APPENDIX

1. Probability Spaces. If we want to describe a random trial mathematically then we first define the sample space Ω, the set of all the possible results or outcomes of the random trial. We shall denote by ω the elements of Ω.

Each event concerning the random trial considered can be represented by a subset of Ω. The impossible event is represented by Θ, the empty set, and the sure event is represented by Ω, the whole sample space. In general, events will be denoted by capital Latin letters $A, B, C, \ldots$.

If the occurrence of A implies the occurrence of B, then we shall write $A \subset B$. The complementary event of an event A will be denoted by $\bar{A}$. The simultaneous occurrence of the events $A, B, C, \ldots$ will be denoted by $ABC\ldots$ or by $A \cap B \cap C \cap \ldots$. The event that at least one event occurs among $A, B, C, \ldots$ will be denoted by $A + B + C + \ldots$ or by $A \cup B \cup C \cup \ldots$. We define $A-B = A\bar{B}$.

We say that $\{A_n\}$ is a monotone sequence of events if either $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots$ or $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$. In the first case we define $\lim_{n \to \infty} A_n = \sum_{n=1}^{\infty} A_n$ and in the second case $\lim_{n \to \infty} A_n = \prod_{n=1}^{\infty} A_n$.

A class of events $\mathcal{A}$ is called an algebra if the following two conditions are satisfied:

(i) If $A \in \mathcal{A}$, then $\bar{A} \in \mathcal{A}$.

(ii) If $A \in \mathcal{A}$ and $B \in \mathcal{A}$, then $A + B \in \mathcal{A}$.

A class of events $\mathcal{B}$ is called a σ-algebra if the following two conditions are satisfied:

(i) If $A \in \mathcal{B}$, then $\bar{A} \in \mathcal{B}$.

(ii) If $A_n \in \mathcal{B}$ for $n = 1,2,\ldots$, then $\sum_{n=1}^{\infty} A_n \in \mathcal{B}$.

A class of events $\mathcal{M}$ is called a monotone class if it satisfies the following requirement:

If $A_n \in \mathcal{M}$ for $n = 1,2,\ldots$ and $\{A_n\}$ is a monotone sequence of events, then $\lim_{n\to\infty} A_n \in \mathcal{M}$.

Theorem 1. Let $\mathcal{A}$ be an algebra of subsets of Ω. Denote by $\mathcal{B}$ the minimal σ-algebra which contains $\mathcal{A}$ and denote by $\mathcal{M}$ the minimal monotone class which contains $\mathcal{A}$. Then $\mathcal{B}$ and $\mathcal{M}$ coincide.

Proof. If $\{A_n\}$ is a monotone sequence of events and $A_n \in \mathcal{B}$, then $\lim_{n\to\infty} A_n \in \mathcal{B}$, that is, $\mathcal{B}$ is a monotone class. Thus $\mathcal{B}$ is a monotone class which contains $\mathcal{A}$. This proves that $\mathcal{M} \subset \mathcal{B}$.

To prove that $\mathcal{B} \subset \mathcal{M}$ for each $A \in \mathcal{M}$ let us define

$$\text{(1)} \qquad \mathcal{M}_A = \{B : B \in \mathcal{M},\ \bar{A}B \in \mathcal{M},\ A\bar{B} \in \mathcal{M},\ A+B \in \mathcal{M}\}.$$

Then $\mathcal{M}_A$ is a monotone class for each $A \in \mathcal{M}$. For if $\{B_n\}$ is a monotone sequence and $B_n \in \mathcal{M}_A$, then $B_n \in \mathcal{M}$, $\bar{A}B_n \in \mathcal{M}$, $A\bar{B}_n \in \mathcal{M}$, $A+B_n \in \mathcal{M}$, and consequently $B = \lim_{n\to\infty} B_n \in \mathcal{M}$, $\bar{A}B = \lim_{n\to\infty} (\bar{A}B_n) \in \mathcal{M}$, $A\bar{B} = \lim_{n\to\infty} (A\bar{B}_n) \in \mathcal{M}$, $A+B = \lim_{n\to\infty} (A+B_n) \in \mathcal{M}$. Therefore $B \in \mathcal{M}_A$.

Now we shall show that if $A \in \mathcal{A}$, then $\mathcal{A} \subset \mathcal{M}_A$ and consequently $\mathcal{M}_A = \mathcal{M}$. If $A \in \mathcal{A}$ and $B \in \mathcal{A}$, then by (1) $B \in \mathcal{M}_A$, that is, $\mathcal{A} \subset \mathcal{M}_A$. Since $\mathcal{M}$ is the minimal monotone class which contains $\mathcal{A}$, and $\mathcal{M}_A$ is a monotone class which contains $\mathcal{A}$, therefore $\mathcal{M} \subset \mathcal{M}_A$. However, by definition $\mathcal{M}_A \subset \mathcal{M}$. Thus $\mathcal{M}_A = \mathcal{M}$ whenever $A \in \mathcal{A}$.

Furthermore, we shall show that $\mathcal{M}_B = \mathcal{M}$ for all $B \in \mathcal{M}$. If $B \in \mathcal{M}$, then $B \in \mathcal{M}_A = \mathcal{M}$ whenever $A \in \mathcal{A}$. Consequently, by symmetry it follows from (1) that $A \in \mathcal{M}_B$ also holds when $A \in \mathcal{A}$ and $B \in \mathcal{M}$. Accordingly, if $A \in \mathcal{A}$, then $A \in \mathcal{M}_B$ for $B \in \mathcal{M}$. This proves that $\mathcal{A} \subset \mathcal{M}_B$ for $B \in \mathcal{M}$. Thus $\mathcal{M} \subset \mathcal{M}_B$ holds and by definition we have $\mathcal{M}_B \subset \mathcal{M}$. Hence $\mathcal{M}_B = \mathcal{M}$ for all $B \in \mathcal{M}$.

Finally, we shall prove that $\mathcal{M}$ is an algebra. If $A \in \mathcal{M}$ and $B \in \mathcal{M}$, then $\mathcal{M}_A = \mathcal{M}$ and by (1) $A+B \in \mathcal{M}$. If $A \in \mathcal{M}$, then $\mathcal{M}_A = \mathcal{M}$ and by (1) $\bar{A}B \in \mathcal{M}$ for all $B \in \mathcal{M}$. If $B = \Omega$, then $B \in \mathcal{M}$ and consequently $\bar{A} \in \mathcal{M}$. This proves that $\mathcal{M}$ is an algebra. Since $\mathcal{M}$ is a monotone class, it follows that $\mathcal{M}$ is necessarily a σ-algebra. Thus $\mathcal{B} \subset \mathcal{M}$. This relation together with $\mathcal{M} \subset \mathcal{B}$ implies that $\mathcal{M} = \mathcal{B}$ which was to be proved.

If we consider a random trial then we suppose that the class of random events is a σ-algebra of subsets of Ω. We use the notation $\mathcal{B}$ for denoting this class.

With every event $A \in \mathcal{B}$ we associate a real number $\underset{\sim}{P}\{A\}$, the probability of A. The probability $\underset{\sim}{P}\{A\}$ is a nonnegative, σ-additive and normed set function defined on $\mathcal{B}$, that is, we assume that

(i) $P\{A\} \geqq 0$ for all $A \in \mathcal{B}$.

(ii) $P\{\Omega\} = 1$.

(iii) If $A_n \in \mathcal{B}$ for $n = 1,2,\ldots$ and $A_iA_j = \theta$ for $i \neq j$,

then

$$P\{\sum_{n=1}^{\infty} A_n\} = \sum_{n=1}^{\infty} P\{A_n\} . \tag{2}$$

In 1914 C. Carathéodory [6] proved an important extension theorem in measure theory. This theorem has many useful applications in the theory of probability. In what follows we shall state and prove this theorem in the terminology of probability theory.

Theorem 2. Let $\mathcal{A}$ be an algebra of subsets of Ω . Let $Q\{A\}$ be a probability defined on $\mathcal{A}$, that is, $Q\{A\} \geqq 0$ for $A \in \mathcal{A}$, $Q\{\Omega\} = 1$ and

$$Q\{\sum_{n=1}^{\infty} A_n\} = \sum_{n=1}^{\infty} Q\{A_n\} \tag{3}$$

whenever $A_n \in \mathcal{A}$ for $n = 1,2,\ldots,$ $\sum_{n=1}^{\infty} A_n \in \mathcal{A}$ and $A_iA_j = \theta$ for $i \neq j$. The probability $Q\{A\}$ defined on $\mathcal{A}$ can uniquely be extended to a probability $P\{A\}$ defined on $\mathcal{B}$, the minimal σ-algebra over $\mathcal{A}$.

Proof. We shall prove that there exists a set function $P\{A\}$ defined on $\mathcal{B}$ which satisfies the conditions (i), (ii), (iii) mentioned above and that $P\{A\}$ is an extension of $Q\{A\}$ that is $P\{A\} = Q\{A\}$ whenever $A \in \mathcal{A}$. Furthermore, we shall prove that $P\{A\}$ for $A \in \mathcal{B}$ is uniquely determined by $Q(A)$ for $A \in \mathcal{A}$.

For any $A \subset \Omega$ let us define

$$(4) \qquad \underset{\sim}{P}^{*}\{A\} = \inf\{\sum_{k=1}^{\infty} \underset{\sim}{Q}\{A_k\} : A \subset \sum_{k=1}^{\infty} A_k \text{ and } A_k \in \mathcal{A}\} .$$

The set function $\underset{\sim}{P}^{*}\{A\}$ satisfies the following properties

(a) $\underset{\sim}{P}^{*}\{A\} \geq 0$ for all A . This follows from the definition (4).

(b) If $A \subset B$, then $\underset{\sim}{P}^{*}\{A\} \leq \underset{\sim}{P}^{*}\{B\}$. This follows from the fact that every covering of B is a covering of A too.

(c) If $A \subset \sum_{n=1}^{\infty} A_n$, then $\underset{\sim}{P}^{*}\{A\} \leq \sum_{n=1}^{\infty} \underset{\sim}{P}^{*}\{A_n\}$.

To prove this let us observe that for any $\varepsilon > 0$ and for each $n = 1,2,\ldots$ we can choose an infinite sequence of sets B_{nj} $(j = 1,2,\ldots)$ such that $B_{nj} \in \mathcal{A}$, $A_n \subset \sum_{j=1}^{\infty} B_{nj}$ and

$$(5) \qquad \sum_{j=1}^{\infty} \underset{\sim}{Q}\{B_{nj}\} \leq \underset{\sim}{P}^{*}\{A_n\} + \frac{\varepsilon}{2^n}$$

for $n = 1,2,\ldots$. Since $A \subset \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} B_{nj}$ and $B_{nj} \in \mathcal{A}$, therefore we have

$$(6) \qquad \underset{\sim}{P}^{*}\{A\} \leq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \underset{\sim}{Q}\{B_{nj}\} \leq \sum_{n=1}^{\infty} \underset{\sim}{P}^{*}\{A_n\} + \varepsilon .$$

Since $\varepsilon > 0$ is arbitrary, this proves (c) .

Now we shall prove that $\underset{\sim}{P}^{*}\{A\}$ is an extension of $\underset{\sim}{Q}\{A\}$, that is, $\underset{\sim}{P}^{*}\{A\} = \underset{\sim}{Q}\{A\}$ if $A \in \mathcal{A}$.

Obviously $\underset{\sim}{P}^*\{A\} \leqq \underset{\sim}{Q}\{A\}$ if $A \in \mathcal{A}$. On the other hand, if $A \in \mathcal{A}$ and $A \subset \sum_{k=1}^{\infty} A_k$ where $A_k \in \mathcal{A}$, then

$$(7) \qquad \underset{\sim}{Q}\{A\} \leqq \sum_{k=1}^{\infty} \underset{\sim}{Q}\{A_k\} .$$

This follows from the σ-additivity of $\underset{\sim}{Q}\{A\}$ on $\mathcal{A}$. If we form the infimum of the right-hand side of (7) for all admissible $\{A_k\}$, then by (4) we obtain that

$$(8) \qquad \underset{\sim}{Q}\{A\} \leqq \underset{\sim}{P}^*\{A\} .$$

Hence $\underset{\sim}{P}^*\{A\} = \underset{\sim}{Q}\{A\}$ for $A \in \mathcal{A}$.

Now denote by $\mathcal{A}^*$ the class of sets S for which for every $\varepsilon > 0$ we can find an $A \in \mathcal{A}$ such that

$$(9) \qquad \underset{\sim}{P}^*\{S \Delta A\} < \varepsilon$$

where $S \Delta A = S\bar{A} + A\bar{S}$, the symmetric difference of S and A.

We shall prove that $\mathcal{A}^*$ is a σ-algebra which contains $\mathcal{A}$.

First, we have $\mathcal{A} \subset \mathcal{A}^*$. For if $S \in \mathcal{A}$, then $A = S$ can be chosen, and hence $\underset{\sim}{P}^*\{S \Delta S\} = \underset{\sim}{P}^*\{\Theta\} = 0$, that is, $S \in \mathcal{A}^*$.

Second, if $S \in \mathcal{A}^*$, then $\bar{S} \in \mathcal{A}^*$. Now for each $\varepsilon > 0$ there is an $A \in \mathcal{A}$ such that $\underset{\sim}{P}^*\{S \Delta A\} < \varepsilon$. If $A \in \mathcal{A}$, then $\bar{A} \in \mathcal{A}$ and $\bar{S} \Delta \bar{A} = S \Delta A$. Thus $\underset{\sim}{P}^*\{\bar{S} \Delta \bar{A}\} = \underset{\sim}{P}^*\{S \Delta A\} < \varepsilon$.

Third, if $S_k \in \mathcal{A}^*$ for $k = 1,2,\ldots,n$, then $S = \sum_{k=1}^{n} S_k \in \mathcal{A}^*$ for all $n = 1,2,\ldots$. In this case for every $\varepsilon > 0$ and $k = 1,2,\ldots, n$,

there is an $A_k \in \mathcal{A}$ such that $\underset{\sim}{P}^*\{S_k \Delta A_k\} < \varepsilon/2^k$. Let $A = \sum_{k=1}^{n} A_k$.
Then $A \in \mathcal{A}$,

$$(10) \qquad S \Delta A \subset \sum_{k=1}^{n} (S_k \Delta A_k)$$

and consequently

$$(11) \qquad \underset{\sim}{P}^*\{S \Delta A\} \leqq \sum_{k=1}^{n} \underset{\sim}{P}^*\{S_k \Delta A_k\} < \varepsilon$$

which proves the statement.

Accordingly, $\mathcal{A}^*$ is an algebra which contains $\mathcal{A}$. Now we shall prove that $\mathcal{A}^*$ is in fact a σ-algebra, that is, if $S_k \in \mathcal{A}^*$ for $k = 1,2,\ldots,$ then $S = \sum_{k=1}^{\infty} S_k \in \mathcal{A}^*$. Since $\bar{S}_1 \ldots \bar{S}_{k-1} S_k \in \mathcal{A}^*$ and $S = S_1 + \bar{S}_1 S_2 + \bar{S}_1 \bar{S}_2 S_3 + \ldots,$ it is sufficient to prove that if $S_k \in \mathcal{A}^*$ for $k = 1,2,\ldots$ and if $S_i S_j = \Theta$ for $i \neq j$, then $S = \sum_{k=1}^{\infty} S_k \in \mathcal{A}^*$.

For every $\varepsilon > 0$ and $k = 1,2,\ldots$ there is an $A_k \in \mathcal{A}$ such that $\underset{\sim}{P}^*\{S_k \Delta A_k\} < \varepsilon/2^{k+1}$. Obviously $\bar{A}_1 \ldots \bar{A}_{k-1} A_k \in \mathcal{A}$ for $k = 1,2,\ldots$ and they are exclusive events. Thus

$$(12) \qquad \sum_{k=1}^{n} \underset{\sim}{Q}\{\bar{A}_1 \ldots \bar{A}_{k-1} A_k\} \leqq \underset{\sim}{Q}\{A_1 + \ldots + A_n\} \leqq 1$$

for $n = 1,2,\ldots$. Hence

$$(13) \qquad \sum_{k=1}^{\infty} \underset{\sim}{Q}\{\bar{A}_1 \ldots \bar{A}_{k-1} A_k\} \leqq 1 ,$$

and consequently

$$(14) \qquad \sum_{k=n+1}^{\infty} \underset{\sim}{Q}\{\bar{A}_1 \ldots \bar{A}_{k-1} A_k\} < \frac{\varepsilon}{2}$$

if n is sufficiently large.

Since

$$(15) \qquad (S \,\Delta \sum_{k=1}^{n} A_k) \subset \sum_{k=1}^{\infty} (S_k \,\Delta\, A_k) + \sum_{k=n+1}^{\infty} \bar{A}_1 \ldots \bar{A}_{k-1} A_k$$

holds for $n = 1,2,\ldots$, therefore if we choose $A_k \in A$ $(k = 1,2,\ldots)$ in such a way that $\underset{\sim}{P}^*\{S_k \,\Delta\, A_k\} < \varepsilon/2^{k+1}$ and if we choose n so large that (14) is satisfied, then by (15) we obtain that

$$(16) \qquad \underset{\sim}{P}^*\{S \,\Delta \sum_{k=1}^{n} A_k\} \leq \sum_{k=1}^{\infty} \underset{\sim}{P}^*\{S_k \,\Delta\, A_k\} + \sum_{k=n+1}^{\infty} \underset{\sim}{Q}\{\bar{A}_1 \ldots \bar{A}_{k-1} A_k\} < \varepsilon .$$

Since $\sum_{k=1}^{n} A_k \in A$ for every $n = 1,2,\ldots$, it follows that $S \subset A^*$ which was to be proved.

Accordingly A^* is a σ-algebra which contains A .

Now we shall prove that $\underset{\sim}{P}^*\{S\}$ is a probability on A^* . By definition $\underset{\sim}{P}^*\{S\} \geq 0$ for all $S \in A^*$ and obviously $\underset{\sim}{P}^*\{\Omega\} = 1$. It remains to prove that $\underset{\sim}{P}^*\{S\}$ is σ-additive on A^* . Suppose that $S = \sum_{k=1}^{\infty} S_k$ where $S_k \in A^*$ for $k = 1,2,\ldots$ and $S_i S_j = \Theta$ for $i \neq j$. Then $S \in A^*$ and by (c) we have

$$(17) \qquad \underset{\sim}{P}^*\{S\} \leq \sum_{k=1}^{\infty} \underset{\sim}{P}^*\{S_k\} .$$

We shall prove that (17) holds also with the reverse inequality. Hence it follows that $\underset{\sim}{P}^*\{S\}$ is σ-additive on A^* .

First we shall prove that if $S_1 \in A^*$, $S_2 \in A^*$ and $S_1 S_2 = \Theta$, then

$$\underset{\sim}{P}^*\{S_1\} + \underset{\sim}{P}^*\{S_2\} \leqq \underset{\sim}{P}^*\{S_1 + S_2\} \leqq \underset{\sim}{P}^*\{S\} \quad . \tag{18}$$

Let us choose $A_1 \varepsilon A^*$ and $A_2 \varepsilon A^*$ in such a way that $\underset{\sim}{P}^*\{S_1 \Delta A_1\} < \varepsilon$ and $\underset{\sim}{P}^*\{S_2 \Delta A_2\} < \varepsilon$ where ε is an arbitrary small positive number.

Since now we have

$$S_1 \subset A_1 + (S_1 \Delta A_1) \quad \text{and} \quad S_2 \subset A_2 + (S_2 \Delta A_2) \ , \tag{19}$$

$$A_1 + A_2 \subset S_1 + S_2 + (S_1 \Delta A_1) + (S_2 \Delta A_2) \tag{20}$$

and

$$A_1 A_2 \subset (S_1 \Delta A_1) + (S_2 \Delta A_2) \ , \tag{21}$$

it follows that

$$\begin{aligned} \underset{\sim}{P}^*\{S_1\} + \underset{\sim}{P}^*\{S_2\} &< \underset{\sim}{Q}\{A_1\} + \underset{\sim}{Q}\{A_2\} + 2\varepsilon = \\ &= \underset{\sim}{Q}\{A_1 + A_2\} + \underset{\sim}{Q}\{A_1 A_2\} + 2\varepsilon \ , \end{aligned} \tag{22}$$

$$\underset{\sim}{Q}\{A_1 + A_2\} = \underset{\sim}{P}^*\{A_1 + A_2\} < \underset{\sim}{P}^*\{S_1 + S_2\} + 2\varepsilon \ , \tag{23}$$

and

$$\underset{\sim}{Q}\{A_1 A_2\} = \underset{\sim}{P}^*\{A_1 A_2\} < 2\varepsilon \quad . \tag{24}$$

By (22), (23), and (24) we have

$$\underset{\sim}{P}^*\{S_1\} + \underset{\sim}{P}^*\{S_2\} < \underset{\sim}{P}^*\{S_1 + S_2\} + 6\varepsilon \quad . \tag{25}$$

Since ε is an arbitrary positive number, this proves (18).

By mathematical induction it follows from (18) that

(26) $$\underset{\sim}{P}^*\{S_1\} + \underset{\sim}{P}^*\{S_2\} + \ldots + \underset{\sim}{P}^*\{S_n\} \leq \underset{\sim}{P}^*\{S\}$$

holds for $n = 2,3,\ldots$. If $n \to \infty$ in (26), then we obtain that

(27) $$\sum_{k=1}^{\infty} \underset{\sim}{P}^*\{S_k\} \leq \underset{\sim}{P}^*\{S\} .$$

By (17) and by (27) it follows that $\underset{\sim}{P}^*\{S\}$ is σ-additive on $\mathcal{A}^*$.

Let $\mathcal{B}$ be the minimal σ-algebra which contains $\mathcal{A}$. Obviously we have $\mathcal{B} \subset \mathcal{A}^*$.

If we define $\underset{\sim}{P}\{A\} = \underset{\sim}{P}^*\{A\}$ on $\mathcal{B}$, then $\underset{\sim}{P}\{A\}$ is a probability on the σ-algebra $\mathcal{B}$ and $\underset{\sim}{P}\{A\}$ is an extension of $\underset{\sim}{Q}\{A\}$, that is, $\underset{\sim}{P}\{A\} = \underset{\sim}{Q}\{A\}$ for $A \in \mathcal{A}$.

Now we shall prove that $\underset{\sim}{P}\{A\}$ is the unique extension of $\underset{\sim}{Q}\{A\}$. To prove this let us suppose that $\underset{\sim}{P}_1\{A\}$ and $\underset{\sim}{P}_2\{A\}$ are both probabilities on $\mathcal{B}$ and both are extensions of $\underset{\sim}{Q}\{A\}$, that is, $\underset{\sim}{P}_1\{A\} = \underset{\sim}{P}_2\{A\} = \underset{\sim}{Q}\{A\}$ for $A \in \mathcal{A}$. We shall prove that $\underset{\sim}{P}_1\{A\} = \underset{\sim}{P}_2\{A\}$ on $\mathcal{B}$.

Define

(28) $$\mathcal{M} = \{A : \underset{\sim}{P}_1\{A\} = \underset{\sim}{P}_2\{A\} \text{ and } A \in \mathcal{B}\} .$$

Then $\mathcal{A} \subset \mathcal{M} \subset \mathcal{B}$. We can easily see that $\mathcal{M}$ is a monotone class. Let $\{A_n\}$ be a monotone sequence of events for which $A_n \in \mathcal{M}$. Then $A = \lim_{n\to\infty} A_n \in \mathcal{M}$. For in this case $\underset{\sim}{P}_1\{A_n\} = \underset{\sim}{P}_2\{A_n\}$ for $n = 1,2,\ldots$ and therefore

$$\underset{\sim}{P}_1\{A\} = \lim_{n\to\infty} \underset{\sim}{P}_1\{A_n\} = \lim_{n\to\infty} \underset{\sim}{P}_2\{A_n\} = \underset{\sim}{P}_2\{A\} \quad . \tag{29}$$

By Theorem 1 it follows that M contains the minimal σ-algebra over A_n, that is, $\mathcal{B} \subset M$. Accordingly $M = \mathcal{B}$, that is, $\underset{\sim}{P}\{A\}$ is the unique extension of $\underset{\sim}{Q}\{A\}$ to the σ-algebra $\mathcal{B}$. This completes the proof of the theorem.

In the mathematical description of a random trial we associate a probability space $(\Omega,\mathcal{B},\underset{\sim}{P})$ with the random trial where Ω is the sample space, the set of all the possible outcomes of the random trial, $\mathcal{B}$ is a σ-algebra of subsets of Ω, the set of random events, and $\underset{\sim}{P}$ is a normed measure defined on $\mathcal{B}$, that is, $\underset{\sim}{P}\{A\}$ is the probability of $A \in \mathcal{B}$.

<u>Theorem 3.</u> <u>If</u> $A_n \in \mathcal{B}$ <u>for</u> $n = 1,2,\ldots,$ <u>and if</u> $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots,$ <u>then</u>

$$A = \lim_{n\to\infty} A_n = \sum_{k=1}^{\infty} A_k \in \mathcal{B} \tag{30}$$

<u>and</u>

$$\underset{\sim}{P}\{A\} = \lim_{n\to\infty} \underset{\sim}{P}\{A_n\} \; . \tag{31}$$

<u>Proof.</u> Since $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots$ we can write that

$$A = A_1 + A_2\bar{A}_1 + \ldots + A_n\bar{A}_{n-1} + \ldots \tag{32}$$

where the events on the right-hand side are exclusive events. Thus we have

$$\begin{aligned}\underset{\sim}{P}\{A\} &= \underset{\sim}{P}\{A_1\} + \underset{\sim}{P}\{A_2\bar{A}_1\} + \ldots + \underset{\sim}{P}\{A_n\bar{A}_{n-1}\} + \ldots = \\ &= \underset{\sim}{P}\{A_1\} + [\underset{\sim}{P}\{A_2\} - \underset{\sim}{P}\{A_1\}] + \ldots + [\underset{\sim}{P}\{A_n\} - \underset{\sim}{P}\{A_{n-1}\}] + \ldots \\ &= \lim_{n\to\infty} \underset{\sim}{P}\{A_n\}\end{aligned} \tag{33}$$

because the n-th partial sum is $\underset{\sim}{P}\{A_n\}$ in the above infinite series.

<u>Theorem 4</u>. <u>If</u> $A_n \in \mathcal{B}$ <u>for</u> $n = 1,2,\ldots$ <u>and if</u> $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$, <u>then</u>

(34) $$A = \lim_{n \to \infty} A_n = \prod_{k=1}^{\infty} A_k \in \mathcal{B}$$

<u>and</u>

(35) $$\underset{\sim}{P}\{A\} = \lim_{n \to \infty} \underset{\sim}{P}\{A_n\} .$$

<u>Proof</u>. Since now $\bar{A}_1 \subset \bar{A}_2 \subset \ldots \subset \bar{A}_n \subset \ldots$ and

(36) $$\bar{A} = \lim_{n \to \infty} \bar{A}_n = \sum_{k=1}^{\infty} \bar{A}_k ,$$

by Theorem 3 we obtain that

(37) $$\underset{\sim}{P}\{\bar{A}\} = \lim_{n \to \infty} \underset{\sim}{P}\{\bar{A}_n\}$$

and this proves (35).

