The Annals of Applied Probability
2003, Vol. 13, No. 4, 1252–1263

ON THE MULTIPLICITY OF THE MAXIMUM IN A DISCRETE RANDOM SAMPLE

BY F. THOMAS BRUSS AND RUDOLF GRÜBEL

Université Libre de Bruxelles and Universität Hannover

Let M_n be the maximum of a sample $X_1, \dots, X_n$ from a discrete distribution and let W_n be the number of i's, $1 \le i \le n$, such that $X_i = M_n$. We discuss the asymptotic behavior of the distribution of W_n as $n \to \infty$. The probability that the maximum is unique is of interest in diverse problems, for example, in connection with an algorithm for selecting a winner, and has been studied by several authors using mainly analytic tools. We present here an approach based on the Sukhatme–Rényi representation of exponential order statistics, which gives, as we think, a new insight into the problem.

1. Introduction and results. Let $X_1, \dots, X_n$ be independent and identically distributed integer valued random variables. Let $M_n := \max\{X_1, \dots, X_n\}$ be the maximum of the sample and let $W_n := \#\{1 \le i \le n : X_i = M_n\}$ be the multiplicity of the maximum, $\rho_n := P(W_n = 1)$ is the probability that the maximum is unique. An example where the multiplicity of the maximum is of interest arises in connection with selection algorithms. If, say, a chairperson is to be determined, the individual committee members could throw a coin repeatedly in successive rounds and leave the competition if they obtain a head; this results in a tie if more than one person is left and they all throw heads in the same round. In this example, X_i corresponds to the time of first appearance of a head in the coin tossed by the ith member of the group, that is, with p the probability for head and $q := 1 - p$ we have $P(X_i = k) = q^{k-1}p$ for all $k \in \mathbb{N}$ and ρ_n is the probability that the procedure does not end in a tie. Bruss and O'Cinneide (1990) showed that ρ_n does not converge as $n \to \infty$ for such geometric distributions, but that

$$\lim_{n\to\infty} \bigl(\rho_n - \Psi(n)\bigr) = 0 \qquad \text{with } \Psi(t) := pt \sum_{k\in\mathbb{Z}} q^k \exp(-tq^k)$$

[equation (13) in Bruss and O'Cinneide (1990) only deals with $p = 1/2$, but the argument given there is easily extended to the case of general p]. Note that Ψ is logarithmically periodic: $\Psi(qt) = \Psi(t)$ for all $t > 0$. In particular, writing $\{x\}$ for the fractional part of $x \in \mathbb{R}$, we see from this very periodicity that $\Psi(n)$ depends on n only through $\{\log_q n\}$ which implies that $(\rho_{n_k})_{k\in\mathbb{N}}$ does converge along specific subsequences $(n_k)_{k\in\mathbb{N}}$.

This somewhat surprising result has been rediscovered several times [see the addendum by Kirschenhofer and Prodinger (1998)]. Considering more general

Received May 2002; revised October 2002.
AMS 2000 subject classifications. 60C05 (60F05, 62G30).
Key words and phrases. Convergence in distribution, exponential distribution, order statistics, probabilistic constructions, quantile transformation, Sukhatme–Rényi representation.

distributions, Baryshnikov, Eisenberg and Stengle (1995) show that a limiting probability of a tie for the maximum (and hence $\lim_{n\to\infty} \rho_n$) exists if and only if $P(X_1 = k)/P(X_1 > k) \to 0$ as $k \to \infty$. Eisenberg, Stengle and Strang (1993) go beyond the study of ρ_n and also discuss the distribution of W_n; Brands, Steutel and Wilms (1994) and Kirschenhofer and Prodinger (1996) obtained rates of convergence in the geometric case. For the asymptotics of the maximum M_n itself in the case of general discrete distributions see Athreya and Sethuraman (2001) and the references given there. Fill, Mahmoud and Szpankowski (1996) give a more detailed description and an analysis of the duration of a variant of the above election algorithm, the base distribution is geometric. In these papers there is often a first probabilistic step, resulting in some equation for the quantities of interest, then analytic machinery is used to obtain the desired result. Kirschenhofer and Prodinger (1996), for example, emphasize the use of complex variable techniques. In the present paper we offer a somewhat more probabilistic approach, based on the idea of representing the situation as a discretization of some continuous (and well understood) background model. This is, of course, one of the standard methods of applied probability; see, for example, Grübel and Reimers (2001b) for a similar strategy in a record-renewal problem. The background model we use here combines the quantile transformation and the Sukhatme–Rényi representation of the order statistics associated with a sample from an exponential distribution.

We believe that this method can lead to a better understanding of the behavior of maxima and their multiplicities in discrete samples, but it can also be used to provide alternative proofs for or to improve upon existing results. From the surprisingly numerous papers that deal with aspects of discrete maxima we have chosen two specific questions in order to support this view, leading to the two theorems below. However, the method should also be applicable in other situations not considered here, for example, in connection with the joint distribution of M_n and W_n or in an asymptotic analysis of the last k rounds for k fixed, $n \to \infty$.

