PROBABILITIES OF EXCESSIVE DEVIATIONS OF SIMPLE LINEAR RANK STATISTICS

Robert P. Clickner, et al
Florida State University

Prepared for:
Army Research Office

February 1975

DISTRIBUTED BY:
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

REPRODUCED FROM BEST AVAILABLE COPY
Probabilities of Excessive Deviations of Simple Linear Rank Statistics

by

Robert P. Clickner\textsuperscript{1,2} and Jayaram Sethuraman\textsuperscript{3}
The Florida State University

FSU Technical Report No. M311

February, 1975

Abbreviated Title: Deviations of Linear Rank Statistics

\textsuperscript{1,3} Research supported by the United States Army Research Office, Durham under Grant No. DAHC04-74-G-1000.

\textsuperscript{2} Research supported in part by the National Institutes of Health Grant No. 5-T01-GM00913 to the Florida State University.

The United States Government is authorized to reproduce and distribute reprints for government purposes.

\textsuperscript{1} Now at Department of Statistics, Temple University, Philadelphia Pa., 19122.

\textsuperscript{3} Now visiting professor at the University of Michigan.

Key words and phrases. Excessive deviations, convergence rates, simple linear rank statistics, asymptotic expansions.

AMS 1970 subject classifications 60F10, 62G99.
Probabilities of Excessive Deviations of Simple Linear Rank Statistics

Robert P. Clickner and Jayaram Sethuraman

Department of Statistics
Florida State University
Tallahassee, Florida 32306

U. S. Army Research Office
Box CH, Duke Station
Durham, North Carolina 27706

Approved for public release; distribution unlimited.

Rates of convergence to 0 of probabilities of excessive deviations of simple linear rank statistics are obtained under the null hypothesis. These results fill the gap between the known results on the asymptotic normality and probabilities of large deviations of such statistics. They are also potentially useful in other applications like Bayes risk efficiency.
SUMMARY

Probabilities of Excessive Deviations of
Simple Linear Rank Statistics

Let \( \{S_N\} \) be a sequence of r.v.'s whose asymptotic distribution is \( N(0, \sigma_N^2) \) and let \( \{x_N\} \) be a sequence of constants with \( x_N \to \infty \). A (righthand) excessive deviation is an event of the form \( \{S_N > x_N\} \).

The asymptotic normality of \( S_N \) tells us \( P(S_N > x_N) \to 0 \), \( N \to \infty \), but not the rate of this convergence. These rates are needed for the evaluation of Bahadur \( (x_N = x\sqrt{N}) \) and Bayes Risk \( (x_N = x\sqrt{\log N}) \) efficiencies. When \( S_N \) is a \( k \)-sample linear rank statistic (see Hâjek and Šidák (1967)), and \( x_N = o(N) \), we show

\[
\log P(S_N > x_N) \sim -x_N^2/2, \quad \text{assuming the null hypothesis.}
\]

In the two-sample case, we establish also that

\[
P(S_N > x_N) = o(\exp[-x_N^2/2 + J(x_N, \lambda_N)]),
\]

when \( N\lambda_N (0 < \lambda_N < 1) \) is the size of the first sample and \( J \) is a function whose behavior is analyzed for various \( x_N \) and \( \lambda_N \). For example, if \( x_N = o(N) \), then

\[
J(x_N, \lambda_N) = o(1) \text{ as } N \to \infty.
\]
Simple linear rank statistics arise in a variety of situations, particularly in the problem of testing the equality of two or more distributions by non-parametric tests. While the asymptotic distributions of simple linear rank statistics have been studied extensively (see, e.g., Hájek and Šidák (1967)), investigations concerning their large deviation properties have been made only recently (Stone (1967, 1968, 1969) and Woodworth (1970)). In this paper we study the rates of convergence to zero of null probabilities of excessive deviations of k-sample ($k \geq 2$) simple linear rank statistics. We begin in this section with discussions of the notions of an excessive deviation and of a simple linear rank statistic. The results are in Section 2 and the proofs are in Section 3.

The concept of an excessive deviation of a random variable will be discussed here in a general setting. Let $\{S_N, N \geq 1\}$ be a sequence of random variables with positive finite variances and
let \( \{ \sigma_N, N \geq 1 \} \) be a sequence of positive constants such that

\[
\sup_N \left| P(S_N \leq k\sigma_N) - \Phi(x) \right| \to 0, \quad N \to \infty,
\]

where \( \Phi(\cdot) \) is the standard normal distribution function.

Let \( \{x_N, N \geq 1\} \) be a sequence of nonnegative real numbers.

We call \( P_N(x_N) = P(S_N > x_N \sigma_N) \) the probability of a (right-hand) deviation of size \( x_N \) of \( S_N \). A trivial consequence of (1.1) is

\[
P_N(x_N) \sim \phi(-x_N), \quad N \to \infty,
\]

provided \( x_N = O(1) \). However, when \( x_N \to \infty \), (1.2) is clearly no longer a direct consequence of (1.1), which leads one to ask: given \( \{S_N\} \) and \( \{x_N\} \), where \( x_N \to \infty \) at some specified rate, does (1.2) hold? If not, what exactly is the asymptotic behavior of \( P_N(x_N) \)? To answer these questions, tools more refined than the central limit theorem are needed.

