CHAPTER 3

THE LIMIT DISTRIBUTION OF U

3.1 DEFINITION OF U

Let

$$
(3.1) \qquad
\begin{array}{c}
X_{11}, X_{12}, \ldots, X_{1n_1} \\
X_{21}, X_{22}, \ldots, X_{2n_2} \\
\cdots\cdots\cdots\cdots \\
X_{l1}, X_{l2}, \ldots, X_{ln_l}
\end{array}
$$

be independent random variables such that the random variables in the same row have the same distribution.

Let $f(x_1, \ldots, x_l)$ be a real-valued function which is bounded and integrable. For $1 \leq i_1 \leq n_1, \ldots, 1 \leq i_l \leq n_l$, put

$$(3.2) \qquad U(i_1, \ldots, i_l) = f(X_{1i_1}, \ldots, X_{ln_l}).$$

In this chapter we shall study the asymptotic distribution of

$$(3.3) \qquad U = \sum_{i_l=1}^{n_l} \cdots \sum_{i_1=1}^{n_1} U(i_1, \ldots, i_l).$$

In section 3.3 we show that U is asymptotically normally distributed when $n_1, \ldots, n_l$ are large. Moreover, if $f_1, \ldots, f_k$ are bounded integrable function of l variables, we also show that asymptotic distribution of the random vector

$$U = (U_1, \ldots, U_k) ,$$

where $$U_p = \sum_{i_1=1}^{n_1} \cdots \sum_{i_l=1}^{n_1} f_p(X_{1i_1}, \ldots, X_{1i_l}), \; p = 1, \ldots, k ,$$

is a k-variate normal distribution.

3.2 MOMENTS OF U

DEFINITION 3.2.1 $U(i_1, \ldots, i_l)$ and $U(i_1', \ldots, i_l')$ are said to be linked if and only if there exists k such that $1 \leq k \leq l$ and $i_k = i_k'$. A pair of such $U(i_1, \ldots, i_l)$, $U(i_1', \ldots, i_l')$ will be called a link. Note that if $U(i_1, \ldots, i_l)$ and $U(i_1', \ldots, i_l')$ are not linked then they are Stochastically independent.

DEFINITION 3.2.2 Any set S of $U(i_1, \ldots, i_l)$'s is said to be a connected if and only if for any U and U' in S there exists a finite sequence $U^{(1)}, \ldots, U^{(n)}$ in S such that $U = U^{(1)}$, $U' = U^{(n)}$ and for each $k = 1, \ldots, n-1$, $U^{(k)}$ and $U^{(k+1)}$ are linked.

DEFINITION 3.2.3 Let S, S' be any sets of $U(i_1, \ldots, i_l)$. We say that S, S' are separated if for each $U \in S$, $U' \in S'$, U and U' are not linked.

DEFINITION 3.2.4 Any family of sets of $U(i_1, \ldots, i_l)$ is said to be a family of separated sets if every two sets in it are separated.

PROPOSITION 3.2.1 If $\{S_1, \ldots, S_u\}$ is a family of separated sets of $U(i_1, \ldots, i_l)$'s, then $E(\prod_{U \in \bigcup_{k=1}^{u} S_k} U) = \prod_{k=1}^{u} E(\prod_{U \in S_k} U)$.

Proof Since $S_1, \ldots, S_u$ are separated, hence $\prod_{U \in S_1} U, \ldots, \prod_{U \in S_u} U$ are functions of disjoint sets of independent random variables. Therefore, they are independent. So, we have

$$E\Big(\prod_{U \in \bigcup_{k=1}^{u} S_k} U\Big) = \Big(E \prod_{U \in S_1} U\Big) \cdot \Big(E \prod_{U \in S_2} U\Big) \cdots \Big(E \prod_{U \in S_u} U\Big)$$

$$= \prod_{k=1}^{u} E\Big(\prod_{U \in S_k} U\Big) .$$

PROPOSITION 3.2.2 Let S be a connected set consisting of t elements, $U(i_{1p},\ldots, i_{lp})$, $p = 1,\ldots, t$. If j_k is the number of distinct values of i_k's occuring among $U(i_{1p},\ldots,i_{lp})$ in S. Then

$$\sum_{k=1}^{l} j_k \leqq (l-1)t+1 .$$

Proof Since S is connected and has t elements, there must exist at least t-1 links among the $U(i_1,\ldots, i_l)$. For each link, there correspond an equal value of i_k in the pair of $U(i_1,\ldots,i_l)$'s that are linked. Hence to the t-1 links there correspond t-1 equal values of i_k's.