Our first theorem relates the asymptotic behavior of ρ_n as $n \to \infty$ to that of the tail ratios $P(X_1 \ge k+1)/P(X_1 \ge k)$ as $k \to \infty$. We assume throughout that $P(X_1 \in \mathbb{N}) = 1$ and that $P(X_1 = k) > 0$ for all $k \in \mathbb{N}$. This simplifies the notation and the generalization to an arbitrary $A \subset \mathbb{R}$ of the form $A = \{a_k : k \in \mathbb{N}\}$ with some strictly increasing sequence $(a_k)_{k\in\mathbb{N}}$ of real numbers is trivial.

THEOREM 1. (a) $\liminf_{n\to\infty} \rho_n \ge \liminf_{k\to\infty} P(X_1 > k | X_1 \ge k)$.

(b) *If* $\liminf_{k\to\infty} P(X_1 > k | X_1 \ge k) < 1$, *then* $\liminf_{n\to\infty} \rho_n < 1$.

(c) $\limsup_{n\to\infty} \rho_n > 0$.

Parts (a) and (b) imply that $\lim_{n\to\infty} \rho_n = 1$ is equivalent to $\lim_{k\to\infty} P(X_1 = k | X_1 \ge k) = 0$; this also follows from the above-mentioned result of Baryshnikov, Eisenberg and Stengle (1995). For (c) we clearly need that there is no $a \in \mathbb{R}$ with $P(X_1 \le a) = 1$ and $P(X_1 = a) > 0$, which is an immediate consequence of our general assumptions; obviously $\rho_n \to 0$ if such a point a exists. Eisenberg, Stengle and Strang (1993) found it "striking" that $\lim_{n\to\infty} \rho_n = 0$ only holds in such degenerate cases (they also obtain the stronger result $\limsup_{n\to\infty} \rho_n > e^{-1}$ by elementary means). In this context the contribution of our method should perhaps be seen as turning this result and others on the qualitative behavior of ρ_n into something intuitively plausible.

Our second theorem deals with the distribution of W_n for large n in the geometric case. The family $\{Q_{p,\eta} : 0 \le \eta < 1\}$ of distributions that arise as limit points if the base distribution is geometric with parameter $p = 1 - q$ is given by

$$Q_{p,\eta}(\{l\}) := \frac{p^l}{l!} \sum_{j\in\mathbb{Z}} q^{l(j+\eta)} e^{-q^{j+\eta}} \qquad \text{for all } l \in \mathbb{N}.$$

Brands, Steutel and Wilms (1994) obtained a rate of convergence result for the individual probabilities, Kirschenhofer and Prodinger (1996) extended this to the first two moments. Here we consider a distance measure between the distribution of W_n and a suitable $Q_{p,\eta}$ that covers the behavior of the moment generating functions in a fixed neighborhood of 0. Our result therefore implies convergence with rate $O(1/n)$ of the total variation distance and of all moments. It can also be used to show that the approximation of the distribution $\mathcal{L}(W_n)$ of W_n by a member of the $Q_{p,\cdot}$-family is precise enough to capture the fluctuations of quantities like $P(W_n \ge \log\log n)$; see Grübel and Reimers (2001a) for a similar situation arising in the analysis of von Neumann addition.

THEOREM 2. *Suppose that X_i, $i \in \mathbb{N}$, are independent and geometrically distributed with parameter $p = 1 - q$. Let $\eta_n := \{\log_q n\}$. Then, for all $\gamma < 1/p$,*

$$\sum_{l=1}^{\infty} \gamma^l \bigl|P(W_n = l) - Q_{p,\eta_n}(\{l\})\bigr| = O\left(\frac{1}{n}\right) \qquad \text{as } n \to \infty.$$

The proofs of these theorems are given in Sections 3 and 4, respectively; the probabilistic construction on which they are based is explained and illustrated in Section 2. We let $\mathrm{Exp}(\lambda)$ denote the exponential distribution with parameter λ, $\Gamma(\alpha, \lambda)$ is the gamma distribution with shape parameter α and scale parameter λ, and $U =_{\mathrm{distr}} V$ means that the random quantities U and V have the same distribution.

2. The construction. Using the general assumptions on the distribution of X_1 we see that

$$x_k := -\log P(X_1 > k), \qquad k \in \mathbb{N}_0,$$

defines a sequence $(x_k)_{k\in\mathbb{N}_0}$ of nonnegative real numbers that strictly increase to ∞. Let $\phi : [0, \infty) \to \mathbb{N}$ be the function that takes the value k on the interval $[x_{k-1}, x_k)$. With this function we have $X_1 =_{\mathrm{distr}} \phi(Y_1)$ if $\mathcal{L}(Y_1) = \mathrm{Exp}(1)$.

