Elect. Comm. in Probab. **12** (2007), 93–105

ELECTRONIC COMMUNICATIONS in PROBABILITY

MAXIMA OF THE CELLS OF AN EQUIPROBABLE MULTINOMIAL

ARUP BOSE
Stat-Math Unit, Indian Statistical Institute, Kolkata 700108 INDIA
email: abose@isical.ac.in

AMITES DASGUPTA
Stat-Math Unit, Indian Statistical Institute, Kolkata 700108 INDIA
email: amites@isical.ac.in

KRISHANU MAULIK
Stat-Math Unit, Indian Statistical Institute, Kolkata 700108 INDIA
email: krishanu@isical.ac.in

Submitted September 27, 2006, *accepted in final form* April 4, 2007

AMS 2000 Subject classification: Primary 60G70, 60F05; Secondary 60F10
Keywords: Random sequences, triangular array, maxima, limit distribution

Abstract

Consider a sequence of multinomial random vectors with increasing number of equiprobable cells. We show that if the number of trials increases fast enough, the sequence of maxima of the cells after a suitable centering and scaling converges to the Gumbel distribution. While results are available for maxima of triangular arrays of independent random variables with certain types of distribution, such results in a dependent setup is new. We also prove that the maxima of a triangular sequence of appropriate Binomial random variables have the same limit distribution. An auxiliary large deviation result for multinomial distribution with increasing number of equiprobable cells may also be of independent interest.

1 Introduction and main result

Let $(Y_{1n}, \ldots, Y_{m_n n})_{n\geq 1}$ be a triangular sequence of random variables. Define the row maximum as $M_n = \max\{Y_{1n}, \ldots, Y_{m_n n}\}$. The question of convergence in distribution of M_n with linear normalization has been addressed under a variety of conditions.

The classical case is when there is one sequence of i.i.d. random variables $\{Y_i\}$ and $M_n = \max\{Y_1, \ldots, Y_n\}$. In this case, necessary and sufficient conditions for the convergence are known. See for example, Fisher and Tippett (1928), Gnedenko (1943), de Haan (1970). In particular, it follows from these results that if $\{Y_i\}$ are i.i.d. Poisson or i.i.d. binomial with fixed parameters, then M_n cannot converge to any non degenerate distribution under any linear normalization (cf. Leadbetter et al., 1983, pp 24–27). On the other hand (cf. Leadbetter et al., 1983, Theorem 1.5.3), if Y_i are i.i.d. standard normal variables then

$$\lim_{n\to\infty} P[M_n \le \alpha_n x + \beta_n] = \exp(-e^{-x}),$$

where

$$\alpha_n = \frac{1}{\sqrt{2\log n}} \tag{1.1}$$

and

$$\beta_n = \sqrt{2\log n} - \frac{\log\log n + \log(4\pi)}{2\sqrt{2\log n}}. \tag{1.2}$$

General triangular schemes under various suitable conditions have been considered by several authors. The classical large deviation results due to Cramér (cf. Petrov, 1975, pg 218) play an important role in the proofs of these results.

Consider, for example, the case where $Y_{m_n n} = \big(\sum_{1\le j\le m_n} U_j - m_n\mu\big)/(\sigma m_n^{1/2})$ and U_j are i.i.d. with mean μ and standard deviation σ. Assuming that U_j has a finite moment generating function in an open interval containing the origin and $\log n = o\big(m_n^{(R+1)/(R+3)}\big)$ for some integer $R \ge 0$, Anderson et al. (1997) showed that

$$\lim_{n\to\infty} P[M_n \le \alpha_n x + \beta_n^{(R)}] = \exp(-e^{-x})$$

for α_n as in (1.1) and some suitable sequences $\beta_n^{(R)}$.

They also consider the following case. Suppose $m_n = n$ and for each n, $Y_{m_n n}$, are independent Poisson with mean λ_n such that for some integer $R \ge 0$, $\log n = o(\lambda_n^{(R+1)/(R+3)})$. Then again

$$\lim_{n\to\infty} P[M_n \le \lambda_n + \lambda_n^{1/2}(\beta_n^{(R)} + \alpha_n x)] = \exp(-e^{-x}),$$

where α_n and $\beta_n^{(R)}$ are as before. In particular, in the above results, if $R = 0$ then we can choose α_n as in (1.1) and $\beta_n^{(0)} = \beta_n$, given by (1.2).

Nadarajah and Mitov (2002) consider the maximum of a triangular array of binomial, negative binomial and discrete uniform. The case of binomial triangular array is discussed with increasing number of trials m_n and fixed probability of success, p. The idea of the proof in this case is again similar to that of Anderson et al. (1997) and uses a large deviation result for binomial distribution.

