ON SOME LIMIT THEOREMS OF PROBABILITY DISTRIBUTIONS*

BY KINSAKU TAKANO

(Received June 1, 1954)

Summary

In Part I the Khintchine's uniqueness theorem for the class convergence of probability distributions is proved in a natural way by making use of inverses of distribution functions; its generalization to the multi-dimensional case is also proved; relations between different paired sequences of scaling constants and centering constants in limit problems of probability distributions are given; and the general method to determine scaling constants and centering constants is presented.

In Part II both an analytical derivation of the P. Lévy's canonical form of the infinitely divisible multi-dimensional probability distribution and a necessary and sufficient condition, for the distributions of sums of asymptoticaly uniformly negligible independent multi-dimensional random variables to converge to a given infinitely divisible probability distribution are given. The logarithms of non-vanishing characteristic functions are treated rigorously.

In Part III various versions of the multi-dimensional central limit theorem on sums of independent random variables are studied.

The results in the last two parts are extensions of the known facts in the one-dimensional case to the multi-dimensional case.

Introduction

Let G_1 and G_2 be two p-variate distribution functions. If there are a positive number a and a p-dimensional vector b such that $G_1(x)=G_2(ax+b)$ for all $x \in R_p$, then G_1 and G_2 are said to belong to the same class. Let us denote by $K[G]$ the class containing a distribution function G.

In the theory of probability it occurs very often, that for a given sequence of p-dimensional random variables $\{S_n\}$, there exist a sequence of positive numbers $\{a_n\}$ and a sequence of vectors $\{b_n\}$ such that the

* Most of the results in the Part I of this paper have been given in the writer's previous papers: On the convergence of classes of distributions, Ann. Inst. Statist. Math., Tokyo, 3, 7-15 (1951); A metrization of class-convergence of distributions, loc. cit., 5, 1-7 (1953); On the many-dimensional distribution functions, loc. cit., 5, 41-58 (1953).

sequence of the distributions of $S_n/a_n - b_n$ converges to some distribution. Let F_n be the distribution function of S_n and let G be the distribution function of the limiting distribution. Then the distribution function of $S_n/a_n - b_n$ is given by $F_n(a_n x + a_n b_n)$ and we have

$$(1) \qquad \lim_{n\to\infty} F_n(a_n x + a_n b_n) = G(x),$$

at every continuity point of $G(x)$. Under these circumstances, for any positive number a and any p-dimensional vector b it holds that

$$\lim_{n\to\infty} F_n(a_n a x + a_n a b + a_n b_n) = G(ax + ab)$$

at every point of continuity of $G(ax+ab)$. Thus, in limit problems of probability distributions, limit classes rather than limit distributions appear. When (1) holds the sequence of classes $\{K[F_n]\}$ will be said to converge to the class $K[G]$, $a_1, a_2 \ldots$ will be called *scaling constants,** and $b_1, b_2, \ldots$ will be called *centering constants* or *centering vectors*. A limit problem of probability distributions is always a limit problem of classes.

In limit problems of probability distributions we have interests in (i) the uniqueness of the limit class of a convergent sequence of classes, (ii) relations between different paired sequences of scaling constants and centering constants, and (iii) a general method to determine sequences of scaling constants and centering constants. (i) was first proved by A. Khintchine [14] in the one-dimensional case, (ii) is known in the one-dimentional case, and (iii) has been treated under more or less restrictive conditions. In Part I of this paper a simple natural proof for the Khintchine's uniqueness theorem is given by making use of inverses of distribution functions, a unified treatment of (ii) is given, (iii) is researched with no restrictive conditions, and furthermore the extensions of these results to the multi-dimensional case are also presented.

In the study of limit distributions of sums of independent random variables, it is natural to put the condition of asymptotic uniform negligibility, and under this condition the limit distributions are proved to be infinitely divisible. Hence it is useful to investigate the conditions for the sequence of distributions of normalized sums of independent random variables to converge to a given infinitely divisible distribution. Although considerable attensions have been paid to this problem in the one-dimensional case, there are few known in the multi-dimensional case

* $a_1, a_2, \ldots$ are called normalizing factors (Normierungsfaktoren) by W. Feller [6].

except for the P. Lévy's canonical form of the infinitely divisible distribution. According to P. Lévy [16], §62, the logarithm of the characteristic function of a infinitely divisible p-dimensional probability distribution is given by

$$(2)\qquad \psi(t)=ia't-\frac{1}{2}t'\sigma t+\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)d\nu$$

where a is a p-dimensional vector, σ is a non-negative definite matrix of pth order, and ν is a p-dimensional measure with

$$\int_{|x|<1}|x|^2\,d\nu<\infty,\qquad \int_{|x|\geq 1}d\nu<\infty$$

In (2), t, a, and x denote column vectors, t', a', and x' their transposes, i.e.,

$$t=\begin{pmatrix}t_1\\ \vdots\\ t_p\end{pmatrix},\qquad a=\begin{pmatrix}a_1\\ \vdots\\ a_p\end{pmatrix},\qquad x=\begin{pmatrix}x_1\\ \vdots\\ x_p\end{pmatrix},$$

$$t'=(t_1,\dots,t_p),\qquad a'=(a_1,\dots,a_p),\qquad x'=(x_1,\dots,x_p),$$

the usual matrix notation is used, and $|x|$ denotes Euclidean length of x. This canonical form was found from the point of view of the theory of the additive process. Part II of this paper gives an analytical derivation of (2) and a necessary and sufficient condition for the sequence of distributions of sums of asymptoticaly uniformly negligible independent p-dimensional random variables to converge to a given infinitely divisible distribution. Our method follows M. Loéve [17]* in the one-dimensional case.

Following A. Khintchine [13] let us put

$$\mu(E)=\int_E\frac{x'x}{1+x'x}\,d\nu,$$

then (2) is rewritten as

$$(3)\qquad \psi(t)=ia't-\frac{1}{2}t'\sigma t+\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu,$$

with

$$\mu(R_p)<\infty,\qquad \mu(\{0\})=0.$$

* Japanese readers may refer to the appendix of Y. Kawada [12], an exposition of M. Loéve [17] by the present writer.

Analytically (3) is preferable to (2) as μ has finite total measure while ν may have infinite total measure. The point, in which the multi-dimensional case differs from the one-dimensional case, is that the integrand in the right side of (3) has no determined limit as x tends to 0.

Part III of this paper investigates the various versions of the multi-dimensional central limit theorem, which states that the sum of asymptotically uniformly negligible independent random variables, under appropriate restrictions, is nearly normally distributed. The general convergence theorem in Part II and the knowledge on scaling constants and centering constants in Part I are applied to this problem. The multi-dimensional central limit theorem has been treated by H. Gramér [1], C. G. Essen [5], and W. Hoeffding & H. Robins [10] etc., but the literature which deals with the complete generalization of the well-known versions in the one-dimensional case seems to be scanty.* On various versions of the central limit theorem in the one-dimensional case, readers may refer to W. Feller [8] and M. Loéve [18].

Acknowledgement

The writer wish to express his thanks to Professor Tatsuo Kawata and Professor Kiyonori Kunisawa for their interest and encouragement which greately aided the work. Thanks are also due to Professor Kameo Matusita for his invaluable help.

CONTENTS

Summary
Introduction
Acknowledgement

Part I Class convergences of distributions

1. Distribution functions and their classes
2. Inverse functions of one-variate distribution functions
3. Uniqueness theorem for class convergences: the one-dimensional case
4. Scaling and centering constants: the one-dimensional case
5. Uniqueness theorem, and scaling and centering constants: the multi-dimensional case

Part II Infinitely divisible distributions

6. Preliminaries
7. Continuous amplitudes of non-vanishing characteristic functions

* After I wrote this paper I came to know a paper of N. A. Sapogov [19] in the Mathematical Reviews, vol. 12. But yet I cannot see it.

8. Infinitely divisible distributions
9. Infinitely divisible distributions in the generalized sense
10. Convergence theorem

Part III Central limit theorem

11. General case
12. Lindeberg's and Liapounov's conditions
13. Generalization of P. Lévy's theorem
14. Feller's criterion and the attraction domains of the normal distributions
15. Reduction to the one-dimensional case

Part I Class convergences of distributions

1. Distribution functions and their classes

A measurable function X, defined on a probability space Ω, and taking values in R_p, p-dimensional Euclidean space, will be called a p-dimensional random variable ($p=1, 2, \dots$). Let us write $X=(X_1, X_2, \dots X_p)$ where X_j is the jth component of X, $j=1, 2, \dots, p$. Then

(1.1) $$F(x)=F(x_1, x_2, \dots, x_p)=\Pr\{X_j(\omega)\leq x_j, j=1, 2, \dots, p\}$$

is defined for all $x=(x_1, x_2, \dots, x_p)$ in R_p, where $\Pr\{\cdots\}$ means the probability of the ω set defined by the condition $\{\cdots\}$. The function F defined by (1.1) is called the distribution function of the p-dimensional random variable $X=(X_1, X_2, \dots, X_p)$.

In case $p=1$, F is monotone non-decreasing, continuous to the right and

$$\lim_{x\to-\infty} F(x)=0, \qquad \lim_{x\to\infty} F(x)=1.$$

Any function F satisfying all these conditions will be called a one-variate distribution function.

In case $p>1$, the function F defined by (1.1) is monotone non-decreasing and continuous to the right in each variable and

$$\lim_{x_j\to-\infty} F(x_1, \dots, x_p)=0, \qquad j=1, 2, \dots, p,$$

$$\lim_{x_1, \dots, x_p\to\infty} F(x_1, \dots, x_p)=1,$$

$$[F]_x^y=[F]_{(x_1, \dots, x_p)}^{(y_1, \dots, y_p)}\geq 0, \quad \text{for} \quad x_j\leq y_j, \quad j=1, 2, \dots, p.$$

Here $[F]_x^y$ is defined by

(1.2) $$[F]_x^y=\sum_{i=1}^{2^p}(-1)^{n(v_i)} F(v_i),$$

where $v_i=(v_{i1}, v_{i2},\dots, v_{ip})$ $(i=1, 2,\dots, 2^p)$, each v_{ij} is either x_j or y_j, and $n(v_i)$ denotes the number of lower simbols x_j among the co-ordinates of v_i. $[F]_x^y$ is the pth difference of F and the evaluation in terms of F of

$$\Pr\{x_j<X_j(\omega)\leq y_j, \qquad j=1, 2,\dots, p\}.$$

Any function F satisfying all these conditions will be called a p-variate distribution function.

A probability measure defined on B_p, the family of all p-dimensional Borel sets, will be called a distribution in R_p or a p-dimensional distribution.

If X is a p-dimensional random variable, the p-dimensional distribution P defined by

$$P(A)=\Pr\{X(\omega)\in A\}, \qquad A\in B_p,$$

is called the distribution of the p-dimensional random variable X.

A p-variate distribution function F defines a p-dimensional distribution P

$$P(A)=\int\cdots\int_A dF(x_1,\dots, x_p)$$

and conversely P defines F

$$F(x)=P(\{y;\ y_j\leq x_j, \qquad j=1, 2,\dots, p\})$$

where $x=(x_1,\dots, x_p)$, $y=(y_1,\dots, y_p)$ and $\{y:C\}$ denotes the set of y satisfying the condition C.

Let $P_1, P_2,\dots$ be p-dimensional measures with distribution functions $F_1, F_2,\dots$, and let P be another p-dimensional measure with distribution function F. Then $\lim_{n\to\infty} P_n(A)=P(A)$ for every set of continuity of P, if and only if $\lim_{n\to\infty} F_n(x)=F(x)$ at every point of continuity of F. When these equivalent conditions hold the sequence $\{P_n\}$ is said to converge to P, $\{F_n\}$ is said to converge to F, and these are written as $\lim P_n=P$ and $\lim F_n=F$, respectively. The definition of distance between two p-variate distribution functions G_1 and G_2 that matches this convergence is the following

(1.3) $$d(G_1, G_2)=\min\{\varepsilon;\ G_1(x-\varepsilon e)-\varepsilon\leq G_2(x)\leq G_1(x+\varepsilon e)+\varepsilon, \text{ for all } x\in R_p\}$$

where $e=(1, 1,\dots, 1)\in R_p$ and $\min\{\varepsilon;\ C\}$ denotes the least ε belonging to the set $\{\varepsilon;\ C\}$. Under this definition the space of all p-dimensional distribution functions is a complete metric space, and $\lim_{n\to\infty} F_n(x)=F(x)$ at all points of continuity of F if and only if $\lim_{n\to\infty} d(F_n, F)=0$ (see [20]).

If a sequence of distribution functions $\{F_n\}$ converges to a distribution function F, then we have $\lim_{n\to\infty} F_n(x-0)=F(x)$ at every point of continuity of F. We shall prove this in the one-dimensional case because of the simplicity of writing. Let x_0 be a point of continuity of F, then for any given $\varepsilon>0$, there exists a $\delta>0$ such that

$$F(x_0)-\varepsilon\leq F(x_0-\delta-0),\qquad F(x_0+\delta-0)\leq F(x_0)+\varepsilon. \tag{1.4}$$

For this δ there exists an integer N such that $d(F, F_n)\leq\delta$ for all $n\geq N$, that is,

$$F(x-\delta)-\delta\leq F_n(x)\leq F(x+\delta)+\delta,\qquad -\infty<x<\infty,\qquad n\geq N,$$

hence

$$F(x-\delta-0)-\delta\leq F_n(x-0)\leq F(x+\delta-0)+\delta,\quad -\infty<x<\infty,\quad n\geq N. \tag{1.5}$$

From (1.4) and (1.5) it follows that

$$F(x_0)-\varepsilon-\delta\leq F_n(x_0-0)\leq F(x_0)+\varepsilon+\delta,\qquad n\geq N.$$

Since $\varepsilon+\delta$ can be chosen arbitrarily small, we have $\lim_{n\to\infty} F_n(x_0-0)=F(x_0)$ which completes the proof.

Let F and G be two p-dimensional distribution functions. If there exist a positive number a and a p-dimensional vector b such that

$$F(ax+b)=G(x) \text{ for all } x\in R_p,$$

then we write $F\sim G$. This relation $\sim$ satisfies the equivalence relation: $F\sim F$; if $F\sim G$ then $G\sim F$; if $F\sim G$ and $G\sim H$ then $F\sim H$. Therefore all p-variate distribution functions are classified by letting F and G belong to the same *class* if and only if $F\sim G$. Throughout this paper *classes of p-variate distribution functions* shall be interpreted in this meaning. Two distributions in R_p are said to belong to the same class if and only if the corresponding distribution functions do so. A class of distribution functions and the corresponding class of distributions will be identified if no confusion occurs.

A sequence of classes $\{K_n\}$ will be called to converge to a class K, if a sequence of distribution functions $\{F_n\}$, each F_n being chosen adequately from K_n for each n, converges to some distribution function F belonging to K. A p-dimensional distribution will be called *unit distribution* if the whole probability 1 is placed in a fixed point in R_p, otherwise *non-unit distribution*. When $p=1$, 'unit' or 'non-unit' may be replaced by 'improper' or 'proper'*. A distribution function will be called *unit* or *non-unit*, according as the corresponding distribution is unit or non-unit. All unit distribution functions form a class which will be called *unit class*. Other classes will be called *non-unit classes*.

Throughout this paper, whenever more than one random variable is involved in a discussion, it will always be assumed, unless the contrary is explicitly stated, that the random variables are all defined on the same probability space Ω.

2. Inverse functions of one-variate distribution functions

Throughout sections 2–4 it will always be assumed that any distribution function is one-variate.

Let F be a distribution function and define f by

$$f(y)=\max\{x;\ F(x-0)\leq y\},\qquad 0<y<1. \tag{2.1}$$

Then f is a finite-valued function defined on the open interval $(0, 1)$, $f(y)$ is non-decreasing with y, and f is continuous to the right: $f(y)=f(y+0)$, $0<y<1$.

To see the last equality hold it is sufficient to show that $f(y)\geq f(y+0)$. Now for any positive number ε such that $y<y+\varepsilon<1$, we have, by the definition of $f(y+\varepsilon)$, $F(f(y+\varepsilon)-0)\leq y+\varepsilon$, hence, $F(f(y+0)-0)\leq y+\varepsilon$. Letting $\varepsilon\downarrow 0$ we have $F(f(y+0)-0)\leq y$, hence, $f(y+0)\leq f(y)$.

The function f defined by (2.1) is called the *inverse function of the distribution function* F.

LEMMA 2.1

(2.2) $x\leq f(y)$ *if and only if* $F(x-0)\leq y$,

(2.3) $x\geq f(y-0)$ *if and only if* $F(x)\geq y$,

(2.4) $x<f(y-0)$ *if and only if* $F(x)<y$,

(2.5) $x>f(y)$ *if and only if* $F(x-0)>y$,

(2.6) $f(y-0)\leq x\leq f(y+0)$ *if and only if* $F(x-0)\leq y\leq F(x+0)$.

Notice that both F *and* f *are continuous to the right.*

* In the previous paper [20], I used also in the multi-dimensional case the term 'improper' or 'proper' in the same meaning as 'unit' or 'non-unit', defined above. But it seems better not to do so, for usualy 'improper' or 'proper' is used in the same sense as 'singular' or 'non-singular', that is, a multi-dimensional distribution is called improper or proper according as there exists, or does not, a hyperplane in which the whole probability 1 is placed.

PROOF: (2.2) and (2.5) are immediate consequences of the definition (2.1). To prove (2.3), assume that $x \geq f(y-0)$. Then for any $\varepsilon > 0$ we have $x+\varepsilon > f(y-\varepsilon)$, hence, $F(x+\varepsilon-0) > y-\varepsilon$ by (2.5). Let $\varepsilon \downarrow 0$, then $F(x) \geq y$. Conversely $F(x) \geq y$ implies $x \geq f(y-0)$. Thus (2.3) is proved. (2.4) is equivalent to (2.3). (2.6) follows from (2.2) and (2.3).

From (2.3) and (2.4) we have the following

THEOREM 2.1 *A distribution function F is uniquely determined by its inverse function f. More explicitly, it holds that*

(2.7) $$F(x)=\max\{y;\ f(y-0) \leq x\}, \qquad -\infty < x < \infty.$$

We notice the following facts.

LEMMA 2.2

(2.8) $$f(y-0)=\min\{x;\ F(x+0) \geq y\}, \qquad 0<y<1.$$

(2.9) $$F(x-0)=\min\{y;\ f(y+0) \geq x\}, \qquad -\infty < x < \infty.$$

(2.8) follows from (2.3) and (2.4); (2.9) follows from (2.2) and (2.5).

LEMMA 2.3

(2.10) *If* $y < F(x)$, *then* $f(y) \leq x$.

(2.11) *If* $x < f(y)$, *then* $F(x) \leq y$.

(2.12) *If* $y > F(x-0)$, *then* $f(y-0) \geq x$.

(2.13) *If* $x > f(y-0)$, *then* $F(x-0) \geq y$.

To prove (2.10), assume that $y<F(x)$. Then $y<F(x+\varepsilon-0)$ for any $\varepsilon>0$, hence $f(y)<x+\varepsilon$, which implies $f(y) \leq x$. The others will be proved similarly.

Now, we have the following theorems.

THEOREM 2.2 *Let f be the inverse of a distribution function F. Then the inverse of $F(a(\cdot+b))$, where a is positive and b is real, is equal to $f/a-b$.*

PROOF: Put $F_1=F(a(\cdot+b))$ and fix y, $0<y<1$. Then

$$F_1(f(y)/a-b-0)=F(f(y)-0) \leq y.$$

Hence

(2.14) $$f_1(y) \geq f(y)/a-b,$$

where f_1 denotes the inverse of F_1. Similarly we have

$$f(y) \geq af_1(y)+ab$$

from which it follows that

(2.15) $$f_1(y) \leq f(y)/a-b.$$

(2.14) and (2.15) imply

$$f_1(y)=f(y)/a-b.$$

THEOREM 2.3 *Let* $f_n(n=1, 2,\dots)$ *and* f *be the inverses of distribution functions* F_n *and* F. *Then* $\lim_{n\to\infty} f_n(y)=f(y)$ *at every continuity point of* f *if and only if* $\lim_{n\to\infty} F_n(x)=F(x)$ *at every continuity point* x *of* F.

The 'if' part is stated in Y. Kawada [12], p. 182, without proof, and its special case is proved in P. Lévy [16], § 43.

PROOF: To prove the 'if' part, suppose that $\lim F_n(x)=F(x)$ at every continuity point x of F and let y_0 be a continuity point of f. Write $x_0=f(y_0)$. As f is continuous at point y_0, for any given $\varepsilon>0$, we have

$$x_0-\varepsilon<f(y_0-0), \qquad f(y_0+0)<x_0+\varepsilon.$$

From Lemma 2.1 it follows that

$$F(x_0-\varepsilon)<y_0<F(x_0+\varepsilon). \tag{2.16}$$

If $x_0\pm\varepsilon$ are both continuity points of F, we have from the assumption

$$\begin{aligned}&\lim_{n\to\infty} F_n(x_0+\varepsilon-0)=F(x_0+\varepsilon),\\ &\lim_{n\to\infty} F_n(x_0-\varepsilon)=F(x_0-\varepsilon).\end{aligned} \tag{2.17}$$

From (2.16) and (2.17) it holds that for sufficiently large N,

$$F_n(x_0-\varepsilon)<y_0<F_n(x_0+\varepsilon-0), \qquad n\geq N,$$

which implies, from Lemma 2.1, that

$$x_0-\varepsilon<f_n(y_0)<x_0+\varepsilon, \qquad n\geq N.$$

Since ε can be chosen arbitrarily small, we have

$$\lim_{n\to\infty} f_n(y_0)=x_0=f(y_0),$$

which completes the proof of the 'if' part. The 'only if' part is proved in the same way.

THEOREM 2.4 *If a distribution function* F *is strictly increasing in an open interval* $a<x<b$, *then its inverse* f *is continuous in the open interval* $F(a-0)<y<F(b+0)$.

PROOF: Fix y_0 such that $F(a-0)<y_0<F(b+0)$. We shall prove that $f(y_0)=f(y_0-0)$. Case (1): when $F(a+0)<y_0<F(b-0)$. Write $x_1=f(y_0-0)$ and $x_2=f(y_0+0)$. Then from Lemma 2.1 we have

$$F(x_1+0)\geq y_0\geq F(x_2-0), \qquad a<x_1\leq x_2<b. \tag{2.18}$$

If $x_1<x_2$, by the hypothesis $F(x_1+0)<F(x_2-0)$ which contradicts with (2.18). Therefore $x_1=x_2$ or $f(y_0-0)=f(y_0)$. Case (2): when $F(a-0)<y_0\leq$

$F(a+0)$. Since for any $\varepsilon<0$, $F(a-0)<y_0<F(a+\varepsilon-0)$, we have $f(y_0)=a$. From $F(a-0)<y_0$ and (2.12) we have $f(y_0-0)\geq a$. Hence $f(y_0-0)=f(y_0)$. Case (3): when $F(b-0)\leq y_0<F(b+0)$. The proof runs in the same way as in Case (2).

As an immediate concequence of the theorem we have the following

COROLLARY *If a distribution function F is strictly increasing in the interval $I=\{x;\ 0<F(x)<1\}$ then its inverse f is continuous in* (0, 1).

3. Uniqueness theorem for class convergences: the one-dimensional case

Throughout this and the next sections we shall use the following notations: F and G with or without a subscript denote distribution functions; f and g denote their inverses; if F and G possess subscripts, the same subscripts will be used for their inverses f and g; U denotes the distribution function of the unit distribution which places the whole probability 1 in the origin; a with or without a subscript denotes a positive number; b with or without a subscript denotes a real number. If not otherwise stated, limits will be considered for $n\to\infty$.

First we shall note the special role of the unit class.

THEOREM 3.1 *Any sequence of classes converges to the unit class. More explicitly, for any sequence of distribution functions $\{F_n\}$ there exists a sequence $\{a_n\}$ such that* $\lim F_n(a_n\cdot)=U$.

This was first proved by A. Khintchine [14] and the proof is easy.

Now the A. Khintchine [14]'s uniqueness theorem for class convergences can be stated as follows.

THEOREM 3.2 *Assume that*

$$\lim F_n=F, \qquad \lim F_n(a_n(\cdot+b_n))=G,$$

and that both F and G are non-unit. Then the limits

$$\lim a_n=a>0, \qquad \lim b_n=b$$

exist and it holds that $G=F(a(\cdot+b))$. Hence, F and G belong to the same class.

PROOF: Let f_n, f, g be the inverses of F_n, F, G, respectively. By Theorem 2.2 the inverse of $F_n(a_n(\cdot+b_n))$ is given by f_n/a_n-b_n. From the hypothesis and Theorem 2.3 it follows that

$$\lim f_n(y)=f(y), \tag{3.1}$$

$$\lim f_n(y)/a_n-b_n=g(y) \tag{3.2}$$

at every continuity point of f and g, respectively. By the assumption,

F and G are non-unit, therefore neither f nor g is constant. Hence, we can choose y_1 and y_2 such that $1>y_1>y_2>0$, $f(y_1)>f(y_2)$, $g(y_1)>g(y_2)$, and that both y_1 and y_2 are common continuity points of f and g, so that both (3.1) and (3.2) hold for both $y=y_1$ and $y=y_2$. By making differences we have

$$\lim \{f_n(y_1)-f_n(y_2)\}=f(y_1)-f(y_2)>0,$$

$$\lim \{f_n(y_1)-f_n(y_2)\}/a_n=g(y_1)-g(y_2)>0.$$

By taking the ratio we have

$$\lim a_n=\frac{f(y_1)-f(y_2)}{g(y_1)-g(y_2)}=a(\text{say})>0. \tag{3.3}$$

Form (3.1)/(3.3)−(3.2), then

$$\lim b_n=f(y)/a-g(y)=b(\text{say}).$$

This holds for every common continuity point y of f and g. Since f and g are both continuous to the right, we have $f(y)/a-g(y)=b$ for all y, $0<y<1$. Therefore we have $g=f/a-b$, and this together with Theorem 2.1 and Theorem 2.2 implies $G=F(a(\cdot+b))$.

THEOREM 3.3 *Assume that* $\lim F_n=F$. (*It makes no difference whether* F *is non-unit or unit*). *Then*

(i) *if* $\lim a_n=a>0$, $\quad \lim F_n(a_n\cdot)=F(a\cdot)$;

(ii) *if* $\lim a_n=+\infty$, $\quad \lim F_n(a_n\cdot)=U$;

(iii) *if* $\lim b_n=b$, $\quad \lim F_n(\cdot+b_n)=F(\cdot+b)$.