<u>Note</u>. If $A_n \in \mathcal{B}$ for $n = 1,2,\ldots$, $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ and $\prod_{k=1}^{\infty} A_k = \Theta$, then by Theorem 4 we have

(38) $$\lim_{n \to \infty} \underset{\sim}{P}\{A_n\} = 0 .$$

It is interesting to observe that if $\underset{\sim}{P}\{A\}$ is finitely additive on $\mathcal{B}$ and if $\underset{\sim}{P}\{A\}$ is continuous at Θ, that is, if (38) holds, then $\underset{\sim}{P}\{A\}$ is σ-additive on $\mathcal{B}$. This can be seen as follows:

Let $B_n \in \mathcal{B}$ for $n = 1,2,\ldots$ and suppose that $B_iB_j = \Theta$ for $i \neq j$. Define $A_n = B_n + B_{n+1} + \ldots$ for $n = 1,2,\ldots$. Then $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ and $\prod_{n=1}^{\infty} A_n = \Theta$. For if $\omega \in B_n$, then $\omega \notin A_{n+1}$ and if $\omega \notin \sum_{n=1}^{\infty} B_n$, then $\omega \notin A_n$ for any $n = 1,2,\ldots$. Thus by (38)

$$\lim_{n \to \infty} P\{A_n\} = 0 . \tag{39}$$

On the other hand $A_1 = B_1 + B_2 + \ldots + B_n + A_{n+1}$ and therefore

$$P\{\sum_{k=1}^{\infty} B_k\} = P\{A_1\} = P\{B_1\} + P\{B_2\} + \ldots + P\{B_n\} + P\{A_{n+1}\} \tag{40}$$

for $n = 1,2,\ldots$. Since by (39) $\lim_{n \to \infty} P\{A_{n+1}\} = 0$, it follows from (40) that

$$P\{\sum_{k=1}^{\infty} B_k\} = \sum_{k=1}^{\infty} P\{B_k\} \tag{41}$$

which proves that $P\{A\}$ is σ-additive on $\mathcal{B}$.

Accordingly, we can state that $P\{A\}$ is a probability defined on $\mathcal{B}$ if it satisfies the following requirements:

(a) $P\{A\} \geq 0$ for $A \in \mathcal{B}$

(b) $P\{\Omega\} = 1$

(c) If $A \in \mathcal{B}$ and $B \in \mathcal{B}$ and $AB = \Theta$, then $P\{A+B\} = P\{A\} + P\{B\}$.

(d) If $A_n \in \mathcal{B}$, $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ and $\prod_{n=1}^{\infty} A_n = \Theta$, then $\lim_{n \to \infty} P\{A_n\} = 0$.

This set of requirements is equivalent to the requirements (i), (ii),

(iii) stated earlier. In particular, it follows that a nonnegative and normed set function $P\{A\}$ defined on a σ-algebra $\mathcal{B}$ is σ-additive if and only if (c) and (d) are satisfied.

Now we shall prove a few basic relations for probabilities. First, we shall prove Boole's inequality.

<u>Theorem 5. Let</u> $(\Omega,\mathcal{B},P)$ <u>be a probability space and</u> $A_1, A_2,\ldots, A_k,\ldots$ <u>an infinite sequence of events. Then we have</u>

$$P\{\sum_{k=1}^{\infty} A_k\} \leq \sum_{k=1}^{\infty} P\{A_k\} . \tag{42}$$

<u>Proof</u>. Let $B_1 = A_1$ and $B_k = \bar{A}_1\ldots\bar{A}_{k-1} A_k$ for $k = 2,3,\ldots$. Then we have

$$\sum_{k=1}^{\infty} A_k = \sum_{k=1}^{\infty} B_k . \tag{43}$$

Since the events $B_1, B_2,\ldots, B_k,\ldots$ are mutually exclusive, it follows that

$$P\{\sum_{k=1}^{\infty} A_k\} = \sum_{k=1}^{\infty} P\{B_k\} \leq \sum_{k=1}^{\infty} P\{A_k\} . \tag{44}$$

Here we used that $B_k \subset A_k$ for $k = 1,2,\ldots$.

<u>Theorem 6. Let</u> $(\Omega,\mathcal{B},P)$ <u>be a probability space and</u> $A_1, A_2,\ldots, A_n,\ldots$ <u>be an infinite sequence of events. Define</u>

$$A^* = \limsup_{n\to\infty} A_n = \prod_{n=1}^{\infty} \sum_{k=n}^{\infty} A_k \tag{45}$$

<u>and</u>

(46) $$A_* = \liminf_{n \to \infty} A_n = \sum_{n=1}^{\infty} \prod_{k=n}^{\infty} A_k .$$

We have

(47) $$\underset{\sim}{P}\{A_*\} \leq \liminf_{n \to \infty} \underset{\sim}{P}\{A_n\} \leq \limsup_{n \to \infty} \underset{\sim}{P}\{A_n\} \leq \underset{\sim}{P}\{A^*\} .$$

Proof. If we apply Theorem 3 to the events $\prod_{k=n}^{\infty} A_k$ $(n = 1,2,\ldots)$ then we obtain that

(48) $$\underset{\sim}{P}\{A_*\} = \lim_{n \to \infty} \underset{\sim}{P}\{\prod_{k=n}^{\infty} A_k\}$$

and if we apply Theorem 4 to the events $\sum_{k=n}^{\infty} A_k$ $(n = 1,2,\ldots)$ then we obtain that

(49) $$\underset{\sim}{P}\{A^*\} = \lim_{n \to \infty} \underset{\sim}{P}\{\sum_{k=n}^{\infty} A_k\} .$$

Since

(50) $$\underset{\sim}{P}\{\prod_{k=n}^{\infty} A_k\} \leq \underset{\sim}{P}\{A_n\}$$

for $n = 1,2,\ldots,$ by (48) we obtain that

(51) $$\underset{\sim}{P}\{A_*\} \leq \liminf_{n \to \infty} \underset{\sim}{P}\{A_n\} ,$$

and since

(52) $$\underset{\sim}{P}\{A_n\} \leq \underset{\sim}{P}\{\sum_{k=n}^{\infty} A_k\}$$

for $n = 1,2,\ldots,$ by (49) we obtain that

(53) $$\limsup_{n \to \infty} \underset{\sim}{P}\{A_n\} \leq \underset{\sim}{P}\{A^*\} .$$

By (51) and (53) we obtain (47).

We note that if $A^* = A_*$, then we say that $\lim_{n \to \infty} A_n$ exists, and $\lim_{n \to \infty} A_n = A^* = A_*$. In this case by (47) we have

(54) $$P\{\lim_{n \to \infty} A_n\} = \lim_{n \to \infty} P\{A_n\} .$$

2. Random Variables and Distribution Functions.

Let $(\Omega, \mathcal{B}, P)$ be a probability space. By a real random variable ξ we understand a real function $\xi = \xi(\omega)$ defined for $\omega \in \Omega$ and measurable with respect to $\mathcal{B}$, that is, for every real x the event $\{\omega : \xi(\omega) \leq x\} \in \mathcal{B}$. A random variable $\xi(\omega)$ may be finite or infinite. If it is not specified otherwise, then by a random variable ξ we mean a finite, measurable, real function $\xi(\omega)$ defined on Ω.

If $\xi = \xi(\omega)$ is a real random variable, then $\{\omega : \xi(\omega) \in A\} \in \mathcal{B}$ for any linear Borel set A and $\mu(A) = P\{\xi \in A\}$ is a probability measure on the class of Borel subsets of the real line.

If $\xi = \xi(\omega)$ is a finite random variable, then the function

(1) $$F(x) = P\{\xi \leq x\}$$

defined for $-\infty < x < \infty$ is called the distribution function of the random variable. We define $F(+\infty) = \lim_{x \to \infty} F(x)$ and $F(-\infty) = \lim_{x \to -\infty} F(x)$.

A distribution function $F(x)$ has the following properties: (i) $F(x)$ is a nondecreasing function of x. (ii) $F(+\infty) = 1$ and $F(-\infty) = 0$.

(iii) $F(x)$ is continuous on the right, that is, $\lim_{y \to x} F(y) = F(x+0) = F(x)$ if $y > x$.

Conversely, if $F(x)$ is a real function of x defined for $-\infty < x < \infty$ and if $F(x)$ satisfies the conditions (i), (ii), (iii), then $F(x)$ can be considered as the distribution function of a real random variable. We shall prove that $F(x)$ induces a probability space $(\Omega, \mathcal{B}, \underset{\sim}{P})$ and we shall define a random variable $\xi = \xi(\omega)$ such that $\underset{\sim}{P}\{\xi \leq x\} = F(x)$.

Theorem 1. Let $F(x)$ be a distribution function, that is, a real function satisfying the conditions (i), (ii), (iii). Then there exists a probability space $(\Omega, \mathcal{B}, \underset{\sim}{P})$ and a real random variable ξ such that $\underset{\sim}{P}\{\xi \leq x\} = F(x)$.

Proof. Let $\Omega = R\ (-\infty, \infty)$, a real line. Let $\mathcal{B}$ be the class of Borel sets in R. Let us define $\underset{\sim}{P}\{A\}$ for $A \in \mathcal{B}$ in the following way: If $I = (a,b]$ where $a \leq b$, then let $\underset{\sim}{P}\{I\} = F(b) - F(a)$. If $I = (a,b)$ where $a \leq b$, then let $\underset{\sim}{P}\{I\} = F(b-0) - F(a)$. If $I = [a,b]$ where $a \leq b$, then let $\underset{\sim}{P}\{I\} = F(b) - F(a-0)$. If $I = [a,b)$ where $a \leq b$, then let $\underset{\sim}{P}\{I\} = F(b-0) - F(a-0)$. Thus $\underset{\sim}{P}\{I\}$ is defined for intervals I. Now let us extend the definition of $\underset{\sim}{P}\{A\}$ for elementary sets A. A set A is called an elementary set if it can be represented as the union of a finite number of intervals. If A is an elementary set, then we can write that $A = I_1 + I_2 + \dots + I_n$ where $I_1, I_2, \dots, I_n$ are disjoint intervals. For the elementary set A let us define $\underset{\sim}{P}\{A\} = \underset{\sim}{P}\{I_1\} + \underset{\sim}{P}\{I_2\} + \dots + \underset{\sim}{P}\{I_n\}$. We can easily see that the class of elementary sets $\mathcal{A}$ is an algebra and $\underset{\sim}{P}\{A\}$ is uniquely determined for $A \in \mathcal{A}$, that is, $\underset{\sim}{P}\{A\}$ is independent of the particular representation of A. We have $\underset{\sim}{P}\{A\} \geqq 0$ for each $A \in \mathcal{A}$, $\underset{\sim}{P}\{\Omega\} = 1$ and $\underset{\sim}{P}\{A\}$ is finitely additive, that is, if $A \in \mathcal{A}$ and $A = A_1 + A_2 + \ldots + A_n$ where $A_i \in \mathcal{A}$ for $i = 1,2,\ldots, n$ and $A_i A_j = \Theta$ for $i \neq j$, then $\underset{\sim}{P}\{A\} = \underset{\sim}{P}\{A_1\} + \underset{\sim}{P}\{A_2\} + \ldots + \underset{\sim}{P}\{A_n\}$.

Now we shall prove that $\underset{\sim}{P}\{A\}$ is σ-additive on $\mathcal{A}$. We shall provide two proofs of this fact.

First proof. Let $A \in \mathcal{A}$ and $A_k \in \mathcal{A}$ for $k = 1,2,\ldots$ and suppose that $A \subset \sum_{k=1}^{\infty} A_k$. Then we have

$$\underset{\sim}{P}\{A\} \leqq \sum_{k=1}^{\infty} \underset{\sim}{P}\{A_k\} . \tag{2}$$

To prove (2) we observe that for every $\varepsilon > 0$ we can find a bounded and closed elementary set $B \subset A$ such that $\underset{\sim}{P}\{B\} \geqq \underset{\sim}{P}\{A\} - \frac{\varepsilon}{2}$. This can easily be seen if we take into consideration that every interval I contains a bounded and closed interval K such that $\underset{\sim}{P}\{I\} - \underset{\sim}{P}\{K\}$ is arbitrarily close to 0. For example, if $I = (a,b)$ where $a < b$ and $K_\varepsilon = [a+\varepsilon, b-\varepsilon]$, then $\lim_{\varepsilon \to 0} \underset{\sim}{P}\{K_\varepsilon\} = \lim_{\varepsilon \to 0} [F(b-\varepsilon) - F(a+\varepsilon-0)] = F(b-0) - F(a) = \underset{\sim}{P}\{I\}$, that is, $\underset{\sim}{P}\{I\} - \underset{\sim}{P}\{K_\varepsilon\}$ is arbitrarily close to 0 if $\varepsilon > 0$ is sufficiently small. In a similar way we can see that for every $\varepsilon > 0$ and $k = 1,2,\ldots$ we can find an open elementary set $B_k \supset A_k$ such that $\underset{\sim}{P}\{B_k\} \leqq \underset{\sim}{P}\{A_k\} + \frac{\varepsilon}{2^{k+1}}$. Then we have $B \subset \sum_{k=1}^{\infty} B_k$. Since B is bounded and closed, by the Heine-Borel theorem there is an n such that $B \subset \sum_{k=1}^{n} B_k$. (See e.g. <u>B. Sz. Nagy</u>

[31 p.39].) Then by the finite additivity of $\underset{\sim}{P}\{A\}$ on $\mathcal{A}$ we obtain that $\underset{\sim}{P}\{B\} \leqq \sum_{k=1}^{n} \underset{\sim}{P}\{B_k\}$. Thus

$$(3) \qquad \underset{\sim}{P}\{A\} \leqq \underset{\sim}{P}\{B\} + \frac{\varepsilon}{2} \leqq \sum_{k=1}^{n} \underset{\sim}{P}\{B_k\} + \frac{\varepsilon}{2} \leqq \sum_{k=1}^{\infty} \underset{\sim}{P}\{A_k\} + \varepsilon .$$

Since $\varepsilon > 0$ is arbitrary, therefore (2) follows.

If $A = \sum_{k=1}^{\infty} A_k$ where $A \in \mathcal{A}$, $A_k \in \mathcal{A}$ for $k = 1,2,\ldots$ and $A_i A_j = \Theta$ for $i \neq j$, then we have

$$(4) \qquad \sum_{k=1}^{\infty} \underset{\sim}{P}\{A_k\} \leqq \underset{\sim}{P}\{A\} .$$

This follows from the relation $A_1 + A_2 + \ldots + A_n \subset A$ which implies that $\underset{\sim}{P}\{A_1\} + \underset{\sim}{P}\{A_2\} + \ldots + \underset{\sim}{P}\{A_n\} \leqq \underset{\sim}{P}\{A\}$ for all $n = 1,2,\ldots$. If $n \to \infty$, then we obtain (4). By (2) it follows that (4) holds also with the reverse inequality. Thus

$$(5) \qquad \underset{\sim}{P}\{A\} = \sum_{k=1}^{\infty} \underset{\sim}{P}\{A_k\} ,$$

that is, $\underset{\sim}{P}\{A\}$ is σ-additive on $\mathcal{A}$.

Second proof. Since $\underset{\sim}{P}\{A\}$ is finitely additive on $\mathcal{A}$, it is sufficient to prove that $\underset{\sim}{P}\{A\}$ is continuous at Θ , that is, if $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ where $A_n \in \mathcal{A}$ and $\lim_{n \to \infty} A_n = \prod_{n=1}^{\infty} A_n = \Theta$, then $\lim_{n \to \infty} \underset{\sim}{P}\{A_n\} = 0$. This implies that $\underset{\sim}{P}\{A\}$ is σ-additive on $\mathcal{A}$. (See the previous section where we proved this for a σ-algebra $\mathcal{B}$.) We shall prove that if $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ where $A_n \in \mathcal{A}$ and $\lim_{n \to \infty} \underset{\sim}{P}\{A_n\} \geqq \varepsilon > 0$, then $\prod_{n=1}^{\infty} A_n$ is not empty.

For each $\varepsilon > 0$ and $n = 1,2,\ldots$ we can find a bounded and closed elementary set $B_n \subset A_n$ such that $\underset{\sim}{P}\{B_n\} \geqq \underset{\sim}{P}\{A_n\} - \frac{\varepsilon}{2^{n+1}}$. Let $C_n = B_1 B_2 \ldots B_n$. Since $C_n \subset B_n \subset A_n$ and $A_n\bar{C}_n = A_n\bar{B}_1 + \ldots + A_n\bar{B}_n \subset A_1\bar{B}_1 + \ldots + A_n\bar{B}_n$, it follows that

$$\underset{\sim}{P}\{A_n\} - \underset{\sim}{P}\{C_n\} \leqq \sum_{k=1}^{n} [\underset{\sim}{P}\{A_k\} - \underset{\sim}{P}\{B_k\}] \leqq \frac{\varepsilon}{2} . \tag{6}$$

Hence

$$\underset{\sim}{P}\{C_n\} \geqq \underset{\sim}{P}\{A_n\} - \frac{\varepsilon}{2} \geqq \frac{\varepsilon}{2} > 0 , \tag{7}$$

that is, C_n is not empty. Thus there exists a real number $x_n \varepsilon C_n$ for each $n = 1,2,\ldots$. Since $C_n \subset C_m$ for $n \geqq m$, it follows that $x_n \varepsilon C_m$ for $n \geqq m$, or $x_n \varepsilon B_m$ for $n \geqq m$. Since B_m is bounded and closed by the Bolzano-Weierstrass theorem $\{x_n\}$ contains a convergent subsequence $\{x_{n_k}\}$ such that $\lim_{k \to \infty} x_{n_k} = x \varepsilon B_m$ for all $m = 1,2,\ldots$. (See e.g. B. Sz. - Nagy [31 p.30].) Thus $x \varepsilon A_m$ for all $m = 1,2,\ldots$ and consequently $\prod_{m=1}^{\infty} A_m$ is not empty. This proves that $\underset{\sim}{P}\{A\}$ is σ-additive on $\mathcal{A}$.

Since $\underset{\sim}{P}\{A\}$ is σ-additive on $\mathcal{A}$ by Carathéodory's extension theorem (Theorem 1.2 in the Appendix) we can extend the definition of $\underset{\sim}{P}\{A\}$ to $\mathcal{B}$, the minimal σ-algebra over $\mathcal{A}$, in such a way that $\underset{\sim}{P}\{A\}$ remains non-negative, normed and σ-additive on $\mathcal{B}$ and the extension is unique.

Thus we demonstrated that every distribution function $F(x)$ induces a probability space $(\Omega, \mathcal{B}, \underset{\sim}{P})$ and if $A_x = \{\omega : \omega \leq x\}$, then $\underset{\sim}{P}\{A_x\} = F(x)$ for all $-\infty < x < \infty$.

If we define $\xi = \xi(\omega) = \omega$ for $\omega \in \Omega$, then ξ is a real random variable and $\underset{\sim}{P}\{\xi \leq x\} = \underset{\sim}{P}\{A_x\} = F(x)$ for all $x \in (-\infty, \infty)$. This completes the proof of the theorem.

Now let us suppose that m real random variables $\xi_1, \xi_2, \ldots \xi_m$ are defined on the probability space $(\Omega, \mathcal{B}, \underset{\sim}{P})$. We can consider the random variables $\xi_1, \xi_2, \ldots, \xi_m$ as the components of a vector random variable $\underset{\sim}{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)$. Then $\{\omega : \underset{\sim}{\xi}(\omega) \in A\} \in \mathcal{B}$ for any m-dimensional Borel set A and $\mu(A) = \underset{\sim}{P}\{\underset{\sim}{\xi} \in A\}$ is a probability measure on the class of Borel subsets of the m-dimensional Euclidean space.

The function

$$F(x_1, x_2, \ldots, x_m) = \underset{\sim}{P}\{\xi_1 \leq x_1, \xi_2 \leq x_2, \ldots, \xi_m \leq x_m\} \tag{8}$$

defined for $x_i \in (-\infty, \infty)$ $(i = 1, 2, \ldots, m)$ is called the joint distribution function of the random variables $\xi_1, \xi_2, \ldots, \xi_m$.

An m-dimensional distribution function $F(x_1, x_2, \ldots, x_m)$ has the following properties: (i) $F(x_1, x_2, \ldots, x_m)$ is a nondecreasing function of x_i for each $i = 1, 2, \ldots, m$. (ii) $F(x_1, x_2, \ldots, x_m) \to 1$ if every $x_i \to +\infty$ $(i = 1, 2, \ldots, m)$ and $F(x_1, x_2, \ldots, x_m) \to 0$ if at least one $x_i \to -\infty$ $(i = 1, 2, \ldots, m)$. (iii) If $x_i \leq y_i$ for $i = 1, 2, \ldots, m$ and if $y_i \to x_i$ for $i = 1, 2, \ldots, m$, then $F(y_1, y_2, \ldots, y_m) \to F(x_1, x_2, \ldots, x_m)$. (iv) If $a_i < b_i$ for $i = 1, 2, \ldots, m$, then

$$\sum_{r_1=0}^{1} \sum_{r_2=0}^{1} \cdots\cdots \sum_{r_m=0}^{1} (-1)^{m-r_1-r_2-\cdots-r_m} F(a_1+r_1(b_1-a_1),\ a_2 + \tag{9}$$

$$+ r_2(b_2-a_2), \ldots, a_m+r_m(b_m-a_m)) \geq 0 .$$

If we evaluate the probability $\underset{\sim}{P}\{a_1 < \xi_1 \leq b_1,\ a_2 < \xi_2 \leq b_2, \ldots, a_m < \xi_m \leq b_m\}$ by using the method of inclusion and exclusion, then we obtain the left-hand side of (9).

Conversely, if a real function $F(x_1, x_2, \ldots, x_m)$ is defined for $x_i \in (-\infty, \infty)$ $(i = 1, 2, \ldots, m)$ and if it satisfies the above conditions (i), (ii), (iii), (iv), then $F(x_1, x_2, \ldots, x_m)$ can be considered as the joint distribution function of m real random variables. We shall prove that $F(x_1, x_2, \ldots, x_m)$ induces a probability space $(\Omega, \mathcal{B}, \underset{\sim}{P})$ and we shall define m real random variables $\xi_1 = \xi_1(\omega)$, $\xi_2 = \xi_2(\omega), \ldots, \xi_m = \xi_m(\omega)$ such that

$$\underset{\sim}{P}\{\xi_1 \leq x_1,\ \xi_2 \leq x_2, \ldots, \xi_m \leq x_m\} = F(x_1, x_2, \ldots, x_m) \tag{10}$$

for all $x_i \in (-\infty, \infty)$ $(i = 1, 2, \ldots, m)$.

<u>Theorem 2</u>. <u>Let</u> $F(x_1, x_2, \ldots, x_m)$ <u>be an m-dimensional distribution function, that is, a real function satisfying the conditions</u> (i), (ii), (iii), (iv). <u>Then there exists a probability space</u> $(\Omega, \mathcal{B}, \underset{\sim}{P})$ <u>and</u> m <u>real random variables</u> $\xi_1, \xi_2, \ldots, \xi_m$ <u>such that</u> (10) <u>holds.</u>

<u>Proof.</u> Let $\Omega = R_m = \{(\omega_1, \omega_2, \ldots, \omega_m) : -\infty < \omega_i < \infty \text{ for } i = 1, 2, \ldots, m\}$ be an m-dimensional Euclidean space. Let $\mathcal{B}$ be the class of Borel sets in R_m , that is, $\mathcal{B}$ is the smallest σ-algebra which contains all those

m-dimensional intervals in R_m, whose sides are parallel to the coordinate axis. Let us define $\underset{\sim}{P}\{A\}$ for $A \in \mathcal{B}$ in the following way: If $I = \{(\omega_1, \omega_2, \ldots, \omega_m) : a_i < \omega_i \leqq b_i \text{ for } i = 1,2,\ldots,m\}$. Then let $\underset{\sim}{P}\{I\}$ be the left-hand side of (9). Let us define in a similar way $\underset{\sim}{P}\{I\}$ for any m-dimensional interval whose sides are parallel to the coordinate axis. Denote by $\mathcal{A}$ the class of elementary sets in R_m, that is, $\mathcal{A}$ is the class of all those sets in R_m which can be represented as the union of a finite number of intervals in R_m. Let us extend the definition of $\underset{\sim}{P}$ from intervals to elementary sets in exactly the same way as in the case of one dimension. Then $\underset{\sim}{P}\{A\} \geqq 0$, $\underset{\sim}{P}\{R_m\} = 1$ and $\underset{\sim}{P}\{A\}$ is finitely additive on $\mathcal{A}$. We can easily see that $\mathcal{A}$ is an algebra. By using the Heine-Borel theorem or the Bolzano-Weierstrass theorem for the m-dimensional Euclidean space, in exactly the same way as in the one-dimensional case, we can prove that $\underset{\sim}{P}\{A\}$ is σ-additive on $\mathcal{A}$. Then by Carathéodory's extension theorem (Theorem 1.2 in the Appendix) we can extend the definition of $\underset{\sim}{P}\{A\}$ to $\mathcal{B}$, the minimal σ-algebra over $\mathcal{A}$, in such a way that $\underset{\sim}{P}\{A\}$ remains nonnegative, normed and σ-additive on $\mathcal{B}$ and the extension is unique.

Thus we demonstrated that every m-dimensional distribution function $F(x_1, x_2, \ldots, x_m)$ induces a probability space $(\Omega, \mathcal{B}, \underset{\sim}{P})$ and if $A_{x_1,x_2,\ldots,x_m} = \{(\omega_1, \omega_2, \ldots, \omega_m) : \omega_i \leqq x_i \text{ for } i = 1,2,\ldots, m\}$, then

(11) $$\underset{\sim}{P}\{A_{x_1,x_2,\ldots,x_m}\} = F(x_1, x_2, \ldots, x_m)$$

for all $x_i \in (-\infty, \infty)$ $(i = 1,2,\ldots,m)$. See also R. Sikorski and B. Znojkiewicz [28].

If we define $\xi_i = \xi_i(\omega) = \omega_i$ for $\omega = (\omega_1, \omega_2, \ldots, \omega_m) \in \Omega$ and $i = 1,2,\ldots, m$, then $\xi_1, \xi_2, \ldots, \xi_m$ are real random variables and

$$(12) \qquad \underset{\sim}{P}\{\xi_1 \leqq x_1, \xi_2 \leqq x_2, \ldots, \xi_m \leqq x_m\} = \underset{\sim}{P}\{A_{x_1,x_2,\ldots,x_m}\} = F(x_1,x_2,\ldots,x_m)$$

for all $x_i \in (-\infty,\infty)$ $(i = 1,2,\ldots,m)$. This completes the proof of the theorem.

In generalizing the above results we can consider random variables belonging to a metric space X.

A space X is called a metric space if for any two points (elements) x and y of X there is defined a single-valued real function $d(x, y)$, the distance from x to y, satisfying the conditions: $d(x, y) \geqq 0$; $d(x, y) = 0$ if and only if $x = y$; $d(x, y) = d(y, x)$; and $d(x, z) \leqq d(x, y) + d(y, z)$ for any $z \in X$.

By using the metric $d(x, y)$ we can introduce topological notions in the space X similarly to Euclidean spaces.

A sequence $\{x_n\}$ in the metric space X is called a Cauchy sequence if and only if for each $\varepsilon > 0$ there is an r such that $d(x_m, x_n) < \varepsilon$ whenever $m \geqq r$ and $n \geqq r$.