In this paper we consider the following dependent situation. Suppose $\boldsymbol{Y}_n = (Y_{1n}, \cdots, Y_{nn})$ follow multinomial $(m_n; 1/n, \ldots 1/n)$ distribution and define $M_n = \max_{1\le i\le n} Y_{in}$ to be the maximum of the n cell variables. If m_n tends to infinity fast enough, then the sequence M_n after a suitable linear normalization, converges to the Gumbel distribution. We summarize this result in the following theorem:

Theorem 1.1. *Suppose that $\boldsymbol{Y}_n$ is distributed as multinomial $(m_n; \frac{1}{n}, \ldots, \frac{1}{n})$ and define as before $M_n = \max_{1\le i\le n} Y_{in}$. If*

$$\lim_{n\to\infty} \frac{\log n}{m_n/n} = 0 \tag{1.3}$$

holds, then, for $x \in \mathbb{R}$,

$$P\left[\frac{M_n - (m_n/n) - \beta_n\sqrt{m_n/n}}{\alpha_n\sqrt{m_n/n}} \le x\right] \to \exp(-e^{-x}), \tag{1.4}$$

where α_n is as in (1.1) *and β_n is the unique solution of*

$$\log z + \frac{1}{2}z^2 + \frac{1}{2}\log(2\pi) - z^2 B\left(\frac{z}{\sqrt{m_n/n}}\right) = \log n \tag{1.5}$$

in the region $\beta_n \sim \sqrt{2\log n}$, where

$$B(z) = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{(i+1)(i+2)} z^i. \tag{1.6}$$

A similar result for the maximum of i.i.d. Binomial random variables is given below. Unlike Theorem 3 of Nadarajah and Mitov (2002), we do not require the probability of success to be constant.

Proposition 1.1. *Let $\{Y_{in} : 1 \le i \le n, n \ge 1\}$ be a triangular array of independent Binomial random variables, with $\{Y_{in} : 1 \le i \le n\}$ having i.i.d. Binomial $(m_n; p_n)$ distribution for each $n \ge 1$. Define $M_n = \max_{1\le i\le n} Y_{in}$. If we have*

$$\lim_{n\to\infty} \frac{\log n}{m_n p_n} = 0 \qquad \textit{and} \qquad \lim_{n\to\infty} p_n (\log n)^r = 0,$$

for all $r > 0$, then we have

$$P\left[\frac{M_n - (m_n p_n) - \beta_n\sqrt{m_n p_n}}{\alpha_n\sqrt{m_n p_n}} \le x\right] \to \exp(-e^{-x}),$$

where α_n and β_n are chosen as in Theorem 1.1.

The large deviation results used by Anderson et al. (1997) or Nadarajah and Mitov (2002) are not directly applicable in our case. For our case, even though the random variables in each row of the array can be written as sum of independent variables, the distributions of the summands depend on the row. Our proof of the theorem is based on the following large deviation result. As we are unable to locate this particular large deviation result in the existing literature, we provide a detailed derivation in the next section.

Theorem 1.2. *Suppose $\boldsymbol{Y}_n$ be distributed as multinomial $(m_n; \frac{1}{n}, \ldots, \frac{1}{n})$, such that the condition* (1.3)

$$\lim_{n\to\infty} \frac{\log n}{m_n/n} = 0$$

holds. For any positive integer k and any sequence

$$v_n \sim \sqrt{2\log n}, \tag{1.7}$$

we have

$$P\left[\frac{\min_{1\le i\le k} Y_{in} - m_n/n}{\sqrt{m_n/n}} > v_n\right] \sim \left[(1-\Phi(v_n))\exp\left(v_n^2 B\left(\frac{v_n}{\sqrt{m_n/n}}\right)\right)\right]^k, \tag{1.8}$$

where $B(z)$ is given by (1.6) *and Φ is the univariate standard normal distribution function.*

2 Proofs

For a real number x, denote

$$y_n = x\alpha_n\sqrt{m_n/n} + \beta_n\sqrt{m_n/n} + (m_n/n). \tag{2.1}$$

We prove Theorem 1.1 using the following lemma.

Lemma 2.1. *For each fixed k, and real number x, we have*

$$n^k P(\cap_{i=1}^k \{Y_{in} > y_n\}) \to e^{-kx}. \tag{2.2}$$

where y_n and x are related as in (2.1).

Proof of Theorem 1.1. For any fixed l, for sufficiently large n, using inclusion-exclusion principle and the identical distribution of the marginals from the multinomial distribution, we have,

$$\begin{aligned}
&1 - \sum_{k=1}^{2l-1}(-1)^{k+1}\frac{n(n-1)\cdots(n-k+1)}{k!}P(\cap_{i=1}^k\{Y_{in} > y_n\}) \\
\le& P(\cap_{i=1}^n\{Y_{in} \le y_n\}) \\
\le& 1 - \sum_{k=1}^{2l}(-1)^{k+1}\frac{n(n-1)\cdots(n-k+1)}{k!}P(\cap_{i=1}^k\{Y_{in} > y_n\}).
\end{aligned} \tag{2.3}$$

Hence using Lemma 2.1, we obtain from (2.2) and (2.3), for each fixed l,

$$\begin{aligned}
1 - \sum_{k=1}^{2l-1}(-1)^{k+1}\frac{e^{-kx}}{k!} &\le \liminf_{n\to\infty} P(\cap_{i=1}^n\{Y_{in} \le y_n\}) \\
&\le \limsup_{n\to\infty} P(\cap_{i=1}^n\{Y_{in} \le y_n\}) \le 1 - \sum_{k=1}^{2l}(-1)^{k+1}\frac{e^{-kx}}{k!},
\end{aligned}$$

which gives the desired result (1.4) since l is arbitrary. □

Remark 2.1. As pointed out by the referee, it can be easily seen, using negative dependence, that $P(Y_{1n} \le y_n)^n$ is another choice of the upper bound in (2.3) and, hence,

$$\limsup_{n\to\infty} P(\cap_{i=1}^n\{Y_{in} \le y_n\}) \le \lim_{n\to\infty} \exp\left(-nP(Y_{1n} > y_n)\right) = \exp(-e^{-x}).$$

However, there appears to be no easy way to obtain an appropriate lower bound from the existing literature.