Moreover, on extending the basic probability space if necessary, we may assume that $X_i = \phi(Y_i)$ for all $i \in \mathbb{N}$, with $(Y_i)_{i\in\mathbb{N}}$ a sequence of independent, Exp(1)-distributed random variables. In particular,

$$(X_1, \dots, X_n) =_{\text{distr}} \big(\phi(Y_1), \dots, \phi(Y_n)\big).$$

As ϕ is increasing, this obviously implies

$$\big(X_{(n:1)}, X_{(n:2)}, \dots, X_{(n:n)}\big) =_{\text{distr}} \big(\phi(Y_{(n:1)}), \phi(Y_{(n:2)}), \dots, \phi(Y_{(n:n)})\big),$$

where $X_{(n:m)}$, $Y_{(n:m)}$, $m = 1, \dots, n$, denote the increasing order statistics associated with $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$, respectively.

The quantile transformation, represented by the function ϕ, reduces the study of distributions related to the order statistics of a sample from an arbitrary distribution to the special case of exponential distributions. The Sukhatme–Rényi representation [see, e.g., Shorack and Wellner (1986), page 721] is a structural result for the order statistics in the exponential case; it says that

$$\big(Y_{(n:1)}, Y_{(n:2)}, \dots, Y_{(n:n)}\big) =_{\text{distr}} (V_n, V_n + V_{n-1}, \dots, V_n + \cdots + V_1),$$

with $V_1, \dots, V_n$ independent and $\mathcal{L}(V_i) = \text{Exp}(i)$, $1 \le i \le n$. Putting these two steps together we can therefore base a construction of the sequence $(W_n)_{n\in\mathbb{N}}$ on a sequence $(V_i)_{i\in\mathbb{N}}$ of independent, Exp(i)-distributed random variables V_i: For all $n \in \mathbb{N}$, $W_n =_{\text{distr}} \tilde{W}_n$, where $\tilde{W}_n$ is the maximal l with the property that

$$V_n + V_{n-1} + \cdots + V_i \in [x_k, x_{k+1}) \qquad \text{for } i = 1, \dots, l$$

for some $k \in \mathbb{N}_0$. Of course, this does not preserve the joint distribution of the W_n's, but joint distributions are not required for the assertions made in Section 1.

The following interpretation of the construction may be helpful: Let $N^{(n)} = (N_t^{(n)})_{t\ge 0}$ be a continuous time Markov chain with state space $\mathbb{N}_0$, transition rates $q_{i,i-1} = i$ for $i > 0$, absorption at 0 and start in n. Markov chains with this simple structure are also known as pure death processes. Regarding the V-variables as holding times we see that the above implies

$$(\#\{1 \le i \le n : X_i \ge k\})_{k\in\mathbb{N}_0} =_{\text{distr}} \big(N_{x_k}^{(n)}\big)_{k\in\mathbb{N}_0}.$$

This sequence represents the whole selection process. Note that $\tilde{W}_n$ refers to just one aspect of $N^{(n)}$, namely $\min\{N_{x_k}^{(n)} : k \in \mathbb{N}_0, N_{x_k}^{(n)} > 0\}$. For geometric distributions the sequence $(x_k)_{k\in\mathbb{N}_0}$ is of the simple form $x_k = k\lambda$ for some $\lambda > 0$. In the terminology of the selection rule mentioned in the introduction this means that the number of candidates that survive the successive rounds can be obtained distributionally from sampling the process $N^{(n)}$ at integer multiples of $\lambda = -\log(1 - p)$, with p the probability for a head. Figure 1 illustrates this with $n = 20$; we see that in this particular case four candidates remain in the competition after the third round and that the sample maximum is not unique.

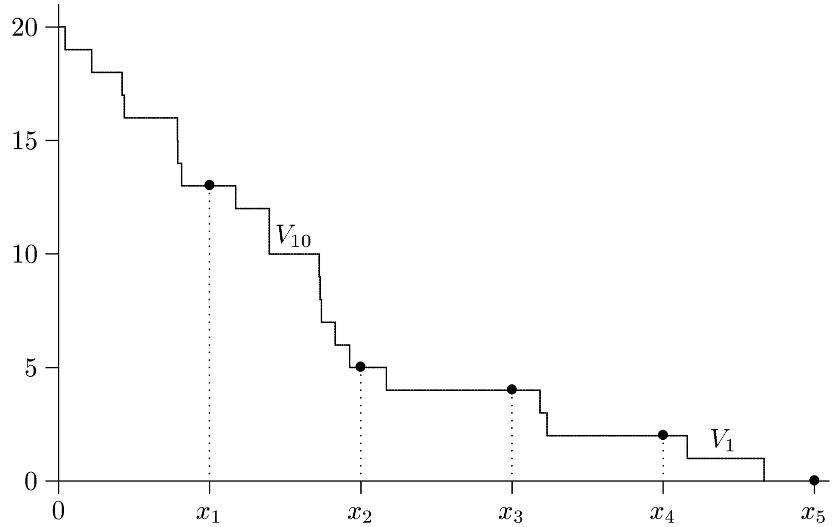

FIG. 1. *Sampling the death process.*

For $\tilde{W}_n = 1$ it is necessary and sufficient that the last interval

$$[V_n + \cdots + V_2, V_n + \cdots + V_2 + V_1)$$

contains one of the x_k's, $k \in \mathbb{N}_0$. In our construction this interval has length V_1 and hence does not depend on n, in contrast to the quantity $Y_{(n:n)} - Y_{(n:n-1)}$ in the original quantile representation. This independence from n makes the non-convergence of ρ_n in the geometric case and other results intuitively obvious, but the formal proofs may still require some care.