We will now introduce some terminology, due to Rubin and Sethuraman (1965a). If \( x_N = O(1) \), the event \( \{S_N > x_N \sigma_N\} \) is called an ordinary deviation of \( S_N \), while, if \( x_N \to \infty \), it is an excessive deviation of \( S_N \). Two special cases of excessive deviations have
separate names, because of the applications they have found in statistics. The case where \( x_n^2/\log N \rightarrow c^2 \), \( 0 < c < \infty \) is a moderate deviation of \( S_N \), which arises in the study of Bayes risk efficiency (see Rubin and Sethuraman (1965b) and Clickner (1972)). The most extensively studied excessive deviation is the large deviation, in which \( x_n^2/N \rightarrow c^2 \). This attention arises from the fact that probabilities of large deviations must be evaluated in order to compute Bahadur efficiencies (see Bahadur (1960)).

Much of the previous work on excessive deviations has been for sums of independent random variables. When \( S_N \) is the mean of i.i.d. random variables with zero mean, unit variance and finite moment generating function, and \( \sigma_N = N^{-\frac{1}{2}} \), Cramér (1938) has shown

\[
P_N(x_N) \sim \phi(-x_N) \exp \left[ x_N^3 N^{-\frac{1}{2}} \lambda(x_N^{-\frac{1}{2}}) \right],
\]

where \( x_N \rightarrow \infty \), \( x_N^2 = o(N) \), and where \( \lambda(z) \) is a function which admits a convergent power series expansion for small \( |z| \). A corollary of (1.3) is that (1.2) holds if \( x_N^6 = o(N) \).
Feller (1943) generalized Cramér's (1938) result to non-
identically distributed random variables. Rubin and Sethuraman
(1965a) considered only moderate deviations but were able to relax
the requirement of a moment generating function. Other work on the
problem has been done by Chernoff (1952), Linnik (1960, 1961, 1962),
and others.

Only right-hand deviations \( \{ S_N > N \sigma_N \} \) have been discussed
here. This practice will be continued throughout this paper.
All results for right-hand excessive deviations have immediate
extensions to left-hand deviations \( \{ S_N < -N \sigma_N \} \) and two-sided
deviations \( \{ |S_N| > N \sigma_N \} \).

We will now define a simple linear rank statistic. Let
\( X_{11}, \ldots, X_{1n_1}, \ldots, X_{k1}, \ldots, X_{k_n_k} \) be a sequence of independ-
ent and continuous random variables, where \( k \geq 2, n_i \geq 1, i = 1, \ldots, k \) and \( n_1 + \ldots + n_k = N \). Let \( R_{11}, \ldots, R_{k_n_k} \) be the ranks
for the combined sample \( X_{11}, \ldots, X_{k_n_k} \). Consider the problem of
testing the null hypothesis \( H: X_{11}, \ldots, X_{k_n_k} \) are identically
distributed versus the k-sample alternative with density
\[
q(x_{11}, \ldots, x_{k_n_k}) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} f(x_{ij} - \Delta_i),
\]
where \( f \) is a known density and \( \Delta_1, \ldots, \Delta_k \) are known constants.
A locally most powerful test of $H$ versus $q$ is based on the \textit{k-sample} \textit{simple linear rank statistic} (see Hájek and Šidák (1967) p. 69).

\begin{equation}
S_N = \sum_{i=1}^{k} \Delta_i \sum_{j=1}^{n_i} a_N(R_{ij})
\end{equation}

where $a_N(1), \ldots, a_N(N)$ are a sequence of constants, called \textit{scores}.

These simple linear rank statistics include many of the more common and important rank statistics. For example, if $a_N(i) = i/N + 1, i = 1, \ldots, N$, we obtain a $k$-sample extension of the classical Wilcoxon rank-sum statistic. When $a_N(i) = 1$ for $i > (N + 1)/2$ and $-1$ for $i \leq (N + 1)/2$, $S_N$ becomes a median test. With $a_N(i) = \mathbb{E}[-(U_N^{(i)})], i = 1, \ldots, N$, where $U_N^{(1)} < \ldots < U_N^{(N)}$ are the order statistics for a sample of size $N$ from the uniform $(0,1)$ distribution, $S_N$ is the Fisher-Yates-Terry-Hoeffding normal scores statistic.

To the authors' knowledge, all previous work on excessive deviations of linear rank statistics has been for the case of large deviations only and under the null hypothesis $H$ (Stone (1967, 1968, 1969) and Woodworth (1970)). In this paper we will be considering only excessive deviations that are not large (i.e. $x_N \to \infty$ but $x_N^2/N \to 0$) and only the null hypothesis $H$. Thus there is no overlap of the present paper with Stone's or Woodworth's work. In fact, our results fill in a gap between their work and the work of these many authors who have studied
the asymptotic normality of simple linear rank statistics (see Hájek and Šidák (1967) for references.