Now we prove Lemma 2.1 using Theorem 1.2.

Proof of Lemma 2.1. Modifying Lemmas 1 and 2 of Anderson et al. (1997), we can find α_n and β_n, so that, for $x_n = \alpha_n x + \beta_n$, we have

$$\log x_n + \frac{1}{2}\log(2\pi) + \frac{1}{2}x_n^2 - x_n^2 B\left(\frac{x_n}{\sqrt{m_n/n}}\right) - \log n \to x. \tag{2.4}$$

Note that the referred lemmas require a polynomial instead of a power series in the defining equation. However, the proofs work verbatim in our case due to the specific form of the coefficients. Also using (16) and (17) of the same reference, we have $\alpha_n \sim 1/\sqrt{2\log n}$ and β_n is the unique solution of (1.5) satisfying $\beta_n \sim \sqrt{2\log n}$. Note that

$$x_n = \frac{y_n - m_n/n}{\sqrt{m_n/n}} = \alpha_n x + \beta_n \sim \sqrt{2\log n}.$$

Thus, using (1.8), we have,

$$\begin{aligned} n^k P\left[\cap_{i=1}^k Y_{in} > y_n\right] &= n^k P\left[\frac{\min_{1\le i\le k} Y_{in} - m_n/n}{\sqrt{m_n/n}} > x_n\right] \\ &\sim n^k \left[(1-\Phi(x_n)) \exp\left(x_n^2 B\left(\frac{x_n}{\sqrt{m_n/n}}\right)\right)\right]^k. \end{aligned}$$

Hence, using

$$1 - \Phi(t) \sim \exp(-t^2/2)/(t\sqrt{2\pi}) \text{ as } t \to \infty, \tag{2.5}$$

(cf. Feller, 1968, Lemma 2, Chapter VII), we have,

$$n^k P\left[\cap_{i=1}^k Y_{in} > y_n\right] \sim e^{-k\left[\log x_n + \frac{1}{2}\log(2\pi) + \frac{1}{2}x_n^2 - x_n^2 B\left(\frac{x_n}{\sqrt{m_n/n}}\right) - \log n\right]} \to e^{-kx}.$$

The last step follows from (2.4). □

Now we prove the large deviation result given in Theorem 1.2.

Proof of Theorem 1.2. Let us consider a random vector $(Z_0, Z_1, \ldots, Z_k)$, which has multinomial $(1; \frac{n-k}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$ distribution. Denote by F_n the distribution of $(Z_1 - \frac{1}{n}, \cdots, Z_k - \frac{1}{n})$. Note that F_n has mean vector $\mathbf{0}$ and its covariance matrix is given by $((a_{ij}))$, $a_{ii} = 1/n - 1/n^2$, $a_{ij} = -1/n^2, i \neq j$. Let $\boldsymbol{U}_n^{(i)} = (U_{1n}^{(i)}, \ldots, U_{kn}^{(i)})$, $1 \le i \le m_n$, be i.i.d. F_n.
Define $\boldsymbol{X}_n = (X_{1n}, \cdots, X_{kn}) = \sum_{i=1}^{m_n} \boldsymbol{U}_n^{(i)}$. We apply Esscher transform or exponential tilting on the distribution of $\boldsymbol{X}_n$. Let $\Psi_n(t_1, \ldots, t_k)$ be the cumulant generating function of F_n:

$$\Psi_n(t_1, \ldots, t_k) = -\frac{t_1 + \cdots + t_k}{n} + \log\left(1 + \frac{e^{t_1} + \cdots + e^{t_k} - k}{n}\right). \tag{2.6}$$

Let s_n be the unique solution of

$$m_n \partial_1 \Psi_n(s, \ldots, s) = v_n \sqrt{m_n/n}. \tag{2.7}$$

Next we define the exponential tilting for the multivariate case as

$$dV_n(w_1, \ldots, w_k) = e^{-\Psi_n(s_n, \ldots, s_n)} e^{s_n(w_1 + \cdots + w_k)} dF_n(w_1, \ldots, w_k). \tag{2.8}$$

Then, the m_n-th convolution power of V_n is given by

$$dV_n^{\star m_n}(w_1, \ldots, w_k) = e^{-m_n \Psi_n(s_n, \cdots, s_k)} e^{s_n(w_1 + \cdots + w_k)} dF_n^{\star m_n}(w_1, \ldots, w_k).$$

