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

# ON THE MODE OF AN EMPIRICAL HISTOGRAM FOR SUMS

PERSI DIACONIS AND DAVID FREEDMAN

**Suppose $S_n$ is a sum of independent, identically distributed random variables, which are integer-valued, with span 1, and have finite fourth moment. If $n$ is large, $S_n$ is approximately normal. An empirical histogram for $k$ copies of $S_n$ will be close to the normal curve provided $k \gg \sqrt{n}\log n$. Suppose now that $\sqrt{n}(\log n)^3 \ll k \ll n^{5/2}$. The object of this paper is to determine the asymptotic joint distribution of the location and size of the mode of this histogram. With overwhelming probability, the mode is unique. Its location and size are asymptotically independent. The location is asymptotically normal, while the size is asymptotically double-exponential. For other $k$'s, the behavior changes. Likewise, the behavior changes if the third moment is finite but the fourth moment infinite.**

**1. Introduction.** The central limit theorem is often used to explain the approximate normality of an empirical histogram. However, even if a random variable $S_n$ is approximately normal because it is a sum of $n$ independent random variables, further theory is required to explain the global closeness of a histogram constructed from $k$ independent copies of $S_n$ to the normal density. As shown in [4], if $n$ and $k$ go to infinity in such a way that $k/(\sqrt{n}\log n) \to \infty$, then the largest deviation of the histogram from the normal curve tends to 0. If the histogram is close to the normal curve, then the maximum of the histogram should be close to the maximum of the normal curve. In this paper, the object is to obtain the joint distribution of the location and size of the maximum of such a histogram. Under suitable conditions on $n$ and $k$:

· with high probability, the maximum of the histogram is taken on at a unique location;

· the size of the maximum is independent of the location of the maximum;

· suitably normalized, the location of the maximum is normally distributed and the size of the maximum has a double-exponential distribution.

To be more specific, suppose the $X_i$ are independent, identically distributed, integer-valued, and have span 1

$$(1.1)\quad \text{g.c.d. } \{j - k: j, k \in C\} = 1, \text{ where } j \in C \text{ iff } P(X_1 = j) > 0 .$$

Suppose the fourth moment is finite:

$$E\{X_1^4\} < \infty \ . \tag{1.2}$$

Let

$$\mu = E\{X_1\} \text{ and } \sigma^2 = \operatorname{Var}(X_1) \ . \qquad S_n = X_1 + \cdots + X_n \ . \tag{1.3}$$

Thus, $(S_n - n\mu)/\sigma\sqrt{n}$ is approximately normal. Take $k$ independent copies of $S_n$, and make an empirical histogram for these $k$ numbers. In [4] it was shown that if $k$ and $n$ approach infinity in such a way that $k/\sqrt{n}\ \log n \to \infty$, the empirical histogram converges uniformly to the normal curve. If $k/\sqrt{n} \to \infty$ but $k = 0(\sqrt{n}\log n)$, the histogram was shown to converge pointwise but not uniformly. Finally, if $k = 0(\sqrt{n})$, the histogram does not even converge pointwise.

This result is refined in [6], which obtains the joint distribution of the location and size of the maximum derivation between the empirical histogram and the probability histogram, using the growth condition

$$k/\sqrt{n}\,(\log n)^3 \longrightarrow \infty \ . \tag{1.4}$$

This paper will borrow several results from [5] and [6].

To state the main result of this paper, let $N_j$ be the number of copies of $S_n$ which are equal to $j$. Up to scaling, $N_j$ is the empirical histogram for the $k$ sums. Let

$$\rho = (2\pi)^{-1/8}\sigma^{-5/4} \tag{1.5}$$

$$l = k/\sigma\sqrt{2\pi n} \tag{1.6}$$

$$m = n^{5/8}/k^{1/4} \tag{1.7}$$

$$\varepsilon_n = m^{-1}(2\log m)^{-1/4} \tag{1.8}$$

$$w_n(x) = \left(2\log\frac{1}{\varepsilon_n} - 2\log\log\frac{1}{\varepsilon_n} + x\right)^{1/2} \tag{1.9}$$

$$\Phi(y) = (2\pi)^{-1/2}\int_{-\infty}^{y}\exp\left(-\frac{1}{2}u^2\right)du \ . \tag{1.10}$$

The main result can now be stated; the proof is deferred to the next section.

