THE THEORY OF EXPERIMENT: OPERATIONAL DEFINITION OF THE PROBABILITY SPACE

By Milton Philip Olson

Lafayette, California

1. Introduction. In his monograph [2], *Grundbegriffe der Wahrscheinlichkeitsrechnung*, A. Kolmogorov gave a brief explanation of the relation between the mathematical objects of modern probability theory and concepts used in the description of experiments. While the correspondence given there is intuitively very satisfying, no instructions are given to the experimenter for the practical determination of the mathematical objects. Kolmogorov's discussion provides a path leading from mathematical theorems to physical interpretations; but the experimenter is often concerned with travel in the opposite direction, from the physical world to the mathematical abstraction.

An obvious problem is that human agents can collect only a finite amount of data; but, in a vague sense, an abstract probability space may contain an uncountably infinite amount of data. The fact that no matter how much data is on hand, more can be collected already suggests that the derivation of the probability space from physical experiments will necessarily involve limit operations. One of the central results of this paper will be the proof that under certain conditions measures derived from samples converge to the probability measure governing the random phenomenon. The Glivenko-Cantelli theorem is a result in this direction. However, the methods used here as well as the results are quite different.

Even so, one might ask why the author would bother with such an investigation, since the theory of experiment as demonstrated by the continuing existence of insurance companies, gambling houses and other, more scientific, applications seems to be reasonable and satisfactory. The first reason is that the theory given here has mathematical consequences through its contact with the theory of statistics and information. The validity of the second reason depends on one's tastes. If probability is purely abstract and, hence, a branch of measure theory, then the scheme for interpretation of the results given in Kolmogorov's monograph entirely suffices. On the other hand, if probability is a physical concept seeking mathematical expression, then it should be subject to Bridgeman's dictum: The only valid definition of a physical concept is in terms of the operations used to measure it. And this criterion of operational definition necessitates the investigation proposed.

2. Preliminary discussion. The attempt to give a direct construction of an abstract probability space from experiment has never been entirely successful. The problem studied here is more limited; its framework is similar to the setting of the Glivenko-Cantelli theorem. It is assumed that the experiment possesses a

Received 24 May 1966.

sample description space S, a σ-algebra of subsets of S denoted Σ, and a probability measure μ defined on Σ. This may impose a considerable limitation on the class of random phenomena considered. For example, the author knows no operationally defined method for determining whether a probability (in a generalized sense) is countably additive on a σ-algebra or merely finitely additive on an algebra. However, some limitations seem necessary on the class of phenomena considered. These are stated in the form of informal assumptions given below.

ASSUMPTION 1. It is assumed that the random phenomenon is described by the triple (S, Σ, μ) where μ is a fixed but unknown probability measure on Σ.

The point of departure from the Glivenko-Cantelli theorem is contained in the following informal definition, also called an assumption.

ASSUMPTION 2. A measuring instrument is a finite disjoint measurable partition $\theta = \{A_1, A_2, \cdots, A_k\}$ of S such that on any trial of the experiment, it can be announced with certainty in which element of the partition the outcome lies. A measuring instrument θ_1 is "finer" than another θ_2 if each $A_j \in \theta_2$ is of the form $A_j = \bigcup_{i=1}^{l_j} B_i$, where $B_i \in \theta_1$ for $i = 1, 2, \cdots, l_j$. This relationship is denoted $\theta_1 \geqq \theta_2$.

A little reflection will convince the reader that this definition is a fairly accurate description of the way in which experimental outcomes are measured. In the Glivenko-Cantelli theorem, it is assumed that the experimental outcomes can be measured exactly.

ASSUMPTION 3. The experimenter can make as many "independent" trials of the experiment as he wishes. Independence here is in the vague physical sense that the outcome on a preceding trial in no way determines the outcome on a following trial. It will be convenient to denote the outcomes of an infinite sequence of such "independent trials" by $s^* = (s_1, s_2, \cdots)$. This is in spite of the fact that the actual outcomes s_j cannot necessarily be known precisely, but only through a partition. (S^*, Σ^*, μ^*) is equal to $(\times_1^\infty S, \times_1^\infty \Sigma, \times_1^\infty \mu)$, which is the probability space of infinite sequences of "independent trials" s^*.