Denote

$$u_n = e^{s_n} - 1. \tag{2.9}$$

Note that V_n has mean vector $\mu_n \mathbf{1}_k$ and covariance matrix $\Sigma_n = a_n I_k - b_n J_k$, where $\mathbf{1}_k$ is the k-vector with all coordinates 1, I_k is the $k \times k$ identity matrix, J_k is the $k \times k$ matrix with all entries 1 and μ_n, a_n and b_n are given as follows:

$$\begin{aligned}
\mu_n &= \partial_1 \Psi_n(s_n, \ldots, s_n) = -\frac{1}{n} + \frac{e^{s_n}}{n + k(e^{s_n} - 1)} \\
&= \frac{(n-k)(e^{s_n}-1)}{n(n+k(e^{s_n}-1))} = \frac{(n-k)u_n}{n(n+ku_n)}, && (2.10) \\
b_n &= -\partial_1 \partial_2 \Psi_n(s_n, \ldots, s_n) = \frac{e^{2s_n}}{(n+k(e^{s_n}-1))^2} \\
&= \left(\frac{1+u_n}{n+ku_n} \right)^2, && (2.11) \\
\tau_n^2 := a_n - b_n &= \partial_1^2 \Psi_n(s_n, \ldots, s_n) = \frac{e^{s_n}(n-k+(k-1)e^{s_n})}{(n+k(e^{s_n}-1))^2} \\
&= \frac{(1+u_n)(n-k+(k-1)(1+u_n))}{(n+ku_n)^2}, && (2.12)
\end{aligned}$$

and we also denote

$$\gamma_n = \Psi_n(s_n, \ldots, s_n). \tag{2.13}$$

With notations as above, the required probability becomes

$$\begin{aligned}
P_n &= P\left[\frac{\min_{1 \le i \le k} Y_{in} - m_n/n}{\sqrt{m_n/n}} > v_n \right] \\
&= P[X_{1n} > v_n \sqrt{m_n/n}, \ldots, X_{kn} > v_n \sqrt{m_n/n}] \\
&= \int_{v_n \sqrt{m_n/n}}^{\infty} \cdots \int_{v_n \sqrt{m_n/n}}^{\infty} dF_n^{\star m_n}(w_1, \ldots, w_k) \\
&= e^{m_n \gamma_n} \int_{m_n \mu_n}^{\infty} \cdots \int_{m_n \mu_n}^{\infty} e^{-s_n(w_1 + \cdots + w_k)} dV_n^{\star m_n}(w_1, \ldots, w_k). && (2.14)
\end{aligned}$$

Now we replace V_n by a k-variate normal with mean vector $\mu_n \mathbf{1}_k$ and covariance matrix $\tau_n^2 I_k$ (i.e., independent coordinates). The result of this change of distribution leads to the approximation (for P_n), given by

$$\begin{aligned}
A_{s_n} =& e^{m_n \gamma_n} \left[\int_{m_n \mu_n}^{\infty} e^{-s_n y} \phi\left(\frac{y - m_n \mu_n}{\tau_n \sqrt{m_n}} \right) \frac{dy}{\tau_n \sqrt{m_n}} \right]^k && (2.15) \\
=& e^{m_n \gamma_n} \left[\int_0^{\infty} \phi(z) e^{-s_n(m_n \mu_n + z \tau_n \sqrt{m_n})} dz \right]^k \\
=& e^{m_n(\gamma_n - k s_n \mu_n)} \rho^k(s_n \tau_n \sqrt{m_n}), && (2.16)
\end{aligned}$$

where $\rho(t) = \int_0^\infty e^{-zt} \phi(z) dz = e^{\frac{t^2}{2}}(1 - \Phi(t))$ and ϕ and Φ are the univariate standard normal density.

Then the proof follows combining the two following propositions. □

The first proposition shows that A_{s_n} has the correct asymptotic behavior and the second one shows that A_{s_n} is a good approximation for P_n.

Proposition 2.1. *If v_n satisfies* (1.7)*, namely, $v_n \sim \sqrt{2\log n}$ and m_n satisfies* (1.3) *given by $\log n = o(m_n/n)$, then*

$$A_{s_n} \sim \left[(1-\Phi(v_n))\exp\left(v_n^2 B\left(\frac{v_n}{\sqrt{m_n/n}}\right)\right)\right]^k. \tag{2.17}$$

Proposition 2.2. *If v_n satisfies* (1.7)*, namely, $v_n \sim \sqrt{2\log n}$ and m_n satisfies* (1.3) *given by $\log n = o(m_n/n)$, then*

$$P_n \sim A_{s_n}. \tag{2.18}$$

The proofs of Propositions 2.1 and 2.2 depend on the following lemma about the rate of growth of u_n, where u_n is defined in (2.9).

Lemma 2.2. *If v_n satisfies* (1.7) *given by $v_n \sim \sqrt{2\log n}$ and if m_n satisfies* (1.3) *given by $\log n = o(m_n/n)$, we have*

$$u_n = \frac{v_n}{\sqrt{m_n/n}}\left(1 + O\left(\frac{1}{n}\right)\right) \to 0. \tag{2.19}$$

Proof. Note that the first partial of Ψ_n is

$$\partial_1 \Psi_n(t_1,\dots,t_k) = -\frac{1}{n} + \frac{e^{t_1}}{e^{t_1}+\cdots+e^{t_k}+n-k}.$$

Hence, using (2.7), we have

$$\frac{v_n}{\sqrt{m_n/n}} = \frac{(n-k)(e^{s_n}-1)}{n+k(e^{s_n}-1)} = \frac{(n-k)u_n}{n+ku_n}. \tag{2.20}$$

Solving, we get,

$$u_n = \left(1-\frac{k}{n}\right)^{-1}\left(1-\frac{k}{n-k}\frac{v_n}{\sqrt{m_n/n}}\right)^{-1}\frac{v_n}{\sqrt{m_n/n}}$$

and, the result follows using $\frac{v_n}{\sqrt{m_n/n}} \sim \sqrt{\frac{2\log n}{m_n/n}} \to 0$, from (1.3) and (1.7). □

The following corollary regarding the asymptotic behavior of μ_n, b_n and τ_n^2 then follows immediately from (2.10)–(2.12).