2. Main Results.

We need some notation. Let $\lambda_{N_i} = n_i/N$, $i=1, \ldots, k$. Let

$$\mu_{\Delta N} = \lambda_{N1} \Delta_1 + \ldots + \lambda_{Nk} \Delta_k$$

and

$$(2.1) \quad \sigma^2_{\Delta N} = \sum_{i=1}^{k} \lambda_{N_i} (\Delta_i - \mu_{\Delta N})^2.$$ 

We shall assume

$$(2.2) \quad 0 < \liminf_{N \to \infty} \lambda_{N_i} \leq \limsup_{N \to \infty} \lambda_{N_i} < 1, \quad i = 1, \ldots, k,$$

and $\sigma^2_{\Delta N} > 0$ for all $N \geq k$. Further, let $\sum_{j=1}^{N} a_N(j) = 0$, and $\sum_{j=1}^{N} a_N^2(j) = N$.

Let $S_N$ be as defined in (1.5). Throughout the rest of this paper we will assume that the null hypothesis $H$ holds. Then $E S_N = 0$ and $\text{var} S_N = N^2 \sigma^2_{\Delta N}/(N-1)$. We define $P_N(X_N)$ to be
so we define $P_N(x_N)$ to be

$$P_N(x_N) = P(S_N > x_N \sigma_{\Delta N} N^{1/2})$$

We can now state our main results.

**Theorem 2.1.** Let $S_N$ be of the form (1.5), assume the null hypothesis $H$ obtains and define $P_N(x_N)$ by (2.3), where $x_N \to \infty$ and $x_N^2/N \to 0$. Assume

$$x_N^2 \max_{1 \leq j \leq N} a_N^2(j) = o(N)$$

and

$$\gamma_N^3 = \sum_{j=1}^{N} |a_N(j)|^3/N = o(N^{3/2}/x_N)$$

Then, as $N \to \infty$,

$$\log P_N(x_N) \sim -x_N^2/2.$$ 

This crude estimate of $P_N(x_N)$ can be improved somewhat in the two-sample case ($k = 2$), provided (2.4) and (2.5) are strengthened a little. When $k = 2$, without loss of generality, we may set $\Delta_1 = 1, \Delta_2 = 0$, which gives (with $n = n_1$),

$$S_N = \sum_{j=1}^{n_1} a_N(R_{1j}).$$

Letting $\lambda_N = \lambda_{N1}$ and $\lambda_N = \lambda_{N2}$, we get $\mu_{\Delta N} = \lambda_N$ and $\sigma_{\Delta N}^2 = \lambda_N \lambda_N$, so

(2.3) becomes
We now state our additional results for the two-sample case.

Observe that (2.9) includes both (2.4) and (2.5).

Theorem 2.2. Let the null hypothesis $H_0$ obtain, let $S_N$ be of the form (2.7), and define $P_N(x_N)$ by (2.8), where $x_N = 0$, $x_N^2/N \to 0$. If

$$ (2.9) \quad \gamma_N^3 = o\left(\frac{N}{x_N^3}\right), $$

then, as $N \to \infty$,

$$ (2.10) \quad P_N(x_N) = o\{\exp[-x_N^2/2 + J(x_N, \lambda_N)]\}, $$

where

$$ (2.11) \quad J(x_N, \lambda_N) = (x_N^2/2)(1 - \lambda_N^2/\lambda_N + \lambda_N q_N) + NI(\lambda_N, p_N), $$

where

$$ (2.12) \quad I(\lambda, p) = \lambda \log(\lambda/p) + \bar{\lambda} \log(\bar{\lambda}/q), $$

where $q = 1-p$, $q_N = 1-p_N$, and $p_N$ is the unique solution of the equation

$$ (2.13) \quad n = \frac{N}{\sum_{j=1}^{N} \{1 + (q_N/p_N)\exp[-x_N^2/2 + J(x_N, \lambda_N)]\}^{-1}. $$

Theorems 2.1 and 2.2 are the best results obtainable with pre-
sently available methods. However, it would be desirable to improve on these results to obtain a better estimate of $P_N(x_N)$, one comparable to Cramér's (1938) result (1.3). To do this in the two-sample case a certain rate of convergence to normality of the normed distribution $\phi_N(\cdot)$ of the mean of certain dependent random variables arising in the proof of Theorem 2.2 is required. This convergence rate has not been established. Let

$$\Delta_N = \sup_x |\phi_N(x) - \phi(x)|,$$

where $\phi_N(\cdot)$ is defined in (3.39). It follows from Hâjek (1964) that $\Delta_N \to 0$ as $N \to \infty$. His result is reproduced here as Lemma 3.7.

The required convergence rate is $\Delta_N x_N = o(1)$ where $x_N \to \infty$ (See (3.46)). If it can be verified that this rate holds then the following estimate is obtainable:

$$(2.15) \quad P_N(x_N) \sim \phi(-x_N) \exp\{J(x_N, \lambda_N)\}, N \to \infty.$$ 

To extend Theorem 2.2 to the case $k > 2$, a multivariate generalization of Hâjek's (1964) asymptotic normality result is required. Unfortunately, it too, is not yet available.

