PACIFIC JOURNAL OF MATHEMATICS
Vol. 100, No. 2, 1982

ON THE MAXIMUM OF SCALED MULTINOMIAL VARIABLES

D. FREEDMAN

Suppose S_n is a sum of n independent, identically distributed, integer-valued random variables. Let $p_j = P(S_n = j)$. Take k independent copies of S_n, and let N_j be the number of these sums which are equal to j. In previous papers Persi Diaconis and I studied

$$\max_j N_j$$
$$\max_j (N_j - kp_j)$$
$$\max_j (N_j - k\tilde{p}_j)\ ,$$

where $\tilde{p}_j$ is the normal approximation to p_j. Likewise, we have studied the histogram as a density estimator. These problems all have a common structure, namely, determining the asymptotic behavior of the maximum of scaled multinomial variables. The object here is to present a general theorem, flexible enough to cover all the cases mentioned above. The form of this theorem may seem a bit arbitrary at first, but it is suggested by the special cases.

1. Introduction. In this section, the theorem will be stated; the proof is deferred to § 3. Section 2 presents the analogous theorem for normal variables, so as to bring out the main ideas in the proof.

For the main theorem, consider a sequence of multinomial distributions, indexed by n. However, this index may be suppressed in later sections, to lighten the notation. At every stage n, there are boxes indexed by the integers j. Associated with each box is a probability $p_j = p_{nj}$; and

$$\sum_{j=-\infty}^{\infty} p_{nj} = 1\ .$$

The p_{nj}'s will tend to 0 as n grows. At stage n, there are $k = k_n$ balls; $k_n \to \infty$ as $n \to \infty$. The balls are dropped independently in turn into the boxes; a ball lands in box j with probability p_{nj}. Let $N_j = N_{nj}$ be the number of balls which land in box j at stage n. Thus, each N_{nj} is binomial with small success probability p_{nj} and large number of trials k_n. Jointly, the variables N_{nj} for $j = 0, \pm 1, \cdots$ have a multinomial distribution.

Next, introduce coefficients $\alpha_j = \alpha_{nj} \geqq 0$ and $\beta_j = \beta_{nj}$. The primary interest is in $\max_j V_{nj}$, where

$$V_{nj} = \alpha_{nj} Z_{nj} + \beta_{nj}\sqrt{2\log 1/\varepsilon_n}\ . \tag{1.1}$$
$$Z_{nj} = (N_{nj} - k_n p_{nj})/\sqrt{k_n p_{nj}}$$

In (1.1),

$$\varepsilon = \varepsilon_n \longrightarrow 0 \tag{1.2}$$

is a scale factor for the coefficients, which also have a center $c = c_n$. To make this precise, introduce

$$t_{nj} = \varepsilon_n(j - c_n)\ . \tag{1.3}$$

and assume:

$$\alpha_{nj} = \alpha_n(t_{nj}) + o\Big(1/\log\frac{1}{\varepsilon_n}\Big)\ ; \tag{1.4}$$

$$\beta_{nj} = \beta_n(t_{nj}) + o\Big(1/\log\frac{1}{\varepsilon_n}\Big)\ ; \tag{1.5}$$

(1.6) the functions $\alpha_n \geqq 0$ and β_n are defined and continuous on a proper compact interval I, which does not depend on n.

Conditions (1.4) and (1.5) are required to hold uniformly over j with $t_{nj} \in I$. Assume further that

(1.7) $\alpha_n \longrightarrow \alpha_\infty$ and $\beta_n \longrightarrow \beta_\infty$ as $n \longrightarrow \infty$, uniformly on I.

(1.8) For $1 \leqq n \leqq \infty$, the function $\alpha_n + \beta_n$ has a unique global maximum at an interior point t_n of I, and $\alpha_n(t_n) > 0$; furthermore, $t_n \to t_\infty$.

Conditions (1.7) and (1.8) imply that

(1.9) $\alpha_n(t_n)$ is bounded below by a positive number.

Assume further that α_n and β_n are locally quadratic at t_n: namely, as $t \to t_n$,

$$\alpha_n(t) = \alpha_n(t_n) + \alpha'_n\cdot(t - t_n) + \frac{1}{2}\alpha''_n\cdot(t - t_n)^2 + o(t - t_n)^2 \tag{1.10}$$

$$\beta_n(t) = \beta_n(t_n) + \beta'_n\cdot(t - t_n) + \frac{1}{2}\beta''_n\cdot(t - t_n)^2 + o(t - t_n)^2 \tag{1.11}$$

where the "o" is uniform in n. Note that

$$\begin{aligned} \alpha'_n &\longrightarrow \alpha'_\infty, & \alpha''_n &\longrightarrow \alpha''_\infty \\ \beta'_n &\longrightarrow \beta'_\infty, & \beta''_n &\longrightarrow \beta''_\infty \end{aligned} \tag{1.12}$$

In (1.10) and (1.11), it is not necessary to assume differentiability anywhere except at t_n: the primes just denote numbers. Necessarily

$$\alpha_n' + \beta_n' = 0 . \tag{1.13}$$

Also, $\alpha_n'' + \beta_n'' \leqq 0$. More is assumed:

$$\alpha_n'' + \beta_n'' < 0 \quad \text{for} \quad 1 \leqq n \leqq \infty \tag{1.14}$$

Abbreviate

$$\rho_n^2 = -(\alpha_n'' + \beta_n'')/\alpha_n(t_n) > 0; \text{ set } \rho = \rho_\infty . \tag{1.15}$$

The following growth conditions will be assumed, as $n \to \infty$:

$$\limsup_{n\to\infty} \sum_j \{p_{nj}: t_{nj} \in I\} < 1 \tag{1.16}$$

$$p_{nj}\cdot\left(\log\frac{1}{\varepsilon_n}\right)^2 \longrightarrow 0 \tag{1.17}$$

$$(k_n p_{nj})\Big/\left(\log\frac{1}{\varepsilon_n}\right) \longrightarrow \infty . \tag{1.18}$$

To avoid "large deviation" terms, a condition stronger than (1.18) is needed:

$$(k_n p_{nj})\Big/\left(\log\frac{1}{\varepsilon_n}\right)^3 \longrightarrow \infty . \tag{1.19}$$

Conditions (1.17)-(1.19) are to hold uniformly in j with $t_{nj} \in I$. Note that (1.17)-(1.18) imply

$$k_n\Big/\left(\log\frac{1}{\varepsilon_n}\right)^3 \longrightarrow \infty . \tag{1.20}$$

To state the main result, let

$$w_n(x) = \left(2\log\frac{1}{\varepsilon_n} - 2\log\log\frac{1}{\varepsilon_n} + x\right)^{1/2} . \tag{1.21}$$

Let Φ be the standard normal distribution function, with density ϕ:

$$\Phi(y) = \int_{-\infty}^{y} \phi(u)du, \quad \text{where} \quad \phi(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) . \tag{1.22}$$

As usual, $\exp(v) = e^v$. Let

$$M_n = \max_j \{V_{nj}: t_{nj} \in I\}, \text{ occurring at index } L_n . \tag{1.23}$$

Here, V_{nj} is defined by (1.1).

In brief, the main result is that L_n and M_n are asymptotically independent, L_n being asymptotically normal and M_n being asymptotically double-exponential.

THEOREM 1.24. *Assume* (1.1)-(1.23). *Let* $n \to \infty$. *With proba-*

bility approaching one, $M_n = \max_j \{V_{nj}: t_{nj} \in I\}$ *is assumed at a unique index* L_n. *Furthermore, the chance that*

$$\rho\sqrt{2\log\frac{1}{\varepsilon_n}}[\varepsilon_n(L_n - c_n) - t_n] < y$$

and

$$M_n < \alpha_n(t_n)w_n(x) + \beta_n(t_n)\sqrt{2\log\frac{1}{\varepsilon_n}}$$

converges to

$$\Phi(y)\cdot\exp\left\{-\frac{1}{2\rho}e^{-x/2}\right\}.$$

Now for some heuristic comments about the theorem, especially the assumed error rate $o(1/\log 1/\varepsilon_n)$ in (1.4)–(1.5). This error rate is critical. Basically, all the action in M_n is over j's with t_{nj} close to t_n. So, M_n can be crudely approximated as the sum of two terms:

$$\alpha_n(t_n)\max_j\{Z_{nj}: t_{nj} \text{ near } t_n\}$$

and

$$\beta_n(t_n)\sqrt{2\log\frac{1}{\varepsilon_n}}.$$

Both terms are of order $\sqrt{\log 1/\varepsilon_n}$. Changing the coefficients by $o(1/\log 1/\varepsilon_n)$ changes M_n by $o(1/\sqrt{\log 1/\varepsilon_n})$. Next, consider the asymptotic distribution function for M_n in (1.24):

$$\alpha_n(t_n)w_n(\cdot) + \beta_n(t_n)\sqrt{2\log\frac{1}{\varepsilon_n}}.$$

This is centered just to the left of

$$[\alpha_n(t_n) + \beta_n(t_n)]\sqrt{2\log\frac{1}{\varepsilon_n}},$$

which may be large. But the spread is of order

$$1\Big/\sqrt{\log\frac{1}{\varepsilon_n}},$$

which is small. So the distribution may move off to infinity, but gets more and more concentrated. And only terms which are $o(1/\sqrt{\log 1/\varepsilon_n})$ can be dropped from M_n without affecting its asymptotic behavior.

Now for a comment on L_n. The action in M_n occurs for j's with t_{nj} near t_n. At first, it might seem that $O(1/\varepsilon_n)$ indices j should be involved, but this is slightly exaggerated: the right order is

$$\frac{1}{\varepsilon_n}\Big/\sqrt{\log\frac{1}{\varepsilon_n}}\,.$$

Thus, it is necessary to work within shrinking neighborhoods of t_∞. Then, it might seem that α and β can be treated as constants. Not so, however; the quadratic terms in (1.10)-(1.11) really matter in the asymptotics, as the presence of ρ in the statement of (1.24) should indicate: in effect, however, the linear terms cancel.