(Assumption 3 serves the mathematical purpose of putting a fixed but unknown probability measure on the space of a denumerable number of trials.)

A procedure for experimental determination of the measure μ can now be sketched. The experimenter chooses a measuring instrument $\theta_1 = \{A_1^1, A_2^1, \cdots, A_{n_1}^1\}$ and makes m_1 trials of the experiment. This determines a sample measure μ_1 such that

$$\mu_1(A_k^1) = \text{number of times } A_k^1 \text{ occurred in } m_1 \text{ trials}/m_1. \tag{2.1}$$

Then a finer measuring instrument $\theta_2 = (A_1^2, A_2^2, \cdots, A_{n_2}^2)$ is chosen and m_2 further trials are made determining a sample measure μ_2 such that

$$\mu_2(A_k^2) = \text{number of times } A_k^2 \text{ occurred in } m_2 \text{ trials}/m_2. \tag{2.2}$$

This procedure is repeated to obtain a sequence (θ_j, m_j, μ_j) of partitions and sample measures. It remains to investigate the conditions under which (θ_j, m_j, μ_j) converges to (Σ, μ).

3. The convergence theorem. Before discussing the convergence problem, it is necessary to create an abstract setting in which the discussion of convergence can take place. Suppose there is given a family of partitions $\{\theta_j\}_{j=1}^{\infty}$ such that $\theta_{j+1} \geqq \theta_j$ and a sequence of positive integers m_j. With each $s^* \varepsilon S^*$, in the manner described above, there is associated a sequence of partitions and sample measures $\{\theta_j, m_j, \mu_j(s^*))\}_{j=1}^{\infty}$. Except for a set $E^* \varepsilon \Sigma^*$ of μ^* measure zero, $\mu_j(s^*)$ is absolutely continuous with respect to μ restricted to θ_j. This means that if $\mu(A_k^j) = 0$, then sample sequences for which A_k^j occurs between the $(m_{j-1} + 1)$st trial and the m_jth trial form a set of μ^* measure zero. Thus, on $S^* - E^*$,

$$\mu_j(s^*, A_k^j) = \int_{A_k^j} X_j(s^*, s)\mu(ds), \qquad A_k^j \varepsilon \theta_j, \tag{3.1}$$

where

$$X_j(s^*, s) = \left(\sum_{i=1}^{n_j} [\mu_j(s^*, A_i^j)/\mu(A_i^j)]\chi_{A_i^j}\right)(s), \tag{3.2}$$

with the convention that the coefficient of $\chi_{A_i^j}$ is zero if $\mu(A_i^j) = 0$. As long as s^* is restricted to $S^* - E^*$, $\mu_j(s^*, \cdot)$ admits the integral representation shown in (3.1). Thus, $X_j : (S^* - E^*) \to L_1(S, \Sigma, \mu)$ and X_j is a μ^* a.e. defined function taking values in the Banach space $L_1(S, \Sigma, \mu)$. It is measurable because it takes only a finite number of different values. Noting that $\|X_j\|_{L_1(S,\Sigma,\mu)}(s^*) = 1$,

$$\int \|X_j\|_{L_1(S,\Sigma,\mu)}(s^*)\mu^*(ds^*) = 1, \qquad j = 1, 2, \cdots. \tag{3.3}$$

In this way, we have that X_j is a μ^* a.e. defined representative of an equivalence class in $L_1(S^*, \Sigma^*, \mu^*, L_1(S, \Sigma, \mu))$, the Banach space of all μ^*-integrable functions taking values in $L_1(S, \Sigma, \mu)$. It is in this latter space that the discussion of convergence will take place.