The conditions of Theorems 2.1 and 2.2 are stronger than those required to establish the asymptotic normality of simple linear rank statistics $S_N$ (See, e.g., Hâjek and Šidák (1967) pp. 193-195). This is to be expected since these theorems give
more detailed and precise information about the asymptotic behavior of $S_N$ than does a statement of asymptotic normality. Even so, it is natural to ask whether these stronger conditions are very restrictive, in terms of potential applications. Many simple linear rank statistics $S_N$ have scores of either of the forms

$$(2.16) \quad a_N(i) = E \phi(u_N^{(1)})$$

or

$$(2.17) \quad \hat{a}_N(i) = \phi(i/N+1),$$

where $\phi$ is a non-constant function on $(0,1)$ and, further, $S_N$ is asymptotically normal if $\int_0^1 \phi^2(u) du < \infty$ (See, e.g., Hájek and Šidák (1967) Chapter V). Clearly, if

$$(2.18) \quad \int_0^1 |\phi(u)|^3 du < \infty,$$

then (2.9) holds for the scores $a_N$ and $\hat{a}_N$ and we may apply Theorems 2.1 and 2.2 to $S_N$. Some simple linear rank statistics that have scores of the form (2.16) or (2.17) where $\phi$ satisfies (2.18) are: median, Wilcoxon, Van der Waerden and Fisher-Yates-Terry-Hoeffding, all tests for location shift, and Capon, Klotz, Ansari-Bradley, quartile, and Savage, all tests for a shift in scale. See Clickner (1972) for details.
It is not immediately obvious from Theorem 2.2 as presented, exactly how fast $P_N(x_N)$ tends to zero when $x_N$ grows at a specified rate, say $x_N^2 = \log N$, or perhaps $x_N^6 = N^3$. It is necessary to analyze $J(x_N, \lambda_N)$ for various $x_N$ and $\lambda_N$ to see more clearly the behavior of $P_N(x_N)$. This is done in Corollaries 2.5 and 2.6 following the preliminary Lemmas 2.3 and 2.4.

**Lemmas 2.3.** Let $p_N$ be the solution of equation (2.13). Let (2.9) hold. Then

\[(2.19) \quad p_N = \lambda_N + \sum_{k=1}^{\infty} c_k(\lambda_N)(x_N^2/\lambda_N)^k,\]

where $(c_k(\lambda), k \geq 1)$ is a sequence of functions of $\lambda$ whose first two elements are

\[(2.20) \quad c_1(\lambda) = 1 - \lambda,\]
\[(2.21) \quad c_2(\lambda) = (\lambda - 1)(1 - 1/(2\lambda)),\]

Further, if $\lambda_N = 1 + O(N^{-1})$, then $p_N = 1 + O(N^{-1})$.

**Lemmas 2.4.** Define $I(\lambda, p)$ by (2.12) for $0 < \lambda < 1$ and $0 < p < 1$. If $|\lambda - p| < \min (p, q)$, then

\[(2.22) \quad I(\lambda, p) = 4(\lambda - p)^2 \left(2pq - \lambda p^2 + \lambda q^2\right) + \sum_{i=3}^{\infty} \frac{(\lambda - p)^i}{i!} \left(\frac{\lambda}{q^i} + \frac{\lambda}{(-p)^i}\right).\]
Corollary 2.5. In Theorem 2.2 assume \( \lambda_N = 1 + o(1/N) \). Then
\[
J(x_N, \lambda_N) = o(1).
\]

Corollary 2.6. Under the conditions of Theorem 2.2,
\[
(2.23) \quad J(x_N, \lambda_N) = \frac{x_N^4}{N} \sum_{i=0}^{\infty} d_i(\lambda_N) \left( \frac{x_N^2}{N} \right)^i,
\]
where \( \{d_i(\lambda), \ i \geq 0\} \) is a sequence of functions of \( \lambda \) whose first element is
\[
(2.24) \quad d_0(\lambda) = -(1-2\lambda)^2/3\lambda - \lambda.
\]
Further, if, for some integer \( k = 1, 2, \ldots, \)
\[
(2.25) \liminf(x_N^{2(k+1)} N^{-k}) > 0 \text{ and } x_N^{2(k+2)} = o(p^{k+1}),
\]
then
\[
(2.26) \quad J(x_N, \lambda_N) = o(1), \text{ if } k = 0
\]
\[
= \frac{x_N^4}{N} \sum_{i=0}^{k-1} d_i(\lambda_N) \left( \frac{x_N^2}{N} \right)^i + o(1), \text{ if } k > 0.
\]
Proofs.

We begin the proofs with some preliminaries and four lemmas which will be used in proving both Theorems 2.1 and 2.2.

To simplify notation, we will often write $a_i$ for $a_N(i)$, $\lambda_i$ for $\lambda_N$, $\sigma_\Delta$ for $\sigma_{\Delta N}$, etc., suppressing the dependence on $N$.

Let $W(1) < \ldots < W(N)$ be the order statistics for the combined sample $X_{11}, \ldots, X_{kn}$, and let

$$C_{i,j} = \begin{cases} 1 & \text{if } W(j) \text{ is from the } i\text{-th sample,} \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, \ldots, k$, $j = 1, \ldots, N$. Clearly, $C_{11} + \ldots + C_{1N} = n_1$,