Hence there are at most $lt-(t-1) = (l-1)t+1$ distinct i_k among the $U(i_1,\ldots, i_l)$'s in S. Hence we have

$$\sum_{k=1}^{l} j_k \leqq (l-1)t + 1 .$$

PROPOSITION 3.2.3 Let $A(t ; n_1,\ldots, n_l)$ denote the number of ways that we can form a connected set consisting of exactly t of the $U(i_1,\ldots, i_l)$, $i_k = 1,\ldots, n_k$; $k = 1,\ldots, l$, not necessary distinct. Then

$$A(t ;n_1,\ldots,n_l) \leqq c(l,t) \sum \cdots \sum_{j_1+\cdots+j_l \leqq (l-1)t+1} n_1^{j_1} n_2^{j_2} \cdots n_l^{j_l} \text{ ways,}$$

where $c(l,t)$ is a number depending only on l and t.

Proof There are j_k values of i_k. We can assign j_k values to i_k in the t U's in

$$\sum_{a_1+\dots+a_{j_k}=t} \frac{t!}{a_1!\dots a_{j_k}!} n_k(n_k-1)\dots(n_k-j_k+1) \text{ ways.}$$

Hence there are at most

$$\sum_{j_1+\dots+j_l\leq(l-1)t+1} \prod_{k=1}^{l} \sum_{a_1+\dots+a_{j_k}=t} \frac{t!}{a_1!\dots a_{j_k}!} n_k(n_k-1)\dots(n_k-j_k+1)$$

$$= \sum_{j_1+\dots+j_l\leq(l-1)t+1} \prod_{k=1}^{l} n_k(n_k-1)\dots(n_k-j_k+1) \sum_{a_1+\dots+a_{j_k}=t} \frac{t!}{a_1!\dots a_{j_k}!}$$

$$= \sum_{j_1+\dots+j_l\leq(l-1)t+1} \prod_{k=1}^{l} n_k(n_k-1)\dots(n_k-j_k+1)(j_k)^t$$

$$= \sum_{j_1+\dots+j_l\leq(l-1)t+1} \left(\prod_{k=1}^{l}(j_k)^t\right)\left(\prod_{k=1}^{l} n_k(n_k-1)\dots(n_k-j_k+1)\right)$$

$$\leq \sum_{j_1+\dots+j_l\leq(l-1)t+1} \prod_{k=1}^{l}(j_k)^t\, n_1^{j_1} n_2^{j_2}\dots n_l^{j_l} .$$

Let $c(l,t) = \max\limits_{j_1+\dots+j_l\leq(l-1)t+1} \prod_{k=1}^{l}(j_k)^t$, then

$$\sum_{j_1+\dots+j_l\leq(l-1)t+1} \prod_{k=1}^{l}(j_k)^t n_1^{j_1}\dots n_l^{j_l} \leq c(l,t) \sum_{j_1+\dots+j_l\leq(l-1)t+1} n_1^{j_1}\dots n_l^{j_l} .$$

Hence there are at most $c(l,t) \sum\limits_{j_1+\cdots+j_l \leq (l-1)t+1} n_1^{j_1} \cdots\cdots n_l^{j_l}$ ways of assingning values to $i_1, \ldots, i_l$ in the t U's to have a connected set with t U's.

Hence $A(t;n_1,\ldots,n_l) = c(l,t) \sum \cdots \sum\limits_{j_1+\cdots+j_l \leq (l-1)t+1} n_1^{j_1} \cdots\cdots n_l^{j_l}$.

<u>LEMMA 3.2.1</u> Let $n_k = a_k n$ where a_k's are constants for $k = 1,\ldots,l$. Let $N(t_1,\ldots,t_u)$ be the number of ways of forming families of u separated connected set $S_1,\ldots, S_u$ such that

(1) $|S_j| = t_j \geq 2$ for all j ,

(2) $|S_j| = t_j \geq 3$ for some j ,

(3) $|\cup S_j| = m$,

then $N(t_1,\ldots, t_u) \leq O(n^{\frac{m(2l-1)-1}{2}})$.

