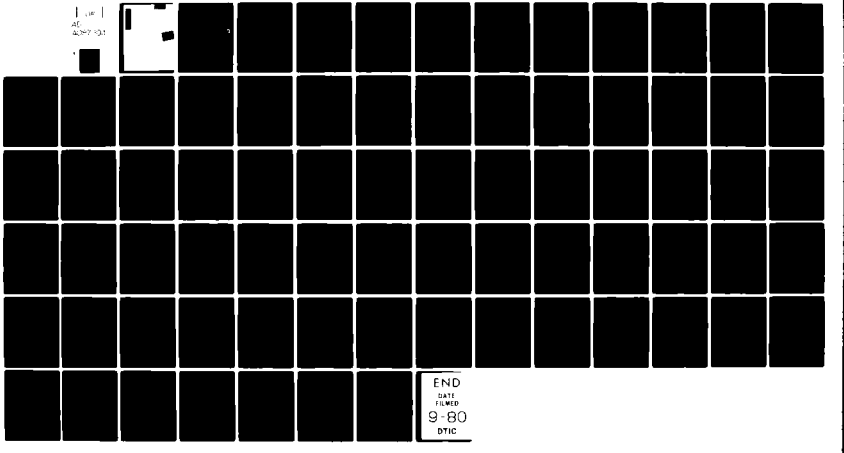
AD-A087 904 STANFORD UNIV CA DEPT OF STATISTICS F/G 12/1
THE DIFFERENCE BETWEEN THE EMPIRICAL HISTOGRAM AND THE PROBABIL--ETC(U)
JUL 80 P DIACONIS, D FREEDMAN N00014-76-C-0475
UNCLASSIFIED TR-288 NL



END
DATE FILMED
9-80
DTIC

12

THE DIFFERENCE BETWEEN THE EMPIRICAL HISTOGRAM AND THE PROBABILITY HISTOGRAM FOR SUMS

By

Persi Diaconis and David Freedman

TECHNICAL REPORT NO. 288

July 2, 1980

Prepared under Contract
N00014-76-C-0475 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

THE DIFFERENCE BETWEEN THE EMPIRICAL HISTOGRAM AND THE PROBABILITY HISTOGRAM FOR SUMS

By

Persi Diaconis, Stanford University
and
David Freedman, University of California, Berkeley[1]

1. Introduction.

The central limit theorem is often used heuristically to justify the approximation of histograms for data by the normal curve. This argument can be made precise through the following model. There is some basic random variable X. Take the sum of n independent copies of X, and then take k independent copies of these sums. Provided X is well-behaved, n is large, and k is large in relation to n, the histogram for the sums will be close to the normal curve. In this way, the model rigorously justifies the use of the normal curve to approximate the data.

In more detail, let $X_1, X_2, \ldots$ be independent, identically distributed random variables. Suppose the X_i are integer-valued and have g.c.d. 1:

$$\text{g.c.d. } \{m: P(X_i = m) > 0\} = 1 . \tag{1.1}$$

Suppose too

$$E|X_1^3| < \infty . \tag{1.2}$$

Let $S_n = X_1 + \cdots + X_n$. Take k independent copies of S_n, and let N_j be the number of these sums which are equal to j: the notation hides the dependence of N_j on n and k. Up to scaling, the counts N_j correspond to the empirical histogram for the k sums. Of course,

Research partially supported by NSF Grant MCS-77-01665.

(1.3) $$E(N_j) = kp_j, \quad \text{where} \quad p_j = P(S_n=j) .$$

Up to scaling, the numbers p_j correspond to the probability histogram for the k sums. When n is large, the local central limit theorem implies that the p_j are uniformly close to an appropriately scaled normal curve. In this paper, we study the random variable

(1.4) $$M_{nk} = \max_j (N_j - kp_j) .$$

Up to scaling, this is the maximum difference between the empirical histogram of the k sums and the probability histogram of the sum. This random variable is the key to understanding the maximum difference between the histogram and the normal curve, as will be shown in a future paper.

We suppose

(1.5) $$n \to \infty \quad \text{and} \quad k/\sqrt{n}\ (\log n)^3 \to \infty .$$

Assumption (1.5) is discussed at the end of this section. Let

(1.6) $$\mu = E(X_1) \quad \text{and} \quad \sigma^2 = \text{Var } X_1 .$$

In Freedman (1977) it is shown that if (1.1-1.6) hold, then

(1.7) $$M_{nk}\sqrt{\frac{k}{\sigma\sqrt{2\pi n}}\ \log n} \to 1 \quad \text{in probability.}$$

Our main object in this paper is to prove the following result, which sharpens (1.7).

(1.8) Theorem. Assume (1.1-1.6). With probability approaching one, $M_{nk} = \max_j (N_j - kp_j)$ is taken on at a unique index L_{nk}. Moreover, L_{nk} and M_{nk} are asymptotically independent, L_{nk} being asymptotically normal and M_{nk} being asymptotically double-exponential. To be more precise, let

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-\tfrac{1}{2}u^2)du \,, \tag{1.9}$$

$$z_n(x) = \sqrt{\frac{k}{\sigma\sqrt{2\pi n}}} \cdot \sqrt{\log n - 2 \log\log n + x} \,. \tag{1.10}$$

Then, the probability that

$$L_{nk} - n\mu < y\sigma\sqrt{2n/\log n} \quad \text{and} \quad M_{nk} < z_n(x)$$

converges to

$$\Phi(y) \quad \exp[-\sigma\sqrt{2}\, e^{-\frac{1}{2}x}] \,.$$

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

This theorem is illustrated in table 1 and figure 1, which reports three computer simulations. In all three, the basic random variable X_1 took the six values 1 through 6 with equal probability 1/6. The table shows the values chosen for n and k, the lower and upper quartiles for

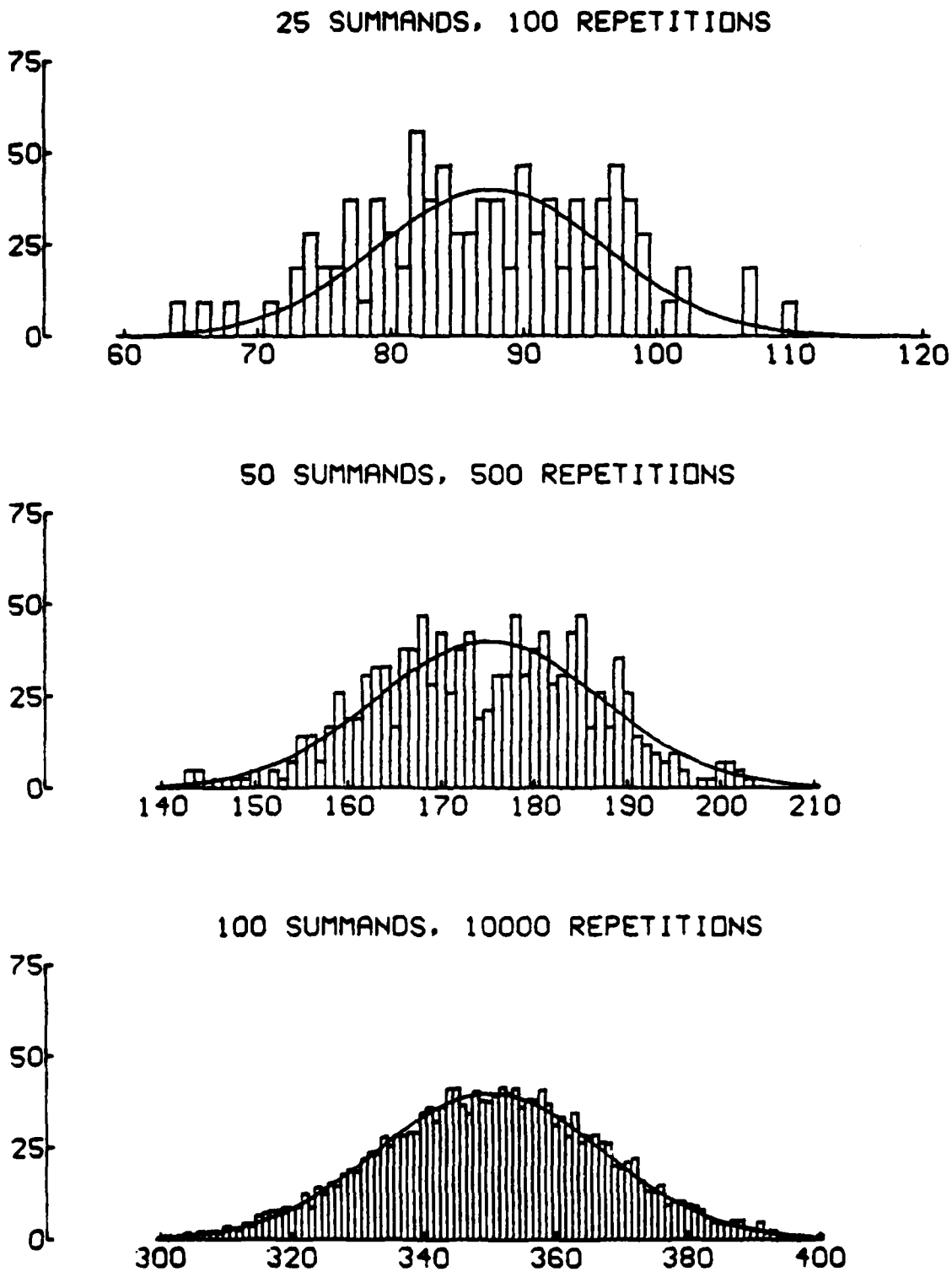
TABLE 1.

Computer simulation illustrating theorem (1.8).

n	k	$k/[\sqrt{n}(\log n)^3]$	lower quartile		upper quartile		observed value	
			L_{nk}	M_{nk}	L_{nk}	M_{nk}	L_{nk}	M_{nk}
25	100	0.6	83	3.0	92	4.9	97	2.5
50	500	1.2	169	6.2	181	9.5	185	8.3
100	10,000	10.2	341	24.9	358	36.8	358	26.9

FIGURE 1.

Computer simulation: the empirical histogram converges to the normal curve.



the asymptotic distributions of L_{nk} and M_{nk}, as computed from the theorem, and the values observed in the simulation.

In the first line of table 1, the location L_{nk} is reported as 97: so $N_j - kp_j$ is largest when $j = 97$. From the figure, the mode of the empirical histogram is at 82: so N_j is largest when $j = 82$. In general, the location of the maximum deviation and the mode are very different. Compare theorem (1.8) with Diaconis and Freedman (1978).

The figure shows the histograms themselves. The horizontal scale shows j, but the three axes are arranged so that the means $n\mu$ line up, and the standard deviations $\sigma\sqrt{n}$ cover the same physical distance. The vertical scale shows $(\sigma\sqrt{n}) \times (N_j/k) \times 100\%$: that is, N_j is converted to a percent (relative to k), and the histograms are rescaled vertically by $\sigma\sqrt{n}$ to have the same physical area. For a discussion of this convention, see pp. 29 ff and 275 ff of Freedman, Pisani, Purves (1978). Thus, the maximum difference between the rescaled empirical and theoretical histograms is

$$\frac{\sigma\sqrt{n}}{k} M_{nk} .$$

In the third line of table 1, for instance, that is about 5%.

In section 2, we argue heuristically that our problem can be reduced to finding the location and size of the maximum of a sequence Z_j of independent, normal variables having mean 0, but variances $1-(j^2/2\sigma^2 n)$ as $n \to \infty$. This problem is solved in proposition (2.2). In section 4, we prove theorem (1.8). Section 3 gives some technical lemmas, including bounds for the probability of an intersection of events, and approximations for binomial tail probabilities.

In this paper and in Freedman (1977) the condition $E|X_1^3| < \infty$ was assumed. In fact lemmas (4.1) and (4.3) of this paper replace lemma (15) in Freedman (1977), and hence prove theorem (5) of that paper assuming only $E(X_1^2) < \infty$.

The main use of the third-moment assumption in this paper is to justify the bound on the error in the local central limit theorem in equation (3.27). In section 5, we construct an example to show that new limiting behavior can occur when $E|X_1^3| = \infty$. The argument involves Edgeworth-like corrections to sums of random variables without third moments, and may be of independent interest.

Concerning condition (1.5), if $k \gg \sqrt{n} \log n$ but $k = 0[\sqrt{n}\ (\log n)^3]$, the form of the norming constants in equation (1.10) changes radically, because the large deviations corrections in the normal approximation to the binomial distribution must be accounted for. This is worked out in detail for the closely related problem of the mode of the histogram, in section 4 of Diaconis and Freedman (1978). If $k = 0\ (\sqrt{n} \log n)$, then Freedman (1977) shows that the maximum deviation does not converge to zero in probability.

2. A Heuristic Argument.

It is convenient to discuss the behavior of $N_j - kp_j$ separately for four zones. The zones are described in terms of constants δ and A, where δ is small but positive, A is large but finite. For definiteness, we will assume $0 < \delta < 1/10$. The zones are:

(2.1a) zone I : $|j-n\mu| < \delta\sigma\sqrt{n}$

(2.1b) zone II : $\delta\sigma\sqrt{n} \leq |j-n\mu| \leq A\sigma\sqrt{n}$

(2.1c) zone III: $A\sigma\sqrt{n} < |j-n\mu| \leq \sigma n^{3/4}$

(2.1d) zone IV : $\sigma n^{3/4} < |j-n\mu|$.