The space X is called complete if for each Cauchy sequence $\{x_n\}$ in X there is a point $x \in X$ such that $d(x, x_n) \to 0$ as $n \to \infty$.

The space X is called separable if it contains a sequence $\{x_n\}$ which is dense everywhere, that is, if for every $x \in X$ there is a subsequence

$\{x_{n_k}\}$ of $\{x_n\}$ such that $d(x, x_{n_k}) \to 0$ as $k \to \infty$.

If $x \in X$ and r is a positive real number, then the set $S(x; r) = \{y : d(x, y) < r, y \in X\}$ is called an open sphere in X with center x and radius r. The set $S^*(x; r) = \{y : d(x, y) \leqq r, y \in X\}$ is called a closed sphere in X with center x and radius r.

A set A in X is called an open set if each $x \in A$ is an interior point of A, that is if for each $x \in A$ there is an $r > 0$ such that $S(x; r) \subset A$.

A set A^* in X is called a closed set if each limit point of A^* belongs to A^*. A point $x \in X$ is a limit point of A^* if there is a sequence of points $x_n \in A^*$ $(n = 1,2,\ldots)$ for which $x_n \neq x$ and $d(x, x_n) \to 0$ as $n \to \infty$. (If $x \in A^*$ and if x is not a limit point of A^*, then x is called an isolated point of A^*.)

Denote by F the smallest σ-algebra which contains all the open sets (closed sets) in X. The elements of F are called Borel sets in X. If X is separable, then F can also be characterized as the smallest σ-algebra which contains all the open spheres (closed spheres) in X.

Let (Ω, B, P) be a probability space. By a random variable ξ taking on values in a metric space X we understand a function $\xi = \xi(\omega)$ which is defined for $\omega \in \Omega$, which takes on values in X, and which is measurable with respect to B, that is, for each open set (closed set) A in X the set $\{\omega : \xi(\omega) \in A\}$ belongs to B. If the metric space X is separable, then in order that $\xi = \xi(\omega)$ be a random variable it is sufficient to require that for each open shpere (closed sphere) S in X the set $\{\omega : \xi(\omega) \in S\}$ belong to B. For this requirement implies that

$\{\omega : \xi(\omega) \in A\} \in \mathcal{B}$ for every open set (closed set) A in X.

If $\xi = \xi(\omega)$ is a random variable taking on values in a metric space X, then $\{\omega : \xi(\omega) \in A\} \in \mathcal{B}$ for every Borel set A in X. Thus $\mu(A) = \mathbf{P}\{\xi \in A\}$ is uniquely determined for each $A \in \mathcal{F}$. The set function $\mu(A)$ is a probability measure on $\mathcal{F}$, the σ-algebra of Borel sets in X.

The converse of this last statement is also true.

Theorem 3. Let X be a complete and separable metric space with distance function $d(x, y)$. Let $\mathcal{F}$ be the σ-algebra of Borel subsets of X and let μ be a probability measure on $\mathcal{F}$. Let $\Omega = (0, 1)$, $\mathcal{B}$ the σ-algebra of Borel subsets of Ω, and $\mathbf{P}$ the Lebesgue measure. Then there exists a random variable $\xi(\omega)$ taking values in X and defined on $(\Omega, \mathcal{B}, \mathbf{P})$ such that

$$\mathbf{P}\{\xi(\omega) \in S\} = \mu(S) \tag{13}$$

for $S \in \mathcal{F}$.

Proof. We observe that if $X = S_1 + S_2 + \ldots$ where $S_1, S_2, \ldots$ are disjoint sets belonging to $\mathcal{F}$ and if we define

$$\xi^{(1)}(\omega) = x_i \quad \text{for} \quad \mu(S_1)+\ldots+\mu(S_{i-1}) < \omega \leq \mu(S_1)+\ldots+\mu(S_i) \tag{14}$$

$(i = 1, 2, \ldots)$ where x_i is an inner point of S_i, then

(15) $$\underset{\sim}{P}\{\xi^{(1)}(\omega) \in S\} = \mu(S)$$ 10

whenever S belongs to the σ-algebra generated by $\{S_i\}$.

Now for each $i = 1,2,\ldots$ let $S_i = S_{i1} + S_{i2} + \ldots$ where $S_{i1}, S_{i2},\ldots.$ are disjoint sets belonging to F and define

(16) $$\xi^{(2)}(\omega) = x_{ij} \quad \text{for} \quad \mu(S_1)+\ldots+\mu(S_{i-1}) + \mu(S_{i,1}) +\ldots+ \mu(S_{i,j-1}) < \omega \leqq \mu(S_1)+\ldots+\mu(S_{i-1}) + \mu(S_{i,1})+\ldots+ \mu(S_{i,j})$$ 10

where x_{ij} is an inner point of S_{ij} . We have

(17) $$\underset{\sim}{P}\{\xi^{(2)}(\omega) \in S\} = \mu(S)$$ 10

whenever S belongs to the σ-algebra generated by $\{S_{ij}\}$.

By repeating the above procedure countably infinitely many times we can define a sequence of functions $\xi^{(1)}(\omega), \xi^{(2)}(\omega),\ldots$ on the interval $(0, 1)$. Denote by $d_1, d_2,\ldots$ the suprema of the diameters of the sets $\{S_i\}$, $\{S_{ij}\},\ldots$ respectively. Obviously,

(18) $$d(\xi^{(r)}(\omega), \xi^{(r+m)}(\omega)) < d_r$$

for m = 1,2,... . If we choose the partitions $\{S_i\}$, $\{S_{ij}\}$,... in such a way that $\lim_{r \to \infty} d_r = 0$, then

(19) $$\lim_{r \to \infty} \xi^{(r)}(\omega) = \xi(\omega)$$

exists for $\omega \in \Omega$ since X is a complete metric space. It is easy to see that $\xi(\omega)$ is a random variable and

$$(20) \qquad \mathbf{P}\{\xi(\omega) \in S\} = \mu(S)$$

for $S \in F$.

3. Weak Convergence of Probability Measures. Let $F_1(x), F_2(x), \ldots, F_n(x), \ldots$ and $F(x)$ be one-dimensional distribution functions. We say that the sequence of distribution functions $\{F_n(x)\}$ converges weakly to the distribution function $F(x)$ if

$$(1) \qquad \lim_{n \to \infty} F_n(x) = F(x)$$

in every continuity point of $F(x)$. In this case we write that $F_n(x) \Rightarrow F(x)$ as $n \to \infty$.

Let $\mu_n(A)$ be the probability measure induced by $F_n(x)$ and let $\mu(A)$ be the probability measure induced by $F(x)$. The set functions $\mu_n(A)$ $(n = 1,2,\ldots)$ and $\mu(A)$ are uniquely determined by $F_n(x)$ $(n = 1,2,\ldots)$ and $F(x)$ for each linear Borel set A .

For any set A let us denote by $A^{(c)}$ the closure of A , that is, $A^{(c)}$ is the set of limit points and isolated points of A , and let us denote by $A^{(i)}$ the interior of A , that is, $A^{(i)}$ is the set of interior points of A . Obviously $A^{(i)} \subset A \subset A^{(c)}$.

If $\mu(A^{(c)}) = \mu(A^{(i)})$ for a linear Borel set A , then we say that A is a continuity set of the measure μ .

We can easily see that $F_n(x) \Rightarrow F(x)$ if and only if

(2) $$\lim_{n \to \infty} \mu_n(A) = \mu(A)$$

for every continuity Borel set of μ or equivalently for every continuity interval of μ . In this case we write $\mu_n \Longrightarrow \mu$ and say that the measures μ_n converge weakly to the measure μ .

We note that in general (1) does not imply that (2) holds for any Borel set A . For example, let us assume that

(3) $$F_n(x) = \begin{cases} 0 & \text{for } x \leqq 0 , \\ \dfrac{[nx]}{n} & \text{for } 0 \leqq x \leqq 1 , \\ 1 & \text{for } x \geqq 1 , \end{cases}$$

and

(4) $$F(x) = \begin{cases} 0 & \text{for } x \leqq 0 , \\ x & \text{for } 0 \leqq x \leqq 1 , \\ 1 & \text{for } x \geqq 1 , \end{cases}$$

then

(5) $$\lim_{n \to \infty} F_n(x) = F(x)$$

for every x , that is, $F_n(x) \Longrightarrow F(x)$ as $n \to \infty$; however, if A denotes the set of irrational numbers in the interval $(0, 1)$, then $\mu_n(A) = 0$ for all $n = 1,2,\ldots$ whereas $\mu(A) = 1$.

By Theorem 41.8 we can easily conclude that $\mu_n \Longrightarrow \mu$ if and only if

(6) $$\lim_{n\to\infty} \int_{-\infty}^{\infty} h(x)d\mu_n = \int_{-\infty}^{\infty} h(x)d\mu$$

holds for every continuous and bounded real function $h(x)$ on the interval $(-\infty, \infty)$.

We can extend the notion of weak convergence of probability measures to more general spaces than the one discussed above. We can replace the real line by a finite dimensional Euclidean space or by a metric space. Now we shall consider the latter one.

Let X be a metric space and denote by F the class of Borel sets in X. Let $\mu_n(A)$ $(n = 1,2,\ldots)$ and $\mu(A)$ be probability measures defined for $A \in F$.

We say that μ_n converges weakly to μ, that is, $\mu_n \Rightarrow \mu$ as $n \to \infty$, if and only if

$$\lim_{n\to\infty} \int_X h(x)d\mu_n = \int_X h(x)d\mu \tag{7}$$

for every continuous and bounded real function $h(x)$ on X. The function $h(x)$ is continuous on X if for every $x \in X$ and for every $\varepsilon > 0$ there is a $\delta > 0$ such that $|h(x) - h(y)| < \varepsilon$ whenever $y \in X$ and $d(x, y) < \delta$. Here $d(x, y)$ denotes the metric in X.

For any set $A \in F$ denote by $A^{(c)}$ the closure of A and by $A^{(i)}$ the interior of A. If $\mu(A^{(c)}) = \mu(A^{(i)})$ for a set $A \in F$, then we say that A is a continuity set of the measure μ.

<u>Theorem 1.</u> <u>Let</u> X <u>be a metric space.</u> <u>Let</u> F <u>be the σ-algebra of Borel subsets of</u> X. <u>Let</u> $\mu_n(A)$ $(n = 1,2,\ldots)$ <u>and</u> $\mu(A)$ <u>be probability measures defined for</u> $A \in F$. <u>The measure</u> μ_n <u>converges weakly to the measure</u> μ <u>if and only if</u>

$$\lim_{n\to\infty} \mu_n(A) = \mu(A) \tag{8}$$

<u>for every continuity set</u> A <u>of the measure</u> μ .

<u>Proof.</u> If $\mu_n \Rightarrow \mu$, then for every $A \in F$ and for every $\varepsilon > 0$ we can find a continuous nonnegative function $h(x)$ such that $h(x) = 1$ for $x \in A^{(c)}$ and

(9) $$\mu(A^{(c)}) \geqq \int_X h(x)d\mu - \varepsilon .$$

Hence we have

(10) $$\mu(A^{(c)}) \geqq \int_X h(x)d\mu - \varepsilon = \lim_{n \to \infty} \int_X h(x)d\mu_n - \varepsilon \geqq \limsup_{n \to \infty} \mu_n(A) - \varepsilon$$

for any $\varepsilon > 0$. This implies that

(11) $$\limsup_{n \to \infty} \mu_n(A) \leqq \mu(A^{(c)})$$

for any $A \in F$. By (11) we can conclude that

(12) $$\mu(A^{(i)}) \leqq \liminf_{n \to \infty} \mu_n(A) \leqq \limsup_{n \to \infty} \mu_n(A) \leqq \mu(A^{(c)})$$

holds for every $A \in F$. If we replace A by $X - A$ in (11), then we obtain the first half of (12). The second half is precisely (11). If A is a continuity set of μ , then (12) implies (8).

Now let us prove the converse statement, that is, that (8) implies

(13) $$\lim_{n \to \infty} \int_X h(x)d\mu_n = \int_X h(x)d\mu$$

for any continuous and bounded real function $h(x)$ on X . Since the set of points $\{c\}$ for which $\mu\{x : h(x) = c\} > 0$ is atmost countable, it follows that for any $\varepsilon > 0$ we can find a finite number of points $c_0, c_1, \ldots, c_m$ such that $0 < c_i - c_{i-1} < \varepsilon$ for $i = 1,2,\ldots, m$, $c_0 < h(x) < c_m$ for every x, and each set $C_i = \{ x : c_{i-1} < h(x) \leqq c_i \}$ is a continuity set of μ. Then

$$(14) \qquad \limsup_{n\to\infty} \left| \int_X h(x)d\mu_n - \int_X h(x)d\mu \right| \leqq \varepsilon \sum_{i=1}^{m} \mu(C_i) \leqq \varepsilon .$$

Since $\varepsilon > 0$ is arbitrary this proves (13).

A sequence of measures $\{\mu_n\}$ is called weakly compact if every subsequence of $\{\mu_n\}$ contains a subsequence which is weakly convergent.

The following theorem was found in 1956 by Yu. V. Prokhorov [25]. See also I. I. Gikhman and A. V. Skorokhod [10 pp. 441-446], and P. Billingsley [3 pp. 35-40], [4].

Theorem 2. Let X be a metric space. Let F be the σ-algebra of Borel subsets of X. Let $\{\mu_n\}$ be a sequence of probability measures on F. If for every $\varepsilon > 0$ there exists a compact set K in X such that

$$(15) \qquad \sup_{1\leqq n<\infty} \mu_n(X-K) < \varepsilon ,$$

then $\{\mu_n\}$ is weakly compact.

Proof. We recall that a set K in the metric space X is compact if every open covering of K contains a finite subclass which is also a covering of K, or equivalently, if every sequence of elements in K

contains a subsequence which converges to some $x \varepsilon K$.

First, we shall prove the theorem in the case where X is a compact space. If X is a compact space, then it is complete and separable. Let $x_1, x_2, \ldots \quad , x_\ell, \ldots$ be a countable everywhere dense set in X .
Denote by R the set of positive rational numbers. Let A be the class of sets which can be represented as finite unions of (disjoint) open spheres $S(x_\ell, r)$ with center x_ℓ $(\ell = 1,2,\ldots)$ and radius $r \varepsilon R$. The class A is countable and F is the smallest σ-algebra which contains A .

By using the diagonal method we can easily prove that every infinite subsequence of $\{\mu_n\}$ contains a subsequence $\{\mu_{n_k}\}$ such that the limit

$$\lim_{k \to \infty} \mu_{n_k}(A) = \overline{\mu}(A) \tag{16}$$

exists for all $A \varepsilon A$.

We can easily see that if $A \varepsilon A$ and $B \varepsilon A$, then $A+B \varepsilon A$ and $\overline{\mu}(A+B) \leqq \overline{\mu}(A) + \overline{\mu}(B)$. If $AB = \theta$, then $\overline{\mu}(A+B) = \overline{\mu}(A) + \overline{\mu}(B)$, and if $A \subset B$, then $\overline{\mu}(A) \leqq \overline{\mu}(B)$. Furthermore, $X \varepsilon A$ and $\overline{\mu}(X) = 1$.

Denote by A^* the set of closed subsets of X and define

$$\mu(A) = \inf\{\sum_{m=1}^{\infty} \overline{\mu}(A_m) : A \subset \sum_{m=1}^{\infty} A_m \text{ and } A_m \varepsilon A\} \tag{17}$$

for $A \varepsilon A^*$. The class A^* is obviously an algebra.

The set function $\mu(A)$ defined on A^* satisfies the following properties: (a) $\mu(A) \geq 0$. (b) If $A \subset B$, then $\mu(A) \leq \mu(B)$. (c) If $A \subset \sum_{n=1}^{\infty} A_n$, then $\mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n)$. These properties immediately follow from the definition (17). For details see the proof of Theorem 1.2 in the Appendix.

Now we shall prove that if $A \in \mathcal{A}^*$ and $B \in \mathcal{A}^*$ and $AB = \Theta$, then

$$\mu(A+B) = \mu(A) + \mu(B) \; . \tag{18}$$

By property (c), it is sufficient to prove that

$$\mu(A+B) \geq \mu(A) + \mu(B) \; . \tag{19}$$

If we suppose that $A \subset \sum_{m=1}^{\infty} A_m$ and $B \subset \sum_{m=1}^{\infty} B_m$ where $A_m \in \mathcal{A}$, $B_m \in \mathcal{A}$ and $A_m B_m = \Theta$ for $m = 1,2,\ldots$, then by (16) we have

$$\overline{\mu}(A_m + B_m) = \overline{\mu}(A_m) + \overline{\mu}(B_m) \tag{20}$$

and thus

$$\inf\{\sum_{m=1}^{\infty} \overline{\mu}(A_m + B_m)\} \geq \inf\{\sum_{m=1}^{\infty} \overline{\mu}(A_m)\} + \inf\{\sum_{m=1}^{\infty} \overline{\mu}(B_m)\} \; . \tag{21}$$

It is not difficult to see that in defining $\mu(A+B)$ by (17) we can restrict ourself to such sums which occur on the left-hand side of (21). By (21) we get (19).

More generally we can prove that if $A_k \in \mathcal{A}^*$ for $k = 1,2,\ldots, n$ and $A_j A_k = \Theta$ for $j \neq k$, then

(22) $$\mu(A_1+A_2+\dots+A_n) = \mu(A_1)+\mu(A_2)+\dots+\mu(A_n) .$$

For $n = 2$, (22) is true by (18), and by mathematical induction it follows that (22) is true for every $n = 2,3,\dots$.

By the above properties it follows that $\mu(A)$ is a nonnegative and σ-additive set function on A^*. To prove this let us suppose that

(23) $$A = \sum_{j=1}^{\infty} A_j$$

where $A \in A^*$, $A_j \in A^*$ for $j = 1,2,\dots$ and $A_jA_k = \Theta$ for $j \neq k$. Then by (c) we have

(24) $$\mu(A) \leqq \sum_{j=1}^{\infty} \mu(A_j) .$$

Since $\sum_{j=1}^{n} A_j \subset A$, by (b) and by (22) we have

(25) $$\sum_{j=1}^{n} \mu(A_j) = \mu\left(\sum_{j=1}^{n} A_j\right) \leqq \mu(A)$$

for $n = 1,2,\dots$. If we let $n \to \infty$ in (25), we get

(26) $$\sum_{j=1}^{\infty} \mu(A_j) \leqq \mu(A) .$$

A comparison of (24) and (26) shows that

(27) $$\mu(A) = \sum_{j=1}^{\infty} \mu(A_j)$$

which was to be proved.

We observe that if $A \in A^*$, then

(28) $$\mu(A) \geqq \limsup_{k\to\infty} \mu_{n_k}(A) .$$

To prove this let us suppose that $A \subset \sum_{m=1}^{\infty} A_m$ where $A_m \varepsilon \mathcal{A}$. Since A is a compact set and A_m $(m = 1,2,\ldots)$ are open sets, there is a finite n such taht $A \subset \sum_{m=1}^{n} A_m$ also holds. Thus it follows that

$$(29) \qquad \begin{aligned} &\sum_{m=1}^{\infty} \bar{\mu}(A_m) \geqq \sum_{m=1}^{n} \bar{\mu}(A_m) = \lim_{k\to\infty} \sum_{m=1}^{n} \mu_{n_k}(A_m) = \\ &\geqq \lim_{k\to\infty} \mu_{n_k}\left(\sum_{m=1}^{n} A_m\right) \geqq \limsup_{k\to\infty} \mu_{n_k}(A) \end{aligned}$$

always holds. Hence by (17) we obtain (28).

If, in particular, $A = X$ in (28), then we obtain that $\mu(X) \geqq 1$. On the other hand it follows from (17) that $\mu(A) \leqq 1$ for any $A \varepsilon \mathcal{A}^*$. This implies that $\mu(X) = 1$.

Accordingly, we proved that $\mu(A)$ is a probability measure on the algebra $\mathcal{A}^*$. By Theorem 1.2 in the Appendix we can uniquely extend the definition of $\mu(A)$ to the σ-algebra $\mathcal{F}$ in such a way that $\mu(A)$ remains a probability measure.

By (28), for any set $A \varepsilon \mathcal{F}$ we have

$$(30) \qquad \mu(A^{(c)}) \geqq \limsup_{k\to\infty} \mu_{n_k}(A^{(c)}) \geqq \limsup_{k\to\infty} \mu_{n_k}(A) .$$

If we apply (30) to the set $X - A^{(i)}$, then we obtain that

$$(31) \qquad \mu(A^{(i)}) \leqq \liminf_{k\to\infty} \mu_{n_k}(A^{(i)}) \leqq \liminf_{k\to\infty} \mu_{n_k}(A) .$$

By (30) and (31) it follows that if $A \in F$ is a continuity set of the probability measure μ, that is, $\mu(A) = \mu(A^{(c)}) = \mu(A^{(i)})$, then

$$\lim_{k \to \infty} \mu_{n_k}(A) = \mu(A) . \tag{32}$$

This proves that if X is a compact metric space, F is the σ-algebra of Borel subsets of X, and $\{\mu_n\}$ is a sequence of probability measures defined on F, then $\{\mu_n\}$ is weakly compact. By a slight change of the above proof we can see that the last statement remains valid unchangeably if instead of $\mu_n(X) = 1$ we assume only $\sup_{1 \leq n < \infty} \mu_n(X) < \infty$ for the sequence of measures $\{\mu_n\}$.

By using the above result we can easily prove Theorem 2 in the general case. Accordingly, let us assume that X is an arbitrary metric space and that (15) is satisfied. In the following proof we may assume that μ is not necessarily a probability measure, but an arbitrary measure for which

$$\sup_{1 \leq n < \infty} \mu_n(X) < \infty . \tag{33}$$

Let us choose a sequence of compact sets $K_1, K_2, \ldots, K_r, \ldots$ in X in such a way that $K_1 \subset K_2 \subset \ldots \subset K_r \subset \ldots$ and

$$\sup_{1 \leq n < \infty} \mu_n(X - K_r) < \frac{1}{r} \tag{34}$$

for $r = 1, 2, \ldots$. By the previous results we can conclude that the sequence of measures $\mu_n(AK_r)$ $(n = 1, 2, \ldots)$ defined for $A \in F$ is weakly compact, that is, there is a measure $\mu^{(r)}(A)$ and a sequence of positive integers $n_k^{(r)}$ $(k = 1, 2, \ldots)$ such that

(35) $$\lim_{k \to \infty} \mu_{n_k^{(r)}}(AK_r) = \mu^{(r)}(A)$$

for every continuity set of $\mu^{(r)}$. Let us choose the sequences $\{n_k^{(r)}\}$ in such a way that $\{n_k^{(r+1)}\}$ is a subsequence of $\{n_k^{(r)}\}$ for each $r = 1,2,\ldots$. In this case, if $s \geqq r$, then the measures $\mu^{(s)}$ and $\mu^{(r)}$ coincide on K_r, that is,

(36) $$\mu^{(s)}(AK_r) = \mu^{(r)}(AK_r)$$

for $A \varepsilon F$. Since

(37) $$|\mu^{(s)}(A)-\mu^{(r)}(A)| \leqq |\mu^{(s)}(AK_r)-\mu^{(r)}(AK_r)| + \mu^{(s)}(X-K_r) + \mu^{(r)}(X-K_r) ,$$

it follows that

(38) $$|\mu^{(s)}(A) - \mu^{(r)}(A)| < \frac{2}{r}$$

for $A \varepsilon F$ and $s \geqq r$. Thus the limit

(39) $$\lim_{s \to \infty} \mu^{(s)}(A) = \mu(A)$$

exists for $A \varepsilon F$ and $\mu(A)$ is a measure on F. By (38) we have

(40) $$|\mu(A) - \mu^{(r)}(A)| < \frac{2}{r}$$

for $A \varepsilon F$ and $r = 1,2,\ldots$. Furthermore, by (36) we get that

(41) $$\mu(AK_r) = \mu^{(r)}(AK_r)$$

for $A \varepsilon F$ and $r = 1,2,\ldots$.

Now we shall prove that if $A \in F$ and if A is a continuity set of $\mu(A)$, then

(42) $$\lim_{k \to \infty} \mu_{n_k^{(k)}}(A) = \mu(A) .$$

If A is a continuity set of $\mu(A)$, then

(43) $$\lim_{k \to \infty} \mu_{n_k^{(s)}} (AK_r) = \mu^{(s)}(AK_r) = \mu^{(r)}(AK_r) = \mu(AK_r)$$

for $s \geqq r$. Hence by the diagonal method we obtain that

(44) $$\lim_{k \to \infty} \mu_{n_k^{(k)}} (AK_r) = \mu(AK_r)$$

for $r = 1,2,\ldots$. Since

(45) $$|\mu_n(AK_r) - \mu_n(A)| \leqq \mu_n(X-K_r) < \frac{1}{r}$$

for all $n = 1,2,\ldots$ and $r = 1,2,\ldots$,and since

(46) $$|\mu(AK_r) - \mu(A)| \leqq \mu(X-K_r) < \frac{1}{r}$$

for $r = 1,2,\ldots$, it follows from (44) that

(47) $$\limsup_{k \to \infty} |\mu_{n_k^{(k)}}(A) - \mu(A)| \leqq \frac{2}{r}$$

for $r = 1,2,\ldots$. If $r \to \infty$, then we obtain (39). Accordingly, the sequence of measures $\{\mu_n\}$ is weakly compact. This completes the proof of the theorem.

In conclusion we mention a related theorem which has several useful applications in weak convergence of probability measures.

Theorem 3. Let X be a separable metric space with distance function $d(x, y)$. Let F be the σ-algebra of Borel subsets of X. Let $\mu_1, \mu_2, \ldots, \mu_n, \ldots$ and μ be probability measures on F. If $\mu_n \Longrightarrow \mu$ as $n \to \infty$, then there exists a probability space $(\Omega, \mathcal{B}, P)$ and random variables $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ and ξ taking values in the space X such that

(48) $$P\{\xi_n \in A\} = \mu_n(A) \quad \text{and} \quad P\{\xi \in A\} = \mu(A)$$

for $A \in F$, and

(49) $$P\{\lim_{n \to \infty} d(\xi_n, \xi) = 0\} = 1 .$$

This theorem was proved in 1956 by A. V. Skorokhod [29 p. 281] in the case when X is a complete and separable metric space. R. M. Dudley [9] demonstrated that is valid for separable metric spaces X. See also R. Pyke [26]. We can easily prove Theorem 3 by using the construction in the proof of Theorem 2.3 in the Appendix.

4. Product Probability Spaces. In defining independent random trials we need the notion of product probability spaces. To introduce this notion let us consider a family of random trials $\mathfrak{S}_t$ defined for $t \in T$ where T is a parameter set. Let $(\Omega_t, \mathcal{B}_t, P_t)$ be the probability space associated with the random trial $\mathfrak{S}_t$.

Denote by $\mathfrak{S}_T$ the compound random trial which consists of the performance of all the random trials $\mathfrak{S}_t$ for $t \in T$. Let $(\Omega_T, \mathcal{B}_T, P_T)$ be the probability space associated with the compound random trial $\mathfrak{S}_T$.