3. Proof of Theorem 1. For the proof of (a) we first note that

$$\limsup_{k\to\infty}(x_k - x_{k-1}) = -\log\kappa \qquad \text{with } \kappa := \liminf_{k\to\infty} P(X_1 > k | X_1 \geq k).$$

Let $\delta > 0$ be given and let $C = C(\delta) < \infty$ be such that $x_{k+1} - x_k < \delta - \log\kappa$ whenever $x_k > C$. In terms of the construction explained in the previous section we then obviously have

$$\{\tilde{W}_n = 1\} \supset \{V_1 \geq \delta - \log\kappa\} \cap \{V_2 + \cdots + V_n \geq C\}.$$

The sequence of events on the right-hand side is increasing; clearly, $V_2 + \cdots + V_n \to \infty$ almost surely as $n \to \infty$. Hence,

$$\liminf_{n\to\infty} \rho_n = \liminf_{n\to\infty} P(\tilde{W}_n = 1) \geq P(V_1 \geq \delta - \log\kappa) = \kappa e^{-\delta},$$

and letting $\delta \downarrow 0$ completes the proof of part (a).

For the proof of (b) let $Z_n := \sum_{k=2}^n V_k$, $a_n := \sum_{k=2}^n k^{-1}$. It is straightforward to show that $Z_n - a_n$ converges in quadratic mean, and hence in probability, to some finite random variable Z_∞. Let r be an element of the support of $\mathcal{L}(Z_\infty)$ so that

$$P(|Z_\infty - r| < \delta) > 0 \qquad \text{for all } \delta > 0.$$

If $\liminf_{k\to\infty} P(X_1 > k | X_1 \ge k) < 1$, then we can find an $\varepsilon > 0$ and a subsequence $(x_{k_j})_{j\in\mathbb{N}}$ of $(x_k)_{k\in\mathbb{N}_0}$ such that

$$x_{k_j+1} - x_{k_j} \ge 6\varepsilon \qquad \text{for all } j \in \mathbb{N}.$$

Since $x_{k_j} \uparrow \infty$, $a_n \uparrow \infty$ and $a_{n+1} - a_n \downarrow 0$ we can further find a $j_0 \in \mathbb{N}$ such that each of the intervals $(x_{k_j} + 2\varepsilon, x_{k_j} + 3\varepsilon)$, $j \ge j_0$, contains at least one member of the sequence $(r + a_n)_{n\in\mathbb{N}}$. This means that there exists a subsequence $(r + a_{n_j})_{j\in\mathbb{N}}$ such that

$$x_{k_j} + 2\varepsilon < r + a_{n_j} < x_{k_j} + 3\varepsilon \qquad \text{for all } j \ge j_0.$$

Consider now the events

$$A = \{V_1 < \varepsilon\}, \qquad B = \{|Z_\infty - r| < \varepsilon\}, \qquad C_n = \{|Z_n - a_n - Z_\infty| < \varepsilon\}.$$

Using $Z_n = r + a_n + Z_\infty - r + Z_n - a_n - Z_\infty$ we see that

$$x_{k_j} < Z_{n_j} = V_2 + \cdots + V_{n_j}, \qquad V_1 + \cdots + V_{n_j} = V_1 + Z_{n_j} < x_{k_j+1}$$

and hence $\tilde{W}_{n_j} \ge 2$ on $A \cap B \cap C_{n_j}$ for all $j \ge j_0$. Since V_1 and Z_n are independent we therefore obtain

$$\begin{aligned}
1 - \liminf_{n\to\infty} \rho_n &= \limsup_{n\to\infty} P(W_n \ge 2) \\
&\ge \limsup_{j\to\infty} P\big(\tilde{W}_{n_j} \ge 2\big) \\
&\ge P(A) \limsup_{j\to\infty} P\big(B \cap C_{n_j}\big) \\
&\ge (1 - e^{-\varepsilon})\Big(P(B) - \lim_{n\to\infty} P(C_n^c)\Big) \\
&> 0.
\end{aligned}$$

For the proof of (c) we can proceed as in (b), forcing Z_{n_k} into the interval $(x_k - 1, x_k)$ for k large enough and then using $P(V_1 > 1) > 0$. Alternatively we can find subsequences $(a_{n_j})_{j\in\mathbb{N}}$ and $(x_{k_j})_{j\in\mathbb{N}}$ such that $\lim_{j\to\infty}(a_{n_j} - x_{k_j}) = 0$ and then use

$$\begin{aligned}
\limsup_{n\to\infty} \rho_n &\ge \limsup_{j\to\infty} P\big(\tilde{W}_{n_j} = 1\big) \\
&\ge \limsup_{j\to\infty} P\big(Z_{n_j} - a_{n_j} < x_{k_j} - a_{n_j} < Z_{n_j} - a_{n_j} + V_1\big) \\
&= P(Z_\infty < 0 < Z_\infty + V_1),
\end{aligned}$$

with the last equality following with Slutsky's lemma and the continuity of the distribution function of Z_∞. We suspect that $P(Z_\infty < 0 < Z_\infty + V_1) = e^{-1}$, but we do not have a proof.