This is known (for instance, see H. Cramér [2], p. 254), and is easily proved, for instance, by making use of the inverses of distribution functions.

According to Theorems 3.1 and 3.2, if a sequence of distribution functions $\{F_n\}$ converges to a non-unit distribution function F, and if for some sequences $\{a_n\}$ and $\{b_n\}$ the sequence $\{F_n(a_n(\cdot+b_n))\}$ converges to a distribution function, then the limit distribution function must be $F(a(\cdot+b))$ or $U(\cdot+b)$ for some a and b. With respect to these circumstances we have the following two theorems.

THEOREM 3.4 *Assume that* $\lim F_n=F$ *and that* F *is non-unit. Then*

$$\lim F_n(a_n(\cdot+b_n))=F(a(\cdot+b)), \tag{3.4}$$

if and only if

$$\lim a_n=a, \qquad \lim b_n=b. \tag{3.5}$$

This was first proved by B. Gnedenko [9], §3.

PROOF: Since the 'if' part is an immediate consequence of Theorem 3.3, it is sufficient to prove the 'only if' part. Assume (3.4). According to Theorem 3.2 the limits

$$\lim a_n = a' > 0, \qquad \lim b_n = b'$$

exist and we have

$$F(a(\cdot + b)) = F(a'(\cdot + b'))$$

from which it follows that

$$f(y)/a - b = f(y)/a' - b', \qquad 0 < y < 1.$$

Since F is non-unit, f can take at least two different values, therefore, it must hold that $a' = a$ and $b' = b$.

THEOREM 3.5 *Assume that* $\lim F_n = F$ *and that* F *is non-unit. Then*

$$\lim F_n(a_n(\cdot + b_n)) = U(\cdot + b)$$

if and only if

$$\lim a_n = +\infty, \qquad \lim b_n = b.$$

The 'if' part follows from Theorem 3.3; the 'only if' part is proved by taking the inverses of distribution functions.

The normal distribution with mean m and variance v is denoted by $N(m, v)$. It is convenient to denote by $N(m, 0)$ the unit distribution which has the whole probability 1 placed in the point m. As an application of Theorem 3.2 the following fact is proved.

If a sequence of normal distributions $N(m_n, v_n)$ *tends to a distribution* L, *then the limits*

$$\lim m_n = m, \qquad \lim v_n = v \tag{3.6}$$

exist and

$$L = N(m, v) \tag{3.7}$$

(K. Ito [11], p. 187)

PROOF: Let G denote the distribution function of the normal distribution $N(0, 1)$, then the distribution function of $N(m_n, v_n)$ is given by $G((\cdot - m_n)/\sqrt{v_n})$. Let F denote the distribution function of the limit distribution L. Moreover, let us write $G_n = G$, $n = 1, 2, \ldots$. Then we have

$$\lim G_n = G, \qquad \lim G_n((\cdot - m_n)/\sqrt{v_n}) = F.$$

If F is non-unit, from Theorem 3.2 the limits (3.6) exist and we have $F = G((\cdot - m)/\sqrt{v})$ from which (3.7) follows. If F is unit, by Theorem 3.5 there exist the limits (3.6) with $v = 0$ and we have $F = U(\cdot - m)$, from which (3.7) follows.

4. Scaling and centering constants: the one-dimensional case

Now we want to determine the sequences of scaling and centering constants. For this purpose, we shall discuss about dispersions and centres of distributions which will play important roles as scaling and centering constants, respectively.

We shall begin with the dispersions of distributions. Let F be a distribution function and let φ be its characteristic function. Then the *mean concentration function* Ψ_F, introduced by K. Kunisawa [15], of F is defined by

$$\Psi_F(l)=l\int_0^\infty e^{-lt}|\varphi(t)|^2\,dt, \qquad 0<l<\infty. \tag{4.1}$$

It is easily shown that

$$\Psi_F(l)=\int_{-\infty}^{\infty}\frac{l^2}{l^2+x^2}d\tilde{F}(x), \qquad 0<l<\infty, \tag{4.2}$$

where $\tilde{F}=F*\{1-F(-\cdot)\}$ is the symmetrization of F. From (4.2) it is seen that Ψ_F is a non-decreasing and continuous function defined on the open interval $(0, \infty)$ and we have

$$\begin{aligned} &0<\Psi_F(l)\leq 1, \qquad \Psi_F(\infty)=1, \\ &\Psi_F(+0)=\tilde{F}(+0)-\tilde{F}(-0)=\sum_\nu p_\nu^2=\Sigma_F \text{ (say)}, \end{aligned} \tag{4.3}$$

where the p_ν's are the jumps of F at all its points of discontinuity. Obviously

$$\begin{aligned} &\Sigma_F<1, \qquad \text{if } F \text{ is non-unit;} \\ &\Sigma_F=1, \qquad \text{if } F \text{ is unit.} \end{aligned} \tag{4.4}$$

Now let us put

$$\begin{aligned} &\Psi_F(0)=\Psi_F(+0), \\ &\Psi_F(l)=0, \quad \text{for } l<0. \end{aligned} \tag{4.5}$$

Then Ψ_F is a distribution function. The inverse function D_F of Ψ_F is called the *dispersion function* of F. The value of D_F at a point α will be called the *α-dispersion* of F.

We shall need the following properties of the dispersions.

LEMMA 4.1 *For any distribution function F, its dispersion function D_F is non-negative and continuous. If F is unit,*

$$D_F(\alpha)=0, \qquad 0<\alpha<1. \tag{4.6}$$

If F is non-unit,

(4.7) $$D_F(\alpha) > 0 \quad \textit{if and only if} \quad \alpha > \Sigma_F,$$

where Σ_F *is defined by* (4.3). *Moreover* Σ_F *is invariant when* F *runs over the same class.*

PROOF: If F is unit $\Psi_F(l)=1,\ l\geq 0,\ =0,\ l<0$, from which (4.6) follows. Next, assume that F is non-unit. In this case Ψ_F is strictly increasing in the open interval $(0, \infty)$, therefore, D_F must be continuous in $(0, 1)$ by Corollary to Theorem 2.4. (4.7) follows from (2.4). The last part of the lemma is clear.

LEMMA 4.2 $\quad D_{F(a\cdot+b)}(\alpha)=a^{-1}D_F(\alpha).$

LEMMA 4.3 $\quad D_{F_1*F_2}(\alpha)\geq D_{F_j}(\alpha),\quad j=1,\ 2,$

where $F_1 * F_2$ *denotes the convolution of* F_1 *and* F_2.

LEMMA 4.4 *If* $\lim F_n=F$,

(4.8) $$\lim D_{F_n}(\alpha)=D_F(\alpha), \qquad 0<\alpha<1.$$

These properties of the dispersions are deduced from the following corresponding properties of the mean concentration functions.

$$\Psi_{F(a\cdot+b)}(l)=\Psi_F(al),$$
$$\Psi_{F_1*F_2}(l)\leq \Psi_{F_j}(l), \qquad j=1,\ 2,$$
$$\lim \Psi_{F_n}(l)=\Psi_F(l), \qquad 0<l<1.$$

Furthermore, we can prove that *if* F_2 *is non-unit then*

$$\Psi_{F_1*F_2}(l)<\Psi_{F_1}(l), \qquad \textit{for all} \quad 0<l<\infty,$$

and

$$D_{F_1*F_2}(\alpha)>D_{F_1}(\alpha), \qquad \textit{for all} \quad \alpha>\Sigma_{F_1*F_2}.$$

Note that (4.8) holds for every point in the interval $0<\alpha<1$, as D_F has no point of discontinuity. This is the reason why we use the Kunisawa's dispersions instead of the P. Lévy's dispersions, inverses of maximal concentration functions.

We must now turn to centres of distributions. For any distribution function F, the real number c defined by

(4.9) $$\int_{-\infty}^{\infty} \operatorname{arc\,tan}(x-c)\, dF(x)=0$$

will be called the *centre* of F and will be denoted by $c=c(F)$. Any distribution function F with $c(F)=0$ is called to be *centered*. As is easily proved, centres have the following properties.

LEMMA 4.5 $\quad c(F(\cdot+b))=c(F)-b.$

LEMMA 4.6 *If* $\lim F_n=F$, $\quad \lim c(F_n)=c(F).$

(See J. L. Doob [4], p. 408).

As it was noted in Lemma 4.1 the Σ_F defined by (4.3) is determined by the class K containing F, so that it can be denoted by Σ_K. Clearly $\Sigma_K=1$, or <1 according as K is unit or non-unit.

In the sequel we shall denote by α a constant such that $0<\alpha<1$. Let F be a distribution function with $\Sigma_F<\alpha$. Then $D_F(\alpha)>0$ by Lemma 4.1. Put $D=D_F(\alpha)$, and $c=c(F(D\cdot))$. Then the $F(D(\cdot+c))$ is a centered distribution function with α-dispersion 1. Let $F(a(\cdot+b))$ be another centered distribution function with α-dispersion 1. Then we have $D/a=1$, $c-b=0$, so that $F(a(\cdot+b))=F(D(\cdot+c))$. Thus we have

LEMMA 4.7 *Assume that K is a class with $\Sigma_K<\alpha$. Then a centered distribution function, belonging to K, with α-dispersion 1 exists and is uniquely determined.*

Now we can determine scaling constants and centering constants in limit problems of distributions.

THEOREM 4.1 *Let K_n's $(n=0, 1, 2,\dots)$ be classes with $\Sigma_{K_n}<\alpha$. For each n let F_n be the centered distribution function, belonging to K_n, with α-dispersion 1. Then* $\lim K_n=K_0$ *if and only if* $\lim F_n=F_0$.

PROOF: It is sufficient to prove the 'only if' part. Assume that $\lim K_n=K_0$. Then there exist G_n's such that $G_n \in K_n$ $(n=0, 1, 2,\dots)$ and $\lim G_n=G_0$. Put $D_n=D_{G_n}(\alpha)$. Then $D_n>0$ $(n=0, 1, 2,\dots)$ and $\lim D_n=D_0$ by Lemma 4.4. According to Theorem 3.3 we have $\lim G_n(D_n\cdot)=G_0(D_0\cdot)$. Put $c_n=c(G_n(D_n\cdot))$. Then $\lim c_n=c_0$ by Lemma 4.6. Hence $\lim G_n(D_n(\cdot+c_n))=G_0(D_0(\cdot+c_0))$ by Theorem 3.3. By Lemma 4.7 we have $G_n(D_n(\cdot+c_n))=F_n$. Therefore $\lim F_n=F_0$.

We shall mean by the *α-dispersion* and the *centre* of a random variable X, the α-dispersion and the centre of the distribution function of the X, respectively, and we shall denote them by $D_X(\alpha)$ and $c(X)$. The dispersion and the centre of a distribution is defined by the corresponding ones of the distribution function of the distribution.

COROLLARY 1 *Let $\{X_n; n=1, 2, \dots\}$ be a sequence of random variables with positive α-dispersions. Assume that for some sequences $\{a_n\}$ and $\{b_n\}$ the distribution of X_n/a_n-b_n converges to a distribution L with positive α-dispersion. Then the distribution of X_n/D_n-c_n converges to the centered distribution, belonging to the same class with L, with α-dispersion 1, where $D_n=D_{X_n}(\alpha)$ and $c_n=c(X_n/D_n)$.*

COROLLARY 2 *Let $\{X_n; n=1, 2,\dots\}$ be a sequence of random variables* *with positive α-dispersions. If for some sequence $\{a_n\}$ the distribution function of X_n/a_n converges to a distribution function F with positive α-dispersion D, then the distribution functions of X_n/D_n converge to the $F(D\cdot)$.*

Let f be the inverse of a distribution function F. Then any number m satisfying $F(m-0)\leq\frac{1}{2}\leq F(m+0)$, or equivalently $f(\frac{1}{2}-0)\leq m\leq f(\frac{1}{2}+0)$, is called a median of F. Let $\{F_n; n=0,1,2,\dots\}$ be a sequence of distribution functions and let m_n be any median of F_n for each n. If $\lim F_n=F_0$ and if the median of F_0 is uniquely determined, then we have $\lim m_n=m_0$ by Theorem 2.3. Let us note that if a distribution function belonging to a class K has the uniquely determined median, that is, if its inverse is continuous at the point $\frac{1}{2}$, then every distribution function belonging to the class K has this property.

As a result of these accounts we have the following

THEOREM 4.2 *Let the hypotheses of Corollary 1 to Theorem 4.1 hold. Moreover, assume that the median of L is uniquely determined, and let m_n be any median of X_n for each n. Then the distribution of $(X_n-m_n)/D_n$ converges to the distribution, belonging to the same class with L, with median 0 and α-dispersion 1.*

5. Uniqueness theorem and scaling and centering constants: the multi-dimensional case

Now let us generalize the results of the preceding two sections to the multi-dimensional case.

Let F, $F(x)=F(x_1,\dots,x_p)$, be a p-dimensional distribution function and let $F_1, F_2,\dots, F_p$ be its one-dimensional marginal distribution functions defined by

$$F_1(\xi)=\lim_{x_2,\cdots,x_p\to\infty}F(\xi,x_2,\dots,x_p),\quad F_2(\xi)=\lim_{x_1,x_3,\cdots,x_p\to\infty}F(x_1,\xi,x_3,\dots,x_p),\dots$$

$$F_p(\xi)=\lim_{x_1,\cdots,x_{p-1}\to\infty}F(x_1,x_2,\dots,x_{p-1},\xi),\qquad -\infty<\xi<\infty.$$

We shall call the convolution of the marginal distribution functions $F^*=F_1*F_2*\cdots*F_p$ the *trace distribution function* (or briefly *trace*) of the p-dimensional distribution function F. Then we have:

(i) *A p-dimensional distribution function F is non-unit if and only if its trace F^* is non-unit.*

(ii) *If the trace of a p-dimensional distribution function F is F^*, the trace of $F(a(\cdot+b))$ is $F^*(a(\cdot+b_1+\cdots+b_p))$, where $a>0$ and $b=(b_1,\dots,b_p)$.*

(iii) *Let* $F_n(n=0, 1, 2, \ldots)$ *be p-dimensional distribution functions and let* F_n^* *be the corresponding traces. Then* $\lim F_n = F_0$ *implies* $\lim F_n^* = F_0^*$.

The last (iii) follows from the fact that if F_n converges to F_0 any marginal distribution function of F_n converges to the corresponding marginal distribution function of F_0 (see, for instance, [20 , Lemma 3) and from the continuity of the convolution.

By the *dispersion function* of a p-dimensional distribution function F, we mean the dispersion function of the trace F^* of F. We denote by D_F the dispersion function of F as in the one-dimensional case, so that $D_F = D_{F^*}$.

Let F be a p-dimensional distribution function, $F_1, \ldots, F_p$ its one-dimensional marginal distribution functions and $c_1, \ldots, c_p$ the centres of $F_1, \ldots, F_p$, respectively. The vector $c=(c_1, \ldots, c_p)$ will be called the *centre* of F. If the centre of F is 0, F is called to be *centered*. We denote by $c(F)$ the centre of F.

We have the following

LEMMA 5.1 *For any p-dimensional distribution function F, its dispersion function* D_F *is non-negative and continuous. If F is unit*

$$D_F(a)=0, \qquad 0<a<1.$$

If F is non-unit.

$$D_F(a)>0 \quad \textit{if and only if} \quad a>\Sigma_F,$$

where Σ_F *is defined by* $\Sigma_F = \Sigma_{F^*}$, F^* *being the trace of F. Moreover* Σ_F *is invariant when F runs over the same class.*

In the remainder of this section, it will always be assumed, unless the contrary is explicitly stated, that F and G with or without a subscript denote p-dimensional distribution functions; a with or without a subscript denotes a positive number; b with or without a subscript denotes a p-dimensional vector; U denotes the distribution function of the unit distribution which has the whole probability 1 placed in the origin, i.e.,

$$U(x_1, \ldots, x_p) = \begin{cases} 1, & \text{if } x_j \geq 0 \text{ for all } j, \\ 0, & \text{if } x_j < 0 \text{ for some } j. \end{cases}$$

We have the following lemmas, extensions of those in the one-dimensional case.

LEMMA 5.2 $D_{F(a\cdot+b)}(a) = a^{-1} D_F(a)$.

LEMMA 5.3 $D_{F_1 * F_2}(a) \geq D_{F_j}(a), \quad j=1, 2.$

LEMMA 5.4 *If* $\lim F_n=F$, $\lim D_{F_n}(\alpha)=D_F(\alpha)$ *at every point* α *in* $0<\alpha<1$.

LEMMA 5.5 $c(F(\cdot+b))=c(F)-b$.

LEMMA 5.6 *If* $\lim F_n=F$, $\lim c(F_n)=c(F)$.

By making use of these lemmas we can generalize the results of sections 3 and 4 to the multi-dimensional case. Before doing so, we notice the following

THEOREM 5.1 *Any sequence of classes converges to the unit class. More explicitly, for any sequence of distribution functions* $\{F_n\}$ *there exists a sequence* $\{a_n\}$ *such that* $\lim F_n(a_n\cdot)=U$.

This is easily proved (see [20], Theorem 4).

THEOREM 5.2 *Assume that* $\lim F_n=F$. *(It makes no difference whether* F *is non-unit or unit). Then:*

(i) *if* $\lim a_n=a>0$, $\lim F_n(a_n\cdot)=F(a\cdot)$;

(ii) *if* $\lim a_n=+\infty$, $\lim F_n(a_n\cdot)=U$;

(iii) *if* $\lim b_n=b$, $\lim F_n(\cdot+b_n)=F(\cdot+b)$.

PROOF: To prove (i) and (iii), we shall prove that if $\lim a_n=a>0$ and $\lim b_n=b$ then $\lim F_n(a_n\cdot+b_n)=F(a\cdot+b)$. Assume that $\lim a_n=a>0$ and $\lim b_n=b$. Let x be a fixed vector. Then $\lim(a_nx+b_n)=ax+b$. Hence for any positive number ε there exists a number N such that for all $n\geq N$

$$ax+b-\varepsilon e<a_nx+b_n<ax+b+\varepsilon e,$$

where $e=(1,1,\dots,1)\in R_p$ and $(x_1,\dots,x_p)<(y_1,\dots,y_p)$ means that $x_j<y_j$ for all j. Then

$$F_n(ax+b-\varepsilon e)\leq F_n(a_nx+b_n)\leq F_n(ax+b+\varepsilon e),\qquad n\geq N.$$

Therefore, if $ax+b\pm\varepsilon e$ are both continuity points of F, it holds that

$$F(ax+b-\varepsilon e)\leq \liminf F_n(a_nx+b_n)\leq \limsup F_n(a_nx+b_n)\leq F(ax+b+\varepsilon e).$$

As ε can be chosen arbitrarily small, we have

$$\lim F_n(a_nx+b_n)=F(ax+b)$$

if x is a continuity point of $F(a\cdot+b)$.

To prove (ii), assume that $\lim a_n=+\infty$. Fix an $x=(x_1,\dots,x_p)$. Case (1): $x_j>0$ for all $j=1,2,\dots,p$. For any positive number a there exists a number N such that $a_nx>ae$ for all $n\geq N$. Then $F_n(a_nx)\geq F_n(ae)$ for $n\geq N$. If ae is a continuity point of F, it holds that $\liminf_{n\to\infty} F_n(a_nx)\geq F(ae)$. Letting $a\to\infty$ we have $\liminf_{n\to\infty} F_n(a_nx)\geq 1$, from which we have $\lim F_n(a_nx)=1$. Case (2): $x_j<0$ for some j. For any positive number a, there exists an N such that $a_n x_j < -a$ for all $n \geq N$. Denote by F_{nj} and $F_{(j)}$ the marginal distribution functions of F_n and F, with respect to the jth component, respectively. Then we have $F_n(a_n x) \leq F_{nj}(-a)$ for $n \geq N$. If $-a$ is a continuity point of $F_{(j)}$ it holds that $\limsup_{n\to\infty} F_n(a_n x) \leq F_{(j)}(-a)$. Letting $a \to \infty$, we have $\limsup_{n\to\infty} F_n(a_n x) \leq 0$, which implies $\lim F_n(a_n x) = 0$.

THEOREM 5.3 *Assume that*

(5.1) $$\lim F_n = F,$$

(5.2) $$\lim F_n(a_n(\cdot + b_n)) = G$$

and that both F and G are non-unit. Then the limits

$$\lim a_n = a > 0, \qquad \lim b_n = b$$

exist and $G = F(a(\cdot + b))$.

PROOF: Since F and G are non-unit, $\Sigma_F < 1$ and $\Sigma_G < 1$. Take an α such that $\max(\Sigma_F, \Sigma_G) < \alpha < 1$, and put

$$D_n = D_{F_n}(\alpha), \qquad D = D_F(\alpha), \qquad D' = D_G(\alpha).$$

Then $D_{F_n(a_n(\cdot + b_n))}(\alpha) = D_n / a_n$. By Lemma 5.4, (5.1) and (5.2) imply

$$\lim D_n = D, \qquad \lim D_n / a_n = D',$$

from which it follows that

(5.3) $$\lim a_n = D/D' = a \text{ (say)} > 0.$$

From (5.1) and (5.3) it follows that

(5.4) $$\lim F_n(a_n \cdot) = F(a \cdot),$$

by THEOREM 5.2. Put $c(F_n(a_n \cdot)) = c_n$, $c(F(a \cdot)) = c$, and $c(G) = c'$. Then (5.4) and (5.2) imply

$$\lim c_n = c, \qquad \lim (c_n - b_n) = c',$$

by LEMMA 5.5 and Lemma 5.6. From the last two equations we have

(5.5) $$\lim b_n = c - c' = b \text{ (say)}.$$

From (5.4) and (5.5) it holds that

(5.6) $$\lim F_n(a_n(\cdot + b_n)) = F(a(\cdot + b))$$

by THEOREM 5.2. Comparing (5.2) and (5.6) it is seen that

$$G = F(a(\cdot + b)).$$

THEOREM 5.4 *Assume that* $\lim F_n = F$ *and that F is non-unit. Then* $\lim F_n(a_n(\cdot + b_n)) = F(a(\cdot + b))$ *if and only if* $\lim a_n = a$ *and* $\lim b_n = b$.

PROOF: The proof runs in the same way as in Theorem 3.4. It is sufficient to prove that if $F(a(\cdot + b)) = F(a'(\cdot + b'))$ for a non-unit F then

$a=a'$ and $b=b'$. Assume that $F(a(\cdot+b))=F(a'(\cdot+b'))$ and that F is non-unit. Take an α such that $D_F(\alpha)>0$, and put $D=D_F(\alpha)$. Then from $F(a(\cdot+b))=F(a'(\cdot+b'))$ it follows that $D/a=D/a'$ which implies $a=a'$. Put $c=c(F(a\cdot))$. Then $F(a(\cdot+b))=F(a(\cdot+b'))$ implies $c-b=c-b'$, hence $b=b'$.

THEOREM 5.5 *Assume that* $\lim F_n=F$ *and that* F *is non-unit. Then* $\lim F_n(a_n(\cdot+b_n))=U(\cdot+b)$ *if and only if* $\lim a_n=+\infty$ *and* $\lim b_n=b$.

PROOF: The 'if' part follows from Theorem 5.2. To prove the 'only if' part, assume that $F_n(a_n(\cdot+b_n))=U(\cdot+b)$. Take an α such that $D_F(\alpha)>0$ and put $D=D_F(\alpha)$, $D_n=D_{F_n}(\alpha)$. Then we have $\lim D_n=D$ and $\lim D_n/a_n=0$, hence, $\lim a_n=\infty$ and $\lim F_n(a_n\cdot)=U$. Put $c_n=c(F_n(a_n\cdot))$, then we have $\lim c_n=0$ and $\lim(c_n-b_n)=-b$, hence $\lim b_n=b$.

THEOREM 5.6 *Let* $K_n(n=0, 1, 2,\dots)$ *be classes with* $\Sigma_{K_n}<\alpha$. *For each* $n(=0, 1, 2,\dots)$, *let* F_n *be the centered distribution function, belonging to* K_n, *with* α*-dispersion 1. Then* $\lim K_n=K_0$ *if and only if* $\lim F_n=F_0$.

The dispersion and the centre of a p-dimensional random variable or a p-dimensional distribution are defined by the corresponding ones of its distribution function.

COROLLARY 1 *Let* $\{X_n;\, n=1, 2,\dots\}$ *be a sequence of random variables with positive* α*-dispersions. Assume that for some sequences* $\{a_n\}$ *and* $\{b_n\}$ *the distribution of* X_n/a_n-b_n *converges to a distribution L with positive* α*-dispersion. Then the distribution of* X_n/D_n-c_n *converges to the centered distribution, belonging to the same class with L, with* α*-dispersion 1, where* $D_n=D_{X_n}(\alpha)$ *and* $c_n=c(X_n/D_n)$.

COROLLARY 2 *Let* $\{X_n;\, n=1, 2,\dots\}$ *be a sequence of random variables with positive* α*-dispersions. If for some sequence* $\{a_n\}$ *the distribution function of* X_n/a_n *converges to a distribution function F with positive* α*-dispersion D, then the distribution function of* X_n/D_n *converges to the* $F(D\cdot)$.

THEOREM 5.7 *Let the hypotheses of Corollary 1 to Theorem 5.6 hold. Moreover, assume that the median vector of L is uniquely determined, and let* m_n *be any median vector of* X_n *for each n. Then the distribution of* $(X_n-m_n)/D_n$ *converges to the distribution, belonging to the same class with L, with median 0 and* α*-dispersion 1.*

Part II Infinitely divisible distributions

6. Preliminaries

A measure, non-negative completely additive set function, which is defined on all p-dimensional Borel sets will be called a p-dimensional measure. Let $\{\mu_n\}$ be a sequence of p-dimensional measures with $\mu_n(R_p)<\infty$ and let μ be another one with $\mu(R_p)<\infty$. If

$$\lim \mu_n(E)=\mu(E) \tag{6.1}$$

for every set E of continuity of μ, $\{\mu_n\}$ will be said to converge to μ and it is written as $\lim \mu_n=\mu$. Note that $\lim \mu_n=\mu$ implies $\lim \mu_n(R_p)=\mu(R_p)$ for R_p is a set of continuity of any bounded measure.

If μ is a p-dimensional measure with $\mu(R_p)<\infty$, its Fourier-Stieltjes transform is defined by

$$\varphi(t)=\int_{R_p} e^{it'x}\,d\mu. \tag{6.2}$$

The Fourier-Stieltjes transform of a p-dimensional distribution will be called a characteristic function of the distribution.