Next, a comment on the asymptotic independence. This is a bit surprising. In the vicinity of t, the maximum is around

$$[\alpha_n(t) + \beta_n(t)]\sqrt{2\log\frac{1}{\varepsilon_n}}$$

which diminishes as t moves away from t_n. Intuition suggests that large values of L_n should be accompanied by small values of M_n. However, this is too hasty. Keeping t_{nj} away from t_n makes V_{nj} smaller; but saying that $L_n = j$ makes V_{nj} larger. So there is some tension here, and (1.24) shows that the two effects balance.

Finally, a comment on the connection between the multinomial problem and the normal problem discussed in § 2. Formula (1.1) involves the scaled variables

$$Z_{nj} = (N_{nj} - k_n p_{nj})/\sqrt{k_n p_{nj}}$$

which are essentially standard normal, and practically independent. So a theorem for normal variables should—and does—go over to the multinomial case. The argument in § 2 is organized so that the estimates can be re-used in § 3. This depends, however, on the growth condition (1.19). If only (1.18) is assumed, the binomials are no longer quite so normal: "large deviations" corrections become relevant. For a discussion of this point, see [2].

Acknowledgment. This paper is an offshoot from an extended collaboration with Persi Diaconis on related problems, and it is a pleasure to acknowledge his help.

2. The normal case. Conditions (1.2)-(1.15) are assumed on the coefficients. Let U_j be independent standard normal variables, and

$$V_j = V_{nj} = \alpha_{nj}U_j + \beta_{nj}\sqrt{2\log\frac{1}{\varepsilon_n}}\,. \tag{2.1}$$

Define M_n and L_n as before, by (1.23). Recall (1.8).

THEOREM 2.2. *Assume* (1.2)-(1.15). *Define* V_{nj} *by* (2.1). *Assume* (1.21)-(1.23). *As* $n \to \infty$, *the chance that*

$$\rho\sqrt{2\log\frac{1}{\varepsilon_n}}[\varepsilon_n(L_n - c_n) - t_n] < y$$

and

$$M_n < \alpha_n(t_n)w_n(x) + \beta_n(t_n)\sqrt{2\log\frac{1}{\varepsilon_n}}$$

converges to

$$\Phi(y)\cdot\exp\left[\frac{1}{2\rho}e^{-1/2(x)}\right].$$

Without loss of generality, assume

$$\alpha_n(t_n) = 1 \quad \text{and} \quad \beta_n(t_n) = 0\,. \tag{2.3}$$

Let δ be a small positive number. It will be shown that j's with $|t_{nj} - t_\infty| \geqq \delta$ make essentially no contribution to the max, because with probability near one, the corresponding V_j's are all less than

$$(1-\theta)\sqrt{2\log\frac{1}{\varepsilon_n}}\,.$$

To make this this precise, only a very weak estimate is needed.

LEMMA 2.4. *Let* $Z_1, \cdots, Z_m$ *be standard normal variables, not necessarily independent. Let* $0 < a < \infty$. *Then*

$$P\{\max_{i=1,\cdots,m} Z_i \geqq \sqrt{2\log(am)}\} \longrightarrow 0\,.$$

Proof. The probability in question is bounded above by

$$mP\{Z_1 \geqq \sqrt{2\log(am)}\} \approx m\cdot\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{2\log am}}\cdot\exp\left\{-\frac{1}{2}\cdot 2\log(am)\right\}$$
$$= o(1)\,. \qquad \square$$

Notation. $y_n \approx x_n$ means $y_n/x_n \to 1$, while $y_n \sim x_n$ means

$$0 < \liminf_{n\to\infty} y_n/x_n \leqq \limsup_{n\to\infty} y_n/x_n < \infty\,.$$

LEMMA 2.5. *Fix any small, positive* δ. *Let*

$$I_\delta = \{t\colon t \in I \quad \textit{and} \quad |t - t_\infty| \geqq \delta\}\,.$$

For some sufficiently small positive θ, *the probability that*

$$\max_j \{V_{nj}: t_{nj} \in I_\delta\} < (1-\theta)\sqrt{2 \log \frac{1}{\varepsilon_n}}$$

approaches one as $n \to \infty$.

Proof. Recall that $\alpha_n(t_n) = 1$ and $\beta_n(t_n) = 0$ by the normalization (2.3). From (1.6)-(1.8), there is some $\theta > 0$ such that for all large n,

$$(1 - 5\theta) \geqq \max_t \{\alpha_n(t) + \beta_n(t): t \in I_\delta\} .$$

Now use compactness to express I_δ as a finite union of intervals J so short that

$$\alpha_n^J + \beta_n^J \leqq 1 - 4\theta ,$$

where

$$\alpha_n^J = \max_t \{\alpha_n(t): t \in J\}$$

and likewise for β_n^J.

Fix J. By (1.4)-(1.5), for all $n \geqq n_0$,

$$\alpha_{nj} \leqq \alpha_n^J + \theta \quad \text{and} \quad \beta_{nj} \leqq \beta_n^J + \theta$$

for all j with $t_{nj} \in J$. Then

$$\max_j \{V_{nj}: t_{nj} \in J\}$$

is bounded above by

$$(\alpha_n^J + \theta) \max_j \{U_j; t_{nj} \in J\} + (\beta_n^J + \theta)\sqrt{2 \log \frac{1}{\varepsilon_n}} .$$

By 2.4, with probability near one, the last display is at most

$$(\alpha_n^J + \theta + \beta_n^J + \theta)\sqrt{2 \log \frac{1}{\varepsilon_n}} \leqq (1 - 2\theta) \sqrt{2 \log \frac{1}{\varepsilon_n}} ,$$

there being only $0(1/\varepsilon_n)$ indices j with $t_{nj} \in J$. Since there only finitely many J's, the proof terminates. □

Note. In this part of the argument, the error terms in (1.4)-(1.5) need only be assumed to be $o(1)$. Also, since $t_n \to t_\infty$, for large n, if $|t_{nj} - t_n| \geqq \delta$ then $|t_{nj} - t_\infty| \geqq \delta/2$. Only j's with $|t_{nj} - t_n| < \delta$ contribute to the max.

Turn now to the j's with $|t_{nj} - t_n| < \delta$. Here, the argument is more complicated, and a sketch of the idea is given.

(2.6) For $-\infty \leqq a < b \leqq \infty$, let I_{ab} be the set of j's with $|t_{nj} - t_n| < \delta$ and $a \leqq \rho\sqrt{2 \log (1/\varepsilon_n)}(t_{nj} - t_n) < b$, and let $M_{ab} = \max_j \{V_{nj}: j \in I_{ab}\}$.

Despite the notation, I_{ab} and M_{ab} depend on n and δ. It will be proved, among other things, that

$$P\{V_{nj} > w_n(x)\} = o(1) . \tag{2.7}$$

Here and later, "o" and "0" errors are as $n \to \infty$, and are uniform over j with $|t_{nj} - t_n| < \delta$.

Clearly, for $-\infty \leqq a < b \leqq c < d \leqq \infty$,

$$P\{M_{ab} \leqq w_n(x) \text{ and } M_{cd} \leqq w_n(y)\} = P\{M_{ab} \leqq w_n(x)\} \cdot P\{M_{cd} \leqq w_n(y)\} .$$

The factors will be estimated, and appeal made to 2.35 below. For now, only a heuristic argument is given. By (2.7),

$$\log P\{M_{ab} \leqq w_n(x)\} \doteq - \sum_{j \in I_{ab}} P\{V_{nj} > w_n(x)\} .$$

The symbol $\doteq$ means approximately equal, and is used only informally. Now

$$P\{V_{nj} > w_n(x)\} = P\left\{U_j > \sqrt{2 \log \frac{1}{\varepsilon_n}} \lambda_{nj}(x)\right\} ,$$

where

$$\lambda_{nj}(x) = \frac{w_n(x) - \beta_{nj}\sqrt{2 \log \frac{1}{\varepsilon_n}}}{\alpha_{nj}\sqrt{2 \log \frac{1}{\varepsilon_n}}} . \tag{2.8}$$

This $\lambda_{nj}(x)$ is a key technical object in future arguments. To proceed,

$$\begin{aligned} P\left\{U_j > \sqrt{2 \log \frac{1}{\varepsilon_n}} \lambda_{nj}(x)\right\} &\doteq \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2 \log \frac{1}{\varepsilon_n}}} \frac{1}{\lambda_{nj}(x)} \\ &\quad \times \exp\left\{-\left(\log \frac{1}{\varepsilon_n}\right) \lambda_{nj}(x)^2\right\} \end{aligned} \tag{2.9}$$

It is necessary to estimate $\lambda_{nj}(x)$: as it turns out,

$$\begin{aligned} \lambda_{nj}(x) \doteq 1 &- \frac{1}{2}\left(\log\log \frac{1}{\varepsilon_n}\right) \Big/ \log \frac{1}{\varepsilon_n} + \frac{1}{4} x \Big/ \log \frac{1}{\varepsilon_n} \\ &+ \frac{1}{2}\rho^2 (t_{nj} - t_n)^2 \end{aligned} \tag{2.10}$$

Now $(1 + q)^2 \doteq 1 + 2q$ for small q, so

$$\begin{aligned} -\left(\log \frac{1}{\varepsilon_n}\right) \lambda_{nj}(x)^2 \doteq &-\log \frac{1}{\varepsilon_n} + \log\log \frac{1}{\varepsilon_n} - \frac{1}{2} x \\ &- \frac{1}{2}\rho^2 \left(2 \log \frac{1}{\varepsilon_n}\right)(t_{nj} - t_n)^2 . \end{aligned} \tag{2.11}$$