THEOREM 1.11. *Assume* (1.1–10). *Let $k$ and $n$ tend to infinity, with $k \ll n^{5/2}$. With probability approaching one,* $M_n = \max_j N_j$ *is taken on at a unique index $L_n$. Furthermore, the chance that*

$$\rho\sqrt{2\log\frac{1}{\varepsilon_n}}\varepsilon_n(L_n - n\mu) < y \textit{ and } (M_n - l)/\sqrt{l} < w_n(x)$$

*converges to*

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

What is the role of the regularity conditions? Assumption (1.2) is that the fourth moment be finite. This can be relaxed somewhat, but preliminary calculations suggest that the conclusions of 1.11 can fail if only a third moment is assumed. For a related discussion, see § 4 of [6].

Assumption (1.4) governs the rate at which $k$ and $n$ tend to infinity. If $\sqrt{n}\log n \ll k = 0[\sqrt{n}(\log n)^3]$, the conclusions of (1.11) fail: in essence, the scale $w_n(x)$ defined by (1.9) must be revised to account for large-deviations corrections to the central limit theorem. This can be accomplished using an expansion developed by Kolchin, Sevastyanov and Chistyakov (1978), in Lemma 5 of their § II.6. For more details see [2].

If $k$ is of order $\sqrt{n}\log n$, the situation changes radically. The maximum will not in general be assumed at a unique location, and its distribution does not converge, but oscillates. For details, see [2]. Related phenomena are discussed in Anderson (1970) or Iglehart (1977). We plan to explore the case $k = 0[\sqrt{n}(\log n)^3]$ elsewhere.

At the other end of the spectrum, if $k$ is of order $n^{5/2}$, the location and size of the maximum are no longer asymptotically independent; and the asymptotic distribution of the location is discrete. If $k$ grows faster than $n^{5/2}$, the maximum can occur only at one or two locations, with probabilities depending on the arithmetic properties of $\mu$, and on higher moments. This will be discussed in § 3.

2. *The proof*. The object in this section is to prove Theorem 1.11, but first, some heuristics. Let $A$ be a large positive constant and $\delta$ a small positive constant, to be chosen later. It is convenient to distinguish three zones:

$$\text{the inner zone, } |j - n\mu| \leqq Am(2\log m)^{1/4} \tag{2.1}$$

$$\text{the midzone, } Am(2\log m)^{1/4} < |j - n\mu| \leqq \delta\sigma\sqrt{n} \tag{2.2}$$

$$\text{the outer zone, } |j - n\mu| > \delta\sigma\sqrt{n}\,. \tag{2.3}$$

Only the inner zone contributes to the maximum, as will be shown later. The inner zone can be handled using [5], but some effort is needed to get into that framework. Clearly,

$$N_j - l = N_j - kp_j + kp_j - l = \sqrt{l}\,[\alpha_{nj} Z_{nj} + \gamma_{nj}] \tag{2.4}$$

where $l$ was defined in (1.6) and $p_j = P\{X_1 = j\}$ and

$$\alpha_{nj}^2 = \sigma\sqrt{2\pi n}\, p_j \tag{2.5}$$

$$\gamma_{nj} = \sqrt{l}\,[\sigma\sqrt{2\pi n}\, p_j - 1] \tag{2.6}$$

$$Z_{nj} = (N_j - k_{pj})/\sqrt{kp_j}\,. \tag{2.7}$$

The Edgeworth expansion shows that

$$\alpha_{nj} \doteq 1 \tag{2.8}$$

$$\gamma_{nj} \doteq -\frac{1}{2}\rho^2 (j - n\mu)^2/m^2 \tag{2.9}$$

where $\rho$ was defined in (1.5) and $m$ in (1.7). To get into the framework of [5], choose $\varepsilon_n$ so that

$$\varepsilon_n^2 \sqrt{2 \log \frac{1}{\varepsilon_n}} \doteq 1/m^2\,; \tag{2.10}$$

this is the motivation for (1.8). Of course, $\varepsilon_n \to 0$ because $m = n^{5/8}/k^{1/4} \to \infty$, due to the assumption that $k \ll n^{5/2}$. In [5, (1.1)], take

$$\beta_{nj} = \gamma_{nj}\Big/\sqrt{2 \log \frac{1}{\varepsilon_n}}\,. \tag{2.11}$$