Corollary 2.1. *Assume that v_n satisfies* (1.7) *given by $v_n \sim \sqrt{2\log n}$ and m_n satisfies* (1.3) *given by $\log n = o(m_n/n)$. If the tilted distribution V_n has mean vector $\mu_n \mathbf{1}_k$ and covariance matrix $\Sigma_n = a_n I_k - b_n J_k$, then*

$$\mu_n \sim \frac{u_n}{n}, \quad b_n \sim \frac{1}{n^2}, \quad \textit{and} \quad \tau_n^2 \sim \frac{1}{n}. \tag{2.21}$$

Now we prove Proposition 2.1 on the asymptotic behavior of A_{s_n}.

Proof of Proposition 2.1. We first treat the exponent in the first factor of the expression (2.16) for A_{s_n}. Since $\gamma_n = -\frac{k}{n}s_n + \log(1+\frac{k}{n}(e^{s_n}-1)) = -\frac{k}{n}\log(1+u_n) + \log(1+\frac{k}{n}u_n)$ using (2.6), it follows from expression (2.10) for μ_n,

$$\begin{aligned}
&m_n(\gamma_n - ks_n\mu_n)\\
=&\frac{m_n}{n} n\log\left(1+\frac{k}{n}u_n\right) - \frac{m_n}{n}\frac{k(1+u_n)\log(1+u_n)}{1+\frac{k}{n}u_n}\\
=&\frac{m_n}{n}\left(1+\frac{k}{n}u_n\right)^{-1}\left[(n+ku_n)\log\left(1+\frac{k}{n}u_n\right) - (k+ku_n)\log(1+u_n)\right]\\
=&k\frac{m_n}{n}\left(1+\frac{k}{n}u_n\right)^{-1}\sum_{r=2}^{\infty}\frac{(-1)^{r-1}}{r(r-1)}\left[1-\left(\frac{k}{n}\right)^{r-1}\right]u_n^r\\
=&-\frac{k}{2}\frac{m_n u_n^2}{n} + \frac{k^2}{2n}\frac{m_n u_n^2}{n}\\
&\quad - k\frac{m_n u_n^2}{n}\sum_{i=1}^{\infty}(-1)^i\sum_{r=0}^{i}\frac{1}{(r+1)(r+2)}\left(\frac{k}{n}\right)^{i-r}\left[1-\left(\frac{k}{n}\right)^{r+1}\right]u_n^i\\
=&-\frac{k}{2}\frac{m_n u_n^2}{n} + k\frac{m_n u_n^2}{n}B(u_n) + E_n^{(0)} + E_n^{(1)},
\end{aligned} \tag{2.22}$$

where, using (2.19),

$$E_n^{(0)} = \frac{k^2}{2n}\frac{m_n u_n^2}{n} \sim k^2\frac{\log n}{n} \to 0 \tag{2.23}$$

and

$$\begin{aligned}
E_n^{(1)} = \frac{m_n u_n^2}{n}\sum_{i=1}^{\infty}(-1)^i u_n^i &\left[\sum_{r=0}^{i-1}\frac{1}{(r+1)(r+2)}\left(\frac{k}{n}\right)^{i-r}\left\{1-\left(\frac{k}{n}\right)^{r+1}\right\}\right.\\
&\left.\quad - \frac{1}{(i+1)(i+2)}\left(\frac{k}{n}\right)^{i+1}\right].
\end{aligned}$$

Thus, $\left|E_n^{(1)}\right| \le S_1 + S_2$, where

$$\begin{aligned}
S_1 =&\frac{m_n u_n^2}{n}\sum_{i=1}^{\infty}u_n^i\sum_{r=0}^{i-1}\frac{1}{(r+1)(r+2)}\left(\frac{k}{n}\right)^{i-r}\left\{1-\left(\frac{k}{n}\right)^{r+1}\right\}\\
\le&\frac{m_n u_n^2}{n}\sum_{i=0}^{\infty}\left(\frac{k}{n}u_n\right)^{i+1}\sum_{r=0}^{i}\frac{1}{(r+1)(r+2)}\left(\frac{k}{n}\right)^{-r}\left\{1-\left(\frac{k}{n}\right)^{r+1}\right\}\\
=&\frac{m_n u_n^2}{n}\sum_{r=0}^{\infty}\frac{1}{(r+1)(r+2)}\frac{k}{n}\left\{1-\left(\frac{k}{n}\right)^{r+1}\right\}u_n^{r+1}\left(1-\frac{k}{n}u_n\right)^{-1}\\
\sim&2k\frac{\log n}{n}u_n\sum_{r=0}^{\infty}\frac{1}{(r+1)(r+2)}\left\{1-\left(\frac{k}{n}\right)^{r+1}\right\}u_n^r \to 0,
\end{aligned} \tag{2.24}$$

and

$$S_2 = \frac{m_n u_n^2}{n}\sum_{i=1}^{\infty}\frac{1}{(i+1)(i+2)}\left(\frac{k}{n}u_n\right)^i \le \frac{m_n u_n^2}{n}\sum_{i=1}^{\infty}\left(\frac{k}{n}u_n\right)^i \sim 2\log n\frac{k}{n}u_n \to 0, \tag{2.25}$$