The following fundamental properties of Fourier-Stieltjes transforms of p-dimensional measures will be used. Let μ with or without a subscript denote a p-dimensional measure with $\mu(R_p)<\infty$ and let φ be its Fourier-Stieltjes transform. If μ possesses a subscript, φ will have the same subscript.

(i) *φ is continuous for all t and*

$$|\varphi(t)|\leq\varphi(0)=\mu(R_p).$$

(ii) *μ is uniquely determined by φ.*

(iii) *If* $\lim \mu_n=\mu$ *and if f is a bounded continuous function defined on R_p, we have*

$$\lim \int_{R_p} f(x)\,d\mu_n=\int_{R_p} f(x)\,d\mu.$$

(iv) $\lim \varphi_n=\varphi$ *if and only if* $\lim \mu_n=\mu$. *And in this case* $\lim \varphi_n(t)=\varphi(t)$ *uniformly in every bounded t set.*

(v) *If* $\lim \varphi_n(t)=k(t)$ *exists for all t and $k(t)$ is continuous at the origin then $\mu_n\to$ some μ and $k=\varphi$.* (For one-dimensional case see, for instance, H. Cramér [1], p. 121, additional note, or M. Loéve [17], section I, Lemma A).

(iv) and (v) are called the P. Lévy's continuity theorem. We shall give a proof of (v) in the end of this section.

(vi) *Let φ_1 and φ_2 be the characteristic functions of distribution functions F_1 and F_2. Then the characteristic function of the convolution F_1*F_2 is given by $\varphi_1\varphi_2$.*

Throughout parts II and III, unless the contrary is explicitly stated, the following notations will be used: F, with or without subscripts, denotes a p-dimensional distribution function; U denotes the distribution function of the unit distribution which places the whole probability 1 in the origin; for a point $x=(x_1,\ldots,x_p)$ in R_p, $|x|$ denotes its Euclidean norm, i.e.,

$$|x|=(x_1^2+x_2^2+\cdots+x_p^2)^{1/2}, \tag{6.3}$$

and $\|x\|$ denotes the greatest of the absolute values of its components, i.e.,

$$\|x\|=\max(|x_1|, |x_2|, \ldots, |x_p|). \tag{6.4}$$

Let $\{F_{nl};\, l=1, 2,\ldots, l_n, n=1, 2, \ldots\}$ be a sequence of distribution functions. If

$$\lim_{n\to\infty} \max_{1\leq l\leq l_n} d(F_{nl}, U)=0, \tag{6.5}$$

where $d(F, U)$ is defined by (1.3), then F_{nl}, $l=1, 2,\ldots, l_n$, will be called to converge to U uniformly in $l(1\leq l\leq l_n)$ as $n\to\infty$. It is easily seen that (6.5) holds if and only if for each $\varepsilon>0$

$$\lim_n \max_l \int_{\|x\|\geq\varepsilon} dF_{nl}(x)=0, \tag{6.6}$$

and this is equivalent to the condition that for each $\varepsilon>0$

$$\lim_n \max_l \int_{|x|\geq\varepsilon} dF_{nl}(x)=0. \tag{6.7}$$

Lastly we shall give a proof of (v) following P. Lévy [16], pp. 49–50, in the one-dimensional case, for the completeness. Let $H(x)=H(x_1,\ldots,x_p)$ be a real-valued function defined on R_p. If $H(x_1,\ldots,x_p)$ is monotone non-decreasing and continuous to the right in each variable, and if $[H]_x^y\geq 0$ (see (1.2)) for every $x=(x_1,\ldots,x_p)$ and every $y=(y_1,\ldots,y_p)$ such that $x_j\leq y_j$, $j=1, 2,\ldots, p$, then H will be called to be *positively monotonic*. We need the following

LEMMA 6.1 *Any sequence of distribution functions $\{F_n\}$ has a subsequence $\{F'_n\}$ which converges to a positively monotonic function $H(x)$ at every continuity point of the latter.*

PROOF: By the well-known diagonal method we can choose a subsequence $\{F'_n\}$ such that for each rational point $r\in R_p$ $\lim F'_n(r)=H_0(r)$ exists. $H_0(r)$ is defined only for rational points r and it is obvious that $0\leq H_0(r)\leq 1$ and $H_0(r)=H_0(r_1, r_2,\ldots, r_p)$ is monotone non-decreasing in each variable. Using this H_0 define a function H by

$$H(x)=\inf_{r>x} H_0(r), \qquad x \in R_p,$$

where $(r_1, \ldots, r_p) > (x_1, \ldots, x_p)$ denotes that $r_j > x_j$ for all j. Then $H(x)=H(x_1, \ldots, x_p)$ is monotone non-decreasing in each variable and $0 \leq H(x) \leq 1$. For any fixed point x and for any positive number ε there exists a rational point s such that $s>x$ and $H_0(s)<H(x)+\varepsilon$, from which it follows that $H(x) \leq H_0(r) < H(x)+\varepsilon$ for any rational point r such that $x<r<s$. Thus we have

$$H(x)=\lim_{r \downarrow x} H_0(r) \tag{6.8}$$

and $H(x)$ is continuous to the right. Let $x=(x_1, \ldots, x_p)$ and $y=(y_1, \ldots, y_p)$ be two points of R_p such that $x_j \leq y_j$ for all j. Choose $\varepsilon=(\varepsilon_1, \ldots, \varepsilon_p)$ and $\delta=(\delta_1, \ldots, \delta_p)$ such that $x+\varepsilon$ and $y+\delta$ are rational points and $0<\varepsilon<\delta$. Then $[F'_n]_{x+\varepsilon}^{y+\delta} \geq 0$. Letting $n \to \infty$, we have $[H_0]_{x+\varepsilon}^{y+\delta} \geq 0$. Further, letting $\delta \downarrow 0$, we have $[H]_x^y \geq 0$ by (6.8). Thus H is positively monotonic. It remains to ascertain that $F'_n(x)$ converges to $H(x)$ at every continuity point x of the latter. To prove this, fix a point x. Choose $\varepsilon=(\varepsilon_1, \ldots, \varepsilon_p)$ and $\delta=(\delta_1, \ldots, \delta_p)$ such that $\varepsilon_j>0$ and $\delta_j>0$ for all j and that both $x-\varepsilon$ and $x+\delta$ are rational points. Since $x-\varepsilon<x<x+\delta$,

$$F'_n(x-\varepsilon) \leq F'_n(x) \leq F'_n(x+\delta).$$

Letting $n \to \infty$, we have

$$H(x-2\varepsilon) \leq \liminf_{n \to \infty} F'_n(x) \leq \limsup_{n \to \infty} F'_n(x) \leq H_0(x+\delta).$$

Letting $\varepsilon \downarrow 0$ and letting $\delta \downarrow 0$, we have

$$H(x-0) \leq \liminf_{n \to \infty} F'_n(x) \leq \limsup_{n \to \infty} F'_n(x) \leq H(x)$$

which implies that if x is a continuity point of H, then

$$\lim F'_n(x)=H(x).$$

PROOF OF (v): If $k(0)=0$, then $k=0$, $\mu_n \to 0$, hence, (v) holds. If $k(0) \neq 0$, then we can assume that all μ_n's are probability measures without loss of generality. Since $k(t)$ is continuous at the origin and $k(0)=\lim \varphi_n(0)=1$, for each positive number ε there exists a $\delta>0$ such that

$$\left| \frac{1}{(2\delta)^p} \int_{\|t\| \leq \delta} k(t)\, dt \right| \geq 1-\frac{\varepsilon}{3}.$$

Since

$$\lim_n \int_{\|t\| \leq \delta} \varphi_n(t)\, dt = \int_{\|t\| \leq \delta} k(t)\, dt,$$

there exists a $N=N(\varepsilon)$ such that

(6.9) $$\left|\frac{1}{(2\delta)^p}\int_{\|t\|\leq\delta}\varphi_n(t)\,dt\right|\geq 1-\frac{2\varepsilon}{3},\quad \text{for all } n\geq N.$$

Now

$$\begin{aligned}\left|\int_{\|t\|\leq\delta}\varphi_n(t)\,dt\right|&=\left|\int_{\|t\|\leq\delta}dt\int_{R_p}e^{it'x}\,d\mu_n\right|=\left|\int_{R_p}d\mu_n\int_{\|t\|\leq\delta}e^{it'x}\,dt\right|\\&\leq\left|\int_{\|x\|<l}d\mu_n\int_{\|t\|\leq\delta}e^{it'x}\,dt\right|+\left|\int_{\|x\|\geq l}d\mu_n\int_{\|t\|\leq\delta}e^{it'x}\,dt\right|\\&\leq(2\delta)^p\int_{\|x\|<l}d\mu_n+\frac{(2\delta)^p}{l\delta}\int_{\|x\|\geq l}d\mu_n,\end{aligned}$$

from which it follows that

(6.10) $$\left|\frac{1}{(2\delta)^p}\int_{\|t\|\leq\delta}\varphi_n(t)\,dt\right|\leq\int_{\|x\|<l}d\mu_n+\frac{1}{l\delta}\,.$$

From (6.9) and (6.10) it follows that for each $\varepsilon>0$

(6.11) $$\int_{\|x\|<l}d\mu_n\geq 1-\varepsilon$$

for all $n\geq N=N(\varepsilon)$ and for all $l\geq L=L(\varepsilon)=3/(\varepsilon\delta)$. Now $F_1, F_2,\ldots$ be distribution functions defined by the probability measures $\mu_1,\mu_2,\ldots$. By Lemma 6.1 there exists a subsequence $\{F_{n(j)}\}$ which converges to a positively monotonic function $H(x)$ at every continuity point of the latter. We can prove, from (6.11), that H is a distribution function. Since H is positively monotonic and $0\leq H(x)\leq 1$ it remains to prove that

(6.12) $$\liminf_{x_1,\ldots,x_p\to\infty} H(x_1,\ldots,x_p)\geq 1,$$

(6.13) $$\limsup_{x_j\to-\infty} H(x_1,\ldots,x_p)\leq 0,\qquad j=1,2,\ldots,p.$$

Take a point $(y_1,\ldots,y_p)$ such that $\min y_j>L$. Then there exists a l such that $\min y_j>l>L$ and that the point $(l,\ldots,l)$ is a continuity point of H. Since

$$F_n(l,\ldots,l)\geq\int_{\|x\|<l}d\mu_n\geq 1-\varepsilon,\qquad n\geq N,$$

letting $n\to\infty$ through the sequence $\{n(j)\}$, we have

$$H(l,\ldots,l)\geq 1-\varepsilon,$$

hence

$$H(y_1,\ldots,y_p)\geq 1-\varepsilon,\qquad \text{if}\quad \min y_j>L,$$

which implies (6.12) Next, fix $y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_p$ arbitrarily and let $y_j < -L$. Then there exists a continuity point $(z_1, \ldots, z_p)$ of H such that $y_k \leq z_k$, $k=1, 2, \ldots, p$ and $z_j < -L$. Then

$$F_n(z_1, \ldots, z_p) \leq \int_{\|x\| \geq L} d\mu_n \leq \varepsilon, \qquad n \geq N,$$

and letting $n \to \infty$ through the sequence $\{n(j)\}$ we have

$$H(z_1, \ldots, z_p) \leq \varepsilon,$$

hence

$$H(y_1, \ldots, y_p) \leq \varepsilon, \qquad \text{if} \quad y_j < -L,$$

which implies (6.13). Thus H must be a distribution function. Write $H=F$, and let φ be the characteristic function of F. Then from $\lim_{j\to\infty} F_{n(j)}=F$ it follows that $\lim_{j\to\infty} \varphi_{n(j)}=\varphi$ by (iv). Hence $k=\varphi$ and, since F is uniquely determined by φ, any convergent subsequence of $\{F_n\}$ must converge to F, hence, $\{F_n\}$ itself must converge to F.

7. Continuous amplitudes of non-vanishing characteristic functions

Let z be a complex number different from 0. Any real number θ such that $z=|z|e^{i\theta}$ is called an amplitude of z and is denoted by $\theta=\operatorname{amp} z$. $\operatorname{amp} z$ is determined up to the multiple of 2π. If $\theta=\operatorname{amp} z$ and $-\pi<\theta\leq\pi$ then θ is called the principal amplitude of z and is denoted by $\operatorname{Amp} z$. Let φ be a characteristic function. A function θ will be called *continuous amplitude* of φ if

(7.1) $\theta(t)=\operatorname{amp}\varphi(t)$ for all $t \in R_p$,
θ is continuous in R_p and $\theta(0)=0$.

LEMMA 7.1 *Any non-vanishing characteristic function φ has a continuous amplitude, which is uniquely determined by φ.*

PROOF: (i) First, we wish to prove that, for each $T>0$, there exists a function θ, $\{\theta(t); |t| \leq T\}$, satisfying (7.1) in the domain $\{t; |t| \leq T\}$. Since any continuous function defined on a compact space attains its minimum value there, we have

$$\eta=\min_{|t|\leq T} |\varphi(t)| > 0.$$

As φ is uniformly continuous in $\{t; |t| \leq T\}$, to this η corresponds a $\delta>0$ such that

$$|\varphi(t)-\varphi(t')|<\eta, \quad \text{for} \quad |t-t'| \leq \delta, \quad |t| \leq T, \ |t'| \leq T.$$

Hence

(7.2) $\quad |\varphi(t)/\varphi(t')-1|<1 \quad \text{for} \quad |t-t'|\leq\delta,\ |t|\leq T,\ |t'|\leq T.$

Fix a t' such that $|t'|\leq T$. Then the function $\varphi(t)/\varphi(t')$, of t, is continuous and takes values in the right half plane in the domain $|t-t'|\leq\delta$, $|t|\leq T$, hence, the function of t

(7.3) $\quad \mathrm{Amp}\,(\varphi(t)/\varphi(t'))$ (t' fixed) is continuous in $|t-t'|\leq\delta$, $|t|\leq T$, and vanishes at $t=t'$. Furthermore,

(7.4) $$\left|\mathrm{Amp}\frac{\varphi(t)}{\varphi(t')}\right|<\frac{\pi}{2}, \qquad |t-t'|\leq\delta, \quad |t|\leq T.$$

Now let t_0 be any fixed unit vector, i.e., $|t_0|=1$. We shall define $\theta(t_0\tau)$ in the interval $0\leq\tau\leq T$, as follows:

$$\theta(t_0\tau)=\sum_{k=1}^{n}\mathrm{Amp}\frac{\varphi(k\delta t_0)}{\varphi((k-1)\delta t_0)}+\mathrm{Amp}\frac{\varphi(\tau t_0)}{\varphi(n\delta t_0)} \quad \text{for } n\delta\leq\tau<(n+1)\delta, \quad n=0,1,2,\dots.$$

If $n=0$, the first term in the right side vanishes. Thus, $\theta(t)$ is defined for all t with $|t|\leq T$ and it is seen that

(7.5) $\quad \theta(t)=\mathrm{amp}\,\varphi(t),\quad$ for all t with $|t|\leq T$,

(7.6) $\quad \theta(\tau_2 t)-\theta(\tau_1 t)=\mathrm{Amp}\,(\varphi(\tau_2 t)/\varphi(\tau_1 t)),\quad$ if $|\tau_1 t-\tau_2 t|\leq\delta$,

where t is a vector and τ_1, τ_2 are real numbers. (7.5) is obvious from the definition of $\theta(t)$ and $\varphi(0)=1$. To show (7.6), we can assume that t is a unit vector without loss of generality. Moreover, we can assume that $|\tau_2|\geq|\tau_1|$ and $\tau_2\geq 0$ as (7.4) is true and t can be replaced by $-t$ if necessary. Then the following three cases can occur:

Case (a) $\quad 0\leq n\delta\leq\tau_1\leq\tau_2<(n+1)\delta$, for some n;

Case (b) $\quad 0\leq n\delta\leq\tau_1<(n+1)\delta\leq\tau_2<(n+2)\delta$, for some n;

Case (c) $\quad -\delta<\tau_1\leq 0\leq\tau_2<\delta$.

For instance, in case (b), we have

$$\theta(\tau_2 t)-\theta(\tau_1 t)=\mathrm{Amp}\frac{\varphi((n+1)\delta t)}{\varphi(n\delta t)}+\mathrm{Amp}\frac{\varphi(\tau_2 t)}{\varphi((n+1)\delta t)}-\mathrm{Amp}\frac{\varphi(\tau_1 t)}{\varphi(n\delta t)}$$

$$\equiv\mathrm{Amp}\frac{\varphi(\tau_2 t)}{\varphi(\tau_1 t)} \qquad (\mathrm{mod}\ 2\pi).$$

But by (7.4)

$$|\text{the second side} - \text{the third side}|<\frac{4\pi}{2}=2\pi.$$

Therefore '$\equiv$' becomes '$=$', and (7.6) holds. In the other cases it will also be proved similarly.

Now we can show that for any t_1 and t_2 such that $|t_1|\leq T$ and $|t_2|\leq T$, (7.7) $\theta(t_1)-\theta(t_2)=\mathrm{Amp}\,(\varphi(t_1)/\varphi(t_2)), \quad \text{if } |t_1-t_2|\leq\delta.$

Fix t_1 and t_2 and take a positive number n such that $|t_1|\leq n\delta$, $|t_2|\leq n\delta$. It is shown by induction that

(7.8) $$\theta\Big(\frac{kt_1}{n}\Big)-\theta\Big(\frac{kt_2}{n}\Big)=\mathrm{Amp}\,\frac{\varphi(kt_1/n)}{\varphi(kt_2/n)} \quad \text{for } k=0,1,2,\dots,n,$$

which becomes (7.7) if $k=n$. In case $k=0$ it is clear. Assume that (7.8) holds for some $k<n$. Then we have

$$\theta\Big(\frac{k+1}{n}t_1\Big)-\theta\Big(\frac{k+1}{n}t_2\Big)$$
$$=\Big[\theta\Big(\frac{k+1}{n}t_1\Big)-\theta\Big(\frac{k}{n}t_1\Big)\Big]+\Big[\theta\Big(\frac{k}{n}t_1\Big)-\theta\Big(\frac{k}{n}t_2\Big)\Big]+\Big[\theta\Big(\frac{k}{n}t_2\Big)-\theta\Big(\frac{k+1}{n}t_2\Big)\Big]$$
$$=\mathrm{Amp}\,\frac{\varphi((k+1)t_1/n)}{\varphi(kt_1/n)}+\mathrm{Amp}\,\frac{\varphi(kt_1/n)}{\varphi(kt_2/n)}+\mathrm{Amp}\,\frac{\varphi(kt_2/n)}{\varphi((k+1)t_2/n)}$$
$$\equiv\mathrm{Amp}\,\frac{\varphi((k+1)t_1/n)}{\varphi((k+1)t_2/n)}.$$

By making use of (7.4), '$\equiv$' becomes '$=$' and we have (7.8) with k replaced by $k+1$. Hence (7.7) holds. From (7.3) and (7.7) it is seen that θ is continuous in $\{t;\,|t|\leq T\}$ which together with (7.5) completes the proof of this step. (ii) We shall show that if both θ_1 and θ_2 are continuous amplitudes in $\{t;\,|t|\leq T\}$, then $\theta_1=\theta_2$. Since $\theta_1(t)\equiv\theta_2(t)$ (mod 2π) for all t, $\theta_1-\theta_2$ can take only values $0, \pm2\pi, \pm4\pi, \dots$. On the other hand, $\theta_1-\theta_2$ is continuous. Therefore, $\theta_1-\theta_2$ must be a constant. Hence for all t $\theta_1(t)-\theta_2(t)=\theta_1(0)-\theta_2(0)=0$, i.e., $\theta_1=\theta_2$. (iii) Denote by θ_T the continuous amplitude in $\{t;\,|t|\leq T\}$ defined in (i). Then from (ii) we have

$$\theta_T=\theta_{T'} \quad \text{for} \quad |t|\leq\min(T,T').$$

Define θ as follows

$$\theta(t)=\theta_n(t) \quad \text{for} \quad n-1\leq|t|\leq n, \quad n=1,2,\dots.$$

Then θ becomes a continuous amplitude of φ. Thus the lemma is completely proved.

Let φ be a non-vanishing characteristic function and let θ be its continuous amplitude. Then

$$\varphi(t)=\exp\{\log|\varphi(t)|+i\theta(t)\}, \quad t\in R_p.$$

$\log|\varphi|+i\theta$ is called *continuous logarithm* of the φ and is denoted by $\log\varphi$:

(7.9) $$\log \varphi = \log |\varphi| + i\theta.$$

By definition $\psi = \log \varphi$ *is characterized by the following conditions:* $\psi(t)$ *is a continuous function of* t, *vanishes at the origin* $t=0$, *and* $\varphi(t) = \exp\{\psi(t)\}$. The continuous logarithm of a non-vanishing characteristic function φ is uniquely determined by φ. Furthermore, for each positive integer n, the nth root of a non-vanishing characteristic function φ is defined by

(7.10) $$\varphi^{1/n} = \exp\{(\log \varphi)/n\},$$

where $\log \varphi$ is the continuons logarithm of φ. If a non-vanishing characteristic function φ_1 is equal to the nth power of another characteristic function φ_2, i.e., if

(7.11) $$\varphi_1 = \varphi_2^n$$

then $\varphi_2 = {\varphi_1}^{1/n}$, for (7.11) is rewritten as $e^{\log \varphi_1} = e^{n \log \varphi_2}$, which implies $\log \varphi_1 = n \log \varphi_2$ by the uniqueness of the continuous logarithm, hence, $\varphi_2 = e^{\log \varphi_2} = e^{(\log \varphi_1)/n} = {\varphi_1}^{1/n}$. In the sequel the continuous logarithm of a non-vanishing characteristic function φ is called simply the logarithm of φ.

We shall need in the sequel the following

LEMMA 7.2 *If a sequence of non-vanishing characteristic functions* $\{\varphi_n(t)\}$ *converges to a non-vanishing characteristic function* $\varphi_0(t)$ *at every point* $t \in R_p$, *then* $\{\log \varphi_n(t)\}$ *converges to* $\log \varphi_0(t)$ *uniformly in every bounded* t *set.*

PROOF: Let T be a fixed positive number. By the well-known theorem we have

(7.12) $$\lim \varphi_n(t) = \varphi_0(t) \quad \text{uniformly in} \quad |t| \leq T,$$

from which it follows that

(7.13) $$\lim |\varphi_n(t)| = |\varphi_0(t)| \quad \text{uniformly in} \quad |t| \leq T,$$

hence,

$$\lim_n \min_{|t| \leq T} |\varphi_n(t)| = \min_{|t| \leq T} |\varphi_0(t)|.$$

Since $\min_{|t| \leq T} |\varphi_n(t)| > 0$ for all n, there exists an $\eta > 0$ such that

(7.14) $$\min_{|t| \leq T} |\varphi_n(t)| > \eta \quad \text{for all} \quad n = 0, 1, 2, \ldots.$$

From (7.13) and (7.14) it follows that

(7.15) $$\lim \log |\varphi_n(t)| = \log |\varphi_0(t)| \quad \text{uniformly in} \quad |t| \leq T.$$

Now, since $\{\varphi_n(t)\}$ is equicontinuous in $\{t; |t| \leq T\}$, to the η corresponds a δ such that

(7.16) $\quad |\varphi_n(t)-\varphi_n(t')|<\eta \quad$ for $\quad |t-t'|\leq\delta, \quad n=0, 1, 2,\ldots.$

From (7.14) and (7.16) it follows that

(7.17) $\quad |\varphi_n(t)/\varphi_n(t')-1|<1 \quad$ for $\quad |t-t'|\leq\delta, \quad n=0, 1, 2,\ldots.$

Therefore for any fixed t' the function $\operatorname{Amp}(\varphi_n(t)/\varphi_n(t'))$, of t, is continuous in $|t-t'|\leq\delta$ and vanishes at $t=t'$ for $n=0, 1, 2,\ldots.$ It is just so with the function $\theta_n(t)-\theta_n(t')$ of t, where θ_n is the continuous amplitude of φ_n for each n. But

$$\theta_n(t)-\theta_n(t')=\operatorname{amp}\varphi_n(t)-\operatorname{amp}\varphi_n(t')\equiv\operatorname{Amp}(\varphi_n(t)/\varphi_n(t')) \pmod{2\pi}.$$

Therefore it must hold that

(7.18) $\quad \theta_n(t)-\theta_n(t')=\operatorname{Amp}(\varphi_n(t)/\varphi_n(t')) \quad$ for $\quad |t-t'|\leq\delta, \; n=0, 1, 2,\ldots.$

From (7.12) and (7.14) it follows that

$$\lim \{\varphi_n(t)/\varphi_n(t')\}=\varphi_0(t)/\varphi_0(t') \quad \text{uniformly in} \quad |t|\leq T, \; |t'|\leq T.$$

And if $|t-t'|\leq\delta$, the value $\varphi_n(t)/\varphi_n(t')$ lies by (7.14) and (7.17) in the domain $\{z; |z-1|<1, \; |z|\geq\eta\}$. From these two facts it follows that for any fixed t'

(7.19) $\quad \lim \operatorname{Amp}\{\varphi_n(t)/\varphi_n(t')\}=\operatorname{Amp}\{\varphi_0(t)/\varphi_0(t')\}$ uniformly in $|t-t'|\leq\delta.$

From (7.18) and (7.19) we have, for any fixed t',

(7.20) $\quad \lim(\theta_n(t)-\theta_n(t'))=\theta_0(t)-\theta_0(t') \quad$ uniformly in $\quad |t-t'|\leq\delta.$

Let the open sphere $S(t, \delta/2)$ with centre t and radius $\delta/2$ correspond to each point t in $|t|\leq T$. According to the Heine-Borel theorem, the compact set, $|t|\leq T$, is covered by the sum of finite number, say m, of such spheres. Denote the finite set of centres of those spheres by M. Let us assume that M contains the origin. Then from (7.20) for each $\varepsilon>0$ there exists an $N=N(\varepsilon)$ such that

(7.21) $$|(\theta_n(t)-\theta_n(t'))-(\theta_0(t)-\theta_0(t'))|<\varepsilon/m$$

for $n\geq N$, $|t-t'|\leq\delta$, $t'\in M$. From this fact we can prove that

(7.22) $\quad |\theta_n(t)-\theta_0(t)|<\varepsilon \quad$ for $\quad n>N=N(\varepsilon), \; |t|\leq T.$

Fix a t, $|t|\leq T$. The segment joining the origin 0 and the point t is covered by the sum of open spheres with radius $\delta/2$ and with centres $0=t_0, t_1, t_2,\ldots, t_k$, such that

$$t_j\in M, \qquad j=0, 1, 2,\ldots, k,$$
$$S(t_{j-1}, \delta/2)\frown S(t_j, \delta/2)\neq 0, \qquad j=1, 2,\ldots, k,$$
$$t\in S(t_k, \delta/2).$$

Then, since

$$|t_j - t_{j-1}| < \delta, \qquad j=1, 2, \ldots, k,$$
$$|t - t_k| < \delta,$$

we have, from (7.21), for $n \geq N$

$$\text{(7.23)} \quad \begin{aligned} \Delta_j &= |(\theta_n(t_j)-\theta_n(t_{j-1}))-(\theta_0(t_j)-\theta_0(t_{j-1}))| < \varepsilon/m, \qquad j=1, 2, \ldots, k. \\ \Delta &= |(\theta_n(t)-\theta_n(t_k))-(\theta_0(t)-\theta_0(t_k))| < \varepsilon/m. \end{aligned}$$

As $\theta_n(t_0)=\theta_0(t_0)=0$, from (7.23) we have

$$|\theta_n(t)-\theta_0(t)| \leq \sum_{j=1}^{k} \Delta_j + \Delta < \frac{k+1}{m}\varepsilon \leq \varepsilon, \quad \text{for } n \geq N.$$

Since N does not depend on t, (7.22) holds, hence

$$\text{(7.24)} \qquad \lim \theta_n(t) = \theta_0(t) \quad \text{uniformly in} \quad |t| \leq T.$$

(7.15) and (7.24) complete the proof, as $\log \varphi_n(t) = \log|\varphi_n(t)| + i\theta_n(t)$ for $n=0, 1, 2, \ldots$

8. Infinitely divisible distributions

A distribution is called *infinitely divisible* if, for each positive integer n, its characteristic function φ is the nth power of a characteristic function ψ, $\varphi = \psi^n$.