As will be shown below, zones II, III, and IV make no contribution to the maximum. In zone I, the $N_j - kp_j$ are approximately independent and normal, with mean 0 and

$$\mathrm{Var}(N_j - kp_j) = kp_j(1-p_j) \doteq kp_j \doteq \frac{k}{\sigma\sqrt{2\pi n}}\left[1 - \frac{(j-n\mu)^2}{2\sigma^2 n}\right] .$$

Thus, $L_{nk} - n\mu$ and $\sqrt{\sigma\sqrt{2\pi n}/k}\, M_{nk}$ should be distributed like the L_n and M_n of the following proposition.

(2.2) Proposition. Let U_j be independent normal random variables, with mean 0 and $\mathrm{Var}(U_j) = 1 - j^2/2\sigma^2 n$, for $|j| < \delta\sigma\sqrt{n}$. Let $M_n = \max_j U_j$, and let L_n be the index at which the maximum is achieved. Then, as n tends to ∞, the probability that

$$L_n < y\sigma\sqrt{2n/\log n} \quad \text{and} \quad M_n < \sqrt{\log n - 2 \log\log n + x}$$

converges to

$$\Phi(y)\exp\left[-\sigma\sqrt{2}\; e^{-\frac{1}{2}x}\right] .$$

As in (1.9), Φ is the standard normal distribution function.

Proof. For $-\infty \le a \le b \le \infty$, let I_{ab} be the set of j with $|j| < \delta\sigma\sqrt{n}$ and

$$a\sigma\sqrt{2n/\log n} \le j < b\sigma\sqrt{2n/\log n} .$$

Thus, I_{ab} is always finite. Let M_{ab} be the max of U_j for $j \in I_{ab}$. Clearly, if $a < b < c < d$, then M_{ab} and M_{cd} are independent. Abbreviate

$$(2.3) \qquad w_n(x) = \sqrt{\log n - 2 \log\log n + x} \quad \text{and} \quad \gamma = \sigma\sqrt{2} .$$

We will show that

$$(2.4) \qquad P\{M_{ab} < w_n(x)\} \to \exp\{-\gamma[\Phi(b)-\Phi(a)]e^{-\frac{1}{2}x}\} \quad \text{for } -\infty < a < b < \infty .$$

$$(2.5) \qquad \limsup_{n\to\infty} P\{M_{-\infty a} < w_n(x)\} \le \exp\{-\gamma\Phi(a)e^{-\frac{1}{2}x}\} \quad \text{for } -\infty < a < 0 .$$

$$(2.6) \qquad \limsup_{n\to\infty} P\{M_{b\infty} < w_n(x)\} \le \exp\{-\gamma[1-\Phi(b)]e^{-\frac{1}{2}x}\} \quad \text{for } 0 < b < \infty .$$

Granting (2.4-6), the proposition can be derived by an elementary argument as follows. First, (2.4) holds even for infinite a or b,

in view of (2.5-6). Now let Y_1 and Y_2 be independent double-exponential random variables:

$$P\{Y_i < x\} = \exp[-\gamma_i e^{-\frac{1}{2}x}]$$

where

$$\gamma_1 = \gamma\Phi(y) \quad \text{and} \quad \gamma_2 = \gamma[1-\Phi(y)]$$

so

$$\gamma_1 + \gamma_2 = \gamma \ .$$

The function $w_n(\cdot)$ defined by (2.3) is continuous and strictly increasing; so is its inverse, w_n^{-1}. Equation (2.4) implies that the joint distribution of $w_n^{-1}(M_{-\infty y})$ and $w_n^{-1}(M_{y\infty})$ converges in distribution to the joint distribution of Y_1 and Y_2. Now $P(Y_1 = x) = P(Y_1 = Y_2) = 0$. So

$$(2.7) \qquad P\{L_n < y\sigma\sqrt{2n/\log n} \quad \text{and} \quad M_n < w_n(x)\}$$

$$= P\{M_{-\infty y} < w_n(x) \quad \text{and} \quad M_{-\infty y} > M_{y\infty}\}$$

$$= P\{w_n^{-1}(M_{-\infty y}) < x \quad \text{and} \quad w_n^{-1}(M_{-\infty y}) > w_n^{-1}(M_{y\infty})\}$$

$$\to P\{Y_1 < x \quad \text{and} \quad Y_1 > Y_2\}$$

$$= \int_{-\infty}^{x} \tfrac{1}{2}\gamma_1 e^{-\frac{1}{2}z} \exp[-\gamma_1 e^{-\frac{1}{2}z}] \exp[-\gamma_2 e^{-\frac{1}{2}z}]dz$$

$$= \Phi(y) \int_{-\infty}^{x} \tfrac{1}{2}\gamma e^{-\frac{1}{2}z} \exp[-\gamma e^{-\frac{1}{2}z}]dz \ .$$

The last displayed expression equals

$$\Phi(y) \exp[-\gamma e^{-\frac{1}{2}x}] \ .$$

Thus, (2.4-6) imply (2.2).

We turn now to the proof of (2.4). We will write v_j for $\mathrm{Var}(U_j) = 1 - j^2/2\sigma^2 n$. Fix an x, and abbreviate $w = w_n(x)$. In essence, the proof of (2.4) is the following computation:

$$\log P\{M_{ab} < w\} = \sum_{j \in I_{ab}} \log P\{U_j < w\}$$

$$\doteq \sum_{j \in I_{ab}} P\{U_j > w\}$$

$$\doteq \sum_{j \in I_{ab}} \sqrt{v_j/2\pi} \cdot \frac{1}{w} \cdot \exp\{-\tfrac{1}{2} w^2/v_j\}, \tag{2.8}$$

by (3.15). The symbol $\doteq$ means "approximately equal" and is only used in heuristic argument. For our purposes

$$\sqrt{v_j}/w \doteq 1/\sqrt{\log n}\ . \tag{2.9}$$

In the exponent, $1/v_j \doteq 1 + j^2/2\sigma^2 n$. So

$$\tfrac{1}{2} w^2/v_j \doteq \tfrac{1}{2} \log n - \log\log n + \tfrac{1}{2} x + \tfrac{1}{2} j^2/(2\sigma^2 n/\log n)\ . \tag{2.10}$$

Thus, $\log P\{M_{ab} < w\}$ is approximately $\sqrt{2}\,\sigma\, e^{-\frac{1}{2}x}$ times

$$\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\sigma^2 n/\log n}} \sum_{j \in I_{ab}} \exp\{-\tfrac{1}{2} j^2/(2\sigma^2 n/\log n)\}\ . \tag{2.11}$$

This last is a Riemann sum for

$$\frac{1}{\sqrt{2\pi}} \int_a^b \exp(-\tfrac{1}{2} u^2)\, du = \Phi(b) - \Phi(a)\ .$$

We now indicate the details required to justify (2.8). All "O" and "o" error terms are uniform over $j \in I_{ab}$. Note that a and b are finite, so $j^2/n = O(1/\log n)$. In (2.9), then, the ratio of the two sides converges to one as $n \to \infty$, uniformly over $j \in I_{ab}$. In (2.10),

$$\frac{1}{2} w^2 \frac{1}{v_j} - \frac{1}{2} w^2 \left(1 + \frac{j^2}{2\sigma^2 n}\right)$$

$$= \frac{1}{2} w^2 \left(\frac{1}{v_j} - 1 - \frac{j^2}{2\sigma^2 n}\right)$$

$$= O(\log n) \cdot o(j^2/b)$$

$$= o(1) .$$

So the difference between the two sides of (2.10) is

$$\frac{1}{2} w^2 \frac{j^2}{2\sigma^2 n} - \frac{1}{2} \log n \cdot \frac{j^2}{2\sigma^2 n} + o(1)$$

$$= \frac{1}{2} (w^2 - \log n) \frac{j^2}{2\sigma^2 n} + o(1)$$

$$= O(\log \log n) \cdot O(1/\log n) + o(1)$$

$$= o(1) .$$

Thus, the ratio of $P(U_j > w)$ to $\sqrt{2}\,\sigma\, e^{-\frac{1}{2}x}$ times the j-th term in (2.11) goes to one as n goes to unfinity, uniformly over $j \in I_{ab}$. As

a result,

$$P(U_j > w) = o(1) \tag{2.12}$$

and

$$\sum_{j \in I_{ab}} P(U_j > w) \to \sqrt{2}\,\sigma\, e^{-\frac{1}{2}x} [\Phi(b) - \Phi(a)] \,. \tag{2.13}$$

Now (2.12-13) imply that

$$\sum_{j \in I_{ab}} \log P(U_j \le w) = \sum_{j \in I_{ab}} [P(U_j > w) + o(P(U_j > w))]$$

also converges to the right side of (2.13).

This completes the proof of (2.4). The proof of (2.5) is very similar. In place of (2.9), we use the estimate

$$\sqrt{v_j}/w \le 1/w \,.$$

To modify (2.10), we use the fact that

$$1/v_j \ge 1 + j^2/2\sigma^2 n$$

so

$$\tfrac{1}{2} w^2/v_j \ge \tfrac{1}{2} \log n - \log\log n + \tfrac{1}{2} x + \tfrac{1}{2} w^2\, j^2/2\sigma^2 n \,;$$

and $\sum_{j \in I_{-\infty a}} P\{U_j > w\}$ is bounded above, to within $o(1)$, by

$$\sqrt{2}\,\sigma\, e^{-\frac{1}{2}x} \quad \text{times} \quad w^{-2}\log n \quad \text{times}$$

$$\frac{1}{\sqrt{2\pi}} \quad \frac{w}{\sqrt{2\sigma^2 n}} \sum_{j \in I_{-\infty\, a}} \exp\{-\tfrac{1}{2} w^2 j^2 / 2\sigma^2 n\} \quad .$$

Now $w^{-2}\log n \to 1$, and the last display is at most

$$\Phi(a) + o(1)$$

because the normal density is monotone increasing on $(-\infty, 0)$. The $o(1)$ corresponds to the term in j at the edge of $I_{-\infty\, a}$. This completes our discussion of (2.5), and (2.6) is symmetric. □

3. Probability Approximations.

The random variables N_j introduced in section 1 have a joint multinomial distribution. We will approximate this distribution by getting upper and lower bounds in terms of independent binomial random variables. The binomial variables are then approximated by an appropriate normal distribution. This section contains the basic bounds and approximations. The main results are (3.2), (3.16), and (3.17).

The first lemma is an upper bound for multinomial probabilities due to Mallows (1968).

(3.1) Lemma. Let $M_1, M_2, \ldots, M_J$ have a joint multinomial distribution. Then, for any real numbers $a_1, a_2, \ldots, a_J$,

$$P(M_j \le a_j \text{ for } 1 \le j \le J) \le \prod_{j=1}^{J} P(M_j \le a_j) .$$

The second lemma will be used to get a lower bound for multinomial probabilities.

(3.2) Proposition. Let $(\Omega, \mathfrak{F}, P)$ be a probability triple, and let $\mathfrak{F}_0 \subset \mathfrak{F}_1 \subset \cdots$ be sub-σ-fields of $\mathfrak{F}$. Let A_j and G_j be events in $\mathfrak{F}_j$. Let λ_j, λ_j^*, and ϵ_j be nonnegative numbers. Suppose $P(G_j) \ge 1-\epsilon_j$ and

$$\lambda_j \le P(A_j | \mathfrak{F}_{j-1}) \le \lambda_j^* \text{ on } G_{j-1} .$$

Then $P(\cap_{j=1}^{J} A_j)$ is bounded between

$$\prod_{j=1}^{J} \lambda_j - \sum_{j=0}^{J-1} \epsilon_j \quad \text{and} \quad \prod_{j=1}^{J} \lambda_j^* + \sum_{j=0}^{J-1} \epsilon_j .$$

Proof. Induction on J. □

Let N be binomial with parameters k and p, where $0 < p < 1$. The next set of results give approximations for the distribution of N which are uniform as k and p vary over a wide range. Readers may be surprised to find us proving versions of the central limit theorem for coin-tossing, at this late date. However, we need bounds which are uniform as p gets small; to stay in the realm of the central limit theorem, we will require kp to be large. Our results involve the function

$$g(x) = (1+x)\log(1+x) - x \quad \text{for} \quad -1 < x < \infty . \tag{3.3}$$

The function g is strictly convex, strictly decreasing on $(-1,0)$, strictly increasing on $(0,\infty)$. It satisfies

$$g(x) = \frac{1}{2}x^2 + 0(x^3) \quad \text{as} \quad x \to 0 . \tag{3.4}$$

More precisely,

$$\frac{1}{2}x^2/1+x < g(x) < \frac{1}{2}x^2 \quad \text{for} \quad x > 0 , \tag{3.5}$$

with the inequalities reversed for $x < 0$. The function $g(x)$ is closely

connected to a function which arises in the theory of large deviations theory. See, for example, pp. 100-101 of Kolchin, Sevastyanov, and Chistyakov (1978).