<u>Proof</u> Let $S_j = \left\{ U^{(k)} \,/\, T_{j-1}+1 \leq k \leq T_j \right\}$

such that

$$T_0 = 0$$

and

$$T_j = T_{j-1}+t_j \quad \text{for } j = 1,\ldots,u \text{ and } k = 1,\ldots,m.$$

From (1) and (2) we obtain

(4) $2u + 1 \leq t_1+\cdots+ t_u$.

Since S_j are separated, hence $\sum\limits_{j=1}^{u} t_j = |\cup S_j|$.

Therefore, from (3) and (4) we obtain

$$2u + 1 \leqq m$$

so we have

(5) $$u \leqq \frac{m-1}{2} .$$

Since connected set S_j has t_j elements, so from Proposition 3.2.2, the number of distinct values of i_k's occuring among U's in S_j is at most $(l-1)t_j + 1$, then the number of ways of forming S_j is $A(t_j;\ n_1,\ldots,n_l)$. Hence the number of ways of partition m U's into u separated sets S_j's is at most $\prod_{j=1}^{u} A(t_j;\ n_1,\ldots,n_l)$, which is less than or equal to $\prod_{j=1}^{u} \left[c(l,t_j) \sum_{j_1+\cdots+j_u \leqq (l-1)t_j+1} \sum n_1^{j_1} \cdots\cdots n_l^{j_l} \right]$.

The last expression is a polynomial in $n_1,\ldots, n_l$ of degree

$$\sum_{j=1}^{u} \left[(l-1)t_j + 1 \right] = (l-1)\sum_{j=1}^{u} t_j + u = \frac{m(2l-1)-1}{2} .$$

When each n_k is replaced by $a_k n$, this expression is a polynomial in n of degree $\frac{m(2l-1)-1}{2}$. Hence we have

$$N(t_1,\ldots, t_u) \leqq O(n^{\frac{m(2l-1)-1}{2}}).$$

<u>LEMMA 3.2.2</u> Let $n_k = a_k n$ where a_k's are constants for $k = 1,\ldots,l$. Define $V(i_1,\ldots, i_l) = U(i_1,\ldots, i_l) - E\left[U(i_1,\ldots, i_l)\right]$

and $$V = \sum_{i_1=1}^{n_1} \cdots \sum_{i_l=1}^{n_l} V(i_1,\ldots,i_l).$$

Then $$\mu_{2k}(V) = \frac{(2k)!\,n^{k(2l-1)}}{2^k(k!)}(a_1\ldots a_l)^{2k}\Big(\sum_{q=1}^{l} \frac{\gamma_q}{a_q}\Big)^k + o(n^{k(2l-1)}),$$

where $$\gamma_q = E\left[V(i_1,\ldots,i_l)V(i_1',\ldots,i_l')\right]$$

with $i_j = i_j'$ if and only if $j = q$.

<u>Proof</u> For convenience, let us introduce some notations.

Let

$$I = I_1 \times I_2 \times \ldots \times I_l ,$$

where $I_q = \{1, 2,\ldots, n_q\}$, $q = 1,\ldots,l$.

For each $\alpha = (i_1,\ldots,i_l) \in I$ we denote $V(i_1,\ldots,i_l)$ by $V(\alpha)$.

Hence $$\mu_{2k}(V) = E(V^{2k})$$

$$= E\Big[\sum_{\alpha\in I} V(\alpha)\Big]^{2k}$$

$$= E\Big[\prod_{j=1}^{2k}\Big(\sum_{\alpha_j\in I} V(\alpha_j)\Big)\Big]$$

$$= \sum_{\alpha_{2k}\in I}\cdots\sum_{\alpha_1\in I} E\Big[\prod_{j=1}^{2k} V(\alpha_j)\Big]$$

$$= \sum^{(1)} E\Big[\prod_{j=1}^{2k} V(\alpha_j)\Big] + \sum^{(2)} E\Big[\prod_{j=1}^{2k} V(\alpha_j)\Big] + \sum^{(3)} E\Big[\prod_{j=1}^{2k} V(\alpha_j)\Big],$$

where

$\sum^{(1)}$ is extended over the expectation of the products of 2k $V'(\alpha_j)$s which can be formed into a family of separated connected sets such that in the family there exists at least one separated set which has one element.

$\sum^{(2)}$ is extended over the expectation of the products of 2k $V'(\alpha_j)$s which can be formed into a family of separated connected sets such that any set in it has exactly two elements.