Now combine (2.9)–(2.11): the factor $1/\lambda_{nj}(x)$ on the right in (2.9) is essentially 1, so

$$P\left\{U_j > \sqrt{2\log\frac{1}{\varepsilon_n}}\lambda_{nj}(x)\right\} \doteq \frac{1}{2\rho}e^{-x/2}\cdot\varepsilon'\cdot\phi(u_j) \tag{2.12}$$

where

$$\varepsilon' = \rho\varepsilon_n\sqrt{2\log\frac{1}{\varepsilon_n}}$$
$$u_j = \rho\sqrt{2\log\frac{1}{\varepsilon_n}}(t_{nj} - t_n)$$

and ϕ was defined in (1.22) as the standard normal density. Some algebra has been omitted here: the $1/2\rho$ on the right of (2.12) is needed to offset the $\rho\sqrt{2}$ in ε' and get back the $1/\sqrt{2}$ on the right in (2.9). Continuing from (2.12),

$$\begin{aligned}\sum_{j\in I_{ab}} P\left\{U_j > \sqrt{2\log\frac{1}{\varepsilon_n}}\lambda_{nj}(x)\right\} &\doteq \frac{1}{2\rho}e^{-x/2}\cdot\sum_{j\in I_{ab}}\varepsilon'\phi(u_j) \\ &\doteq \frac{1}{2\rho}e^{-x/2}\cdot\int_a^b\phi(u)du\ ,\end{aligned} \tag{2.13}$$

because $u_{j+1} - u_j = \varepsilon'$.

This argument will now be made rigorous: it is (2.10) and (2.11) which take most of the work. Notice that $\lambda_{nj}(x)$ must be estimated to within $o(1/\log 1/\varepsilon_n)$, because its square gets multiplied by $\log 1/\varepsilon_n$. The assumptions are (1.2)–(1.15), (2.1) and (2.3). Two estimates will be needed on α_n and β_n; these estimates must be uniform in n. The proofs are omitted as routine.

LEMMA 2.14. *Fix $\eta_1 > 0$. There is a small positive δ such that for all n, and $|t - t_n| < \delta$,*

$$1 - \alpha_n(t) - \beta_n(t) = \frac{1}{2}\rho^2(1 \pm \eta_1)(t - t_n)^2\ .$$

More explicitly, the display means

$$\frac{1}{2}\rho^2(1 - \eta_1)(t - t_n)^2 < 1 - \alpha_n(t) - \beta_n(t) < \frac{1}{2}\rho^2(1 + \eta_1)(t - t_n)^2\ .$$

LEMMA 2.15. *Let $(1/2)K > 1 + \sup_n|\alpha_n'| + 1/2\sup_n|\alpha_n''|$. There is a small positive δ such that for all n, and $|t - t_n| < \delta$,*

$$\left|\frac{1}{\alpha_n(t)} - 1\right| < K|t - t_n|\ .$$

It is now time to estimate the $\lambda_{nj}(x)$ defined in (2.8), making (2.10)-(2.11) precise.

LEMMA 2.16. *Fix $\eta > 0$. For small positive δ, the following estimates will apply as $n \to \infty$, uniformly over j with $|t_{nj} - t_n| < \delta$, and uniformly over x in compact intervals.*

(a) $|\lambda_{nj}(x) - 1| < \eta + o(1)$

(b) $(\log 1/\varepsilon_n)\lambda_{nj}(x)^2 \leqq \log 1/\varepsilon_n - \log\log 1/\varepsilon_n + (x + \eta)/2 + (\rho^2/2)(1 + \eta)^2(2\log 1/\varepsilon_n)(t_{nj} - t_n)^2 + o(1)$

(c) $(\log 1/\varepsilon_n)\lambda_{nj}(x)^2 \geqq \log 1/\varepsilon_n - \log\log 1/\varepsilon_n + (x - \eta)/2 + (\rho^2/2)(1 - \eta)^2(2\log 1/\varepsilon_n)(t_{nj} - t_n)^2 + o(1)$

Proof. Choose δ so small that the estimates in (2.14)-(2.15) apply, with η_1 to be chosen later. Also choose δ so small that

$$(2.17)\quad \frac{3}{4} < \alpha_n(t) < \frac{5}{4} \quad\text{and}\quad -\frac{1}{4} < \beta_n(t) < \frac{1}{4} \quad\text{for}\quad |t - t_n| < \delta\,.$$

And so, for $n \geqq n_0$,

$$(2.18)\quad \frac{1}{2} < \alpha_{nj} < \frac{3}{2} \quad\text{and}\quad -\frac{1}{2} < \beta_{nj} < \frac{1}{2} \quad\text{for}\quad |t_{nj} - t_n| < \delta\,.$$

In (2.8), replace β_{nj} by $\beta_n(t_{nj})$ and α_{nj} by $\alpha_n(t_{nj})$. This gives a new quantity, to be denoted by $\lambda^*_{nj}(x)$. The first job is to show that

$$(2.19)\qquad \lambda_{nj}(x) = \lambda^*_{nj}(x) + o\Big(1/\log\frac{1}{\varepsilon_n}\Big)\,.$$

However, the first move only caused an error of

$$[\beta_{nj} - \beta_n(t_{nj})]/\alpha_{nj} = o\Big(1/\log\frac{1}{\varepsilon_n}\Big)$$

by (2.18) and assumption (1.5). Likewise, the second move only caused an error of

$$\left[\frac{w_n(x)}{\sqrt{2\log\frac{1}{\varepsilon_n}}} - \beta_n(t_{nj})\right]\left[\frac{1}{\alpha_{nj}} - \frac{1}{\alpha_n(t_{nj})}\right] = o\Big(1/\log\frac{1}{\varepsilon_n}\Big)\,.$$

This completes the proof of (2.19).

To proceed, let

$$r_n = \Big(\log\log\frac{1}{\varepsilon_n}\Big)\Big/\Big(\log\frac{1}{\varepsilon_n}\Big) \quad\text{and}\quad s_n(x) = x\Big/\Big(2\log\frac{1}{\varepsilon_n}\Big)$$

$$q_{nj}(x) = \lambda^*_{nj}(x) - 1$$

$$= \frac{[1 - r_n + s_n(x)]^{1/2} - \beta_n(t_{nj})}{\alpha_n(t_{nj})} - 1$$

$$= \frac{[1 - r_n + s_n(x)]^{1/2} - \alpha_n(t_{nj}) - \beta_n(t_{nj})}{\alpha_n(t_{nj})} .$$

This trivial bit of algebra is the key to the proof. Expand the square root and use (2.14):

$$q_{nj}(x) = \frac{1 - \alpha_n(t_{nj}) - \beta_n(t_{nj}) - \frac{1}{2}r_n + \frac{1}{2}s_n(x)}{\alpha_n(t_{nj})} + o\left(1\Big/\log\frac{1}{\varepsilon_n}\right)$$

$$= \frac{\frac{1}{2}\rho^2(1 \pm \eta_1)(t_{nj} - t_n)^2 - \frac{1}{2}r_n + \frac{1}{2}s_n(x)}{\alpha_n(t_{nj})} + o\left(1\Big/\log\frac{1}{\varepsilon_n}\right) .$$

Now use (2.15) to estimate $\alpha_n(t_{nj})^{-1} - 1$:

$$q_{nj}(x) = \frac{1}{2}\rho^2(1 \pm \eta_1)(t_{nj} - t_n)^2 - \frac{1}{2}r_n + \frac{1}{2}s_n(x) + \tau_1 + \tau_2 + \tau_3 + o\left(1\Big/\log\frac{1}{\varepsilon_n}\right)$$

where

$$|\tau_1| < \frac{1}{2}\rho^2(1 + \eta_1)(t_{nj} - t_n)^2 \cdot K|t_{nj} - t_n|$$

$$|\tau_2| < \frac{1}{2}r_n \cdot K|t_{nj} - t_n|$$

$$|\tau_3| < \frac{1}{2}s_n(|x|) \cdot K|t_{nj} - t_n| .$$

Now $|t_{nj} - t_n| < \delta$ by assumption, and $K\delta$ gets small with δ, so τ_1 merges into the first term: for small enough δ,

$$\frac{1}{2}\rho^2(1 \pm \eta_1)(t_{nj} - t_n)^2 + \tau_1 = \frac{1}{2}\rho^2(1 \pm 2\eta_1)(t_{nj} - t_n)^2 .$$

This uses (1.14)-(1.15) to force $\rho > 0$. Likewise, τ_3 merges into $s_n(x)$:

$$\frac{1}{2}s_n(x) + \tau_3 = \frac{1}{2}s_n(x \pm \eta) .$$

This leaves τ_2. With respect to this error, the claim is

$$\tau_2 = o(t_{nj} - t_n)^2 + o\left(1\Big/\log\frac{1}{\varepsilon_n}\right) . \tag{2.20}$$

Indeed, if $|t_{nj} - t_n| \leqq 1/(\log 1/\varepsilon_n)^{1/2}$ then

$$|\tau_2| \leqq \frac{1}{2}K\Big(\log\log\frac{1}{\varepsilon_n}\Big)\Big/\Big(\log\frac{1}{\varepsilon_n}\Big)^{3/2} = o\Big(1\Big/\log\frac{1}{\varepsilon_n}\Big)\,.$$

On the other hand, if $|t_{nj} - t_n| > 1/(\log 1/\varepsilon_n)^{1/2}$ then

$$|\tau_2|/(t_{nj} - t_n)^2 \leqq \frac{1}{2}K\Big(\log\log\frac{1}{\varepsilon_n}\Big)\Big/\Big(\log\frac{1}{\varepsilon_n}\Big)^{1/2} = o(1)\,.$$

