}\right)\right\}, \tag{42}$$

$$L_k=\left\{(u,v):\frac{k-1}{\nu}\leq u<\frac{k}{\nu},\ \frac{\lambda-k-1}{\nu}\leq v<\frac{\lambda-k-1}{\nu}+l_k\right\}, \tag{43}$$

and

$$\hat{W}_k=\left\{(u,v):\frac{k}{\nu}-\nu\left|F_{k+1}\right|\leq u<\frac{k}{\nu},\ \frac{\lambda-k-1}{\nu}-\nu\left|F_{k+1}\right|\leq v<\frac{\lambda-k-1}{\nu}\right\}.$$

Here l_k is the difference between the height of $h(T_{k+1})$ and that of $h(R_k)$. Then $h(W_k)\cap\mathcal{W}\subset L_k\cup\hat{W}_k$.

Also the area of $\hat{W}_k$ satisfies

$$\left|\hat{W}_k\right|=\nu^2\left|F_{k+1}\right|^2=\frac{\left|F_{k+1}\right|^2}{2\left|T_{k+1}\right|}\leq\frac{\left|F_{k+1}\right|}{2}.$$

We prove that $|L|=O(\gamma_n\log n/n)$, where $L:=\cup L_k$. Note that if $\nabla u_k\geq\nabla u_{k+1}$ then $|L_k|=0$, and if $\nabla u_k<\nabla u_{k+1}$ then

$$\begin{aligned}|L_k|&=\frac{1}{\nu^2}\left(1-\frac{\nabla u_k}{\nabla u_{k+1}}\right)\\&=\left(\nabla f_k-\nabla f_{k+1}\right)\nabla u_k\\&\leq\left(\frac{\nabla f_k}{\nabla u_k}-\frac{\nabla f_{k+1}}{\nabla u_{k+1}}\right)\left(\nabla u_k\right)^2\\&\leq\left|f''(\zeta_k)\right|\left(\nabla u_k\right)^3,\end{aligned} \tag{44}$$

where $u_{k-1}\leq\zeta_k\leq u_{k+1}$. Summing over k yields

$$|L|\leq\sum_{2\leq k\leq\lambda-3}\frac{|f''(\zeta_k)|}{\nu^2\left|f'(\xi_k)\right|}\nabla u_k+O\left(\frac{\log n}{n}\right)=O\left(\frac{\gamma_n\log n}{n}\right).$$

This completes the proof of (41).

We now estimate the difference between the number of maxima in the region $\mathcal{Z}$ and that in $h(T\cup R)$. It is obvious that

$$\mathcal{M}_n(h(Q_k))=\mathcal{M}_n(h(T_k))$$

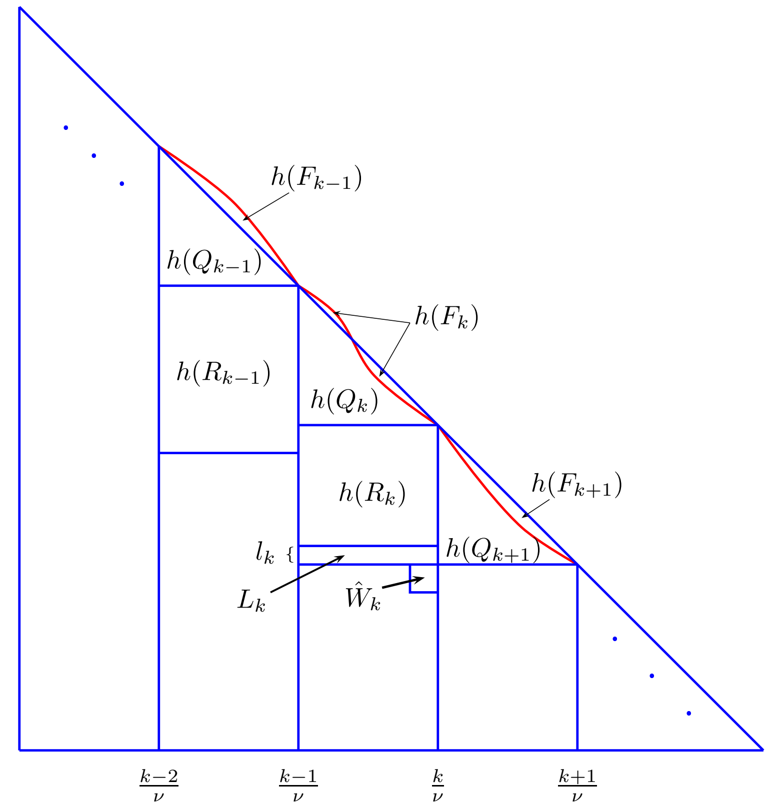

Figure 4: *A possible configuration of* $h(\mathcal{D})$.