We can define the <u>sample space</u> Ω_T in the following way. Since every outcome of the compound random trial $\mathcal{G}_T$ can be represented by the sample element (point) $\omega_T = \{\omega_t, t \in T\}$ where $\omega_t \in \Omega_t$ we may assume that

(1) $$\Omega_T = \{\omega_T : \omega_T = \{\omega_t, t \in T\} , \omega_t \in \Omega_t\} .$$

We shall write

(2) $$\Omega_T = \underset{t \in T}{X} \Omega_t$$

and call Ω_T the <u>product sample space.</u>

If $A_t \subset \Omega_t$ for ~~the~~ $t \in T$, then we define

(3) $$A_T = \underset{t \in T}{X} A_t$$

as the set of sample elements $\omega_T = \{\omega_t, t \in T\}$ such that $\omega_t \in A_t$ for $t \in T$. The set A_T will be called a product set in Ω_T .

Let $T_n = (t_1, t_2, \ldots, t_n)$ be a finite subset of the parameter set T .

We say that A_T is a <u>product cylinder</u> with sides $A_{t_1}, A_{t_2}, \ldots, A_{t_n}$ if

(4) $$A_T = A_{t_1} X A_{t_2} X \ldots X A_{t_n} X \Omega_{T-T_n}$$

where $A_{t_i} \subset \Omega_{t_i}$ for $i = 1,2,\ldots, n$. If $A_{t_i} \in \mathcal{B}_{t_i}$ for $i = 1,2,\ldots, n$, then we say that A_T is a <u>measurable product cylinder.</u>

Next, let us define $\mathcal{B}_T$, the <u>class of random events</u> in the compound random trial $\mathfrak{S}_T$. Let C_T be the class of all measurable product cylinders in Ω_T . Denote by A_T the class of all finite unions of disjoint measurable product cylinders in Ω_T . If $A_T \varepsilon A_T$ and $B_T \varepsilon A_T$, then $A_T B_T \varepsilon A_T$ and $A_T - B_T \varepsilon A_T$. Thus A_T is an algebra of sets.

Let us suppose that $\mathcal{B}_T$ is the minimal σ-algebra which contains all the sets in A_T . The σ-algebra $\mathcal{B}_T$ is called the <u>product σ-algebra</u> and is denoted by

(5) $$\mathcal{B}_T = \underset{t \varepsilon T}{X} \mathcal{B}_t \ .$$

<u>Definition. Let $A_T \subset \Omega_T$ and $T_n = (t_1, t_2, \ldots, t_n)$ be a finite subset of T . We define</u>

(6) $$A_T(\omega_{T_n}) = \{\omega_{T-T_n} : \omega_T \varepsilon A_T\}$$

<u>as the section of A_T at ω_{T_n}</u> .

<u>Theorem 1. If $A_T \varepsilon \mathcal{B}_T$, then $A_T(\omega_{T_n}) \varepsilon \mathcal{B}_{T-T_n}$</u> .

<u>Proof</u>. If $A_T \varepsilon C_T$, then obviously $A_T(\omega_{T_n}) \varepsilon C_{T-T_n}$. This implies that if $A_T \varepsilon A_T$, then $A_T(\omega_{T_n}) \varepsilon A_{T-T_n}$ and consequently $A_T(\omega_{T_n}) \varepsilon \mathcal{B}_{T-T_n}$.

Let

(7) $$\mathcal{S} = \{A_T : A_T \varepsilon \mathcal{B}_T \text{ and } A_T(\omega_{T_n}) \varepsilon \mathcal{B}_{T-T_n}\}$$

for some fixed $\omega_{T_n} \varepsilon \Omega_{T_n}$. Evidently $A_T \subset \mathcal{S} \subset \mathcal{B}_T$. Now we shall prove that S is a σ-algebra. Since $\mathcal{B}_T$ is the minimal σ-algebra over $\mathcal{A}_T$, it follows that $S = \mathcal{B}_T$ which implies the theorem.

First, if $A_T \in S$, then $A_T(\omega_{T_n}) \in \mathcal{B}_{T-T_n}$. Thus $\overline{A}_T(\omega_{T_n}) \in \mathcal{B}_{T-T_n}$ and this implies that $\overline{A}_T \in S$.

Second, if $A_T = \sum_{k=1}^{\infty} A_T^{(k)}$ where $A_T^{(k)} \in S$, then $A_T^{(k)}(\omega_{T_n}) \in \mathcal{B}_{T-T_n}$. Thus

$$A_T(\omega_T) = \sum_{k=1}^{\infty} A_T^{(k)}(\omega_T) \in \mathcal{B}_{T-T_n} . \tag{8}$$

Consequently, S is indeed a σ-algebra. This completes the proof of the theorem.

It remains to define the probability $\underset{\sim}{P}_T\{A\}$ for $A \in \mathcal{B}_T$. We can define probabilities in various ways on $\mathcal{B}_T$, but the so-called product probabilities have a special importance. Now we are going to define this notion.

Let $T_n = (t_1, t_2, \ldots, t_n)$ be a finite subset of T and let

$$A_T = A_{t_1} \times A_{t_2} \times \ldots \times A_{t_n} \times \Omega_{T-T_n} \tag{9}$$

be a measurable product cylinder. Define

$$\underset{\sim}{P}_T\{A_T\} = \underset{\sim}{P}_{t_1}\{A_{t_1}\}\underset{\sim}{P}_{t_2}\{A_{t_2}\} \ldots \underset{\sim}{P}_{t_n}\{A_{t_n}\} \tag{10}$$

for any $A_T \in \mathcal{C}_T$.

If $A_T \varepsilon \mathcal{A}_T$, then

$$A_T = \sum_{k=1}^{m} A_T^{(k)} \tag{11}$$

where $A_T^{(k)}$ $(k = 1,2,\ldots, m)$ are disjoint sets belonging to C_T . Define

$$\underset{\sim}{P}_T\{A_T\} = \sum_{k=1}^{m} \underset{\sim}{P}_T\{A_T^{(k)}\} \tag{12}$$

for $A_T \varepsilon \mathcal{A}_T$. We can easily see that $\underset{\sim}{P}_T\{A_T\}$ is independent of the particular representation (11), it is uniquely determined by A_T .

The set function $\underset{\sim}{P}_T\{A_T\}$ is finitely additive on $\mathcal{A}_T$, $\underset{\sim}{P}_T\{A_T\} \geqq 0$ and $\underset{\sim}{P}_T\{\Omega_T\} = 1$.

<u>Theorem 2.</u> <u>The set function</u> $\underset{\sim}{P}_T\{A_T\}$ <u>is</u> σ-<u>additive on the algebra</u> $\mathcal{A}_T$.

<u>Proof.</u> We shall prove that $\underset{\sim}{P}_T\{A_T\}$ is continuous at Θ , that is, if $A_T^{(k)} \varepsilon \mathcal{A}_T$ for $k = 1,2,\ldots,$ $A_T^{(1)} \supset A_T^{(2)} \supset \cdots \supset A_T^{(k)} \supset \cdots,$ and $\lim_{k\to\infty} A_T^{(k)} = \prod_{k=1}^{\infty} A_T^{(k)} = \Theta$, then $\lim_{k\to\infty} \underset{\sim}{P}_T\{A_T^{(k)}\} = 0$. Equivalently, we can prove that if $\lim_{k\to\infty} \underset{\sim}{P}_T\{A_T^{(k)}\} > 0$, then $\prod_{k=1}^{\infty} A_T^{(k)}$ is not empty. Finite additivity and continuity imply σ-additivity on $\mathcal{A}_T$.

Accordingly, let us assume that $A_T^{(k)} \varepsilon \mathcal{A}_T$ for $k = 1,2,\ldots,$ $A_T^{(1)} \supset A_T^{(2)} \supset \cdots \supset A_T^{(k)} \supset \cdots$ and

$$\underset{\sim}{P}_T\{A_T^{(k)}\} \geqq \varepsilon > 0 \tag{13}$$

for $k = 1,2,\ldots$. We shall prove that $\prod_{k=1}^{\infty} A_T^{(k)}$ is not empty.

For each $k = 1,2,\ldots$ we can write that

(14) $$A_T^{(k)} = A_{T_k^*} \times \Omega_{T-T_k^*}$$

where $A_{T_k^*} \in \mathcal{A}_{T_k^*}$ and T_k^* is a finite subset of T. Let

(15) $$T^* = \bigcup_{k=1}^{\infty} T_k^* .$$

Then T^* is a countable set, and we can write that

(16) $$A_T^{(k)} = A_{T^*}^{(k)} \times \Omega_{T-T^*}$$

where $A_{T^*} \in \mathcal{A}_{T^*}$. (If $T^* = T$, then $A_T^{(k)} = A_{T^*}^{(k)}$.)

Let $T^* = \{t_1, t_2, \ldots\}$. Thus it is sufficient to prove that if $A_{T^*}^{(k)} \in \mathcal{A}_{T^*}$ for $k = 1,2,\ldots$, $A_{T^*}^{(1)} \supset A_{T^*}^{(2)} \supset \ldots \supset A_{T^*}^{(k)} \supset \ldots$ and

(17) $$\underset{\sim}{P}_{T^*}\{A_{T^*}^{(k)}\} \geqq \varepsilon > 0$$

for $k = 1,2,\ldots$, then $\prod_{k=1}^{\infty} A_{T^*}^{(k)}$ is not empty, that is, there exists an element (point)

(18) $$\bar{\omega}_{T^*} = \{\bar{\omega}_{t_1}, \bar{\omega}_{t_2}, \ldots\} \in A_{T^*}^{(k)}$$

for all $k = 1,2,\ldots$.

If $T^* = \{t_1\}$, then the statement is trivial because $\underset{\sim}{P}_{t_1}\{A\}$ is σ-additive on $\mathcal{A}_{t_1}$. Thus let us suppose that T^* contains more than one element.

Let

$$B_{T*}^{(k)} = \{\omega_{t_1} : \underset{\sim}{P}_{T*-\{t_1\}}\{A_{T*}^{(k)}(\omega_{t_1})\} > \frac{\varepsilon}{2}\}. \tag{19}$$

Since $B_{T*}^{(k)}$ is the finite union of sets belonging to A_{t_1}, therefore $B_{T*}^{(k)} \varepsilon A_{t_1}$. We have

$$0 < \varepsilon < \underset{\sim}{P}_{T*}\{A_{T*}^{(k)}\} = \int_{\Omega_{t_1}} \underset{\sim}{P}_{T*-\{t_1\}}\{A_{T*}^{(k)}(\omega_{t_1})\}d\underset{\sim}{P}_{t_1} \leqq \tag{20}$$

$$\leqq [1 - \underset{\sim}{P}_{t_1}\{B_{T*}^{(k)}\}]\frac{\varepsilon}{2} + \underset{\sim}{P}_{t_1}\{B_{T*}^{(k)}\} .$$

Hence

$$\underset{\sim}{P}_{t_1}\{B_{T*}^{(k)}\} > \frac{\varepsilon}{2} \tag{21}$$

for $k = 1,2,\ldots$.

Accordingly, $B_{T*}^{(k)} \varepsilon A_{t_1}$ for $k = 1,2,\ldots$, $B_{T*}^{(1)} \supset B_{T*}^{(2)} \supset \ldots \supset B_{T*}^{(k)} \supset \ldots$ and since $\underset{\sim}{P}_{t_1}\{A\}$ is σ-additive on A_{t_1}, by (21) we have

$$\lim_{k \to \infty} P_{t_1}\{B_{T*}^{(k)}\} = P_{t_1}\{\prod_{k=1}^{\infty} B_{T*}^{(k)}\} \geqq \frac{\varepsilon}{2} > 0 . \tag{22}$$

Consequently, $\prod_{k=1}^{\infty} B_{T*}^{(k)}$ is not empty, that is, there is a point $\bar{\omega}_{t_1} \varepsilon B_{T*}^{(k)} \subset \Omega_{t_1}$ for $k = 1,2,\ldots$ and

$$\underset{\sim}{P}_{T*-\{t_1\}}\{A_{T*}^{(k)}(\bar{\omega}_{t_1})\} \geqq \frac{\varepsilon}{2} > 0 \tag{23}$$

for all $k = 1,2,\ldots$.

If T^* contains more than two elements, if we replace $A_{T*}^{(k)}$ by

$A_{T^*}^{(k)}(\bar{\omega}_{t_1})$ and if we take into consideration that $\underset{\sim}{P}_{t_2}\{A\}$ is σ-additive on A_{t_2}, then by repeating the previous argument we obtain that there exists a point $\bar{\omega}_{t_2} \in \Omega_{t_2}$ such that

$$(24) \qquad \underset{\sim}{P}_{T^*-\{t_1,t_2\}}\{A_{T^*}^{(k)}(\bar{\omega}_{t_1}, \bar{\omega}_{t_2})\} \geqq \frac{\varepsilon}{4} > 0 .$$

By continuing this procedure we obtain that as long as $(t_1, t_2, \ldots, t_n)$ is a proper finite subset of T^*, there exist points $\bar{\omega}_{t_1} \in \Omega_{t_1}, \ldots, \bar{\omega}_{t_n} \in \Omega_{t_n}$ such that

$$(25) \qquad \underset{\sim}{P}_{T^*-\{t_1,\ldots,t_n\}}\{A_{T^*}^{(k)}(\bar{\omega}_{t_1}, \ldots, \bar{\omega}_{t_n})\} \geqq \frac{\varepsilon}{2^n} > 0 .$$

If $T^* = \{t_1, t_2, \ldots, t_p\}$ where $p \geqq 2$, then (25) holds for $n = p - 1$. The sets $A_{T^*}^{(k)}(\bar{\omega}_{t_1}, \ldots, \bar{\omega}_{t_{p-1}}) \in A_{t_p}$ form a decreasing sequence of sets. Since $\underset{\sim}{P}_{t_p}\{A\}$ is σ-additive on A_{t_p} we have

$$(26) \qquad \lim_{k \to \infty} \underset{\sim}{P}_{t_p}\{A_{T^*}^{(k)}(\bar{\omega}_{t_1}, \ldots, \bar{\omega}_{t_{p-1}})\} = \underset{\sim}{P}_{t_p}\{\prod_{k=1}^{\infty} A_{T^*}^{(k)}(\bar{\omega}_{t_1}, \ldots, \bar{\omega}_{t_{p-1}})\} \geqq \frac{\varepsilon}{2^{p-1}} > 0 .$$

Consequently, there is an $\bar{\omega}_{t_p} \in \Omega_{t_p}$ such that $\bar{\omega}_{t_p} \in A_{T^*}^{(k)}(\bar{\omega}_{t_1}, \ldots, \bar{\omega}_{t_{p-1}})$, that is, $(\bar{\omega}_{t_1}, \ldots, \bar{\omega}_{t_p}) \in A_{T^*}^{(k)}$ for all $k = 1, 2, \ldots$. This completes the proof of the theorem in the case when T^* is a finite set.

If $T^* = \{t_1, t_2, \ldots\}$ is an infinite sequence, then (25) holds for $n = 1, 2, \ldots$, and

$$(\bar{\omega}_{t_1}, \bar{\omega}_{t_2}, \ldots) \in A^{(k)}_{T^*} \tag{27}$$

for all $k = 1,2,\ldots$. To prove this, for each $k = 1,2,\ldots$, let us write

$$A^{(k)}_{T^*} = A^{(k)}_{\{t_1,t_2,\ldots,t_n\}} \times \Omega_{t_{n+1}} \times \Omega_{t_{n+2}} \times \ldots \tag{28}$$

where n is some positive integer. By (25) $A^{(k)}_{T^*}(\bar{\omega}_{t_1},\ldots, \bar{\omega}_{t_n})$ is not empty, and by (28) it is necessarily equal to $\Omega_{t_{n+1}} \times \Omega_{t_{n+2}} \times \ldots$. Thus we have

$$(\bar{\omega}_{t_{n+1}}, \bar{\omega}_{t_{n+2}}, \ldots) \in A^{(k)}_{T^*}(\bar{\omega}_{t_1},\ldots, \bar{\omega}_{t_n}) \tag{29}$$

which implies (27). Accordingly (27) holds for all $k = 1,2,\ldots$. This completes the proof of the theorem.

By Theorem 1.2 in the Appendix we can extend the definition of $\underset{\sim}{P}_T\{A_T\}$ to the σ-algebra $\mathcal{B}_T$. The extension is unique and $\underset{\sim}{P}_T\{A_T\}$ is a probability measure on $\mathcal{B}_T$. We shall call $\underset{\sim}{P}_T\{A_T\}$ the <u>product probability measure</u> and $(\Omega_T, \mathcal{B}_T, \underset{\sim}{P}_T)$ the <u>product probability space</u>.

The product probability $\underset{\sim}{P}_T\{A_T\}$ has the property that for any finite set $T_n = (t_1, t_2,\ldots, t_n) \subset T$ and for any $A_{t_i} \in \mathcal{B}_{t_i}$ $(i = 1,2,\ldots, n)$

$$\underset{\sim}{P}_T\{A_{t_1} \times A_{t_2} \times \ldots \times A_{t_n} \times \Omega_{T-T_n}\} = \underset{\sim}{P}_{t_1}\{A_{t_1}\}\underset{\sim}{P}_{t_2}\{A_{t_2}\}\ldots \underset{\sim}{P}_{t_n}\{A_{t_n}\}. \tag{30}$$

Conversely, $\underset{\sim}{P}_T\{A_T\}$ is uniquely determined for $\mathcal{B}_T$ by this property.

In particular, we have

$$(31) \qquad \underset{\sim}{P}_T\{A_{t_i} \times \Omega_{T-\{t_1\}}\} = \underset{\sim}{P}_{t_i}\{A_{t_i}\}$$

for $A_{t_i} \varepsilon \mathcal{B}_{t_i}$ and thus (30) can be expressed as follows:

$$(32) \qquad \underset{\sim}{P}_T\{A_{t_1} \times \ldots \times A_{t_n} \times \Omega_{T-T_n}\} = \prod_{i=1}^{n} \underset{\sim}{P}_T\{A_{t_i} \times \Omega_{T-\{t_i\}}\}.$$

Accordingly, if the compound random trial $\mathfrak{S}_T$ is described by the product probability space $(\Omega_T, \mathcal{B}_T, \underset{\sim}{P}_T)$, then the following n events: $\mathfrak{S}_{t_1}$ results in $A_{t_1}, \ldots, \mathfrak{S}_{t_n}$ results in A_{t_n} will be mutually independent events in $\mathfrak{S}_T$ for any $n = 2,3,\ldots$ and $A_{t_i} \varepsilon \mathcal{B}_{t_i}$. Furthermore, any event $A_t \varepsilon \mathcal{B}_t$, where $t \varepsilon T$, has the same probability in the compound random trial $\mathfrak{S}_T$ as in the constituent random trial $\mathfrak{S}_t$.

If the probability space associated with the compound random trial $\mathfrak{S}_T$ is the product probability space $(\Omega_T, \mathcal{B}_T, \underset{\sim}{P}_T)$, then we say that the constituent random trials $\mathfrak{S}_t$ $(t \varepsilon T)$ are mutually independent.

Conversely, if we consider a family of random trials $\mathfrak{S}_t$ $(t \varepsilon T)$, and if any outcome of any random trial has no influence on the outcome of any other random trial, then we assume that the random trials are mutually independent and with the compound random trial we associate the product probability space.

Let $\mathfrak{S}_t = (\Omega_t, \mathcal{B}_t, \underset{\sim}{P}_t)$ $(t \varepsilon T)$ be a family of mutually independent random trials and let $\xi_t(\omega_t)$ $(t \varepsilon T)$ be random variables defined on $\mathfrak{S}_t$ $(t \varepsilon T)$. If we define the random variables $\xi_t(\omega_T) = \xi_t(\omega_t)$ for $t \varepsilon T$ on the product probability space $\mathfrak{S}_T = (\Omega_T, \mathcal{B}_T, \underset{\sim}{P}_T)$, then $\xi_t(\omega_T)$ $(t \varepsilon T)$

will be mutually independent random variables.

The existence of a product probability space for an arbitrary family of random trials was proved in 1939 by B. Jessen [12]. In the particular case when each Ω_t is the real line and $\mathcal{B}_t$ is the class of linear Borel sets, the existence of a product probability space follows from a more general theorem found in 1933 by A. N. Kolmogorov [19]. (Theorem 47.1 .) See also J. v. Neumann [24 pp. 122-148], S. Kakutani [13] and E. S. Andersen and B. Jessen [2]. Actually, the general existence theorem was stated in 1934 by Z. Łomnicki and S. Ulam [22], but their proof contains an error which was pointed out in 1943 by S. Kakutani [13], and in 1946 by E. S. Andersen and B. Jessen [2].

5. Conditional Probabilities and Conditional Expectations. The general notions of conditional probabilities and conditional expectations were introduced in 1933 by A. N. Kolmogorov [49] . These notions are based on the integral in an abstract space and on the Radon-Nikodym theorem.

Let $(\Omega, \mathcal{A}, \underset{\sim}{P})$ be a probability space, $A \in \mathcal{A}$, and $\xi(\omega)$, a real random variable defined on Ω . If the series

$$(1) \qquad \sum_{j=-\infty}^{\infty} j\lambda \underset{\sim}{P}\{j\lambda < \xi(\omega) \leq (j+1)\lambda \text{ and } \omega \in A\}$$

is absolutely convergent for some $\lambda > 0$, then it is convergent for every $\lambda > 0$ and has a finite limit as $\lambda \to 0$. This limit is, by definition, the integral of the random variable $\xi(\omega)$ over A , and is denoted by

$$(2) \qquad \underset{\sim}{Q}(A) = \int_A \xi(\omega) d\underset{\sim}{P} .$$

If $\underset{\sim}{Q}(\Omega)$ exists, then it is called the ecpectation of the random variable which we denote by $\underset{\sim}{E}\{\xi\}$.

If $\underset{\sim}{Q}(\Omega)$ exists, then $\underset{\sim}{Q}(A)$ exists for every $A \varepsilon \mathcal{A}$, and $\underset{\sim}{Q}(A)$ is a finite and σ-additive set function, that is, if $A = A_1 + A_2 + \ldots + A_k + \ldots$ where $A_k \varepsilon \mathcal{A}$ for $k = 1,2,\ldots$ and $A_i A_j = \Theta$ for $i \neq j$,then

$$(3) \qquad \underset{\sim}{Q}(A) = \sum_{k=1}^{\infty} \underset{\sim}{Q}(A_k)$$

and the right-hand side is absolutely convergent.

The set function $\underset{\sim}{Q}(A)$ is absolutely continuous with respect to $\underset{\sim}{P}\{A\}$, that is, if $A \varepsilon \mathcal{A}$ and $\underset{\sim}{P}\{A\} = 0$, then $\underset{\sim}{Q}(A) = 0$.

The celebrated Radon-Nikodym theorem states that if $\underset{\sim}{Q}(A)$ possesses the mentioned properties, then it can be represented in the form (2) and $\xi(\omega)$ is determined up to an equivalence. More precisely we have the following result.

<u>Theorem 1.</u> <u>Let</u> Ω <u>be an (abstract) set,</u> $\mathcal{A}$ <u>a</u> σ-<u>algebra of subsets of</u> Ω , $\underset{\sim}{P}\{A\}$, <u>a probability measure defined on</u> $\mathcal{A}$. <u>Let</u> $\underset{\sim}{Q}(A)$ <u>be a finite and</u> σ-<u>additive set function defined on</u> $\mathcal{A}$. <u>Suppose that</u> $\underset{\sim}{Q}(A)$ <u>is absolutely continuous with respect to</u> $\underset{\sim}{P}\{A\}$, <u>that is,</u> $A \varepsilon \mathcal{A}$ <u>and</u> $\underset{\sim}{P}\{A\} = 0$ <u>imply</u> $\underset{\sim}{Q}(A) = 0$. <u>Then there is a random variable</u> $\xi(\omega)$ <u>defined on</u> Ω , <u>and integrable over</u> Ω <u>such that</u>

$$(4) \qquad \underset{\sim}{Q}(A) = \int_A \xi(\omega) d\underset{\sim}{P}$$

<u>for all</u> $A \varepsilon \mathcal{A}$. <u>If</u> $\eta(\omega)$ <u>is another random variable which is integrable</u>

over Ω and for which

$$(5) \qquad Q(A) = \int_A \eta(\omega)dP ,$$

and if $D = \{\omega : \eta(\omega) \neq \xi(\omega)\}$, then $P\{D\} = 0$.

For the proof of this theorem we refer to P. Halmos [11]

Theorem 1 makes possible the following definitions.

Let $(\Omega, \mathcal{A}, P)$ be a probability space, $A \in \mathcal{A}$ an event, and $\mathcal{B}$ a σ-algebra of sets belonging to $\mathcal{A}$ (σ-subalgebra of $\mathcal{A}$) . The conditional probability of A relative to $\mathcal{B}$, denoted by $P\{A|\mathcal{B}\}$, is defined as any function of ω which is measurable with respect to $\mathcal{B}$ and which satisfies the equation

$$(6). \qquad \int_B P\{A|\mathcal{B}\}dP = P\{AB\}$$

for all $B \in \mathcal{B}$. By Theorem 1 it follows that such a function exists, $P\{A|\mathcal{B}\}$ is a random variable, and is determined up to an equivalence.

If $\chi = \chi(\omega)$ is a real and finite-valued random variable, then $P\{A|\chi\}$ is defined as any one version of $P\{A|\mathcal{B}\}$ where $\mathcal{B}$ is the σ-algebra generated by χ , that is, $\mathcal{B}$ is the smallest σ-algebra which contains the sets $\{\omega : \chi(\omega) \leq x\}$ for all x . In this case $P\{A|\chi\}$ is a Baire-function of χ and we use the notation $P\{A|\chi = x\} = P\{A|\chi\}\Big|_{\chi(\omega)=x}$.

The following formula

$$(7) \qquad P\{A\} = \int_{-\infty}^{\infty} P\{A|\chi = x\}dP\{\chi \leq x\}$$

is called the <u>theorem of total probability</u>.

Let $(\Omega, A, \underset{\sim}{P})$ be a probability space, η a real and finite-valued random variable whose expectation exists, and $\mathcal{B}$ a σ-algebra of sets belonging to A. The conditional expectation of η relative to $\mathcal{B}$, denoted by $\underset{\sim}{E}\{\eta|\mathcal{B}\}$, is defined as any function of ω which is measurable with respect to $\mathcal{B}$ and which satisfies the equation

$$(8) \qquad \int_B \underset{\sim}{E}\{\eta|\mathcal{B}\}d\underset{\sim}{P} = \int_B \eta d\underset{\sim}{P}$$

for all $B \in \mathcal{B}$. By Theorem 1 it follows that such a function exists, $\underset{\sim}{E}\{\eta|\mathcal{B}\}$ is a random variable, and is determined up to an equivalence.

If $\chi = \chi(\omega)$ is a real and finite-valued random variable, then $\underset{\sim}{E}\{\eta|\chi\}$ is defined as any one version of $\underset{\sim}{E}\{\eta|\mathcal{B}\}$ where $\mathcal{B}$ is the σ-algebra generated by χ. In this case $\underset{\sim}{E}\{\eta|\chi\}$ is a Baire-function of χ and we use the notation $\underset{\sim}{E}\{\eta|\chi = x\} = \underset{\sim}{E}\{\eta|\chi\}_{\chi(\omega)=x}$.