The center $c_n$ is $n\mu$, so in [5, (1.3)]

$$t_{nj} = \varepsilon_n (j - n\mu)\,. \tag{2.12}$$

In [5, (1.4)], put $\alpha_n(t) \equiv 1$. In [5, (1.5)]

$$\beta_n(t) = \gamma_n \beta(t) \tag{2.13}$$

where

$$\gamma_n^2 = \log m/\log \frac{1}{\varepsilon_n} \longrightarrow 1 \tag{2.14}$$

and

$$\beta(t) = -\frac{1}{2}\rho^2 t^2\,. \tag{2.15}$$

Clearly, $\beta_n$ and $\beta$ take their maximum at $t = 0$, where they vanish.

For I, take the interval $[-A, A]$. Conditions [5, (1.6–23)] are easily verified, [5, (1.19)] being the present (1.4): note that

(2.16) For any $\varepsilon > 0$, there is a $\delta > 0$ such that for $n$ large, $|j - n\mu|/\sigma\sqrt{n} < \delta$ entails $(1-\varepsilon) < \sigma\sqrt{2\pi n}p_j < 1 + \varepsilon$ .

Now [5, (1.24)] establishes the conclusions of the present Theorem 1.11, once conditions [5, (1. 4–5)] are verified. That is the point of the next lemma.

LEMMA 2.17. *Uniformly over $j$ with $t_{nj} \in I$,*
(a) $\alpha_{nj} = 1 + o(1/\log 1/\varepsilon_n)$
(b) $\beta_{nj} = \beta_n(t_{nj}) + o(1/\log 1/\varepsilon_n)$.

*Proof*. This follows from the Edgeworth expansion. As $n \to \infty$, uniformly in $j$, because there is a fourth moment and the span is 1,

$$\sigma\sqrt{2\pi n}\, p_j = \exp\left(-\frac{1}{2}x_{nj}^2\right)[1 + aH_3(x_{nj})/\sqrt{n}] + 0(1/n) \tag{2.18}$$

where

$$x_{nj} = (j - n\mu)/\sigma\sqrt{n}$$
$$a = \mu_3/6\sigma^3$$
$$\mu_3 = E[(X_1 - \mu)^3]$$
$$H_3(x) = x^3 - 3x \ .$$

For a discussion of this result, see page 205 of Petrov (1975).

The argument for claim (a) is relatively easy and is omitted. For claim (b), recall (2.6) and (2.11). An error $\eta$ in estimating $\sigma\sqrt{2\pi n}p_j$ is harmless, provided

$$\sqrt{l}\,\eta\Big/\sqrt{2\log\frac{1}{\varepsilon_n}} = o\left(1/\log\frac{1}{\varepsilon_n}\right) .$$

In other terms,

$$\eta = o\left(1\Big/\sqrt{l\log\frac{1}{\varepsilon_n}}\right) .$$

or from (1.6) and (1.8),

$$\eta = o[k^{-1/2}n^{1/4}(\log m)^{-1/2}]$$

We claim

$$1/n = o[k^{-1/2}n^{1/4}(\log m)^{-1/2}] \ . \tag{2.19}$$

Squaring and reorganizing, the assertion is that

$$k/n^{5/2} \ll 1/\log m$$

or

$$1/m^4 \ll 1/\log m \ ,$$

proving (2.19). In particular, the $O(1/n)$ in (2.18) is harmless.