$i = 1, \ldots, k$. Then, by rearranging the sum,

$$S_N = \sum_{j=1}^{N} a_j \sum_{i=1}^{k} \lambda_i C_{i,j}.$$ 

Let $\Psi = \{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$ be a set consisting of $k$ points with $z = (z_1, \ldots, z_k)$ a typical element and let $Z^{(1)}, \ldots, Z^{(N)}$ be a sequence of independent and identically distributed random vectors taking values in $\Psi$ with probabilities $p_1, \ldots, p_k$, $p_i > 0$, $i = 1, \ldots, k$, $p_1 + \ldots + p_k = 1$, that is,

$$P\{Z^{(1)} = z\} = \prod_{i=1}^{k} p_i^{z_i}.$$
Let $Z(N) = Z^{(1)} + \ldots + Z^{(N)} = [Z_1(N), \ldots, Z_k(N)]$. Define a statistic $T_N$ by

$$T_N = \sum_{j=1}^{N} u_j \sum_{i=1}^{k} a_i z_i^{(j)},$$

and let $t_N$ be a realization of $T_N$. The key to these proofs is the observation that

$$P(S_N = t_N) = P(T_N = t_N | Z(N) = n),$$

where $n = (n_1, \ldots, n_k)$, and, further, (3.5) is an identity in the probabilities $p_1, \ldots, p_k$. For each $h > 0$, define a new joint distribution for the random vectors $Z^{(1)}, \ldots, Z^{(N)}$ by

$$Q(Z^{(j)} = z^{(j)}; j = 1, \ldots, N) = \prod_{j=1}^{N} \prod_{i=1}^{k} q_{ij}^{h a_j a_i},$$

where

$$q_{ij} = p_i e^{-\frac{h a_j a_i}{\sum_{i=1}^{k} p_i e^{-h a_j a_i}}},$$

Observe that, under $Q$, $Z^{(1)}, \ldots, Z^{(N)}$ are independent random vectors taking values in $\mathcal{Y}$, but are not identically distributed.
Lemma 3.1. Let $S_N$ be given by (3.7), $T_N$ by (3.4), and $P_N(x_N)$ by
(2.3). For any $h > 0$ and any $p_1, \ldots, p_k$ with $p_1 > 0$, $i = 1, \ldots, k$, $p_1 + \ldots + p_k = 1$, we have

$$P_N(x_N) = \frac{Q(Z(N) = n)}{P(Z(N) = n)} \prod_{j=1}^{k} \{ \sum_{i=1}^{N} \frac{h_i \Delta_i}{p_i} \} = \Lambda_N,$$

where

$$A_N = \prod_{T_N = T_i} Q(T_N = T_i \mid Z(N) \equiv n),$$

and $x_{N_T}$ denotes summation over those $t_N$ satisfying $t_N \geq x_N \Delta_i$.\(^{11}\)

Proof. From (3.6) and (3.7),

$$P_N(x_N) = \left[ \prod_{T_N = T_i} Q(T_N = T_i \mid Z(N) = n) \right]$$

$$= \frac{\prod_{j=1}^{k} \frac{h_j \Delta_j}{p_j}}{\prod_{T_N = T_i} Q(T_N = T_i \mid Z(N) = n)} \prod_{x_{N_T}} Q(T_N = T_i \mid Z(N) = n).$$

which is equal to the right-hand side of (3.8). \( \therefore \) 3.1 is now proved.

Observe that (3.8) is an identity in the $k+1$ arbitrary variables $p_1, \ldots, p_k$ and $h$. We will later exploit this fact by making convenient choices for these quantities. But first we will obtain asymptotic approximations to all the factors in the right-hand side of (3.8) except $A_N$. We begin with $P(Z(N) = n)$.\(^{11}\)
Lemma 3.2. For any \( p_1, \ldots, p_k \) such that \( p_i > 0, \ i = 1, \ldots, k \) and \( p_1 + \ldots + p_k = 1 \), we have, under (2.2), as \( N \to \infty \),

\[
(3.11) \quad P(Z(N) = \mathbf{n}) \to (2\pi N)^{(k-1)/2} \prod_{i=1}^{k} \lambda_i^{-1/2} \exp \left[ -N \sum_{i=1}^{k} \lambda_i \log (\lambda_i / p_i) \right].
\]

Proof. Apply Stirling's formula.

Lemma 3.3. Let \( q_{ij} = (q_{1i} + \ldots + q_{Nj}) / N \) and \( \Sigma_j = (\sigma_{ij}) \), \( i, \ell = 1, \ldots, k-1 \), where

\[
\sigma_{ij} = q_{ij} (1 - q_{ij}) \quad \text{if } i = \ell,
\]

\[
= -q_{ij} q_{ij} \quad \text{if } i \neq \ell.
\]

Assume \( \lim_{N \to \infty} \left[ \frac{1}{N} \sum_{j=1}^{N} (n_j^2 - \Sigma_j) \right] = \Sigma \) (say) is a positive definite matrix. Then we have, uniformly in \( n \),

\[
(3.12) \quad (2\pi N)^{(k-1)/2} \left| \det \Sigma \right| (Z(N) = \mathbf{n}) \to \exp \left[ -\frac{1}{2} \mathbf{n}^t \Sigma^{-1} \mathbf{n} \right] = 0, \ N \to \infty,
\]

where \( \Sigma^{-1} = (n_1^2 - \Sigma_{11}), \ldots, n_{k-1}^2 - \Sigma_{k-1,k-1}, n_k^2 - \Sigma_{k-1,k-1}) \).

Proof. Let \( E_N = (N^{-1/2} \sum_{j=1}^{N} (Z_j(N) - \Sigma_{1j})), \ldots, n_{k-1}^{-1/2} \sum_{j=1}^{N} (Z_{k-1,j}(N) - \Sigma_{k-1,k-1}) \).