$\sum^{(3)}$ is extended over the expectation of the products of 2k $V'(\alpha_j)$s which can be formed into a family of separated connected sets such that every set in it has at least two elements and there exists at least one separated connected set which has at least three elements.

Note that each term in the summation $\sum^{(1)}$ has $E(V(\alpha_j))$, which is zero, as a factor. Hence

$$(1) \qquad \sum^{(1)} E\left[\prod_{j=1}^{2k} V(\alpha_j)\right] = 0 .$$

Each term $E\left[\prod_{j=1}^{2k} V(\alpha_j)\right]$ in $\sum^{(2)}$ can be factored into the form

$$E\left[\prod_{j=1}^{2k} V(\alpha_j)\right] = E\left[V(\alpha_{j_1})V(\alpha_{j_2})\right] E\left[V(\alpha_{j_3})V(\alpha_{j_4})\right] \ldots E\left[V(\alpha_{j_{2t-1}})V(\alpha_{j_{2t}})\right]$$

where $V(\alpha_{j_{2t-1}})$, $V(\alpha_{j_{2t}})$ are linked. Here $j_1, \ldots, j_{2k}$ is a permutation of $1, 2, \ldots, 2k$. To each permutation

$$p = \begin{pmatrix} 1 & 2 & \ldots & 2k \\ j_1 & j_2 & \ldots & j_{2k} \end{pmatrix},$$

let

$$\sum^{(p)} \prod_{t=1}^{k} E\left[V(\alpha_{j_{2t-1}})V(\alpha_{j_{2t}})\right]$$

denote the sum of all $\prod_{t=1}^{k} E\left[V(\alpha_{j_{2t-1}})V(\alpha_{j_{2t}})\right]$ such that each pair $V(\alpha_{2t-1})$, $V(\alpha_{2t})$ are linked but each of them is not linked with any other $V(\alpha_j)$.

Note that distinct permutations may give rise to the same sum in $\sum^{(2)}$, for example

$$p_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \ldots & 2k \\ a & b & c & d & j_5 & j_6 \ldots & j_{2k} \end{pmatrix},$$

$$p_2 \neq \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \ldots & 2k \\ b & a & c & d & j_5 & j_6 \ldots & j_{2k} \end{pmatrix},$$

$$p_3 \neq \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \ldots & 2k \\ c & d & b & a & j_5 & j_6 \ldots & j_{2k} \end{pmatrix},$$

will give rise the same sum. Observe that

$$p = \begin{pmatrix} 1 & 2 & \cdots & 2k \\ j_1 & j_2 & \cdots & j_{2k} \end{pmatrix},$$

and

$$p' = \begin{pmatrix} 1 & 2 & \cdots & 2k \\ j_1' & j_2' & \cdots & j_{2k}' \end{pmatrix},$$

will give rise to distinct sums $\sum^{(p)}$, $\sum^{(p')}$ if and only if $\left\{\{j_1,j_2\},\ldots,\{j_{2k-1},j_{2k}\}\right\}$ and $\left\{\{j_1',j_2'\},\ldots,\{j_{2k-1}',j_{2k}'\}\right\}$ are distinct permutations of $\{1,2,\ldots,2k\}$. Since there are precisely $\frac{(2k)!}{2^k(k!)}$ such partitions of $\{1,2,..,2k\}$, hence we have exactly $\frac{(2k)!}{2^k(k!)}$ permutation p that give rise to different $\sum^{(p)}$. Let P denote a set of such permutations. Then

$$\sum^{(2)} E\left[\prod_{j=1}^{2k} V(\alpha_j)\right] = \sum_{p\in P} \sum^{(p)} \prod_{t=1}^{k} E\, V(\alpha_{j_{2t-1}})V(\alpha_{j_{2t}}) \ .$$

Observe that $\sum^{(p)} \prod_{t=1}^{k} E\left[V(\alpha_{j_{2t-1}})V(\alpha_{j_{2t}})\right]$ are the same for all $p \in P$. Hence

$$\sum^{(2)} E\left[\prod_{j=1}^{2k} V(\alpha_j)\right] = \frac{(2k)!}{2^k(k!)} \sum^{(p_o)} \prod_{j=1}^{k} E\left[V(\alpha_{2j-1})V(\alpha_{2j})\right],$$

where p_o is the identity permutation. The summation $\sum^{(p_o)}$ is taken over all $\alpha_1, \ldots, \alpha_{2k} \in I$ such that $V(\alpha_{2j-1})$, $V(\alpha_{2j})$ are linked but each is not linked to any other α's.