LEMMA 3.1. *Let $\{\theta_j, m_j\}_{j=1}^{\infty}$ be as described above and X be the function on S^* whose value everywhere is the function $1 \varepsilon L_1(S, \Sigma, \mu)$. Then,*

$$\|X - X_j\|_{L_1(\mu^*, L_1(\mu))} \leqq [(n_j - 1)/m_j]^{\frac{1}{2}}, \qquad j = 1, 2, \cdots, \tag{3.4}$$

where n_j is the number of elements in the partition θ_j.

Proof.

$$\begin{aligned} \|X - X_j\|_{L_1(\mu^*, L_1(\mu))} &\leqq \int_{S^*} \|X - X_j\|_{L_1(\mu)}(s^*)\mu^*(ds^*) \\ &= \int_{S^*} \int_S |X(s) - X_j(s)|(s^*)\mu(ds)\mu(ds^*) \\ &= \int_{S^*} \sum_{j=1}^{n_j} |\mu(A_i^j) - \mu_j(s^*, A_i^j)|\mu^*(ds^*) \\ &= \sum_{i=1}^{n_j} \int_{S^*} |\mu(A_i^j) - \mu_j(s^*, A_i^j)|\mu^*(ds^*). \end{aligned} \tag{3.5}$$

$|\mu(A_i^j) - \mu_j(s^*, A_i^j)|$ takes on only a finite number of values $|\mu(A_i^j)|$, $|\mu(A_i^j) - m_j^{-1}|$, $\cdots$, $|\mu(A_i^j) - 1|$. Because of the "independence" of the trials, the μ^* probability that $\mu_j(s^*, A_i^j) = k/m_j$ is none other than $\binom{m_j}{k}\mu(A_i^j)^k\mu(S - A_i^j)^{m_j-k}$: i.e., k successes in m Bernoulli trials with probability of success equal to $\mu(A_i^j)$. Hence, we have

$$\begin{aligned} &\|X - X_j\|_{L_1(\mu^*, L_1(\mu))} \\ &\quad \leqq \sum_{i=1}^{n_j} \sum_{k=0}^{m_j} |\mu(A_i^j) - k/m_j|\binom{m_j}{k}\mu(A_i^j)^k\mu(S - A_i^j)^{m_j-k}. \end{aligned} \tag{3.6}$$

But $\mu(A_i^j)$ is just the expected value of k/m_j with respect to the binomial distribution shown. The inner sum is the L_1-norm of a random variable defined on 0, 1, 2, $\cdots$, m_j with respect to a probability measure. It is dominated by L_2-norm; and, hence,

$$(3.7)\quad \|X - X_j\|_{L_1(\mu^*, L_1(\mu))} \leqq \sum_{i=1}^{n_j} \{ \sum_{k=0}^{m_j} (\mu(A_i^j) - k/m_j)^2 \binom{m_j}{k} \mu(A_i^j) \mu(S - A_i^j)^{m_j-k} \}^{\frac{1}{2}}.$$

The right side of (3.7) is the sum of the standard deviations of the binomial distributions shown divided by m_j, and we have

$$(3.8)\quad \|X - X_j\|_{L_1(\mu^*, L_1(\mu))} \leqq m_j^{-1} \sum_{i=1}^{n_j} \{m_j(\mu(A_i^j)(1 - \mu(A_i^j)\}^{\frac{1}{2}} \leqq \sum_{i=1}^{n_j} \{\mu(A_i^j)\mu(S - A_i^j)/m_j\}^{\frac{1}{2}} \leqq \{(n_j - 1)/m_j\}^{\frac{1}{2}}.$$

(This same result could be calculated directly from the multinomial distribution.) The last inequality follows from a simple maximization argument. Q.E.D.