The following formula

$$(9) \qquad \underset{\sim}{E}\{\eta\} = \int_{-\infty}^{\infty} \underset{\sim}{E}\{\eta|\chi = x\}d\underset{\sim}{P}\{\chi \leq x\}$$

is called the <u>theorem of total expectation.</u>

6. <u>Wald's Theorem.</u> Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of real or complex random variables. Write $\zeta_n = \xi_1 + \xi_2 + \ldots + \xi_n$ for $n = 1, 2, \ldots$ and $\zeta_0 = 0$. The results of this section are concerned with the random variable ζ_ν where ν is a discrete random variable taking on positive integers only. We shall assume that one of the following conditions is

satisfied:

(A) For every $n = 1,2,\ldots$ the event $\{\nu = n\}$ depends only on the random variables $\xi_1, \xi_2, \ldots, \xi_n$, that is, the indicator variable of the event $\{\nu = n\}$ is a Baire-function of the random variables $\xi_1, \xi_2, \ldots, \xi_n$.

(B) For every $n = 1,2,\ldots$ the event $\{\nu = n\}$ and the random variables $\xi_{n+1}, \xi_{n+2}, \ldots$ are independent, that is, the indicator variable of the event $\{\nu = n\}$ and the random variables $\xi_{n+1}, \xi_{n+2}, \ldots$ are independent.

Obviously, condition (A) implies (B), whereas the converse is not true. This can be illustrated by the following example. Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be mutually independent and identically distributed random variables for which $\underset{\sim}{P}\{\xi_n = 1\} = \underset{\sim}{P}\{\xi_n = -1\} = 1/2$ and define $\nu = (\xi_1 \xi_2 + 3)/2$. Then (A) is not satisfied, whereas (B) is satisfied.

In 1944 <u>A. Wald</u> [41], [42] considered the case where $\{\xi_n\}$ is a sequence of mutually independent and identically distributed real random variables and defined ν as the smallest n for which ξ_n does not lie in the interval (a, b) . He deduced a fundamental identity for the random variable ξ_ν and this identity made it possible to find the distribution and the moments of ξ_ν . Actually, Wald's main interest was to find the distribution and the moments of ν .

The following theorems are generalizations of some of Wald's results.

<u>Theorem 1.</u> <u>Let us suppose that</u> $\{\xi_n\}$ <u>is a sequence of real random variables for which</u> $\underset{\sim}{E}\{\xi_n\} = a$ <u>exists and is independent of</u> n . <u>If</u> (B)

is satisfied, $E\{\nu\} < \infty$ and

$$(1) \qquad \sum_{n=1}^{\infty} E\{|\xi_n|\}P\{\nu \geq n\} < \infty ,$$

then

$$(2) \qquad E\{\zeta_\nu\} = a\,E\{\nu\} .$$

Proof. Let us define $\delta_n = 1$ if $\nu \geq n$ and $\delta_n = 0$ otherwise. Then we can write that

$$(3) \qquad \zeta_\nu = \sum_{n=1}^{\infty} \xi_n \delta_n .$$

Since by assumption ξ_n and the event $\{\nu < n\}$ are independent, it follows that ξ_n and the event $\{\nu \geq n\}$ are also independent. Thus ξ_n and δ_n are independent random variables and

$$(4) \qquad E\{\xi_n \delta_n\} = a\,E\{\delta_n\} = a\,P\{\nu \geq n\}$$

for $n = 1,2,\ldots$. Accordingly, by (3) we obtain that

$$(5) \qquad E\{\zeta_\nu\} = \sum_{n=1}^{\infty} E\{\xi_n \delta_n\} = a \sum_{n=1}^{\infty} P\{\nu \geq n\} = a\,E\{\nu\} .$$

In (3) we can form the expectation term by term because

$$(6) \qquad E\{|\zeta_\nu|\} \leq \sum_{n=1}^{\infty} E\{|\xi_n \delta_n|\} = \sum_{n=1}^{\infty} E\{|\xi_n|\}P\{\nu \geq n\} < \infty .$$

This proves (2).

For the proof of (2) see also D. Blackwell [33], J. Wolfowitz [44], A. N. Kolmogorov and Yu. V. Prokhorov [37] and N. L. Johnson [36].

Theorem 2. Let us suppose that $\{\xi_n\}$ is a sequence of mutually independent real random variables for which $E\{\xi_n\} = a$ and $\operatorname{Var}\{\xi_n\} = \sigma^2$ exist and are independent of n. If (A) is satisfied, $E\{\nu\} < \infty$, and

(7) $\sum_{n=1}^{\infty} P\{\nu \geq n\} E\{|\zeta_{n-1} - (n-1)a| \mid \nu \geq n\} < \infty,$

then

(8) $$E\{(\zeta_\nu - a\nu)^2\} = \sigma^2 E\{\nu\} .$$

Proof. In proving this theorem we may assume that $a = 0$. If $a \neq 0$, then instead of $\{\xi_n\}$ we can consider the sequence $\{\xi_n - a\}$. Let us define the random variables δ_n $(n = 1,2,\ldots)$ in exactly the same way as in the previous proof. Then $\delta_1 = 1$ and δ_n depends only on $\xi_1, \xi_2, \ldots, \xi_{n-1}$ for $n = 2,3,\ldots$. Let us write

(9) $$\chi_n = \xi_n^2 + 2\xi_n\zeta_{n-1}$$

for $n = 1,2,\ldots$. We note that $\zeta_0 = 0$. Then $\zeta_n^2 = \chi_1 + \ldots + \chi_n$ for $n = 1,2,\ldots$ and

(10) $$\zeta_\nu^2 = \sum_{n=1}^{\infty} \chi_n\delta_n .$$

By (9) we have

(11) $$E\{\chi_n\delta_n\} = E\{\xi_n^2\}E\{\delta_n\} + 2E\{\xi_n\}E\{\zeta_{n-1}\delta_n\} .$$

For ξ_n and δ_n are independent random variables, and also ξ_n and $\zeta_{n-1}\,\delta_n$ are independent random variables. Since $E\{\xi_n\} = 0$, by (11) we obtain that

(12) $$E\{\chi_n\delta_n\} = \sigma^2 E\{\delta_n\} = \sigma^2 P\{\nu \geq n\}$$

for $n = 1,2,\ldots$. Thus by (10) we have

(13) $$E\{\zeta_\nu^2\} = \sigma^2 \sum_{n=1}^{\infty} P\{\nu \geq n\} = \sigma^2 E\{\nu\} .$$

In (10) we can form the expectation term by term because

(14) $$E\{|\xi_n||\zeta_{n-1}|\delta_n\} = E\{|\xi_n|\}E\{|\zeta_{n-1}|\delta_n\} ,$$

$E\{|\xi_n|\} \leq [E\{\xi_n^2\}]^{1/2} = \sigma < \infty$, and

(15) $$\sum_{n=1}^{\infty} E\{|\zeta_{n-1}|\delta_n\} = \sum_{n=1}^{\infty} P\{\delta_n = 1\}E\{|\zeta_{n-1}|\,|\,\delta_n = 1\} < \infty$$

by (7) .

We note that if in Theorem 2 we replace the condition (A) by condition (B), then (8) is not valid anymore. This was demonstrated by J. Seitz and K. Winkelbauer [39] when they pointed out that several results are erroneous in the paper of A. N. Kolmogorov and Yu. V. Prokhorov [37]. Indeed, if we consider the example mentioned at the beginning of this section, then $a = 0$, $\sigma^2 = 1$, $E\{\zeta_\nu^2\} = \frac{5}{2}$ and $E\{\nu\} = \frac{3}{2}$, that is, (8) is not satisfied. For the proof of (8) see also N. L. Johnson [36]. The higher moments of ζ_ν have been investigated by J. Wolfowitz [44] and K. Winkelbauer [43].

In 1949 A. N. Kolmogorov and Yu. V. Prokhorov [37] considered the case when $\xi_n = (\xi_n^{(1)}, \xi_n^{(2)})$ ($n = 1,2,\ldots$) are independent vector random variables. Let $\zeta_n^{(i)} = \xi_1^{(i)} + \ldots + \xi_n^{(i)}$ for $n = 1,2,\ldots$ and $i = 1,2$, and $\zeta_0^{(i)} = 0$ for $i = 1,2$.

Theorem 3. Let $\xi_n = (\xi_n^{(1)}, \xi_n^{(2)})$ $(n = 1,2,\ldots)$ be mutually independent vector random variables for which $E\{\xi_n^{(1)}\} = a_1$, $E\{\xi_n^{(2)}\} = a_2$ and $E\{(\xi_n^{(1)} - a_1)(\xi_n^{(2)} - a_2)\} = \sigma_{12}$ exist and are independent of n. If (A) is satisfied, $E\{\nu\} < \infty$, and

$$\sum_{n=1}^{\infty} P\{\nu \geq n\} E\{|\zeta_{n-1}^{(1)} - (n-1)a_1| + |\zeta_{n-1}^{(2)} - (n-1)a_2| \,\big|\, \nu \geq n\} < \infty , \tag{16}$$

then

$$E\{(\zeta_\nu^{(1)} - a_1\nu)(\zeta_\nu^{(2)} - a_2\nu)\} = \sigma_{12} E\{\nu\} . \tag{17}$$

Proof. Without loss of generality we may assume that $a_1 = a_2 = 0$. If we define δ_n in the same way as in the previous proofs and if now

$$\chi_n = \xi_n^{(1)} \xi_n^{(2)} + \xi_n^{(1)} \zeta_{n-1}^{(2)} + \xi_n^{(2)} \zeta_{n-1}^{(1)} \tag{18}$$

for $n = 1,2,\ldots$, then we can write that

$$\zeta_\nu^{(1)} \zeta_\nu^{(2)} = \sum_{n=1}^{\infty} \chi_n \delta_n . \tag{19}$$

If we form the expectation of (19), then we obtain that

$$E\{\zeta_\nu^{(1)} \zeta_\nu^{(2)}\} = \sigma_{12} \sum_{n=1}^{\infty} E\{\delta_n\} = \sigma_{12} E\{\nu\} . \tag{20}$$

In a similar way as in the proof of Theorem 2 we can easily see that if (16) is satisfied then we can form the expectation term by term in (19).

Further generalizations of the above results have been given by K. Winkelbauer [43].

6. Interchangeable Random Variables. In generalizing the notions of independent events and independent random variables in 1930 B. De Finetti [52], [53], [54], [55], [56] introduced the notions of interchangeable events and interchangeable random variables.

We say that $A_1, A_2, \ldots, A_n$ are interchangeable events if

$$(1) \qquad \mathbf{P}\{A_{i_1} A_{i_2} \ldots A_{i_j}\} = \mathbf{P}\{A_1 A_2 \ldots A_j\}$$

holds for all $1 \leqq i_1 < i_2 < \ldots < i_j \leqq n$, and we say that $A_1, A_2, \ldots, A_i, \ldots$ is an infinite sequence of interchangeable events if (1) holds for all $n = 1, 2, \ldots$.

We have several classical examples for a finite number of interchangeable events. See e.g. reference [77]. In 1923 E. Eggenberger and G. Pólya [59] found an interesting example for an infinite sequence of interchangeable events.

We say that $\xi_1, \xi_2, \ldots, \xi_n$ are interchangeable random variables if all the $n!$ permutations of $(\xi_1, \xi_2, \ldots, \xi_n)$ have the same joint distribution. If $\xi_1, \xi_2, \ldots, \xi_n$ are real random variables, then they are interchangeable if and only if

$$(2) \qquad \mathbf{P}\{\xi_{i_1} \leqq x_1, \xi_{i_2} \leqq x_2, \ldots, \xi_{i_n} \leqq x_n\} = \mathbf{P}\{\xi_1 \leqq x_1, \xi_2 \leqq x_2, \ldots, \xi_n \leqq x_n\}$$

holds for all the $n!$ permutations $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$ and for all $x_1, x_2, \ldots, x_n$.

We say that $\xi_1, \xi_2, \ldots, \xi_i, \ldots$ is an infinite sequence of interchangeable random variables, if $(\xi_1, \xi_2, \ldots, \xi_n)$ are interchangeable random variables for all $n = 2,3,\ldots$. An infinite sequence of real random variables $\xi_1, \xi_2, \ldots, \xi_i, \ldots$ form a sequence of interchangeable random variables if (2) holds for all $n = 1,2,\ldots$.

Let us define the indicator variable of an event A_i by χ_i , that is, $\chi_i = 1$ if A_i occurs and $\chi_i = 0$ if A_i does not occur. If $A_1, \ldots, A_n$ are interchangeable events, then $\chi_1, \chi_2, \ldots, \chi_n$ are interchangeable random variables, and if $A_1, \ldots, A_i, \ldots$ is an infinite sequence of interchangeable events, then $\chi_1, \chi_2, \ldots, \chi_i, \ldots$ is an infinite sequence of interchangeable random variables.

By generalizing the notion of homogeneous processes with independent increments we can introduce the notion of stochastic processes with interchangeable increments. We say that the process $\{\xi(u), 0 \leqq u \leqq t\}$ has interchangeable increments if for all $n = 2,3,\ldots$

$$\xi\left(\frac{rt}{n}\right) - \xi\left(\frac{rt-t}{n}\right) \quad (r = 1,2,\ldots, n) \tag{3}$$

are interchangeable random variables. The process $\{\xi(u), 0 \leqq u < \infty\}$ is said to have interchangeable increments if the random variables (3) are interchangeable random variables for all finite t .

If we choose m points at random on the interval $(0, 1)$ in such a way that independently of the others each point has a uniform distribution on the interval $(0, 1)$, and if $\nu_m(u)$ denotes the number of points in the interval $(0, u)$, then $\{\nu_m(u) , 0 \leq u \leq 1\}$ is a stochastic process with interchangeable increments.

In 1930 B. De Finetti [52] discovered that an infinite sequence of interchangeable random events can be represented as the outcomes of a sequence of randomized Bernoulli trials. In what follows we shall prove this result.

Let us suppose that $A_1, A_2, \ldots, A_i, \ldots$ is an infinite sequence of interchangeable events. Denote by χ_i the indicator variable of the event A_i . Let $\nu_n = \chi_1 + \chi_2 + \ldots + \chi_n$, that is, ν_n is the number of events occurring among $A_1, A_2, \ldots, A_n$.

Let $\pi_0 = 1$ and

(4) $$\pi_j = \underset{\sim}{P}\{A_1 A_2 \ldots A_j\}$$

for $j = 1, 2, \ldots$.

Since

(5) $$\binom{\nu_n}{j} = \sum_{1 \leq i_1 < i_2 < \ldots < i_j \leq n} \chi_{i_1} \chi_{i_2} \ldots \chi_{i_j}$$

for $0 \leq j \leq n$, we obtain that the j-th binomial moment of ν_n is

(6) $$B_j(n) = \underset{\sim}{E}\{\binom{\nu_n}{j}\} = \binom{n}{j}\pi_j$$

for $j = 0, 1, \ldots, n$. If we take into consideration that

(7) $$B_j(n) = \sum_{k=j}^{n} \binom{k}{j} \underset{\sim}{P}\{\nu_n = k\}$$

for $0 \leq j \leq n$, then we obtain easily that

$$\mathrm{P}\{\nu_n = k\} = \sum_{j=k}^{n} (-1)^{j-k}\binom{j}{k}B_j(n) = \binom{n}{k} \sum_{j=k}^{n} (-1)^{j-k}\binom{n-k}{n-j}\pi_j \tag{8}$$

for $0 \leq k \leq n$.

The following theorems are due to B. De Finetti [55]. See also A. Ya. Khintchine [69], [70], W. Feller [60 p. 225-227], and D. G. Kendall [68].

Theorem 1. *There exists a distribution function* $F(x)$ *such that* $F(x) = 0$ *for* $x < 0$, $F(x) = 1$ *for* $x \geq 1$, *and*

$$\lim_{n \to \infty} \mathrm{P}\{\frac{\nu_n}{n} \leq x\} = F(x) \tag{9}$$

in every continuity point of $F(x)$. *The distribution function* $F(x)$ *is uniquely determined by*

$$\pi_j = \int_0^1 x^j dF(x) \tag{10}$$

for $j = 0,1,2,\ldots$.

Proof. First we note that

$$x^j = \sum_{r=1}^{j} \mathfrak{S}_j^r \, r!\binom{x}{r} \tag{11}$$

for every x and $j = 1,2,\ldots$ where $\mathfrak{S}_j^r$ $(1 \leq r \leq j)$ are Stirling numbers of the second kind. Hence by (6) we obtain that

$$\mathrm{E}\{\nu_n^j\} = \sum_{r=1}^{j} \mathfrak{S}_j^r \, r!\binom{n}{r}\pi_r \tag{12}$$

for $j = 1,2,\ldots$ and $n = 1,2,\ldots$. Since $\mathfrak{S}_j^j = 1$ for $j \geqq 1$, it follows from (12) that

$$(13) \qquad \lim_{n \to \infty} \underset{\sim}{E}\{(\frac{\nu_n}{n})^j\} = \pi_j$$

for $j = 0,1,2,\ldots$. The sequence $\{\pi_j\}$ satisfies the following properties: $\pi_0 = 1$ and

$$(14) \qquad \sum_{j=k}^{n} (-1)^{j-k}\binom{n-k}{n-j}\pi_j \geqq 0$$

for $0 \leqq k \leqq n$. This last inequality is a consequence of (8). Thus by a theorem of F. Hausdorff [65], [66] we can conclude that there exists a distribution function $F(x)$ on the interval $[0, 1]$ such that (10) holds for $j = 0,1,2,\ldots$ and $F(x)$ is uniquely determined by (10). Hence (9) follows by Theorem 41.11 . This completes the proof of the theorem.

From (8) and (10) it follows that

$$(15) \qquad \underset{\sim}{P}\{\nu_n = k\} = \int_0^1 \binom{n}{k}x^k(1-x)^{n-k}dF(x)$$

for $0 \leqq k \leqq n$. From (15) we can also conclude that (9) holds.

Now let us suppose that $(\Omega,\mathcal{A},\underset{\sim}{P})$ is a probability space and $A_1, A_2,\ldots,A_i,\ldots$ is an infinite sequence of interchangeable events such that $A_i \in \mathcal{B}$ for $i = 1,2,\ldots$ and define π_j by (4) and $F(x)$ by (10). Denote by $\mathcal{B}_n$ the minimal σ-algebra which contains the events $A_n, A_{n+1},\ldots$ and let

$$(16) \qquad \mathcal{B}^* = \bigcap_{n=1}^{\infty} \mathcal{B}_n$$

be the so-called tail σ-algebra.

Theorem 2. There exists a random variable θ defined on the probability space $(\Omega, A, \underset{\sim}{P})$ such that

(17) $$\underset{\sim}{P}\{\lim_{n \to \infty} \frac{\nu_n}{n} = \theta\} = 1 .$$

Proof. We shall use formula (12) and we shall need the Stirling numbers $\mathfrak{S}_j^r$ for $1 \leqq r \leqq j \leqq 4$. These are given in Table I.

$\mathfrak{S}_j^r$

j \ r	1	2	3	4
1	1			
2	1	1		
3	1	3	1	
4	1	7	6	1

Table I.

If $n \geqq 1$ and $q \geqq 1$, then by (12) we obtain that

(18) $$\underset{\sim}{E}\{(\frac{\nu_n}{n} - \frac{\nu_{n+q}}{n+q})^2\} = \frac{q(\pi_1 - \pi_2)}{n(n+q)} \leqq \frac{q}{4n(n+q)} .$$

If $q \to \infty$ and $n \to \infty$, then the extreme right member in (18) tends to 0, and therefore we can conclude that there exists a random variable θ such that

(19) $$\lim_{n \to \infty} \underset{\sim}{E}\{(\frac{\nu_n}{n} - \theta)^2\} = 0 .$$

By (12) we can also prove that

$$(20)\quad \lim_{q\to\infty} \underset{\sim}{E}\{(\frac{\nu_n}{n} - \frac{\nu_{n+q}}{n+q})^4\} = \frac{(3\pi_2-6\pi_3+3\pi_4)n^2+(\pi_1-7\pi_2+12\pi_3-6\pi_4)n}{n^4} < \frac{3}{16\,n^2}$$

for $n \geq 1$. Thus b y (20) we have

$$(21)\qquad \underset{\sim}{E}\{(\frac{\nu_n}{n} - \theta)^4\} < \frac{3}{16\,n^2}$$

for all $n \geq 1$. Since for every $\varepsilon > 0$

$$(22)\qquad \underset{\sim}{P}\{|\frac{\nu_n}{n} - \theta| > \varepsilon\} \leq \frac{1}{\varepsilon^4}\underset{\sim}{E}\{(\frac{\nu_n}{n} - \theta)^4\} < \frac{3}{16n^2\varepsilon^4} ,$$

it follows that

$$(23)\qquad \sum_{n=1}^{\infty} \underset{\sim}{P}\{|\frac{\nu_n}{n} - \theta| > \varepsilon\} < \frac{3}{16\varepsilon^4}\sum_{n=1}^{\infty}\frac{1}{n^2} < \frac{3}{8\varepsilon^4}$$

and this implies (17) by Lemma 43.1 .

Obviously $0 \leq \theta \leq 1$ with probability 1 and the random variable θ is measurable with respect to the tail σ-algebra $\mathcal{B}^*$, that is, $\{\theta \leq x\} \in \mathcal{B}^*$ for every x .

By (9) and (17) (or (19)) it follows that necessarily

$$(24)\qquad \underset{\sim}{P}\{\theta \leq x\} \overset{!}{=} F(x) .$$

<u>Theorem 3.</u> <u>We have</u>

$$(25)\qquad \underset{\sim}{P}\{A_1A_2\ldots A_j|\theta\} = \theta^j$$

<u>with probability</u> 1 .

Proof. Since

$$P\{A_1A_2\ldots A_j|\theta\} = E\{\chi_1\chi_2\ldots\chi_j|\theta\} \tag{26}$$

for $j = 1,2,\ldots,$ it follows from (5) that

$$P\{A_1A_2\ldots A_j|\theta\} = \frac{1}{\binom{n}{j}} E\{\binom{\nu_n}{j}|\theta\} \tag{27}$$

for $1 \leq j \leq n$. If $n \to \infty$, then the right-hand side of (27) converges to $E\{\theta^j|\theta\} = \theta^j$ with probability 1 . This completes the proof of the theorem.

By (25) we have

$$\pi_j = E\{\theta^j\} \tag{28}$$

for $j = 0,1,2,\ldots$. This proves once again that π_j can be represented in the form (10).

By (8) and (28) we can write that

$$P\{\nu_n = k\} = E\{\binom{n}{k}\theta^k(1-\theta)^{n-k}\} \tag{29}$$

for $0 \leq k \leq n$. This formula reflects the result that an infinite sequence of interchangeable events $A_1, A_2,\ldots, A_i,\ldots$ can be represented in the following way: We perform an infinite sequence of Bernoulli trials with probability θ for success where θ is a random variable with distribution function $F(x)$ and A_i denotes the event that the i-th trial results in success.

We note that, in general, for a finite number of interchangeable events

$A_1, A_2, \ldots, A_n$, the probabilities π_j $(0 \leq j \leq n)$ cannot be represented in the form (10).

The above results can also be expressed by using the indicator variables $\chi_1, \chi_2, \ldots, \chi_i, \ldots$.

If $A_1, A_2, \ldots, A_i, \ldots$ are interchangeable events, then $\chi_1, \chi_2, \ldots, \chi_i, \ldots$ are interchangeable random variables. By (17) we have

(30) $$\underset{\sim}{P}\{\lim_{n \to \infty} \frac{\chi_1 + \chi_2 + \ldots + \chi_n}{n} = \theta\} = 1 .$$

If $\mathcal{B}$ denotes the σ-algebra generated by the random variable θ, then $\mathcal{B} \subset \mathcal{B}^*$ and

(31) $$\underset{\sim}{P}\{\chi_1 = \varepsilon_1, \chi_2 = \varepsilon_2, \ldots, \chi_k = \varepsilon_k | \mathcal{B}\} = \prod_{i=1}^{k} \underset{\sim}{P}\{\chi_i = \varepsilon_i | \mathcal{B}\}$$

holds for all $k = 1,2,\ldots$ and $\varepsilon_i = 0$ or $\varepsilon_i = 1$ $(i = 1,2,\ldots,k)$ with probability 1 . In (31)

(32) $$\underset{\sim}{P}\{\chi_i = \varepsilon_i | \mathcal{B}\} = \underset{\sim}{P}\{\chi_1 = \varepsilon_i | \mathcal{B}\}$$

with probability 1 .

Accordingly, we can represent $\chi_1, \chi_2, \ldots, \chi_i, \ldots$ as a sequence of conditionally independent random variables with a common distribution. This last result can be extended for more general sequences of interchangeable random variables as was demonstrated in 1937 by B. De Finetti [56]. See also E. B. Dynkin [58], E. Hewitt and L. J. Savage [67], and M. Loève [71 pp. 364-365, and p. 400].

Theorem 4. If (Ω, A, P) is a probability space and ξ_i $(i = 1,2,\ldots)$ is an infinite sequence of interchangeable real random variables, then there exists a nontrivial σ-subalgebra B of A such that

$$(33) \qquad P\{\xi_1 \leq x_1, \xi_2 \leq x_2, \ldots, \xi_k \leq x_k | B\} = \prod_{i=1}^{k} P\{\xi_i \leq x_i | B\}$$

for all $k = 1,2,\ldots$ and $x_1, x_2, \ldots, x_k$ with probability 1, and in (33)

$$(34) \qquad P\{\xi_i \leq x_i | B\} = P\{\xi_1 \leq x_i | B\}$$

for all $i = 1,2,\ldots$ with probability 1.

Proof. We can reduce the proof of this theorem to the results proved above. Let us define

$$(35) \qquad \chi_i(u) = \begin{cases} 1 & \text{if } \xi_i \leq u, \\ 0 & \text{if } \xi_i > u, \end{cases}$$

for $i = 1,2,\ldots$ and

$$(36) \qquad \nu_n(u) = \chi_1(u) + \chi_2(u) + \ldots + \chi_n(u)$$

for $n = 1,2,\ldots$ and all u. Since in this case $\{\chi_i(u)\}$ are interchangeable indicator variables, by Theorem 2 for every u there exists a random variable $\theta(u)$ such that

$$(37) \qquad P\{\lim_{n \to \infty} \frac{\nu_n(u)}{n} = \theta(u)\} = 1 .$$

With probability 1, the random variable $\theta(u)$ is a nondecreasing function of u, $\theta(u) \to 1$ as $u \to \infty$, and $\theta(u) \to 0$ as $u \to -\infty$. We can choose $\theta(u)$ such that for every $\omega \in \Omega$ the function $\theta(u) = \theta(u; \omega)$ is a distribution function in u.

Denote by $\mathcal{B}$ the minimal σ-algebra generated by the random variables $\{\theta(u), -\infty < u < \infty\}$. Then we have

$$\underset{\sim}{P}\{\xi_i \leq x|\mathcal{B}\} = \underset{\sim}{E}\{\chi_i(x)|\mathcal{B}\} = \frac{1}{n}\underset{\sim}{E}\{\nu_n(x)|\mathcal{B}\} \tag{38}$$

for any $i = 1,2,\ldots$ and $n = 1,2,\ldots$. If $n \to \infty$, then by (37) the right-hand side converges to $\theta(x)$ with probability 1. Thus we have

$$\underset{\sim}{P}\{\xi_i \leq x|\mathcal{B}\} = \theta(x) \tag{39}$$

for $i = 1,2,\ldots$ and every x with probability 1. This proves (34).