This completes the proof of (2.20). But (2.20) shows that τ_2 merges into the lead term as well:

$$\frac{1}{2}\rho^2(1 \pm 2\eta_1)(t_{nj} - t_n)^2 + \tau_2 = q_{nj}^*(x) + o\Big(1\Big/\log\frac{1}{\varepsilon_n}\Big)$$

where

$$q_{nj}^*(x) = \frac{1}{2}\rho^2(1 \pm 3\eta_1)(t_{nj} - t_n)^2\,. \tag{2.21}$$

Combining this with (2.19): for small δ,

$$\lambda_{nj}(x) = 1 + q_{nj}^*(x) + o\Big(1\Big/\log\frac{1}{\varepsilon_n}\Big)\,. \tag{2.22}$$

Now (2.22) proves claim (a) of the lemma, because $(t_{nj} - t_n)^2 \leqq \delta^2$, so $q_{nj}^*(x)$ as defined in (2.21) is small with δ. Turning to claims (b)-(c),

$$\lambda_{nj}(x)^2 = 1 + 2q_{nj}^*(x) + q_{nj}^*(x)^2 + o\Big(1\Big/\log\frac{1}{\varepsilon_n}\Big)\,.$$

But $q_{nj}^*(x)^2$ merges into $q_{nj}^*(x)$, because the latter is small: referring to (2.21),

$$2q_{nj}^*(x) + q_{nj}^*(x)^2 = \rho^2(1 \pm 4\eta_1)(t_{nj} - t_n)^2\,.$$

To complete the proof, choose η_1 so that

$$(1 - \eta)^2 < 1 - 4\eta_1 < 1 + 4\eta_1 < (1 + \eta)^2\,. \qquad \square$$

This made (2.10)-(2.11) rigorous. Next, take up (2.13). Introduce

$$\psi_{nj}(x) = \Big(4\pi\log\frac{1}{\varepsilon_n}\Big)^{-1/2}\lambda_{nj}(x)^{-1}\cdot\exp\Big\{-\Big(\log\frac{1}{\varepsilon_n}\Big)\lambda_{nj}(x)^2\Big\}\,. \tag{2.23}$$

LEMMA 2.24. *If δ is small, $\psi_{nj}(x) = o(1)$ as $n \to \infty$, uniformly over j with $|t_{nj} - t_n| < \delta$, and uniformly over x in compact intervals.*

Proof. This is immediate from (2.16a), because $\lambda_{nj}(x)$ is essentially one, and $\log 1/\varepsilon_n$ goes to infinity. □

LEMMA 2.25. *Fix $\eta > 0$. If δ is small, the following estimates apply as $n \to \infty$, uniformly over extended real a and b, and x in compact intervals.*

(a) $\sum_{j \in I_{ab}} \psi_{nj}(x)$ *is bounded below by*

$$\frac{1}{(1+\eta)^2}\{\Phi[b(1+\eta)] - \Phi[a(1+\eta)]\}\cdot\frac{1}{2\rho}\cdot\exp\left\{-\frac{1}{2}(x+\eta)\right\} + o(1) .$$

(b) $\sum_{j \in I_{ab}} \psi_{nj}(x)$ *is bounded above by*

$$\frac{1}{(1-\eta)^2}\{\Phi[b(1-\eta)] - \Phi[a(1-\eta)]\}\cdot\frac{1}{2\rho}\cdot\exp\left\{-\frac{1}{2}(x-\eta)\right\} + o(1) .$$

Proof. This follows from (2.16). Claim (a) will be argued in some detail, and (b) is similar. Define

$$\varepsilon' = \rho(1+\eta)\varepsilon_n\left(2\log\frac{1}{\varepsilon_n}\right)^{1/2}$$

$$u_j = \rho(1+\eta)\left(2\log\frac{1}{\varepsilon_n}\right)(t_{nj} - t_n) .$$

The dependence of ε' and u_j on n is suppressed. Recall that ϕ is the standard normal density. As (2.16a) implies,

$$\lambda_{nj}(x)^{-1} \geqq [1 + o(1)]/(1+\eta) .$$

As (2.16b) now implies, $\psi_{nj}(x)$ is bounded below by

$$[1+o(1)]\frac{1}{(1+\eta)^2}\cdot\frac{1}{2\rho}\cdot\exp\left\{-\frac{1}{2}(x+\eta)\right\}\cdot\varepsilon'\phi(u_j) .$$

The "$o(1)$" is as $n \to \infty$, and is uniform over j with $|t_{nj} - t_n| < \delta$, for small enough δ.

Consider

$$\sum_{j \in I_{ab}} \varepsilon'\phi(u_j) . \tag{2.26}$$

Suppose $0 \leqq a < b < \infty$. Clearly, $u_{j+1} - u_j = \varepsilon'$; and ϕ is monotone decreasing on $[0, \infty)$. So

$$\varepsilon'\phi(u_j) > \int_{u_j}^{u_{j+1}} \phi(u)du .$$

As definition (2.6) shows, provided n is sufficiently large, $j \in I_{ab}$ iff $a(1+\eta) \leqq u_j < b(1+\eta)$. Then the sum in (2.26) is bounded below by

$$\int_{a(1+\eta)+\varepsilon'}^{b(1+\eta)} \phi(u)du .$$

The ε' in the lower limit of the integral takes care of an edge effect: j is discrete, u is continuous. Clearly, the displayed integral exceeds

$$\int_{a(1+\eta)}^{b(1+\eta)} \phi(u)du - \int_0^{\varepsilon'} \phi(u)du .$$

This completes the argument for (a) in case $0 \leqq a < b < \infty$; the case $b = \infty$ then follows; the case $-\infty \leqq a < b \leqq 0$ is symmetric; and the general case follows by addition. This disposes of claim (a), and (b) is similar. □

The η in (2.16) and (2.25) is a nuisance. Over the interesting j, with

$$\rho\sqrt{2 \log \frac{1}{\varepsilon_n}}\,|t_{nj} - t_n| < B < \infty ,$$

the argument in (2.16) is sharp enough to establish the results with $\eta = 0$. However, something needs to be done to cover the j's with, for instance,

$$B < \rho\sqrt{2 \log \frac{1}{\varepsilon_n}}(t_{nj} - t_n) < \infty$$
$$(t_{nj} - t_n) < \delta .$$

To do that, η was needed. Now, however, this technical nuisance can be eliminated. The interval I_{ab} was defined in (2.6), and depends on δ. This dependence matters in the next result, so the interval will be denoted $I_{ab}(\delta)$.

PROPOSITION 2.27. *If δ is small, uniformly over extended real a and b, and uniformly over x in compact intervals, as $n \to \infty$,*

$$\sum_{j \in I_{ab}(\delta)} \psi_{nj}(x) \longrightarrow [\Phi(b) - \Phi(a)]\frac{1}{2\rho}\exp\left\{-\frac{1}{2}x\right\} . \tag{2.28}$$

Proof. Denote the left side of (2.28) by $S_n(a, b, \delta, x)$; and the right side by $T(a, b, x)$. The first thing to show is that the tails don't matter. Fix η in (2.25) at any convenient value, say, $\eta=1/2$. This generates a corresponding δ, for which the estimate in (2.25b) is valid. This is the δ to use. Let $0 < B < \infty$, but large. Then

$$\limsup_{n\to\infty} S_n(B, \infty, \delta, x)$$

is bounded above according to (2.25b), and this bound is small for large B. Likewise for

$$\limsup_{n\to\infty} S_n(-\infty, -B, \delta, x) .$$

Let

$$\hat{a} = (-B) \wedge a \quad \text{and} \quad \hat{b} = b \wedge B .$$

Then $S_n(a, b, \delta, x)$ is bounded below by $S_n(\hat{a}, \hat{b}, \delta, x)$, and above by

$$S_n(\hat{a}, \hat{b}, \delta, x) + S_n(-\infty, -B, \delta, x) + S_n(B, \infty, \delta, x) .$$

As a result, it is only necessary to prove the lemma for a and b with

$$-B \leqq a < b \leqq B .$$

Now, if $j \in I_{ab}(\delta)$, then $j \in I_{ab}(\delta')$ for $n \geqq n_0$, where n_0 depends on δ' but on j. Here, δ' is positive but much smaller than δ. As a result, (2.25) applies with η arbitrarily small. □

Step (2.9) in the heuristic argument is easy to rigorize, in view of (2.16).

For small δ: uniformly in j with $|t_{nj} - t_n| < \delta$, as $n \to \infty$,

$$P\left\{U_j > \sqrt{2 \log \frac{1}{\varepsilon_n}}\lambda_{nj}(x)\right\} = [1 + o(1)]\psi_{nj}(x) . \tag{2.29}$$

Now (2.13) can be finished.

PROPOSITION 2.30. *If δ is small:*

(a) $P\{U_j > \sqrt{2 \log (1/\varepsilon_n)}\lambda_{nj}(x)\} = o(1)$ *as $n \to \infty$, uniformly over j with $|t_{nj} - t_n| < \delta$, and uniformly over x in compact intervals;*

(b) $\sum_{j \in I_{ab}} P\{U_j > \sqrt{2 \log (1/\varepsilon_n)}\lambda_{nj}(x)\}$ *converges to*

$$[\Phi(b) - \Phi(a)] \cdot \frac{1}{2\rho} \cdot e^{-x/2}$$

as $n \to \infty$, uniformly over extended real numbers a and b, and x in compact intervals.

Proof. Claim (a) is immediate from (2.29) and (2.24). Likewise, claim (b) is immediate from (2.29) and (2.27). □

This completes the rigorous discussion of (2.13). Recall M_{ab} from (2.6). The next step is to determine the joint distribution of M_{ab} and M_{cd}.