The lemma shows that $X_l \rightarrow X$ in $L_1(S^*, \Sigma^*, \mu^*, L_1(S, \Sigma, \mu))$ if $\{(n_j - 1)/m_j\}^{\frac{1}{2}} \rightarrow 0$ as $j \rightarrow \infty$. Let $C \varepsilon \Theta_l$, the σ-algebra generated by the partition θ_l and completed with respect to μ-null sets. Then given $\epsilon > 0$,

$$(3.9)\quad \lim_{l\to\infty} \mu^*\{s^* \mid |\mu(C) - \mu_l(s^*; C)| \geqq \epsilon\} = \lim_{l\to\infty} \mu^*\{s^* \mid \int_C |X(s^*; s) - X_l(s^*; s)|\mu(ds) \geqq \epsilon\}.$$

In (3.9), we have tacitly extended X_l from $S^* - E^*$ to all of S^*. The nature of this extension is unimportant in formulas like (3.9) because E^* is a set of μ^* measure zero. Therefore

$$(3.10)\quad \begin{aligned} &\lim_{l\to\infty} \mu^*\{s^* \mid |\mu(C) - \mu_l(s^*; C)| \geqq \epsilon\} \\ &= \lim_{l\to\infty} \mu^*\{s^* \mid \int_C |X(s^*; s) - X_l(s^*; s)|\mu(ds) \geqq \epsilon\} \\ &\leqq \lim_{l\to\infty} \mu^*\{s^* \mid \int_S |X(s^*; s) - X_l(s^*; s)|\mu(ds) \geqq \epsilon\} \\ &\leqq \lim_{l\to\infty} \|\epsilon^{-1}\|X - X_l\|_{L_1(\mu)}\|_{L_1(S^*, \Sigma^*, \mu^*, R)} = 0. \end{aligned}$$

For, $\|X - X_l\|_{L_1(\mu)}(s^*)$ is just a real-valued integrable function on (S^*, Σ^*, μ^*); and, in the course of the proof of the lemma, we have shown that

$$(3.11)\quad \int_{S^*} \|X - X_l\|_{L_1(\mu)} (s^*)\mu^*(ds^*) \rightarrow 0 \quad \text{as} \quad l \rightarrow \infty.$$

Observe that no matter how close $\mu_l(s^*, C)$ is to $\mu(C)$, $\mu_{l+1}(s^*, C)$ may have any of the values 0, $1/m_{l+1}$, $\cdots$, $(m_{l+1} - 1)/m_{l+1}$, 1 for sufficiently improbable sample sequences.

DEFINITION 3.1. A sequence $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ of partitions, integers, and sample measures is said to converge $L_1(\mu^*, L_1(\mu))$ to (Σ_1, μ), where Σ_1 is the smallest σ subalgebra containing $\bigcup_{j=1}^{\infty} \theta_j$ completed with respect to μ null sets in Σ, if $X_l \rightarrow X$ in $L_1(\mu^*, L_1(\mu))$.

In the following lemma, if Λ is the smallest σ-algebra completed with respect to μ null sets in Σ over a family of sets in Σ, then Λ' denotes the smallest (uncompleted) σ algebra over the same family.

Lemma 3.2. *Let* $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ *converge* $L_1(\mu^*, L_1(\mu))$ *to* (Σ_1, μ) *and* $\theta_j \leqq \theta_{j+1}$, $j = 1, 2, \cdots$. *Then* $C \varepsilon \Sigma_1$ *if and only if there exists a sequence of sets* $C_l \varepsilon \Theta_l'$, $l = 1, 2, \cdots$, *such that for each* $\epsilon > 0$,

$$\lim_{l\to\infty} \mu^*\{s^* \mid \|\chi_C(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)} \geqq \epsilon\} = 0. \tag{3.12}$$

($\chi_c(s^*, s) = \chi_c(s)X(s^*, s)$ is another notation for the function on S^* whose value at each point is the set indicator function $\chi_c(s)$ as an element of $L_1(\mu)$.)