LEMMA 8.1 *The characteristic function of a infinitely divisible distribution cannot take the value* 0.

PROOF: Let φ be the characteristic function. Then to each n corresponds a characteristic function φ_n such that $\varphi = \varphi_n^n$, from which it follows that $|\varphi_n|^2 = |\varphi|^{2/n}$. Hence

$$\lim |\varphi_n(t)|^2 = \begin{cases} 1, & \text{if} \quad \varphi(t) \neq 0, \\ 0, & \text{if} \quad \varphi(t) = 0. \end{cases}$$

Put $\psi = \lim |\varphi_n|^2$. As $|\varphi_n|^2$ is a characteristic function and $\psi(t)=1$ in a neighborhood of the origin, by the continuity theorem ψ is a characteristic function. Thus ψ is continuous hence $\psi(t)=1$, $\varphi(t) \neq 0$ for every t.

In this case $\varphi_n = \varphi^{1/n} = e^{(\log \varphi)/n}$, from which it follows that $\lim \varphi_n(t) = 1$ at every t. Hence, the distribution function with the characteristic function φ_n tends to the unit distribution function U.

In the remainder of this part the following notations will be used: a with or without a subscript denotes a vector in R_p; σ with or without a subscript denotes a non-negative definite matrix of pth order, the (j, k)th elements, i.e., elements in the jth row and kth column, of

σ and σ_n are denoted by σ_{jk} and $\sigma_{jk}^{(n)}$, respectively, so that $\sigma=(\sigma_{jk})$ and $\sigma_n=(\sigma_{jk}^{(n)})$; and μ with or without a subscript denotes a p-dimensional measure with $\mu(R_p)<\infty$ and $\mu(\{0\})=0$.

LEMMA 8.2 *The function* ψ *defined by*

$$\psi(t)=ia't-\frac{1}{2}t'\sigma t+\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}d\mu \tag{8.1}$$

is the logarithm of the characteristic function of a infinitely divisible distribution.

PROOF: First we shall show that $e^{\psi(t)}$ is a characteristic function. It is easily seen that the integrand in the right side of (8.1)

$$\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}$$

is bounded and continuous with respect to (x, t) in the domain $|x|>0$, $|t|\leq T$ for each positive number T. Let us take an ε such that $0<\varepsilon<1$ and let us consider the integral

$$I_\varepsilon(t)=\int_{\varepsilon<\|x\|<1/\varepsilon}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}d\mu.$$

Divide the integration domain into disjoint subintervals J_k $(k=1, 2,\dots, n)$, choose a point $x_{(k)}$ from each J_k, and make an approximation sum

$$S=\sum_k\left(e^{it'x_{(k)}}-1-\frac{it'x_{(k)}}{1+x'_{(k)}x_{(k)}}\right)\frac{1+x'_{(k)}x_{(k)}}{x'_{(k)}x_{(k)}}\mu(J_k).$$

Then we have

$$S=\sum_k\lambda_k(e^{it'x_{(k)}}-1)+ib't,$$

where

$$\lambda_k=\frac{1+x'_{(k)}x_{(k)}}{x'_{(k)}x_{(k)}}\mu(J_k)\geq 0,$$

$$b=\sum_k\frac{x_{(k)}}{x'_{(k)}x_{(k)}}\mu(J_k).$$

It is easily verified that $\lambda_k(e^{it'x_{(k)}}-1)$ is the logarithm of the characteristic function of a p-dimensional random variable $x_{(k)}y$, where y is a real random variable whose distribution is Poissonian with mean λ_k, hence, e^S is also a p-dimensional characteristic function. Letting $\max_k$ (diameter of J_k) tend to 0, S converges to a continuous function $I_\varepsilon(t)$. Therefore, by the continuity theorem, $eI^{\varepsilon(t)}$ is a characteristic function. By the same theorem, as

$$I_0(t)=\lim_{\varepsilon\downarrow 0} I_\varepsilon(t)=\int_{\|x\|>0}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu$$
$$=\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu,$$

$e^{I_0(t)}$ is a characteristic function. On the other hand $e^{ia't-2^{-1}t'\sigma t}$ is a normal characteristic function. Therefore $e^{\psi(t)}$ is a characteristic function. Then, for each positive integer n, since $\psi(t)/n$ is written in the form (8.1) with a, σ, μ replaced by a/n, σ/n, μ/n, $e^{\psi(t)/n}$ is also a characteristic function, hence $e^{\psi(t)}$ is the characteristic function of a infinitely divisible distribution. Furthermore, $\psi(t)$ is the logarithm of $e^{\psi(t)}$, as it is continuous and vanishes at $t=0$.

LEMMA 8.3 *In* (8.1) a, σ, *and* μ *are uniquely determined by* ψ.

PROOF: Form

$$\varphi(t)=\psi(t)-\frac{1}{2^p}\int_{\|s-t\|\leq 1}\psi(s)\,ds \tag{8.2}$$
$$=\frac{1}{6}\sum_j\sigma_{jj}+\int_{R_p}e^{it'x}\left(1-\prod_j\frac{\sin x_j}{x_j}\right)\frac{1+x'x}{x'x}\,d\mu,$$

where it is supposed that $(\sin\xi)/\xi=1$ for $\xi=0$.

Write

$$\nu(E)=\int_E\left(1-\prod_j\frac{\sin x_j}{x_j}\right)\frac{1+x'x}{x'x}\,d\mu+\frac{1}{6}\sum_j\sigma_{jj}\chi(E),\qquad E\in B_p, \tag{8.3}$$

where

$$\chi(E)=\begin{cases}1, & \text{if } 0\in E,\\ 0, & \text{if } 0\notin E.\end{cases}$$

Then ν is a p-dimensional measure with $\nu(R_p)<\infty$ and it holds that

$$\varphi(t)=\int_{R_p}e^{it'x}\,d\nu. \tag{8.4}$$

Write

$$g(x)=\left(1-\prod_{j=1}^{p}\frac{\sin x_j}{x_j}\right)\frac{1+x'x}{x'x}\quad\text{for}\quad x\neq 0,\quad g(0)=\frac{1}{6}\,. \tag{8.5}$$

Note that g is continuous everywhere and there exist two constants c_1 and c_2 such that

$$0<c_1<g(x)<c_2\qquad\text{for all}\quad x\in R_p, \tag{8.6}$$

since

$$\lim_{x\to\infty} g(x)=1,$$

and

(8.7) $$\lim_{x\to 0} g(x)=1/6.$$

(See the end of this section). From (8.4) ν is uniquely determined by φ, hence, by ψ. Write

$$\nu_0(E)=\nu(E-\{0\})=\nu(E)-\frac{1}{6}\sum_j \sigma_{jj}\chi(E),$$

which is uniquely determined by ν. Then (8.3) becomes

$$\nu_0(E)=\int_E\Big(1-\prod_j \frac{\sin x_j}{x_j}\Big)\frac{1+x'x}{x'x}\,d\mu,\quad E\in B_p,$$

from which we see that μ is represented by

$$\mu(E)=\int_E\Big[\Big(1-\prod_j \frac{\sin x_j}{x_j}\Big)\frac{1+x'x}{x'x}\Big]^{-1}d\nu_0,\quad E\in B_p,$$

and μ is uniquely determined by ν_0. After all μ is uniquely determined by ψ. Hence it is just so with

$$ia't-\frac{1}{2}t'\sigma t=\psi(t)-\int_{R_p}\Big(e^{it'x}-1-\frac{it'x}{1+x'x}\Big)\frac{1+x'x}{x'x}\,d\mu.$$

By taking its real and imaginary parts it is seen that a and σ are uniquely determined by ψ.

LEMMA 8.4 *If the function*

(8.8) $$\psi_n(t)=ia_n't-\frac{1}{2}t'\sigma_n t+\int_{R_p}\Big(e^{it'x}-1-\frac{it'x}{1+x'x}\Big)\frac{1+x'x}{x'x}\,d\mu_n$$

converges to a function $l(t)$, continuous at the origin, at every point $t\in R_p$, then there exist a, σ, and μ such that

(8.9) $$\lim a_n=a,\qquad \lim(\sigma_n+\tau_n)=\sigma+\tau,$$
$$\lim \mu_n(E)=\mu(E)$$

for every continuity set E of μ whose closure $\bar{E}$ does not contain the origin, and l coincides with ψ determined by

(8.10) $$\psi(t)=ia't-\frac{1}{2}t'\sigma t+\int_{R_p}\Big(e^{it'x}-1-\frac{it'x}{1+x'x}\Big)\frac{1+x'x}{x'x}\,d\mu,$$

where $\tau_n=(\tau_{jk}^{(n)})$ and $\tau=(\tau_{jk})$ are matrices with the elements

(8.11) $$\tau_{jk}^{(n)}=\int_{R_p}\frac{x_jx_k}{x'x}\,d\mu_n,\qquad \tau_{jk}=\int_{R_p}\frac{x_j\,x_k}{x'x}\,d\mu.$$

Conversely if (8.9) *holds, then* ψ_n *converges to* ψ.

PROOF: Assume the hypotheses of the direct part. Then $e^{l(t)}=\lim e^{\psi_n(t)}$ is, by the continuity theorem, a characteristic function. Let $l_1(t)$ be the logarithm of the characteristic function $e^{l(t)}$, defined by (7.9). Then by Lemma 7.2 $\psi_n(t)$ converges to $l_1(t)$ uniformly in every bounded t set. Hence we have $l_1=l$ and

$$\lim \psi_n(t)=l(t) \tag{8.12}$$

uniformly in every bounded t set. Let us form

$$\varphi_n(t)=\psi_n(t)-\frac{1}{2^p}\int_{\|s-t\|\leq 1}\psi_n(s)\,ds, \qquad t\in R_p. \tag{8.13}$$

Then we have

$$\varphi_n(t)=\int_{R_p} e^{it'x}\,d\nu_n, \tag{8.14}$$

with

$$\nu_n(E)=\int_E\left(1-\prod_j\frac{\sin x_j}{x_j}\right)\frac{1+x'x}{x'x}\,d\mu_n+\frac{1}{6}\sum_j\sigma_{jj}^{(n)}\chi(E), \quad E\in B_p. \tag{8.15}$$

Let us write

$$\varphi(t)=l(t)-\frac{1}{2^p}\int_{\|s-t\|\leq 1}l(s)\,ds. \tag{8.16}$$

From (8.12), (8.13), and (8.16) it follows that

$$\lim \varphi_n(t)=\varphi(t), \quad \text{for all} \quad t\in R_p. \tag{8.17}$$

As φ is continuous, by the continuity theorem there exists a p-dimensional measure ν such that

$$\lim \nu_n=\nu. \tag{8.18}$$

Now put

$$\bar\nu_n(E)=\nu_n(E)-\frac{1}{6}\sum_j\sigma_{jj}^{(n)}\chi(E)=\int_E\left(1-\prod_j\frac{\sin x_j}{x_j}\right)\frac{1+x'x}{x'x}\,d\mu_n, \quad E\in B_p, \tag{8.19}$$

then from (8.18) and (8.19) it follows that

$$\lim \bar\nu_n(E)=\nu(E) \tag{8.20}$$

for every continuity set E of ν which does not contain the origin. Owing to (8.19) μ_n is represented by

$$\mu_n(E)=\int_E\left\{\left(1-\prod_j\frac{\sin x_j}{x_j}\right)\frac{1+x'x}{x'x}\right\}^{-1}d\bar\nu_n, \quad E\in B_p. \tag{8.21}$$

Hence by (8.20) we have

(8.22) $$\lim \mu_n(E)=\mu_0(E),$$

for every continuity set E, not containing 0, of ν, where

$$\mu_0(E)=\int_E \left\{\left(1-\prod_j \frac{\sin x_j}{x_j}\right)\frac{1+x'x}{x'x}\right\}^{-1} d\nu, \quad E \in B_p.$$

Since any continuity set of μ_0 is a continuity set of ν, (8.22) holds for every continuity set E, not containing 0, of μ_0. Let us put

$$\mu(E)=\mu_0(E-\{0\}), \qquad E \in B_p.$$

Then we have

(8.23) $$\lim \mu_n(E)=\mu(E)$$

for every continuity set E of μ, whose closure $\bar{E}$ does not contain the origin. We can show that $\{\mu_n(R_p);\, n=1, 2, \dots\}$ is bounded. From (8.6), (8.21), and (8.19) we have

$$\mu_n(R_p)< c_1^{-1}\,\bar{\nu}_n(R_p)< c_1^{-1}\,\nu_n(R_p),$$

and (8.18) implies that

$$\lim \nu_n(R_p)=\nu(R_p).$$

Hence $\{\mu_n(R_p)\}$ is bounded. Now ψ_n can be written as

(8.24) $$\psi_n(t)=ia_n't-\frac{1}{2}t'(\sigma_n+\tau_n)t+\int_{R_p} h(x, t)\, d\mu_n,$$

where

$$h(x, t)=\left(e^{it'x}-1-\frac{it'x}{1+x'x}+\frac{(t'x)^2}{2(1+x'x)}\right)\frac{1+x'x}{x'x}, \quad x \neq 0,$$

$$h(0, t)=0.$$

It is easily seen that for each $t_0 \in R_p$,

$$\lim_{x\to 0,\, t\to t_0} h(x, t)=0,$$

and $h(x, t)$ is bounded and continuous in $x \in R_p$, $|t| \leq T$ for each $T>0$. Hence, by Lemma 8.5 (below) we have

(8.25) $$\lim \int_{R_p} h(x, t)\, d\mu_n=\int_{R_p} h(x, t)\, d\mu.$$

From (8.24), (8.25) and (8.12) it is seen that the limit

$$\lim \left\{ia_n't-\frac{1}{2}\, t'(\sigma_n+\tau_n)t\right\}$$

exists and hence also the limits

(8.26) $$\lim a_n=a \quad \text{(say)}, \qquad \lim (\sigma_n+\tau_n)=\sigma^* \quad \text{(say)}$$

exist. After all we have

$$l(t)=ia't-\frac{1}{2}t'\sigma^*t+\int_{R_p}h(x,t)\,d\mu$$

which can be written as

$$(8.27)\qquad l(t)=ia't-\frac{1}{2}t'\sigma t+\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu,$$

where

$$(8.28)\qquad \sigma=\sigma^*-\tau,$$

τ being defined by (8.11). Next we must show that σ is non-negative definite. Now, $\psi_n(t)$ can be written, for each $\varepsilon>0$, as

$$\psi_n(t)=ia_n't+\int_{|x|>\varepsilon}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu_n$$
$$+\int_{|x|\le\varepsilon}h(x,t)\,d\mu_n-\frac{1}{2}\left\{t'\sigma_nt+\int_{|x|\le\varepsilon}\frac{(t'x)^2}{x'x}\,d\mu_n\right\}.$$

Assume that the set $\{x\,;\,|x|>\varepsilon\}$ is a continuity set of μ. Since each term, except the last in the right side, converges as n tends to ∞ (cf. Lemma 8.5), so is also the case with the latter and we have

$$l(t)=ia't+\int_{|x|>\varepsilon}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu$$
$$+\int_{|x|\le\varepsilon}h(x,t)\,d\mu-\frac{1}{2}\lim_{n\to\infty}\left\{t'\sigma_nt+\int_{|x|\le\varepsilon}\frac{(t'x)^2}{x'x}\,d\mu_n\right\}.$$

Let ε tend to 0. Then the third term in the right side converges to 0 and we have

$$l(t)=ia't+\int_{|x|>0}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu$$
$$-\frac{1}{2}\lim_{\varepsilon\to0}\lim_{n\to\infty}\left\{t'\sigma_nt+\int_{|x|\le\varepsilon}\frac{(t'x)^2}{x'x}\,d\mu_n\right\}.$$

Comparing this with (8.27) we have

$$t'\sigma t=\lim_{\varepsilon\to0}\lim_{n\to\infty}\left\{t'\sigma_nt+\int_{|x|\le\varepsilon}\frac{(t'x)^2}{x'x}\,d\mu_n\right\}$$

from which it is seen that σ is non-negative definite. This completes the proof of the direct part.

To prove the converse part, let us rewrite (8.8) and (8.10) as

$$\psi_n(t)=ia_n't-\frac{1}{2}t'(\sigma_n+\tau_n)t+\int_{R_p} h(x,t)\,d\mu_n,$$

$$\psi(t)=ia't-\frac{1}{2}t'(\sigma+\tau)t+\int_{R_p} h(x,t)\,d\mu.$$

We can show the boundedness of the sequence $\{\mu_n(R_p);\, n=1,2,\dots\}$, which together with (8.9) and Lemma 8.5 implies that $\lim\psi_n=\psi$ and completes the proof. Now notice that

$$\mu_n(R_p)=\sum_j\int_{R_p}\frac{x_j^2}{x'x}\,d\mu_n=\sum_j\tau_{jj}^{(n)}. \tag{8.29}$$

From the hypothesis (8.9) we have

$$\lim\sum_j(\sigma_{jj}^{(n)}+\tau_{jj}^{(n)})=\sum_j(\sigma_{jj}+\tau_{jj}),$$

hence, there exists a constant K, independent of n, such that

$$\sum_j(\sigma_{jj}^{(n)}+\tau_{jj}^{(n)})<K,\qquad n=1,2,\dots. \tag{8.30}$$

Since $\sigma_{jj}^{(n)}\geq 0$, (8.29) and (8.30) imply that

$$\mu_n(R_p)<K,\qquad n=1,2,\dots.$$

LEMMA 8.5 *Assume that*

$$\mu_n(\{0\})=0,\qquad \mu_n(R_p)\leq K\qquad (n=0,1,2,\dots),$$
$$\lim\mu_n(E)=\mu_0(E)$$

for every continuity set E of μ_0 whose closure $\bar{E}$ does not contain the origin. If $h(x)$ is bounded and continuous in the whole space, and if

$$\lim_{x\to 0}h(x)=0 \tag{8.31}$$

then

$$\lim\int_{R_p}h(x)d\mu_n=\int_{R_p}h(x)d\mu_0. \tag{8.32}$$

Moreover, if $\{x;\,|x|<\varepsilon\}$ is a continuity set of μ_0,

$$\lim_n\int_{|x|\leq\varepsilon}h(x)\,d\mu_n=\int_{|x|\leq\varepsilon}h(x)\,d\mu_0. \tag{8.33}$$

PROOF: For any $\delta>0$ we have

$$\left|\int_{R_p}h(x)\,d\mu_n-\int_{R_p}h(x)\,d\mu_0\right|$$
$$\leq\left|\int_{|x|>\delta}h(x)\,d\mu_n-\int_{|x|>\delta}h(x)d\mu_0\right|+\left|\int_{|x|\leq\delta}h(x)\,d\mu_n\right|+\left|\int_{|x|\leq\delta}h(x)\,d\mu_0\right|.$$

If the set $\{x; |x| > \delta\}$ is a continuity set of μ_0, the first term in the right side tends to zero, hence,

$$\limsup_{n\to\infty}\left|\int_{R_p} h(x)\, d\mu_n - \int_{R_p} h(x)\, d\mu_0\right| \le 2K \sup_{|x|\le\delta} |h(x)|.$$

This together with (8.31) implies (8.32) as δ may be chosen arbitrarily small. (8.33) follows from (8.32) and the fact that

$$\lim \int_{|x|>\varepsilon} h(x)\, d\mu_n = \int_{|x|>\varepsilon} h(x)\, d\mu_0.$$

THEOREM 8.1 *Let $d(t)$ be the characteristic function of any infinitely divisible distribution. Then $d(t)$ cannot take the value 0, hence, has the continuous logarithm $\log d(t)$, and it is uniquely represented in the form* (8.1). *Conversely, any function $\psi(t)$ defined by* (8.1) *is the logarithm of the characteristic function of some infinitely divisible distribution.*

PROOF: Let $d(t)$ be the characteristic function of an infinitely divisible distribution. It is shown in Lemma 8.1, that $d(t)$ cannot take the value 0. Hence the continuous logarithm, $\log d(t)$, exists. Then we have

$$\log d(t) = \lim_{n\to\infty} n\,(e^{\{\log d(t)\}/n} - 1).$$

Since $e^{\{\log d(t)\}/n} = d^{1/n}(t)$ is a characteristic function, it is written as

$$e^{\{\log d(t)\}/n} = \int_{R_p} e^{it'x}\, d P_n,$$

where P_n is a p-dimensional distribution. Hence we have

$$\begin{aligned} n(e^{\{\log d(t)\}/n} - 1) &= n\int_{R_p} (e^{it'x} - 1)\, d P_n \\ &= n\int_{R_p} \frac{it'x}{1+x'x}\, d P_n + n\int_{R_p} \left(e^{it'x} - 1 - \frac{it'x}{1+x'x}\right) d P_n \\ &= ia_n't + \int_{R_p} \left(e^{it'x} - 1 - \frac{it'x}{1+x'x}\right) \frac{1+x'x}{x'x}\, d\mu_n, \end{aligned}$$

where

$$a_n = n\int_{R_p} \frac{x}{1+x'x}\, d P_n, \qquad \mu_n(E) = n\int_E \frac{x'x}{1+x'x}\, d P_n, \qquad E \in B_p.$$

Consequently we have

$$\log d(t) = \lim_{n\to\infty} \left\{ ia_n't + \int_{R_p} \left(e^{it'x} - 1 - \frac{it'x}{1+x'x}\right) \frac{1+x'x}{x'x}\, d\mu_n \right\}$$

with

$$\mu_n(R_p)<\infty, \qquad \mu_n(\{0\})=0.$$

According to Lemma 8.4, $\log d(t)$ can be represented in the form (8.1). The uniqueness of the representation is proved in Lemma 8.3. The converse is shown in Lemma 8.2.

THEOREM 8.2 *If a sequence of infinitely divisible distributions converges to some distribution, the limit distribution is also infinitely divisible.*

PROOF: Let $\{\varphi_n\}$ be the corresponding sequence of the characteristic functions and let φ_0 be the characteristic function of the limiting distribution. Then $\lim \varphi_n(t)=\varphi_0(t)$ for all t, hence, for each positive integer k, $\lim |\varphi_n(t)|^{2/k}=|\varphi_0(t)|^{2/k}$ for all t. Since φ_n is infinitely divisible, $|\varphi_n|^{2/k}$ is also a characteristic function. And $|\varphi_0|^{2/k}$ is continuous. Hence, by the continuity theorem, $|\varphi_0|^{2/k}$ must be a characteristic function. As this holds for every $k=1, 2,\dots$, φ_0 cannot take the value 0 (see the proof of Lemma 8.1). According to Lemma 7.2

$$\lim \log \varphi_n(t)=\log \varphi_0(t), \qquad t \in R_p,$$

from which we see that, for any positive integer k, $\varphi_n{}^{1/k}(t)=e^{\{\log \varphi_n(t)\}/k}$ converges to $\varphi_0{}^{1/k}(t)=e^{\{\log \varphi_0(t)\}/k}$. By the continuity theorem $\varphi_0{}^{1/k}(t)$ is also a characteristic function. As this holds for every $k=1, 2,\dots$, φ_0 must be infinitely divisible.

THEOREM 8.3 *Let L_n, $n=0, 1, 2,\dots$, be infinitely divisible distributions defined by*

$$\psi_n(t)=ia_n't-\frac{1}{2}t'\sigma_n t+\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}d\mu_n.$$

Then L_n converges to L_0 if and only if the following three conditions hold:

$$\lim a_n=a_0, \qquad \lim(\sigma_n+\tau_n)=\sigma_0+\tau_0,$$

$$\lim \mu_n(E)=\mu_0(E)$$

for every continuity set E of μ_0 whose closure $\bar{E}$ does not contain the origin, where τ_n and τ_0 are defined by (8.11).

This follows from Lemmas 7.2, 8.3, and 8.4.