PROPOSITION 2.31. *If δ is small: uniformly over a, b, c, d with*

$-\infty \leqq a < b \leqq c < d \leqq \infty$, *and uniformly over* x, y *in compact intervals,*

$$P\{M_{ab} \leqq w_n(x) \quad and \quad M_{cd} \leqq w_n(y)\}$$

converges to $\exp\{Q(x, y)\}$ *as* $n \to \infty$, *where*

$$-2\rho Q(x, y) = [\Phi(b) - \Phi(a)]e^{-(1/2)x} + [\Phi(d) - \Phi(c)]e^{-(1/2)y} .$$

Proof. Clearly, the logarithm of the probability is

$$\sum_{j \in I_{ab}} \log P\{V_{nj} \leqq w_n(x)\} + \sum_{j \in I_{cd}} \log P\{V_{nj} \leqq w_n(y)\} . \tag{2.32}$$

Take the first sum, for instance. Definition (2.8) of $\lambda_{nj}(x)$ was set up so that

$$V_{nj} \leqq w_n(x) \quad \text{iff} \quad U_j \leqq \sqrt{2 \log \frac{1}{\varepsilon_n}}\lambda_{nj}(x) .$$

Expanding $\log p = \log[1 - (1 - p)] \doteq -(1 - p)$, the first sum in (2.32) can be estimated as

$$-[1 + o(1)] \sum_{j \in I_{ab}} P\left\{U_j > \sqrt{2 \log \frac{1}{\varepsilon_n}}\lambda_{nj}(x)\right\} \longrightarrow -[\Phi(b) - \Phi(a)]\frac{1}{2\rho}e^{-(1/2)x} ,$$

using (2.30). The other sum is similar. □

In (2.31), the index j was restricted so that $|t_{nj} - t_n| < \delta$. This was part of the definition of $I_{ab} = I_{ab}(\delta)$, even for infinite a and b, in (2.6). As a result, M_{ab} depends on δ too; write $M_{ab}(\delta)$ to indicate this dependence. The restriction on j was necessary, to make the estimates in (2.16). However, it can now be eliminated.

COROLLARY 2.33. *The conclusions of* (2.31) *apply, whatever* δ *may be.*

Proof. Let δ be small, so that (2.31) applies, and let δ' be large. Let

$$M^* = \max\{V_{nj}: |t_{nj} - t_n| \geqq \delta\} .$$

Clearly,

$$M_{ab}(\delta) \leqq M_{ab}(\delta') \leqq M_{ab}(\delta) \vee M^* .$$

But (2.5) and (2.31) show

$$P\left\{M^* \leqq (1 - \theta)\sqrt{2 \log \frac{1}{\varepsilon_n}} \leqq M_{ab}(\delta)\right\} \longrightarrow 1 ,$$

so

$$P\{M_{ab}(\delta') = M_{ab}(\delta)\} \longrightarrow 1 . \qquad \square$$

In particular, δ can be chosen so large that $\{t: |t - t_n| < \delta\}$ includes all of I. The proof of Theorem 2.2 is then accomplished by appeal to (2.35) below.

It is helpful, at times, to take $\max_j V_{nj}$ not only over the j with $t_{nj} \in I$, but over all integers j. This can be done if α_{nj} and β_{nj} are defined for all j, and j's with $t_{nj} \notin I$ do not count. This can be made precise, as follows.

COROLLARY 2.34. *Assume* (1.2)–(1.15). *Define* V_{nj} *by* (2.1), *for all integers* j. *Assume* (1.21)–(1.23). *Assume further that for some* $\theta > 0$,

$$\max_j \{V_{nj}: t_{nj} \notin I\} < (1 - \theta)\sqrt{2 \log \frac{1}{\varepsilon_n}}$$

with probability approaching one as $n \to \infty$. *Then the conclusions of* (2.2) *apply to*

$$M_n = \max_j \{V_{nj}: \textit{all integer } j\} .$$

Proof. Use the argument of (2.33). $\square$

Note. The condition here is that α_{nj} should get small, or β_{nj} should become large and negative, or both, as $j \to \pm\infty$.

Theorem 2.35 below will be used repeatedly, and so it is given here in some generality.

Framework for 2.35. Let V_{nj} be a random variable, defined for integers j in a finite, non-empty (non-random) interval J_n. For (2.2), take J_n to be the set of j's with $t_{nj} \in I$, and V_{nj} is defined by (2.1). Let $v_n(y)$ be a strictly increasing function of y, with $v_n(-\infty) = -\infty$ and $v_n(\infty) = \infty$. For (2.2),

$$v_n(y) = \rho\sqrt{2 \log \frac{1}{\varepsilon_n}}[\varepsilon_n(y - c_n) - t_n] .$$

Likewise, let $w_n(x)$ be a continuous and strictly increasing function of x, with $w_n(-\infty) = -\infty$ and $w_n(\infty) = \infty$. For (2.2), this function is defined by (1.21) down to

$$x = -2 \log \frac{1}{\varepsilon_n} + 2 \log \log \frac{1}{\varepsilon_n} ;$$

it may be extended back to $-\infty$ in any convenient way, subject to the conditions given above. Let

$$I_{ab}^n = \{j: j \in J_n \text{ and } v_n(a) \leqq j < v_n(b)\}$$
$$M_{ab}^n = \max_j \{V_{nj}: j \in I_{ab}^n\}$$

with $M_{ab}^n = -\infty$ if I_{ab}^n is empty. Let

$$M_n = \max_j \{V_{nj}: j \in J_n\}$$

with L_n being the leftmost j at which the max occurs, and $\hat{L}_n$ the rightmost. Let Φ be a distribution function: for (2.2), the standard normal. Let ψ be a monotone increasing function on $(-\infty, \infty)$, with $\psi(-\infty) = -\infty$, and $\psi(\infty) = 0$; suppose too that ψ has a continuous derivative ψ'. For (2.2), take $\psi(x) = -(1/2\rho)e^{-x/2}$.

THEOREM 2.35. *Under the conditions given above, suppose that for* $-\infty \leqq a < b \leqq c < d \leqq \infty$

$$P\{M_{ab}^n \leqq w_n(x) \quad \textit{and} \quad M_{cd}^n \leqq w_n(y)\}$$

converges to $\exp\{Q(x, y)\}$ *as* $n \to \infty$, *where*

$$Q(x, y) = [\Phi(b) - \Phi(a)]\psi(x) + [\Phi(d) - \Phi(c)]\psi(y) .$$

Then

$$P\{v_n(a) \leqq L_n \leqq \hat{L}_n \leqq v_n(b) \textit{ and } M_n \leqq w_n(x)\}$$

converges to

$$[\Phi(b) - \Phi(a)]\cdot\exp\{\psi(x)\} .$$

Proof. Begin with the case $a = -\infty$. Since $v_n(-\infty) = -\infty$, there is no condition on L_n. Now

$$\hat{L}_n \leqq v_n(b) \quad \text{and} \quad M_n > w_n(x)$$

iff

$$X_n > x \quad \text{and} \quad X_n > Y_n$$

where

$$X_n = w_n^{-1}(M_{-\infty b}^n) \quad \text{and} \quad Y_n = w_n^{-1}(M_{b\infty}^n) .$$

By assumption, w_n is strictly increasing, and its range is the whole line, so w_n^{-1} is well defined.

As is given in the statement of the proposition, (X_n, Y_n) converges in law to (X, Y), where

$$P\{X \leqq x \quad \text{and} \quad Y \leqq y\} = \exp\{Q(x, y)\} .$$

In particular, X has the probability density

$$\Phi(b)\psi'(u)\exp\{\Phi(b)\psi(u)\}$$

while

$$P\{Y < u\} = \exp\{[1 - \Phi(b)]\psi(u)\} ;$$

and X, Y are idependent. Note that ψ is negative, but ψ' is positive. Now

$$P\{X_n > x \text{ and } X_n > Y_n\} \longrightarrow P\{X > x \text{ and } X > Y\} .$$

The limiting probability is

$$\begin{aligned}
&\int_x^\infty \Phi(b)\psi'(u) \exp\{\Phi(b)\psi(u)]\}\cdot\exp\{[1 - \Phi(b)]\psi(u)\}du \\
&\quad = \Phi(b)\int_x^\infty \psi'(u)\exp\{\psi(u)\}du \\
&\quad = \Phi(b)[\exp\{\psi(\infty)\} - \exp\{\psi(x)\}] \\
&\quad = \Phi(b)[1 - \exp\{\psi(x)\}]
\end{aligned}$$

This proves the theorem when $a = -\infty$ and $-\infty < b < \infty$. A similar argument goes through when $-\infty < a < \infty$ and $b = \infty$. In particular, L_n and $\hat{L}_n$ have the same asymptotic distribution, namely, the law of $v_n^{-1}(L_n)$ converges weak-star to Φ, and likewise for $v_n^{-1}(\hat{L}_n)$. Since $v_n^{-1}(L_n) \leqq v_n^{-1}(\hat{L}_n)$, it follows that

$$v_n^{-1}(\hat{L}_n) - v_n^{-1}(L_n) \longrightarrow 0 \text{ in probability.} \tag{2.36}$$

The balance of the argument is omitted as routine. □

3. The multinomial case. In this section, Theorem 1.24 will be proved. We are back in the multinomial situation: (1.1)-(1.23) are in force. Without further loss of generality, assume the normalization (2.3). Again, let δ be a small positive number. The j's with $|t_{nj} - t_\infty| \geqq \delta$ make essentially no contribution to the max, because with probability near one, the corresponding V_{nj}'s are all less than

$$(1 - \theta)\sqrt{2\log\frac{1}{\varepsilon}} .$$

This will be seen in (3.3). Here $\varepsilon = \varepsilon_n$; the subscript n was dropped to lighten the notation.