4. Proof of Theorem 2. We first collect some properties of the background construction. As in Section 2 let $Y_{(n:l)}$, $1 \le l \le n$, be the order statistics associated with a sample of size n from Exp(1).

LEMMA 3. (a) *$Y_{(n:n-l)}$ and $Y_{(n:n)} - Y_{(n:n-l)}$ are independent.*

(b) *$P(Y_{(n:n)} - Y_{(n:n-l)} \le y) = (1 - e^{-\lambda y})^l$ for all $y \ge 0$.*

PROOF. With $V_1, \dots, V_n$ as in Section 2,

$$Y_{(n:l)} = V_n + \cdots + V_{n-l+1}, \qquad Y_{(n:n)} - Y_{(n:l)} = V_{n-l} + \cdots + V_1,$$

from which the independence follows immediately. Further, the second of these equalities implies that $Y_{(n:n)} - Y_{(n:l)}$ is equal in distribution to $Y_{(n-l:n-l)}$. This is the maximum of a sample of size $n - l$ from Exp(1), which gives (b). $\square$

From the familiar formula for the density of order statistics [see, e.g., David (1981), page 9] it follows easily that

$$f_{nl}(z) := \frac{e^{-lz}}{(l-1)!} \frac{n!}{(n-l)!\,n^l} \left(1 - \frac{e^{-z}}{n}\right)^{n-l}, \qquad z > -\log n,$$

is a density of $Y_{(n:n-l+1)} - \log n$. Further, if $\mathcal{L}(S_l) = \Gamma(l, 1)$ then

$$g_l(z) := \frac{e^{-lz}}{(l-1)!} \exp(-e^{-z}), \qquad z \in \mathbb{R},$$

is a density for $-\log S_l$. It is easy to see that $f_{nl}(z) \to g_l(z)$ for all $z \in \mathbb{R}$ as $n \to \infty$ for each fixed l, but we need more. We require an upper bound for the L^1-distance of these functions that holds uniformly in $l = O(\log n)$.

LEMMA 4. *For each $C_1 < \infty$ there exists a $C_2 < \infty$ such that*

$$\int |f_{nl}(z) - g_l(z)|\, dz \le C_2 \frac{l^2}{n} \qquad \textit{for all } n \in \mathbb{N}, l \le C_1 \log n.$$

PROOF. We split the integral and consider the intervals $(-\infty, -(\log n)/3]$ and $(-(\log n)/3, \infty)$ separately. For the first of these we use the crude bound

$$\int_{-\infty}^{-(\log n)/3} |f_{nl}(z) - g_l(z)|\, dz \le \int_{-\infty}^{-(\log n)/3} f_{nl}(z)\, dz + \int_{-\infty}^{-(\log n)/3} g_l(z)\, dz.$$

The second term on the right-hand side leads to

$$\int_{-\infty}^{-(\log n)/3} g_l(z)\, dz = P\bigl(-\log Z_l \le -(\log n)/3\bigr) = P\bigl(Z_l \ge n^{1/3}\bigr)$$

with $\mathcal{L}(Z_l) = \Gamma(l, 1)$ as above. The function $l \mapsto P(Z_l \ge z)$ is increasing for fixed z, a fact that follows easily from the convolution property $\Gamma(\alpha, \lambda) \star \Gamma(\beta, \lambda) = \Gamma(\alpha + \beta, \lambda)$ of the gamma family of distributions. It is therefore enough to bound the tail for $l = l_n := C_1 \log n$. For this we use a familiar argument: The moment generating function associated with $\Gamma(\alpha, 1)$ is given by $t \mapsto (1 - t)^{-\alpha}$, $t < 1$. Hence with $t = 1 - l_n/n^{1/3}$,

$$\begin{aligned} P\big(Z_{l_n} \ge n^{1/3}\big) &\le e^{-tn^{1/3}}(1-t)^{-l_n} \\ &= \exp\big(-n^{1/3} + l_n - l_n \log l_n + \tfrac{1}{3} l_n \log n\big), \end{aligned}$$

which is obviously $O(1/n)$ as $l_n = O(\log n)$. For the f_{nl}-part we use $n! \le (n-l)!\,n^l$ and $\log(1-x) \le -x$ for $x > -1$ to obtain, with the change of variable $y = e^{-z}$,