In a similar way we obtain that

$$\begin{aligned}\underset{\sim}{P}\{\xi_1 \leq x_1,\ldots, \xi_k \leq x_k|\mathcal{B}\} &= \underset{\sim}{E}\{\chi_1(x_1)\ldots \chi_k(x_k)|\mathcal{B}\} = \\ &= \frac{1}{\binom{n}{k}}\underset{\sim}{E}\{\sum_{1\leq i_1<i_2<\ldots<i_k\leq n} \chi_{i_1}(x_1)\ldots\chi_{i_k}(x_k)|\mathcal{B}\}\end{aligned} \tag{40}$$

for $n \geq k$. If $n \to \infty$, then the right-hand side of (40) has the limit

$$\lim_{n\to\infty} \frac{1}{\binom{n}{k}}\underset{\sim}{E}\{\nu_n(x_1)\nu_n(x_2)\ldots \nu_n(x_k)|\mathcal{B}\} = \theta(x_1)\theta(x_2)\ldots\theta(x_k) \tag{41}$$

with probability 1. Accordingly,

$$\underset{\sim}{P}\{\xi_1 \leq x_1, \xi_2 \leq x_2,\ldots, \xi_k \leq x_{i_k}|\mathcal{B}\} = \theta(x_1)\theta(x_2)\ldots\theta(x_k) \tag{42}$$

with probability 1 . By (39) and (42) we obtain (33) which was to be proved.

Finally, we mention that in 1960 H. Bühlmann [50] demonstrated that a stochastic process $\{\xi(u) , 0 \leq u < \infty\}$ with interchangeable increments can be represented as a homogeneous stochastic process with conditionally independent increments.

Accordingly, if (Ω, A, P) is a probability space and $\{\xi(u) , 0 \leq u < \infty\}$ is a stochastic process with interchangeable increments, then there is a nontrivial σ-subalgebra B of A such that with probability 1 the process $\{\xi(u) , 0 \leq u < \infty\}$ is homogeneous and has independent increments with respect to B .

8. Slowly Varying Functions. A real function $L(x)$ defined for $x \geq a$ where a is some positive number is called a slowly varying function at $x \to \infty$ if it is positive for $x \geq a$, measurable on any finite interval in $[a, \infty)$ and if

$$\lim_{x \to \infty} \frac{L(\omega x)}{L(x)} = 1 \tag{1}$$

for every $\omega > 0$.

An example for slowly varying functions is

$$L(x) = (\log x)^{c_1} (\log_2 x)^{c_2} \ldots (\log_n x)^{c_n} \tag{2}$$

where $\log_2 x = \log \log x$ and $\log_k x = \log \log_{k-1} x$ for $k = 3, 4, \ldots$ and $c_1, c_2, \ldots, c_n$ are real numbers.

Slowly varying functions play an important role in mathematics in proving various limit theorems. In particular, they have important applications in the theory of probability in obtaining various limiting distributions.

As early as 1904 A. Pringsheim [105] was concerned with monotone slowly varying functions. See also E. Landau [97], G. Pólya [102], [103], G. Pólya and G. Szegő [104 pp. 67-69], and R. Schmidt [107]. In 1930 J. Karamata [92] found the most general form of a continuous slowly varying function. In 1949 J. Korevaar, T. van Aardenne-Ehrenfest and N. G. de Bruijn [96] studied positive, measurable functions $L(x)$ satisfying (1) .

We have the following representation theorem.

Theorem 1. *If* $L(x)$ *is a slowly varying function, then there exists some positive constant* a *such that*

$$L(x) = c(x)e^{\int_a^x \frac{\varepsilon(u)}{u}\,du} \tag{3}$$

for $x \geq a$ *where* $c(x)$ *and* $\varepsilon(x)$ *are bounded measurable functions on the interval* $[a, \infty)$ *and satisfy the conditions* $\lim_{x\to\infty} c(x) = c$ *where* c *is a positive constant and* $\lim_{x\to\infty} \varepsilon(x) = 0$.

This theorem was found in 1930 by J. Karamata [92], [94] for continuous $L(x)$, in 1949 J. Korevaar, T. van Aardenne-Ehrenfest, and N. G. de Bruijn [96] proved it for the case when $\log L(x)$ is integrable on every compact subinterval of $[a,\infty)$ and in 1959 N. G. de Bruijn [84] proved it for measurable $L(x)$.

Theorem 2. If $L(x)$ is a slowly varying function, then (1) holds uniformly for $\omega \in [a_1, a_2]$ where $[a_1, a_2]$ is any finite subinterval of $(0, \infty)$.

For continuous $L(x)$ this theorem has been proved by J. Karamata [92], [94] and for measurable $L(x)$ by J. Korevaar, T. van Aardenne-Ehrenfest, and N. G. de Bruijn [96]. See also G. H. Hardy and W. W. Rogosinski [90], H. Delange [85], W. Matuszewska [99], [100], and R. Bojanić and E. Seneta [82].

It is interesting to mention the following result which was found in 1967 by C. C. Heyde [91]. See also B. A. Rogozin [106].

Theorem 3. Let $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ be a sequence of nonnegative numbers. If

$$\lim_{n\to\infty} \frac{\alpha_1 + \alpha_2 + \ldots + \alpha_n}{n} = \alpha \tag{4}$$

where α is a finite nonnegative number, then

$$\exp\{-\sum_{n=1}^{\infty} \frac{\alpha_n}{n} x^n\} \sim (1-x)^{\alpha} L(\frac{1}{1-x}) \tag{5}$$

as $x \to 1-0$ where $L(x)$ is a slowly varying function of x at $x \to \infty$.

In (5) the left-hand side is asymptotically equal to the right-hand side, that is, the ratio of the two sides tends to 1 as $x \to 1-0$.

The following theorem was found in 1959 by N. G. de Bruijn [84].

Theorem 4. *If* $L(x)$ *is a slowly varying function, then there exists a slowly varying function* $L^*(x)$ *such that*

(6) $$\lim_{x \to \infty} L^*(xL(x))L(x) = 1$$

and

(7) $$\lim_{x \to \infty} L(xL^*(x))L^*(x) = 1 .$$

Moreover, $L^*(x)$ *is asymptotically uniquely determined by* $L(x)$.

In several cases we can easily construct a function $L^*(x)$ occurring in Theorem 4 by using the following procedure of A. Békéssy [79]. Let $k_1(x) = 1/L(x)$ and define $k_2(x), k_3(x), \ldots$ recursively by the formula

(8) $$k_{n+1}(x) = k_1(x\, k_n(x)) .$$

If $k_{n+1}(x) \sim k_n(x)$ for some n as $x \to \infty$, then $L^*(x) \sim k_n(x)$ as $x \to \infty$.

The notion of slowly varying functions is strongly related to the notion of regularly varying functions. See R. Schmidt [107].

A real function $U(x)$ defined for $x \geq a$ where a is some positive number is called a regularly varying function at $x \to \infty$ if it is positive for $x \geq a$, measurable on any finite interval in $[a, \infty)$ and if

(9) $$\lim_{x \to \infty} \frac{U(\omega x)}{U(x)} = \omega^q$$

for every $\omega > 0$ where q is a constant.

We can easily see that $U(x)$ is a regularly varying function if and only if it is of the form

(10) $$U(x) = x^q L(x)$$

where $L(x)$ is a slowly varying function.

We note that instead of (9) it is sufficient to require only that the limit $\lim_{x \to \infty} U(\omega x)/U(x) = V(\omega)$ exists for $\omega > 0$ and $V(\omega) \neq 0$. Since $V(\omega_1 \omega_2) = V(\omega_1)V(\omega_2)$ for any $\omega_1 > 0$ and $\omega_2 > 0$, and since $V(\omega)$ is measurable on $(0, \infty)$ it follows by a result of G. Hamel [89] (see also B. Blumberg [81]) that $V(\omega) = \omega^q$ where q is some real constant.

In conclusion we mention one problem which frequently occurs in the theory of probability. Let $T(x)$ be a slowly varying function defined on the interval $[a, \infty)$ where $a > 0$. Let α be a positive real number. The problem is to find a function $B(x)$ which satisfies the requirements $\lim_{x \to \infty} B(x) = \infty$ and

(11) $$\lim_{x \to \infty} \frac{xT(B(x))}{[B(x)]^\alpha} = 1 .$$

If we write

(12) $$L(x) = [T(x)]^{-1/\alpha}$$

and

(13) $$L^*(x) = \frac{B(x^\alpha)}{x} ,$$

then by (11) we obtain that

$$\lim_{x \to \infty} L(xL^*(x))L^*(x) = 1 \ . \tag{14}$$

Since obviously $L(x)$ is a slowly varying function, it follows from Theorem 4 that $L^*(x)$ is also a slowly varying function of x and is asymptotically uniquely determined by $L(x)$.

Let $k_1(x) = [T(x)]^{1/\alpha}$ and $k_{n+1}(x) = k_1(xk_n(x))$ for $n = 1,2,\ldots$. If $k_{n+1}(x) \sim k_n(x)$ for some n as $x \to \infty$, then $L^*(x) \sim k_n(x)$ as $x \to \infty$.

Finally, by (13) it follows that

$$B(x) = x^{1/\alpha} L^*(x^{1/\alpha}) \tag{15}$$

satisfies all the requirements, and furthermore $B(x)$ is asymptotically uniquely determined by (11).

9. Abelian and Tauberian Theorems. The Abelian and Tauberian theorems for power series and for Laplace-Stieltjes transforms mentioned in this section have many useful applications in the theory of probability.

The Abelian theorems for power series are concerned with a sequence of real or complex numbers $a_0, a_1, \ldots, a_n, \ldots$. Let us define

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{1}$$

as the generating function of the sequence $\{a_n\}$.

First, in 1826 N. H. Abel [108 p. 314] proved the following theorem.

Theorem 1. *If the series*

$$\sum_{n=0}^{\infty} a_n \tag{2}$$

is convergent, then the power series (1) *is convergent in the unit circle* $|z| < 1$ *and*

$$\lim_{z \to 1-0} f(z) = \sum_{n=0}^{\infty} a_n \tag{3}$$

whenever z *approaches* 1 *through the real axis in the unit circle* $|z| < 1$.

In 1875 O. Stolz [224] demonstrated that if (2) is convergent, then (3) holds whenever z approaches 1 through a straight line lying in the circle $|z| < 1$.

In 1920 G. H. Hardy and J. E. Littlewood [155] remarked that if (2) is convergent, then (3) holds whenever z approaches 1 through a Jordan curve which lies between two chords of the unit circle, meeting at $z = 1$.

However, (4) does not necessarily hold even if z approaches 1 through a Jordan curve which lies in the unit circle and which possesses a continuously turning tangent at every point except $z = 1$, but $\arg(1-z)$ tends either to $\pi/2$ or $-\pi/2$.

If we apply Theorem 1 to the sequence $a_0, a_1 - a_0, a_2 - a_1, \dots$, then we obtain the following version of Theorem 1.

Theorem 2. *If the limit*

(4) $$\lim_{n \to \infty} a_n = A$$

exists, then (1) *is convergent in the unit circle* $|z| < 1$, *and*

(5) $$\lim_{z \to 1-0} (1-z)f(z) = A .$$

In 1880 G. Frobenius [137] demonstrated that in Theorem 2 the condition (4) could be replaced by the weaker condition

(6) $$\lim_{n \to \infty} \frac{a_0 + a_1 + \dots + a_n}{n} = A .$$

In 1878 P. Appell [113] proved the following generalization of Theorem 2.

Theorem 3. *If the limit*

(7) $$\lim_{n \to \infty} \frac{a_n}{n^\alpha} = \frac{A}{\Gamma(\alpha+1)}$$

exists for some $\alpha > -1$, *then* (1) *is convergent in the unit circle* $|z| < 1$

and

(8) $$\lim_{z \to 1-0} (1-z)^{\alpha+1} f(z) = A \ .$$

In 1901 A. Pringsheim [211] proved that if in Theorem 3 we replace the condition (7) by the weaker condition

(9) $$\lim_{n \to \infty} \frac{a_0 + a_1 + \ldots + a_n}{n^{\alpha}} = \frac{A}{\Gamma(\alpha+1)}$$

for $\alpha > -1$, then we have

(10) $$\lim_{z \to 1-0} (1-z)^{\alpha} f(z) = A \ .$$

We note that Theorem 3 remains valid for complex α with $\mathrm{Re}(\alpha) > -1$.

Theorem 3 does not cover the case $\alpha = -1$; however, if

(11) $$\lim_{n \to \infty} na_n = B \ ,$$

then by a result of E. Lasker [194] we have

(12) $$\lim_{z \to 1-0} \frac{f(z)}{\log \frac{1}{1-z}} = B \ .$$

We can generalize Theorem 3 in the following way.

Theorem 4. *If the limit*

(13) $$\lim_{n \to \infty} \frac{a_n}{n^{\alpha} L(n)} = \frac{A}{\Gamma(\alpha+1)}$$

exists for some $\alpha > -1$ *where* $L(z)$ *is a slowly varying function of* z

at $z \to \infty$, then (1) is convergent in the unit circle $|z| < 1$, and

(14) $$\lim_{z \to 1-0} \frac{(1-z)^{\alpha+1} f(z)}{L(1/(1-z))} = A .$$

In 1901 E. Lasker [194] proved that Theorem 4 is true if

(15) $$L(z) = (\log z)^{\alpha_1} (\log_2 z)^{\alpha_2} \ldots (\log_r z)^{\alpha_r}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_r$ arbitrary real numbers. He also proved that if

(16) $$\lim_{n \to \infty} \frac{na_n}{L(n)} = B$$

where $L(z)$ is given by (15) with $\alpha_1 = \alpha_2 = \ldots = \alpha_{k-1} = -1$ and $\alpha_k > -1$ $(1 \leq k \leq r)$, then

(17) $$\lim_{z \to 1-0} (\log_k \frac{1}{1-z})^{-\alpha_k-1} (\log_{k+1} \frac{1}{1-z})^{-\alpha_{k+1}} \ldots (\log_r \frac{1}{1-z})^{-\alpha_r} f(z) = \frac{B}{\alpha_k+1} .$$

In 1904 A. Pringsheim [212] proved Theorem 4 for increasing and decreasing slowly varying functions $L(z)$.

By adopting Lasker's method and by referring to Theorem 8.2 in the Appendix we can easily prove Theorem 4 in the general case. We shall only sketch a proof. If $z \to 1$ through real numbers <1, then by (13) we have

(18) $$f(z) \sim \frac{A}{\Gamma(\alpha+1)} \sum_{n=a}^{\infty} L(n)n^{\alpha}z \sim \frac{A}{\Gamma(\alpha+1)} \int_{a}^{\infty} L(u)u^{\alpha}z^{u}\, du$$

where a is some positive number. Let us put $z = (p-1)/p$ and $u = pv$ in (18). If $z \to 1$, then $p \to \infty$, and by (18) we have

(19)
$$f(z) \sim \frac{AL(p)p^{\alpha+1}}{\Gamma(\alpha+1)} \int_{a/p}^{\infty} \frac{L(pv)}{L(p)} (1-\frac{1}{p})^{pv} v^{\alpha} dv \sim$$
$$\sim \frac{AL(p)p^{\alpha+1}}{\Gamma(\alpha+1)} \int_{0}^{\infty} e^{-v} v^{\alpha} dv = AL(p)p^{\alpha+1} = \frac{AL(\frac{1}{1-z})}{(1-z)^{\alpha+1}} .$$

This proves (14). We note that Theorem 4 remains valid for complex α with $\mathrm{Re}(\alpha) > -1$.

The Tauberian theorems for power series are concerned with the converse of the Abelian theorems.

First, in 1897 A. Tauber [236] proved the converse of Theorem 1.

Theorem 5. *If the series*

(20)
$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

is convergent for $|z| < 1$, *if the limit*

(21)
$$\lim_{z \to 1-0} f(z)$$

exists as $z \to 1$ *through real numbers* < 1, *and if*

(22)
$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} k a_k = 0 ,$$

then

(23)
$$\sum_{n=0}^{\infty} a_n$$

is convergent and is equal to the limit (21).

The example $a_0 = 1$, $a_1 = -1$, $a_2 = 1$, $a_3 = -1, \ldots$ shows that the converse of Theorem 1 is false without making some additional restrictions on the sequence $\{a_n\}$. Actually A. Tauber proved that the conditions in Theorem 5 are necessary and sufficient. If (23) is convergent, then by Theorem 1 (20) is convergent for $|z| < 1$, and the limit (21) exists and by a theorem of L. Kronecker [187] (22) is satisfied too.

If we apply Theorem 5 to the sequence $a_0, a_1 - a_0, a_2 - a_1, \ldots$, then we obtain the following version of Theorem 5.

Theorem 6. *If the series*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{24}$$

is convergent for $|z| < 1$, *if the limit*

$$\lim_{z \to 1-0} (1-z)f(z) = A \tag{25}$$

exists as $z \to 1$ *through real numbers* < 1, *and if*

$$\lim_{n \to \infty} \left(\frac{a_0 + a_1 + \ldots + a_n}{n} - a_n\right) = 0 \tag{26}$$

then

$$\lim_{n \to \infty} a_n = A . \tag{27}$$

Obviously (26) is satisfied if

$$\lim_{n \to \infty} n(a_n - a_{n-1}) = 0 . \tag{28}$$

In 1900 A. Pringsheim [210] proved that in Theorem 6 the condition (26) can be replaced by the condition that $a_{n-1} \leqq a_n$ for $n = 1,2,\ldots$.

In 1911 J. E. Littlewood [195] proved that in Theorem 6, (26) can be replaced by the condition that

$$(29) \qquad |n(a_n - a_{n-1})| < K$$

for $n = 1,2,\ldots$ where K is some positive constant. The condition (29) was suggested in 1910 by G.H. Hardy [144 p.308].

In 1912 G. H. Hardy and J. E. Littlewood [151] remarked that in Theorem 6 the condition (26) can be replaced by the hypothesis that a_n $(n = 0,1,2,\ldots)$ are real and

$$(30) \qquad n(a_n - a_{n-1}) > -K$$

for $n = 1,2,\ldots$ where K is some positive constant. This was already observed in 1910 by E. Landau [191].

For this last result a simple proof was given in 1930 by J. Karamata [172]. (See also E. C. Titchmarsh [238 pp. 227-229].) For other proofs see H. Wielandt [242] and S. Izumi [169] .

Further generalizations of Theorem 6 have been given by E. Landau [192] and R. Schmidt [222]. In 1925 R. Schmidt [222] proved that in Theorem 6 the condition (26) can be replaced by the hypothesis that a_n $(n = 0,1,2,\ldots)$ are real and

$$(31) \qquad \liminf (a_n - a_m) \geqq 0$$

when $n > m$ and m and n tend to infinity in such a way that $m/n \to 1$.

See also T. Vijayaraghavan [239].

In 1914 G. Hardy and J. E. Littlewood [153] proved the following converse of Theorem 4.

Theorem 7. *Let us suppose that* a_n $(n = 0,1,2,\ldots)$ *are real numbers, the series*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{32}$$

is convergent for $|z| < 1$, *and the limit*

$$\lim_{z \to 1-0} \frac{(1-z)^{\alpha+1} f(z)}{L(1/(1-z))} = A \tag{33}$$

exists as $z \to 1$ *through real numbers* < 1 *for some* $\alpha \geqq 0$ *where* $L(z)$ *is a slowly varying function of* z *at* $z \to \infty$.

If either

$$a_n - a_{n-1} \geqq 0 \tag{34}$$

for $n = 1,2,\ldots$ *and* $\alpha \geqq 0$, *or*

$$n(a_n - a_{n-1}) > -K n^{\alpha} L(n) \tag{35}$$

for $n = 1,2,\ldots,$ $\alpha > 0$ *and* K *is a positive constant, then*

$$\lim_{n \to \infty} \frac{a_n}{n^{\alpha} L(n)} = \frac{A}{\Gamma(\alpha+1)} . \tag{36}$$

In proving this theorem G. H. Hardy and J. E. Littlewood [153] assumed that the function $L(z)$ is of the form of (15). However, their proof can easily be extended to the general case.

Abelian and Tauberian theorems have been proved for Dirichlet's series too. In his studies in the theory of numbers P. G. L. Dirichlet [125], [127 p. 252 and pp. 371-379] encountered the following type of series

$$\mu(s) = \sum_{n=0}^{\infty} a_n e^{-\lambda_n s} \tag{37}$$

where $\{\lambda_n\}$ is an increasing sequence of nonnegative real numbers for which $\lim_{n \to \infty} \lambda_n = \infty$ and s is a complex number. If $\lambda_n = n$, then $\mu(s)$ reduces to a power series in e^{-s}. If $\lambda_n = \log(n+1)$, then (37) is called an ordinary Dirichlet's series. For the theory of Dirichlet's series we refer to E. Landau [190 pp. 721-882] and G. H. Hardy and M. Riesz [160].

By the investigation of E. Landau [188], [190], J. E. Littlewood [195], and G. H. Hardy and J. E. Littlewood [153], [154], [156], [159] and others we have several Abelian and Tauberian theorems for Dirichlet's series.

Theorem 1 has the following extension for Dirichlet's series.

Theorem 8. *If the series*

$$\sum_{n=0}^{\infty} a_n \tag{38}$$

is convergent, then (37) *is convergent in the half-plane* $Re(s) > 0$ *and*

$$(39)\qquad \lim_{s\to+0} \mu(s) = \sum_{n=0}^{\infty} a_n$$

whenever s approaches 0 through positive real numbers or through complex numbers lying in the sector $\{s = re^{i\phi} : r > 0 ,\ |\phi| < c < \frac{\pi}{2}\}$.

For the proof of this theorem see R. Dedekind ([127 p. 374]), E. Cahen [122] and E. Landau [190 pp. 737-742].

The converse of Theorem 8 is not valid without making some restrictions on the sequence $\{a_n\}$. As an extension of Theorem 5 we have the following result.

Theorem 9. If the series (37) is convergent for $Re(s) > 0$, if the limit $\lim_{s\to+0} \mu(s) = A$ exists as $s \to 0$ through positive real numbers, if $\lim_{n\to\infty} \lambda_n = \infty$ and if

$$(40)\qquad \lim_{n\to\infty} \frac{\lambda_n a_n}{\lambda_n - \lambda_{n-1}} = 0 ,$$

then

$$(41)\qquad \sum_{n=0}^{\infty} a_n$$

is convergent and is equal to A.

This theorem is due to E. Landau [188]. In 1911 J. E. Littlewood [195] proved that (40) can be replaced by the conditions $\lim_{n\to\infty} \lambda_{n-1}/\lambda_n = 1$ and

$$(42)\qquad |a_n| < K \frac{\lambda_n - \lambda_{n-1}}{\lambda_n}$$

for $n = 1,2,\ldots$ where K is a positive constant. In 1914 G.H. Hardy and J.E. Littlewood [153],[154] proved that if $\{a_n\}$ is a sequence of real numbers, then in Theorem 9 the hypothesis (41) can be replaced by $\lim_{n \to \infty} \lambda_{n-1}/\lambda_n = 1$ and

$$a_n > -K \frac{\lambda_n - \lambda_{n-1}}{\lambda_n} \tag{43}$$

for $n = 1,2,\ldots$ where K is some positive constant.

Further generalizations of Theorem 9 have been given by E. Landau [192] and R. Schmidt [222].

Theorem 4 and Theorem 7 have also analog extensions for Dirichlet's series.

We can consider the Dirichlet's series (37) as a particular Laplace-Stieltjes integral. If we define

$$m(x) = \begin{cases} 0 & \text{for } x < \lambda_0 \\ a_0 + a_1 + \ldots + a_n & \text{for } \lambda_n \leqq x < \lambda_{n+1}, \ (n = 0,1,2,\ldots) \end{cases}, \tag{44}$$

then (37) can be expressed as

$$\mu(s) = \int_{-0}^{\infty} e^{-sx}\, dm(x) . \tag{45}$$

Most of the Abelian and Tauberian theorems valid for step functions $m(x)$ can be extended to more general functions $m(x)$.

In what follows we assume that $m(x)$ is a real function defined on the interval $[0, \infty)$ and is of bounded variation in every finite interval. In this case the integral

$$\mu(s) = \int_{-0}^{\infty} e^{-sx}\, dm(x) = s \int_{0}^{\infty} e^{-sx}\, m(x) dx \tag{46}$$

is called the Laplace-Stieltjes transform of $m(x)$.

For the Laplace-Stieltjes transform $\mu(s)$ we have the following Abelian and Tauberian theorems.

Theorem 10. If

(47) $$\lim_{x \to \infty} m(x) = A$$

exists, then $\mu(s)$ is convergent for $\mathrm{Re}(s) > 0$ and

(48) $$\lim_{s \to +0} \mu(s) = A$$

whenever s approaches 0 through positive real numbers or through complex numbers lying in the sector $\{s = re^{i\phi} : r > 0 \ , \ |\phi| < c < \frac{\pi}{2}\}$.

This theorem is analogous to Theorem 1 and is an easy extension of Theorem 8. More generally we have the following theorem.

Theorem 11. If

(49). $$\lim_{x \to \infty} \frac{m(x)}{x^{\alpha}} = \frac{A}{\Gamma(\alpha+1)}$$

exists for some $\alpha > -1$, then $\mu(s)$ is convergent for $\mathrm{Re}(s) > 0$ and

(50) $$\lim_{s \to +0} s^{\alpha}\mu(s) = A$$

whenever s approaches 0 through positive real numbers or through complex numbers lying in the sector $\{s = re^{i\phi} : r > 0 \ , \ |\phi| < c < \frac{\pi}{2}\}$.

This theorem is an **easy** extension of Theorem 3. For its proof see G. H. Hardy and J. E. Littlewood [159 p.27], D. V. Widder [241 p. 182] and G. Doetsch [131 p. 456].

Theorem 4 has the following extension for Laplace- Stieltjes transforms. (See G. Doetsch [131 p. 460].

Theorem 12. If

(51) $$\lim_{x \to \infty} \frac{m(x)}{x^{\alpha} L(x)} = \frac{A}{\Gamma(\alpha+1)}$$

exists for some $\alpha > -1$, where $L(x)$ is a slowly varying function of x at $x \to \infty$, then $\mu(s)$ is convergent in the domain $\mathrm{Re}(s) > 0$, and

(52) $$\lim_{s \to +0} \frac{s^{\alpha} \mu(s)}{L(1/s)} = A$$

whenever s approaches 0 through positive real numbers or through complex numbers lying in the sector $\{s = re^{i\phi} : r > 0 , |\phi| < c < \frac{\pi}{2}\}$.

If we make some specific restrictions on the function $m(x)$ then the converse of Theorems 10, 11, and 12 is also true. In what follows we shall consider only the case when $m(x)$ is a nonnegative and nondecreasing function of x.