LEMMA 3.1. *Let $\zeta > 0$. Then*

$$P\left\{\max_j Z_{nj} > (1 + \zeta)\sqrt{\log\frac{1}{\varepsilon}}\right\} \longrightarrow 0$$

as $n \to \infty$, the max being taken over all j with $t_{nj} \in I$.

Proof. The chance that

$$N_j - kp_j > (1 + \zeta)\sqrt{kp_j}\sqrt{2 \log \frac{1}{\varepsilon}} \tag{3.2}$$

is at most

$$\exp\left\{-\left(\log \frac{1}{\varepsilon}\right)(1 + \zeta)^2\gamma\right\}$$

where

$$\gamma = 1\Big/\left[1 + (1 + \zeta)\sqrt{\left(2 \log \frac{1}{\varepsilon}\right)\Big/kp_j}\right]$$

is uniformly close to 1 by condition (1.18). Eventually, $\gamma > 1/(1+\zeta)$, and then the probability of the event (3.2) will be bounded above by

$$\exp\left\{-\left(\log \frac{1}{\varepsilon}\right)(1 + \zeta)\right\} = \varepsilon^{1+\zeta} .$$

However, there are only $0(1/\varepsilon)$ indices j with $t_{nj} \in I$. The version of Bernstein's inequality used above appears as theorem (4) in [5]. □

LEMMA 3.3. *Fix any small positive* δ. *Let*

$$I_\delta = \{t\colon t \in I \quad \text{and} \quad |t - t_\infty| \geqq \delta\} .$$

For sufficiently small positive θ, *the probability that*

$$\max_j \{V_{nj}\colon t_{nj} \in I_\delta\} < (1 - \theta)\sqrt{2 \log \frac{1}{\varepsilon}}$$

approaches one as n *tends to infinity.*

Proof. Argue exactly as in (2.5), but use (3.1) instead of (2.4). For $n \geqq n_0$,

$$\max_j \{V_{nj}\colon t_{nj} \in J\}$$

is bounded above by

$$(\alpha_n^J + \theta) \max_j \{Z_{nj}\colon t_{nj} \in J\} + (\beta_n^J + \theta)\sqrt{2 \log \frac{1}{\varepsilon}} .$$

By (3.1), with probability approaching one, the last display is at most

$$\begin{aligned}&[(\alpha_n^J + \theta)(1 + \zeta) + (\beta_n^J + \theta)]\sqrt{2 \log \frac{1}{\varepsilon}} \\ &\quad < (1 + \zeta)(\alpha_n^J + \beta_n^J + 2\theta)\sqrt{2 \log \frac{1}{\varepsilon}}\end{aligned}$$

$$< (1 + \zeta)(1 - 2\theta)\sqrt{2 \log \frac{1}{\varepsilon}}$$

$$< (1 - \theta)\sqrt{2 \log \frac{1}{\varepsilon}}$$

for small ζ. □

Note. Again, in this part of the argument, the error terms in (1.4)-(1.5) need only be assumed to be $o(1)$. The result disposes of the j with $|t_{nj} - t_n| \geqq \delta$.

It is now time to deal with the j's for which $|t_{nj} - t_n| < \delta$. Define I_{ab} and M_{ab} by (2.6), with V_{nj} from (1.1). Despite the notation, I_{ab} and M_{ab} depend on n and on δ. This will be made explicit only when it matters.

PROPOSITION 3.4. *Assume* (1.1)-(1.23). *If δ is small, and $-\infty \leqq a < b \leqq c < d \leqq \infty$, then*

$$P\{M_{ab} \leqq w_n(x) \quad \textit{and} \quad M_{cd} \leqq w_n(y)\}$$

converges to $\exp\{Q(x, y)\}$ *as $n \to \infty$, where*

$$-2\rho Q(x, y) = [\Phi(b) - \Phi(a)]e^{-(1/2)x} + [\Phi(d) - \Phi(c)]e^{-(1/2)y} .$$

Granting (3.4), the condition on δ can be eliminated by (3.3), just as in (2.33). Then the proof of (1.24) can be completed by appealing to (2.35). Thus, (1.24) reduces to (3.4). Before going on to the proof of (3.4), note that it may sometimes be helpful to take $\max_j V_{nj}$ not only over the j with $t_{nj} \in I$, but over all integer j. This can be done, for α_{nj}, β_{nj}, p_{nj} and N_{nj} are defined for all j, as is V_{nj} by (1.1).

COROLLARY 3.5. *Assume* (1.1)-(1.23). *Suppose further that for some $\theta > 0$,*

$$\max_j \{V_{nj}: t_{nj} \notin I\} < (1 - \theta)\sqrt{2 \log \frac{1}{\varepsilon_n}}$$

with probability approaching one as n tends to infinity. Then the conclusions of (1.24) *apply as well to*

$$M_n = \max_j \{V_{nj}: \textit{all integer } j\} .$$

Proof. Use the argument of (2.33). □

Turn now to the proof of (3.4). It will be necessary to estimate the probability above and below. The upper bound is easier. In essence, an inequality of Mallows (1968) shows that the probability is at most

$$\prod_{j \in I_{ab}} P\{V_{nj} \leqq w_n(x)\} \cdot \prod_{j \in I_{cd}} P\{V_{nj} \leqq w_n(y)\} .$$

Of course, $\log p = \log [1 - (1 - p)] < -(1 - p)$ for $0 < p < 1$, so the logarithm of the displayed product is at most

$$-\sum_{j \in I_{ab}} P\{V_{nj} > w_n(x)\} - \sum_{j \in I_{cd}} P\{V_{nj} > w_n(y)\} .$$

Due to the minus-sign, $P\{V_{nj} > w_n(x)\}$ must be estimated from below. Recall $\lambda_{nj}(x)$ and $\psi_{nj}(x)$ from (2.8) and (2.23). Then

$$\begin{aligned} P\{V_{nj} > w_n(x)\} &= P\left\{\{N_j > kp_j + \sqrt{kp_j}\sqrt{2 \log \frac{1}{\varepsilon} \lambda_{nj}(x)}\right\} \\ &= [1 + o(1)]\psi_{nj}(x) , \end{aligned} \tag{3.6}$$

the "$o(1)$" being uniform over j with $|t_{nj} - t_n| < \delta$, provided δ is small. The argument for (3.6) is omitted, being very similar to one below. The lim sup of the probability in (3.4) is then at most $\exp\{Q(x, y)\}$, by (2.27). Further details on the lim sup are omitted.

For the lim inf,

(3.7) Under the conditions of (3.4), the probability in (3.4) is bounded below by $\gamma + o(1)$, where $\log \gamma$ is in turn bounded below by

$$-[1 + o(1)][\sum_{j \in I_{ab}} \psi_{nj}(x) + \sum_{j \in I_{cd}} \psi_{nj}(y)] + o(1) .$$

Granting (3.7), an appeal to (2.27) shows that the lim inf of the probability in (3.4) is at least $\exp\{Q(x, y)\}$. This completes the proof of (3.4), and hence of the main theorem (1.24).

Thus, (1.24) is reduced to (3.7). Now begins a series of calculations designed to prove (3.7). Eventually, lemma (3.2) of [1] will be used. Let

$$\begin{aligned} A_j &= \{V_{nj} \leqq w_n(x)\} \quad \text{for} \quad j \in I_{ab} \\ &= \{V_{nj} \leqq w_n(y)\} \quad \text{for} \quad j \in I_{cd} \end{aligned} \tag{3.8}$$

(3.9) K_j be the set of i with $t_n - \delta \leqq t_{ni} \leqq t_{nj}$

$$g_j = \sum_{i \in K_j} p_i \tag{3.10}$$

$$R = 2\left(k \log \frac{1}{\varepsilon}\right)^{1/2} \tag{3.11}$$

$$G_j = \{\sum_{i \in K_j} N_i > kg_j - R\} \tag{3.12}$$

In (3.9), note that $t_{ni} \leqq t_{nj}$ iff $i \leqq j$. In (3.10), assumption (1.16) implies

$$\limsup_n \max_j \{g_j \colon t_{nj} \in I\} < 1 . \tag{3.13}$$

LEMMA 3.14. *$\sum_j \{1 - P(G_j)\colon t_{nj} \in I\} \to 0$ as $n \to \infty$.*

Proof. Clearly, $\sum_{i \in K_j} N_i$ is binomial, with number of trials k and success probability $g_j < 1$. So

$$\begin{aligned} 1 - P(G_j) &< \exp\left\{-\frac{1}{2}R^2\Big/kg_j\right\} \\ &< \exp\left\{-\frac{1}{2}R^2\Big/k\right\} \\ &= \exp\left\{-2\log\frac{1}{\varepsilon}\right\} \quad \text{by (3.11)} \\ &= \varepsilon^2 . \end{aligned}$$

But there are only $o(1/\varepsilon)$ indices j with $t_{nj} \in I$, completing the proof. The version of Bernstein's inequality used above appears as theorem (4) in [5]. □

LEMMA 3.15. *Let N'_j be binomial, with success probability $p'_j = p_j/(1 - g_{j-1})$ and number of trials k'_j, the integer part of $k(1-g_{j-1}) + R$. Let*

$$\begin{aligned} V'_j = V'_{nj} &= \alpha_{nj}(kp_j)^{-1/2}(N'_j - kp_j) + \beta_{nj}\left(2\log\frac{1}{\varepsilon}\right)^{1/2} \pi_j(x) \\ &= \pi_{nj}(x) = P\{V'^{n}_j > w_n(x)\} . \end{aligned}$$