Clearly \( E_N \) is asymptotically normal with zero mean vector and covariance matrix \( \Sigma \). Kvačka (1954) has proven a local limit theorem for sums of i.i.d. lattice vectors. To prove (3.12) we follow her argument in outline, varying the details to handle our non-identically distributed vectors and taking advantage of the special structure of \( Z(N) \). See Clickner (1972) for the details.
We now need some notation. Let $E_Q$, $\text{var}_Q$, etc., denote expectation, variance, etc. under $Q$. Then

\[
E_Q T_N = \frac{N}{k} \sum_{j=1}^{k} \frac{1}{j} \Delta_1 q_{ij}
\]

and

\[
\text{var}_Q T_N = \frac{N}{k} \sum_{j=1}^{k} \Delta_1^2 q_{ij} (1 - q_{ij}).
\]

Also, define $\nu_\Delta = p_1 \Delta_1 + \cdots + p_k \Delta_k$ and

\[
\tau_\Delta^2 = \sum_{i=1}^{k} p_i (\Delta_1 - \nu_\Delta)^2.
\]

The motivation for the bound on $h$ in Lemma 3.4 will be made clear in (3.24) and (3.25).

Lemma 3.4. Let $h = h_N$, where $h_N = K \chi_N / N^\delta$, for some $K$, $0 < K < \infty$.

Assume $x_N \to \infty$, $x_N^2 / N \to 0$, and (2.4) obtains. Then, as $N \to \infty$,

\[
\sum_{j=1}^{k} \log( \sum_{i=1}^{k} p_i e^{h_N a_i \Delta_1} ) = Nh_N 2^{\Delta}/2 + O(Nh_N^3 N^2),
\]

\[
q_{ij} = p_i + h_i p_i h_N^2 (\Delta_1 - \nu_\Delta)^2 - \tau_\Delta^2] + O(h_N^3 N^2),
\]

\[
E_Q T_N = Nh_N^2 \Delta_1^2 + O(Nh_N^2 N^2),
\]

\[
\text{var}_Q T_N = N \sum_{i=1}^{k} \Delta_1^2 p_i (1 - p_i) + O(Nh_N^2 N^2),
\]
Proof. Consider the left-hand side of (3.16):

\[ \sum_{i=1}^{k} \frac{h_{N} a_{i} q_{i}}{\sum_{j=1}^{N} p_{i} e^{h_{N} a_{j} \Delta_{i}}} = 1 + h_{N} a_{j} \nu_{\Delta} + h_{N}^{2} a_{j}^{2} \sum_{i=1}^{k} p_{i} a_{i}^{2} + \frac{\theta}{6} h_{N}^{3} a_{j}^{3} \sum_{i=1}^{k} h_{N} a_{j}^{\Delta_{i}} \]

where \( |\theta| \leq 1 \). Then

\[ \sum_{j=1}^{N} \log \left( \sum_{i=1}^{k} p_{i} e^{h_{N} a_{j} \Delta_{i}} \right) = \sum_{j=1}^{N} \left[ h_{N} a_{j} \nu_{\Delta} + h_{N}^{2} a_{j}^{2} + O(h_{N}^{3} |a_{j}|^{3}) \right] \]

\[ = h_{N}^{2} N \nu_{\Delta} + O(h_{N}^{3} \sum_{j=1}^{N} |a_{j}|^{3}) \]

\[ = N h_{N}^{2} \nu_{\Delta} / 2 + O(h_{N}^{3} |a_{j}|^{3}). \]

Expressions (3.17) - (3.21) follow similarly. The proof of Lemma 3.4 is complete.
In addition to Lemmas 3.1-3.4, we need two lemmas specifically for Theorem 2.1.

**Lemma 3.5.** Define $A_N$ by (3.9). Then for each $N$ and any $h > 0$,

$$0 \leq \log A_N + \frac{h}{2} \sigma^2 N^1$$

$$\geq -4V_T h + \log(1 - E_Q[(T_N - 2V_T - x_N^D)^2 | Z(N) = n]) / 4V_T^2).$$

**Proof.** Clearly $A_N \leq \exp[-hx_N^D N^1]$. On the other hand,

$$A_N \geq \exp[-hx_N^D N^1 - 4V_T h] Q(x_N^D N^1 \leq T_N \leq x_N^D N^1 + 4V_T | Z(N) = n)$$

$$\geq \exp[-hx_N^D N^1 - 4V_T h] (1 - E_Q[(T_N - x_N^D N^1 - 2V_T)^2 | Z(N) = n]) / 4V_T^2),$$

by Chebyshev's inequality. Lemma 3.5 follows.