Theorem 13. If $m(x)$ $(0 \leq x < \infty)$ is a nonnegative and nondecreasing function of x, if $\mu(s)$ is convergent for $\mathrm{Re}(s) > 0$, and if for some $\alpha \geq 0$

(53) $$\lim_{s \to +0} s^{\alpha} \mu(s) = A$$

whenever s approaches 0 through positive real numbers, then

(54) $$\lim_{x \to \infty} \frac{m(x)}{x^{\alpha}} = \frac{A}{\Gamma(\alpha+1)} .$$

This theorem was proved in 1907 by E. Landau [188] in the particular case where $\alpha = 0$. See also E. Landau [193]. In 1914 G. H. Hardy and J. E. Littlewood [153], [154] proved this theorem in the case where $m(x)$ is given by (44) where $a_0, a_1, a_2, \ldots$ are nonnegative real numbers and $\lim_{n \to \infty} \lambda_{n+1}/\lambda_n = 1$. In 1921 G. Doetsch [128] proved Theorem 13 for $\alpha = 1$. From the above mentioned results of Hardy and Littlewood in 1927 E. C. Titchmarsh [237] draw the conclusion that Theorem 13 is generally true. This was proved in 1929 by O. Szász [229], [230] and in 1930 by G. H. Hardy and J. E. Littlewood [159]. O. Szász [230] demonstrated also that if we assume only that

(55) $$\liminf \, [m(y) - m(x)] \geq 0$$

when $y > x$ and x and y tend to infinity in such a way that $y/x \to 1$, then Theorem 13 remains valid unchangeably. For the proof of Theorem 13 see also D. V. Widder [241 p. 192].

In 1931 J. Karamata [174] generalized Theorem 13. As a particular case of Karamata's theorem we have the following result. See also G. H. Hardy [148 p. 166] and G. Doetsch [131 p. 511].

Theorem 14. *If* $m(x)$ $(0 \leq x < \infty)$ *is a nonnegative and nondecreasing function of* x, *if* $\mu(s)$ *is convergent for* $\mathrm{Re}(s) > 0$, *and if*

(56) $$\lim_{s \to +0} \frac{s^{\alpha}\,\mu(s)}{L(1/s)} = A$$

exists as s approaches 0 through positive real numbers for some $\alpha \geq 0$ where $L(x)$ is a slowly varying function of x at $x \to \infty$, then

(57) $$\lim_{x \to \infty} \frac{m(x)}{x^{\alpha} L(x)} = \frac{A}{\Gamma(\alpha+1)} .$$

In generalizing a result of E. Landau [190 p. 874] in 1930 S. Ikehara [165] proved a useful Tauberian theorem which, according to N. Wiener [246 pp. 44-45] and S. Bochner [117], can be formulated in the following way.

Theorem 15. If $m(x)$ $(0 \leq x < \infty)$ is a nonnegative and nondecreasing function of x, if $\mu(s)$ is convergent for $\mathrm{Re}(s) > 1$ and if there exists a constant A such that the function

(58) $$\mu(s) - \frac{A}{s-1}$$

approaches a finite limit uniformly on every finite interval of the line $\mathrm{Re}(s) = 1$ as $\mathrm{Re}(s) \to 1 + 0$, then

(59) $$\lim_{x \to \infty} \frac{m(x)}{e^{x}} = A .$$

See also H. Heilbronn and E. Landau [162], N. Wiener and H. R. Pitt [249], D. A. Raikov [213] and N. I. Achieser [109 p. 238].

In 1928 N. Wiener [244] introduced a new method for proving Tauberian theorems. His fundamental theorem is as follows:

Theorem 16. Let $f(x)$ be a bounded measurable function, defined over $(-\infty, \infty)$. Let $K_1(x)$ be a function in L_1, and let

$$\int_{-\infty}^{\infty} e^{iux} K_1(x)dx \neq 0 \tag{60}$$

for every real u. Let

$$\lim_{x \to \infty} \int_{-\infty}^{\infty} f(u)K_1(u-x)du = A \int_{-\infty}^{\infty} K_1(u)du . \tag{61}$$

Then if $K_2(x)$ is any function in L_1,

$$\lim_{x \to \infty} \int_{-\infty}^{\infty} f(u)K_2(u-x)du = A \int_{-\infty}^{\infty} K_2(u)du . \tag{62}$$

Conversely, let $K_1(x)$ be a function of L_1, and let

$$\int_{-\infty}^{\infty} K_1(x)dx \neq 0 . \tag{63}$$

Let (61) imply (62) whenever $K_2(x)$ belongs to L_1 and $f(x)$ is bounded. Then (60) holds.

See N. Wiener [246 p. 25]. For some extensions and applications of Theorem 16 see H. R. Pitt [201], [202], [203], [204], [206].

10. Rouché's Theorem and Lagrange's Expansion.

In 1861 E. Rouché [267] found a very useful theorem in the theory of complex variables which we present here in a slightly more general form.

Theorem 1. If $f(z)$ and $g(z)$ are regular functions of z in a domain D (open connected set), continuous on the closure of D and satisfy $|g(z)| < |f(z)|$ on the boundary of D, then $f(z)$ and $f(z) + g(z)$ have the same number of zeros in D.

For the proof of this theorem we refer to S. Saks and A. Zygmund [269 p. 157] and E. C. Titchmarsh [272 pp. 116-117].

In 1768 J. L. Lagrange [261] proved the following expansion.

Theorem 2. Let $g(z)$ be a regular function of z in the domain D and continuous on the closure of D. Let a be a point of D and let w be such that the inequality

(1) $$|wg(z)| < |z-a|$$

is satisfied on the boundary of D. Then the equation

(2) $$\zeta = a + wg(\zeta)$$

regarded as an equation in ζ, has exactly one root in D. If $f(z)$ is a regular function of z in D, then we have

(3) $$f(\zeta) = f(a) + \sum_{n=1}^{\infty} \frac{w^n}{n!} \frac{d^{n-1}\{f'(a)[g(a)]^n\}}{da^{n-1}} .$$

For the proof of this theorem we refer to E. T. Whittaker and G. N. Watson [273 p. 132]. Some generalizations of this theorem were given in 1799 by H. Bürmann [262] and in 1900 by F. G. Teixeira [271]. See also C. A. Dixon [256], H. Bateman [250], and E. T. Whittaker and G. N. Watson [273 pp. 128-133].

Finally, we mention the following simple but useful theorem.

Theorem 3. *If* $f(z)$ *is regular for all finite values of* z *and*

(4) $$\lim_{|z| \to \infty} \frac{f(z)}{z^k} = 0$$

for some $k > 0$, *then* $f(z)$ *is a polynomial of degree* $< k$.

The above theorem was found in 1892 by J. Hadamard [257], [258 pp. 118-119]. See also E. C. Titchmarsch [272 pp. 85-86]. This theorem is a generalization of the following theorem: If a function is regular for all finite values of z and is bounded, then it is constant. This latter theorem was found in 1844 by A. Cauchy [254]. C. W. Borchardt [252] named it Liouville's theorem because he heard it in a lecture of J. Liouville in 1847. See also A. Cauchy [255].

REFERENCES

Probability Measures

[1] Alexandroff, A. D., "Additive set-functions in abstract spaces," Matematicheskii Sbornik (Recueil Mathématique) 8 (1940) 307-348, 9 (1941) 563-628, 13 (1943) 169-238.

[2] Andersen, E. S., and B. Jessen, "Some limit theorems on integrals in an abstract set," D. Kgl. Danske Videnks. Selskab, Mat.-fys. Medd. 22 No. 14 (1946) 29 pp.

[3] Billingsley, P., Convergence of Probability Measures. John Wiley and Sons, New York, 1968.

[4] Billingsley, P., Weak Convergence of Measures:Applications in Probability. Regional Conference Series in Applied Mathematics. No. 5. SIAM, 1971.

[5] Blake, L. H., "Canonical extension of measures and the extension of regularity of conditional probabilities," Pacific Journal of Mathematics 41 (1972) 25-32.

[6] Carathéodory, C., "Über das lineare Mass von Punktmengen - eine Verallgemeinerung des Langenbegriffs," Nachrichten der K. Geselschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse (1914) 404-426. [Constantin Carathéodory: Gesammelte Mathematische Schriften. Vol. 4. C. H. Beck, München, 1956, pp. 249-277.]

[7] De Bruijn, N. G., and A. C. Zaanen, "Non σ-finite measures and product measures," Nederl. Akad. Wetensch. Proceedings Ser. A. Math. Sci. 57 (1954) 456-466. [Indagationes Mathematicae 16 (1954) 456-466.]

[8] Doob, J. L., Stochastic Processes. John Wiley and Sons, New York, 1953.

[9] Dudley, R. M., "Distances of probability measures and random variables," The Annals of Mathematical Statistics 39 (1968) 1563-1572.

[10] Gikhman, I. I., and A. V. Skorokhod, Introduction to the Theory of Random Processes. W. B. Saunders, Philadelphia, 1969. [English translation of the Russian original published by Izdat Nauka, Moscow, 1965.]

[11] Halmos, P., Measure Theory, D. van Nostrand, Princeton, 1950.

[12] Jessen, B., "Abstrakt maal- og integralteori. 4," Matematisk Tidskrift B (1939) 7-21.

[13] Kakutani, S., "Notes on infinite product measure spaces, I." Proc. Imper. Acad. Tokyo 19 (1943) 148-151.

[14] Kellerer, H. G., "Funktionen auf Produkträumen mit vorgegebenen Marginal--funktionen," Mathematische Annalen 144 (1961) 323-344.

[15] Kellerer, H. G., "Masstheoretische Marginalprobleme," Mathematische Annalen 153 (1964) 168-198.

[16] Kellerer, H. G., "Marginalprobleme für Funktionen," Mathematische Annalen 154 (1964) 147-156.

[17] Kellerer, H. G., "Schnittmass-Funktionen in mehrfachen Produkträumen," Mathematische Annalen 155 (1964) 369-391.

[18] Kellerer, H. G., "Bemerkung zu einem Satz von H. Richter," Archiv der Mathematik 15 (1964) 204-207.

[19] Kolmogorov, A. N., Foundations of the Theory of Probability. Chelsea, New York, 1950. [English translation of A. Kolmogoroff: Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer, Berlin, 1933.]

[20] Kolmogorov, A. N., and S. V. Fomin, Elements of the Theory of Functions and Functional Analysis. Vol. 2. Measure. The Lebesque Integral. Hilbert Space. Graylock Press, Albany, N. Y., 1961. [English translation of the Russian original published in 1960.]

[21] Landers, D., "The convergence determining class of connected open sets in product spaces," Proceedings of the American Mathematical Society 33 (1972) 529-533.

[22] Łomnicki, Z., and S. Ulam, "Sur la théorie de la mesure dans les espaces combinatoires et son application au calcul des probabilités. I Variables indépendantes," Fundamenta Mathematicae 23 (1934) 237-278.

[23] Mařík, J., "The Baire and Borel measures," Czechoslovak Mathematical Journal 7 (1957) 248-253.

[24] Neumann, J. v., Functional Operators. Vol. I. Measures and Integrals. Princeton University Press, 1950.

[25] Prokhorov, Yu. V., "Convergence of random processes and limit theorems in probability theory," Theory of Probability and its Applications 1 (1956) 157-214.

[26] Pyke, R., "Applications of almost surely convergent constructions of weakly convergent processes," Probability and Information Theory. Proceedings of the International Symposium at McMaster University, Canada, April 1968. Lecture Notes in Mathematics. No. 89, Springer, Berlin, 1969, pp. 187-200.

[27] Sikorski, R., Boolean Algebras. Third edition. Springer-Verlag, New York, 1969.

[28] Sikorski, R., and B. Znojkiewicz, "A proof of the representation theorem for the space of random variables," Prace Matematyczne 11 (1967) 175-178.

[29] Skorokhod, A. V., "Limit theorems for stochastic processes," Theory of Probability and its Applications 1 (1956) 261-290.

[30] Strassen, V., "The existence of probability measures with given marginals," The Annals of Mathematical Statistics 36 (1965) 423-439.

[31] Sz.- Nagy, B., Introduction to Real Functions and Orthogonal Expansions. Oxford University Press, New York, 1965. [English translation

Wald's Theorem

[32] Bellman, R., "On a generalization of the fundamental identity of Wald," Proceedings of the Cambridge Philosophical Society 53 (1957) 257-259.

[33] Blackwell, D., "On an equation of Wald," The Annals of Mathematical Statistics 17 (1946) 84-87.

[34] Blackwell, D., and M. A. Girshick, "On functions of sequences of independent chance vectors with applications to the problem of the 'random walk' in k dimensions," The Annals of Mathematical Statistics 17 (1946) 310-317.

[35] Harris, T. E., "Note on differentiation under the expectation sign in the fundamental identity of sequential analysis," The Annals of Mathematical Statistics 18 (1947) 294-295.

[36] Johnson, N. L., "A proof of Wald's theorem of cumulative sums," The Annals of Mathematical Statistics 30 (1959) 1245-1247.

[37] Kolmogorov, A. N., and Yu. V. Prokhorov, "On sums of a random number of random variables," (Russian) Uspekhi Matem. Nauk 4 (1949) 168-172.

[38] Miller, H. D., "A generalization of Wald's identity with applications to random walks," The Annals of Mathematical Statistics 32 (1961) 549-560.

[39] Seitz, J., and K. Winkelbauer, "Remark concerning a paper of Kolmogorov and Prochorov," (Russian) Czechoslovak Mathematical Journal 3 (1953) 89-91.

[40] Stein, Ch., "A note on cumulative sums," The Annals of Mathematical Statistics 17 (1946) 498-499.

[41] Wald, A., "On cumulative sums of random variables," The Annals of Mathematical Statistics 15 (1944) 283-296.

[42] Wald, A., "Differentiation under the expectation sign in the fundamental identity of sequential analysis," The Annals of Mathematical Statistics 17 (1946) 493-497. [Reprinted in Selected Papers in Statistics and Probability by Abraham Wald. Edited by T. W. Andersen et al., McGraw-Hill, New York, 1955, pp. 469-373.]

[43] Winkelbauer, K., "Moments of cumulative sums of random variables," (Russian) Czechoslovak Mathematical Journal 3 (1953) 93-108.

[44] Wolfowitz, J., "The efficiency of sequential estimates and Wald's equation for sequential processes," The Annals of Mathematical Statistics 18 (1947) 215-230.

of the second Hungarian edition published by Tankönyvkiadó, Budapest, 1961. First edition 1954.]

Interchangeable Random Variables

[45] Benczur, A., "On sequences of equivalent events and the compound Poisson process," Studia Scientiarum Mathematicarum Hungarica 3 (1968) 451-458.

[46] Berman, S. M., "An extension of the arc sine law," The Annals of Mathematical Statistics 33 (1962) 681-684.

[47] Blom, G., "On the asymptotic distribution of linear combinations of interchangeable random variables," Mathematica Scandinavica 7 (1959) 321-332.

[48] Blum, J. R., H. Chernoff, M. Rosenblatt, and H. Teicher, "Central limit theorems for interchangeable processes," Canadian Journal of Mathematics 10 (1958) 222-229.

[49] Bühlmann, H., "Le problème 'limite central' pour les variables aléatoires échangeables," Comptes Rendus Acad. Sci. Paris 246 (1958) 534-536.

[50] Bühlmann, H., "Austauschbare stochastische Variabeln und ihre Grenzwertsaetze," University of California Publications in Statistics 3 No. 1 (1960) 1-36.

[51] Chernoff, H., and H. Teicher, "A central limit theorem for sums of interchangeable random variables," The Annals of Mathematical Statistics 29 (1958) 118-130.

[52] De Finetti, B., "Funzione caratteristica di un fenomeno aleatorio," Memorie R. Accad. Naz. Lincei Cl. Sci. Fis. (6) 4 (1930) 86-133.

[53] De Finetti, B., "Classi di numeri aleatori equivalenti," Atti R. Accad. Naz. Lincei, Rendiconti, Cl. Sci. Fis. Mat. Nat. (6) 18 (1933) 107-110.

[54] De Finetti, B., "La legge dei grandi numeri nel caso dei numeri equivalenti," Atti R. Accad. Naz. Lincei, Rendiconti, Classe di Scienze, Fisiche, Matematiche e Naturali. (6) 18 (1933) 203-207.

[55] De Finetti, B., "Sulla legge di distribuzione dei valori in una successione di numeri aleatori equivalenti," Atti Reale Accademia Naz. Lincei, Rendiconti, Classe di Scienze, Fisiche, Matematiche e Naturali (6) 18 (1933) 279-284.

[56] De Finetti, B., "La prévision: ses lois logiques, ses sorces subjectives," Annales de l'Institut Henri Poincaré 7 (1937) 1-68. [English translation: "Foresight: Its logical laws, its subjective sources," in Studies in Subjective Probability. Ed. H. E. Kyburg, Jr., and H. E. Smokler. John Wiley and Sons, New York, 1964, pp. 93-158.]

[57] Dwass, M., and S. Karlin, "Conditioned limit theorems," The Annals of Mathematical Statistics 34 (1963) 1147-1167.

[58] Dynkin, E. B., "Classes of equivalent random variables," (Russian) Uspechi Mat. Nauk. 8 (1953) 125-130.

[59] Eggenberger, F., and G. Pólya, "Über die Statistik verketteter Vorgänge," Zeitschrift für angewandte Mathematik und Mechanik 3 (1923) 279-289.

[60] Feller, W., An Introduction to Probability Theory. Vol. II. John Wiley and Sons, New York, 1966. [Second edition 1971.]

[61] Freedman, D. A., "Invariants under mixing which generalize de Finetti's theorem," The Annals of Mathematical Statistics 33 (1962) 916-923.

[62] Freedman, D., "L'urne de Bernard Friedman," Comptes Rendus Acad. Sci. Paris 257 (1963) 3809.

[63] Freedman, D. A., "Bernard Friedman's urn," The Annals of Mathematical Statistics 36 (1965) 956-970.

[64] Friedman, B., "A simple urn model," Communications on Pure and Applied Mathematics 2 (1949) 59-70.

[65] Hausdorff, F., "Summationsmethoden und Momentfolgen. I-II," Mathematische Zeitschrift 9 (1921) 74-109 and 280-299.

[66] Hausdorff, F., "Momentprobleme für ein endliches Intervall," Mathematische Zeitschrift 16 (1923) 220-248.

[67] Hewitt, E., and L. J. Savage, "Symmetric measures on Cartesian products," Transactions of the American Mathematical Society 80 (1955) 470-501.

[68] Kendall, D. G., "On finite and infinite sequences of exchangeable events," Studia Scientiarum Mathematicarum Hungarica 2 (1967) 319-327.

[69] Khintchine, A., "Sur les classes d'événements équivalents," Matematicheskii Sbornik 39 (1932) 40-43.

[70] Khintchine, A. Ya., "On classes of equivalent events," (Russian) Doklady Akad. Nauk SSSR 85 (1952) 713-714.

[71] Loève, M., Probability Theory. Third edition. Van Nostrand, Princeton, 1963. [First edition 1955. Second edition 1960.]

[72] Milicer-Gruzewska, H., "Sulla legge limite delle variabili casuali equivalenti," Atti Accad. Naz. Lincei Memorie Cl. Sci. Fis. Mat. e Nat. Sez. I. (8) 2 (1948) 25-33.

[73] Milicer-Gruzewska, H., "On the law of probability and the characteristic function of the standardized sum of equivalent variables," Comptes Rendus des Séances de la Soc. des Sciences et des Lettres de Varsovie Cl. III. Sci. Math. Phys. 42 (1949) 99-143.

[74] Ryll- Nardzewski, C., "On stationary sequences of random variables and the De Finetti's equivalence," Colloquium Mathematicum 4 (1957) 149-156.

[75] Rényi, A., "On stable sequences of events," Sankhyā: The Indian Journal of Statisitcs. Ser. A 25 (1963) 293-302.

[76] Rényi, A., and P. Révész, "A study of sequences of equivalent events as special stable sequences," Publicationes Mathematicae (Debrecen) 10 (1963) 319-325.

[77] Takács, L., "On the method of inclusion and exclusion," Journal of the American Statistical Association 62 (1967) 102-113.

Slowly Varying Functions

[78] Aljančić, S., R. Bojanić and M. Tomić, "Sur la valeur asymptotique d'une classe des intégrales définies," Publications de l'Institut Mathématique de l'Académie Serbe de Sciences 7 (1954) 81-94.

[79] Békéssy, A., "Eine Verallgemeinerung der Laplaceschen Methode," Publications of the Mathematical Institute of the Hungarian Academy of Sciences 2 (1957) 105-125.

[80] Birnbaum, Z. W., and W. Orlicz, "Über die Verallgemeinerung des Begriffes der zueinander konjugierten Potenzen," Studia Mathematica 3 (1931) 1-67.

[81] Blumberg, H., "On convex functions," Transactions of the American Mathematical Society 20 (1919) 40-44.

[82] Bojanić, R., and E. Seneta, "Slowly varying functions and asymptotic relations," Journal of Mathematical Analysis and Applications 34 (1971) 302-315.

[83] Csiszár, I., and P. Erdös, "On the function $g(t) = \limsup_{x\to\infty}(f(x+t)-f(x))$," Publications of the Mathematical Institute of the Hungarian Academy of Sciences 9 (1965) 603-606.

[84] De Bruijn, N. G., "Pairs of slowly oscillating functions occurring in asymptotic problems concerning the Laplace transform," Nieuw Archief voor Wiskunde (3) 7 (1959) 20-26.

[85] Delange, H., "Sur un théorème de Karamata," Bulletin des Sciences Mathématiques 79 (1955) 9-12.

[86] Feller, W., "On regular variation and local limit theorems," Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Vol. II. Part I. University of California Press, 1967, pp. 373-388.

[87] Feller, W., "One-sided analogues of Karamata's regular variation," L'Enseignement Mathématique 15 (1969) 107-121.

[88] Haan, L. de, "A form of regular variation and its application to the domain of attraction of the double exponential distribution," Zeitschrift für Wahrscheinlichkeitstheorie und verw. Gebiete 17 (1971) 241-258.

[89] Hamel, G., "Eine Basis aller Zahlem und die unstetige Lösungen der Funktionalgleichung: $f(x+y) = f(x)+f(y)$," Mathematische Annalen 60 (1905) 459-462.

[90] Hardy, G. H., and W. W. Rogosinski, "Notes on Fourier series (III): Asymptotic formulae for the sums of certain trigonometrical series," The Quarterly Journal of Mathematics (Oxford) 16 (1945) 49-58.

[91] Heyde, C. C., "Some local limit results in fluctuation theory," Journal of the Australian Mathematical Society 7 (1967) 455-464.

[92] Karamata, J., "Sur un mode de croissance régulière des fonctions," Mathematica (Cluj) 4 (1930) 38-53.

[93] Karamata, J., "Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacesche und Stieltjessche Transformation betreffen," Journal für reine und angewandte Mathematik 164 (1931) 27-39.

[94] Karamata, J., "Sur un mode de croissance régulière théorèmes fondamentaux," Bulletin de la Societé Mathématique de France 61 (1933) 55-62.

[95] Kohlbecker, E. E., "Weak asymptotic properties of partitions," Transactions of the American Mathematical Society 88 (1958) 346-365.

[96] Korevaar, J., T. van Aardenne-Ehrenfest, and N. G. de Bruijn, "A note on slowly oscillating functions," Nieuw Archief voor Wiskunde (2) 23 (1949) 77-86.

[97] Landau, E., "Sur les valeurs moyennes de certaines fonctions arithmétiques," Bulletin de l'Académie Royale de Belgique Cl. Sci. (1911) 443-472.

[98] Marcus, M., and M. Pinsky, "On the domain of attraction of $e^{-e^{-x}}$," Journal of Mathematical Analysis and Applications 28 (1969) 440-449.

[99] Matuszewska, W., "Regularly increasing functions in connection with the theory of L*φ-spaces," Studia Mathematica 21 (1962) 317-344.

[100] Matuszewska, W., "A remark on my paper Regularly increasing functions in connection with the theory of L*φ-spaces," Studia Mathematica 25 (1965) 265-269.

[101] Parameswaran, S., "Partition functions whose logarithms are slowly oscillating," Transactions of the American Mathematical Society 100 (1961) 217-240.

[102] Pólya, G., "Über eine neue Weise bestimmte Integrale in der analytischen Zahlentheorie zu gebrauchen," Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse (1917) 149-159.

[103] Pólya, G., "Bemerkungen über unendliche Folgen und ganze Funktionen," Mathematische Annalen 88 (1923) 169-183.

[104] Pólya, G., and G. Szegö, Aufgaben und Lehrsatze aus der Analysis. Band I. Springer, Berlin, 1925. [English translation: G. Pólya and G. Szegö, Problems and Theorems in Analysis. Vol. I. Springer, New York, 1972.]

[105] Pringsheim, A., "Über den Divergenz-Character gewisser Potenzreihen an der Convergenzgrenze," Acta Mathematica 28 (1904)1-30.

[106] Rogozin, B. A., "The distribution of the first ladder moment and height and fluctuation of a random walk," Theory of Probability and its Applications 16 (1971) 575-595.

[107] Schmidt, R., "Über divergente Folgen und lineare Mittelbildungen," Mathematische Zeitschrift 22 (1925) 89-152.

Abelian and Tauberian Theorems

[108] Abel, N. H., "Untersuchungen über die Reide: $1+\frac{m}{1}x+\frac{m(m-1)}{1.2}x^2+\frac{m(m-1)(m-2)}{1.2.3}x^3+\ldots$ u.s.w.," Journal für die reine und angewandte Mathematik 1 (1826) 311-339. [English translation of a part in D. E. Smith: A Source Book in Mathematics. Vol. 1, Dover, New York, 1959, pp. 286-291.]

[109] Achieser, N. I., Theory of Approximation. F. Ungar, 1956. [English translation of the Russian original published by Gostehizdat, Moscow, 1947.]

[110] Ananda Rau, K., "A note on a theorem of Mr. Hardy's," Proceedings of the London Mathematical Society (2) 17 (1918) 334-336.

[111] Ananda-Rau, K., "On the converse of Abel's theorem," Journal of the London Mathematical Society 3 (1928) 200-205.

[112] Ananda-Rau, K., "An example in the theory of summation of series by Riesz's typical means," Proceedings of the London Mathematical Society (2) 30 (1930) 367-372.

[113] Appell, P., "Sur certaines séries ordonnées par rapport aux puissances d'une variable," Comptes Rendus Acad. Sci. Paris 87 (1878) 689-692.

[114] Axer, A., "Beitrag zur Kenntnis der zahlentheoretischen Funktionen $\mu(n)$ and $\lambda(n)$," Prace matematyczno-fizyczne 21 (1910) 65-95.

[115] Beneš, V. E., "Extensions of Wiener's Tauberian theorem for positive measures," Journal of Mathematical Analysis and Applications 2 (1961) 1-20.

[116] Beurling, A., "Un théorème sur les fonctions bornées et uniformément continues sur l'axe réel," Acta Mathematica 77 (1945) 127-136.

[117] Bochner, S., "Ein Satz von Landau und Ikehara," Mathematische Zeitschrift 37 (1933) 1-9.

[118] Bochner, S., "An extension of a Tauberian theorem on series with positive terms," Journal of the London Mathematical Society 9 (1934) 141-148.

[119] Bosanquet, L. S., "Note on the converse of Abel's theorem," Journal of the London Mathematical Society 19 (1944) 161-168.

[120] Bosanquet, L. S., and M. L. Cartwright, "Some Tauberian theorems," Mathematische Zeitschrift 37 (1933) 416-423.

[121] Bromwich, T. J. I'A, and G. H. Hardy, "Some extensions to multiple series of Abel's theorem on the continuity of power series," Proceedings of the London Mathematical Society (2) 2 (1905) 161-189.

[122] Cahen, E., "Sur la fonction $\zeta(s)$ de Riemann et sur des fonctions analogues," Annales Scientifiques de l'École Normale Supérieure (3) 11 (1894) 75-164.

[123] Delange, H., "Sur la réciproque du théorème d'Abel sur les séries entières," Comptes Rendus Acad. Sci. Paris 224 (1947) 436-438.