when $\mathcal{N}_n(F_k) = 0$. Summing over k, we have

$$|\mathcal{M}_n(h(Q)) - \mathcal{M}_n(h(T))| \le \sum_{k\,:\,\mathcal{N}_n(F_k)>0} \mathcal{N}_n(T_k \cup Q_k).$$

Note that $|T_k \cup Q_k| = O(\log n/n)$ and

$$\begin{aligned} & P\left(\sum_{k\,:\,\mathcal{N}_n(F_k)>0} \mathcal{N}_n(T_k \cup Q_k) \ge c\gamma_n \log^2 n\right) \\ & \quad \le P\left(\mathcal{N}_n(F) \ge c\gamma_n \log n\right) + P\left(\mathcal{N}_n(T_k \cup Q_k)\right) \ge c \log n \text{ for some } k) \\ & \quad = O(n^{-1/2}). \end{aligned}$$

It follows that

$$P\left(|\mathcal{M}_n(h(Q)) - \mathcal{M}_n(h(T))| \ge c\gamma_n \log^2 n\right) = O(n^{-1/2}).$$

On the other hand, if $\mathcal{N}_n(F_k \cup F_{k+1}) = 0$, then

$$\mathcal{M}_n(h(R_k)\,|\mathcal{Z}\,) = \mathcal{M}_n(h(R_k)\,|\mathcal{T}\,),$$

where $\mathcal{M}_n(\mathcal{A}_1\,|\mathcal{A}_2\,)$ denotes the number of maximal points of $\mathcal{X} \cap \mathcal{A}_2$ that also lie in $\mathcal{A}_1$. Now if

$$\sup_{1 \le k < \lambda} |h(R_k)| = O(\log n/n), \tag{45}$$

then similar arguments as above leads to

$$P\left(|\mathcal{M}_n(h(R)\,|\mathcal{Z}) - \mathcal{M}_n(h(R)\,|\mathcal{T})| \geq c\gamma_n \log^2 n\right) = O(n^{-1/2}). \tag{46}$$

However, if (45) fails, then we define the subset of $h(R_k)$:

$$\hat{R}_k := \{p \ : \ p \in h(R_k) \text{ and } |\triangle(p, \mathcal{T} \cap \mathcal{Z})| \leq 2\log n/n\},$$

which satisfies $|\hat{R}_k| \leq 8 \log n/n$ for all $1 \leq k < \lambda$. Also the region

$$\{p \ : \ p \in h(R_k) \text{ and } |\triangle(p, \mathcal{T} \cap \mathcal{Z})| > 2\log n/n\},$$

is negligible (by a similar argument). Thus (46) also holds and this proves (32).

Note that the proof can be largely simplified if f is known to be convex.

Proof of (33). Let $N_h(\mathcal{A})$ be the number of points in $\{h(X_1), ..., h(X_n)\} \cap \mathcal{A}$. Note that $|\mathcal{Z}| = O(\sqrt{\log n/n})$. Then $M_h(\mathcal{Z}) \overset{d}{=} \mathcal{M}_n(\mathcal{Z})$ when conditioning on $N_h(\mathcal{Z}) = \mathcal{N}_n(\mathcal{Z})$. It follows that

$$\begin{aligned}
&\sup_x \left| P\left(\frac{M_h(\mathcal{Z}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\right) - P\left(\frac{\mathcal{M}_n(\mathcal{Z}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\right)\right| \\
&= \sup_x \left| \begin{array}{l} \displaystyle\sum_k P\left(\frac{M_h(\mathcal{Z}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\,|N_h(\mathcal{Z}) = k\right) P\left(N_h(\mathcal{Z}) = k\right) \\ \displaystyle -\sum_k P\left(\frac{\mathcal{M}_n(\mathcal{Z}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\,|\mathcal{N}_n(\mathcal{Z}) = k\right) P\left(\mathcal{N}_n(\mathcal{Z}) = k\right) \end{array} \right| \\
&= \sup_x \left| \sum_k P\left(\frac{M_h(\mathcal{Z}) - \alpha\sqrt{\pi n}}{\sqrt{\alpha\sigma^2\sqrt{n}}} < x\,|N_h(\mathcal{Z}) = k\right) \left(P\left(N_h(\mathcal{Z}) = k\right) - P\left(\mathcal{N}_n(\mathcal{Z}) = k\right)\right)\right| \\
&\leq \sum_k \left|P\left(N_h(\mathcal{Z}) = k\right) - P\left(\mathcal{N}_n(\mathcal{Z}) = k\right)\right| \\
&= O\left(\sqrt{\log n/n}\right),
\end{aligned}$$