$$\begin{aligned} \int_{-\infty}^{-(\log n)/3} f_{nl}(z)\,dz &\le \int_{-\log n}^{-(\log n)/3} \frac{e^{-lz}}{(l-1)!}\left(1 - \frac{e^{-z}}{n}\right)^{n-l} dz \\ &= \int_{n^{1/3}}^{n} \frac{y^{l-1}}{(l-1)!} \exp\left((n-l)\log\left(1 - \frac{y}{n}\right)\right) dy \\ &\le \int_{n^{1/3}}^{n} \frac{1}{(l-1)!} \exp\left(-y + \frac{l}{n} y + (l-1)\log y\right) dy \\ &\le \exp(-n^{1/3})\, n \exp\big(C_1 \log n + C_1 (\log n)^2\big), \end{aligned}$$

where we used in the last inequality the assumption on l as specified in the lemma. The upper bound is of order $O(1/n)$, as required.

For the integral over the range $(-(\log n)/3, \infty)$ we again use the change of variable $y = e^{-z}$ to obtain

$$\begin{aligned} \int_{-(\log n)/3}^{\infty} |f_{nl}(z) - g_l(z)|\,dz &= \int_0^{n^{1/3}} \frac{1}{y} |f_{nl}(-\log y) - g_l(-\log y)|\,dy \\ &= \int_0^{n^{1/3}} \frac{y^{l-1}e^{-y}}{(l-1)!} \big|\exp\big(h_{nl}(y)\big) - 1\big|\,dy \end{aligned}$$

with

$$h_{nl}(y) = \sum_{k=1}^{l-1} \log\left(1 - \frac{k}{n}\right) + (n-l)\log\left(1 - \frac{y}{n}\right) + y.$$

As $l_n = o(n)$ we can use the fact that $|\log(1-x)| \le 2|x|$ in a neighborhood of $x = 0$ to obtain

$$\left|\sum_{k=1}^{l-1} \log\left(1 - \frac{k}{n}\right)\right| \le 2 \sum_{k=1}^{l-1} \frac{k}{n} = \frac{l(l-1)}{n}$$

for n large enough. Similarly, since $|\log(1-x)+x| \le x^2$ near $x=0$, for n large enough,

$$\begin{aligned}\left|(n-l)\log\left(1-\frac{y}{n}\right)+y\right| &\le (n-l)\left|\log\left(1-\frac{y}{n}\right)+\frac{y}{n}\right|+\frac{ly}{n}\\ &\le \frac{y^2}{n}+\frac{ly}{n},\end{aligned}$$

uniformly over $l \le C_1 \log n$, $0 \le y \le n^{1/3}$. Continuing in this vein we use that $|e^x - 1| \le 2x$ in a neighborhood of $x=0$, which on putting pieces together leads to

$$\left|\exp\bigl(h_{nl}(y)\bigr)-1\right| \le 2\left(\frac{l(l-1)}{n}+\frac{y^2}{n}+\frac{ly}{n}\right)$$

uniformly over $l \le C_1 \log n$ and $0 \le y \le n^{1/3}$, for n large enough. [It is here that the exponent $1/3$ is used as we need $y^2/n = o(1)$.]

We can therefore bound the remaining integral as

$$\begin{aligned}&\int_0^{n^{1/3}} \frac{y^{l-1}e^{-y}}{(l-1)!}\left|\exp\bigl(h_{nl}(y)\bigr)-1\right|dy\\ &\quad\le \frac{2}{n}\left(l(l-1)\int_0^\infty \frac{y^{l-1}e^{-y}}{(l-1)!}\,dy+\int_0^\infty \frac{y^{l+1}e^{-y}}{(l-1)!}\,dy+l\int_0^\infty \frac{y^{l}e^{-y}}{(l-1)!}\,dy\right)\\ &\quad= \frac{2}{n}\bigl(l(l-1)+l(l+1)+l^2\bigr),\end{aligned}$$

which is of the desired form. $\square$

The next lemma connects the distributions $Q_{p,\eta}$ to the functions g_l. Let $q = 1-p$ as before and $\lambda := -\log q$.

LEMMA 5. *If Z_l is a random variable with density g_l, then*

$$\sum_{k=l}^{\infty} Q_{p,\eta}(\{k\}) = E\bigl(1-q\exp(\lambda\{\lambda^{-1}Z_l-\eta\})\bigr)^{l-1}.$$