We begin with a special case of theorem (4b) in Freedman (1973), restated here for ease of reference.

(3.6) Lemma. Let $u > 0$. Then

$$P\{N > kp(1+u)\} \leq \exp\{-g(u)kp\} .$$

Combining (3.5) and (3.6) gives a version of Bernstein's inequality:

$$(3.7) \qquad P\{N > kp+m\} \leq \exp\{-\tfrac{1}{2}m^2/(kp+m)\} \quad \text{for} \quad m > 0 .$$

Parenthetically, theorem (4a) in Freedman (1973) implies

$$(3.8) \qquad P\{N < kp-m\} \leq \exp\{-\tfrac{1}{2}m^2/kp\} \quad \text{for} \quad m > 0 .$$

The next result is a variation on lemma (7) of Freedman (1977), and is proved the same way. (The condition $b > 1/\delta$ in that lemma is superfluous.)

(3.9) Lemma. Fix $\epsilon > 0$. There is a positive δ such that the ratio $P(N = a+b)/P(N = a)$ is bounded between $(1 \pm \epsilon)$ times

$$(\frac{a}{a+b})^{\frac{1}{2}} \exp[-g(\tfrac{b}{a})a][\frac{(k-a)p}{a(1-p)}]^b ,$$

for all integers a, b and k satisfying

$$a > 1/\delta, \quad a+b > 1/\delta, \quad k > 1/\delta, \quad a+b < \delta k, \quad |b| < \delta\sqrt{k-a} \ .$$

The next result gives a bound for binomial probabilities $P(N=\nu)$ in terms of the function g. Informally: for small values of p, large values of kp, and ν's not too far from kp, $P(N=\nu)$ is asymptotic to

$$\frac{1}{\sqrt{2\pi kp}} \exp\{-g[(\nu-kp)/kp]kp\} \ .$$

This is a combination local central limit – large deviation result, holding uniformly in small p.

(3.10) Lemma. Fix $\epsilon > 0$. There is a positive δ that $P(N=\nu)$ is bounded between

$$(1 \pm \epsilon) \frac{1}{\sqrt{2\pi kp}} \exp\{-g[(\nu-kp)/kp]kp\}$$

for all k, p and ν satisfying

$$p < \delta, \quad kp > 1/\delta, \quad |\nu-kp| < \delta kp, \quad |\nu-kp| < \delta\sqrt{k} \ .$$

Proof. First, suppose ν is the integer part of kp, so $\nu = kp-\theta$ with $0 \le \theta < 1$. Abbreviate $\bar{p} = 1-p$, so $k-\nu = k\bar{p}+\theta$. By Stirling's formula, $P(N=\nu)$ is asymptotic to

$$\frac{1}{\sqrt{2\pi k}} \sqrt{\frac{k}{\nu} \frac{k}{k-\nu}}$$

times

$$(1 - \frac{\theta}{kp})^{-kp+\theta} \quad (1 + \frac{\theta}{k\bar{p}})^{-k\bar{p}-\theta}$$

The first expression is asymptotic to $1/\sqrt{2\pi kp}$. The second is asymptotic to

$$(1 - \frac{\theta}{kp})^{-kp} \quad (1 + \frac{\theta}{k\bar{p}})^{-k\bar{p}} \quad .$$

By taking logs and expanding, the last is seen to be nearly 1. Thus,

$$P(N=[kp]) \approx 1/\sqrt{2\pi kp} \quad . \tag{3.11}$$

More explicitly, the ratio of the two sides of (3.11) converges to 1 as $\delta \to 0$, uniformly in k and p satisfying our conditions, namely, $p < \delta$ and $kp > 1/\delta$.

We now apply (3.9), with $a = [kp]$ and $b = \nu - a$ to see

$$P(N=\nu)/P(N=[kp]) \approx \exp[-g(b/a)a] \quad . \tag{3.12}$$

First, as easily verified, the conditions of (3.9) hold. Second,

$$(\frac{a}{a+b})^{\frac{1}{2}} \approx 1 \quad ; \tag{3.13}$$

indeed $a = kp + 0(1)$ and $b = \nu - kp + 0(1)$ is small relative to kp, by assumption. (The "0(1)" terms are bounded as $\delta \to 0$, uniformly in k, p and ν satisfying our conditions.) Likewise,

(3.14) $$\left[\frac{(k-a)p}{a(1-p)}\right]^b \approx 1 ;$$

indeed, this expression is

$$(1 + \frac{\theta}{k\bar{p}})^b / (1 - \frac{\theta}{kp})^b ,$$

and $b = \nu - a = \nu - kp + O(1)$ is small relative to kp and to $k\bar{p}$. This completes the proof of (3.12).

Recall that $a = [kp]$ and $b = \nu - a$. The final step in proving (3.10) is to show that

$$|g(\frac{\nu-kp}{kp})kp - g(\frac{b}{a})a|$$

is uniformly small. This tedious piece of calculus is omitted. □

We now recall a bound on the tail probabilities of a normal random variable:

(3.15) Suppose U is a normal random variable with mean 0 and variance 1. Let $w > 0$. Then $P\{U > w\}$ is bounded between

$$\frac{1}{\sqrt{2\pi}} \frac{1}{w + \frac{1}{w}} \exp(-\tfrac{1}{2} w^2) \text{ and } \frac{1}{\sqrt{2\pi}} \frac{1}{w} \exp(-\tfrac{1}{2} w^2) .$$

The following result is a generalization, as one sees by taking $f(x) = \frac{1}{2} x^2$.

(3.16) Proposition. (a) Let f be convex, and suppose $f(\infty) = \infty$. Then

$$\int_x^\infty \exp[-f(u)]\, du \le \frac{1}{f'(x)} \exp[-f(x)] .$$

(b) Suppose $1/f'$ is convex. Then

$$\int_x^\infty \exp[-f(u)]\, du \ge \frac{1}{f'(x) + \frac{f''(x)}{f'(x)}} \exp[-f(x)] .$$

Proof. Claim (a). Let $F = \exp(-f)$. Then $F(\infty) = 0$ and $F' = -f'F$. Now f' is monotone so

$$f'(x) \int_x^\infty F(u)\, du \le \int_x^\infty -F'(u)\, du = F(x) .$$

Claim (b). Let $\tilde{F} = (1/f')\exp(-f)$. Then $\tilde{F}(\infty) = 0$ and $\tilde{F}' = -(1 + \frac{f''}{f'^2})\exp(-f)$. So

$$[1 + \frac{f''(x)}{f'(x)^2}] \int_x^\infty F(u)\, du \ge \int_x^\infty -\tilde{F}'(u)\, du = \tilde{F}(x) . \square$$

Note. If f is convex and f' is concave, then $1/f'$ is convex. This is the case for our function g. The concept of complete monotonicity is relevant here: see section XIII.4 of Feller (1971).

The next lemma extends corollary (8) of Freedman (1977). It provides a bound for binomial tail probabilities, in terms of g, analogous to (3.15).

(3.17) Proposition. Fix $\epsilon > 0$. There is a $\delta > 0$ such that $P(N > kp+m)$ is bounded between

$$(1 \pm \epsilon) \frac{1}{\sqrt{2\pi}} \frac{\sqrt{kp}}{m} \exp[-g(m/kp)kp]$$

for all k, p and m satisfying

$$p < \delta, \quad kp > 1/\delta, \quad \sqrt{kp}/\delta < m < \delta kp, \quad m < \delta\sqrt{k} .$$

Proof. We break the tail probability up into two parts:

$$P\{N > kp+m\} = P_1 + P_2 ,$$

where

$$P_1 = P\{kp+m < N \leq kp+2m\} \text{ and } P_2 = P\{N > kp+2m\} .$$

Later, we will show that P_2 is negligible, but that

$$P_1 \approx \frac{1}{\sqrt{2\pi}} \frac{\sqrt{kp}}{m} \exp[-g(m/kp)kp] . \tag{3.18}$$

First, some preliminaries. To simplify the notation, let

$$G(u) = \exp[-g(u/kp)kp] .$$

Clearly,

$$\frac{1}{\sqrt{2\pi kp}} G(m) \text{ is small relative to } \frac{1}{\sqrt{2\pi}} \frac{\sqrt{kp}}{m} G(m) , \tag{3.19}$$

because m is small relative to kp, by assumption. Likewise, in view of (3.5)

(3.20) $G(2m)$ is small relative to $\frac{1}{\sqrt{2\pi}} \frac{\sqrt{kp}}{m} G(m)$,

because m is small relative to kp, but large relative to $\sqrt{kp}$.

Next, we claim

(3.21) $\sum_{i=m+1}^{\infty} G(i)$ is bounded above by $(1+\epsilon_1) \frac{kp}{m} G(m)$.

Indeed, G is monotone decreasing, so

$$\sum_{i=m+1}^{\infty} G(i) < \int_m^{\infty} G(u)\, du < \frac{1}{g'(m/kp)} G(m)$$

by (3.16). But

$$g'(\frac{m}{kp}) = \log(1 + \frac{m}{kp}) \approx \frac{m}{kp}$$

because m is small relative to kp, by assumption. This proves (3.21).

We are now ready to prove part of (3.18), by estimating P_1 from above. Let ν range from $[kp]+m+1$ to $[kp]+2m$, so $P_1 = \sum_\nu P(N=\nu)$. As (3.10) implies, P_1 is bounded above by

(3.22) $$(1+\epsilon_2) \frac{1}{\sqrt{2\pi kp}} \sum_\nu G(\nu - kp).$$

Because G is decreasing, the expression (3.22) is bounded above by

$$\frac{1}{\sqrt{2\pi kp}} \sum_{i=m}^{2m-1} G(i) .$$

In view of (3.19) and (3.21),

(3.23) P_1 is bounded above by

$$(1+\epsilon_3) \frac{1}{\sqrt{2\pi}} \frac{\sqrt{kp}}{m} \exp\{-g(m/kp)kp\} .$$

We are now ready to prove the rest of (3.18), by estimating P_2 from below. Indeed, as (3.10) implies P_2 is bounded below by $(1-\epsilon_4)$ times

$$(3.24) \qquad \frac{1}{\sqrt{2\pi kp}} \sum_{\nu} G(\nu-kp) \geq \frac{1}{\sqrt{2\pi kp}} \sum_{i=m+1}^{2m} G(i) .$$

The right side of (3.24) is $T_1-T_2-T_3$, where:

$$T_1 = \frac{1}{\sqrt{2\pi kp}} \sum_{i=m}^{\infty} G(i)$$

$$\geq \frac{1}{\sqrt{2\pi kp}} \int_m^{\infty} G(u)\, du$$

$$\approx \frac{1}{\sqrt{2\pi}} \frac{\sqrt{kp}}{m} G(m) \quad \text{by} \quad (3.16) ,$$

$$T_2 = \frac{1}{\sqrt{2\pi kp}} G(m) = o\left[\frac{\sqrt{kp}}{m} G(m)\right] \quad \text{by} \quad (3.19)$$

$$T_3 = \frac{1}{\sqrt{2\pi kp}} \sum_{i=2m+1}^{\infty} G(i)$$

$$\leq (1+\epsilon_1) \frac{1}{\sqrt{2\pi}} \frac{\sqrt{kp}}{2m} G(2m) \quad \text{by (3.21)}$$

$$= O\left[\frac{\sqrt{kp}}{m} G(m)\right] \quad \text{by (3.20)},$$

on recollecting that $\sqrt{kp}/m$ is small by assumption. This completes the proof of (3.18).

Our last job is to show that P_2 is negligible by comparison with P_1. But

$$P_2 = P(N > kp+2m) \leq G(2m)$$

by (3.6); now (3.18) and (3.20) can be used to complete the argument. □

(3.25) Remark. Under the conditions of (3.17), $P(N < kp-m)$ can be bounded by the same expression. The argument is almost identical, because (3.10) is symmetric around kp. The only difference comes at the end:

$$P_2 = P(N < kp-2m) < \exp(-2m^2/kp) \text{ by (3.8)}$$

$$< G(2m) \text{ by (3.5)}.$$

Many of our arguments use the local central limit theorem. We state two versions for ease of reference. For the first version, assuming only a finite second moment, see page 517 of Feller (1971). For the second version, assuming a finite third moment, see page 197 of Petrov (1975).

(3.26) Suppose (1.1), and $E(X_1^2) < \infty$, and (1.6). Then uniformly in j,

$$P(S_n=j) = \frac{1}{\sigma\sqrt{2\pi n}} \exp\{-\tfrac{1}{2}[(j-n\mu)/\sigma\sqrt{n}\,]^2\} + o(1/\sqrt{n})$$

(3.27) Suppose (1.1), (1.2), and (1.6). Then uniformly in j

$$P(S_n=j) \sim \frac{1}{\sigma\sqrt{2\pi n}} \exp\{-\tfrac{1}{2}[(j-n\mu)/\sigma\sqrt{n}]^2\} + O(1/n)$$

4. Proof of theorem (1.8).

The first step in the proof is to show that, with probability approaching one, the maximum does not occur in zones II, III, or IV as defined in (2.1). This part of the argument does not require third moments nor the full force of (1.5). Thus, for lemmas (4.1), (4.3), and (4.5), we assume (1.1)-(1.6) with (1.2) replaced by $E(X_1^2) < \infty$ and (1.5) replaced by $n \to \infty$ and $k/\sqrt{n} \log n \to \infty$. We begin with zone IV.