Let J denote the set of all $(\alpha_1, \ldots, \alpha_{2k})$ such that α_{2j-1}, α_{2j} have at least one common component but each has no common component with any other α's. Hence we may write

$$\sum^{(p_o)} \prod_{j=1}^{k} E\left[V(\alpha_{2j-1})V(\alpha_{2j})\right] = \sum \cdots \sum_{(\alpha_1, \ldots, \alpha_{2k}) \in J} \prod_{j=1}^{k} E\left[V(\alpha_{2j-1})V(\alpha_{2j})\right].$$

To each $(\alpha_1, \ldots, \alpha_{2k}) \in J$, let k_q be the number of j's such that α_{2j-1}, α_{2j} have common q^{th} component, $q = 1, \ldots, 2k$. Hence for $(\alpha_1, \ldots, \alpha_{2k}) \in J$ we have

$$\sum_{q=1}^{l} k_q \geqq k .$$

Let J_1 denote the set of $(\alpha_1, \ldots, \alpha_{2k})$ such that

$$\sum_{q=1}^{l} k_q = k ,$$

and J_2 denote the set of $(\alpha_1, \ldots, \alpha_{2k})$ such that

$$\sum_{q=1}^{l} k_q > k .$$

So that we have

$$\sum^{(p_o)} \prod_{j=1}^{k} E\left[V(\alpha_{2j-1})V(\alpha_{2j})\right] = \sum_{(\alpha_1,\ldots,\alpha_{2k})\in J_1} \cdots \sum \prod_{j=1}^{k} E\left[V(\alpha_{2j-1})V(\alpha_{2j})\right]$$

$$+ \sum_{(\alpha_1,\ldots,\alpha_{2k})\in J_2} \cdots \sum \prod_{j=1}^{k} E\left[V(\alpha_{2j-1})V(\alpha_{2j})\right].$$

When $2k = 2$, we have

$$\sum\sum_{(\alpha_1,\alpha_2)\in J_1} E\left[V(\alpha_1)V(\alpha_2)\right] = \sum_{i_1=1}^{n_1} \cdots \sum_{i_l=1}^{n_l} \sum_{i_1'=1}^{n_1} \cdots \sum_{i_l'=1}^{n_l} E\left[V(i_1,\ldots,i_l)V(i_1',\ldots,i_l')\right],$$

where the summation is taken over all values of $i_1,\ldots,i_l,i_1',\ldots,i_l'$ such that one and only one of the following's occurs :

$$i_1 = i_1', \ldots, i_l = i_l'$$

There are

$n_1n_2(n_2-1)\ldots\, n_l(n_l-1)$ terms in which $i_1 = i_1'$,

$n_1(n_1-1)n_2n_3(n_3-1)\ldots\, n_l(n_l-1)$ terms in which $i_2 = i_2'$,

. . . .

. . . .

$n_1(n_1-1)\ldots\ldots\ldots\, n_{l-1}(n_{l-1}-1)n_l$ terms in which $i_l = i_l'$.

Hence, we have

$$\sum\sum_{(\alpha_1,\alpha_2)\in J_1} E\left[V(\alpha_1)V(\alpha_2)\right] = n_1n_2(n_2-1)\ldots n_l(n_l-1)\,\gamma_1$$

$$+ n_1(n_1-1)n_2n_3(n_3-1)\ldots n_l(n_l-1)\,\gamma_2$$

$$+ \ldots\ldots\ldots\ldots$$

$$+ n_1(n_1-1)\ldots\ldots\ldots n_{l-1}(n_{l-1}-1)n_l\,\gamma_l .$$

Setting $n_1 = a_1 n, \; n_2 = a_2 n, \ldots, \; n_l = a_l n$, we have

$$\sum_{(\alpha_1,\alpha_2)}\sum_{J_1} E[V(\alpha_1)V(\alpha_2)] = [n^{2l-1}+o(n^{2l-1})][a_1 a_2^2 \cdots a_l^2 \gamma_1 + \cdots + a_1^2 \cdots a_{l-1}^2 a_l \gamma_l]$$

$$= [n^{2l-1}+o(n^{2l-1})][(a_1^2 \cdots a_l^2)\frac{\gamma_1}{a_1} + \cdots + (a_1^2 \cdots a_l^2)\frac{\gamma_l}{a_l}$$

$$= n^{2l-1}(a_1 \cdots a_l)^2 \sum_{q=1}^{l} \frac{\gamma_q}{a_q} + o(n^{2l-1}).$$