since $m_n u_n^2/n \sim v_n^2 \sim 2\log n$, using (2.19) and (1.3). Hence, we have

$$E_n^{(1)} \to 0. \tag{2.26}$$

Thus, using (2.22), (2.23) and (2.26), we have

$$m_n(\gamma_n - ks_n\mu_n) = -\frac{k}{2}\frac{m_n u_n^2}{n} + k\frac{m_n u_n^2}{n}B(u_n) + o(1)$$

and hence, we have,

$$e^{m_n(\gamma_n - ks_n\mu_n)} \sim \exp\left(-\frac{k}{2}\frac{m_n u_n^2}{n} + k\frac{m_n u_n^2}{n}B(u_n)\right). \tag{2.27}$$

Further, observe from (2.19) that, $m_n u_n^2/n - v_n^2 = O(v_n^2/n) = O(\log n/n) \to 0$, using (1.7). Also, $B(u_n) \sim u_n/6 \to 0$ and

$$B(u_n) - B\left(\frac{v_n}{\sqrt{m_n/n}}\right) = O\left(u_n - \frac{v_n}{\sqrt{m_n/n}}\right) = O\left(\frac{v_n}{\sqrt{m_n/n}}\frac{1}{n}\right) = o(1/v_n^2),$$

using (2.19), (1.3) and (1.7). Hence, $m_n u_n^2/n$ in (2.27) can be replaced by v_n giving

$$e^{m_n(\gamma_n - ks_n\mu_n)} \sim \exp\left(-\frac{k}{2}v_n^2 + ku_n^2 B\left(\frac{v_n}{\sqrt{m_n/n}}\right)\right). \tag{2.28}$$

Using the asymptotic expression (2.21) for τ_n, the fact $u_n = e^{s_n} - 1 \sim s_n$ and (2.19), we have

$$\tau_n s_n \sqrt{m_n} \sim u_n \sqrt{m_n/n} = v_n.$$

The proof is then completed using (2.5), which gives

$$\rho^k(s_n\tau_n\sqrt{m_n}) \sim \frac{1}{(v_n\sqrt{2\pi})^k}. \tag{2.29}$$

□

Next we prove Proposition 2.2.

Proof of Proposition 2.2. Let $\Phi_{\boldsymbol{\mu},A}$ denote the k-variate normal distribution function with mean vector $\boldsymbol{\mu}$ and covariance matrix A. Using (2.14) and (2.15), we easily see that

$$\frac{P_n - A_{s_n}}{e^{m_n\gamma_n}} = \int_{m_n\mu_n}^{\infty} \cdots \int_{m_n\mu_n}^{\infty} e^{-s_n(u_1+\cdots+u_k)} d(V_n^{\star m_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \tau_n^2 I_k})(u_1, \ldots, u_k).$$

Denote the distribution function of the signed measure $V_n^{\star m_n} - \Phi^{\star m_n}_{\mu \mathbf{1}_k, \tau_n^2 I_k}$ by H_n. Then, using Theorem 3.1 in Appendix and (2.16), we have

$$|P_n - A_{s_n}| \le 2^k \|H_n\|_\infty e^{m_n(\gamma_n - ks_n\mu_n)} = 2^k A_{s_n}\rho^{-k}(s_n\tau_n\sqrt{m_n})\|H_n\|_\infty, \tag{2.30}$$

where $\|H_n\|_\infty$ is the sup norm. Hence, using (2.29) and the fact that $z_n = \sqrt{m_n/n}u_n \sim v_n$, using (2.19), we have

$$\frac{P_n}{A_{s_n}} = 1 + O\left(v_n^k \|H_n\|_\infty\right). \tag{2.31}$$

So, to complete the proof, we need to study $\|H_n\|_\infty$. We write H_n as sum of two signed measures by introducing the normal distribution with covariance matrix, Σ_n, same as that of V_n:

$$H_n = \left(\Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \tau_n^2 I_k}\right) + \left(V_n^{\star m_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n}\right). \tag{2.32}$$

We estimate the first part directly and the second part by Berry-Esseen theorem.
Observe that

$$\|\Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \tau_n^2 I_k}\|_\infty = \|\Phi_{\mathbf{0}, \tau_n^{-2}\Sigma_n} - \Phi_{\mathbf{0}, I_k}\|_\infty,$$

which is estimated easily using normal comparison lemma, attributed to Slepian, Berman and others (see, for example, Leadbetter et al., 1983, Theorem 4.2.1). Observe that $\tau_n^{-2}\Sigma_n = \frac{a_n}{a_n - b_n} I_k - \frac{b_n}{a_n - b_n} J_k$, using (2.11) and (2.12). Hence, from normal comparison lemma and asymptotic behavior of a_n and b_n in (2.21), we have

$$\begin{aligned}\|\Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \tau_n^2 I_k}\|_\infty &= \|\Phi_{\mathbf{0}, \tau_n^{-2}\Sigma_n} - \Phi_{\mathbf{0}, I_k}\|_\infty \\ &\le \frac{1}{2\pi} \frac{k(k-1)}{2} \frac{b_n}{\sqrt{a_n(a_n - 2b_n)}} \sim O(1/n).\end{aligned}$$