Now we shall verify (8.7) by induction. When $p=1$, it is easily verified. Assume that (8.7) holds for some p. Then for any given ε such that $0<\varepsilon<1/6$, there exists a $\delta>0$ such that

$$\text{(8.34)} \quad \frac{1}{6}-\varepsilon<\left(1-\prod_{j=1}^{p}\frac{\sin x_j}{x_j}\right)\frac{1+\sum_1^p x_j^2}{\sum_1^p x_j^2}<\frac{1}{6}+\varepsilon, \qquad \text{if} \quad 0<\sum_1^p x_j^2<\delta,$$

(8.35) $$\frac{1}{6}-\varepsilon<\left(1-\frac{\sin x_{p+1}}{x_{p+1}}\right)\frac{1+x_{p+1}^2}{x_{p+1}^2}<\frac{1}{6}+\varepsilon, \qquad \text{if} \quad 0<x_{p+1}^2<\delta.$$

Now assume that

$$\sum_{j=1}^{p+1} x_j^2<\delta, \qquad \sum_1^p x_j^2 \neq 0, \qquad x_{p+1}\neq 0,$$

and write

$$s_p=\sum_1^p x_j^2.$$

Then both (8.34) and (8.35) hold and can be rewritten as

(8.36) $$1-\left(\frac{1}{6}+\varepsilon\right)\frac{s_p}{1+s_p}<\prod_1^p \frac{\sin x_j}{x_j}<1-\left(\frac{1}{6}-\varepsilon\right)\frac{s_p}{1+s_p},$$

(8.37) $$1-\left(\frac{1}{6}+\varepsilon\right)\frac{x_{p+1}^2}{1+x_{p+1}^2}<\frac{\sin x_{p+1}}{x_{p+1}}<1-\left(\frac{1}{6}-\varepsilon\right)\frac{x_{p+1}^2}{1+x_{p+1}^2}.$$

Form (8.36)×(8.37), then

$$1-\left(\frac{1}{6}+\varepsilon\right)\left(\frac{s_p}{1+s_p}+\frac{x_{p+1}^2}{1+x_{p+1}^2}\right)<\prod_1^{p+1}\frac{\sin x_j}{x_j}$$
$$<1-\left(\frac{1}{6}-\varepsilon\right)\left(\frac{s_p}{1+s_p}+\frac{x_{p+1}^2}{1+x_{p+1}^2}\right)+\left(\frac{1}{6}-\varepsilon\right)^2\frac{s_p}{1+s_p}\cdot\frac{x_{p+1}^2}{1+x_{p+1}^2}.$$

On the other hand we have

$$\frac{s_p+x_{p+1}^2}{1+s_p+x_{p+1}^2}<\frac{s_p}{1+s_p}+\frac{x_{p+1}^2}{1+x_{p+1}^2}<(1+\delta)\frac{s_p+x_{p+1}^2}{1+s_p+x_{p+1}^2},$$
$$\frac{s_p}{1+s_p}\cdot\frac{x_{p+1}^2}{1+x_{p+1}^2}<\frac{\delta}{2}\cdot\frac{s_p+x_{p+1}^2}{1+s_p+x_{p+1}^2}.$$

Hence we have

$$1-\left(\frac{1}{6}+\varepsilon\right)(1+\delta)\frac{s_{p+1}}{1+s_{p+1}}<\prod_1^{p+1}\frac{\sin x_j}{x_j}<1-\left\{\left(\frac{1}{6}-\varepsilon\right)-\left(\frac{1}{6}-\varepsilon\right)^2\frac{\delta}{2}\right\}\frac{s_{p+1}}{1+s_{p+1}},$$

where $s_{p+1}=s_p+x_{p+1}^2=\sum_1^{p+1}x_j^2$. Assume that δ is so small that

$$\left(\frac{1}{6}+\varepsilon\right)(1+\delta)<\frac{1}{6}+2\varepsilon, \quad \left(\frac{1}{6}-\varepsilon\right)-\left(\frac{1}{6}-\varepsilon\right)^2\frac{\delta}{2}>\frac{1}{6}-2\varepsilon.$$

Then we have

$$1-\left(\frac{1}{6}+2\varepsilon\right)\frac{s_{p+1}}{1+s_{p+1}}<\prod_1^{p+1}\frac{\sin x_j}{x_j}<1-\left(\frac{1}{6}-2\varepsilon\right)\frac{s_{p+1}}{1+s_{p+1}}$$

from which it follows that

(8.38) $$\frac{1}{6}-2\varepsilon<\left(1-\prod_{1}^{p+1}\frac{\sin x_j}{x_j}\right)\frac{1+\sum_{1}^{p+1}x_j^2}{\sum_{1}^{p+1}x_j^2}<\frac{1}{6}+2\varepsilon.$$

Thus, for any given ε, $0<\varepsilon<1/6$, there exists a $\delta>0$ such that (8.38) holds if $\sum_{1}^{p+1}x_j^2<\delta$. Therefore (8.7) holds with p replaced by $p+1$.

9. Infinitely divisible distributions in the generalized sense

A distribution is called *infinitely divisible in the generalized sense* (following J. L. Doob [4], p. 129) if, for each $\eta>0$, its distribution function F can be written as a convolution of distribution functions $F_1, F_2, \dots, F_n$,

(9.1) $$F=F_1 * F_2 * \cdots * F_n$$

with

(9.2) $$\int_{|x|\geq\eta} dF_j(x) \leq \eta, \qquad j=1, 2, \dots, n.$$

If F is an infinitely divisible distribution function, for each n, F can be written as a convolution

$$F=G_n * G_n * \cdots * G_n \qquad (n \text{ times})$$

with G_n tending to the unit distribution function U. Thus any infinitely divisible distribution is infinitely divisible in the generalized sense. In this section it is shown that the converse is also true. We shall begin with the following lemmas.

LEMMA 9.1 *Let* $\{F_{nl}; l=1, 2, \dots, l_n, n=1, 2, \dots\}$ *be a sequence of one-dimensional distribution functions. Write*

$$a_{nl}=\int_{|x|<1} x\, dF_{nl}(x).$$

If, for each $\varepsilon>0$,

$$\lim_{n} \max_{l} \int_{|x|\geq\varepsilon} dF_{nl}(x)=0,$$

then we have

$$\lim_{n} \max_{l} \int_{|x|<1} (x-a_{nl})^2\, dF_{nl}(x)=0.$$

Moreover, if the convolution $F_{n1} * F_{n2} * \cdots * F_{nl}$ *converges to a distribution function and if* ε_1 *is a positive constant, then there exists a constant* c, *independent of* n, *such that*

$$\sum_l \int_{|x| \geq \varepsilon_1} dF_{nl}(x) \leq c, \qquad n=1, 2, \ldots,$$

$$\sum_l \int_{|x|<1} (x-a_{nl})^2 \, dF_{nl}(x) \leq c, \qquad n=1, 2, \ldots .$$

c depends on the given sequences $\{F_{nl}\}$ *and* ε_1. (see M. Loève [17], section II, 4, and section IV, 2*.)

LEMMA 9.2 *Let* $\{F_{nl};\ l=1, 2, \ldots, l_n,\ n=1, 2, \ldots\}$ *be a sequence of p-dimensional distribution functions. Assume that for each* $\varepsilon>0$

(9.3) $$\lim_n \max_l \int_{|x| \geq \varepsilon} dF_{nl}(x)=0$$

and that the convolution $F_{n1} * F_{n2} * \cdots * F_{nl_n}$ *converges to a distribution function. Let* φ_{nl} *be the characteristic function of* F_{nl}, *and put*

$$a_{nl}=\int_{\|x\|<1} x \, dF_{nl}(x),$$

$$\gamma_{nl}(t)=\varphi_{nl}(t)\, e^{-it' a_{nl}}-1.$$

Then we have

(9.4) $$\lim_n \sum_l |\gamma_{nl}(t)|^2=0.$$

PROOF: Let us denote the one-dimensional marginal distribution functions of F_{nl} by F_{nlj} $(j=1, 2, \ldots, p)$, and let us put

$$b_{nlj}=\int_{|x|<1} x \, dF_{nlj}(x),$$

$$b_{nl}=(b_{nl1}, b_{nl2}, \ldots, b_{nlp}),$$

$$a_{nl}=(a_{nl1}, a_{nl2}, \ldots, a_{nlp}).$$

Then we have

$$|b_{nlj}-a_{nlj}|=\left|\int_{|x_j|<1} x_j \, dF_{nl}(x)-\int_{\|x\|<1} x_j \, dF_{nl}(x)\right|$$

$$=\left|\int_{|x_j|<1,\ \|x\| \geq 1} x_j \, dF_{nl}(x)\right| \leq \int_{\|x\| \geq 1} dF_{nl}(x),$$

and therefore

(9.5) $$|b_{nl}-a_{nl}|^2=\sum_j |b_{nlj}-a_{nlj}|^2 \leq p \int_{\|x\| \geq 1} dF_{nl}(x).$$

---

* Japanese readers may also refer to Y. Kawada [12], Appendix II, lemma 6 and III, lemma 8.

Now we have

$$|\gamma_{nl}(t)|=\left|\int_{R_p}(e^{it'(x-a_{nl})}-1)\,dF_{nl}(x)\right|$$

$$\leq 2\int_{\|x\|\geq 1} dF_{nl}(x)+\left|\int_{\|x\|<1} it'(x-a_{nl})\,dF_{nl}(x)\right|+\frac{1}{2}\int_{\|x\|<1}|t'(x-a_{nl})|^2\,dF_{nl}(x)$$

$$=2I_1+I_2+I_3 \quad \text{(say)},$$

where we used the well-known formula

$$|e^{i\xi}-1-i\xi|\leq \xi^2/2, \qquad -\infty<\xi<\infty.$$

But

$$I_2=|t'a_{nl}-t'a_{nl}\int_{\|x\|<1} dF_{nl}(x)|=|t'a_{nl}\int_{\|x\|\geq 1} dF_{nl}(x)|$$

$$\leq |t|\cdot|a_{nl}|I_1\leq p^{1/2}|t|I_1,$$

$$I_3\leq \frac{|t|^2}{2}\int_{\|x\|<1}|x-a_{nl}|^2\,dF_{nl}(x)$$

$$\leq |t|^2\left(\int_{\|x\|<1}|x-b_{nl}|^2\,dF_{nl}(x)+|b_{nl}-a_{nl}|^2\right)$$

$$\leq |t|^2\left(\int_{\|x\|<1}|x-b_{nl}|^2\,dF_{nl}(x)+pI_1\right),$$

by (9.5). Hence, we have

$$|\gamma_{nl}(t)|\leq (2+p^{1/2}|t|+p|t|^2)I_1+|t|^2\int_{\|x\|<1}|x-b_{nl}|^2\,dF_{nl}(x). \tag{9.6}$$

Now, by the inequality

$$\int_{\|x\|\geq 1} dF_{nl}(x)\leq \int_{|x|\geq 1} dF_{nl}(x)$$

and the hypothesis (9.3) we have

$$\lim_n \max_l I_1=\lim_n \max_l \int_{\|x\|\geq 1} dF_{nl}(x)=0. \tag{9.7}$$

According to Lemma 9.1

$$\lim_n \max_l \int_{|x|<1}(x-b_{nlj})^2\,dF_{nlj}(x)=0, \qquad j=1, 2, \ldots, p.$$

From this and the inequality

$$\int_{\|x\|<1}|x-b_{nl}|^2\,dF_{nl}(x)\leq \sum_j \int_{|x|<1}(x-b_{nlj})^2\,dF_{nlj}(x)$$

we have

$$\lim_n \max_l \int_{\|x\|<1} |x-b_{nl}|^2 dF_{nl}(x)=0. \tag{9.8}$$

From (9.6), (9.7), and (9.8) it follows that

$$\lim_n \max_l |\gamma_{nl}(t)|=0. \tag{9.9}$$

On the other hand, for each $j(=1, 2\ldots)$, the sequence $\{F_{nlj};\ l=1,2,\ldots, l_n,\ n=1, 2,\ldots\}$ of one-dimensional distribution functions satisfies the hypotheses of Lemma 9.1 and therefore has a constant c_j, the existence of which is mentioned in Lemma 9.1 with $\varepsilon_1=1$. Put $c=\max\limits_{1\leq j\leq p} c_j$. Then

$$\sum_l \int_{\|x\|\geq 1} dF_{nl}(x) \leq \sum_l \sum_j \int_{|x|\geq 1} dF_{nlj}(x)=\sum_j \sum_l \int_{|x|\geq 1} dF_{nlj}(x) \leq pc,$$

$$\sum_l \int_{\|x\|<1} |x-b_{nl}|^2 dF_{nl}(x) \leq \sum_l \sum_j \int_{|x|<1} (x-b_{nlj})^2 dF_{nlj}(x) \leq pc.$$

These together with (9.6) imply that

$$\sum_l |\gamma_{nl}(t)| \leq \{2+p^{1/2}|t|+(p+1)|t|^2\}\, pc=K(t) \quad \text{(say)}. \tag{9.10}$$

Hence,

$$\sum_l |\gamma_{nl}(t)|^2 \leq \max_l |\gamma_{nl}(t)|\cdot\sum_l |\gamma_{nl}(t)| \leq K(t) \max_l |\gamma_{nl}(t)|. \tag{9.11}$$

(9.9) and (9.11) imply (9.4) q.e.d.

LEMMA 9.3. *Assume the hypotheses in Lemma 9.2. Let V be a neighborhood of the origin, and put*

$$\tilde{a}_{nl}=\int_V x\, dF_{nl}(x),$$

$$\tilde{\gamma}_{nl}(t)=\varphi_{nl}(t)\, e^{-it'\tilde{a}_{nl}}-1.$$

Then it holds that

$$\lim_n \sum_l |\tilde{\gamma}_{nl}(t)|^2=0. \tag{9.12}$$

By a neighborhood of the origin we mean a bounded Borel set which contains a sphere with the origin as its centre.

PROOF: As V is a neighborhood of the origin there exist two positive numbers ε_1 and ε_2 such that

$$S(0, \varepsilon_1) \subset V \subset S(0, \varepsilon_2),$$

where $S(0, \varepsilon)$ denotes the open sphere with centre 0 and radius ε. $\{x;\ |x|<\varepsilon\}$. We may assume that $\varepsilon_1 \leq 1$. Write

$$\tilde{a}_{nl}=(\tilde{a}_{nl1}, \tilde{a}_{nl2}, \ldots, \tilde{a}_{nlp}).$$

Then for each $j=1, 2, \ldots, p$, we have

$$|a_{nlj}-\tilde{a}_{nlj}|=\left|\int_{\|x\|<1} x_j\, dF_{nl}(x)-\int_V x_j\, dF_{nl}(x)\right|$$

(9.13)
$$\leq\int_{\|x\|<1,\, V^c} |x_j|\, dF_{nl}(x)+\int_{V,\, \|x\|\geq 1} |x_j|\, dF_{nl}(x)$$

$$\leq(1+\varepsilon_2)\int_{|x|\geq\varepsilon_1} dF_{nl}(x),$$

where V^c denotes the complement of V. This together with (9.3) implies that

(9.14)
$$\lim_n \max_l |a_{nlj}-\tilde{a}_{nlj}|=0, \quad j=1, 2, \ldots, p.$$

By Lemma 9.1 there exists a constant c such that

$$\sum_l\int_{|x|\geq\varepsilon_1} dF_{nl}(x)<c, \qquad \text{for all } n.$$

Hence from (9.13) it follows that

(9.15)
$$\sum_l |a_{nlj}-\tilde{a}_{nlj}| \leq (1+\varepsilon_2)c, \qquad \text{for all } n \text{ and } j.$$

From (9.14) and (9.15) it follows that

$$\lim_n \sum_l |a_{nlj}-\tilde{a}_{nlj}|^2=0, \qquad j=1, 2, \ldots, p.$$

Add these from $j=1$ to $j=p$. Then

(9.16)
$$\lim_n \sum_l |a_{nl}-\tilde{a}_{nl}|^2=0.$$

On the other hand

$$|\gamma_{nl}(t)-\tilde{\gamma}_{nl}(t)|=|\varphi_{nl}(t)(e^{-it'a_{nl}}-e^{-it'\tilde{a}_{nl}})|$$
$$=|e^{-it'(a_{nl}-\tilde{a}_{nl})}-1| \leq |t'(a_{nl}-\tilde{a}_{nl})| \leq |t|\cdot|a_{nl}-\tilde{a}_{nl}|.$$

This together with (9.16) implies that

$$\lim_n \sum_l |\gamma_{nl}(t)-\tilde{\gamma}_{nl}(t)|^2=0,$$

and this together with (9.4) implies (9.12). Thus the lemma is proved.

THEOREM 9.1 *Let $\{F_{nl};\ l=1, 2, \ldots, l_n,\ n=1, 2, \ldots\}$ be a sequence of distribution functions. If F_{nl} converge to the unit distribution function U uniformly in l as n tends to ∞, and the convolution $F_{n1} * F_{n2} * \cdots * F_{nl_n}$ converges to a distribution function F, then F must be infinitely divisible.*

The proof runs in the same way as in the one-dimensional case, but we shall give a proof, for the completeness following M. Loève [17].

PROOF: Let φ_{nl} be the characteristic function of F_{nl}, let V be a neighborhood of the origin, and put

(9.17)
$$\begin{cases} a_{nl}=\int_V x\,d\,F_{nl}(x), \\ \gamma_{nl}(t)=\varphi_{nl}(t)\,e^{-ia'_{nl}t}-1, \\ \log\varphi^*_{nl}(t)=ia'_{nl}\,t+\gamma_{nl}(t). \end{cases}$$

Since $\log\varphi^*_{nl}(t)$ is written as

$$\log\varphi^*_{nl}(t)=ia'_{nl}\,t+\int_{R_p}(e^{it'x}-1)\,d\,F_{nl}(x+a_{nl}),$$

$\varphi^*_{nl}(t)$ is the characteristic function of an infinitely divisible distribution (see the proof of Lemma 8.2). Now from (9.17) it follows that

$$\begin{aligned}|\varphi_{nl}(t)-\varphi^*_{nl}(t)|&=|e^{ia'_{nl}t}(1+\gamma_{nl}(t)-e^{\gamma_{nl}(t)})|\\ &\leq(1/2)\,e^{|\gamma_{nl}(t)|}\,|\gamma_{nl}(t)|^2\leq 5\,|\gamma_{nl}(t)|^2,\end{aligned}$$

as $|\gamma_{nl}(t)|\leq 2$. Hence

(9.18)
$$|\varphi_{nl}(t)-\varphi^*_{nl}(t)|\leq 5\,|\gamma_{nl}(t)|^2.$$

Put

$$\varphi^*_n(t)=\prod_l\varphi^*_{nl}(t),\qquad \varphi_n(t)=\prod_l\varphi_{nl}(t).$$

Since the convolution of any finite sequence of infinitely divisible distributions is also infinitely divisible, φ^*_n must be the characteristic function of an infinitely divisible distribution. On the other hand

(9.19)
$$|\varphi_n(t)-\varphi^*_n(t)|\leq\sum_l|\varphi_{nl}(t)-\varphi^*_{nl}(t)|,$$

by using the fact, easily proved by induction, that $|a_l|\leq 1$ and $|b_l|\leq 1$, $l=1,2,\ldots,l_n$, imply $|\Pi a_l-\Pi b_l|\leq\sum|a_l-b_l|$. From (9.18) and (9.19) it follows that

$$|\varphi_n(t)-\varphi^*_n(t)|\leq 5\sum_l|\gamma_{nl}(t)|^2.$$

According to Lemma 9.3

$$\lim_n\sum_l|\gamma_{nl}(t)|^2=0.$$

Therefore

(9.20)
$$\lim_n|\varphi_n(t)-\varphi^*_n(t)|=0.$$

Denote by $\varphi(t)$ the characteristic function of F. Then from the hypothesis it holds that

(9.21)
$$\lim\varphi_n(t)=\varphi(t).$$

(9.20) and (9.21) imply that

$$\lim \varphi_n^*(t)=\varphi(t).$$

Therefore, according to Theorem 8.2, F must be infinitely divisible.

COROLLARY *An infinitely divisible distribution in the generalized sense is infinitely divisible.*

10. Convergence theorem

THEOREM 10.1 *Assume that F_{nl}, $l=1, 2, \ldots, l_n$, converge to U uniformly in $1 \leq l \leq l_n$ as $n \to \infty$. Then the convolution $F_{n1} * F_{n2} * \cdots * F_{nl_n}$ converges to the infinitely divisible distribution function defined by*

$$(10.1) \qquad \psi(t)=ia't-\frac{1}{2}t'\sigma t+\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)\frac{1+x'x}{x'x}\,d\mu,$$

if and only if the following three conditons hold:

$$(10.2) \qquad \lim_n \sum_l \left(a_{nl}+\int_{R_p}\frac{x}{1+x'x}\,dF_{nl}(x+a_{nl})\right)=a,$$

$$(10.3) \quad \lim_n \sum_l \int_{R_p}\frac{x_j x_k}{1+x'x}\,dF_{nl}(x+a_{nl})=\sigma_{jk}+\int_{R_p}\frac{x_j x_k}{x'x}\,d\mu, \quad j, k=1, 2, \ldots, p,$$

$$(10.4) \qquad \lim_n \sum_l \int_E \frac{x'x}{1+x'x}\,dF_{nl}(x+a_{nl})=\mu(E),$$

for every continuity set E, with closure not containing the origin, of μ, where

$$a_{nl}=\int_V x\,dF_{nl}(x),$$

and V is an arbitrarily fixed neighborhood of the origin.

PROOF: The notations adopted in the proof of Theorem 9.1 are used. By (9.20)

$$\lim \varphi_n(t)=e^{\psi(t)}$$

if and only if

$$\lim \varphi_n^*(t)=e^{\psi(t)}.$$

Now we have

$$\log \varphi_n^*(t)=\sum_l\left\{ia_{nl}'t+\int_{R_p}(e^{it'x}-1)\,dF_{nl}(x+a_{nl})\right\}$$

$$=\sum_l\left\{ia_{nl}'t+\int_{R_p}\frac{it'x}{1+x'x}\,dF_{nl}(x+a_{nl})+\int_{R_p}\left(e^{it'x}-1-\frac{it'x}{1+x'x}\right)dF_{nl}(x+a_{nl})\right\},$$

hence,

$$\log \varphi_n^*(t) = i a_n' t + \int_{R_p} \left(e^{it'x} - 1 - \frac{it'x}{1+x'x} \right) \frac{1+x'x}{x'x} d\mu_n, \tag{10.5}$$

where

$$a_n = \sum_l \left\{ a_{nl} + \int_{R_p} \frac{x}{1+x'x} d F_{nl}(x+a_{nl}) \right\},$$

$$\mu_n(E) = \sum_l \int_E \frac{x'x}{1+x'x} d F_{nl}(x+a_{nl}), \qquad E \in B_p.$$

From Theorem 8.3 $\varphi_n^*(t)$ converges to $e^{\psi(t)}$ if and only if (10.2)–(10.4) simultaneously hold. Thus, the proof is completed.

In the above proof, let $\{b_n\}$ be any sequence of vectors. Then we have

$$| \varphi_n(t) e^{-ib_n't} - \varphi_n^*(t) e^{-ib_n't} | = | \varphi_n(t) - \varphi_n^*(t) |,$$

and

$$\log (\varphi_n^*(t) e^{-ib_n't}) = i(a_n - b_n)'t + \int_{R_p} \left(e^{it'x} - 1 - \frac{it'x}{1+x'x} \right) \frac{1+x'x}{x'x} d\mu_n.$$

Therefore we have the following

COROLLARY *Assume that* F_{nl}, $l=1, 2, \ldots, l_n$, *converge to* U *uniformly in* $1 \leq l \leq l_n$ *as* $n \to \infty$, *and let* $\{b_n\}$ *be a sequence of vectors. Then the convolution* $F_{n1} * F_{n2} * \cdots * F_{nl_n} * U(\cdot + b_n)$ *converges to the distribution function defined by* (10.1) *if and only if*

$$\lim \left\{ \sum_l \left(a_{nl} + \int_{R_p} \frac{x}{1+x'x} d F_{nl}(x+a_{nl}) \right) - b_n \right\} = a, \tag{10.6}$$

(10.3) *and* (10.4) *hold simultaneously.*

Let us notice that, *in Theorem 10.1, if the limiting distribution is non-unit the sequence* $\{l_n\}$ *must be tend to* ∞. To prove this, assume that $\{l_n\}$ has a bounded subsequence $\{l_{n(j)}\}$, $l_{n(j)} < L$, $j=1, 2, \ldots$, and put

$$G_n = F_{n1} * F_{n2} * \cdots * F_{nl_n}.$$

Then, for each $\varepsilon > 0$, it holds that

$$\int_{|x| \geq \varepsilon} d G_{n(j)}(x) \leq \sum_{l=1}^{l_{n(j)}} \int_{|x| \geq \varepsilon/L} d F_{n(j)l}(x) \leq L \max_l \int_{|x| \geq \varepsilon/L} d F_{n(j)l}(x) \to 0.$$

Hence

$$\lim_{j \to \infty} G_{n(j)} = U,$$

this contradicts the hypothesis.

Part III Central limit theorem

11. General case

The normal distribution with mean vector m and covariance matrix $\sigma=(\sigma_{jk})$ (matrix of the second order central moments σ_{jk}), i.e., the distribution with characteristic function $\exp(im't-\frac{1}{2}t'\sigma t)$ is denoted by $N(m, \sigma)$. Unit distributions, though often denoted by $N(m, 0)$, are not called normal in this paper. The most general version of the p-dimensional central limit theorem is given by the following theorem, from which various versions of the central limit theorem can be deduced. In Theorem 11.1 the same notations as in the preceding three sections are used.

THEOREM 11.1 *Assume that*

$$\lim_n \max_{1\leq l\leq l_n} \int_{|x|\geq\varepsilon} dF_{nl}(x)=0, \quad \text{for each } \varepsilon>0. \tag{11.1}$$

Then the distribution defined by the convolution

$$F_{n1} * F_{n2} * \cdots * F_{nl_n} * U(\cdot+b_n)$$

converges to a normal distribution $N(0, \sigma)$, *if and only if the following three conditions hold:*

$$\lim_n \sum_l \int_{|x|\geq\varepsilon} dF_{nl}(x)=0, \quad \text{for each } \varepsilon>0, \tag{11.2}$$

$$\lim_n \Big(\sum_l \int_V x\, dF_{nl}(x)-b_n\Big)=0, \tag{11.3}$$

$$\lim_n \sum_l \Big(\int_V x_j x_k\, dF_{nl}(x)-\int_V x_j\, dF_{nl}(x)\int_V x_k\, dF_{nl}(x)\Big)=\sigma_{jk}, \quad j, k=1, 2,\ldots, p, \tag{11.4}$$

where V is an arbitrarily fixed neighborhood of the origin.

Let $X_{n1}, X_{n2}, \ldots, X_{nl_n}$ be independent random variables with distribution functions $F_{n1}, F_{n2},\ldots, F_{nl_n}$ for each $n=1, 2,\ldots$. Then (11.1) means that X_{nl}, $l=1, 2,\ldots, l_n$, converge to 0 in probability uniformly in $1\leq l\leq l_n$ as $n\to\infty$, i.e., that each term is asymptotically individually negligible. Under (11.1), (11.2) is equivalent to the condition that $\max_l |X_{nl}|$ converges to 0 in probability as $n\to\infty$, i.e., that the greatest term is asymptotically negligible (P. Lévy [16], §34). Let us put

$$X'_{nl}(\omega)=\begin{cases} X_{nl}(\omega), & \text{if } X_{nl}(\omega)\in V, \\ 0, & \text{if } X_{nl}(\omega)\notin V. \end{cases}$$

Then (11.2) implies that the difference between the distribution functions of $X_{n1}+\cdots+X_{nl_n}-b_n$ and $X'_{n1}+\cdots+X'_{nl_n}-b_n$ is asymptotically negligible (see section 13), (11.3) means that the mean vector of $X'_{n1}+\cdots+X'_{nl_n}-b_n$ converges to the mean vector of the limit distribution, and (11.4) represents that the covariance matrix of $X'_{n1}+\cdots+X'_{nl_n}-b_n$ converges to the covariance matrix of the limit distribution.