(4.1) Lemma. Let y be any positive number. Let M_4 be the maximum of $N_j - kp_j$ over zone IV, namely the set of j's satisfying $|j - n\mu| > \sigma n^{3/4}$. Let θ_{nk} be the probability that $M_4 > y(k \log n)^{1/2}/n^{1/4}$. Then $\theta_{nk} \to 0$ as $n \to \infty$.

Proof. To begin with, θ_{nk} is bounded by the sum over j in zone IV of

$$P\{N_j - kp_j > y(k \log n)^{1/2}/n^{1/4}\} .$$

By Chebychev's inequality, the last displayed probability is at most

$$y^{-2} n^{1/2} p_j / \log n .$$

Summing over j in zone IV shows

$$\theta_{nk} \leq (y^{-2} n^{1/2}/\log n) \cdot P(|S_n - n\mu| > \sigma n^{3/4}) .$$

Using Chebyshev's inequality again,

$$P(|S_n - n\mu| > \sigma n^{3/4}) \leq n^{-1/2} .$$

So

$$\theta_{nk} \leq y^{-2}/\log n . \quad \square$$

We turn now to zone III. It will be convenient to abbreviate

$$m = \sqrt{\frac{k}{\sigma\sqrt{2\pi n}} \log n} . \tag{4.2}$$

This m appears in (1.7) and is the leading term in (1.10); it will have some connection with the variable m in section 3.

(4.3) Lemma. Let $0 < A < \infty$. Let M_A be the maximum of $N_j - kp_j$ over j in zone III, namely the set S_A of j with $A\sigma n^{1/2} < |j - n\mu| \leq \sigma n^{3/4}$. Fix any ϵ with $0 < \epsilon < 1$. Then there is an $A = A(\epsilon)$ so large that

$$P\{M_A > (1-\epsilon)m\} \to 0$$

as n and k tend to infinity, satisfying the growth condition $k/\sqrt{n} \log n \to \infty$.

Proof. Let θ satisfy $0 < \theta < \frac{1}{4}(1-\epsilon)^2$. Recall that $p_j = P(S_n = j)$. In view of the local central limit theorem (3.26), there is an $A > 0$ and a finite n_1 such that $p_j < \theta/\sigma\sqrt{2\pi n}$ for all j in S_A provided $n > n_1$. From inequality (3.7),

$$P\{N_j > kp_j + (1-\epsilon)m\} \leq \exp[-\tfrac{1}{2}(1-\epsilon)^2 \frac{m^2}{kp_j + (1-\epsilon)m}] . \tag{4.4}$$

We must now bound $m^2/[kp_j+(1-\epsilon)m]$ from below. The choice of A forces $kp_j \le k\theta/\sigma\sqrt{2\pi n}$. The growth condition on n and k forces $(1-\epsilon)m \le k\theta/\sigma\sqrt{2\pi n}$ eventually. Then,

$$\frac{m^2}{kp_j+(1-\epsilon)m} \ge \frac{m^2}{2k\theta/\sigma\sqrt{2\pi n}} = \frac{\log n}{2\theta},$$

and the right side of (4.4) is smaller than 1/n. Finally, $P\{M_A > (1-\epsilon)m\}$ is bounded above by the sum over $j\epsilon S_A$ of

$$P\{N_j > kp_j+(1-\epsilon)m\}.$$

This sum comprises at most $2\sigma n^{3/4}$ terms, each bounded by 1/n. This proves (4.3). □

We turn now to zone II.

(4.5) Lemma. Fix $A > \delta > 0$. Let M_2 be the maximum of $N_j - kp_j$ for j in zone II, namely the set S_2 of j with $\delta\sigma\sqrt{n} \le |j-n\mu| \le A\sigma\sqrt{n}$. Let m be defined by (4.2). Then, for $\epsilon > 0$ sufficiently small,

$$P\{M_2 > (1-\epsilon)m\} \to 0$$

as n and k tend to infinity satisfying the growth condition $k/\sqrt{n}\log n \to \infty$.

Proof. In view of the local central limit theorem (3.26), there is a $\delta_1 > 0$ and a finite n_1 such that for $n > n_1$,

(4.6) $p_j < (1-\delta_1)/\sigma\sqrt{2\pi n}$ for all j satisfying $|j-n\mu| > \delta\sigma\sqrt{n}$.

Choose ϵ positive with $(1-\epsilon)^2/(1-\delta_1) > 1$. Inequalities (3.7) and (4.6) imply, after some algebra:

$$(4.7) \qquad P\{N_j > kp_j + (1-\epsilon)m\} < \exp[-\tfrac{1}{2}\beta \frac{\log n}{1+\theta_n}]$$

where $\beta = (1-\epsilon)^2/(1-\delta_1) > 1$ by our choice of ϵ, and

$$\theta_n = \frac{1-\epsilon}{1-\delta_1} \sigma^{\frac{1}{2}}(2\pi)^{\frac{1}{4}} n^{\frac{1}{4}}(\log n)^{\frac{1}{2}}/k^{\frac{1}{2}} \to 0$$

by the growth condition on n and k.

Choose β' with $1 < \beta' < \beta$. For large n, the right side of (4.7) is at most $n^{-\beta'/2}$. Sum (4.7) over j in S_2, comprising at most $2A\sigma\sqrt{n}$ terms, to see that

$$P\{M_2 > (1-\epsilon)m\} < 2A\sigma\sqrt{n}\ n^{-\beta'/2} \to 0\ .\ \square$$

We now proceed to zone I. Fix δ with $0 < \delta < 1/10$. For $-\infty \le a \le b \le \infty$, let I_{ab} be the set of j with $|j-n\mu| < \delta\sigma\sqrt{n}$ and

$$(4.8) \qquad a\sigma\sqrt{2n/\log n} \le j-n\mu < b\sigma\sqrt{2n/\log n}\ .$$

These intervals are all finite and nonrandom; $I_{-\infty\,\infty}$ corresponds to all of zone I. Let

$$W_{ab} \text{ be the max of } N_j - kp_j \text{ for } j \text{ in } I_{ab}.$$

To state the main result concerning zone I, recall $z_n(x)$ from (1.10). As in (1.9), let Φ be the standard normal distribution function. We will now need finite third moments and the growth condition (1.5).

(4.9) Proposition. Assume (1.1)-(1.6). Let $-\infty \le a \le b \le c \le d \le \infty$. Then

$$P\{W_{ab} \le z_n(x) \text{ and } W_{cd} \le z_n(y)\}$$

converges to $\exp\{-\sigma\sqrt{2}\ Q\}$, where

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

The proof of (4.9) is a bit complicated. Here are some preliminary estimates.

(4.10) Lemma. The bounds given below apply uniformly to $j \in I_{-\infty\,\infty}$, and the "o" errors are uniform in $j \in I_{-\infty\,\infty}$ as $n \to \infty$.

(a) $(j-n\mu)^2/\sigma^2 n < \delta^2 < 1/100$.

(b) p_j is bounded between

$$\frac{1}{\sigma\sqrt{2\pi n}}\left[1 - \frac{1}{2}\left(1 \pm a_\delta\right)\frac{(j-n\mu)^2}{\sigma^2 n}\right]\left[1 + o\left(\frac{1}{\log n}\right)\right]$$

where $a_\delta = \frac{1}{2}\delta^2$.

(c) $1/\left[1 - \frac{1}{2}\left(1 \pm a_\delta\right)\frac{(j-n\mu)^2}{\sigma^2 n}\right]$ is bounded between

$$1 + \frac{1}{2}\left(1 \pm b_\delta\right)\frac{(j-n\mu)^2}{\sigma^2 n}$$

where b_δ is a function of δ only and $b_\delta = O(\delta^2)$ as $\delta \to 0$.

Note. The notation in (b), although not standard, is convenient for our purposes and will be used throughout. To spell (b) out, there is a sequence $\epsilon_n > 0$ with $\epsilon_n \log n \to 0$, such that for all n, and all j in $I_{-\infty\,\infty}$, the probability p_j is bounded above by

$$\frac{1}{\sigma\sqrt{2\pi n}}\left[1 - \frac{1}{2}\left(1 - a_\delta\right)\frac{(j-n\mu)^2}{\sigma^2 n}\right]\left(1+\epsilon_n\right)$$

and below by

$$\frac{1}{\sigma\sqrt{2\pi n}}\left[1 - \frac{1}{2}(1+a_\delta)\,\frac{(j-n\mu)^2}{\sigma^2 n}\right](1-\epsilon_n)\,.$$

Proof. Part (a) is trivial, and then (b) follows from the local central limit theorem (3.26), using the estimate

$$1-x < e^{-x} < 1-x+\frac{1}{2}x^2\,.$$

Here, $x = (j-n\mu)^2/2\sigma^2 n < \frac{1}{2}\delta^2$, so

$$x-\frac{1}{2}x^2 = (1-\tfrac{1}{2}x)x > (1-\tfrac{1}{4}\delta^2)\,x\,.$$

Part (c) follows from the identity

$$\frac{1}{1-x} = 1 + x + \frac{x^2}{1-x}\,. \quad \square$$

The probability in (4.9) will now be estimated from above and below. An upper bound derives from Mallows' inequality given in (3.1):

(4.11) The probability that $W_{ab} \le z_n(x)$ and $W_{cd} \le z_n(y)$ is at most

$$\Pi_{j\in I_{ab}}\lambda_j(x)\cdot\Pi_{j\in I_{cd}}\lambda_j(y)$$

where

$$\lambda_j(x) = P[N_j \le kp_j + z_n(x)]\,.$$

The rest of the upper-bounding is very similar to the lower-bounding, so details are omitted.

The lower bound will be derived from lemma (3.2). In that lemma take

$$A_j = \{N_j \le kp_j + z_n(x)\} \quad \text{for} \quad j \in I_{ab}$$

(4.12)

$$= \{N_j \le kp_j + z_n(y)\} \quad \text{for} \quad j \in I_{cd} .$$

To define G_j, let

(4.13a) K_j be the set of i with $n\mu - \delta\sigma\sqrt{n} < i \le j$,

(4.13b) $$g_j = \sum_{i \in K_j} p_i$$

(4.13c) $$G_j = \{\sum_{i \in K_j} N_i > kg_j - M\} ,$$

where

(4.13d) $$M = (k \log n)^{\frac{1}{2}} .$$

Since $I_{-\infty\,\infty}$ includes at most $2\delta\sigma n^{\frac{1}{2}}$ indices, (4.10b) shows

(4.14) $g_j < 0.88$ for all $j \in I_{-\infty\,\infty}$, for all large n.

Lemma (3.2) also involves σ-fields $\mathfrak{V}_j$, defined as follows:

(4.15) $\mathfrak{F}_j$ is the σ-field spanned by the N_i as i ranges over K_j.

To apply lemma (3.2), we must estimate $P(G_j)$.

(4.16) Lemma. Define G_j by (4.13). Then the sum of $1-P(G_j)$ over all $j \in I_{-\infty\,\infty}$ tends to 0 as n and k tend to infinity satisfying (1.5).

Proof. Clearly, $\sum_{i \in K_j} N_i$ is binomial: the number of trials is k, and the success probability is g_j. From (3.8),

$$1-P(G_j) \leq \exp[-\tfrac{1}{2}\, M^2/kg_j] \ .$$

Replace g_j by the upper bound (4.14): eventually,

$$1-P(G_j) \leq \exp(-\tfrac{1}{2}\, M^2/k\delta) = n^{-1/2\delta} < n^{-5}$$

for all $j \in I_{-\infty\,\infty}$. But $I_{-\infty\,\infty}$ only comprises $0(n^{\frac{1}{2}})$ terms. □

We are now ready to establish the basic lower bound.

(4.17) Lemma. Define g_j and M by (4.13), and $z_n(x)$ by (1.10). Let N'_j be binomial, with success probability $p'_j = p_j/(1-g_{j-1})$ and the number of trials k'_j, the integer part of $k(1-g_{j-1})+M$. Let

$$\lambda_j(x) = P\{N'_j \leq kp_j + z_n(x)\} \ .$$

Then $P\{W_{ab} \leq z_n(x)$ and $W_{cd} \leq z_n(y)\}$ is bounded below by

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

Proof. Given $\mathfrak{F}_{j-1}$, as defined in (4.15), the conditional distribution of N_j is still binomial; the conditional success probability is p_j', as defined above; the number of trials T_j is an $\mathfrak{F}_{j-1}$-measurable random variable:

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

On G_{j-1}, however, $T_j \le k_j'$. So

$$P\{N_j \le kp_j + z_n(x) | \mathfrak{F}_{j-1}\} \ge \lambda_j(x) \quad \text{on} \quad G_{j-1} :$$

heuristically, the more you toss the coin, the more heads you get.

Lemma (3.2) completes the proof. □

We must now estimate λ_j. Here are some preliminaries.