We claim

$$x_{nj}/\sqrt{n} = o[k^{-1/2}n^{1/4}(\log m)^{-1/2}] \quad \text{uniformly in } j \text{ with } t_{nj} \in I \ . \tag{2.20}$$

Indeed, $x_{nj} = O[n^{-1/2}m(\log m)^{1/4}]$. Taking fourth powers and reorganizing, the assertion is that

$$1/m^4 = k/n^{5/2} \ll 1/(\log m)^3 \ ,$$

proving (2.20). Thus, the term in $H_3(x)/\sqrt{n}$ is harmless.

Finally, we claim

$$x_{nj}^4 = o[k^{-1/2}n^{1/4}(\log m)^{-1/2}] \quad \text{uniformly in } j \text{ with } t_{nj} \in I \ . \tag{2.21}$$

This boils down to the assertion

$$k \gg \sqrt{n}\,(\log m)^3 \ ,$$

which follows from the growth condition (1.4). So $\exp(-(1/2)x_{nj}^2)$ can be replaced by $1 - (1/2)x_{nj}^2$.

To sum up, (2.18) implies that uniformly over $j$ with $t_{nj} \in I$,

$$\sigma\sqrt{2\pi n}p_j = 1 - (j - n\mu)^2/(2\sigma^2 n) + o[k^{-1/2}n^{1/4}(\log m)^{-1/2}] \ . \tag{2.22}$$

From this, claim (b) is immediate. □

This completes the argument for the inner zone, and shows that $\max_j (N_j - l)/\sqrt{l}$ is of order $w_n(x)$, where $j$ is restricted to the inner zone. We must now deal with the midzone, and show that for any $x$,

$$P\{\max_j (N_j - l)/\sqrt{l} > w_n(x)\} \longrightarrow 0 \ , \quad \text{where } j \text{ is restricted to the midzone (2.2).} \tag{2.23}$$

It will be convenient to make a more general argument, for use in § 3. In particular, $k$ is allowed to be of order $n^{5/2}$ or more, so $m$ may converge to a finite limit, or even to zero.

LEMMA 2.24. *Fix $\varepsilon > 0$. There is a large positive $L$ and a small positive $\delta$ such that: for all large $n$,*

$$L < |j - n\mu| \leqq \delta\sigma\sqrt{n}$$

*entails*

$$(1 - \varepsilon)(j - n\mu)^2/2\sigma^2 n < 1 - \sigma\sqrt{2\pi n}p_j < (1 + \varepsilon)(j - n\mu)^2/2\sigma^2 n .$$

*Proof.* This follows from (2.18). The remainder term $O(1/n)$ can be merged into the $\varepsilon(j - n\mu)^2/2\sigma^2 n$, because $1/n$ is only a small multiple of

$$(j - n\mu)^2/2\sigma^2 n ,$$

because $|j - n\mu|$ is large by assumption. Likewise for the term in $H_3(x_{nj})/\sqrt{n}$:

$$x_{nj}^3/\sqrt{n} \leqq \delta^2 x_{nj}/\sqrt{n}$$

because $|x_{nj}| \leqq \delta$, and $x_{nj}/\sqrt{n}$ is a small multiple of $x_{nj}^2$ by previous reasoning. Finally

$$\left|\exp\left(-\frac{1}{2}x^2\right) - 1 + \frac{1}{2}x^2\right| \leqq \frac{1}{2}\varepsilon x^2$$

for $|x| \leqq \delta$, if $\delta$ is small enough. □

Recall $\gamma_{nj}$ from (2.6), and the definition of $\rho, l$ and $m$ from (1.5–7). With present assumptions, $m$ may converge to 0.