Let V_1 and V_2 be two neighborhoods of the origin. Then under (11.2), (11.3) and (11.4) with $V=V_1$ are equivalent to (11.3) and (11.4) with $V=V_2$, respectively (see Lemma 15.1).

PROOF OF THEOREM 11.1 According to Corollary to Theorem 10.1 the distribution defined by $F_{n1}*F_{n2}*\cdots*F_{nl_n}*U(\cdot+b_n)$ converges to $N(0,\sigma)$ if and only if the following three conditions hold:

(11.5) $$\lim_n\sum_l\int_E\frac{x'x}{1+x'x}\,dF_{nl}(x+a_{nl})=0, \qquad \text{if } \bar{E}\not\ni 0,$$

(11.6) $$\lim_n\left[\sum_l\left(a_{nl}+\int_{R_p}\frac{x}{1+x'x}\,dF_{nl}(x+a_{nl})\right)-b_n\right]=0,$$

(11.7) $$\lim_n\sum_l\int_{R_p}\frac{x_jx_k}{1+x'x}\,dF_{nl}(x+a_{nl})=\sigma_{jk}, \quad j,k=1,2,\ldots,p,$$

where

(11.8) $$a_{nl}=\int_V x\,dF_{nl}(x),$$

Now (11.5) can be replaced by

$$\lim_n\sum_l\int_{|x|\geq\varepsilon}\frac{x'x}{1+x'x}\,dF_{nl}(x+a_{nl})=0, \qquad \text{for each } \varepsilon>0,$$

hence, by

(11.5′) $$\lim_n\sum_l\int_{|x|\geq\varepsilon}dF_{nl}(x+a_{nl})=0, \qquad \text{for each } \varepsilon>0.$$

And this implies that

$$\lim_n\sum_l\int_{V^c}dF_{nl}(x+a_{nl})=0,$$

from which it follows that

$$\lim_n \sum_l \int_{V^c} \frac{x}{1+x'x} dF_{nl}(x+a_{nl})=0,$$

$$\lim_n \sum_l \int_{V^c} \frac{x_j x_k}{1+x'x} dF_{nl}(x+a_{nl})=0, \qquad j, k=1, 2, \ldots, p.$$

Therefore, under (11.5′), (11.6) and (11.7) can, respectively, be replaced by

$$(11.6') \qquad \lim_n \Big[\sum_l \Big(a_{nl}+\int_V \frac{x}{1+x'x} dF_{nl}(x+a_{nl})\Big)-b_n\Big]=0,$$

$$(11.7') \qquad \lim_n \sum_l \int_V \frac{x_j x_k}{1+x'x} dF_{nl}(x+a_{nl})=\sigma_{jk}, \qquad j, k=1, 2, \ldots, p.$$

Thus, (11.5)-(11.7) are equivalent to (11.5′)-(11.7′). Next we shall show that under (11.5′), (11.7′) is equivalent to the following condition:

$$(11.7'') \qquad \lim_n \sum_l \int_V x_j x_k \, dF_{nl}(x+a_{nl})=\sigma_{jk}, \qquad j, k=1, \ldots, p.$$

To prove this, since

$$x_j x_k - \frac{x_j x_k}{1+x'x} = x_j x_k \frac{x'x}{1+x'x},$$

it is sufficient to show that each of (11.7′) and (11.7″), together with (11.5′), implies

$$(11.9) \qquad \lim_n \sum_l \int_V x_j x_k \frac{x'x}{1+x'x} dF_{nl}(x+a_{nl})=0.$$

Notice that there exists two positive numbers $\varepsilon_1<\varepsilon_2$ such that

$$S(0, \varepsilon_1) \subset V \subset S(0, \varepsilon_2).$$

Assume (11.5′) and (11.7′). Then dividing $\int_V$ into $\int_{V, |x|<\eta} + \int_{V, |x|\geq\eta}$ we have

$$\limsup_{n\to\infty} \Big|\sum_l \int_V x_j x_k \frac{x'x}{1+x'x} dF_{nl}(x+a_{nl})\Big|$$

$$\leq \limsup_{n\to\infty} \Big[\eta^2 \sum_l \int_V \frac{x'x}{1+x'x} dF_{nl}(x+a_{nl}) + \varepsilon_2^2 \sum_l \int_{|x|\geq\eta} dF_{nl}(x+a_{nl})\Big]$$

$$= \eta^2 \sum_j \sigma_{jj},$$

from which (11.9) is deduced as η may be chosen arbitrarily small. Thus (11.5′) and (11.7′) imply (11.9). This holds even if (11.7′) is exchanged by (11.7″). Therefore under (11.5′), (11.7′) is equivalent to

(11.7″). It is similarly proved that under (11.5′) and (11.7″), (11.6′) is equivalent to the following condition:

(11.6″) $$\lim_n \left[\sum_l \left(a_{nl} + \int_V x\, dF_{nl}(x+a_{nl})\right) - a_n\right] = 0.$$

By now it has been shown that (11.5), (11.6), and (11.7) are equivalent to (11.5′) (11.6″) and (11.7″). On the other hand, (11.3) is rewritten by (11.8) as

(11.3′) $$\lim_n \left(\sum_l a_{nl} - b_n\right) = 0.$$

And under (11.2), (11.4) is equivalent to

(11.4′) $$\lim_n \sum_l \int_V (x_j - a_{nlj})(x_k - a_{nlk})\, dF_{nl}(x) = \sigma_{jk}, \quad j, k = 1, 2, \ldots, p.$$

where $$a_{nl} = (a_{nl1}, a_{nl2}, \ldots, a_{nlp}).$$

It is left to show that (11.5′), (11.6″), and (11.7″) are equivalent to (11.2), (11.3′) and (11.4′). Now, from (11.1) and (11.8) it is easily proved that

(11.10) $$\lim_n \max_l |a_{nl}| = 0.$$

By (11.10), (11.5′) is equivalent to (11.2). This follows from the fact that for each $\varepsilon > 0$ there exists an $N = N(\varepsilon)$ such that

$$\max_l |a_{nl}| \leq \varepsilon/2, \qquad \text{for} \quad n \geq N,$$

and hence

$$\int_{|x|\geq 2\varepsilon} dF_{nl}(x) \leq \int_{|x-a_{nl}|\geq \varepsilon} dF_{nl}(x) \leq \int_{|x|\geq \varepsilon/2} dF_{nl}(x), \qquad \text{for} \quad n \geq N,$$

where the second side is equal to $\int_{|x|\geq\varepsilon} dF_{nl}(x+a_{nl})$. Under (11.5′) or equivalently under (11.2), (11.6″) is equivalent to (11.3′). To prove this we shall deduce

(11.11) $$\lim_n \sum_l \int_V x\, dF_{nl}(x+a_{nl}) = 0$$

from (11.2):

$$\left|\int_V x_j\, dF_{nl}(x+a_{nl})\right| = \left|\int_{V+a_{nl}} (x_j - a_{nlj})\, dF_{nl}(x)\right|$$

$$\leq \left|\int_{V+a_{nl}} (x_j - a_{nlj})\, dF_{nl}(x) - \int_V (x_j - a_{nlj})\, dF_{nl}(x)\right| + \left|\int_V (x_j - a_{nlj})\, dF_{nl}(x)\right|$$

$$\leq \int_{(V+a_{nl})\frown V^c} |x_j - a_{nlj}|\, dF_{nl}(x) + \int_{V\frown(V+a_{nl})^c} |x_j - a_{nlj}|\, dF_{nl}(x) + \left|a_{nlj}\int_{V^c} dF_{nl}(x)\right|$$

$$\leq \varepsilon_2 \int_{|x|\geq\varepsilon_1} dF_{nl}(x) + 2\varepsilon_2 \int_{|x|\geq\varepsilon_1/2} dF_{nl}(x) + \varepsilon_2 \int_{|x|\geq\varepsilon_1} dF_{nl}(x)$$

$$\leq 4\varepsilon_2 \int_{|x|\geq\varepsilon_1/2} dF_{nl}(x), \qquad n \geq N(\varepsilon_1),$$

where $V+a=\{x;\, x=y+a,\ y\in V\}$, hence,

$$\left|\sum_l \int_V x_j\, dF_{nl}(x+a_{nl})\right| \leq 4\varepsilon_2 \sum_l \int_{|x|\geq\varepsilon_1/2} dF_{nl}(x), \quad n \geq N(\varepsilon_1),\ j=1,2,\ldots,p,$$

from which (11.11) follows by (11.2). Under (11.5′) or equivalently under (11.2), (11.7″) is equivalent to (11.4′). To prove this we shall deduce

$$\text{(11.12)} \qquad \lim_n \left[\sum_l \int_{V-a_{nl}} x_j x_k\, dF_{nl}(x+a_{nl}) - \sum_l \int_V x_j x_k\, dF_{nl}(x+a_{nl})\right] = 0$$

from (11.2) and (11.5′):

$$\left|\sum_l \int_{V-a_{nl}} x_j x_k\, dF_{nl}(x+a_{nl}) - \sum_l \int_V x_j x_k\, dF_{nl}(x+a_{nl})\right|$$

$$\leq \sum_l \int_{(V-a_{nl})\frown V^c} |x_j x_k|\, dF_{nl}(x+a_{nl}) + \sum_l \int_{V\frown(V-a_{nl})^c} |x_j x_k|\, dF_{nl}(x+a_{nl})$$

$$\leq 4\varepsilon_2^2 \sum_l \int_{|x|\geq\varepsilon_1} dF_{nl}(x+a_{nl}) + \varepsilon_2^2 \sum_l \int_{|x|\geq\varepsilon_1} dF_{nl}(x) \to 0,$$

hence, (11.12) holds. Thus the proof is completed.

In the sequel, unless the contrary is explicitly stated, the following notations will be used: $X_{n1}, X_{n2}, \ldots, X_{nl_n}$ denote independent p-dimensional random variables with distribution functions $F_{n1}, F_{n2}, \ldots, F_{nl_n}$ for each $n=1,2,\ldots$;

$$S_n = X_{n1} + X_{n2} + \cdots + X_{nl_n};$$

a_n denotes a positive number and b_n denotes a p-dimensional vector for each $n=1,2,\ldots$; X, with or without a subscript, denotes a p-dimensional random variable and F denotes a p-dimensional distribution function.

COROLLARY 1 *The distribution of* $(S_n-b_n)/a_n$ *converges to a normal distribution* N(0, σ) *and* X_{nl}/a_n *converges to* 0 *in probability uniformly in* l, $1\leq l\leq l_n$, *if and only if the following three conditions hold:*

$$\text{(11.13)} \qquad \lim_n \sum_l \int_{|x|\geq\varepsilon a_n} dF_{nl}(x) = 0, \qquad \text{for each } \varepsilon > 0,$$

$$\text{(11.14)} \qquad \lim_n \frac{1}{a_n}\left(\sum_l \int_{a_n V} x\, dF_{nl}(x) - b_n\right) = 0,$$

$$\lim_n \frac{1}{a_n^2}\sum_l\Big(\int_{a_nV} x_jx_k\,dF_{nl}(x)-\int_{a_nV} x_j\,dF_{nl}(x)\int_{a_nV} x_k\,dF_{nl}(x)\Big)=\sigma_{jk} \qquad j,k=1,2,\dots,p, \tag{11.15}$$

where V is any neighborhood of the origin and $a_nV=\{x;\,x=a_ny,\ y\in V\}$.

This is a generalization of W. Feller [6]'s Satz 1. Usually a_nV is taken as $|x|<a_n$ or $\|x\|<a_n$.

PROOF: Apply the theorem with F_{nl} and b_n replaced by $F_{nl}(a_n\cdot)$ and b_n/a_n, respectively.

Let $\{X_{nl};\ l=1,2,\dots,l_n,\ n=1,2,\dots\}$ be a sequence of random variables. If there exist sequences $\{a_n\}$ and $\{b_n\}$ such that the distribution of $(X_{n1}+\cdots+X_{nl_n}-b_n)/a_n$ converges to a normal distribution and X_{nl}/a_n, $1\le l\le l_n$, converge to 0 in probability uniformly in l as $n\to\infty$, then it is said that $\{X_{nl}\}$ *obeys the central limit theorem.* As before the α-dispersion of a random variable X is denoted by $D_X(\alpha)$.

COROLLARY 2 *Assume that there exists an α such that $D_{S_n}(\alpha)>0$ for all n, and put $D_n=D_{S_n}(\alpha)$. Then $\{X_{nl}\}$ obeys the central limit theorem if and only if*

$$\lim_n \sum_l \int_{\|x\|\ge\varepsilon D_n} dF_{nl}(x)=0, \qquad \textit{for each}\quad \varepsilon>0, \tag{11.16}$$

and the limits

$$\lim_n \frac{1}{D_n^2}\sum_l\Big(\int_{\|x\|<D_n} x_jx_k\,dF_{nl}(x)-\int_{\|x\|<D_n} x_j\,dF_{nl}(x)\int_{\|x\|<D_n} x_k\,dF_{nl}(x)\Big)=\sigma_{jk}, \ \text{(say)} \qquad j,k=1,2,\dots,p, \tag{11.17}$$

exist. In this case the distribution of

$$\frac{1}{D_n}\sum_l\Big(X_{nl}-\int_{\|x\|<D_n} x\,dF_{nl}(x)\Big) \tag{11.18}$$

converges to the normal distribution $N(0,\sigma)$ *with mean vector* 0 *and second order central moments σ_{jk} defined by* (11.17).

PROOF: If (11.16) and (11.17) hold, (11.13)–(11.15) hold with

$$a_n=D_n,\ b_n=\sum_l\int_{\|x\|<D_n} x\,dF_{nl}(x), \qquad V=\{x;\,\|x\|<1\},$$

hence, the distribution of (11.18) converges to $N(0,\sigma)$ by Corollary 1; and X_{nl}/D_n converges to 0 in probability uniformly in $1\le l\le l_n$ by (11.16). Conversely assume that, for some sequences $\{a_n\}$ and $\{b_n\}$, the distribution of $(X_{n1}+\cdots X_{nl_n}-b_n)/a_n$ converges to a normal distribution and X_{nl}/a_n converges to 0 in probability uniformly in l. Then, since D_n/a_n converges to the α-dispersion D of the limit distribution by Lemma 4.4 and since $0<D<1$, the above assumptions hold with a_n replaced by D_n from Theorem 5.2. Therefore (11.16) and (11.17) must hold by Corollary 1.

12. Lindeberg's and Liapounov's conditions

Let X be a random variable with $\mathrm{E}(|X|^2)<\infty$*. Then

$$\mathrm{E}(|X-\mathrm{E}X|^2)$$

will be called the *variance* of X and will be denoted by $v(X)$. Then we have

$$v(X)=\mathrm{E}(|X-\mathrm{E}X|^2)=\mathrm{E}(|X|^2)-|\mathrm{E}X|^2,$$

Let F be a distribution function with $\int_{R_p}|x|^2\,dF(x)<\infty$, and put $m=\int_{R_p}x\,dF(x)$. Then

$$\int_{R_p}|x-m|^2\,dF(x)$$

will be called the *variance* of the distribution defined by F. The variance of a multi-dimensional distribution is equal to the sum of the variances of its one-dimensional marginal distributions. A normal distribution $\mathrm{N}(0,\sigma)$ with mean vector 0 and variance 1, $\sum_j \sigma_{jj}=1$, will be called a *normalized normal distribution*.

Theorem 12.1 *Suppose that each of X_{nl} has the vanishing mean vector and the finite variance and put*

$$v_{nl}=v(X_{nl}),$$
$$s_n=(v_{n1}+v_{n2}+\cdots+v_{nl_n})^{\frac{1}{2}}.$$

Let $\mathrm{N}(0,\sigma)$ be a normalized normal distribution. Then the distribution of $(X_{n1}+\cdots+X_{nl_n})/s_n$ converges to the $\mathrm{N}(0,\sigma)$ and X_{nl}/s_n converges to 0 in probability uniformly in $1\leq l\leq l_n$ as $n\to\infty$, if and only if for each $\varepsilon>0$

$$\lim_n \frac{1}{s_n^2}\sum_l \int_{|x|<\varepsilon s_n} x_j x_k\,dF_{nl}(x)=\sigma_{jk},\quad j,k=1,2,\ldots,p. \tag{12.1}$$

The condition (12.1) is a generalization of the Lindeberg's condition well-known in the one-dimensional case. The theorem can be proved

* E denotes 'mean value of' or 'mean vector of'.

by making use of characteristic functions in the same way as in the one-dimensional case, but here it is proved as a consequence of the results of the preceeding section.

PROOF: According to Corollary 1 to Theorem 11.1 it is sufficient to prove that (12.1) is equivalent to the set of the following three conditions:

$$\lim_n \sum_l \int_{|x|\geq\varepsilon s_n} dF_{nl}(x)=0, \qquad \text{for each} \quad \varepsilon>0, \tag{12.2}$$

$$\lim_n \frac{1}{s_n}\sum_l \int_{|x|<s_n} x\,dF_{nl}(x)=0, \tag{12.3}$$

$$\lim_n \frac{1}{s_n^2}\sum_l \Big(\int_{|x|<s_n} x_j x_k\,dF_{nl}(x)-\int_{|x|<s_n} x_j\,dF_{nl}(x)\int_{|x|<s_n} x_k\,dF_{nl}(x)\Big)=\sigma_{jk}, \tag{12.4}$$
$$j,\ k=1,\ 2,\ \ldots,\ p.$$

Assume (12.1). Add (12.1) with $j=k$ from $j=1$ to $j=p$. Then

$$\lim_n \frac{1}{s_n^2}\sum_l \int_{|x|<\varepsilon s_n} |x|^2\,dF_{nl}(x)=\sum_j \sigma_{jj}=1, \quad \text{for each} \quad \varepsilon>0,$$

therefore, since from the definition of s_n

$$\lim_n \frac{1}{s_n^2}\sum_l \int_{R_p} |x|^2\,dF_{nl}(x)=1, \tag{12.5}$$

it holds that

$$\lim_n \frac{1}{s_n^2}\sum_l \int_{|x|\geq\varepsilon s_n} |x|^2\,dF_{nl}(x)=0, \quad \text{for each} \quad \varepsilon>0. \tag{12.6}$$

(12.2) follows from (12.6) and the following inequality:

$$\sum_l \int_{|x|\geq\varepsilon s_n} dF_{nl}(x)\leq \frac{1}{\varepsilon^2 s_n^2}\sum_l \int_{|x|\geq\varepsilon s_n} |x|^2\,dF_{nl}(x).$$

(12.3) follows from (12.6) and the following inequality:

$$\frac{1}{s_n}\sum_l \left|\int_{|x|<s_n} x\,dF_{nl}(x)\right|\leq \frac{1}{s_n^2}\sum_l \int_{|x|\geq s_n} |x|^2\,dF_{nl}(x),$$

which follows from

$$\left|\int_{|x|<s_n} x\,dF_{nl}(x)\right|^2=\left|\int_{|x|\geq s_n} x\,dF_{nl}(x)\right|^2\leq \int_{|x|\geq s_n}|x|^2\,dF_{nl}(x)\int_{|x|\geq s_n} dF_{nl}(x)\leq \frac{1}{s_n^2}\Big(\int_{|x|\geq s_n}|x|^2\,dF_{nl}(x)\Big)^2,$$

where we used that the mean vector of F_{nl} is 0. (12.4) follows from (12.1) with $\varepsilon=1$ and the fact that

$$\left|\frac{1}{s_n^2}\sum_l\int_{|x|<s_n} x_j\,dF_{nl}(x)\int_{|x|<s_n} x_k\,dF'_{nl}(x)\right|\leq\frac{1}{s_n}\sum_l\left|\int_{|x|<s_n} x_j\,dF_{nl}(x)\right|\to 0.$$

Conversely assume that (12.2)–(12.4) hold. From (12.4) we have

$$\lim_n\left(\frac{1}{s_n^2}\sum_l\int_{|x|<s_n}|x|^2\,dF_{nl}(x)-\frac{1}{s_n^2}\sum_l\left|\int_{|x|<s_n} x\,dF_{nl}(x)\right|^2\right)=\sum_j\sigma_{jj}=1. \tag{12.7}$$

Form the difference (12.5)–(12.7). Then

$$\lim_n\left(\frac{1}{s_n^2}\sum_l\int_{|x|\geq s_n}|x|^2\,dF_{nl}(x)+\frac{1}{s_n^2}\sum_l\left|\int_{|x|<s_n} x\,dF_{nl}(x)\right|^2\right)=0,$$

from which it follows that

$$\lim_n\frac{1}{s_n^2}\sum_l\left|\int_{|x|<s_n} x\,dF_{nl}(x)\right|^2=0. \tag{12.8}$$

On the other hand

$$\begin{aligned}&\left|\frac{1}{s_n^2}\sum_l\int_{|x|<s_n} x_j\,dF_{nl}(x)\int_{|x|<s_n} x_k\,dF_{nl}(x)\right|\\ &\leq\frac{1}{s_n^2}\left[\sum_l\left(\int_{|x|<s_n} x_j\,dF_{nl}(x)\right)^2\right]^{1/2}\cdot\left[\sum_l\left(\int_{|x|<s_n} x_k\,dF_{nl}(x)\right)^2\right]^{1/2}\\ &\leq\frac{1}{s_n^2}\sum_l\left|\int_{|x|<s_n} x\,dF_{nl}(x)\right|^2,\end{aligned} \tag{12.9}$$

hence, by (12.8), the left side of (12.9) tends to 0, and (12.4) becomes

$$\lim_n\frac{1}{s_n^2}\sum_l\int_{|x|<s_n} x_jx_k\,dF_{nl}(x)=\sigma_{jk}. \tag{12.10}$$

Moreover

$$\begin{aligned}&\left|\frac{1}{s_n^2}\sum_l\int_{|x|<\varepsilon s_n} x_jx_k\,dF_{nl}(x)-\frac{1}{s_n^2}\sum_l\int_{|x|<s_n} x_jx_k\,dF_{nl}(x)\right|\\ &\leq\frac{1}{s_n^2}\sum_l\int_{|x|<\varepsilon s_n,\ |x|\geq s_n}|x_jx_k|\,dF_{nl}(x)+\frac{1}{s_n^2}\sum_l\int_{|x|<s_n,\ |x|\geq\varepsilon s_n}|x_jx_k|\,dF_{nl}(x)\\ &\leq\varepsilon^2\sum_l\int_{|x|\geq s_n}dF_{nl}(x)+\sum_l\int_{|x|\geq\varepsilon s_n}dF_{nl}(x),\end{aligned} \tag{12.11}$$

hence, by (12.2), the left side of (12.11) tends to 0, and this together with (12.10) implies (12.1)

Next we shall consider the case when $X_{n1}, X_{n2}, \ldots, X_{nl_n}$ are uniformly bounded.

COROLLARY 1 *Let* $M_n = \max_{n,l} \sup_{\omega} |X_{nl}(\omega)|$. *Suppose that* $\mathrm{E}\,X_{nl}=0$ *for all n, l and that*

$$\lim_n \frac{M_n}{s_n} = 0. \tag{12.12}$$

Let $\mathrm{N}(0, \sigma)$ *be a normalized normal distribution. Then the distribution of* $(X_{n1}+\cdots+X_{nl_n})/s_n$ *converges to the* $\mathrm{N}(0, \sigma)$ *if and only if*

$$\lim_n \frac{1}{s_n^2}\sum_l \int_{R_p} x_j x_k \, d\,F_{nl}(x) = \sigma_{jk}, \qquad j, k = 1, 2, \ldots, p. \tag{12.13}$$

PROOF: From (12.12) X_{nl}/s_n, $l=1, 2, \ldots, l_n$, converge to 0 uniformly in l with probability 1, hence, in probability. Next, for each $\varepsilon>0$ there exists an N such that $M_n<\varepsilon s_n$ for all $n \geq N$, hence,

$$\int_{|x|<\varepsilon s_n} x_j x_k \, d\,F_{nl}(x) = \int_{R_p} x_j x_k \, d\,F_{nl}(x), \qquad \text{for} \quad n \geq N.$$

Therefore, (12.1) becomes equivalent to (12.13) and the corollary follows from the theorem.

COROLLARY 2 *Suppose that* $\mathrm{E}\,X_{nl}=0$, *and for some* $\rho>0$,

$$\int_{R_p} |x|^{2+\rho} \, d\,F_{nl}(x) < \infty$$

for all n and l, and that

$$\lim_n \frac{1}{s_n^{2+\rho}} \sum_l \int_{R_p} |x|^{2+\rho} \, d\,F_{nl}(x) = 0. \tag{12.14}$$

Let $\mathrm{N}(0, \sigma)$ *be a normalized normal distribution. Then the distribution of* $(X_{n1}+\cdots+X_{nl_n})/s_n$ *converges to the* $\mathrm{N}(0, \sigma)$ *and* X_{nl}/s_n, $1 \leq l \leq l_n$, *converge to* 0 *in probability uniformly in l, if and only if* (12.13) *holds.* (cf. W. Hoeffding and H. Robbins [10], Appendix).

PROOF: For each $\varepsilon>0$

$$\left|\frac{1}{s_n^2}\sum_l \int_{|x|\geq\varepsilon s_n} x_j x_k \, d\,F_{nl}(x)\right| \leq \frac{1}{s_n^2}\sum_l \int_{|x|\geq\varepsilon s_n} |x|^2 \, d\,F_{nl}(x)$$

$$\leq \frac{1}{\varepsilon^\rho s^{2+\rho}} \sum_l \int_{R_p} |x|^{2+\rho} \, d\,F_{nl}(x).$$

This together with (12.14) implies that for each $\varepsilon>0$

$$\lim_n \frac{1}{s_n^2}\sum_l \int_{|x|\geq\varepsilon s_n} x_j x_k \, d\,F_{nl}(x) = 0, \qquad j, k = 1, 2, \ldots, p.$$

Hence, (12.1) becomes equivalent to (12.13) and the corollary follows from the theorem.

(12.14) is called Liapounov's condition.