In general, it can be shown that

$$\sum_{(\alpha_1,\ldots,\alpha_{2k})} \cdots \sum_{J_1} \prod_{j=1}^{k} E[V(\alpha_{2j-1})V(\alpha_{2j})] = n^{k(2l-1)}(a_1 \cdots a_l)^{2k}\left(\sum_{q=1}^{l}\frac{\gamma_q}{a_q}\right)^k + o(n^{k(2l-1)}).$$

By a similar argument, it can be shown that

$$\sum_{(\alpha_1,\ldots,\alpha_{2k})} \cdots \sum_{J_2} \prod_{j=1}^{k} E[V(\alpha_{2j-1})V(\alpha_{2j})] = o(n^{k(2l-1)}).$$

Hence

$$(2) \quad \sum^{(2)} E \prod_{j=1}^{2k} V(\alpha_j) = \frac{(2k)!}{2^k (k!)} n^{k(2l-1)}(a_1 \cdots a_l)^{2k}\left(\sum_{q=1}^{l}\frac{\gamma_q}{a_q}\right)^k + o(n^{k(2l-1)}).$$

By Lemma 3.2.1, we see that the number of terms in $\sum^{(3)}$ is at most $O(n^{\frac{2k(2l-1)-1}{2}})$. Hence

$$(3) \quad \sum^{(3)} E\left[\prod_{j=1}^{2k} V(\alpha_j)\right] \leq O(n^{\frac{2k(2l-1)-1}{2}}) = o(n^{k(2l-1)}).$$

From (1), (2) and (3) we obtain

$$\mu_{2k}(V) = \frac{(2k)!}{2^k(k!)} n^{k(2l-1)}(a_1\cdots a_l)^{2k}\left(\sum_{q=1}^{l} \frac{\gamma_q}{a_q}\right)^k + o(n^{k(2l-1)}).$$

LEMMA 3.2.3 Let $n_k = a_k n$ where a_k's are constants for $k = 1,\ldots, l$.

We claim that $\mu_{2k+1}(V) = o(n^{\frac{(2k+1)(2l-1)}{2}})$,

where $V = \sum_{i_1=1}^{n_1} \cdots \sum_{i_l=1}^{n_l} V(i_1,\ldots,i_l)$

and $V(i_1,\ldots,i_l) = U(i_1,\ldots,i_l) - E[U(i_1,\ldots,i_l)]$.

Proof Let I, α, $V(\alpha)$ be the same as in Lemma 3.2.2.

Hence

$$\mu_{2k+1}(V) = E[V^{2k+1}]$$

$$= E\left[\sum_{\alpha\in I} V(\alpha)\right]^{2k+1}$$

$$= E\left[\prod_{j=1}^{2k+1} \sum_{\alpha_j\in I} V(\alpha_j)\right]$$

$$= \sum_{\alpha_{2k+1}\in I} \cdots \sum_{\alpha_1\in I} E\left[\prod_{j=1}^{2k+1} V(\alpha_j)\right]$$

$$= \sum^{(1)} E\left[\prod_{j=1}^{2k+1} V(\alpha_j)\right] + \sum^{(2)} E\left[\prod_{j=1}^{2k+1} V(\alpha_j)\right],$$

where

$\sum^{(1)}$ is extended over the expectation of the products of 2k+1 $V'(\alpha_j)$s which can be formed into a family of separated connected sets such that in the family there exists at least one separated set which has exactly one element.

$\sum^{(2)}$ is extended over the expectation of the products of 2k+1 $V'(\alpha_j)$s which can be formed into a family of separated connected sets such that any set in it has at least two elements.

Note that each term in the summation $\sum^{(1)}$ has $E[V(\alpha_j)]$, which is zero, as a factor. Hence $\sum^{(1)} E\left[\prod_{j=1}^{2k+1} V(\alpha_j)\right] = 0$.