Proof. Let $C \varepsilon \Sigma_1$. By definition, Σ_1 is the smallest σ algebra over the σ algebras $\{\Theta_j\}_{j=1}^{\infty}$ and the completion of the σ algebra Σ_1', the smallest σ algebra over $\bigcup_{j=1}^{\infty} \theta_j$. Σ_1' is also the smallest σ algebra over the algebra Σ_0, the collection of all finite unions of the sets $A_i^j \varepsilon \theta_j$; $i = 1, \cdots, n_j$; $j = 1, 2, \cdots$. The set indicator functions for elements of Σ_0 are then $L_1(\mu)$-dense in the set indicator functions for elements of Σ_1. The assertion of the lemma is trivial unless $C \not\varepsilon \Sigma_0$. Hence, if $C \varepsilon \Sigma_1$ and $C \not\varepsilon \Sigma_0$, then there exists a sequence $\{D_k\}_{k=1}^{\infty} \subset \Sigma_0$ such that given $\epsilon > 0$, there exists n such that for all $k \geqq n$, $\|\chi_C - \chi_{D_k}\|_{L_1(\mu)} < \epsilon$. Using the fact that $\Theta_1' \subset \Theta_2' \subset \cdots$, a smallest positive integer t_k exists such that $D_k \varepsilon \Theta_{t_k}'$. Define $C_l = S \varepsilon \Theta_l'$ for $l < t_1$. Let k_2 be the smallest k such that $t_k > t_1$. Define $C_l = D_{t_1}$ for $t_1 \leqq l < t_{k_2}$. Continue in this fashion. The resulting sequence $\{C_l\}_{l=1}^{\infty}$ is such that $C_l \varepsilon \Theta_l'$ and $\{\chi_{C_l}\}_{l=1}^{\infty}$ converges $L_1(\mu)$ to χ_C. Now $\chi_{C_l}(s^*, s) \to \chi_C(s^*, s)$ in $L_1(\mu)$ norm for each $s^* \varepsilon S^*$. (Note that this convergence is uniform in s^* because the functions are constants.) $X_l(s^*, s) \to X(s^*, s)$ in $L_1(\mu^*, L_1(\mu))$ norm. Hence

$$\begin{aligned}
&\{s^* \mid \|\chi_C(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)} < \epsilon\}\\
&\quad \supset \{s^* \mid \|\chi_C(s^*, s) - \chi_{C_l}(s^*, s)\|_{L_1(\mu)} < \epsilon/2\}\\
&\quad \cap \{s^* \mid \|\chi_{C_l}(s)X(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)} < \epsilon/2\}\\
&\quad \supset \{s^* \mid \|\chi_C(s^*, s) - \chi_{C_l}(s^*, s)\|_{L_1(\mu)} < \epsilon/2\}\\
&\quad \cap \{s^* \mid \|X(s^*, s) - X_l(s^*, s)\|_{L_1(\mu)} < \epsilon/2\}.
\end{aligned} \tag{3.13}$$

Then, taking complements, we have

$$\begin{aligned}
&\mu^*\{s^* \mid \|\chi_C(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)} \geqq \epsilon\}\\
&\quad \leqq \mu^*\{s^* \mid \|\chi_C(s^*, s) - \chi_{C_l}(s^*, s)\|_{L_1(\mu)} \geqq \epsilon/2\}\\
&\qquad + \mu^*\{s^* \mid \|X(s^*, s) - X_l(s^*, s)\|_{L_1(\mu)} \geqq \epsilon/2\}.
\end{aligned} \tag{3.14}$$

Letting $l \to \infty$, the first term on the right hand side of (3.14) goes to zero because $\chi_{C_l} \to \chi_C$ in $L_1(\mu)$-norm and the second term goes to zero because convergence in $L_1(\mu^*, L_1(\mu))$-norm implies convergence in μ^* measure.