where the last estimate holds by the usual Poison approximation of binomial distributions:

$$\sum_{k\geq 0} \left| \binom{n}{k} p^k (1-p)^{n-k} - e^{-np}\frac{(np)^k}{k!}\right| = O(p),$$

for small p; see Prohorov (1953).

Proof of (34). Recall that $\mathcal{D} = Q \cup R \cup W$. Obviously, $M_n(W)$ is negligible and

$$M_n(Q) = M_h(h(Q)),$$

since h preserves the dominance relation inside all strips $u_{k-1} \leq u < u_k$ and $M_n(Q_k) = M_h(h(Q_k))$ for all $1 \leq k \leq \lambda$.

We now construct another mapping h_1 such that $M_{h_1}(\mathcal{Z}) \overset{d}{=} M_h(\mathcal{Z})$ and

$$P(M_{h_1}(h(R)) - M_n(R) > c\gamma_n \log n) = O(n^{-1}).$$

Note that the height of $h(T_k)$ is $1/\nu$ and that of $h(R_k)$ is $u_k/(\nu u_{k+1})$. So that if $p_0 \in R_k$ and $\nabla u_k \geq \nabla u_{k+1}$, then

$$h\left(\triangle(p_0, \mathcal{D})\right) \subseteq \triangle\left(h(p_0), h(\mathcal{D})\right),$$

that is, $p_0 \in R_k$ is a maximal point of $\{X_1, ..., X_n\}$ if $h(p_0)$ is a maxima point of $\{h(X_1), ..., h(X_n)\}$, and thus

$$M_h(h(R_k)) \leq M_n(R_k).$$

Consider now the case when $\nabla u_k < \nabla u_{k+1}$. Define (recall the definitions of (42) and (43))

$$g_1(u,v) = \begin{cases} (u, v - l_k), & \text{on } h(R_k), \\ (u, v + 1/\nu - l_k), & \text{on } L_k, \\ (u,v), & \text{elsewhere.} \end{cases}$$

Let $h_1 = g_1 \circ h$. Then h only differs from h_1 on $h(R_k)$ and L_k when $\nabla u_k < \nabla u_{k+1}$. For $p_0 \in R_k$,

$$h_1(\triangle(p_0, \mathcal{D})) \subseteq \triangle\left(h_1(p_0), h_1(\mathcal{D})\right),$$

that is, p_0 is a maximal point of $\{X_1, ..., X_n\}$ if $h_1(p_0)$ is a maximal point of $\{h_1(X_1), ..., h_1(X_n)\}$. Thus the relation

$$M_{h_1}(h(R_k)) \leq M_n(R_k) + M_{h_1}(L_k),$$

holds for all $1 \leq k < \lambda$. So that

$$M_{h_1}(h(R)) \leq M_n(R) + M_{h_1}(L).$$

Since $|L| = O(\gamma_n \log n/n)$, we have $P\left(M_{h_1}(L) \geq c\gamma_n \log n\right) \leq O(n^{-1})$. This completes the proof of (34).

The estimate (35) is proved similarly.

The proof of Theorem 3 is now complete.

Remark. Theorem 3 holds for more general f. For example, it holds when f is twice differentiable except at a finite number of points and the number of components in which f satisfies $\{u : f'(u) = 0\}$ is finite (γ_n has to be suitably modified). The method of proof is to split the unit interval into finite number of subintervals in which either f is twice differentiable with $f'(u) < 0$ or $f'(u) = 0$ in each subinterval. The contribution of the rectangles in the subinterval for which $f'(u) = 0$ is negligible (being at most of order $\log n$; see Bai et al. 2001). Then we define γ_n in each subinterval in which $f' < 0$ as above and argue similarly.

4 Conclusions

Dominance is an extremely useful notion in diverse fields, and stochastic problems associated with it introduce concrete, intriguing, challenging problems for probabilists.