COROLLARY 2.25. *There is a large positive $L$ and a small positive $\delta$ such that: for all large $n$,*

$$L < |j - n\mu| \leqq \delta\sigma\sqrt{n}$$

*entails $0 < \alpha_{nj} < 1$ and $\gamma_{nj} < -\rho^2(j - n\mu)^2/4m^2$.*

COROLLARY 2.26. *Fix $x$ and $\beta$ nonnegative. There is a large positive $L$ and a small positive $\delta$ such that: for all large $n$,*

$$L < |j - n\mu| \leqq \delta\sigma\sqrt{n}$$

*entails*

$$P\{\alpha_{nj}Z_{nj} + \gamma_{nj} > -x - m^{-2}\beta\} \leqq \exp\{-\lambda^4/256\} ,$$

*where*

$$\lambda = \rho|j - n\mu|/m .$$

*Proof.* As 2.25 implies, $-\gamma_{nj} > \lambda^2/4$. For $L$ large, $\lambda^2/8 > x + m^{-2}\beta$. Then

$$\alpha_{nj}Z_{nj} + \gamma_{nj} > -x - m^{-2}\beta$$

entails $Z_{nj} > \lambda^2/8$. By [3], this last has probability at most

$$\exp\left\{-\frac{1}{128}\lambda^4\Big/\left[1 + \frac{1}{8}(kp_j)^{-1/2}\lambda^2\right]\right\}.$$

We must now get rid of the term $(kp_j)^{-1/2}\lambda^2$. In view of (2.16), there is a $\delta$ so small that $kp_j > (1/2)k/\sigma\sqrt{2\pi n}$ for all $j$ with $|j - n\mu| \leqq \delta\sigma\sqrt{n}$. Then

$$\lambda^2 \leqq (\rho^2/m^2)\delta^2\sigma^2 n$$

and

$$(kp_j)^{-1/2}\lambda^2 \leqq 2^{1/2}(2\pi)^{1/4}\sigma^{5/2}\delta^2 \leqq 8$$

for small $\delta$. □

We can now prove a result sharper than (2.23).

LEMMA 2.27. *Suppose $k \ll n^{5/2}$, so $m \to \infty$. Then $\sum_j P\{N_j > l\} \to 0$, where $j$ is restricted to the midzone* (2.2), *provided $A$ is sufficiently large and $\delta$ is sufficiently small.*

*Proof.* Start from the identities (2.4–7). The idea is to bound

$$\sum_j P\{\alpha_{nj}Z_{nj} + \gamma_{nj} > 0\},$$

using (2.26) with $x = \beta = 0$. The bound is

$$\text{(2.28)} \quad \sum_j \exp\{-\rho^4(j - n\mu)^4/256m^4\} \leqq 2m\int_{A_m} \exp\{-\rho^4u^4/256\}du$$

where for $m$ large,

$$A_m = A(2\log m)^{1/4} - m^{-1} > \frac{1}{2}A(2\log m)^{1/4}.$$

To verify (2.28), let $j_0$ be the least integer exceeding

$$n\mu + A_m(2\log m)^{1/4}.$$

The sum on the left includes all $j \geqq j_0$; the rest of the $j$'s are similar and will not be discussed. Due to monotonicity, the $j$th term in the sum is at most

$$\int_{j-1}^{j} \exp\{-\rho^4(t-n\mu)^4/256m^4\}dt\ ,$$

so our half of the sum is at most

$$\int_{j_0-1}^{\infty} \exp\{-\rho^4(t-n\mu)^4/256m^4\}dt\ .$$

Now make the change of variables $u = (t - n\mu)/m$. To upper bound the right hand side of (2.28), multiply the integrand by $u^3 \geqq 1$ for $m$ large. Provided $\rho^4 A^4 > 2^3 \cdot 256$, the bound tends to zero. □

Finally, we must dispose of the outer zone (2.3), and it is convenient to do this even if $k$ is of order $n^{5/2}$ or bigger.