We will now choose values for $p_1, \ldots, p_k$ and $h$ for the $k$-sample case. Let

$$p_i = \lambda, \quad i = 1, \ldots, k$$

and let $h = h_N$ be the unique solution of

$$E_Q T^{*h} = x_N^D N^1 + 2V_T.$$
This choice of $h_N$ maximizes a term in the lower bound of $A_N$ in Lemma 3.5. A simple argument shows that $h_N$ is well defined by (3.24), and, further

\[(3.25)\quad h_N \sim n_N/\sigma^{1/2}\Delta.\]

Lemma 3.6. Let (3.23) hold and let $h_N$ be the solution of (3.24), satisfying (3.25). Then, as $N \to \infty$,

\[(3.26)\quad E_Q(T_N|Z(N) = n) \sim E^Q_{n_N} \quad \text{and} \quad \text{var}_Q(T_N|Z(N) = n) \sim \nu^2_T.\]

Proof. Consider, for $i = 1, \ldots, k$, $j = 1, \ldots, N$,

\[(3.28)\quad Q[Z(I) = 1|Z(N) = n] = q_{ij} Q[Z(N) = n],\]

where $E_{n} = (n_1, \ldots, n_{i-1}, n_i, n_{i+1}, \ldots, n_k)$, $i = 1, \ldots, k$.

From (3.17), (3.20), (3.21) and (3.23), $\Sigma = \lim N^{-1} \sum_{j=1}^{N} E_{ij}$ has elements of the form $\lambda_1 (1 - \lambda_1)$ on the main diagonal and $-\lambda_1 \lambda_2$ off the main diagonal. Since $\prod_{i=1}^{k} \lambda_i = 1$, det $\Sigma = \prod_{i=1}^{k} \lambda_i$. Hence, by Lemma 3.3,

\[(3.29)\quad \frac{Q[Z(N) = E_4]}{Q[Z(N) = E_2]} \sim \exp\left( \frac{N}{2\lambda_1} \left[ (\lambda_1 - q_{1,1})^2 - (\lambda_1 - 1/N \cdot q_{1,1})^2 \right] \right)
\]

\[= \exp\left( \frac{(\lambda_1 - q_{1,1})}{\lambda_1} - \frac{1}{2\lambda_1} \right)\]

\[= \exp\left( \frac{\sigma^2_{1,1}}{N/N} \right),\]
uniformly in $i = 1, \ldots, k, j = 1, \ldots, N$. Hence,

$$
(3.30) \quad E_Q(T_N | Z(N) = n) = \sum_{j=1}^{N} a_j \sum_{j_1=1}^{k} \Delta_1 Q(Z_{j_1}^j = 1 | Z(N) = n) \sim E_Q T_N.
$$

Similarly, for $i, i' = 1, \ldots, k, j, j' = 1, \ldots, N, j \neq j'$,

$$
(3.31) \quad Q(Z_{j_1}^j = 1, Z_{j_1}^{j'} = 1 | Z(N) = n) \sim q_{i_1} q_{i'_1},
$$

uniformly in $i, i', j$ and $j'$. The joint probability on the left-hand side of (3.31) is zero if $j = j'$. Hence

$$
(3.32) \quad E_Q(T_{N}^2 | Z(N) = n) = \sum_{j=1}^{N} a_j^2 \sum_{j_1=1}^{k} \Delta_1^2 Q(Z_{j_1}^j = 1 | Z(N) = n)
$$

$$
+ \sum_{j \neq j'} \sum_{j_1 \neq j_1'} a_{j,j_1} a_{j,j_1'} Q(Z_{j_1}^j = 1, Z_{j_1'}^{j'} = 1 | Z(N) = n) \sim E_Q T_{N}^2.
$$

Lemma 3.6 follows.

We can now prove Theorem 2.1. This will be done by substituting Lemmas 3.2-3.6 in Lemma 3.1. More specifically, let

(3.23) hold and let $h = h_N$ solve (3.24) which entails

$$
h_N \sim x_N / N^2 \sigma \alpha. \quad \text{From Lemmas 3.2 and 3.3,}
$$
(3.33) \[ \frac{Q(z(n) = n)}{P(z(n) = n)} \sim \exp\left\{- \frac{N}{2} \sum_{i=1}^{k} \frac{(\lambda_i - q_i)^2}{\lambda_i} \right\} \]

\[ = \exp\left\{0\left(\frac{x_N^h}{N}\right)\right\}, \]

using (3.17). From Lemma 3.4, specifically (3.16), we have

(3.34) \[ \frac{1}{N} \sum_{j=1}^{k} \log(\frac{1}{\lambda_j} e^{N^q_j a_j}) \sim \frac{x_N^2}{2}. \]

By the definition of \( h_N \),

\[ E_Q\left[ (T_N - x_N^h A_N^h - 2V_T)^2 | Z(N) = n \right] = \text{var}_Q(T_N | Z(N) = n) \sim V_T^2, \]

by Lemma 3.6. Hence, from Lemma 3.5,

(3.35) \[ \log A_N \sim -x_N^2. \]

Theorem 2.1 now follows by substituting (3.33), (3.34), and (3.35) in Lemma 3.1.

**Proof of Theorem 2.2**

The main difference between the proofs of Theorems 2.1 and 2.2 lies in the treatment of the sum \( A_N \), defined in (3.9). Here,
we write $A_N$ as $A_N = \exp[-hQ_{T_N}]C_N$.

where

$$C_N = \int_0^\infty e^{-hV_y} d\phi_N(y),$$

where

$$\bar{a} = \frac{\sum_{j=1}^N s_j q_{1j}(1-q_{1j})}{\sum_{j=1}^N q_{1j}(1-q_{1j})},$$

$$v^2 = \frac{\sum_{j=1}^N (s_j - \bar{a})^2 q_{1j}(1-q_{1j})}{\sum_{j=1}^N q_{1j}(1-q_{1j})},$$

$$\phi_N(y) = Q(T_N \leq y + E_{Q_{T_N}}Z(N) = m),$$

and

$$B = [x_N(\lambda N)^{1/2} - E_{Q_{T_N}}]/V.$$ 

Observe that $\phi_N(\cdot)$ is the conditional distribution referred to in the discussion following Theorem 2.2 (see equation (2.14)).