(4.18) Lemma. Define λ_j as in (4.17). So,

$$\lambda_j(x) = P\{N_j' \le k_j'p_j' + h\}$$

where

$$h = z_n(x) + kp_j - k_j'p_j' .$$

The bounds and error terms given below are uniform over $j\in I_{-\infty\infty}$; the "O" and "o" are as k and n approach infinity, satisfying condition

(1.5) that $k/\sqrt{n}\,(\log n)^3 \to \infty$.

(a) $k'_j p'_j$ is bounded between

$$\frac{k}{\sigma\sqrt{2\pi n}}\left[1 - \frac{1}{2}\,(1 \pm a_\delta)\,\frac{(j-n\mu)^2}{\sigma^2 n}\right]\left[1 + o\left(\frac{1}{\log n}\right)\right]$$

with a_δ of (4.10b).

(b) $h = z_n(x)[1 + O(n^{-1/4})]$

(c) $h^3 = o(k'_j p'_j)^2$

(d) $h^2/k'_j p'_j$ is bounded between

$$\log n - 2 \log\log n + x + \frac{1}{2}\,(1 \pm 2b_\delta)\,\frac{(j-n\mu)^2}{\sigma^2 n/\log n} + o(1)$$

with the b_δ of (4.10c).

(e) Recall the definition of $g(x)$ at (3.3).

Then $g(h/k'_j p'_j)k'_j p'_j$ is bounded between

$$\frac{1}{2}\log n - \log\log n + \frac{1}{2}\,x + \frac{1}{4}\,(1 \pm 2b_\delta)\,\frac{(j-n\mu)^2}{\sigma^2 n/\log n} + o(1) \ .$$

Proof. Part (a). From the definitions of p'_j and k'_j given in (4.17),

$$k'_j p'_j = kp_j\left[1 + \frac{M-\theta}{k(1-g_{j-1})}\right] , \tag{4.19}$$

where $0 \le \theta \le 1$. By (4.13d) and condition (1.5), $(M-\theta)/k = o(1/\log n)$; and $g_j < \delta$ eventually, by (4.14). Then (4.10b) completes the argument for part (a).

Part (b). Recall h as defined in the statement of the lemma. Continuing from (4.19),

$$\frac{h}{z_n(x)} - 1 = -\frac{(M-\theta)}{z_n(x)} \frac{p_j}{1-g_{j-1}} = O(n^{-1/4})$$

as in (a), using the definition (1.10) of $z_n(x)$.

Part (c). This is immediate from (a), (b), and condition (1.5). This is the first time that the full force of (1.5) has been used.

Part (d). Using (a), (b), and (4.10c), $h^2/k'_j p'_j$ is bounded between

$$(\log n - 2 \log\log n + x)[1 + \frac{1}{2}(1 \pm b_\delta) \frac{(j-n\mu)^2}{\sigma^2 n}][1 + o(\frac{1}{\log n})] .$$

Of course, $(2 \log\log n - x)/\log n < b_\delta$ eventually, proving (d).

Part (e). This follows from (c) and (d), by using inequality (3.4). □

The proof of proposition (4.9). We pick up the argument from (4.17). The problem is to put a lower bound on $\Pi_{j \in I_{ab}} \lambda_j(x)$, the other factor being similar. Defining h as in (4.18),

$$1-\lambda_j(x) = P\{N'_j > kp_j + z_n(x)\}$$

$$= P\{N'_j > k'_j p'_j + h\} .$$

Fix an $\epsilon > 0$. Then proposition (3.17) implies that eventually, for all $j \in I_{-\infty,\infty}$, $1 - \lambda_j(x)$ is bounded between

(4.20) $$(1\pm\epsilon)\ (k'_j p'_j/2\pi)^{1/2}\ \frac{1}{h}\ \exp[-g(h/k'_j p'_j)k'_j p'_j]\ .$$

"Eventually" means for all k and n, with n and $k/\sqrt{n}\ \log n$ sufficiently large.

We pause to verify the conditions of the lemma being appealed to: $p'_j = p_j/(1-g_{j-1})$ is small by (4.10b) and (4.14); $k'_j p'_j$ is large by (4.18d); $h^2/k'_j p'_j$ is large by (4.18d); $h/k'_j p'_j$ is small by (4.18c); and $h/k^{1/2}$ is small by (4.18b). Thus, the bound (4.20) is established.

The exponent in (4.20) was estimated in (4.18e); the factor $k'_j p'_j$ in (4.18a); and the factor h in (4.18b). We conclude that eventually, for all j in $I_{-\infty\infty}$, $1-\lambda_j(x)$ is bounded between

$$(1\pm\epsilon)(1\pm c_\delta)\sigma\sqrt{2}\ e^{-x/2}$$

times

$$\frac{1}{\sqrt{2\pi}}\ \frac{1}{\sqrt{2\sigma^2 n/\log n}}\ \exp\{-\frac{1}{4}\ (1\pm 2b_\delta)\ \frac{(j-n\mu)^2}{\sigma^2 n/\log n}\}$$

where c_δ is a function of δ only and $c_\delta = O(\delta^2)$ as $\delta \to 0$.

Arguing as in proposition (2.2), we get a lower bound for the probability in (4.9), of the form $(1-3\epsilon)(1-d_\delta)\exp(-Q_\delta)$, where

$$Q_\delta = \sigma\sqrt{2}\ e^{-x/2}\ [\Phi(f_\delta b) - \Phi(f_\delta a)]$$

$$+ \sigma\sqrt{2}\ e^{-y/2}\ [\Phi(f_\delta d) - \Phi(f_\delta c)]$$

with $d_\delta = O(\delta^2)$ and $f_\delta = 1+O(\delta^2)$. It is at this point that the uniformity in j was critical.

An upper bound of similar form can be obtained by essentially the same argument, starting from (4.11) instead of (4.17). Then the proof can be completed as in (2.2). □

Our last main job in proving (1.8) is showing that $\max(N_j - kp_j)$ is assumed at a unique index j, with probability approaching one as n and $k/\sqrt{n}\ (\log n)^3$ tend to ∞. The following heuristic discussion may make the argument easier to follow. Inspection of (1.8) suggests that the maximum is likely to occur only for j's within $O(n/\log n)^{1/2}$ of $n\mu$. Call these the critical j's.

(4.21) There are $O(n/\log n)^{1/2}$ critical j's.

Recall m from (4.2). The range of likely values for the maximum is from

$$m(1 - \frac{2\ \text{loglog}\ n}{\log n} - \frac{b}{\log n})^{1/2} \doteq m(1 - \frac{2\ \text{loglog}\ n}{2\ \log n} - \frac{b}{2\ \log n})$$

to

$$m(1 - \frac{2\ \text{loglog}\ n}{\log n} + \frac{b}{\log n})^{1/2} \doteq m(1 - \frac{2\ \text{loglog}\ n}{2\ \log n} + \frac{b}{2\ \log n}) .$$

Call these the critical values. In what follows we write $a_n \sim b_n$ if $\underline{\lim}\ a_n/b_n > 0$ and $\overline{\lim}\ a_n/b_n < \infty$.

(4.22) There are the order of $m/\log n \sim k^{\frac{1}{2}}\ n^{-\frac{1}{4}}\ (\log n)^{-\frac{1}{2}}$ critical values for the maximum.

Each of these critical values for the maximum corresponds to some value i for an N_j of around

$$kp_j + m(1 - \frac{2 \text{ loglog } n}{\log n})^{1/2}$$

$$\doteq kp_j + (kp_j)^{1/2} (\log n - 2 \text{ loglog } n)^{1/2} .$$

Now $N_j - kp_j$ is essentially normal with mean 0 and variance $kp_j \sim k/\sqrt{n}$, so

$$P(N_j = i) \doteq \frac{1}{\sqrt{2\pi kp_j}} \exp[-\frac{1}{2} \frac{(i-kp_j)^2}{kp_j}]$$

$$\sim k^{-1/2} n^{1/4} \exp[-\frac{1}{2} \log n + \text{loglog } n]$$

$$= k^{-1/2} n^{-1/4} \log n .$$

Furthermore, the N_j are nearly independent. The chance that the maximum occurs at two different indices is bounded by the sum, over the critical j, j', i and i' satisfying

(4.23) $$i - kp_j = i' - kp_{j'},$$

of

$$P(N_j = i \text{ and } N_{j'} = i') \sim k^{-1} n^{-\frac{1}{2}} (\log n)^2 .$$

The number of critical i's was estimated in (4.22): to each, there corresponds at most one i' by (4.23), "at most" because i' has to

be an integer. The number of pairs j, j' is $O(n/\log n)$ by (4.21). So the chance that the maximum occurs at two different indices is of order

$$n(\log n)^{-1} \cdot k^{1/2} n^{-1/4} (\log n)^{-1/2} \cdot k^{-1} n^{-1/2} (\log n)^2$$

$$= k^{-1/2} n^{1/4} (\log n)^{1/2} .$$

This last quantity tends to zero because of the growth condition (1.5).

Returning to rigorous argument, recall $z_n(x)$ defined in (1.10) and I_{ab} defined by (4.8). The main estimate is the following.

(4.24) Lemma. Assume (1.1)-(1.6). Fix positive finite numbers a and b. Then, uniformly over pairs of indices $j \neq j'$ in I_{-aa}, and pairs of integers i, i' satisfying

$$z_n(-b) \leq i - kp_j = i' - kp_{j'} \leq z_n(b) \tag{4.25}$$

we have

$$P(N_j = i) = O(k^{-1/2} n^{-1/4} \log n) \tag{4.26}$$

and

$$P(N_{j'} = i' \mid N_j = i) = O(k^{-1/2} n^{-1/4} \log n) . \tag{4.27}$$

Proof. The first assertion (4.26) follows from lemma (3.10), the requisite estimates for p_j being given by (4.10). To make this

valid, the conditions of (3.10) must be verified. Then, one argues from (4.25) and (1.5) that $(i-kp_j)^3 = o(kp_j)^2$, so

$$g[(i-kp_j)/kp_j]\, kp_j = \frac{1}{2}\,(i-kp_j)^2/kp_j + o(1)\ .$$

But, using (4.22) again,

$$(i-kp_j)^2 = \frac{k}{\sigma\sqrt{2\pi n}}\,[\log n - 2\ \mathrm{loglog}\ n + O(1)]\ .$$

And by (4.10),

$$kp_j = \frac{k}{\sigma\sqrt{2\pi n}}\,[1 + O(\frac{1}{\log n})]\ .$$

So

$$g[(i-kp_j)/kp_j]\, kp_j = \frac{1}{2}\log n - \mathrm{loglog}\, n + O(1)\ .$$

We omit the other details in the proof of (4.26).

For the second assertion (4.27), 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 lemma (3.10). All "O" and "o" estimates are uniform over $j \neq j'$ in I_{-aa} and i,i' satisfying (4.25), as n and $k/\sqrt{n}\,(\log n)^3$ tend to infinity. We will show that

(4.28) $$\hat{\hat{kp}} = kp_{j'}[1+o(1/\log n)]$$

(4.29) $$(i'-\hat{\hat{kp}})^2/\hat{\hat{kp}} = \log n - 2 \log\log n + O(1) .$$

Assume these bounds for the moment. Another application of lemma (3.10) shows that $P(N_{j'}=i' \mid N_j=i)$ is of order

$$(\hat{\hat{kp}})^{-1/2} \exp[-\tfrac{1}{2} (i'-\hat{\hat{kp}})^2/\hat{\hat{kp}}] .$$

Now use (4.28) and (4.29) as before to complete the proof.

We now prove (4.28) and (4.29). Let

$$\theta = \hat{\hat{kp}} - kp_{j'} = (kp_j - i)p_{j'}/(1-p_j) .$$

Abbreviate $h = i'-kp_{j'} = i-kp_j$. Then (b) of Lemma (4.10) implies $\theta/h = -p_{j'}/(1-p_j) = O(1/\sqrt{n})$. Condition (4.25) makes h of order $k^{1/2}n^{-1/2}(\log n)^{1/2}$, and then $\theta = O(k^{1/2}n^{-3/4}(\log n)^{1/2})$. Finally $kp_{j'}$ is of order $k//\overline{n}$, again by (b) of lemma (4.10). To summarize:

(4.30) $$\theta = O[k^{1/2}n^{-3/4}(\log n)^{1/2}],\ h \approx k^{1/2}n^{-1/4}(\log n)^{1/2},\ kp_{j'} \approx kn^{-1/2} .$$

Now we can prove (4.28):

$$\begin{aligned}\hat{\hat{kp}} &= kp_{j'} + \theta \\ &= kp_{j'} + O[k^{1/2}n^{-3/4}(\log n)^{1/2}] \\ &= kp_{j'}\{1+O[k^{-1/2}n^{-1/4}(\log n)^{1/2})\} \\ &= kp_{j'}[1+o(1/\log n)] .\end{aligned}$$

To prove (4.29), note that $i' - \hat{k}\hat{p} = h - \theta$. Now

$$(i' - \hat{k}\hat{p})^2 / \hat{k}\hat{p} = (h^2 - 2\theta h + \theta^2)/\hat{k}\hat{p} = h^2/\hat{k}\hat{p} + o(1) .$$

From (4.25)

$$h^2 = \frac{k}{\sigma\sqrt{2\pi n}} [\log n - 2 \log\log n + O(1)] .$$

In view of (4.10),

$$\hat{k}\hat{p} = \frac{k}{\sigma\sqrt{2\pi n}} [1 + O(1/\log n)] .$$

So

$$h^2/\hat{k}\hat{p} = \log n - 2 \log\log n + O(1) . \quad \square$$

We can now prove the uniqueness assertion in (1.8).