Since 2k+1 is odd, we may apply Lemma 3.2.1 to obtain an upper bound for the number of terms in $\sum^{(2)}$. It can be seen that there are atmost $O(n^{\frac{(2k+1)(2l-1)-1}{2}})$ terms in $\sum^{(2)}$. Hence

$$\sum^{(2)} E\left[\prod_{j=1}^{2k+1} V(\alpha_j)\right] \lessapprox O(n^{\frac{(2k+1)(2l-1)-1}{2}}) = o(n^{\frac{(2k+1)(2l-1)}{2}}).$$

Therefore,

$$\mu_{2k+1}(V) = o(n^{\frac{(2k+1)(2l-1)}{2}}).$$

3.3 THE ASYMPTOTIC DISTRIBUTION OF U

THEOREM 3.3.1 For each positive integer n, let F_n be the distribution function of $X = \frac{U - E(U)}{\sqrt{Var(U)}}$, where U is defined as in (3.3) with $n_q = a_q n$, where $a_1, \ldots, a_l$ are positive constants. Let γ_q be as defined in Lemma 3.2.2. If $\sum_{q=1}^{l} \frac{\gamma_q}{a_q} \neq 0$, then

$$\lim_{n \to \infty} F_n(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt .$$

Proof Let $V = U - E(U)$. Hence $Var(V) = Var(U)$.

Therefore
$$\mu_m(X) = E\left[\frac{V}{\sqrt{Var(V)}}\right]^m$$

$$= \frac{E(V)^m}{(Var(V))^{\frac{m}{2}}} .$$

Note that Lemma 3.2.2 and Lemma 3.2.3 are applicable to our random variable V. When m is replaced by 2k, we obtain

$$\mu_{2k}(X) = \frac{\frac{(2k)!}{2^k(k!)} n^{k(2l-1)} (a_1 \ldots a_l)^{2k} \left(\sum_{q=1}^{l} \frac{\gamma_q}{a_q}\right)^k + o(n^{k(2l-1)})}{\left[n^{(2l-1)} (a_1 \ldots a_l)^2 \sum_{q=1}^{l} \frac{\gamma_q}{a_q} + o(n^{2l-1})\right]^k}$$

$$= \frac{(2k)!}{2^k(k!)} + o(1) ,$$

hence $\mu_m(X) = \frac{m!}{2^{\frac{m}{2}}(\frac{m}{2})!} + o(1)$, when m is an even integer.

When m is replaced by 2k+1 , we obtain

$$\mu_{2k+1}(X) = \frac{o(n^{\frac{(2k+1)(2l-1)}{2}})}{\left[n^{(2l-1)}(a_1 \dots a_l)^2 \sum_{q=1}^{l} \frac{\gamma_q}{a_q} + o(n^{(2l-1)})\right]^{\frac{2k+1}{2}}}$$

$$= o(1) ,$$

hence $\mu_m(X) = o(1)$ when m is an odd integer.

So we have

$$\lim_{n\to\infty} \mu_m(X) = \begin{cases} \dfrac{m!}{2^{\frac{m}{2}}(\frac{m}{2})!} & \text{when m is even integer,} \\ 0 & \text{when m is odd integer,} \end{cases}$$

so by Corollary 2.1 we obtain

$$\lim_{n\to\infty} F_n(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt .$$

THEOREM 3.3.2 Let f_p, $p = 1,\dots,k$, be bounded integrable function of l variables and let (X_{ji}), $i = 1,\dots,n_j$, $j = 1,\dots,l$ be a syetem of independent random variables such that for each j, X_{ji} are identically distributed. For $p = 1,\dots,k$, let $U_p = \sum_{i_1=1}^{n_1} \dots \sum_{i_l=1}^{n_l} f_p(X_{1i_1},\dots,X_{li_l})$. Assume that the covariance matrix (σ_{pq}) of $(U_1,\dots, U_k)$ is non-singular.

Let $\gamma_q(t_1,\ldots,t_k)$ be γ_q of Lemma 3.2.2 with

$$f(x_1,\ldots,x_l) = \sum_{p=1}^{k} \frac{t_p}{\sqrt{\sigma_{pq}}} f_p(x_1,\ldots,x_l) .$$

If $n_q = a_q n$, where $a_1,\ldots,a_l$ are positive constants such that

$$\sum_{q=1}^{l} \frac{\gamma_q(t_1,\ldots,t_k)}{a_q} \neq 0$$

for all $(t_1,\ldots,t_k) \neq (0,\ldots,0)$, then $(U_1,\ldots,U_k)$ is asymptotically normally distributed when n is large.