If δ is small and n is large, then

$$P\{M_{ab} \leqq w_n(x) \quad and \quad M_{cd} \leqq w_n(y)\}$$

is bounded below by

$$\{\prod_{j \in I_{ab}} [1 - \pi_j(x)]\}\cdot\{\prod_{j \in I_{cd}} [1 - \pi_j(y)]\} - \sum_{j \in I_{-\infty\infty}} [1 - P(G_j)] .$$

By definition (2.6), *$I_{-\infty\infty}$ is the set of j's with $|t_{nj} - t_n| < \delta$.*

Proof. This follows from (3.2) of [1]. Indeed, let $\mathscr{F}_j$ be the σ-field spanned by N_i for $i \in K_j$. Given $\mathscr{F}_{j-1}$, the variable N_j is conditionally binomial, with success probability p'_j; the number of trials T_j is an $\mathscr{F}_{j-1}$-measurable random variable:

$$T_j = k - \sum_{i \in K_{j-1}} N_i .$$

On G_{j-1}, however, $T_j \leqq k'_j$. So

$$\begin{aligned}
&P\{V_j \leqq w_n(x) \mid \mathscr{F}_{j-1}\} \\
&\quad = P\left\{N_j \leqq kp_j + \sqrt{kp_j}\sqrt{2\log\frac{1}{\varepsilon}}\lambda_{nj}(x) \,\middle|\, \mathscr{F}_{j-1}\right\} \\
&\quad \geqq P\left\{N'_j \leqq kp_j + \sqrt{kp_j}\sqrt{2\log\frac{1}{\varepsilon}}\lambda_{nj}(x)\right\} \\
&\quad = P\{V'_j \leqq w_n(x)\} \\
&\quad = 1 - \pi_j(x) .
\end{aligned}$$

The heuristic: the more often you toss the coin, the more heads come up.

In this argument, it is tacitly supposed that $\alpha_{nj} > 0$, which will be the case for all j with $|t_{nj} - t_n| < \delta$, if δ is small and n is large. □

This proves the first part of (3.7): the product of the $(1-\pi)$'s in (3.15) serves for γ, and the sum of the $[1 - P(G_j)]$'s is $o(1)$ by (3.14). For the second part of (3.7), estimate

$$\log \gamma \doteq -\sum_{j \in I_{ab}} \pi_j(x) - \sum_{j \in I_{cd}} \pi_j(y) .$$

This will now be made rigorous.

LEMMA 3.16. *If δ is small, then*

$$\pi_j(x) = [1 + o(1)]\psi_{nj}(x)$$

uniformly over j with $|t_{nj} - t_n| < \delta$, as $n \to \infty$.

Proof. To begin with, for some ζ with $0 \leqq \zeta < 1$,

$$k'_j = k(1 - g_{j-1}) + R - \zeta ,$$

so

$$k'_j p'_j = kp_j + \frac{(R-\zeta)p_j}{1 - g_{j-1}}$$

and

$$k'_j p'_j = kp_j + 0(Rp_j) \tag{3.17}$$

by (3.13). Consequently,

$$\begin{aligned}
(k'_j p'_j)/(kp_j) &= 1 + O\left[\left(\log\frac{1}{\varepsilon}\right)^{1/2} \Big/ k^{1/2}\right] \quad \text{by (3.11)} \\
&= 1 + o\left(1 \Big/ \log\frac{1}{\varepsilon}\right) \quad \text{by (1.20)} .
\end{aligned} \tag{3.18}$$

Of course, $\pi_j(x)$ is the chance that N_j' exceeds

$$kp_j + \sqrt{kp_j}\sqrt{2 \log \frac{1}{\varepsilon}}\lambda_{nj}(x) . \tag{3.19}$$

Now, the expression (3.19) can be rewritten as

$$k_j'p_j' + \sqrt{k_j'p_j'}\sqrt{2 \log \frac{1}{\varepsilon}}\left\{\lambda_{nj}(x) + o\left(1\Big/ \log \frac{1}{\varepsilon}\right)\right\} . \tag{3.20}$$

Granting this transformation, proposition (3.17) of [1] can be applied and completes the proof. The conditions of that lemma are satisfied by assumptions (1.2) and (1.17–18), and the properties of k_j' and p_j' developed obove. Note that $\lambda_{nj}(x)$ is nearly 1, by (2.16a). Also, (3.4) of [1] can be used to simplify the expression in (3.17) of [1].

To get from (3.19) to (3.20), note that

$$\begin{aligned} \frac{kp_j - k_j'p_j'}{\sqrt{k_j'p_j'}\sqrt{2 \log \frac{1}{\varepsilon}}} &\approx \frac{kp_j - k_j'p_j'}{\sqrt{kp_j}\sqrt{2 \log \frac{1}{\varepsilon}}} && \text{by (3.18)} \\ &= O(\sqrt{p_j}) && \text{by (3.11) and (3.17)} \\ &= o\left(1\Big/ \log \frac{1}{\varepsilon}\right) && \text{by (1.17)} \end{aligned} \tag{3.21}$$

So, it is harmless to replace the first term kp_j in (3.19) by $k_j'p_j'$.

Now replace the factor $\sqrt{kp_j}$ in the second term of (3.19) by $\sqrt{k_j'p_j'}$. The error is $\gamma_1\gamma_2\gamma_3$, where

$$\begin{aligned} \gamma_1 &= \sqrt{k_j'p_j'}\sqrt{2 \log \frac{1}{\varepsilon}} \\ \gamma_2 &= [(kp_j)/(k_j'p_j')]^{1/2} - 1 \\ \gamma_3 &= \lambda_{nj}(x) \end{aligned}$$

But $\gamma_2 = o(1/\log 1/\varepsilon)$ by (3.18), and $\gamma_3 = O(1)$ by (2.16a). This completes the move from (3.19) to (3.20), and hence the proof. □

Proof of 3.7. The probability in (3.14) is bounded below by $\gamma - \gamma'$, where

$$\begin{aligned} \gamma &= \prod_{j \in I_{ab}} [1 - \pi_j(x)] \cdot \prod_{j \in I_{cd}} [1 - \pi_j(y)] \\ \gamma' &= \sum_{j \in I_{-\infty\infty}} [1 - P(G_j)] , \end{aligned} \tag{3.22}$$

in view of (3.15). And $\gamma' \to 0$ by (3.14). Next,

$$\pi_j(x) = [1 + o(1)]\psi_{nj}(x) = o(1) \tag{3.23}$$

uniformly over j with $|t_{nj} - t_n| < \delta$, by (3.16) and (2.24). Of course, $\log(1-u) = -[1+o(1)]u$ as $u \to 0$, and this completes the proof. $\square$

This completes the proof of the distributional assertions in (1.24), and it remains only to prove that the maximum occurs at a unique index. First, some heuristics. Inspection of (1.24) suggests that the maximum is likely to occur only for $O((1/\varepsilon)/\sqrt{\log 1/\varepsilon})$ indices j near $c_n + \varepsilon_n^{-1}t_n$. Call these the critical j's. If N_j is maximal, it must take a critical value i, of order

$$kp_j + \sqrt{kp_j}\sqrt{2\log\frac{1}{\varepsilon}}\lambda_{nj}(x)\,, \tag{3.24}$$

where

$$\lambda_{nj}(x) \doteq 1 + \frac{1}{2}\rho^2(t_{nj}-t_n)^2 - \frac{1}{2}\frac{\log\log 1/\varepsilon}{\log 1/\varepsilon} + \frac{1}{4}\frac{x}{\log 1/\varepsilon}\,. \tag{3.25}$$

In particular, there are $O[\sqrt{kp_j}/\sqrt{\log 1/\varepsilon}]$ values i which are critical for the index j.

Suppose the maximum occurs at indices j and j'. Then

$$N_j = i \quad \text{and} \quad N_{j'} = i'$$

with

$$\begin{aligned}\alpha_{nj}\frac{i-kp_j}{\sqrt{kp_j}} &+ \beta_{nj}\sqrt{2\log\frac{1}{\varepsilon}} \\ &= \alpha_{nj'}\frac{i'-kp_{j'}}{\sqrt{kp_{j'}}} + \beta_{nj'}\sqrt{2\log\frac{1}{\varepsilon}}\,.\end{aligned} \tag{3.26}$$

Thus, i and i' are rigidly linked. The chance that the maximum occurs at two distinct indices j and j' is then

$$\sum_{j,j'}\sum_{i,i'} P\{N_j = i \quad \text{and} \quad N_{j'} = i'\} \tag{3.27}$$

This sum extends only over critical indices j and j'; for each j, the inner sum is over i's critical for j. Now N_j and $N_{j'}$ are nearly independent, so

$$P\{N_j = i \quad \text{and} \quad N_{j'} = i'\} \doteq P\{N_j = i\}\cdot P\{N'_j = i'\}\,.$$

The probabilities on the right can be estimated by (3.10) of [1]:

$$\begin{aligned}P\{N_j = i\} &\doteq \frac{1}{\sqrt{2\pi kp_j}}\exp\left\{-\frac{1}{2}\frac{(i-kp_j)^2}{kp_j}\right\} \\ &\sim \frac{1}{\sqrt{kp_j}}\,\varepsilon\left(\log\frac{1}{\varepsilon}\right).\end{aligned}$$

See (2.11). Likewise,

$$P\{N_{j'} = i'\} \sim \frac{1}{\sqrt{kp_{j'}}}\varepsilon\left(\log\frac{1}{\varepsilon}\right).$$

In the inner sum of (3.27), the index i determines i' by (3.26); so the inner sum is

$$O\left[\frac{\sqrt{kp_j}}{\sqrt{\log 1/\varepsilon}}\cdot\frac{1}{\sqrt{kp_j}}\cdot\frac{1}{\sqrt{kp_{j'}}}\varepsilon^2\left(\log\frac{1}{\varepsilon}\right)^2\right].$$

But $kp_{j'} \gg (\log 1/\varepsilon)^3$ by (1.19), so the inner sum is $o(\varepsilon^2)$. Coming to the outer sum, the number of terms is $O((1/\varepsilon^2)/\log 1/\varepsilon)$.