Now select $p_1$ and $h$. Choose $p_1 = p_N$ to be the solution of
(3.41) \[ n = \prod_{i=1}^{N} \left( 1 + \frac{q_i}{p_N} \right) \exp \left( -x_N^2 \left( \lambda_N^2 / N \right)^{1/2} / p_N^2 N \right) \].

This is the \( p_N \) of (2.13). Choose

(3.42) \[ h = h_N = x_N^2 (\lambda_N^2 / N)^{1/2} / p_N. \]

Note that \( p_N \) is well-defined by (3.41) and we have

\( p_N \sim \lambda_N \) and \( h_N \sim x_N / (\lambda_N^2 / N)^{1/2} \), so these choices are only slightly different from those in the proof of Theorem 2.1 -- but are more convenient for the proof of Theorem 2.2. The reasons for these choices will become apparent in Lemma 3.7 and display (3.45).

Lemmas 3.7 and 3.8 constitute the analysis of \( A_N \).

**Lemma 3.7.** (Hájek (1964)). Let (3.41) hold. A necessary and sufficient condition for \( \gamma_N \to 0 \), where

\[ \gamma_N = \sup_\lambda |\phi_N(\lambda) - \phi(\lambda)| \]

and \( \phi_N(\cdot) \) is defined in (3.39) is
Proof. This is Theorem 7.1 of Hájek (1964).

**Lemma 3.8.** Let \( p_N \) and \( h_N \) be given by (3.41) and (3.42), respectively, and let (2.9) obtain. With \( C_N \) as in (3.36), we have, as \( N \to \infty \),
\[ C_N \to 0. \]

**Proof.** We can write
\begin{align*}
(3.44) \quad C_N &= \int_B e^{-h_N Y V} \, d\phi(y) \\
&= \int_B e^{-h_N Y B} \left[ \phi_N(B) - \phi(B) \right] + hV \int_B e^{-h_N Y V} \left[ \phi_N(y) - \phi(y) \right] dy.
\end{align*}
Techniques similar to those of Lemma 3.4 yield \( V^2 \sim N \lambda_{N, N} \). Consider
\begin{align*}
(3.45) \quad B &= V^{-1} \left[ x_N (\lambda_{N, N} H) \right]_{B_{-1} (T_N)} \\
&= x_N - p q h (N / \lambda_{N, N})^k + o((Nh^2)^{-k}) \\
&= o(x_N^{-1}).
\end{align*}
with \( h \) given by (3.41). Hence

\[
(3.46) \quad C_N = [1+o(1)][(2\pi x_N^2)^{-1} + \Delta_N].
\]

It follows from (2.9) that \( A_c \) is empty for large \( N \); hence \( \Delta_N \to 0 \) by Lemma 3.7. Since \( x_N \to \infty \), Lemma 3.8 follows.

To complete the proof of Theorem 2.2, apply condition (2.11) and the selections of \( p_1 \) and \( h \), (3.42) and (3.43), to Lemma 3.1 - 3.4 with \( k=2 \) to obtain

\[
(3.47) \quad P_N(x_N) \approx C_N \exp[-x_N^2/2 + J(x_N, \lambda_N)].
\]

Proof of Lemmas 2.3 and 2.4

Lemma 2.4 is proved by expanding the logarithms.

In Lemma 2.3, the case \( \lambda_N = t_1 + 0(N^{-1}) \) is proved by substituting \( p_N = t_1 + 0(N^{-1}) \) to verify that it is a solution of (2.13). Otherwise, recall that

\[
(3.48) \quad \lambda_N = p + (p-t_1)\lambda_N \frac{\bar{w}^2/pqN + o(N^{-1})}{N}.
\]

using (3.17) and (3.42). Now, suppose \( x_N^4 = o(N) \) and propose

\[
(3.49) \quad p = \lambda_N + c_1 x_N^2/N + o(N^{-1})
\]

as a solution of (3.48). Substitute (3.49) in (3.48) and solve for
to obtain \( c_1 = c_1(\lambda_N) \), \( \lambda_N \), as in (2.20). Next, suppose

\[ x_N^3 = o(N) \]  

and try

\[ p = \lambda_N c_1(\lambda_N)x_N^2/\Pi + c_2(\lambda_N)^{1/2}/\Pi^2 + o(\Pi^{-1}) \]  

as a solution of (3.48). Again, solve for \( c_2 \) to obtain \( c_2 = c_2(\lambda_N) \), as in (2.21). The higher order coefficients \( c_k(\lambda_N) \) can be found successively by continuing this iterative procedure. This is not done here because the algebra becomes very cumbersome. The proof of Lemma 2.3 is complete.

**Proof of Corollaries 2.5 and 2.6**

Corollary 2.5 is an immediate consequence of Lemmas 2.3 and 2.4, and Corollary 2.6 is proved in essentially the same manner as Lemma 2.3.
REFERENCES