PROOF. If Z_{l+1} has density g_{l+1}, then $S_{l+1} := \exp(-Z_{l+1})$ has density $y \mapsto y^l e^{-y}/l!$. With $b_j := q^{j+\eta}$ we have

$$j \le \lambda^{-1}Z_{l+1}-\eta < j+1 \quad \Longleftrightarrow \quad b_{j+1} < S_{l+1} \le b_j$$

so that, for $k = 1, \dots, l$,

$$
\begin{aligned}
E \exp\bigl(\lambda k\{\lambda^{-1} Z_{l+1} - \eta\}\bigr) &= \sum_{j\in\mathbb{Z}} E \mathbb{1}_{(b_{j+1}, b_j]}(S_{l+1}) \exp\bigl(\lambda k(-\lambda^{-1} Z_{l+1} - \eta - j)\bigr) \\
&= \sum_{j\in\mathbb{Z}} q^{k(\eta+j)} E \mathbb{1}_{(b_{j+1}, b_j]}(S_{l+1}) S_{l+1}^{-k} \\
&= (q^{-k} - 1) q^{k\eta} \sum_{j\in\mathbb{Z}} q^{kj} E \mathbb{1}_{(b_j, \infty)}(S_{l+1}) S_{l+1}^{-k},
\end{aligned}
$$

where we used partial summation for the last equality. Since

$$
E \mathbb{1}_{(b,\infty)}(S_{l+1}) S_{l+1}^{-k} = \int_b^\infty \frac{y^{l-k}}{l!} e^{-y}\, dy = \frac{(l-k)!}{l!} \left(\sum_{i=0}^{l-k} \frac{b^i}{i!} \right) e^{-b},
$$

we obtain

$$
\begin{aligned}
1 &- E\bigl(1 - q \exp(\lambda\{\lambda^{-1} Z_{l+1} - \eta\})\bigr)^l \\
&= -\sum_{k=1}^{l} \binom{l}{k} (-1)^k q^k E \exp\bigl(\lambda k\{\lambda^{-1} Z_{l+1} - \eta\}\bigr) \\
&= -\sum_{j\in\mathbb{Z}} e^{-q^{j+\eta}} \sum_{k=1}^{l} \sum_{i=0}^{l-k} \frac{(-1)^k}{k!\, i!} (1 - q^k) q^{(j+\eta)(k+i)}.
\end{aligned}
$$

The two inner sums can be rearranged,

$$
\sum_{k=1}^{l} \sum_{i=0}^{l-k} \frac{(-1)^k}{k!\, i!} (1 - q^k) q^{(j+\eta)(k+i)} = \sum_{i=1}^{l} \sum_{k=1}^{i} \frac{(-1)^k}{k!\,(i-k)!} (1 - q^k) q^{(\eta+j)i},
$$

so that

$$
\begin{aligned}
E&\bigl(1 - q \exp(\lambda\{\lambda^{-1} Z_l - \eta\})\bigr)^{l-1} - E\bigl(1 - q \exp(\lambda\{\lambda^{-1} Z_{l+1} - \eta\})\bigr)^l \\
&= -\sum_{j\in\mathbb{Z}} e^{-q^{j+\eta}} \sum_{k=1}^{l} \frac{(-1)^k}{k!(l-k)!} (1 - q^k) q^{(\eta+j)l} \\
&= \frac{(1-q)^l}{l!} \sum_{j\in\mathbb{Z}} q^{(\eta+j)l} e^{-q^{j+\eta}} \\
&= Q_{p,\eta}(\{l\}). \qquad \square
\end{aligned}
$$

After these preparations the proof of Theorem 2 can now be carried out. We prove the tail version, that is,

$$
\sum_{l=1}^{\infty} \gamma^l \left| P(W_n \ge l) - \sum_{k=l}^{\infty} Q_{p,\eta}(\{k\}) \right| = O\left(\frac{1}{n}\right) \qquad \text{as } n \to \infty,
$$

for all $\gamma < 1/p$; the two versions are easily seen to be equivalent.

For $W_n \geq l$ it is necessary and sufficient that $Y_{(n:n-l+1)}$ and $Y_{(n:n)}$ are both contained in the same interval $[x_{k-1}, x_k)$ for some $k \in \mathbb{N}$. With $\lambda = -\log q$ we have $x_k = \lambda k$ for all $k \in \mathbb{N}_0$, hence this leads to

$$P(W_n \geq l) = P\big(Y_{(n:n)} - Y_{(n:n-l+1)} \leq \lambda(1 - \{\lambda^{-1} Y_{(n:n-l+1)}\})\big).$$

Obviously,

$$\begin{aligned}\{\lambda^{-1} Y_{(n:n-l+1)}\} &= \{\lambda^{-1} Y_{(n:n-l+1)} + \lfloor \log_q n \rfloor\} \\ &= \big\{\lambda^{-1} Y_{(n:n-l+1)} + \log_q n - \{\log_q n\}\big\} \\ &= \{\lambda^{-1}(Y_{(n:n-l+1)} - \log n) - \eta_n\}.\end{aligned}$$

Together with Lemma 3 this implies

$$\begin{aligned}P(W_n \geq l) &= \int P\big(Y_{(n:1)} - Y_{(n:n-l+1)} \leq \lambda(1-y)\big) P^{\{\lambda^{-1} Y_{(n:n-l+1)}\}}(dy) \\ &= \int \big(1 - e^{-\lambda(1-y)}\big)^{l-1} P^{\{\lambda^{-1} Y_{(n:n-l+1)}\}}(dy) \\ &= E\big(1 - \exp\big(-\lambda(1 - \{\lambda^{-1} Y_{(n:n-l+1)}\})\big)\big)^{l-1} \\ &= E\big(1 - q\exp\big(\lambda\{\lambda^{-1}(Y_{(n:n-l+1)} - \log n) - \eta_n\}\big)\big)^{l-1}.\end{aligned}$$