Returning to rigorous argument, the main estimate is the following.

LEMMA 3.28. *Assume* (1.1)-(1.23). *Fix positive, finite numbers a and b. Then, uniformly over pairs of indices $j \neq j'$ in I_{-aa}, and values i, i' satisfying*

$$kp_j + \sqrt{kp_j}\sqrt{2\log\frac{1}{\varepsilon}}\lambda_{nj}(-b) \leqq i \leqq kp_j + \sqrt{kp_j}\sqrt{2\log\frac{1}{\varepsilon}}\lambda_{nj}(b) \tag{3.29}$$

$$kp_{j'} + \sqrt{kp_{j'}}\sqrt{2\log\frac{1}{\varepsilon}}\lambda_{nj'}(-b) \leqq i' \leqq kp_{j'} + \sqrt{kp_{j'}}\sqrt{2\log\frac{1}{\varepsilon}}\lambda_{nj'}(b) \tag{3.30}$$

we have

$$P\{N_j = i\} = O\left[(kp_j)^{-1/2}\varepsilon\log\frac{1}{\varepsilon}\right] \tag{3.31}$$

$$P\{N_{j'} = i' \mid N_j = i\} = O\left[(kp_{j'})^{-1/2}\varepsilon\log\frac{1}{\varepsilon}\right]. \tag{3.32}$$

Proof. The first assertion (3.31) is more or less immediate from (3.10) of [1], and the present (2.16). Details are omitted; also see (3.4) of [1]. For the second assertion (3.32), given $N_j = i$, the conditional distribution of $N_{j'}$ is binomial with success probability $\hat{p} = p_{j'}/(1 - p_j)$ and number of trials $\hat{k} = k - i$. Some preliminary estimates are needed before appealing to (3.10) of [1]. All "O" and "o" estimates are uniform over j, j', i and i' satisfying the conditions of the lemma.

It will be shown that

$$\hat{k}\hat{p} = kp_{j'}\left[1 + O\left(k^{-1/2}\left(\log\frac{1}{\varepsilon}\right)^{-3/2}\right)\right] \tag{3.33}$$
$$= kp_{j'}\left[1 + o\left(1\Big/\log\frac{1}{\varepsilon}\right)\right]$$

$$(i' - \hat{k}\hat{p})^2/\hat{p}\hat{k} = 2\log\frac{1}{\varepsilon} - 2\log\log\frac{1}{\varepsilon} + O(1)\,. \tag{3.34}$$

Granting (3.33)-(3.34), the second assertion too follows from (3.10) of [1].

Turn now to proving (3.33)-(3.34). First,

$$\hat{k}\hat{p} - kp_{j'} = (kp_j - i)p_{j'}/(1 - p_j)\,.$$

Now $p_j \to 0$ by (1.17), so $1/(1 - p_j) = O(1)$. Next, $\lambda_{nj}(x) = O(1)$ uniformly over j with $|t_{nj} - t_n| < \delta$ and x with $|x| \leqq b$, by (2.16a). So

$$i - kp_j = O\left[\sqrt{kp_j}\sqrt{2\log\frac{1}{\varepsilon}}\right] \tag{3.35}$$

and

$$(\hat{k}\hat{p} - kp_{j'})/kp_{j'} = O\left[\sqrt{p_j}\sqrt{\log\frac{1}{\varepsilon}}\Big/\sqrt{k}\right]\,. \tag{3.36}$$

Now (1.20) proves (3.33).

For (3.34),

$$i' - \hat{k}\hat{p} = (i' - kp_{j'}) + (kp_{j'} - \hat{k}\hat{p})$$

so

$$(i' - \hat{k}\hat{p})^2 = (i' - kp_{j'})^2 + 2(i' - kp_{j'})(kp_{j'} - \hat{k}\hat{p}) + (kp_{j'} - \hat{k}\hat{p})^2\,.$$

Now

$$(i' - kp_{j'})^2/\hat{k}\hat{p} = [(i' - kp_{j'})^2/kp_{j'}]\left[1 + o\left(1\Big/\log\frac{1}{\varepsilon}\right)\right]$$
$$= \left[2\log\frac{1}{\varepsilon} - 2\log\log\frac{1}{\varepsilon} + O(1)\right]\left[1 + o\left(1\Big/\log\frac{1}{\varepsilon}\right)\right]$$
$$= 2\log\frac{1}{\varepsilon} - 2\log\log\frac{1}{\varepsilon} + O(1)$$

where (3.33) was used in the first line, condition (3.30) and estimate (2.16b-c) of λ_{nj} in the second. Likewise,

$$(i' - kp_{j'})(kp_{j'} - \hat{k}\hat{p})/\hat{k}\hat{p} \sim (i' - kp_{j'})(kp_{j'} - \hat{k}\hat{p})/kp_{j'} \text{ by (3.33)}$$
$$\sim \sqrt{p_{j'}p_j}\cdot\log\frac{1}{\varepsilon} \text{ by (3.35-6)}$$
$$\to 0 \quad \text{by (1.17)}$$

Finally,

$$\begin{aligned}(kp_{j'} - \hat{k}\hat{p})^2/\hat{k}\hat{p} &\sim (kp_{j'} - \hat{k}\hat{p})^2/kp_{j'} \quad \text{by (3.33)}\\ &\sim p_{j'}p_j \log \frac{1}{\varepsilon} \quad \text{by (3.36)}\\ &\sim 0 \quad \text{by (1.17)}\end{aligned}$$ □

4. A generalization. Some cases of interest do not quite fit into the framework of conditions (1.4–15). Assume (1.1–3), for j's such that t_{nj} falls in the compact interval I. Let t_n and t_∞ be interior points of I, with $t_n \to t_\infty$. Let α_n, α'_n and β_n, β'_n be numbers, for $1 \leqq n < \infty$; let α''_∞ and β''_β be numbers. For any positive η_1 and η_2, suppose there exist positive, finite numbers $\delta_1 = \delta_1(\eta_1, \eta_2)$ and $n^* = n^*(\eta_1, \eta_2)$ such that $|t_{nj} - t_n| < \delta_1$ and $n > n^*$ entail

$$\begin{aligned}&\left|\alpha_{nj} - \alpha_n - \alpha'_n\cdot(t_{nj} - t_n) - \frac{1}{2}\alpha''_\infty\cdot(t_{nj} - t_n)^2\right|\\ &\qquad < \eta_1\cdot(t_{nj} - t_n)^2 + \eta_2\Big/\log\frac{1}{\varepsilon_n}\end{aligned} \tag{4.1}$$

$$\begin{aligned}&\left|\beta_{nj} - \beta_n - \beta'_n\cdot(t_{nj} - t_n) - \frac{1}{2}\beta''_\infty(t_{nj} - t_n)^2\right|\\ &\qquad < \eta_1\cdot(t_{nj} - t_n)^2 + \eta_2\Big/\log\frac{1}{\varepsilon_n}\end{aligned} \tag{4.2}$$

Suppose furthermore

$$\alpha_n \to \alpha_\infty \text{ positive, and } \alpha'_n \to \alpha'_\infty, \text{ both finite} \tag{4.3}$$

$$\alpha'_n + \beta'_n = 0 \tag{4.4}$$

$$\alpha''_\infty + \beta''_\beta < 0\,. \tag{4.5}$$

Let

$$\rho^2 = -(\alpha''_\infty + \beta''_\infty)/\alpha_\infty > 0\,. \tag{4.6}$$

THEOREM 4.7. *Suppose* (1.1–3) *and* (1.16–23), *but* (4.1–5) *in place of* (1.4–15). *Let* $n \to \infty$. *There is some small positive* δ *such that, with probability approaching one,* $M_n = \max_j [V_{nj}\colon |t_{nj} - t_n| < \delta]$ *is assumed at a unique index* L_n. *Furthermore, the chance that*

$$\rho\sqrt{2\log\frac{1}{\varepsilon_n}}[\varepsilon_n(L_n - c_n) - t_n] < y$$

and

$$M_n < \alpha_n w_n(x) + \beta_n\sqrt{2\log\frac{1}{\varepsilon_n}}$$

converges to

$$\Phi(y)\cdot\exp\left[-\frac{1}{2\rho}e^{-(1/2)x}\right].$$

The prrof is omitted, being identical to that for (1.24). Note, however, that (4.7) gives no control over j's with t_{nj} remote from t_∞. If, for example, $\alpha_{nj}=\alpha(t_{nj})$ and $\beta_{nj}=\beta(t_{nj})$ where α and β are smooth functions, then (4.7) can be used separately in the vicinity of each local maximum of $\alpha+\beta$.

References

1. P. Diaconis and D. Freedman, *On the maximum difference between the empirical and expected histograms for sums*, Pacific J. Math., **100** (1982a), 287–327.
2. ———, *On the mode of an empirical histogram for sums.*, Pacific J. Math., **100** (1982d), 373–385.
3. ———, *On the maximum difference between a histogram for data and the underlying density*, Z. Wahrscheinlichkeitstheorie, **58** (1981), 139–169.
4. ———, *On the maximum difference between the empirical and expected histograms for sums, Part II*, Pacific J. Math., **100** (1982c), 359–371.
5. D. Freedman, *Another note on the Borel-Cantelli lemma and the strong law, with the Poisson approximation as a by-product*, Ann. Prob., **1**, No. 6, (1973), 920–925.
6. C. Mallows, *An inequality involving multinomial probabilities*, Biometrika, **55** (1968), 422–424.

Received January 7, 1981. Research partially supported by National Science Foundation Grant MCS-77-01665.

University of California
Berkeley, CA 94720