LEMMA 2.29. *For any* $\delta > 0$, *there is a* $\theta_\delta < 1$ *such that, confining* $j$ *to the outer zone* (2.3),

$$P\{\max_j N_j < \theta_\delta l\} \longrightarrow 1\ .$$

*Proof.* The basic idea is that the empirical histogram is close to the normal curve, and hence falls off quite rapidly. To be more precise, define $x_{nj}$ as in (2.18). From that result, or the local Berry Esseen theorem, there is a $C < \infty$ such that

$$\left|\sigma\sqrt{2\pi n}p_j - \exp\left(-\frac{1}{2}x_{nj}^2\right)\right| < C/\sqrt{n} \quad \text{for all } j\ . \tag{2.30}$$

Fix $D > \sigma^{-1/2}(2\pi)^{-1/4}$. By [3, (5)], with probability approaching one,

$$N_j < kp_j + T \text{ for all } j, \text{ where } T = Dk^{1/2}(\log n)^{1/2}/n^{1/4}\ . \tag{2.31}$$

We will use (2.30) to force the right hand side of (2.31) below $l$, for all $j$ in the outer zone. Indeed, $|x_{nj}| \geqq \delta$, so

$$\exp\left\{-\frac{1}{2}x_{nj}^2\right\} \leqq \theta = \exp\left\{-\frac{1}{2}\delta^2\right\} < 1\ ,$$

and from (2.30),

$$kp_j \leqq \theta l + \sigma^{-1}(2\pi)^{-1/2}Ck/n\ .$$

But $k/n \ll l$, for the latter is of order $k/\sqrt{n}$. And $T \ll l$ in (2.31), because $k \gg \sqrt{n} \log n$. □

**3. When $k$ is large.** The object in this section is to indicate how Theorem 1.11 breaks down when $k$ is of order $n^{5/2}$ or larger. To state the first result, let

$$\begin{aligned} &\gamma(t) = \frac{1}{2}\rho^2(-t^2 + ct + d), \text{ where} \\ &c = \mu_3/3\sigma^2 \text{ and } d = (\mu_4 - 3\sigma^4)/4\sigma^2 \\ &\mu_i = E\{(X_1 - \mu)^i\} \text{ and } \mu = E\{X_1\}\ . \end{aligned} \tag{3.1}$$

Let $W_i$ be independent, with common $N(0, 1)$ distribution, for $i=0, \pm1, \pm2, \cdots$.

(3.2) For real $\alpha$, let $L_{\lambda\alpha}$ and $M_{\lambda\alpha}$ be the location and size respectively of

$$\max_i \{W_i + \lambda^{-2}\gamma(i - \alpha)\}\ .$$

(3.3) Let $I(x)$ and $F(x)$ denote the integer part and fractional part of $x$.

THEOREM 3.4. *Assume* (1.1–7), *except that* (1.4) *is replaced by the condition that $k/n^{5/2}$ converges to a finite positive limit* $1/\lambda^4$. *Thus, $m \to \lambda$. Suppose* (3.1–3). *Suppose, by passing to a subsequence if necessary, that $F(n\mu) \to \alpha$. With probability approaching one, $M_n = \max_j N_j$ is taken on at a unique index $L_n$. Furthermore, the joint distribution of $L_n - I(n\mu)$ and $(M_n - l)/\sqrt{l}$ converges weakly to that of $L_{\lambda\alpha}$ and $M_{\lambda\alpha}$.*

*Proof.* The argument will only be sketched. Fix a large, positive number $L$. The $j$'s which count are those satisfying $|j - n\mu| \leqq L$. Refer back to (2.4–7). For $j = I(n\mu)+i$ and $|i|\leqq L$, the $Z_{nj}$ are asymptotically distributed like the $W_i$. This follows from (3.17) below.

The Edgeworth expansion (2.18) can be taken out to the term of order $1/n$, which cannot be dropped; but the remainder $o(1/n)$ is negligible. The conclusion: for $|j - n\mu| \leqq L$,

$$\begin{aligned} &\alpha_{nj} \longrightarrow 1 \\ &\gamma_{nj} - \lambda^{-2}\gamma(j - n\mu) \longrightarrow 0\ . \end{aligned} \tag{3.5}$$

Clearly, for $|j - n\mu| \leqq L$,

$$\gamma(j - n\mu) - \gamma(j - I(n\mu) - \alpha) \longrightarrow 0\ . \tag{3.6}$$