[124] Delange, H., "The converse of Abel's theorem on power series," Annals of Mathematics (2) 50 (1949) 94-109.

[125] Dirichlet, G. L., "Recherches sur diverses applications de l'analyse infinitésimale à la théorie des nombres. I-II." Journal für die reine und angewandte Mathematik 19 (1839) 324-369, and 21 (1840) 1-12 and 134-155.

[126] Dirichlet, L., "Démonstration d'un théorème d'Abel," Journal de Mathématiques Pures et Appliquées (2) 7 (1862) 253-255.

[127] Dirichlet, P. G. L., Vorlesungen über Zahlentheorie. Supplemented by R. Dedekind. (First edition 1863.) Second edition 1871. F. Vieweg, Braunschweig.

[128] Doetsch, G., "Ein Konvergenzkriterium für Integrale," Mathematische Annalen 82 (1921) 68-82.

[129] Doetsch, G., "Sätze von Tauberschem Character im Gebiet der Laplace- und Stieltjes - Transformation," Sitzungsberichte der Preussischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse. Berlin Akademie der Wissenschaften (1930) 144-157.

[130] Doetsch, G., Theorie und Anwendung der Laplace-Transformation. Springer, Berlin, 1937.

[131] Doetsch, G., Handbuch der Laplace-Transformation. Band I. Theorie der Laplace-Transformation. Birkhäuser, Basel, 1950.

[132] Doetsch, G., Handbuch der Laplace-Transformation. Band II and III. Anwendungen der Laplace-Transformation. Birkhäuser, Basel, 1955, 1956.

[133] Drasin, D., "Tauberian theorems and slowly varying functions," Transactions of the American Mathematical Society 133 (1968) 333-356.

[134] Feller, W., "On the Classical Tauberian theorems," Archiv der Mathematik 14 (1963) 317-322.

[135] Franel, J., "Sur la théorie des séries," Mathematische Annalen 52 (1899) 529-549.

[136] Freud, G., "Restglied eines Tauberschen Satzes. I, II, III," Acta Mathematica Acad. Sci. Hungaricae 2 (1951) 299-308, 3 (1952) 299-307, 5 (1954) 275-289.

[137] Frobenius, G., "Ueber die Leibnitzsche Reihe," Journal für die reine und angewandte Mathematik 89 (1880) 262-264.

[138] Ganelius, T., "Un théorème taubérien pour la transformation de Laplace," Comptes Rendus Acad. Sci. Paris 242 (1956) 719-721.

[139] Ganelius, T., "Tauberian theorems for the Stieltjes transform," Mathematica Scandinavica 14 (1964) 213-219.

[140] Ganelius, T. H., Tauberian Theorems. Lecture Notes in Mathematics. No. 232. Springer, Berlin, 1971.

[141] Haar, A., "Über asymptotische Entwicklungen von Funktionen," Mathematische Annalen 96 (1926) 69-107. [Alfréd Haar:Gesammelte Arbeiten. Akadémiai Kiadó, Budapest, 1959, pp. 174-212.]

[142] Hardy, G. H., "On certain oscillating series," Quarterly Journal of Mathematics 38 (1907) 269-288.

[143] Hardy, G. H., "Some theorems connected with Abel's theorem on the continuity of power series," Proceedings of the London Mathematical Society (2) 4 (1907) 247-265.

[144] Hardy, G. H., "Theorems relating to the summability and convergence of slowly oscillating series," Proceedings of the London Mathematical Society (2) 8 (1910) 301-320.

[145] Hardy, G. H., "An extension of a theorem on oscillating series," Proceedings of the London Mathematical Society (2) 12 (1913) 174-180.

[146] Hardy, G. H., "Sur la sommation des séries de Dirichlet," Comptes Rendus Acad. Sci. Paris 162 (1916) 463-466.

[147] Hardy, G. H., "The application of Abel's method of summation to Dirichlet's series," Quarterly Journal of Mathematics 47 (1916) 176-192.

[148] Hardy, G. H., Divergent Series. Oxford University Press, 1949.

[149] Hardy, G. H., and S. Chapman, "A general view of the theory of summable series," The Quarterly Journal of Pure and Applied Mathematics 42 (1911) 181-215.

[150] Hardy, G. H., and J. E. Littlewood, "The relations between Borel's and Cesàro's methods of summation," Proceedings of the London Mathematical Society (2) 11 (1912-13) 1-16.

[151] Hardy, G. H., and J. E. Littlewood, "Contributions to the arithmetic theory of series," Proceedings of the London Mathematical Society (2) 11 (1912-13) 411-478.

[152] Hardy, G. H., and J. E. Littlewood, "Tauberian theorems concerning series of positive terms," Messenger of Mathematics 42 (1913) 191-192.

[153] Hardy, G. H., and J. E. Littlewood, "Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive," Proceedings of the London Mathematical Society (2) 13 (1914) 174-191.

[154] Hardy, G., and J. E. Littlewood, "Some theorems concerning Dirichlet's series," Messenger of Mathematics 43 (1914) 134-147.

[155] Hardy, G. H., and J. E. Littlewood, "Abel's theorem and its converse," Proceedings of the London Mathematical Society (2) 18 (1920) 205-235.

[156] Hardy, G. H., and J. E. Littlewood, "On a Tauberian theorem for Lambert's series, and some fundamental theorems in the analytic theory of numbers," Proceedings of the London Mathematical Society (2) 19 (1921) 21-29.

[157] Hardy, G. H., and J. E. Littlewood, "Abel's theorem and its converse (II)," Proceedings of the London Mathematical Society (2) 22 (1923) 254-269.

[158] Hardy, G. H., and J. E. Littlewood, "A further note on the converse of Abel's theorem," Proceedings of the London Mathematical Society (2) 25 (1926) 219-236.

[159] Hardy, G. H., and J. E. Littlewood, "Notes on the theory of series (XI): On Tauberian theorems," Proceedings of the London Mathematical Society (2) 30 (1930) 23-37.

[160] Hardy, G. H., and M. Riesz, The General Theory of Dirichlet's Series. Cambridge University Press, 1915. [Reprinted by Stechert-Hafner, New York, 1964.]

[161] Heilbronn, H., and E. Landau, "Bemerkungen zur vorstehenden Arbeit von Herrn Bochner," Mathematische Zeitschrift 37 (1933) 10-16.

[162] Heilbronn, H., and E. Landau, "Anwendungen der N. Wienerischen Methode," Mathematische Zeitschrift 37 (1933) 18-21.

[163] Hölder, O., "Grenzwerthe von Reihen an der Convergenzgrenze," Mathematische Annalen 20 (1882) 535-549.

[164] Hurwitz, W. A., "A trivial Tauberian theorem," Bulletin of the American Mathematical Society 32 (1926) 77-82.

[165] Ikehara, S., "An extension of Landau's theorem in the analytical theory of numbers," Journal of Mathematics and Physics M.I.T. 10 (1930-1931) 1-12.

[166] Ikehara, S., "On Tauberian theorems of Hardy and Littlewood and a note on Wintner's paper," Journal of Mathematics and Physics M.I.T. 10 (1930-1931) 75-83.

[167] Ingham, A. E., "On the 'high-indices theorem' of Hardy and Littlewood. Quarterly Journal of Mathematics (Oxford) 8 (1937) 1-7.

[168] Ingham, A. E., "On Tauberian theorems," Proceedings of the London Mathematical Society (3) 14a (1965) 157-173.

[169] Izumi, S., "A simple proof of Littlewood's Tauberian theorem," Proceedings of the Japan Academy 30 (1954) 927-929.

[170] Jensen, J. L. W. V., "Om raekkers konvergens," Tidskrift for Matematik (5) 2 (1884) 63-72.

[171] Kalatalova, M. A., "Tauberian type theorems for double series," (Russian) Ukrain. Mat. Zhur. 23 (1971) 733-744. [English translation: Ukrain Mathematical Journal 23 (1971) 597-606.]

[172] Karamata, J., "Über die Hardy-Littlewoodschen Umkehrungen des Abelschen Stetigkeitssatzes," Mathematische Zeitschrift 32 (1930) 319-320.

[173] Karamata, J., "Neuer Beweis und Verallgemeinerung einiger Tauberian-Sätze," Mathematische Zeitschrift 33 (1931) 294-299.

[174] Karamata, J., "Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplaceshce und Stieltjessche Transformation betreffen," Journal für reine und angewandte Mathematik 164 (1931) 27-39.

[175] Karamata, J., "Über einen Satz von Vijayaraghavan," Mathematische Zeitschrift 34 (1932) 737-740.

[176] Karamata, J., "Über die O-Inversionssätze der Limitierungsverfahren," Mathematische Zeitschrift 37 (1933) 582-588.

[177] Karamata, J., "Weiterführung der N. Wienerschen Methode," Mathematische Zeitschrift 38 (1934) 701-708.

[178] Karamata, J., "Sur les théorèmes inverses des procédés de sommabilité," (La théorie des fonctions VI.) Actualités Scientifiques et Industrielles N° 450, Hermann et C^ie^, Paris, 1937, pp. 1-47.

[179] Knopp, K., Theory and Applications of Infinite Series. Blackie and Son, London and Glasgow, 1951. [Second English edition translated from the Fourth German edition published in 1947.]

[180] König, H., "Neuer Beweis eines klassischen Tauber-Satzes," Archiv der Mathematik 11 (1960) 278-279.

[181] Korevaar, J., "An estimate of the error in Tauberian theorems for power series," Duke Mathematical Journal 18 (1951) 723-734.

[182] Korevaar, J., "Best L_1 approximation and the remainder in Littlewood's theorem," Indagationes Mathematicae 15 (1953) 281-293. [Nederl. Akad. Wetensch. Proc. Ser. A. Math. Sci. 56 (1953) 281-293.]

[183] Korevaar, J., "A very general form of Littlewood's theorem." Indagationes Mathematicae 16 (1954) 36-45. [Nederl. Akad. Wetensch. Proceedings Ser. A. Math. Sci. 57 (1954) 36-45.]

[184] Korevaar, J., "Another numerical Tauberian theorem for power series," Indagationes Mathematicae 16 (1954) 46-56. [Nederl. Akad. Wetensch. Proceedings Ser. A Math. Sci. 57 (1954) 46-56.]

[185] Korevaar, J., "Numerical Tauberian theorems for Dirichlet and Lambert series," Indagationes Mathematicae 16 (1954) 152-160. [Nederl. Akad. Wetensch. Proceedings. Ser. A. Math. Sci. 57 (1954) 152-160.]

[186] Korevaar, J., "Numerical Tauberian theorems for power series and Dirichlet series, I, II," Indagationes Mathematicae 16 (1954) 432-443 and 444-455. [Nederl. Akad. Wetensch. Proceedings. Ser. A. Math. Sci. 57 (1954) 432-443 and 444 -455.]

[187] Kronecker, L., "Quelques remarques sur la détermination des valeurs moyennes," Comptes Rendus Acad. Sci. Paris 103 (1886) 980-987.

[188] Landau, E., "Über die Konvergenze einiger Klassen von unendlichen Reihen am Rande des Konvergenzgebietes," Monatshefte für Mathematik und Physik 18 (1907) 8-28.

[189] Landau, E., "Über das Konvergenzproblem der Dirichlet'schen Reihen," Rendiconti del Circolo Matematico di Palermo 28 (1909) 113-151.

[190] Landau, E., Handbuch der Lehre von der Verteilung der Primzahlen. I - II.B.G. Teubner, Leipzig and Berlin, 1909.

[191] Landau, E., "Über die Bedeutung einiger neuen Grenzwertsätze der Herren Hardy and Axer," Prace matematyczno-fizyczne 21 (1910) 97-177.

[192] Landau, E., "Über einen Satz des Herrn Littlewood," Rendiconti del Circolo Matematico di Palermo 35 (1913) 265-276.

[193] Landau, E., "Ein neues Konvergenzkriterium für Integrale," Sitzungsberichte der mathematisch-physikalischen Klasse der Kgl. Bayerischen Akademie der Wissenschaften (1913) 461-467.

[194] Lasker, E., "Über Reihen auf der Convergenzgrenze," Philosophical Transactions of the Royal Society of London, Ser. A 196 (1901) 431-477.

[195] Littlewood, J. E., "The converse of Abel's theorem on power series," Proceedings of the London Mathematical Society (2) 9 (1911) 434-448.

[196] Malliavin, P., "Un théorème taubérien avec reste pour la transformée de Stieltjes," Comptes Rendus Acad. Sci. Paris 255 (1962) 2351-2352.

[197] Minakshi Sundaram, S., "On generalized Tauberian theorems," Mathematische Zeitschrift 45 (1939) 495-506.

[198] Minakshisundaram, S., "A Tauberian theorem on (λ,k) - process of summation," Journal of the Indian Mathematical Society 3 (1939) 127-130.

[199] Minakshisundaram, S., "On V. Ramaswami's Tauberian theorem on oscillation," Journal of the Indian Mathematical Society 3 (1939) 131-135.

[200] Neder, L., "Über Taubersche Bedingungen," Proceedings of the London Mathematical Society (2) 23 (1924) 172-184.

[201] Pitt, H. R., "A remark on Wiener's general Tauberian theorem," Duke Mathematical Journal 4 (1938) 437-440.

[202] Pitt, H. R., "An extension of Wiener's general Tauberian theorem," American Journal of Mathematics 60 (1938) 532-534.

[203] Pitt, H. R., "General Tauberian theorems," Proceedings of the London Mathematical Society 44 (1938) 243-288.

[204] Pitt, H. R., "General Tauberian theorems, II," Journal of the London Mathematical Society 15 (1940) 97-112.

[205] Pitt, H. R., "A note on some elementary Tauberian theorems," Quarterly Journal of Mathematics 19 (1948) 177-180.

[206] Pitt, H. R., Tauberian Theorems. Oxford University Press, 1958.

[207] Pleijel, Å., "On a theorem by P. Malliavin," Israel Journal of Mathematics 1 (1963) 166-168.

[208] Postnikov, A. G., "The remainder term in the Tauberian theorem of Hardy and Littlewood," (Russian) Doklady Akademii Nauk SSSR (N.S.) 77 (1951) 193-196.

[209] Pringsheim, A., "Allgemeine Theorie der Divergenz und Convergenz von Reien mit positiven Glidern," Mathematische Annalen 35 (1890) 297-394.

[210] Pringsheim, A., "Über das Verhalten von Potenzreihen auf dem Convergenzkreise," Sitzungsberichte der mathematisch-physikalischen Klasse der Kgl. Bayerischen Akademie der Wissenschaften zu München 30 (1900) 37-100.

[211] Pringsheim, A., "Ueber die Divergenz gewisser Potenzreihen an der Convergenzgrenze," Sitzungsberichte der mathematisch-physikalischen Klasse der Kgl. Bayerischen Akademie der Wissenschaften zu München 31 (1901) 505-524.

[212] Pringsheim, A., "Über den Divergenz-Character gewisser Potenzreihen an der Convergenzgrenze," Acta Mathematica 28 (1904) 1-30.

[213] Raikov, D., "Generalization of the Ikehara-Landau theorem," (Russian) Matematicheskii Sbornik (Recueil Mathématique) 3 (1938) 559-568.

[214] Rajagopal, C. T., "On converse theorems of summability," Mathematical Gazette 30 (1946) 272-276.

[215] Rajagopal, C. T., "On a generalization of Tauber's theorem," Commentarii Mathematici Helvetici 24 (1950) 219-231.

[216] Rajagopal, C. T., "Two one-sided Tauberian theorems," Archiv der Mathematik 3 (1952) 108-113.

[217] Rajagopal, C. T., "A generalization of Tauber's theorem and some Tauberian constants," Mathematische Zeitschrift 57 (1953) 405-414.

[218] Ramaswami, V., "Some Tauberian theorems on oscillation," Journal of the London Mathematical Society 10 (1935) 294- 308.

[219] Ramaswami, V., "The generalized Abel-Tauber theorem," Proceedings of the London Mathematical Society (2) 41 (1936) 408-417.

[220] Rényi, A., "On a Tauberian theorem of O. Szász," Acta Scientiarum Mathematicarum (Szeged) 11 (1946-1948) 119-123.

[221] Riesz, M., "Über die Summierbarkeit durch typische Mittel," Acta Scientiarum Mathematicarum (Szeged) 2 (1924) 18-31.

[222] Schmidt, R., "Über divergente Folgen und lineare Mittelbildungen," Mathematische Zeitschrift 22 (1925) 89-152.

[223] Shtshegloff, M., "To the question of the behaviour of a power series on the circle of convergence," (Russian) Matematicheskii Sbornik (Recueil Mathématique) 14 (1944) 109-132.

[224] Stolz, O., "Beweis einiger Sätze über Potenzreihen," Zeitschrift für Mathematik und Physik 20 (1875) 369-376.

[225] Stolz, O., "Nachtrag zur Mittheilung 'Beweis einiger Sätze über Potenzreihen," Zeitschrift für Mathematik und Physik 29 (1884) 127-128.

[226] Szász, O., "Ein Grenzwertsatz über Potenzreihen," Sitzungsberichte der Berliner Math. Gesellschaft 21 (1921) 25-29. [Reprinted in Otto Szász: Collected Mathematical Papers. Editor: H. D. Lipsich. Hafner, New York, 1955, pp. 515-519.]

[227] Szász, O., "Ueber Dirichletsche Reihen an der Konvergenzgrenze," Atti del Congresso Internazionale dei Matematici, September 3-10, 1928, Bologna, pp. 269-276. [Reprinted in Otto Szász: Collected Mathematical Papers. Edited by H. D. Lipsich. Hafner, New York, pp. 594-601.]

[228] Szász, O., "Verallgemeinerung eines Littlewoodschen Satzes über Potenzreihen," Journal of the London Mathematical Society 3 (1928) 254-262. [Reprinted in Otto Szász: Collected Mathematical Papers. Edited by H. D. Lipsich. Hafner, New York, 1955, pp. 556-564.]

[229] Szász, O., "Über neuere Untersuchungen im Zusammenhange mit dem Ablel'-schen Potenzreihen-Satz," (Hungarian) Matematikai és Fizikai Lapok 36 (1929) 10-22. [Reprinted in Otto Szász: Collected Mathematical Papers. Edited by H. D. Lipsich. Hafner, New York, 1955, pp. 565-577.]

[230] Szász, O., "Verallgemeinerung und neuer Beweis einige Sätze Tauberscher Art," Sitzungsber. der Bayer. Akad. der Wiss. Math.- Phys. Kl. (1929) 325-340. [Reprinted in Otto Szász: Collected Mathematical Papers. Ed. H. D. Lipsich. Hafner, New York, 1955, pp. 578-593.]

[231] Szász, O., "Über einige Sätze von Hardy und Littlewood," Nachr. d. Ges. der Wiss. zu Göttingen. I. Mathematik. (1930) 315-333. [Reprinted in Otto Szász: Collected Mathematical Papers. Ed. H. D. Lipsich. Hafner, New York, 1955. pp. 606-624.]

[232] Szász, O., "Über einen Satz von Hardy und Littlewood," Sitzungsberichten der Preuss. Akademie der Wissenschaften. Physikalisch-mathematische Klasse (1930) 470-473. [Reprinted in Otto Szász: Collected Mathematical Papers. Ed. H. D. Lipsich. Hafner, New York, 1955, pp. 602-605.]

[233] Szász, O., "Generalization of two theorems of Hardy and Littlewood on power series," Duke Mathematical Journal 1 (1935) 105-111. [Reprinted in Otto Szász: Collected Mathematical Papers. Ed. H. D. Lipsich. Hafner, New York, 1955. pp. 625-631.]

[234] Szász, O., "Converse theorems of summability for Dirichlet series," Transactions of the American Mathematical Society 39 (1936) 117-130. [Reprinted in Otto Szász: Collected Mathematical Papers. Edited by H. D. Lipsich, Hafner, New York, 1955, pp. 632-645.]

[235] Szász, O., Collected Mathematical Papers. Ed. H. D. Lipsich. Hafner, New York, 1955.

[236] Tauber, A., "Ein Satz aus der Theorie der unendlichen Reihen," Monatshefte für Mathematik und Physik 8 (1897) 273-277.

[237] Titchmarsh, E. C., "On integral functions with real negative zeros," Proceedings of the London Mathematical Society (2) 26 (1927) 185-200.

[238] Titchmarsh, E. C., The Theory of Functions. Second edition. Oxford University Press, 1939. (First edition 1932.)

[239] Vijayaraghavan, T., "A Tauberian theorem," Journal of the London Mathematical Society 1 (1926) 113-120.

[240] Vijayaraghavan, T., "Converse theorems on summability," Journal of the London Mathematical Society 2 (1927) 215-222.

[241] Widder, D. V., The Laplace Transform. Princeton University Press, 1941.

[242] Wielandt, H., "Zur Umkehrung des Abelschen Stetigkeitssatzes," Mathematische Zeitschrift 56 (1952) 206-207.

[243] Wiener, N., Une méthode nouvelle pour la démonstration des théorèmes de M. Tauber," Comptes Rendus Acad. Sci. Paris 184 (1927) 793-795.

[244] Wiener, N., "A new method in Tauberian theorems," Journal of Mathematics and Physics of the Massachusetts Institute of Technology 7 (1928) 161-184.

[245] Wiener, N., "A type of Tauberian theorem applying to Fourier series," Proceedings of the London Mathematical Society 30 (1929) 1-8.

[246] Wiener, N., "Tauberian theorems," Annals of Mathematics 33 (1932) 1-100. (Corrections p. 787.) [Reprinted in Selected Papers of Norbert Wiener. The M.I.T. Press, Cambridge, Massachusetts, 1964, pp. 261-360, and in Norbert Wiener: Generalized Harmonic Analysis and Tauberian Theorems. The M.I.T. Press, Cambridge, Massachusetts, 1966, pp. 143-242.]

[247] Wiener, N., "A one-sided Tauberian theorem," Mathematische Zeitschrift 36 (1933) 787-789.

[248] Wiener, N., The Fourier Integral and Certain of its Applications. Cambridge University Press, New York, 1933. [Reprinted by Dover, New York, 1959.]

[249] Wiener, N., and H. R. Pitt, "A generalization of Ikehara's theorem," Journal of Mathematics and Physics M.I.T. 17 (1938) 247-258.

Rouché's Theorem and Lagrange's Expansion

[250] Bateman, H., "An extension of Lagrange's expansion," Transactions of the American Mathematical Society 28 (1926) 346-356.

[251] Boas, R. P., Entire Functions. Academic Press, New York, 1954.

[252] Borchardt, C. W., "Leçons sur les fonctions doublement périodiques faites en 1847 par M. J. Liouville," Journal für die reine und angewandte Mathematik 88 (1880) 277-310.

[253] Bugajew, N. W., "Verallgemeinerte Form der Lagrangeschen Reihe," (Russian) Matematicheskii Sbornik 22 (1901) 219-224.

[254] Cauchy, A., "Mémoire sur les fonctions complémentaires," Comptes Rendus Acad. Sci. Paris 19 (1844) 1377-1378.

[255] Cauchy, A., "Rapport sur un mémoire présenté à l'Académie par M. Hermite, et relatif aux fonctions à double période." (Remarques de M. Liouville. Note de M. Augustin Cauchy relative aux observations présentées à l'Académie par M. Liouville.) Comptes Rendus Acad. Sci. Paris 32 (1851) 442-454.

[256] Dixon, A. C., "On Bürmann's theorem," Proceedings of the London Mathematical Society 34 (1901-1902) 151-153.

[257] Hadamard, J., "Sur les fonctions entières de la forme $e^{G(x)}$," Comptes Rendus Acad. Sci. Paris 114 (1892) 1053-1055.

[258] Hadamard, J., "Étude sur les propriétés des fonctions entières et en particulier d'une fonction considerée par Riemann," Journal de Mathématiques Pures et Appliquées (4) 9 (1893) 171-215. [Reprinted in Oeuvres de Jacques Hadamard. Tome I. Centre National de la Recherche Scientifique. Paris, 1968, pp. 103-147.]

[259] Johnson, A. R., "Two general theorems, of which Lagrange's and Laplace's theorems form particular cases," The Messenger of Mathematics 14 (1885) 76-87.

[260] Kössler, M., "On a generalization of Lagrange's series," Proceedings of the London Mathematical Society (2) 20 (1922) 365-373.

[261] Lagrange, J. L., "Nouvelle méthode pour résoudre les équations littérales par le moyen des séries," Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Berlin, année 1768, Tom. 24 (1770) 5-73. [Oeuvres de Lagrange. Tom. III. Gauthier-Villars, Paris, 1869, pp. 5-73.]

[262] Lagrange, J. L., and A.-M. Legendre, "Rapport sur deux mémoires d'analyse du Professeur Burmann," Mémoires de la Classe des Sciences Mathématiques et Physiques de l'Institut 2 (1799) 13-17.

[263] Lerch, M., "Über eine neue Verallgemeinerung der Taylorschen und der Lagrangeschen Reihe," (Czech) Rozpravy české Akademie císaře Františka Josefa pro vědy, slovesnost a umění, II. Cl. 20 No. 36 (1911) 14 pp.

[264] Meech, L. W., "New demonstration and forms of Lagrange's theorem. The general theorem." The Analyst. A Nonthly Journal of Pure and Applied Mathematics (Des Moines, Iowa) 3 (1876) 33-42.

[265] McClintock, E., "On certain expansion theorems," American Journal of Mathematics 4 (1881) 16-24.

[266] Pellet, A.-E., "Sur un mode de séparation des racines des équations et la formule de Lagrange," Bulletin des Sciences Mathématiques et Astronomiques 5 (1881) 393-395.

[267] Rouché, E., "Mémoire sur la série de Lagrange," Comptes Rendus Acad, Sci. Paris 52 (1861) 295-296. [Journal de l'École Polytechnique 39e cah. (1862) 193-224.]

[268] Rouché, E., "Mémoire sur la série de Lagrange," Mémoires présentés par divers savants à l'Académie des Sciences de l'Institut Impérial de France. Sci. Math. et Phys. (Paris) 18 (1868) 457-487.

[269] Saks, S., and A. Zygmund, Analytic Functions. Warszawa, 1952.

[270] Teixeira, F. G., "Sur le développement des fonctions implicites en une série," Journal de Mathématiques Pures et Appliquées (3) 7 (1881) 277-282.

[271] Teixeira, F. G., "Sur les séries ordonnées suivant les puissances d'une fonction donnée," Journal für die reine und angewandte Mathematik 122 (1900) 97-123.

[272] Titchmarsch, E. C., The Theory of Functions. Second edition. Oxford University Press, 1939. (First edition 1932.)

[273] Whittaker, E. T., and G. N. Watson, A Course of Modern Analysis. Fourth edition. Cambridge University Press, 1927. (First edition 1902.)

[274] Zolotareff, G., "Sur la série de Lagrange," Nouvelles Annales de Mathématiques (2) 15 (1876) 422-423.

Additional References

[275] Cauchy, A., "Extrait du Mémoire sur quelques séries analogues à la série de Lagrange," Mémoires de l'Académie des Sciences 9 (1830) 104. [Oeuvres complètes d'Augustin Cauchy. I Sér. Tome II. Gauthier-Villars, Paris, 1908, pp. 73-78.]

[276] Good, I. J., "Generalizations to several variables of Lagrange's expansion, with applications to stochastic processes," Proceeding of the Cambridge Philosophical Society 56 (1960) 367-380.