We conclude this paper with a few questions. First what is the optimal rate in (29)? Is it $n^{-1/4}\gamma_n$ or $n^{-1/4}\gamma_n \log n$ for smooth f? Second, what can one say about large deviations? Almost no results are known along this direction. Third, how to derive optimal Berry-Esseen bounds for maxima in higher dimensions? Even the simplest case of hypercubes remains unknown, although one expects a rate of order $(\log n)^{-(d-1)/2}$ for the d-dimensional hypercube (see Bai et al., 1998). Finally, we can address the same questions for some structural parameters (like the number of hull points) in the convex hull of a random sample chosen from some planar regions. Central limit theorems have been derived but no convergence rates are known; see Groeneboom (1988), Cabo and Groeneboom (1994).

A seemingly more natural problem is "what regularity conditions on f implies the asymptotic normality of $M_n(\mathcal{D})$?"

Acknowledgements

We thank Louis Chen for pointing out another approach to Berry-Esseen bound based on Stein's method and concentration inequality. Part of the work was done while the second and the third authors were visiting Department of Statistics and Applied Probability, National University of Singapore; they thank the Department for its hospitality.

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, 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**, paper no. 3, 41 pages; available at www.math.washington.edu/∼ejpecp/EjpVol6/paper3.pdf

[3] A. D. Barbour and A. Xia (2001), The number of two dimensional maxima, *Advances in Applied Probability*, **33**, 727–750.

[4] Y. Baryshnikov (2000), Supporting-points processes and some of their applications, *Probability Theory and Related Fields*, **117**, 163–182.

[5] R. A. Becker, L. Denby, R. McGill and A. R. Wilks (1987), Analysis of data from the "Places Rated Almanac", *The American Statistician*, **41**, 169–186.

[6] A. J. Cabo and P. Groeneboom (1994), Limit theorems for functionals of convex hulls, *Probability Theory and Related Fields*, **100**, 31–55.

[7] T. M. Chan (1996), Output-sensitive results on convex hulls, extreme points, and related problems, *Discrete and Computational Geometry*, **16**, 369–387.

[8] S. N. Chiu and M. P. Quine, Central limit theory for the number of seeds in a growth model in $\mathbb{R}^d$ with inhomogeneous Poisson arrivals, *Annals of Applied Probability*, **7** (1997), 802–814.

[9] A. Datta and S. Soundaralakshmi (2000), An efficient algorithm for computing the maximum empty rectangle in three dimensions, *Information Sciences*, **128**, 43–65.

[10] L. Devroye (1993), Records, the maximal layer, and the uniform distributions in monotone sets, *Computers and Mathematics with Applications*, **25**, 19–31.

[11] M. E. Dyer and J. Walker (1998), Dominance in multi-dimensional multiple-choice knapsack problems, *Asia-Pacific Journal of Operational Research*, **15**, 159–168.

[12] I. Z. Emiris, J. F. Canny and R. Seidel (1997), Efficient perturbations for handling geometric degeneracies, *Algorithmica*, **19**, 219–242.

[13] J. L. Ganley (1999), Computing optimal rectilinear Steiner trees: A survey and experimental evaluation, *Discrete Applied Mathematics*, **90**, 161–171.

[14] M. J. Golin (1993), Maxima in convex regions, in *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, (Austin, TX, 1993), 352–360, ACM, New York.

[15] P. Groeneboom (1988), Limit theorems for convex hulls, *Probability Theory and Related Fields*, **79**, 327–368.

[16] H.-K. Hwang (2003), Second phase changes in random m-ary search trees and generalized quicksort: convergence rates, *Annals of Probability*, **31**, 609–629.

[17] R. E. Johnston and L. R. Khan (1995), A note on dominance in unbounded knapsack problems, *Asia-Pacific Journal of Operational Research*, **12**, 145–160.

[18] S. Martello and P. Toth (1990), *Knapsack Problems: Algorithms and Computer Implementations*, John Wiley & Sons, New York.

[19] R. Neininger and L. Rüschendorf (2002), A general contraction theorem and asymptotic normality in combinatorial structures, *Annals of Applied Probability*, accepted for publication (2003); available at www.stochastik.uni-freiburg.de/homepages/neininger.

[20] V. V. Petrov (1975), *Sums of Independent Random Variables*, Springer-Verlag, New York.

[21] 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 of: *Uspehi Matematičeskih Nauk*, **8** (1953), no. 3 (35), 135–142.

[22] M. Zachariasen (1999), Rectilinear full Steiner tree generation, *Networks*, **33**, 125–143.