13. Generalization of P. Lévy's theorem

Let $\{X_n\}$ and $\{Y_n\}$ be sequences of p-dimensional random variables. If for any pair of sequences $\{a_n\}$ and $\{b_n\}$ the convergence of $\{(X_n-b_n)/a_n\}$ in distribution implies the convergence of $\{(Y_n-b_n)/a_n\}$ in distribution and conversely, then $\{X_n\}$ and $\{Y_n\}$ are said to be *equivalent with respect to the convergence in distribution.*

LEMMA 13.1 *If* $\lim_n \Pr\{X_n \neq Y_n\}=0$, *$\{X_n\}$ and $\{Y_n\}$ are equivalent with respect to the convergence in distribution.*

PROOF: This is obvious from the inequality

$$\sup_{x \in R_p} | F_n(a_n x+b_n)-G_n(a_n x+b_n) | \leq \Pr\{X_n \neq Y_n\},$$

where F_n and G_n are the distribution functions of X_n and Y_n, respectively.

LEMMA 13.2 *Let X and Y be p-dimensional random variables and let D_X and D_Y be the dispersion functions of X and Y, respectively. If* $\Pr(X \neq Y) \leq \delta < 1/(4p)$,

$$D_Y(\alpha-2p\delta) \leq D_X(\alpha) \leq D_Y(\alpha+2p\delta) \quad \textit{for} \quad 2p\delta<\alpha<1-2p\delta. \tag{13.1}$$

PROOF: (i) Case $p=1$. Let X' and Y' be other real random variables such that the two-dimensional random variable (X', Y') is independent of (X, Y) and has the same distribution as (X, Y) has. Let Ψ_X and Ψ_Y be the mean concentration functions of X and Y, respectively. Then by definition,

$$| \Psi_X(l)-\Psi_Y(l) | = \left| \mathrm{E}\left(\frac{l^2}{l^2+(X-X')^2}\right)-\mathrm{E}\left(\frac{l^2}{l^2+(Y-Y')^2}\right) \right|$$
$$\leq \mathrm{E}\left| \frac{l^2}{l^2+(X-X')^2}-\frac{l^2}{l^2+(Y-Y')^2} \right| \leq \Pr(X-X' \neq Y-Y')$$
$$\leq \Pr(X \neq Y)+\Pr(X' \neq Y') \leq 2\delta,$$

Fix an α such that $0<\alpha<\alpha+2\delta<1$, and put $D_X(\alpha)=l$. If $l>0$ then, since Ψ_X is continuous at the point l, $\Psi_X(l)=\alpha$ and

$$\Psi_Y(l) \leq \Psi_X(l)+2\delta=\alpha+2\delta,$$

hence, $l \leq D_Y(\alpha+2\delta)$. This holds also when $l=0$. Therefore

$$D_X(\alpha) \leq D_Y(\alpha+2\delta) \qquad \text{for} \quad 0<\alpha<\alpha+2\delta<1.$$

Interchanging X and Y, we have

$$D_Y(\alpha) \leq D_X(\alpha+2\delta), \qquad \text{for} \quad 0<\alpha<\alpha+2\delta<1.$$

Thus, (13.1) holds for $p=1$. (ii) Case: $p>1$. Let $X=(X_1, X_2, \ldots, X_p)$ and $Y=(Y_1, Y_2, \ldots, Y_p)$. Let $X'=(X_1', X_2', \ldots, X_p')$ and $Y'=(Y_1', Y_2', \ldots, Y_p')$ be other p-dimensional random variables such that $X_1', X_2', \ldots, X_p'$ are independent, $Y_1', Y_2', \ldots, Y_p'$ are independent and (X_j', Y_j') has the same distribution as $(X_j, Y_j$ has for each $j=1, 2, \ldots, p$. Put $X^*=X_1'+X_2'+\cdots+X_p'$ and $Y^*=Y_1'+Y_2'+\cdots+Y_p'$. Then

$$\Pr(X^* \neq Y^*) \leq \sum_{j=1}^{p} \Pr(X_j' \neq Y_j') = \sum_{j=1}^{p} \Pr(X_j \neq Y_j)$$
$$\leq p \Pr(X \neq Y) \leq p\delta.$$

Hence, by the one-dimensional case, (13.1) holds with X and Y replaced by X^* and Y^*, respectively, which implies that (13.1) holds since by definition $D_X=D_{X^*}$ and $D_Y=D_{Y^*}$.

LEMMA 13.3 *Let v and D be the variance and the dispersion function of a random variable X respectively. Then*

$$\sqrt{v} \geq \sqrt{(1-\alpha)/2}\; D(\alpha), \qquad 0<\alpha<1. \tag{13.2}$$

PROOF: It is sufficient to prove the lemma when X is one-dimensional. Denote by F the distribution function, and by Ψ the mean concentration function of X. Then

$$1-\Psi(l)=\int_{-\infty}^{\infty} \frac{x^2}{l^2+x^2}\, d\widetilde{F}(x) \leq \frac{1}{l^2}\int_{-\infty}^{\infty} x^2\, d\widetilde{F}(x)=\frac{2v}{l^2},$$

hence

$$\Psi(l) \geq 1-2v/l^2, \qquad 0<l<\infty. \tag{13.3}$$

Fix an α, $0<\alpha<1$, and put $l_1=D(\alpha)$. If $l_1>0$ then $\Psi(l_1)=\alpha$, hence from (13.3)

$$\alpha \geq 1-2v/D^2(\alpha)$$

which implies (13.2). If $l_1=0$, (13.2) is trivial.

THEOREM 13.1 *Assume that there exists an α such that $D_{S_n}(\alpha)>0$ for all n, and put $D_n=D_{S_n}(\alpha)$. Let $S_n=(S_{n1}, S_{n2}, \ldots, S_{np})$ and let $D(S_{nj})$ and $D(S_{nj}+S_{nk})$ be the α-dispersions of S_{nj} and $S_{nj}+S_{nk}$, respectively. Assume that*

$$\lim_n \max_l \int_{\|x\|>\varepsilon D_n} dF_{nl}(x)=0, \quad \textit{for each} \quad \varepsilon>0. \tag{13.4}$$

Then, for some sequence $\{a_n\}$ and $\{b_n\}$ the distribution of $(S_n-b_n)/a_n$ converges to a normal distribution, if and only if the following (13.5) *holds and the limits* (13.6) *and* (13.7) *exist:*

(13.5) $$\lim_n \sum_l \int_{\|x\|>\varepsilon D_n} dF_{nl}(x)=0, \quad \textit{for each} \quad \varepsilon>0,$$

(13.6) $$\lim_n \sum_l \frac{D(S_{nj})}{D_n}, \quad j=1, 2, \ldots, p,$$

(13.7) $$\lim_n \frac{D(S_{nj}+S_{nk})}{D_n}, \quad j, k=1, 2, \ldots, p.$$

The condition (13.4) represents that X_{nl}/D_n, $1\leq l\leq l_n$, converge to 0 in probability uniformly in l as $n\to\infty$, i.e., that each term of X_{nl} is asymptotically individually negligible with respect to the dispersion of the sum $S_n=X_{n1}+\cdots+X_{nl_n}$. Under (13.4), (13.5) is equivalent to the condition that $\max_l |X_{nl}|/D_n$ converges to 0 in probability as $n\to\infty$, i.e., that the greatest term of X_{nl}, $1\leq l\leq l_n$, is asymptotically negligible with respect to the dispersion of the sum (cf. section 11).

PROOF: Assume that there exist $\{a_n\}$ and $\{b_n\}$ such that the distribution of $(S_n-b_n)/a_n$ converges to a normal distribution. Then so do the distribution of $(S_n-b_n)/D_n$ by Corollary 2 to Theorem 5.6. Moreover, X_{nl}/D_n, $1\leq l\leq l_n$, converge to 0 in probability uniformly in l by (13.4). Therefore (13.5) holds from Corollary 2 to Theorem 11.1. Now write $b_n=(b_{n1}, b_{n2}, \ldots, b_{np})$. Then the distributions of real random variables

$$(S_{nj}-b_{nj})/D_n \quad \text{and} \quad \{(S_{nj}+S_{nk})-(b_{nj}+b_{nk})\}/D_n$$

converge as $n\to\infty$, hence, the limits (13.6) and (13.7) exist.

Conversely, assume (13.5) and the existence of the limits (13.6) and (13.7). Since (13.5) implies

$$\lim_n \inf_{\varepsilon>0} \{\varepsilon+\sum_l \int_{\|x\|>\varepsilon D_n} dF_{nl}(x)\}=0,$$

there exists a sequence $\{\varepsilon_n\}$ such that $\varepsilon_n\to 0$ and

(13.8) $$\sum_l \int_{\|x\|>\varepsilon_n D_n} dF_{nl}(x)=\delta_n \text{ (say)} \to 0.$$

Define X'_{nl} by

$$X'_{nl}(\omega)=\begin{cases} X_{nl}(\omega), & \text{if } \|X_{nl}(\omega)\|\leq \varepsilon_n D_n, \\ 0 \quad , & \text{if } \|X_{nl}(\omega)\|>\varepsilon_n D_n, \end{cases}$$

and put $S'_n=X'_{n1}+X'_{n2}+\cdots+X'_{nl_n}$. Then we have

(13.9) $$\Pr(S_n \neq S'_n)\leq \sum_l \Pr(X_{nl}\neq X'_{nl})=\delta_n\to 0,$$

hence, by Lemma 13.1, $\{S_n\}$ and $\{S'_n\}$ are equivalent with respect to the convergence in distribution. Let D'_n be the dispersion function of S'_n. Then from Lemma 13.2 we have

$$D'_n(\alpha+2p\delta_n) \geq D_n(\alpha)=D_n, \quad \text{if} \quad \alpha+2p\delta_n<1.$$

Fix an α' such that $\alpha<\alpha'<1$. For sufficiently large n, we have $\alpha'>\alpha+2p\delta_n$, hence

$$D'_n(\alpha') \geq D'_n(\alpha+2p\delta_n) \geq D_n.$$

Put $\sqrt{v(S'_n)}=s_n$. By Lemma 13.3

$$s_n \geq \sqrt{(1-\alpha')/2}\, D'_n(\alpha'),$$

hence,

$$s_n \geq \sqrt{(1-\alpha')/2}\, D_n, \tag{13.10}$$

for sufficiently large n. Moreover, put

$$M_n=\max_l \sup_\omega \| X'_{nl}(\omega)-\mathrm{E}\, X'_{nl} \|.$$

Then

$$M_n \leq 2\varepsilon_n D_n. \tag{13.11}$$

(13.10), (13.11) and lim $\varepsilon_n=0$ imply

$$\lim_n (M_n/s_n)=0.$$

From Corollary 1 to Theorem 12.1 any subsequence of the distributions of $(S'_n-\mathrm{E}\, S'_n)/s_n$ has a subsequence converging to a normal distribution with mean vector 0. Therefore, by (13.9) the case is the same with $(S_n-\mathrm{E}\, S'_n)/s_n$, and hence, also with $(S_n-b_n)/D_n$ by Corollary 2 to Theorem 5.6 where $b_n=\mathrm{E}\, S'_n$. We shall show that the limit distribution does not depend on subsequences. Assume that the distribution of $(S_{n(i)}-b_{n(i)})/D_{n(i)}$ converges to a normal distribution $\mathrm{N}(0, \sigma)$ as $n(i)\to\infty$. Then the distribution of $(S_{n(i)j}-b_{n(i)j})/D_{n(i)}$ converges to $\mathrm{N}(0, \sigma_{jj})$, hence by Lemma 4.4 the limit

$$\lim \frac{D(S_{n(i)j})}{D_{n(i)}}=d$$

exists, where d is the α-dispersion of $\mathrm{N}(0, \sigma_{jj})$. From the existence of (13.6) d, the α-dispersion of $\mathrm{N}(0, \sigma_{jj})$, does not depend on the subsequence. Hence, σ_{jj} is independent of the subsequence for each $j=1, 2,\dots, p$. Next, the distribution of $[(S_{n(i)j}+S_{n(i)k})-(b_{n(i)j}+b_{n(i)k})]/D_{n(i)}$ converges to $\mathrm{N}(0, \sigma_{jj}+2\sigma_{jk}+\sigma_{kk})$, this together with the existence of (13.7) implies that $\sigma_{jj}+2\sigma_{jk}+\sigma_{kk}$ does not depend on the subsequence, hence σ_{jk} does not depend on the subsequence. After all, σ is independent of the subsequence, and the distribution of $(S_n-b_n)/D_n$ converges to the normal distribution $N(0, \sigma)$.

COROLLARY *Assume that each term of* X_{nl}, $1\le l\le l_n$, *is asymptotically individually negligible with respect to the dispersion of the sum* $S_n=X_{n1}+\cdots+X_{nl_n}$ *and that the class of the distribution of* S_n *converges to a class. Then the limiting class is normal if and only if the greatest term of* X_{nl}, $1\le l\le l_n$, *is asymptotically negligible with respect to the dispersion of* S_n.

14. Feller's criterion and the attraction domains of normal distributions

In this section we want to generalize some results of W. Feller [7] to the multi-dimensional case.

THEOREM 14.1 *Suppose that* 0 *is a median vector for each* X_{nl}. *Then* $\{X_{nl}\}$ *obeys the central limit theorem if and only if there exists a sequence of positive numbers* $\{\lambda_n\}$ *such that*

$$\lim_n \sum_l \int_{|x|>\lambda_n} dF_{nl}(x)=0, \tag{14.1}$$

$$\lim_n \frac{1}{\lambda_n^2}\sum_l \int_{|x|\le\lambda_n} |x|^2\, dF_{nl}(x)=\infty \tag{14.2}$$

and the limits

$$\lim_n \frac{1}{a_n^2}\sum_l\Big(\int_{|x|\le\lambda_n} x_j x_k\, dF_{nl}(x)-\int_{|x|\le\lambda_n} x_j\, dF_{nl}(x)\int_{|x|\le\lambda_n} x_k\, dF_{nl}(x)\Big)=\sigma_{jk}\ \text{(say)},\quad j,k=1,2,\ldots,p. \tag{14.3}$$

exist, where

$$a_n^2=\sum_l\Big\{\int_{|x|\le\lambda_n}|x|^2\, dF_{nl}(x)-\Big|\int_{|x|\le\lambda_n} x\, dF_{nl}(x)\Big|^2\Big\}. \tag{14.4}$$

In this case the distribution of $(S_n-b_n)/a_n$ *converges to* $N(0,\sigma)$ *where* $\sigma=(\sigma_{jk})$ *is defined by* (14.3) b_n *is defined by*

$$b_n=\sum_l\int_{|x|\le\lambda_n} x\, dF_{nl}(x). \tag{14.5}$$

Note that (14.3) *and* (14.4) *imply that*

$$\sum_j \sigma_{jj}=1. \tag{14.6}$$

The theorem holds even if '>' *and* '≤' *are simultaneously replaced by* '≥' *and* '<', *respectively.*

PROOF: To prove the 'if' part, assume (14.1)–(14.3) and define $a_n>0$

and b_n by (14.4) and (14.5), respectively. By the Schwarz inequality and the fact that 0 is a median we have

$$\Big(\int_{|x|\le\lambda_n,\ x_j>0} x_j\,dF_{nl}(x)\Big)^2 \le \int_{|x|\le\lambda_n} x_j^2\,dF_{nl}(x)\int_{x_j>0} dF_{nl}(x) \le \frac{1}{2}\int_{|x|\le\lambda_n} x_j^2\,dF_{nl}(x).$$

Similary

$$\Big(\int_{|x|\le\lambda_n,\ x_j<0} x_j\,dF_{nl}(x)\Big)^2 \le \frac{1}{2}\int_{|x|\le\lambda_n} x_j^2\,dF_{nl}(x).$$

Therefore, since

$$\int_{|x|\le\lambda_n,\ x_j>0} x_j\,dF_{nl}(x)\ge 0 \quad\text{and}\quad \int_{|x|<\lambda_n,\ x_j<0} x_j\,dF_{nl}(x)\le 0,$$

we get

$$\Big(\int_{|x|\le\lambda_n} x_j\,dF_{nl}(x)\Big)^2 \le \frac{1}{2}\int_{|x|\le\lambda_n} x_j^2\,dF_{nl}(x).$$

Summing over j we obtain

$$\Big|\int_{|x|\le\lambda_n} x\,dF_{nl}(x)\Big|^2 \le \frac{1}{2}\int_{|x|\le\lambda_n} |x|^2\,dF_{nl}(x),$$

this together with (14.4) implies

$$a_n^2 \ge \frac{1}{2}\sum_l\int_{|x|\le\lambda_n} |x|^2\,dF_{nl}(x). \tag{14.7}$$

From (14.2) and (14.7) it follows that

$$\lim_n \frac{a_n}{\lambda_n}=\infty. \tag{14.8}$$

Define X'_{nl} by

$$X'_{nl}(\omega)=\begin{cases} X_{nl}(\omega), & \text{if } |X_{nl}(\omega)|\le\lambda_n,\\ 0\quad, & \text{otherwise,}\end{cases}$$

and put

$$S'_n=X'_{n1}+X'_{n2}+\cdots+X'_{nl_n},$$
$$M_n=\max_l\sup_\omega|X'_{nl}(\omega)-\mathrm{E}X'_{nl}|.$$

Then $X'_{n1}, X'_{n2}, \dots, X'_{nl_n}$ are independent for each n,

$$M_n\le 2\lambda_n,\quad b_n=\mathrm{E}S'_n,\quad a_n^2=v(S'_n), \tag{14.9}$$

and (14.3) means that the covariance matrix of S'_n/a_n converges to σ. From (14.8) and (14.9) we have

$$\lim_{n\to\infty}\frac{M_n}{a_n}=0. \tag{14.10}$$

Then from Corollary 1 to Theorem 12.1 the distribution of $(S'_n-b_n)/a_n$ converges to $N(0, \sigma)$. Moreover

$$\Pr(S_n \neq S'_n) \leq \sum_l \Pr(X_{nl} \neq X'_{nl}) = \sum_l \Pr(|X_{nl}| > \lambda_n)$$
$$= \sum_l \int_{|x|>\lambda_n} dF_{nl}(x),$$

this together with (14.1) implies that

$$\lim \Pr(S_n \neq S'_n) = 0.$$

Therefore, from Lemma 13.1, the distribution of $(S_n-b_n)/a_n$ converges to $N(0, \sigma)$. The asymptotic uniform negligibility of X_{nl}/a_n follows from (14.1) and (14.8). This completes the proof of the 'if' part.

To prove the 'only if' part, assume that $\{X_{nl}\}$ obeys the central limit theorem. Then from Corollary 1 to Theorem 11.1 we have for some sequence of positive numbers $\{\alpha_n\}$ and for each $\varepsilon > 0$

$$\text{(14.11)} \qquad \lim_n \sum_l \int_{|x|>\varepsilon\alpha_n} dF_{nl}(x) = 0,$$

$$\text{(14.12)} \qquad \lim_n \frac{1}{\alpha_n^2} \sum_l \left(\int_{|x|\leq\varepsilon\alpha_n} x_j x_k \, dF_{nl}(x) - \int_{|x|\leq\varepsilon\alpha_n} x_j \, dF_{nl}(x) \int_{|x|\leq\varepsilon\alpha_n} x_k \, dF_{nl}(x) \right) = \sigma_{jk},$$
$$j, k = 1, 2, \ldots, p.$$

We may assume that $\sum_j \sigma_{jj} = 1$ without loss of generality by Theorem 5.2. From (14.11) and (14.12), for each positive integer m there exists an n_m such that, for all $n \geq n_m$,

$$\sum_l \int_{|x|>\alpha_n/m} dF_{nl}(x) < \frac{1}{m},$$

$$\left| \frac{1}{\alpha_n^2} \sum_l \left(\int_{|x|\leq\alpha_n/m} x_j x_k \, dF_{nl} - \int_{|x|\leq\alpha_n/m} x_j \, dF_{nl} \int_{|x|\leq\alpha_n/m} x_k \, dF_{nl} \right) - \sigma_{jk} \right| < \frac{1}{m}, \quad j, k = 1, 2, \ldots, p.$$

It may be assumed that

$$n_1 < n_2 < n_3 < \cdots .$$

Define λ_n, for $n \geq n_1$, as follows:

$$\text{(14.13)} \qquad \lambda_n = \frac{\alpha_n}{m} \qquad \text{for} \quad n_m \leq n < n_{m+1}.$$

Then if $n \geq n_m$,

$$\sum_l \int_{|x|>\lambda_n} dF_{nl}(x) < \frac{1}{m},$$

$$\left|\frac{1}{a_n^2}\sum_l\left(\int_{|x|\le\lambda_n} x_j x_k\,dF_{nl}-\int_{|x|\le\lambda_n} x_j\,dF_{nl}\int_{|x|\le\lambda_n} x_k\,dF_{nl}\right)-\sigma_{jk}\right|<\frac{1}{m}\,.$$

Hence, we have (14.1) and

$$\lim\frac{1}{\alpha_n^2}\sum_l\left(\int_{|x|\le\lambda_n} x_j x_k\,dF_{nl}-\int_{|x|\le\lambda_n} x_j\,dF_{nl}\int_{|x|\le\lambda_n} x_k\,dF_{nl}\right)=\sigma_{jk},\qquad j,k=1,2,\ldots,p. \tag{14.14}$$

Sum (14.14) with $j=k$ from $j=1$ to $j=p$. Then $\sum\sigma_{jj}=1$ being assumed,

$$\lim_n\frac{a_n^2}{\alpha_n^2}=1, \tag{14.15}$$

where a_n^2 is defined by (14.4). From (14.13) $\lim(\alpha_n/\lambda_n)=\infty$, this together with (14.15) implies $\lim(a_n^2/\lambda_n^2)=\infty$, hence we have (14.2). From (14.14) and (14.15) it follows that (14.3) holds. Thus we have deduced (14.1)–(14.3).

It is possible to derive the 'if' part from Corollary 1 to Theorem 11.1. The proof proceeds as follows. Assume (14.1)–(14.3) and define $a_n>0$ and b_n by (14.4) and (14.5), respectively. It is sufficient to prove (11.13)–(11.15) with $V=\{x;\ |x|<1\}$ from these assumptions. From (14.1) and (14.8) it follows that (11.13) holds for each $\varepsilon>0$. From (14.8), $a_n>\lambda_n$ for sufficiently large n. Write $b_n=(b_{n1},b_{n2},\ldots,b_{np})$ where b_n is defined by (14.5). Then, assuming $a_n>\lambda_n$, we have

$$\left|\frac{1}{a_n}\left(\sum_l\int_{|x|<a_n} x_j\,dF_{nl}(x)-b_{nj}\right)\right|=\left|\frac{1}{a_n}\sum_l\int_{\lambda_n<|x|<a_n} x_j\,dF_{nl}(x)\right|$$
$$\le\sum_l\int_{|x|>\lambda_n} dF_{nl}(x),$$

this together with (14.1) implies (11.14) with $V=\{x;|x|<1\}$. Now for n such that $a_n>\lambda_n$

$$\left|\frac{1}{a_n^2}\sum_l\left(\int_{|x|<a_n} x_j x_k\,dF_{nl}(x)-\int_{|x|\le\lambda_n} x_j x_k\,dF_{nl}(x)\right)\right|$$
$$\le\frac{1}{a_n^2}\sum_l\int_{\lambda_n<|x|<a_n}|x_j x_k|\,dF_{nl}(x)\le\sum_l\int_{|x|>\lambda_n} dF_{nl}(x),$$

and

$$\left|\frac{1}{a_n^2}\sum_l\left(\int_{|x|<a_n} x_j\,dF_{nl}\int_{|x|<a_n} x_k\,dF_{nl}-\int_{|x|\le\lambda_n} x_j\,dF_{nl}\int_{|x|\le\lambda_n} x_k\,dF_{nl}\right)\right|$$
$$=\left|\frac{1}{a_n^2}\sum_l\left(\int_{\lambda_n<|x|<a_n} x_j\,dF_{nl}(x)\int_{|x|<a_n} x_k\,dF_{nl}(x)+\int_{|x|\le\lambda_n} x_j\,dF_{nl}(x)\int_{\lambda_n<|x|<a_n} x_k\,dF_{nl}(x)\right)\right|$$

$$\leq 2\sum_{l}\int_{|x|>\lambda_n} dF_{nl}(x),$$

these together with (14.1) imply

$$\text{(14.16)}\qquad \lim\Big[\frac{1}{a_n^2}\sum_{l}\Big(\int_{|x|<a_n} x_jx_k\,dF_{nl}-\int_{|x|<a_n} x_j\,dF_{nl}\int_{|x|<a_n} x_k\,dF_{nl}\Big)$$
$$-\frac{1}{a_n^2}\sum_{l}\Big(\int_{|x|\leq\lambda_n} x_jx_k\,dF_{nl}-\int_{|x|\leq\lambda_n} x_j\,dF_{nl}\int_{|x|\leq\lambda_n} x_k\,dF_{nl}\Big)\Big]=0.$$

(11.15) with $V=\{x;\ |x|<1\}$ follows from (14.16) and (14.3). Thus the proof is completed.

We shall need the following

LEMMA 14.1 *Let $X_1, X_2,\ldots$ be independent random variables with the same distribution function F and assume that for some sequences $\{a_n\}$ and $\{b_n\}$, the distribution $(X_1+\cdots+X_n-b_n)/a_n$ converges to a normal distribution* $\mathrm{N}(0, \sigma)$. *Then we have*

$$\text{(14.17)}\qquad \lim a_n=\infty,\qquad \lim(a_{n+1}/a_n)=1.$$

PROOF: Take an α such that $D_F(\alpha)>0$ and denote by D_n the α-dispersion of $X_1+\cdots+X_n$ $(n=1, 2, \ldots)$. Then we have

$$\text{(14.18)}\qquad 0<D_1<D_2<\cdots,$$

$$\text{(14.19)}\qquad \lim(D_n/a_n)=D,\qquad 0<D<\infty,$$

where D is the α-dispersion of $\mathrm{N}(0\ \sigma)$. From (14.18) it follows that $\lim D_n=\infty$ or $0<\lim D_n<\infty$ exists. If the latter occurs then from (14.19) finite $\lim a_n=a$ exists, $0<a<\infty$, and from Theorem 5.2 the distribution of $X_1+\cdots+X_n-b_n$ converges to $\mathrm{N}(0, a^2\sigma)$. Let φ and ψ be the characteristic functions of F and $\mathrm{N}(0, a^2\sigma)$. Then we have

$$\lim_{n\to\infty}\varphi^n(t)\,e^{-ib_n't}=\psi(t),$$

hence,

$$\lim|\varphi(t)|^n=|\psi(t)|=\begin{cases}1, & \text{if } |\varphi(t)|=1,\\ 0, & \text{otherwise,}\end{cases}$$

which contradicts the fact that $\psi(t)=\exp\Big[-\frac{a^2}{2}t'\sigma t\Big]$. Therefore it must hold that $\lim D_n=\infty$, and this together with (14.19) implies that $\lim a_n=\infty$. Now, since the distribution of $(X_1+\cdots+X_{n+1}-b_{n+1})/a_{n+1}$ converges to $\mathrm{N}(0, \sigma)$ and X_{n+1}/a_{n+1} converges to 0 in probability, the distribution of $(X_1+\cdots+X_n-b_{n+1})/a_{n+1}$ converges to $\mathrm{N}(0, \sigma)$ (see H. Cramér [2], p. 254). As the distribution of $(X_1+\cdots+X_n-b_n)/a_n$ converges also to $\mathrm{N}(0, \sigma)$ it follows that $\lim(a_{n+1}/a_n)=1$ from Theorem 5.4.