(4.31) Proposition. Assume (1.1-1-6). Then $\max_j (N_j - kp_j)$ is assumed at a unique index j, with probability approaching one.

Proof. We first establish:

(4.32) Given $\epsilon > 0$, there are large, finite numbers a and b, such that eventually, except for probability ϵ, $\max_j (N_j - kp_j)$ is assumed only for indices in I_{-aa}, and is between $z_n(-b)$ and $z_n(b)$.

"Eventually" means for sufficiently large values of n and $k/n^{1/2}(\log n)^3$. Indeed, for any b, except for probability less than $\epsilon/5$, the maximum over zone IV is eventually smaller than $z_n(-b)$ by (4.1). Likewise for zone III by (4.3), and zone II by (4.5). As a matter of notation, zone I is $I_{-\infty\,\infty}$ and can be dealt with by (4.9). There are a and b so large that eventually, except for probability $\epsilon/5$,

$$z_n(-b) < \max\{N_j - kp_j \text{ for } j \text{ in } I_{-\infty\alpha}\} < z_n(b) .$$

Finally, choose a so large that eventually, except for probability $\epsilon/5$, $N_j - kp_j < z_n(-b)$ for all $j \in I_{-\infty\infty} \setminus I_{-aa}$. This completes the proof of (4.32).

The next two estimates are easily checked.

(4.33) The number of pairs (j,j') with $j' \neq j$ in I_{-aa} is $O(n/\log n)$.

(4.34) The number of integers i with $z_n(-b) \leq i - kp_j < z_n(b)$ is at most

$$z_n(b) - z_n(-b) = O[k^{1/2}n^{-1/4}(\log n)^{-1/2}] .$$

Now, lemma (4.24) and relations (4.33-34) enable us to estimate the chance of finding two distinct indices $j' \neq j$ in I_{-aa} with $N_j - kp_j = N_{j'} - kp_{j'}$ in the critical range from $z_n(-b)$ to $z_n(b)$. Take the number of pairs $j \neq j'$ in I_{-aa}, and multiply by the number of integers i with $i - kp_j$ in the critical range. Each i is a possible value for N_j, and associated with it is at most one possible value i' for $N_{j'}$ satisfying $i - kp_j = i' - kp_{j'}$. Then, we multiply by $P(N_j = i) \cdot P(N_{j'} = i' \mid N_j = i)$. The result is $O[k^{-1/2}n^{-1/4}(\log n)^{1/2}] \to 0$. □

The proof of theorem (1.8) can now be given.

Proof of theorem (1.8). As proposition (4.31) implies, M_{nk} is assumed at a unique location L_{nk}, with probability approaching one. Lemmas (4.1), (4.3), and (4.5) imply that with probability approaching one, L_{nk} is in zone I. Now, proposition (4.9) implies (1.8) by the same argument which showed that equation (2.4) implies proposition (2.2). □

5. Moment Assumptions.

We have been assuming (1.3) that $E|X_1^3| < \infty$. An argument sketched later in this section will prove theorem (1.8) under the weaker hypothesis

$$(5.1) \qquad E\{|X_1|^2[\log(1+|X_1|)]^{2+\delta}\} < \infty \text{ for some } \delta > 0 .$$

However, if only

$$(5.2) \qquad E\{|X_1|^2[\log(1+|X_1|)]^{\delta}\} < \infty \text{ for some } \delta \text{ with } 0 < \delta < 1 ,$$

the conclusions of the main theorem (1.8) become false. This is demonstrated by theorem (5.4) below, which gives the correct asymptotic formulae for a particular random variable satisfying (5.2) but not (5.1). The present discussion leaves unresolved the case $E\{|X_1|^2[\log(1+|X_1|)]^{1+\delta}\} < \infty$ for some δ with $0 \leq \delta \leq 1$.

Theorem (5.4) uses an Edgeworth-like correction in the local central limit theorem for a random variable without third moments, given in (5.23). The techniques used are similar to those in Pitman (1968). Some related work can be found in Cramer (1963), or Takeuchi and Akahira (1976a,b).

We now define a class of probability distributions on the integers. Let $a > 1$ be given. Let $q = q_a$ be the symmetric probability on the integers with

$$(5.3) \qquad \begin{cases} q(0) = q(-1) = q(1) = 0 \\ q(j) = b/[j^3(\log|j|)^a] \text{ for } j = \pm 2, \pm 3, \ldots , \\ b \text{ is chosen so that } \sum_j q(j) = 1 . \end{cases}$$

As for theorem (1.8), let σ^2 be the variance of q. Let $X_1, X_2, \ldots$ be independent with common distribution (5.3), and write $S_n = X_1 + \cdots + X_n$. Take k independent copies of S_n. Let N_j be the number of these sums which are equal to j. Let $M_{nk} = \max_j [N_j - kP(S_n = j)]$.

(5.4) Theorem. Let $X_1, X_2, \ldots$ be independent, having common distribution (5.3), with $1.5 < a \leq 2$. Let $c = b2^{a-1}/(a-1)$. Define

$$(5.5) \qquad z_n(x,a) = \sqrt{\frac{k}{\sigma\sqrt{2\pi n}}}\ \sqrt{\log n + c(\log n)^{2-a} - 2 \log\log n + x}\ .$$

Then, with probability approaching one, M_{nk} is taken on at a unique location L_{nk}, and the probability that

$$L_{nk} < y\sigma\sqrt{2n/\log n} \quad \text{and} \quad M_{nk} < z_n(x,a)$$

converges to

$$\Phi(y)\ \exp[-\sigma\sqrt{2}\ e^{-x/2}]$$

as n and k tend to infinity satisfying the growth condition (1.5).

Note that q in (5.3) has a second moment, but just barely. In particular

$$\sum j^2[\log(1+|j|)]^{\delta} q(j)$$

is finite provided $\delta < a-1$, but infinite for $\delta \geq a-1$. Note too that $z_n(x,a)$ defined in (5.5) is different from $z_n(x)$ of (1.10). When $a = 2$, the difference is not so dramatic--a constant under the square root sign. But when $a < 2$, there is an extra term going to infinity, namely $c(\log n)^{2-a}$.

To prove (5.4), sharp estimates of $P(S_n=j)$ are needed. This is done by approximating the characteristic function of X_1. We begin with some preliminary estimates. The first lemma is implicit in Pitman (1968).

(5.6) Lemma. Suppose the real numbers u_n are nonnegative, nonincreasing and (weakly) convex:

$$u_n \geq 0, \qquad u_n \geq u_{n+1}, \qquad u_n - u_{n+1} \geq u_{n+1} - u_{n+2}.$$

Suppose too that $u_n \to 0$ as $n \to \infty$. Let S be the (conditionally convergent) sum $u_0 - u_1 + u_2 - u_3 + \cdots$. Then

$$\frac{1}{2} u_0 \leq S \leq u_0 - \frac{1}{2} u_1 .$$

Proof. Clearly,

$$\begin{aligned} S &= (u_0 - u_1) + (u_2 - u_3) + \cdots \\ &\geq (u_1 - u_2) + (u_3 - u_4) + \cdots \quad \text{by convexity} \\ &= u_0 - S . \end{aligned}$$

This proves the first inequality. For the second,

$$S = u_0 - (u_1 - u_2 + u_3 - \cdots) \leq u_0 - \frac{1}{2} u_1 ,$$

because $u_1 - u_2 + u_3 - \cdots \geq \frac{1}{2} u_1$ by the first inequality applied to the sequence $u_1, u_2, \ldots$. □

The next lemma is at the heart of the approximations in this section. It is abstracted from Theorem 1 of Pitman (1968). We work with conditionally convergent Riemann integrals. Such integrals are denoted by using an arrow over the integral sign.

(5.7) Lemma. Let $H(x)$ be a convex function on $[0,\infty)$ which decreases to 0. Then, for any $t > 0$,

$$(5.8) \qquad \frac{1}{2} I_0 \le t\int_0^{\to\infty} H(x)\sin(tx)\,dx \le I_0 - \frac{1}{2} I_1$$

where

$$(5.9) \quad I_0 = \int_0^{\pi} H(x/t)\sin x\,dx \text{ and } I_1 = \int_0^{\pi} H[(x+\pi)/t]\sin x\,dx .$$

Proof. By changing variables, it is enough to do the case $t = 1$. Abbreviate $h_j(x) = H(x+j\pi)$. Split at multiples of π to see that the integral to be estimated equals the conditionally convergent sum

$$\sum_{j=0}^{\infty} \int_{j\pi}^{(j+1)\pi} H(x)\sin x\,dx = \sum_{j=0}^{\infty} \int_0^{\pi} (-1)^j h_j(x)\sin x\,dx .$$

Group the terms as $(u_0-u_1) + (u_2-u_3) + \cdots$ to get

$$\sum_{k=0}^{\infty} \int_0^{\pi} [h_{2k}(x)-h_{2k+1}(x)]\sin x\,dx .$$

Because H is monotone decreasing, $h_{2k} \ge h_{2k+1}$: so the last sum is absolutely convergent, and we can take $\sum$ inside the integral by Fubini's theorem. As a result,

$$\int_0^{\to\infty} H(x)\sin x\,dx = \int_0^{\pi} \sum_{k=0}^{\infty} [h_{2k}(x)-h_{2k+1}(x)]\sin x\,dx .$$

Now H is convex, so for each x the sequence $h_0(x), h_1(x), \cdots$ is

convex. By (5.6),

$$\frac{1}{2} h_{(0)}(x) \leq \sum_{k=0}^{\infty} (h_{(k)}(x) - h_{(k+1)}(x)) \leq h_{(0)}(x) - \frac{1}{2} h_{(1)}(x) .$$

This completes the proof. □

We will now apply this result to H's of order $1/(\log x)^{\alpha}$, at infinity, and evaluate the corresponding integrals I_0 and I_1 in (5.7). A calculus estimate will be needed.

(5.11) Lemma. Let α be a positive number and x a real number. Then,

(a) $1/(1+x)^{\alpha} \geq 1 - \alpha x$ for $x \geq -1$.

(b) $1/(1+x)^{\alpha} \leq 1 - \alpha x + 2^{\alpha+1} \alpha(\alpha+1)x^2$ for $x \geq -\frac{1}{2}$.

(c) $1/(1-x)^{\alpha} \leq 1 + kx$ for $0 \leq x \leq \frac{1}{2}$, with k finite but depending on α.

Proof. Expand $1/(1+x)^{\alpha}$ in a Taylor series with remainder. □

(5.12) Lemma. Let α be a positive number. Then, as $t \to 0^+$,

$$\int_{\sqrt{t}}^{\pi} (\log \frac{x}{t})^{-\alpha} \sin x \, dx$$

and

$$\int_{\sqrt{t}}^{\pi} (\log \frac{x+t}{t})^{-\alpha} \sin x \, dx$$

are both equal to

$$\Gamma(1-\alpha)(\log \frac{1}{t})^{-\alpha} + O(\log \frac{1}{t})^{-\alpha-1} .$$

Proof. We will only do the first integral, the second being similar. Write the first integrand as

$$(\log \tfrac{1}{t})^{-\alpha} \quad (1 + \frac{\log x}{\log \frac{1}{t}})^{-\alpha} \quad \sin x \ .$$

Now $x \geq \sqrt{t}$, so $(\log x)/(\log \frac{1}{t}) \geq -\frac{1}{2}$, and (5.11a-b) can be used to estimate the middle factor in the display. The integrals which result are evaluated as follows:

$$\int_{\sqrt{t}}^{\pi} \sin x \ dx = 2 + O(t)$$

$$\int_0^{\pi} \sin x \ |\log x|^j \ dx < \infty \quad \text{for} \quad j = 1 \quad \text{or} \quad 2. \quad \square$$

The following calculus facts are needed to estimate the tails of the distribution (5.3).

(5.13) Lemma. Let $\alpha > 0$ and $\beta > 1$.

(a) Define

$$f(\alpha,\beta,u) = \int_u^{\infty} x^{-\beta}(\log x)^{-\alpha} \ dx \ .$$

Then $f(\alpha,\beta,u)$ is bounded below by

$$\frac{1}{\beta-1} u^{-\beta+1}(\log u)^{-\alpha} - \frac{\alpha}{(\beta-1)^2} u^{-\beta+1}(\log u)^{-\alpha-1} \ .$$

Likewise, $f(\alpha,\beta,u)$ is bounded above by

$$\frac{1}{\beta-1} u^{-\beta+1}(\log u)^{-\alpha} \ .$$

(b) $\sum_{j=n}^{\infty} j^{-\beta} \log j)^{-\alpha} = \frac{1}{\beta-1} n^{-\beta+1}(\log n)^{-\alpha} + O[n^{-\beta+1}(\log n)^{-\alpha-1}]$.