Hence, corresponding to the first term of (2.32), we have, using (1.7),

$$v_n^k \|\Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \tau_n^2 I_k}\|_\infty = O(v_n^k/n) \to 0. \tag{2.33}$$

Next we study the second term of (2.32). Suppose $\boldsymbol{\xi}_j$ are i.i.d. V_n with mean $\mu_n \mathbf{1}_k$, covariance Σ_n. Then

$$\begin{aligned}&V_n^{\star m_n}(u_1, \ldots, u_k) - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n}(u_1, \ldots, u_k)) \\ &\qquad = P\left[\frac{1}{\sqrt{m_n}} \sum_{j=1}^{m_n} (\boldsymbol{\xi}_j - \mu_n \mathbf{1}_k) \le \frac{\boldsymbol{u} - m_n \mu_n \mathbf{1}_k}{\sqrt{m_n}}\right] - \Phi_{\mathbf{0}, \Sigma_n}\left(\frac{\boldsymbol{u} - m_n \mu_n \mathbf{1}_k}{\sqrt{m_n}}\right),\end{aligned}$$

and hence, by multivariate Berry-Esseen theorem, (see, e.g., Bhattacharya and Ranga Rao, 1976, Corollary 17.2, pg. 165)

$$\begin{aligned}\|V_n^{\star m_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n}\|_\infty &= \sup_{\boldsymbol{u}} \left| P\left[\frac{1}{\sqrt{m_n}} \sum_{j=1}^{m_n} (\boldsymbol{\xi}_j - \mu_n \mathbf{1}_k) \le \boldsymbol{u}\right] - \Phi_{0, \Sigma_n}(\boldsymbol{u})\right| \\ &\le \frac{C_3}{\sqrt{m_n}} \frac{\kappa_n}{\lambda_n^{3/2}},\end{aligned} \tag{2.34}$$

where $\kappa_n = E\|\boldsymbol{\xi}_1 - \mu_n \mathbf{1}_k\|_2^3$, (the norm being Euclidean one),

$$\lambda_n = a_n - k b_n \sim \frac{1}{n}, \tag{2.35}$$

by (2.11) and (2.12), is the smallest eigenvalue of $\Sigma_n = a_n I - b_n J$, and C_3 is a universal constant. So, to complete the proof we need to estimate κ_n. Using the definition of V_n, (2.8), we have,

$$\kappa_n = e^{-m_n \gamma_n} \int \cdots \int e^{s_n(u_1 + \cdots + u_k)} \left(\sum_{j=1}^k (u_j - \mu_n)^2\right)^{3/2} dF_n(u_1, \ldots, u_k).$$

Recall that F_n is the distribution of the last k coordinates of the centered multinomial $(1; (n-k)/n, 1/n, \ldots, 1/n)$ distribution, which puts mass $1/n$ at each of the k vectors which have all coordinates $-1/n$ except the ith one being $(n-1)/n$, for $i = 1, \ldots, k$, and $(n-k)/n$ at $(-1/n, \ldots, -1/n)$. Thus,

$$e^{m_n\gamma_n}\kappa_n = \frac{n-k}{n} e^{-\frac{ks_n}{n}} k^{\frac{3}{2}} \Big(\frac{1}{n} + \mu_n\Big)^3 + \frac{k}{n} e^{\frac{(n-k)s_n}{n}} \left[(k-1)\Big(\frac{1}{n} + \mu_n\Big)^2 + \Big(1 - \frac{1}{n} - \mu_n\Big)^2\right]^{\frac{3}{2}}.$$

Since, from (2.21) we have $\mu_n \sim \frac{u_n}{n}$, and by (2.19), we have $s_n = \log(1+u_n) \sim u_n \to 0$,

$$\kappa_n \sim e^{-m_n\gamma_n}\frac{k}{n}.$$

Thus, using (2.34) and (2.35), we have,

$$v_n^k \|V_n^{\star m_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n}\|_\infty \le kC_3 \frac{v_n^k}{\sqrt{\frac{m_n}{n}}} e^{-m_n\gamma_n}. \tag{2.36}$$

Also, from (2.6), we get, for fixed k,

$$\begin{aligned} m_n\gamma_n = m_n\Psi(s_n, \ldots, s_n) &= -\frac{km_ns_n}{n} + m_n \log\left[1 + \frac{k(e^{s_n}-1)}{n}\right] \\ &= \frac{m_n}{n}[-k\log(1+u_n) + n\log(1+\frac{k}{n}u_n)] \sim \frac{k}{2}\frac{m_n}{n}u_n^2 \sim \frac{k}{2}v_n^2 \to \infty \end{aligned}$$

using (2.19). Hence, from (2.36), we have

$$\lim_{n\to\infty} v_n^k \|V_n^{\star m_n} - \Phi^{\star m_n}_{\mu_n \mathbf{1}_k, \Sigma_n}\|_\infty = 0. \tag{2.37}$$

Combining (2.33) and (2.37), we get,

$$\lim_{n\to\infty} v_n^k \|H_n\|_\infty = 0$$

and the result follows from (2.31). □

Finally we prove Proposition 1.1.