When the hypothesis of Lemma 14.1 holds, F is said to *belong to the attraction domain of the normal distribution* $\mathrm{N}(0, \sigma)$.

THEOREM 14.2 *Let F be a distribution function with the vanishing median vector and let $\mathrm{N}(0, \sigma)$ be a normalized normal distribution, $\sum_j \sigma_{jj}=1$. Then F belongs to the attraction domain of $\mathrm{N}(0, \sigma)$ if and only if the following conditions hold:*

$$\lim_{u\to\infty} \frac{u^2 \int_{|x|>u} dF(x)}{\int_{|x|\leq u} |x|^2 dF(x)}=0, \tag{14.20}$$

$$\lim_{u\to\infty} \frac{\sigma_{jk}(u)}{v(u)}=\sigma_{jk}, \qquad j, k=1, 2, \ldots, p, \tag{14.21}$$

where

$$\sigma_{jk}(u)=\int_{|x|\leq u} x_j x_k \, dF(x)-\int_{|x|\leq u} x_j \, dF(x) \int_{|x|\leq u} x_k \, dF(x), \tag{14.22}$$

$$v(u)=\int_{|x|\leq u} |x|^2 dF(x)-\left|\int_{|x|\leq u} x \, dF(x)\right|^2. \tag{14.23}$$

Let X be any random variable with distribution function F, and define X_u by

$$X_u(\omega)=\begin{cases} X(\omega), & \text{if } |X(\omega)|\leq u, \\ 0 , & \text{otherwise}. \end{cases}$$

Then $\sigma_{jk}(u)$, $j, k=1, 2, \ldots, p$, and $v(u)$ are the second order central moments and the variance of X_u, and $\sigma_{jk}(u)/v(u)$ are the second order central moments of the normalized variable $(X_u-\mathrm{E}\, X_u)/\sqrt{v(X_u)}$.

PROOF OF THEOREM 14.2 To prove the 'if' part, it is sufficient, by Theorem 14.1, to deduce from (14.20) the existence of a sequence of positive numbers $\{\lambda_n\}$ such that

$$\begin{cases} \lim n \int_{|x|>\lambda_n} dF(x)=0, \\ \lim \frac{n}{\lambda_n^2} \int_{|x|\leq\lambda_n} |x|^2 dF(x)=\infty, \\ \lim \lambda_n=\infty. \end{cases} \tag{14.24}$$

If for some u_0

$$\int_{|x|>u_0} dF(x)=0,$$

then a sequence $\{\lambda_n\}$ satisfying (14.24) is obtained by

$$\lambda_n = n^{1/4}.$$

Therefore we may assume that for all $u > 0$

$$\int_{|x|>u} dF(x) > 0.$$

For each positive integer m, define μ_m by

$$\mu_m = \min\left\{\mu ; \int_{|x|>\mu} dF(x) \leq \frac{1}{m}\right\}.$$

Then we have

$$\lim \mu_m = \infty, \tag{14.25}$$

$$\frac{\int_{|x|\leq\mu_m} |x|^2 dF(x)}{\left(\frac{\mu_m}{2}\right)^2 \cdot \frac{1}{m}} \geq \frac{\int_{|x|\leq\mu_m/2} |x|^2 dF(x)}{\left(\frac{\mu_m}{2}\right)^2 \int_{|x|>\mu_m/2} dF(x)}. \tag{14.26}$$

From (14.20), (14.25) and (14.26) we have

$$\lim_{m\to\infty} \frac{m}{\mu_m^2} \int_{|x|\leq\mu_m} |x|^2 dF(x) = \infty.$$

Hence we can choose a sequence of positive numbers $n_1 < n_2 < n_3 < \cdots$ such that for each $p = 1, 2, \ldots,$

$$\frac{np}{\mu_{np}^2} \int_{|x|\leq\mu_{np}} |x|^2 dF(x) \geq p^2, \qquad \text{for} \quad n \geq n_p.$$

Define λ_n by

$$\lambda_n = \mu_{np} \qquad \text{for} \quad n_p \leq n < n_{p+1}.$$

Then we have

$$\lim \lambda_n = \infty,$$

$$\left.\begin{aligned} &\frac{n}{\lambda_n^2} \int_{|x|\leq\lambda_n} |x|^2 dF(x) \geq p \\ &n \int_{|x|>\lambda_n} dF(x) \leq \frac{1}{p} \end{aligned}\right\} \quad \text{for} \quad n_p \leq n < n_{p+1},$$

hence, (14.24) holds

Conversely, assume that F belongs to the attraction domain of $N(0, \sigma)$ and let $\{a_n\}$ be a sequence which satisfies the condition of Lemma 14.1. Then from Corollary 1 to Theorem 11.1 we have

(14.27) $$\lim_{n\to\infty} n\int_{|x|>a_n} dF(x)=0,$$

(14.28) $$\lim_{n\to\infty}\frac{n\sigma_{jk}(a_n)}{a_n^2}=\sigma_{jk}, \qquad j, k=1, 2, \ldots, p.$$

From (14.28) we have

(14.29) $$\lim \frac{nv(a_n)}{a_n^2}=\sum_j \sigma_{jj}=1.$$

This together with (14.27) implies

$$\lim_{n\to\infty}\frac{a_n^2\int_{|x|>a_n} dF(x)}{v(a_n)}=0,$$

hence,

(14.30) $$\lim\frac{a_n^2\int_{|x|>a_n} dF(x)}{\int_{|x|\leq a_n}|x|^2\,dF(x)}=0.$$

From (14.17) there exists N such that

$$a_{n+1}<2a_n \qquad \text{for all} \quad n\geq N.$$

We may assume that $a_n=D_n/D$ (see the proof of Lemma 14.1), hence that $a_n<a_{n+1}$ for all n. Let $\{u_n\}$ be an arbitrary sequence such that

(14.31) $$a_n\leq u_n\leq a_{n+1}.$$

Then for $n\geq N$, as $a_n\leq u_n<2a_n$, we have

$$\frac{u_n^2\int_{|x|>u_n} dF(x)}{\int_{|x|\leq u_n}|x|^2\,dF(x)}\leq\frac{4a_n^2\int_{|x|>a_n} dF(x)}{\int_{|x|\leq a_n}|x|^2\,dF(x)},$$

this together with (14.30) implies

$$\lim_{n\to\infty}\frac{u_n^2\int_{|x|>u_n} dF(x)}{\int_{|x|\leq u_n}|x|^2\,dF(x)}=0.$$

Since u_n is arbitrary except for (14.31) and since $\lim a_n=\infty$, we have (14.20). From (14.28) and (14.29) we have

(14.32) $$\lim_{n\to\infty}\frac{\sigma_{jk}(a_n)}{v(a_n)}=\sigma_{jk}, \qquad j, k=1, 2, \ldots, p.$$

From (14.29) it follows that the sequence

$$\left\{\frac{a_n^2}{nv(a_n)};\ n=1, 2, \dots\right\} \text{ is bounded.} \tag{14.33}$$

From (14.27), (14.33) and $a_n \leq u_n < 2a_n$ it can be proved that

$$\lim_n \left(\frac{\sigma_{jk}(u_n)}{v(a_n)} - \frac{\sigma_{jk}(a_n)}{v(a_n)}\right)=0. \tag{14.34}$$

(See the proof of (14.16)). (14.32) and (14.34) imply

$$\lim \frac{\sigma_{jk}(u_n)}{v(a_n)}=\sigma_{jk}, \qquad j, k=1, 2, \dots, p.$$

Sum these with $j=k$ from $j=1$ to $j=p$, then

$$\lim \frac{v(u_n)}{v(a_n)}=1.$$

From the last two equations it follows that

$$\lim \frac{\sigma_{jk}(u_n)}{v(u_n)}=\sigma_{jk}, \qquad j, k=1, 2, \dots, p,$$

hence, (14.21) holds.

THEOREM 14.3 *If F belongs to the attraction domain of a normalized normal distribution* $\mathrm{N}(0, \sigma)$, *then for each α such that $0<\alpha<2$, the α-th absolute moment of F is finite:*

$$\int_{R_p} |x|^\alpha dF(x)<\infty,$$

hence, the mean vector of F

$$m=\int_{R_p} x\, dF(x)$$

is well defined and the distribution of

$$(X_1+\cdots+X_n-nm)/a_n$$

converges to $\mathrm{N}(0, \sigma)$ *for a sequence* $\{a_n\}$.

The proof runs in the same way as in the one-dimensional case.

15. Reduction to the one-dimensional case

By now, in this part, the various versions of the multi-dimensional central limit theorem have been studied from Theorem 11.1, which was proved from the general convergence theorem on the infinitely divisible multi-dimensional distributions. In this section we wish to show that Theorem 11.1 in the multi-dimensional case can be reduced to that in the one-dimensional case.

First let us notice the following fact.

LEMMA 15.1 *Let V_1 and V_2 be two neighborhoods of the origin. Then under* (11.2), (11.3) *and* (11.4) *with $V=V_1$ are equivalent to* (11.3) *and* (11.4) *with $V=V_2$, respectively.*

PROOF: This follows from the following facts:

$$\lim_n \Big(\sum_l \int_{V_1} x\, dF_{nl}(x) - \sum_l \int_{V_2} x\, dF_{nl}(x)\Big)=0, \tag{15.1}$$

$$\lim_n \Big[\sum_l\Big(\int_{V_1} x_j x_k\, dF_{nl}(x) - \int_{V_1} x_j\, dF_{nl}(x) \int_{V_1} x_k\, dF_{nl}(x)\Big) - \sum_l\Big(\int_{V_2} x_j x_k\, dF_{nl}(x) - \int_{V_2} x_j\, dF_{nl}(x) \int_{V_2} x_k\, dF_{nl}(x)\Big)\Big]=0, \quad j, k=1, 2, \dots, p. \tag{15.2}$$

To prove these, note that there exist positive numbers ε and δ such that

$$S(0, \varepsilon)\subset V_j \subset S(0, \delta) \qquad \text{for } j=1, 2.$$

Now

$$\begin{aligned}&\Big|\sum_l\Big(\int_{V_1} x_j\, dF_{nl}(x) - \int_{V_2} x_j\, dF_{nl}(x)\Big)\Big| \\ &\leq \sum_l\Big(\int_{V_1\cap V_2^c} |x_j|\, dF_{nl}(x) + \int_{V_2\cap V_1^c} |x_j|\, dF_{nl}(x)\Big) \\ &\leq 2\delta \sum_l \int_{|x|\geq\varepsilon} dF_{nl}(x), \qquad j=1, 2, \dots, p,\end{aligned}$$

and this together with (11.2) implies (15.1). Similary

$$\Big|\sum_l\Big(\int_{V_1} x_j x_k\, dF_{nl}(x) - \int_{V_2} x_j x_k\, dF_{nl}(x)\Big)\Big| \leq \delta^2 \sum_l \int_{|x|\geq\varepsilon} dF_{nl}(x),$$

$$\begin{aligned}&\Big|\sum_l\Big(\int_{V_1} x_j\, dF_{nl}(x) \int_{V_1} x_k\, dF_{nl}(x) - \int_{V_2} x_j\, dF_{nl}(x) \int_{V_2} x_k\, dF_{nl}(x)\Big)\Big| \\ &=\Big|\sum_l\Big[\Big(\int_{V_1} x_j\, dF_{nl} - \int_{V_2} x_j\, dF_{nl}\Big)\int_{V_1} x_k\, dF_{nl} + \int_{V_2} x_j\, dF_{nl}\Big(\int_{V_1} x_k\, dF_{nl} - \int_{V_2} x_k\, dF_{nl}\Big)\Big]\Big| \\ &\leq 4\delta^2 \sum_l \int_{|x|\geq\varepsilon} dF_{nl}(x),\end{aligned}$$

and these together with (11.2) imply (15.2).

The following lemma is useful for the reduction of the multi-dimensional central limit theorem to the one-dimensional case.

LEMMA 15.2 *Let* $\{X_n\}$ *be a sequence of p-dimensional random variables. Then the distribution of* X_n *converges to a normal distribution* $N(0, \sigma)$, *if and only if for each* $t \in R_p$ *the distribution of* $t'X_n$ *converges to the distribution* $N(0, t'\sigma t)$.

This is well-known and is easily proved by making use of the characteristic functions. Notice that if $t'\sigma t=0$, $N(0, t'\sigma t)$ denotes the unit distribution which has the whole probability 1 placed in the origin.

Now, in the one-dimensional case, Theorem 11.1 becomes

THEOREM 11.1′ *Let* $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nl_n}$ *be independent real random variables with distribution functions* $H_{n1}, H_{n2}, \ldots, H_{nl_n}$ *for each positive integer* n, *let* $\{\beta_n\}$ *be a sequence of real numbers, and let* v *be a non-negative real numbers. Furthermore assume that*

$$\lim_n \max_l \int_{|x|\geq \varepsilon} d\,H_{nl}(x)=0, \qquad \textit{for each} \quad \varepsilon>0.$$

Then the distribution of $\xi_{n1}+\xi_{n2}+\cdots+\xi_{nl_n}-\beta_n$ *converges to the distribution* $N(0, v)$, *if and only if the following three conditions hold:*

$$\lim_n \sum_l \int_{|x|\geq \varepsilon} d\,H_{nl}(x)=0, \qquad \textit{for each} \quad \varepsilon>0,$$

$$\lim_n \Big(\sum_l \int_{|x|<1} x\,d\,H_{nl}(x)-\beta_n\Big)=0,$$

$$\lim_n \sum_l \Big(\int_{|x|<1} x^2\,d\,H_{nl}(x)-\Big(\int_{|x|<1} x\,d\,H_{nl}(x)\Big)^2\Big)=v.$$

If $v>0$, this theorem is a slightly modified form of a well-known version of the one-dimensional central limit theorem (see W. Feller [6], Satz 1), for the distribution of $\xi_{n1}+\cdots+\xi_{nl_n}-\beta_n$ converges to $N(0, v)$ if and only if the distribution of $(\xi_{n1}+\cdots+\xi_{nl_n}-\beta_n)/\sqrt{v}$ converges to $N(0, 1)$. If $v=0$ Theorem 11.1′ becomes a version of the law of large numbers.

LEMMA 15.3 *If* (11.1) *holds, then for each* $t \in R_p$ *and for each* $\varepsilon>0$

$$\lim_n \max_l \int_{|t'x|\geq \varepsilon} d\,F_{nl}(x)=0.$$

This is obvious from

$$\int_{|t'x|\geq \varepsilon} d\,F_{nl}(x) \leq \int_{|x|\geq \varepsilon/|t|} d\,F_{nl}(x), \qquad \text{for} \quad t \neq 0,$$

which follows from $|t'x| \leq |t| \cdot |x|$.

Now we can reduce Theorem 11.1 to Theorem 11.1′. This follows from Lemma 15.2, 15.3, and the following

LEMMA 15.4

$$\lim_n \sum_l \int_{|t'x|\geq \varepsilon} dF_{nl}(x) = 0, \qquad \textit{for each} \quad \varepsilon > 0, \tag{15.3}$$

$$\lim_n \Big(\sum_l \int_{|t'x|<1} t'x\, dF_{nl}(x) - t'b_n \Big) = 0. \tag{15.4}$$

$$\lim_n \sum_l \Big[\int_{|t'x|<1} (t'x)^2\, dF_{nl}(x) - \Big(\int_{|t'x|<1} t'x\, dF_{nl}(x) \Big)^2 \Big] = t'\sigma t, \tag{15.5}$$

hold for each $t \in R_p$, *if and only if* (11.2)–(11.4) *hold.*

To prove the 'if' part assume (11.2)–(11.4) and fix a $t \neq 0$ (if $t=0$, (15.3)–(15.5) are trivial). Then (15.3) follows from (11.2) and the following inequality:

$$\sum_l \int_{|t'x|\geq \varepsilon} dF_{nl}(x) \leq \sum_l \int_{|x|\geq \varepsilon/|t|} dF_{nl}(x).$$

We may assume that $V = \{x;\, |x| < 1/|t|\}$. Now

$$\begin{aligned} &\sum_l \int_{|t'x|<1} t'x\, dF_{nl}(x) - t'b_n \\ &= \Big(\sum_l \int_{|x|<1/|t|} t'x\, dF_{nl}(x) - t'b_n \Big) + \sum_l \int_{|t'x|<1,\ |x|\geq 1/|t|} t'x\, dF_{nl}(x) \\ &= t' \Big(\sum_l \int_{|x|<1/|t|} x\, dF_{nl}(x) - b_n \Big) + \theta \sum_l \int_{|x|\geq 1/|t|} dF_{nl}(x), \end{aligned}$$

where $|\theta| \leq 1$, and this together with (11.3) and (11.2) implies (15.4). Furthermore

$$\begin{aligned} &\sum_l \Big\{ \int_{|t'x|<1} (t'x)^2\, dF_{nl}(x) - \Big(\int_{|t'x|<1} t'x\, dF_{nl}(x) \Big)^2 \Big\} \\ &= \sum_l \Big\{ \int_{|x|<1/|t|} (t'x)^2\, dF_{nl}(x) - \Big(\int_{|x|<1/|t|} t'x\, dF_{nl}(x) \Big)^2 \Big\} \\ &\quad + \sum_l \Big\{ \int_{|t'x|<1,\ |x|\geq 1/|t|} (t'x)^2\, dF_{nl}(x) - \Big(\int_{|t'x|<1,\ |x|\geq 1/|t|} t'x\, dF_{nl}(x) \Big)^2 - 2 \int_{|x|<1/|t|} t'x\, dF_{nl}(x) \int_{|t'x|<1,\ |x|\geq 1/|t|} t'x\, dF_{nl}(x) \Big\} \\ &= T_1 + T_2 \ \text{(say)}. \end{aligned}$$

T_1 tends to $t'\sigma t$ by (11.4), and T_2 tends to 0 as

$$|T_2| \leq 3\sum_l \int_{|x|\geq 1/|t|} dF_{nl}(x).$$

Therefore (15.5) holds.

To prove the 'only if' part, assume that (15.3)–(15.5) hold for each $t \in R_p$. Then (11.2) follows from (15.3) and the following inequality:

$$\sum_l \int_{|x|\geq \varepsilon} dF_{nl}(x) \leq \sum_j \sum_l \int_{|x_j|\geq \varepsilon/\sqrt{p}} dF_{nl}(x).$$

Write $b_n=(b_{n1}, b_{n2}, \ldots, b_{np})$. Then for each j, we have

$$\begin{aligned}
&\sum_l \int_{|x|<1} x_j\, dF_{nl}(x)-b_{nj} \\
&=\Big(\sum_l \int_{|x_j|<1} x_j\, dF_{nl}(x)-b_{nj}\Big)-\sum_l \int_{|x_j|<1,\ |x|\geq 1} x_j dF_{nl}(x) \\
&=\Big(\qquad\qquad ,, \qquad\qquad\Big)+\theta\sum_l \int_{|x|\geq 1} dF_{nl}(x) \\
&=U_1+U_2 \ \text{(say)}.
\end{aligned}$$

U_1 tends to 0 by (15.4), U_2 tends to 0 by (11.2), and hence (11.3) with $V=\{x\,;\,|x|<1\}$ holds. Furthermore from (11.2) and (15.5) it follows that

$$\lim_n \sum_l \left\{\int_{|x|<1} (t'x)^2\, dF_{nl}(x)-\Big(\int_{|x|<1} t'x\, dF_{nl}(x)\Big)^2\right\}=t'\sigma t$$

(see the proof of Lemma 15.1). Since this holds for each t, (11.4) with $V=\{x: |x|<1\}$ must hold.

REFERENCES

[1] H. Cramér, *Random variables and probability distributions*, Cambridge Tracts in Math. no. 36 (1937).
[2] H. Cramér, *Mathematical methods of statistics*, Princeton Univ. Press (1946).
[3] W. Doeblin, Sur l'ensemble de puissances d'une loi de probabilité, *Studia Math.* 9, 71–96 (1940).
[4] J. L. Doob, *Stochastic Processes*, John Wiley and Sons (1953).
[5] C. G. Essen, Fourier analysis of distribution functions, *Acta Math.* 77, 1–125 (1945).
[6] W. Feller, Über den zentralen Grentzwertzatz der Wahrscheinlichkeitzrechnung, *Math. Zeit.* 40, 521–559 (1935).
[7] W. Feller, Über den zentralen Grentzwertzatz der Wahrscheinlichkeitzrechnung II, *Math. Zeit.* 42, 301–312 (1937).
[8] W. Feller, The fundamental limit theorems in probability, *Bull. Amer. Math. Soc.* 51, 800–832 (1945).

[9] B. Gnedenko, Sur la distribution limit du terme maximum d'une serie aléatoire, *Ann. Math.* 44, 423–453 (1943).

[10] W. Hoeffding and H. Robbins, The central limit theorem for dependent random variables, *Duke Math. Jour.* 15, 773–780 (1948).

[11] Kiyosi Ito, *The theory of probability* (in Japanese), Iwanami, Tokyo (1953).

[12] Yukiyoshi Kawada, *The theory of probability* (in Japanese), 4th ed., Kyoritsusha, Tokyo (1952).

[13] A. Khintchine, Déduction nouvelle d'une formule de M. Paul Lévy, *Bull. Univ. d'Etat Moscou, Sér. Internat., Sect. A. Math. et Méca.* I, 1–5 (1937).

[14] A. Khintchine, Ueber Klassenkonvergenz von Verteilungsgesetzen, *Izvestiya Naučno-Issled. Inst. Math. u. Mech. Univ. Tomsk,* 1, 258–261 (1937).

[15] Kiyonori Kunisawa, On an analytical method in the theory of independent random variables, *Ann. Inst. Statist. Math.* Tokyo, 1, 1–77 (1949).

[16] P. Lévy, *Théorie de l'addition des variables aléatoire,* Gautier-Villars (1937).

[17] M. Loéve, On sets of probability laws and their limit elements, *Univ. Calif. Publ. Statist.* 1, 53–87 (1950).

[18] M. Loéve, Fundamental limit theorems of probability theory, *Ann. Math. Stat.* 21, 321–338 (1950).

[19] N. A. Sapogov, On a multidimensional limit theorem of the theory of probability (in Russian), *Uspehi Matem. Nauk* (N.S.), vol. 5, no. 3 (1950).

[20] Kinsaku Takano, On the many-dimensional distribution functions, *Ann. Inst. Statist. Math.* Tokyo, 5, 41–58 (1953).

ERRATA

These Annals, Vol. II, No. 1, 1950. P. 18 insert after last line of section 2

"In these cases we have not the maximum value but only the stationary value just as the minimax solution. If we want to obtain the maximum value, we must estimate the rational rate k_1 and k_2 from experiences in the past time. This fact holds also in the following sections."

Vol. V, No. 1, 1953

Page	line				
27,	9,	read	$M-1-\frac{(M-1)(R-1)}{N}$	instead of	the right hand side of (6)
27,	12,	insert	under the assumption after "we have" $N_i=N/R$		
27,	14,	read	$2(M-1)+O\left(\frac{1}{N}\right)+O\left(\frac{1}{n}\right)$	instead of	the right hand side of (8)
27,	last,	read	(strike off the table)		
28,	5–6,	read	(strike off the sentence "under the condition $M=R(R-1)$ and $R \neq 1$"		

Vol. VI, No. 1, 1954

Page	line				
13,	12,	read	$\binom{M}{Mp_i}p^{Mp_i}q^{Mq_i}$	instead of	$\binom{M}{Mp_i}p^{Mp_i}p^{Mq_i}$
14,	3,	read	$0.96N$	instead of	$096N$
15,	6,	read	$\ldots k\sqrt{\varepsilon^* D^2(\bar{X})}\} \leqq \frac{1}{k^2}$	instead of	$\ldots k\sqrt{\varepsilon^* D^2(\bar{X})} \leqq \frac{1}{k^2}$
15,	23,	read	$X_{(i)}$	instead of	$X_{i)}$
24,	7,	read	$-\mu_{11}(2)\mu_{20}(2)\ldots$	instead of	$-\mu_{11}(1)\mu_{20}(2)\ldots$
24,	10,	read	$\frac{N_1^2N_2}{N^3}((\bar{X}_1-\bar{X}_2)\ldots$	instead of	$\frac{N_1^2N_2}{N^3}(\bar{X}_1-\bar{X}_2)\ldots$
25,	9,	read	$\frac{2N_1N_2}{N^3}(\bar{Y}_1-\bar{Y}_2)^2\cdots$	instead of	$\frac{2N_1N_2}{N^2}(\bar{Y}_1-\bar{Y}_2)^2\cdots$
			$+\frac{N_1N_2}{N^5}(N_1^3+N_2^3)\cdots$	instead of	$+\frac{N_1N_2}{N^5}(N_1^2+N_2^3)\cdots$
28,	2 from the bottom,		$+O(n^{-3/2})$	instead of	$+O(^{-3/2})$
30,	11,	read	$-\frac{4\mu_{31}}{\mu_{11}\mu_{20}}-\frac{4\mu_{13}}{\mu_{11}\mu_{02}}+\cdots$	instead of	$-\frac{4\mu_{31}}{\mu_{11}\mu_{21}}-\frac{4\mu_{13}}{\mu_{11}\mu_{12}}+\cdots$
36,	3 from the bottom,		the coming issue	instead of	this issue

Page	line				
54,	6,	read	[20], Lemma	instead of	[20 , Lemma
68,	28,	read	$e^{I\varepsilon(t)}$	instead of	$eI^{\varepsilon(t)}$
97,	6,	read	(X_j, Y_j) has	instead of	$(X_j, Y_j$ has
98,	2,	read	$\lim_n \frac{D(S_{nj})}{D_n}$	instead of	$\lim_n \sum_l \frac{D(S_{nj})}{D_n}$