(c) Let $\alpha > 1$. Then

$$\int_u^{\infty} x^{-1}(\log x)^{-\alpha} dx = \frac{1}{\alpha-1} (\log u)^{-\alpha+1} .$$

Proof. Claim (a). Clearly,

$$f(\alpha,\beta,u) = \frac{1}{\beta-1} u^{-\beta+1} \log u)^{-\alpha} - \frac{\alpha}{\beta-1} f(\alpha+1,\beta,u) .$$

In particular,

$$f(\alpha,\beta,u) \leq \frac{1}{\beta-1} u^{-\beta+1}(\log u)^{-\alpha} . \tag{5.14}$$

Using (5.14) with $\alpha+1$ in place of α,

$$f(\alpha,\beta,u) \geq \frac{1}{\beta-1} u^{-\beta+1}(\log u)^{-\alpha} - \frac{\alpha}{(\beta-1)^2} u^{-\beta+1}(\log u)^{-\alpha-1} .$$

Claim (b). Clearly, $x^{-\beta}(\log x)^{-\alpha}$ is monotone. So, the sum in (b) is bounded between $f(\alpha,\beta,n)$ and $f(\alpha,\beta,n) + n^{-\beta}(\log n)^{-\alpha}$.

Claim (c). This is clear. □

Of course, the estimates in (5.13a-b) can be developed into asymptotic series.

We now begin to investigate the probability density q defined by (5.3). The following repeated integrals of the tail of q will be useful.

(5.15) Lemma. Define $q(j)$ and b by (5.3). For $x > 0$, set

$$F(x) = \sum_{|j| > x} q(j)$$

$$G(x) = \int_x^\infty F(u)\ du$$

$$H(x) = \int_x^\infty G(u)\ du .$$

Then, as $x \to \infty$,

(a) $$F(x) = bx^{-2}(\log x)^{-a} + O[x^{-2}(\log x)^{-a-1}]$$

(b) $$G(x) = bx^{-1}(\log x)^{-a} + O[x^{-1}(\log x)^{-a-1}]$$

(c) $$H(x) = \frac{b}{a-1}(\log x)^{-a+1} + O[(\log x)^{-a}] .$$

Proof. This is immediate from (5.13). □

Note. Claims (b) and (c) follow from (a). In fact, theorem (5.9) and the supporting results (5.17-20-23) hold provided the X's are symmetric, interger-valued, aperiodic and satisfy (5.15a), namely,

$$P(|X| > x) = bx^{-2}(\log x)^{-a} + O[x^{-2}(\log x)^{-a-1}]$$

as $x \to \infty$. It is the precise bound on the remainder, namely $O[x^{-2}(\log x)^{-a-1}]$, which enables us to push the calculations through. In this generality, there is no connection between a and b, or $\sigma^2 = \text{Var } X$; the latter is finite by our condition on $P(|X| > x)$.

Let $\theta(t)$ be the characteristic function of q. Since q is symmetric,

$$\theta(t) = \sum_j \cos(jt)q(j). \tag{5.16}$$

The next lemma gives an approximation for θ. Recall that σ^2 is the variance of q: namely, $\sigma^2 = \sum_j j^2 q(j)$.

(5.17) Lemma. Fix $a > 1$. As t tends to 0,

$$\theta(t) = 1 - \frac{1}{2}\sigma^2 t^2 + r(t)$$

where

$$r(t) = \frac{b}{a-1} t^2 (\log \frac{1}{|t|})^{-a+1} + o[t^2 (\log \frac{1}{|t|})^{-a}]$$

and b is defined in (5.3).

Proof. By symmetry, it is enough to do the case where $t \to 0^+$. Integrating by parts three times gives

$$\theta(t) = 1 - \frac{1}{2}\sigma^2 t^2 + r(t) \tag{5.18}$$

where

$$r(t) = t^3 \int_0^\infty \sin(tx)\, H(x)\, dx\,,$$

and H is defined in (5.15). Clearly, H is convex and $H(0) < \infty$ because $\sigma^2 < \infty$. Now (5.7) shows that

(5.19) $$t^2 \frac{1}{2} I_0 \leq r(t) \leq t^2 (I_0 - \frac{1}{2} I_1)$$

with I_0 and I_1 defined as in (5.9). To approximate I_0, break the range of integration into two zones $[0,\sqrt{t}]$ and $[\sqrt{t},\pi]$. For the first zone,

$$0 \leq \int_0^{\sqrt{t}} H(x/t) \sin x \, dx < H(0) \int_0^{\sqrt{t}} x \, dx = \frac{1}{2} H(0) t .$$

For the second zone, use the approximation (5.15c) for $H(x/t)$, along with (5.12). Combining the two zones we have:

$$I_0 = \frac{2b}{a-1} (\log \frac{1}{t})^{-a+1} + O[(\log \frac{1}{t})^{-a}] .$$

The same estimate for I_1 can be obtained by the same argument. Substituting into (5.19) completes the proof. □

We will use lemma (5.17) in the following form:

(5.20) Lemma. Define q and b by (5.3). Let $\theta(t)$ be the characteristic function of q defined in (5.16). Let σ^2 be the variance of q. Then, as $t \to 0$,

$$\exp(\frac{1}{2} \sigma^2 t^2) \theta(t) = 1 + \gamma(t) + \rho(t)$$

where

$$\gamma(t) = \frac{b}{a-1} t^2 (\log \frac{1}{|t|})^{-a+1}$$

and

$$\rho(t) = O[t^2(\log \tfrac{1}{|t|})^{-a}] .$$

The next two facts are well known, but are recorded here for ease of reference.

(5.21) Lemma. Let z_j and z'_j be complex numbers of absolute value at most A. Then

$$| \Pi_{j=1}^n z_j - \Pi_{j=1}^n z'_j | \leq A^n \sum_{j=1}^n |z_j - z'_j| .$$

(5.22) Lemma. Let n be a positive integer, and z a complex number. Then

$$|(1+z)^n - 1 - nz| \leq \tfrac{1}{2} n^2 |z|^2 e^{n|z|} .$$

We are now ready to give an approximation for the probability density of S_n. We will abbreviate the normal density with mean 0 and variance σ^2 by φ_σ:

$$\varphi_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\tfrac{1}{2} x^2/\sigma^2) .$$

Also, we introduce the Hermite polynomial:

$$h_\sigma(x) = (x/\sigma)^2 - 1 .$$

(5.23) Proposition. Fix $a > 1$. Let $X_1, X_2, \ldots$ be independent with common distribution (5.3). Let $S_n = X_1 + \cdots + X_n$. Then, as $n \to \infty$,

$$\sqrt{n}\, P(S_n = j) = \varphi_\sigma(j/\sqrt{n})[1 - \epsilon_n h_\sigma(j/\sqrt{n})] + r_{jn} ,$$

where

$$\sigma^2 = \text{Var } X_1 ,$$

$$\epsilon_n = \frac{b}{a-1} / (\log \sqrt{n})^{a-1} \quad \text{with} \quad b \quad \text{as in (5.3)} ,$$

and

$$d = \min(a, 2a-2) > a-1$$

$$r_{jn} = O(\log n)^{-d} \quad \text{uniformly in} \quad j .$$

Note. Similar techniques would give more terms in an Edgeworth-like expansion.

Proof. Our argument is adapted from section 16.2 of Feller (1971). By the Fourier inversion formula,

$$\varphi_\sigma(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx) \exp(-\tfrac{1}{2} \sigma^2 t^2) \, dt . \tag{5.24}$$

Differentiate twice with respect to x:

$$h_\sigma(x)\varphi_\sigma(x) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx) \, t^2 \exp(-\tfrac{1}{2} \sigma^2 t^2) \, dt . \tag{5.25}$$

Likewise,

$$P(S_n = j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-itj) \quad \theta(t)^n \, dt .$$

Changing variables,

$$\sqrt{n} \, P(S_n = j) = \frac{1}{2\pi} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp(-itj/\sqrt{n}) \quad \theta(t/\sqrt{n})^n \, dt \tag{5.26}$$

Combining (5.24-25-26), we have

$$r_{jn} = \sqrt{n}\, P(S_n = j) - \varphi_\sigma(j/\sqrt{r}) + \epsilon_n \varphi_\sigma(j/\sqrt{n}) h_\sigma(j//\overline{n}) .$$

$$= J_1 + J_2$$

where

$$J_1 = \frac{1}{2\pi} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp(-itj/\sqrt{n})\, \psi_n(t)\, dt$$

with

(5.27) $$\psi_n(t) = \varphi(t/\sqrt{n})^n - (1+\epsilon_n t^2) \exp(-\tfrac{1}{2}\sigma^2 t^2)$$

and

$$J_2 = -\frac{1}{2\pi} \int_{|t|>\pi\sqrt{n}} \exp(-itj/\sqrt{n})\, (1+\epsilon_n t^2) \exp(-\tfrac{1}{2}\sigma^2 t^2)\, dt .$$

Clearly,

$$|J_2| \le \frac{1}{2\pi} \int_{|t|>\pi\sqrt{n}} (1+\epsilon_n t^2) \exp(-\tfrac{1}{2}\sigma^2 t^2)\, dt ,$$

so J_2 is negligible.

Likewise,

(5.28) $$|J_1| \le \frac{1}{2\pi} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} |\psi_n(t)|\, dt .$$

We split the interval of integration in (5.28) into three sub-intervals:

R_1, the set of t with $|t| < n^{1/4}$

R_2, the set of t with $n^{1/4} < |t| \leq \delta n^{1/2}$

R_3, the set of t with $\delta n^{1/2} < |t| \leq \pi n^{1/2}$.

We will choose $\delta > 0$ later.

Consider R_3. The distribution $q(j)$ of (5.3) is aperiodic, so $\sup_{|t|>\delta} |\theta(t)| = \lambda < 1$. Of course

$$|\psi_n(t)| \leq \theta(t/\sqrt{n})^n + (1+\epsilon_n t^2) \exp(-\tfrac{1}{2}\sigma^2 t^2) .$$

So the contribution from R_3 to the integral in (5.28) is at most

$$\lambda^n \sqrt{n} + \int_{|t|>\delta\sqrt{n}} (1+\epsilon_n t^2) \exp(-\tfrac{1}{2}\sigma^2 t^2)\, dt .$$

This is negligible.

Next consider R_2. Using (5.17), we may choose $\delta > 0$ so small that

$$|\theta(u)| < 1 - \tfrac{1}{4}\sigma^2 u^2 < \exp(-\tfrac{1}{4}\sigma^2 u^2)$$

for $|u| < \delta$. This fixes the δ defining R_2. In particular,

$$|\theta(t/\sqrt{n})|^n < \exp(-\tfrac{1}{4}\sigma^2 t^2)$$

for all t in R_2. As a result, the contribution from R_2 to the integral in (5.28) is negligible.

We now show that the contribution from R_1 to the integral in 5.28) is $O(\log n)^{-a} + O(\log n^{-a})$. Let

$$\beta_n(t) = \exp(\tfrac{1}{2}\sigma^2 t^2/n)\ \theta(t/\sqrt{n}) - 1 .$$

Recall the definition of ψ_n from (5.27). Clearly

$$\psi_n(t) = \{[1+\beta_n(t)]^n - 1 - \epsilon_n t^2\} \exp(-\tfrac{1}{2}\sigma\ t^2) . \tag{5.29}$$

Recall γ and ρ from (5.20). As that result shows,

$$\beta_n(t) = \gamma(t/\sqrt{n}) + \rho(t/\sqrt{n}) . \tag{5.30}$$

Use (5.20) again: there are positive, finite constants K_i such that for all t in R_1

$$|n\rho(t/\sqrt{n})| \le K_1 t^2/(\log n)^a \tag{5.31}$$

and then

$$|n\beta_n(t)| \le K_2 t^2/(\log n)^{a-1} ; \tag{5.32}$$

so

$$|n\beta_n(t)| \le \tfrac{1}{4}\sigma^2 t^2 \quad \text{for all large } n . \tag{5.33}$$

We are estimating

$$J_1^* = \frac{1}{2\pi} \int_{R_1} |\psi_n(t)|dt .$$

Using (5.29), (5.30), and the triangle inequality, we have

$$J_1^* \le J_a + J_b + J_c$$

where

$$J_a = \frac{1}{2\pi} \int_{R_1} |n\rho(t/\sqrt{n})| \exp(-\tfrac{1}{2}\sigma^2 t^2)\, dt$$

$$J_b = \frac{1}{2\pi} \int_{R_1} |[1+\beta_n(t)]^n - [1+n\beta_n(t)]| \exp(-\tfrac{1}{2}\sigma^2 t^2)\, dt$$

$$J_c = \frac{1}{2\pi} \int_{R_1} |n\gamma(t/\sqrt{n}) - \epsilon_n t^2| \exp(-\tfrac{1}{2}\sigma^2 t^2)\, dt\ .$$

We estimate J_a, using (5.31):

$$J_a \le \frac{1}{2\pi} (\log n)^{-a} \int t^2 \exp(-\tfrac{1}{2}\sigma^2 t^2)\, dt$$

$$= O(\log n)^{-a}\ .$$

We estimate J_b, using (5.22):

$$|[1+\beta_n(t)]^n - 1 - n\beta_n(t)| \le \tfrac{1}{2} |n\beta_n(t)|^2 \exp(|n\beta_n(t)|)\ .$$

In view of (5.32-33), this bound is at most

$$\tfrac{1}{2} K_2^2 (\log n)^{-2a} t^4 \exp(\tfrac{1}{4}\sigma^2 t^2)\ .$$

So

$$J_b \le \frac{1}{4\pi} K_2^2 (\log n)^{2-2a} \int t^4 \exp(-\tfrac{1}{4}\sigma^2 t^2)\, dt$$

$$= O(\log n)^{2-2a}\ .$$

Finally, we estimate J_c. This is more delicate, and it is convenient to treat two zones separately. Let

R_1^+ be the set of t with $1 < |t| \le n^{1/4}$

R_1^- be the set of t with $|t| \le 1$.