Proof of Proposition 1.1. If $p_n = 1/n$, the condition $p_n(\log n)^r \to 0$ holds for all $r > 0$ and the result follows from (2.2) with $k = 1$, since

$$-\log(P[Y_{1n} \le y_n])^n \sim nP[Y_{1n} > y_n] \to e^{-x}.$$

For general p_n, the argument is exactly same as that for derivation of (2.2). The extra condition is required to prove the convergence in (2.23)–(2.25) and (2.33). In (2.33), the bound is of the order $v_n^k p_n \sim (2\log n)^{k/2} p_n$, necessitating the extra condition. In (2.23)–(2.25), the bound for $E_n^{(0)}$ in (2.23) is of the highest order, which in this case becomes $p_n \log n$. It is still negligible under the assumption. □

Acknowledgements. We thank an anonymous referee for helpful suggestions leading to a significantly clearer presentation and pointing out the interesting work of Nadarajah and Mitov (2002).

3 Appendix

Here we prove the result on integration by parts, which was used in the proof of Proposition 2.2, see (2.30). Let H be the distribution function of a finite signed measure on $\mathbb{R}^k$. For any subset I of $\{1,\dots,k\}$ and $a \in \mathbb{R}$, define,

$$y_i^I = \begin{cases} a, & i \in I \\ y_i, & i \notin I \end{cases}, \quad \text{for } 1 \le i \le k,$$

$$H^I(a; y_1,\dots,y_k) = H(y_1^I,\dots,y_k^I)$$

and

$$H_{\boldsymbol{y}}^I(y_i; i \in I) = H(y_1,\dots,y_k)$$

considered as a function in coordinates indexed by I only.

Theorem 3.1. *For $1 \le l \le k$ and $I \subseteq \{1,\dots,l\}$, we have,*

$$\int_a^\infty \cdots \int_a^\infty e^{-s(y_1+\cdots+y_k)} dH_{y_1,\dots,y_k}^{\{1,\dots,l\}}(y_1,\dots,y_k)$$
$$= \sum_{I\subset\{1,\dots,l\}} \int_a^\infty \cdots \int_a^\infty (-1)^{|I|} s^l e^{-s(y_1+\cdots+y_k)} H^I(a; y_1,\dots,y_k) dy_1\cdots dy_k. \quad (3.1)$$

The bound (2.30) then follows immediately by considering $l = k$.

Proof. We prove (3.1) by induction on l. For $l = 1$, (3.1) is the usual integration by parts formula. Assume (3.1) for l. Then

$$\int_a^\infty \cdots \int_a^\infty e^{-s(y_1+\cdots+y_{l+1})} H_{y_1,\dots,y_k}^{\{1,\dots,l+1\}}(dy_1,\dots,dy_{l+1})$$
$$= \sum_{I\subset\{1,\dots,l\}} (-1)^{|I|} \int_a^\infty \cdots \int_a^\infty s^l e^{-s(y_1+\cdots+y_l)} \int_a^\infty e^{-sy_{l+1}} H_{y_1^I,\dots,y_k^I}^{\{l+1\}}(dy_{l+1}) dy_1\cdots dy_l$$
$$= \sum_{I\subset\{1,\dots,l\}} (-1)^{|I|} \int_a^\infty \cdots \int_a^\infty s^l e^{-s(y_1+\cdots+y_l)} \Big[e^{-sa} H^{I\cup\{l+1\}}(a; y_1,\dots,y_k)$$
$$+ \int_a^\infty s e^{-sy_{l+1}} H^I(a; y_1,\dots,y_k) dy_{l+1} \Big] dy_1\cdots dy_l$$
$$= \sum_{I\subset\{1,\dots,l\}} \int_a^\infty \cdots \int_a^\infty e^{-s(y_1+\cdots+y_{l+1})} s^{l+1} \times$$
$$\Big[(-1)^{|I|+1} H^{I\cup\{l+1\}}(a; y_1,\dots,y_k) + (-1)^{|I|} H^I(a; y_1,\dots,y_k) \Big] dy_1\cdots dy_{l+1}$$

where we use the induction hypothesis for the first step and the usual integration by parts for the second step, and the final step is the required sum, since any subset of $\{1,\dots,l+1\}$ either contains $l+1$ or does not and the remainder is a subset of $\{1,\dots,l\}$. This completes the inductive step and the proof of the theorem. ∎

References

C. W. Anderson, S. G. Coles, and J. Hüsler. Maxima of Poisson-like variables and related triangular arrays. *Ann. Appl. Probab.*, 7(4):953–971, 1997. ISSN 1050-5164.

R. N. Bhattacharya and R. Ranga Rao. *Normal approximation and asymptotic expansions.* John Wiley & Sons, New York-London-Sydney, 1976. Wiley Series in Probability and Mathematical Statistics. MR0436272

W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968. MR0228020

R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest and smallest members of a sample. *Proc. Camb. Philos. Soc.*, 24:180–190, 1928.

B. Gnedenko. Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math. (2)*, 44:423–453, 1943. ISSN 0003-486X. MR0008655

L. de Haan. *On regular variation and its application to the weak convergence of sample extremes*, volume 32 of *Mathematical Centre Tracts.* Mathematisch Centrum, Amsterdam, 1970. MR0286156

M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes.* Springer Series in Statistics. Springer-Verlag, New York, 1983. ISBN 0-387-90731-9. MR0691492

S. Nadarajah and K. Mitov. Asymptotics of maxima of discrete random variables. *Extremes*, 5(3):287–294, 2002. ISSN 1386-1999. MR1995780

V. V. Petrov. *Sums of independent random variables.* Springer-Verlag, New York, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82. MR0388499