Thus, the joint distribution of

$$\alpha_{nj} Z_{nj} + \gamma_{nj}\colon\ j = I(n\mu) + i, \quad |i| \leqq L \tag{3.7}$$

converges to that of

$$W_i + \lambda^{-2}\gamma(i - \alpha)\colon |i| \leqq L\ . \tag{3.8}$$

This completes the argument for $j$'s with $|j - n\mu| \leqq L$. And $\max_j (N_j - l)/\sqrt{l}$ over such $j$'s has a proper limiting distribution. What remains is to show that $j$'s with $|j - n\mu| > L$ do not contribute to the maximum, with probability approaching one as $L \to \infty$. For $j$'s with $|j - n\mu| > \delta\sigma\sqrt{n}$, Lemma 2.29 applies. For $j$'s with

$$L < |j - n\mu| \leqq \delta\sigma\sqrt{n}\ , \tag{3.9}$$

Corollary 2.26 can be used. Let $0 < x < \infty$. We have to show that

$$\sum_j P\{\alpha_{nj} Z_{nj} + \gamma_{nj} > -x\} \tag{3.10}$$

is small, where $j$ is restricted to satisfy (3.9). Use (2.26) with $\beta = 0$, to bound (3.10) by

$$2\int_{L-1} \exp\{-\rho^4 u^4/256 m^4\} du\ . \tag{3.11}$$

This is small for $L$ large. □

*Note.* $L_{\lambda\alpha}$ is discrete; $L_{\lambda\alpha}$ and $M_{\lambda\alpha}$ are dependent. Thus, the behavior is qualitatively different from that described in 1.11. It is also interesting that different subsequences can produce different limits, due to the presence of $\alpha$, the limiting fractional part of $n\mu$.

When $k$ increases faster than $n^{5/2}$, the situation changes again.

THEOREM 3.12. *Assume* (1.1–7), *except that* (1.4) *is replaced by the condition*

$$n^{5/2} \ll k \ll n^{7/2}\ ;$$

*and* (1.2) *is strengthened to* $E\{|X_1|^5\} < \infty$. *Suppose* (3.1–3). *Suppose, by passing to a subsequence, that* $\gamma(j - n\mu)$, *as a function of the integer* $j$, *takes its maximum at the unique integer* $j = j_n$. *Then, with probability approaching one,* $\max_j N_j$ *is taken on at* $j = j_n$.

*Proof.* The argument, like that in (3.4), is only sketched. The Edgeworth expansion (2.18) must be carried out to the term in $1/n$, with a remainder $O(1/n^{3/2})$ which is negligible. Confine $j$ to the range $|j - n\mu| \leqq L$. Then $\alpha_{nj} \to 1$ and

$$\gamma_{nj} - \gamma(j - n\mu)/m^2 \longrightarrow 0 .$$

The $Z_{nj}$ are asymptotically independent standard normals, as in 3.4. An elementary argument shows that for $|j-n\mu| \leqq L$, $\max_j \{\alpha_{nj}Z_{nj} + \gamma_{nj}\}$ is assumed at $j_n$, with probability approaching one. This max is essentially

$$\beta_n/m^2 + W + o(1) \tag{3.13}$$

where $\beta_n = \gamma(j_n - n\mu)$ is bounded, and $W$ is standard normal. Note that $\beta_n$ may be positive, zero, or negative, and $m > 0$ but $m \to 0$.

In any case, for $|j - n\mu| \leqq L$, $\max_j N_j$ is essentially

$$l + \sqrt{l}\{\beta_n/m^2 + W + o(1)\} . \tag{3.14}$$

Since $\sqrt{l}/m^2 \ll l$, the display (3.14) is of order $l$, and $j$'s with $|j - n\mu| > \delta\sigma\sqrt{n}$ do not contribute to the maximum, by (2.29).