As f_{nl} is the density of $Y_{(n:n-l+1)} - \log n$ we can rewrite this as

$$P(W_n \geq l) = \int \big(1 - \psi_n(z)\big)^{l-1} f_{nl}(z)\, dz$$

with ψ_n defined by

$$\psi_n(z) := q \exp\big(\lambda\{\lambda^{-1} z - \eta_n\}\big) \qquad \text{for all } z \in \mathbb{R}.$$

Note that $0 \leq 1 - \psi_n(z) \leq 1 - q$ and that

$$\sum_{k=l}^{\infty} Q_{p,\eta}(\{k\}) = \int \big(1 - \psi_n(z)\big)^{l-1} g_l(z)\, dz$$

by Lemma 5. Let $l_n := C_1 \log n$ with $C_1 := -1/\log(\gamma p)$. Using Lemma 4 we obtain

$$\begin{aligned}&\sum_{l \leq l_n} \gamma^l \left| P(W_n \geq l) - \sum_{k=l}^{\infty} Q_{p,\eta}(\{k\}) \right| \\ &\qquad \leq \sum_{l \leq l_n} \gamma^l \int \big(1 - \psi_n(z)\big)^{l-1} \big| f_{nl}(z) - g_l(z) \big|\, dz \\ &\qquad \leq \sum_{l \leq l_n} \gamma^l p^l \int |f_{nl}(z) - g_l(z)|\, dz \leq \sum_{l \leq l_n} \gamma^l p^l C_2 \frac{l^2}{n} = O\Big(\frac{1}{n}\Big).\end{aligned}$$

The proof of Theorem 2 will therefore be complete once we have shown that

$$\sum_{l>l_n} \gamma^l P(W_n \geq l) = O\left(\frac{1}{n}\right), \qquad \sum_{l>l_n} \gamma^l \sum_{k=l}^{\infty} Q_{p,\eta}(\{k\}) = O\left(\frac{1}{n}\right).$$

This, however, is obvious from the above representation of the tails as integrals of $(1-\psi_n)^{l-1}$, which is $O(p^l)$ uniformly in $n \in \mathbb{N}$, and the definition of l_n.

REFERENCES

ATHREYA, J. S. and SETHURAMAN, S. (2001). On the asymptotics of discrete order statistics. *Statist. Probab. Lett.* **54** 243–249.

BARYSHNIKOV, Y., EISENBERG, B. and STENGLE, G. (1995). A necessary and sufficient condition for the existence of the limiting probability of a tie for first place. *Statist. Probab. Lett.* **23** 203–209.

BRANDS, J. J. A. M., STEUTEL, F. W. and WILMS, R. J. G. (1994). On the number of maxima in a discrete sample. *Statist. Probab. Lett.* **20** 209–218.

BRUSS, F. T. and O'CINNEIDE, C. A. (1990). On the maximum and its uniqueness for geometric random samples. *J. Appl. Probab.* **27** 598–610.

DAVID, H. A. (1981). *Order Statistics*, 2nd ed. Wiley, New York.

EISENBERG, B., STENGLE, G. and STRANG, G. (1993). The asymptotic probability of a tie for first place. *Ann. Appl. Probab.* **3** 731–745.

FILL, J. A., MAHMOUD, H. M. and SZPANKOWSKI, W. (1996). On the distribution for the duration of a randomized leader election algorithm. *Ann. Appl. Probab.* **6** 1260–1283.

GRÜBEL, R. and REIMERS, A. (2001a). On the number of iterations required by von Neumann addition. *Theor. Inform. Appl.* **35** 187–206.

GRÜBEL, R. and REIMERS, A. (2001b). On the total time spent in records by a discrete uniform sequence. *J. Appl. Probab.* **38** 768–775.

KIRSCHENHOFER, P. and PRODINGER, H. (1996). The number of winners in a discrete geometrically distributed sample. *Ann. Appl. Probab.* **6** 687–694.

KIRSCHENHOFER, P. and PRODINGER, H. (1998). Addendum. The number of winners in a discrete geometrically distributed sample. *Ann. Appl. Probab.* **8** 647.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

DÉPARTEMENT DE MATHÉMATIQUE ET ISRO
UNIVERSITÉ LIBRE DE BRUXELLES
CP 210, BOULEVARD DU TRIOMPHE
B-1050 BRUXELLES
BELGIUM
E-MAIL: tbruss@ulb.ac.be

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT HANNOVER
POSTFACH 60 09
D-30060 HANNOVER
GERMANY
E-MAIL: rgrubel@stochastik.uni-hannover.de