Proof Let $\varphi_n(t_1,\ldots,t_k)$ be the characteristic function of the joint distribution of $\dfrac{U_p - E(U_p)}{\sqrt{Var(U_p)}}$ for $p = 1,\ldots,k$.

We obtain

$$\varphi_n(t_1,\ldots,t_k) = E\left\{ e^{i\left[\frac{U_1-E(U_1)}{\sqrt{Var(U_1)}} t_1+\ldots+\frac{U_k-E(U_k)}{\sqrt{Var(U_k)}} t_k\right]}\right\}$$

$$= E\,(e^{iXt}) ,$$

where

$$X = \frac{\frac{t_1U_1}{\sqrt{\sigma_{11}}} +\ldots+ \frac{t_kU_k}{\sqrt{\sigma_{kk}}} - E\left[\frac{t_1U_1}{\sqrt{\sigma_{11}}} +\ldots+ \frac{t_kU_k}{\sqrt{\sigma_{kk}}}\right]}{\sqrt{Var\left(\frac{t_1U_1}{\sqrt{\sigma_{11}}} +\ldots+ \frac{t_kU_k}{\sqrt{\sigma_{kk}}}\right)}}$$

and $$t = \sqrt{Var\left(\frac{t_1U_1}{\sqrt{\sigma_{11}}} +\ldots+ \frac{t_kU_k}{\sqrt{\sigma_{kk}}}\right)} .$$

I16606735-

Note that

$$\frac{t_1U_1}{\sqrt{\sigma_{11}}}+\dots+\frac{t_kU_k}{\sqrt{\sigma_{kk}}} = \sum_{i_1=1}^{n_1}\dots\sum_{i_1=1}^{n_1}\left[\frac{t_1}{\sqrt{\sigma_{11}}} f_1(X_{1i_1},\dots,X_{1i_1})+\dots+\frac{t_k}{\sqrt{\sigma_{kk}}}f_k(X_{1i_1},\dots,X_{1i_1}\right.$$

$$= \sum_{i_1=1}^{n_1}\dots\sum_{i_1=1}^{n_1} f(X_{1i_1},\dots, X_{1i_1}) ,$$

where $f(x_1,\dots,x_1) = \frac{t_1}{\sqrt{\sigma_{11}}} f_1(x_1,\dots,x_1)+\dots+ \frac{t_k}{\sqrt{\sigma_{kk}}} t_k(x_1,\dots, x_1)$

which is also bounded and integrable.

Let $U = \sum_{i_1=1}^{n_1}\dots\sum_{i_1=1}^{n_1} f(X_{1i_1},\dots,X_{1i_1})$, then

$$X = \frac{U - E(U)}{\sqrt{Var(U)}} .$$

Hence by Theorem 2.1.2, Theorem 2.2.4, Theorem 2.2.6 and Theorem 3.3.1, we see that

$$\lim_{n\to\infty} E(e^{iXt}) = e^{-\frac{t^2}{2}}$$

i.e. $\lim_{n\to\infty} \varphi_n(t_1,\dots,t_k) = e^{-\frac{1}{2}Var\left[\frac{t_1U_1}{\sqrt{\sigma_{11}}}+\dots+\frac{t_kU_k}{\sqrt{\sigma_{kk}}}\right]}$.

But $Var\left[\frac{t_1U_1}{\sqrt{\sigma_{11}}}+\dots+\frac{t_kU_k}{\sqrt{\sigma_{kk}}}\right] = \sum_{q=1}^{k}\sum_{p=1}^{k}\frac{t_pt_q\sigma_{pq}}{\sqrt{\sigma_{pp}\sigma_{qq}}}$, hence $\lim_{n \to \infty} \varphi_n(t_1, \ldots, t_k) = e^{-\frac{1}{2} \sum_{q=1}^{k} \sum_{p=1}^{k} \frac{t_p t_q \sigma_{pq}}{\sqrt{\sigma_{pp} \sigma_{qq}}}}$,

which is the characteristic function of the multivariate normal distribution.

We see that $\varphi_n(t_1, \ldots, t_k)$, which is the characteristic function of $\frac{U_p - E(U_p)}{\sqrt{Var(U_p)}}$ for $p = 1, \ldots, k$, has its limit as the characteristic function of multivariate normal distribution, hence $(U_1, \ldots, U_k)$ is asymptotically normally distributed when n is large.