This leaves only the problem of eliminating $j$'s with

$$L < |j - n\mu| \leqq \delta\sigma\sqrt{n} . \tag{3.15}$$

It is enough to prove that for any positive $x$ and $\beta$,

$$\sum_j P\{\alpha_{nj}Z_{nj} + \gamma_{nj} > -x - m^{-2}\beta\} \tag{3.16}$$

is small for $L$ large, $j$ being restricted to the midzone (3.15). This can be argued as in (3.4). Since $m \to 0$, the expression (3.11) tends to zero for any $L > 1$. However, the bound in (2.25) is valid only for large $L$, thus (3.11) can be used as a bound on (3.16) only for large $L$. □

If $\gamma(j - n\mu)$ takes its maximum at two $j$'s, then $\max_j N_j$ can be assumed at either one, with probabilities computable from the Edgeworth expansion. Likewise, if $k$ is $n^{9/2}$ or larger, more moments are needed, and more terms in the Edgeworth expansion.

It may be useful to give 3.17 in a bit more generality. Let $J$ be a finite set, and $f \notin J$. Let $\pi_{nj}$: $j \in J \cup \{f\}$ be positive numbers whose sum is one. Let $k_n$ be a positive integer. Let $M_{nj}$: $j \in J \cup \{f\}$ be multinomial, with parameters $k_n$ and $\pi_{nj}$. That is, $k_n$ balls are dropped independently into boxes labelled by $J \cup \{f\}$; each ball lands in box $j$ with probability $\pi_{nj}$; and $M_{nj}$ is the total number in box $j$. Let $W_{nj} = (M_{nj} - k_n\pi_{nj})/\sqrt{k_n\pi_{nj}}$.

PROPOSITION 3.17. *Suppose $\pi_{nj} \to 0$ and $k_n\pi_{nj} \to \infty$ for each $j \in J$. Then the joint distribution of $W_{nj}$: $j \in J$ converges weakly to* *that of independent standard normals.*

*Proof.* Fix constants $c_j$, and consider $S = \sum_{j \in J} c_j W_{nj}$. As is easily verified, $S$ is the sum of $k_n$ independent, identically distributed random variables, each bounded by $\sum_{j \in J} |c_j| / \sqrt{k_n \pi_{nj}} \to 0$: there is one variable in the sum for each ball. An elementary computation shows that $E(S) = 0$ and

$$\operatorname{Var} S = \sum_{j \in J} c_j^2 - \eta$$

where

$$\eta = 2 \sum c_j c_{j'} \sqrt{\pi_{nj} \pi_{nj'}}\,,$$

the last sum extending over $j \neq j'$, both in $J$. But $\eta \to 0$ because $\pi_{nj} \to 0$. Now $S$ is asymptotically normal, with mean 0 and variance $\sum c_j^2$, for instance by the Berry-Esseen bound. □

## REFERENCES

1. C. W. Anderson, *Extreme value theory for a class of discrete distributions with applications to some stochastic processes*, J. Appl. Prob., **7** (1970), 99-103.

2. P. Diaconis and D. Freedman, *The Distribution of the Mode of an Empirical Histogram*, Bell Laboratories Technical Report, 1979.

3. D. Freedman, *Another note on the Borel-Cantelli lemma and the strong law, with the Poisson approximation as a byproduct*, Ann. Prob., **6** (1973), 910-915.

4. ———, *A central limit theorem for empirical histograms*, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 1977.

5. ———, *On the maximum of scaled multinomial variables*, Pacific J. Math., **100** (1982b), 329-358.

6. D. Freedman and P. Diaconis, *On the maximum differences between empirical and expected histograms for sums*, Pacific J. Math., **100** (1982a), 287-327.

7. D. L. Iglehart, *Regenerative simulation for extreme values*, Stanford University, Dept. of Operations Research, Technical Report No. 43, 1977.

8. A. Kolchin, B. Sevastyanov and C. Chistyakov, *Random Allocations*, Wiley, New York, 1978.

9. V. V. Petrov, *Sums of Independent Random Variables*, Springer, New York, 1975.

Received August 8, 1980, First author's research partially supported by NSF Grant MCS 77-01665. Second research partially supported by NSF Grant MPS 74-21416 and ERDA Contract EY-76-C-03-0515.

UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720
AND
STANFORD UNIVERSITY
STANFORD, CA 94305