So $R_1 = R_1^+ \cup R_1^-$. Using the definition of γ in (5.20), and the definition of ε_n in (5.34), the contribution of J_c from R_1^+ is

$$\frac{1}{2\pi} \frac{b}{a-1} \frac{1}{(\log \sqrt{n})^{a-1}} \tag{5.54}$$

times

$$\int_{R_1^+} \left| \left(1 - \frac{\log |t|}{\log \sqrt{n}}\right)^{1-a} - 1 \right| t^2 \exp\left(-\frac{1}{2}\sigma^2 t^2\right) dt . \tag{5.55}$$

The integrand in (5.55) can be estimated by (5.11c), because $|t| < n^{1/4}$ in R_1, so $\log|t|/\log\sqrt{n} < \frac{1}{2}$ in R_1. Hence (5.55) is at most

$$\frac{K}{\log \sqrt{n}} \int_{|t| > 1} (\log|t|)\, t^2 \exp\left(-\frac{1}{2}\sigma^2 t^2\right) dt .$$

This is $O(\log n)^{-1}$. Thus, the contribution to J_c from R_1^+ is at most $O(\log n)^{-a}$. Likewise, the contribution to J_c from R_1^- is the factor (5.54) times

$$\int_{R_1^-} \left| \left[1 + \frac{\log(1/|t|)}{\log \sqrt{n}}\right]^{1-a} - 1 \right| t^2 \exp\left(-\frac{1}{2}\sigma^2 t^2\right) dt .$$

By (5.11b), this is at most

$$\frac{a-1}{\log \sqrt{n}} \int_{-1}^{1} (\log \frac{1}{|t|}) t^{\cdot} \exp(-\frac{1}{\cdot} \sigma^{\cdot 2} t^{2}) \, dt$$

$$+ \cdot^{a} \frac{a(a-1)}{(\log \sqrt{n})^{\cdot}} \int_{-1}^{1} (\log \frac{1}{|t|})^{2} t^{\cdot} \exp(-\frac{1}{\cdot} \sigma^{2} t^{2}) \, dt \, .$$

This too is $\cdot(\log n)^{-1}$. Thus, the contribution to J_c from R_1^- is also $\cdot(\log n)^{-a}$. □

We are now in position to finish the proof of (5.4). The argument is very similar to the pro f of theorem (1.8) and we just indicate the changes.

Proof of (5.4). The bounds on the maximum deviation in zones IV, III, and II, as given in (4.1), (4.3), and (4.5) go through without change. The estimates used in (4.10) are valid as stated, except that (4.10b) becomes:

(5.36) $p_j = P(S_n = j) = \frac{1}{\sigma \sqrt{2\pi n}} \pi_{jn}$, where

$$\pi_{jn} = 1 - [1 + \cdot(\delta^{\cdot})] \, j^{\cdot 2}/(2\sigma^{\cdot} n) + \frac{b}{a-1} (\log \sqrt{n})^{1-a} + \cdot(\log n)^{2-2a})$$

uniformly in j ‹ zone I; the "$\cdot(\delta^{\cdot})$" is uniform in j and n. This follows from (5.[illegible]). Indeed, since j is confined to zone I, $j^{\cdot}/n^{2} = \cdot(\delta^{2})$, and

$$\exp[-j^{\cdot}/(2\sigma^{\cdot} n)] = 1 - [1 + \cdot(\delta^{2})] \, j^{\cdot}/(\cdot\sigma^{2} n) \, .$$

We must now multiply by

$$1 - \frac{b}{a-1} (\log \sqrt{n})^{1-a} \left(\frac{j^{2}}{2\sigma^{\cdot} n} - 1\right) .$$

We get

$$1 - [1 + O(\delta^2)]j^2/2\sigma^2 n + \frac{b}{a-1}(\log\sqrt{n})^{a-1} + T$$

where

$$T = \frac{b}{a-1}(\log\sqrt{n})^{1-a}\ \frac{j^2}{2\sigma^2 n}\,[1 + [1+O(\delta^2)](\frac{j^2}{2\sigma^2 n} - 1)]$$

$$= o(1) \cdot \frac{j^2}{\sigma^2 n}$$

because $a > 1$; and T can be merged into the lead term of π_{jn}. Since $a \leq 2$,

$$d = \min(a, 2a-2) = 2a- \ ,$$

accounting for the error term $O(\log n)^{2-2a}$ in π_{jn}. Instead of (4.1 c), we must estimate

$$1 - \pi_{nj} = 1 + [1 + O(\delta^2)]\, j^2/(2\sigma^2 n)$$

$$- \frac{b}{a-1}(\log\sqrt{n})^{1-a} + O(\log n)^{2-2a} \ ,$$

the higher powers of the lead term merging into the $O(\delta^2)$, because for $j \in$ zone I,

$$(j^2/2\sigma^2 n)^2 \quad (\delta^2) \cdot j^2/\sigma^2 n \ .$$

Lemmas (4.16-17) go through as stated, with $z_n(x)$ replaced by $z_n(x,a)$. In (4.18a), the bound on $k'_j p'_j$ is

$$\frac{k}{\sigma\sqrt{2\pi n}}\{1 - [1+O(\delta^2)]\frac{j^2}{2\sigma^2 n} + \frac{b}{a-1}\frac{1}{(\log\sqrt{n})^{1-a}} + O(\log n)^{2-2a}\}[1+o(\frac{1}{\log n})] .$$

In (4.18b), $z_n(x)$ should be replaced by $z_n(x,a)$. There is no problem with (4.18c). Remarkably enough, the bound on $h^2/k'_j p'_j$ in (4.18d) goes through unchanged, with $\mu = 0$, and some different $b_\delta = O(\delta^2)$. It is here that we use the condition $a > 1.5$, as well as the definition of the constant c in (5.4). In essence, h^2 and $k'_j p'_j$ have the same lead factor $k/\sigma\sqrt{2\pi n}$, which cancels in $h^2/k'_j p'_j$, leaving the product UVW, where

$$U = \log n + c(\log n)^{2-a} - 2\log\log n + x$$

$$V = 1 + [1+O(\delta^2)]\frac{j^2}{2\sigma^2 n} - \frac{b}{a-1}\frac{1}{(\log\sqrt{n})^{a-1}} + O(\log n)^{2-2a}$$

$$W = 1 + o(1/\log n) .$$

The factor W can be ignored. In multiplying out UV, the term

$$-\log n \cdot \frac{b}{a-1}\frac{1}{(\log\sqrt{n})^{a-1}} = -2^{a-1}\frac{b}{a-1}(\log n)^{2-a}$$

cancels the term

$$c(\log n)^{2-a} \cdot 1 ,$$

because c was chosen to be $2^{a-1}\, b/(a-1)$. Furthermore,

$$U \cdot O(\log n)^{2-2a} = o(1) ,$$

because we imposed the condition $a > 1.5$, so $3-2a < 0$.

The remaining arguments used in (1.8) go through with only minor changes. □

Finally, we indicate why theorem (1.8) continues to hold when the third-moment condition (1.2) is replaced by the log-moment condition (5.1). Let

$$F(x) = P(X_1 > x) \quad \text{for} \quad x \geq 0$$

$$= P(X_1 < x) \quad \text{for} \quad x < 0 .$$

By Chebychev's inequality,

$$F(x) = O[x^2(\log|x|)^{-2-\delta}] \quad \text{as} \quad x \to \pm\infty . \tag{5.37}$$

Let

$$G(x) = \int_x^{\infty} F(u)\, du \quad \text{for} \quad x \geq 0$$

$$= \int_{-\infty}^{x} F(u)\, du \quad \text{for} \quad x < 0 .$$

Using (5.13),

(5.38) $$G(x) = O[|x|^{-1}(\log|x|)^{-2-\delta}] \quad \text{as} \quad x \to \pm\infty .$$

Let

$$H(x) = \int_x^\infty G(u)\,du \quad \text{for} \quad x > 0$$

$$= \int_{-\infty}^x G(u)\,du \quad \text{for} \quad x < 0 .$$

Using (5.13) again,

(5.39) $$H(x) = O[(\log|x|)^{-1-\delta}] \quad \text{as} \quad x \to \pm\infty .$$

Now, if $\theta(t)$ is the characteristic function of X_1, and $\mu = E(X_1)$, we can argue as in (5.17) to show that

$$\theta(t) = 1 + i\mu t - \frac{1}{2}E(X_1^2)t^2 + r(t)$$

where as $t \to 0$,

$$r(t) = O[|t|^2(\log\frac{1}{|t|})^{-1-\delta}] .$$

Consequently, by an argument like the one used for (5.23) but without the Hermite polynomial,

(5.40) $$\sqrt{n}\,P(S_n=j) = \varphi_\sigma[j-n\mu] + O(\log n)^{-1-\delta} ,$$

where $\sigma^2 = \text{Var } X_1$. The approximation (5.40) can be used in place of the local central limit theorem (.27) in proving (1.8). An approximation similar to (5.40) based on fractional moments (rather then log moments) is given by Ibragimov (1966).

References

Cramer, H. (1963), On Asymptotic Expansions for Sums of Independent Random Variables with a Limiting Stable Distribution. Sankhya, 25, Series A, 13-24

Diaconis, P. and Freedman, D. (1979), The Distribution of the Mode of an Emperical Histogram. Submitted to Zeitschrift für Wahrscheinlichkeitstheorie.

Feller, W. (1971), An Introduction to Probability Theory and Its Applications, Vol. II, 2nd Ed. Wiley, New York.

Freedman, D. (1973), Another Note on the Borel-Cantelli Lemma and the Strong Law with the Poisson Approximation as a By-Product. Ann. Prob., 1, 910-925.

Freedman, D. (1977), A Central Limit Theorem for Emperical Histograms. Zeitschrift für Wahrscheinlichkeitstheorie, 41, 1-11.

Freedman, D., Pisiani, R., and Purves, R. (1978), Statistics. W. W. Norton, New York.

Ibragimov, I.A. (1966), On the Accuracy of Approximation to the Distribution Functions of Sums of Independent Variables. Th. Prob. and its Applications II, 559-579.

Kolchin, A., Sevastyanov, B., and Chistyakov, C. (1978). Random Allocations, Wiley, New York.

Mallows, C.L. (1968), An Inequality Involving Multinomial Probabilities. Biometrika, 55, 422-424.

Petrov, V.V. (1975), Sums of Independent Random Variables. Springer, New York.

Pitman, E.J.G. (1968), On the Behavior of the Characteristic Function of a Probability Distribution in the Neighborhood of the Origin. Jour. Austral. Math., Soc. 8, 423-443.

Takeuchi, K. and Akahira, M. (1976a), On Gram-Charlier-Edgeworth Type Expansions of the Sums of Random Variables (I). Rep. Univ. Electro-Comm., 27, 95-115.

Takeuchi, K. and Akahira, M. (1976b), On Gram-Charlier-Edgeworth Type Expansions of the Sums of Random Variables (II). Rep. Univ. Electro-Comm., 27, 117-[illegible].

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 288	2. GOVT ACCESSION NO.: AD-A087904
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): THE DIFFERENCE BETWEEN THE EMPIRICAL HISTOGRAM AND THE PROBABILITY HISTOGRAM FOR SUMS	5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Persi Diaconis and David Freedman	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0475; NSF-MCS 77-01665
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Statistics, Stanford University, Stanford, CA 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Statistics and Probability Program Code 436, Arlington, Virginia 22217	12. REPORT DATE: JUL[redacted]1980
	13. NUMBER OF PAGES: 69
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Research partially supported by NSF Grant MCS-77-01665.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Histogram; Multinomial distribution; Edgeworth expansion.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE REVERSE SIDE.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

#288

THE DIFFERENCE BETWEEN THE EMPIRICAL HISTOGRAM AND THE PROBABILITY HISTOGRAM FOR SUMS

Suppose S_n is a sum of n independent and identically distributed random variables with $E|X_1^3| < \infty$. If n is large, S_n is approximately normal. A histogram of k copies of S_n will be close to the normal curve if k is large relative to $\sqrt{n} \log n$. This paper derives the joint distribution of the location and size of the maximum deviation between this histogram and the probability histogram for S_n. When k is large relative to $\sqrt{n}\ (\log n)^3$, the maximum deviation is taken on at a unique location. The location is normally distributed and independent of the size of the maximum deviation, which has a double-exponential distribution. We construct an example, involving Edgeworth-like expansions, to show the behavior changes if $E(X_1^2) < \infty$ but $E|X_1^3| = \infty$.


DATE
FILME
8 - 8