Conversely, let Σ_2 be the collection of all sets satisfying (3.12). We have shown above that $\Sigma_1 \subset \Sigma_2$, and the proof is completed by proving the reverse inclusion. If $C \varepsilon \Sigma_2$, then by hypothesis there exists a sequence $\{C_l\}_{l=1}^{\infty} \subset \bigcup_{j=1}^{\infty} \Theta_j' \subset \Sigma_1$ such that (3.12) is satisfied.

$$\begin{aligned}
\|\chi_C(s) - \chi_{C_l}(s)\|_{L_1(\mu)} &\leqq \|\chi_C(s)X(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)}\\
&\quad + \|\chi_{C_l}(s)X_l(s^*, s) - \chi_{C_l}(s)X(s^*, s)\|_{L_1(\mu)}.
\end{aligned} \tag{3.15}$$

Relation (3.15) holds for μ^* almost all $s^* \varepsilon S^*$. On the right hand side are two sequences of functions on S^* converging in μ^* measure to zero. Given $\epsilon > 0$, $\{s_l^*\}_{l=l_0}^{\infty}$ and $l_0(\epsilon)$ can be found such that both terms on the right side of (3.15) are less than $\epsilon/2$ at s_l^* for all $l \geqq l_0(\epsilon)$. Hence,

$$\lim_{l\to\infty} \|\chi_C(s) - \chi_{C_l}(s)\|_{L_1(\mu)} = 0. \tag{3.16}$$

This shows that $C \varepsilon \Sigma_1$. Q.E.D.

Let $\chi_C(s^*, s) = \chi_C(s)$ for all $s^* \varepsilon S^*$ and recall that

$$\|\chi_C(s^*, s) - \chi_{C_l}(s) X_l(s^*, s)\|_{L_1(\mu)} = |\mu(C) - \mu_l(s^*, C_l)|. \tag{3.17}$$

Thus, we see that Lemma 3.2 describes a sense in which the sample measures on the partition σ algebras get "close" to the unknown measure μ. However, the sense of convergence here is too weak to operationally define the measure. Rather, we seek an assertion about the convergence of the sample measures valid for all $s^* \varepsilon S^*$ except possibly a set of μ^* measure zero.

LEMMA 3.3. *Let* $\{(\theta_j, m_j)\}_{j=1}^{\infty}$, X, *and* $\{X_j\}_{j=1}^{\infty}$ *be as in Lemma* 3.1. *Then given* $\epsilon_j > 0$,

$$\mu^*\{s^* \mid \|X - X_j\|_{L_1(\mu)}(s^*) \geqq \epsilon_j\} \leqq (n_j - 1)^{\frac{1}{2}}/\epsilon_j m_j^{\frac{1}{2}}. \tag{3.18}$$

PROOF. This is an immediate consequence of a basic inequality [3], p. 158, and the fact that $(n_j - 1)^{\frac{1}{2}}/m_j^{\frac{1}{2}}$ is an upper bound of $\int_{S^*} \|X - X_j\|_{L_1(\mu)}(s^*)\mu^*(ds^*)$ as shown in (3.5). Q.E.D.

Observe that the function $\|X - X_j\|_{L_1(\mu)}(S^*)$ has only a finite number of different values as s^* varies. These values depend only on the trial $(\sum_{l=1}^{j-1} m_l) + 1$ through $\sum_{l=1}^{j} m_l$. Because of the independence of the trials, $\{\|X - X_j\|_{L_1(\mu)}(s^*)\}_{j=1}^{\infty}$ is a family of independent random variables on the probability space (S^*, Σ^*, μ^*).

Therefore, we have

$$\begin{aligned} &\mu^* \textstyle\bigcup_{k=1}^{\infty} \{s^* \mid \|X - X_{j+k}\|_{L_1(\mu)}(s^*) \geqq \epsilon\} \\ &\quad= \mu^*(S^* - \textstyle\bigcap_{k=1}^{\infty} \{s^* \mid \|X - X_{j+k}\|_{L_1(\mu)}(s^*) < \epsilon\} \\ &\quad= 1 - \textstyle\prod_{k=1}^{\infty} \mu^*\{s^* \mid \|X - X_{j+k}\|_{L_1(\mu)}(s^*) < \epsilon\} \\ &\quad= 1 - \textstyle\prod_{k=1}^{\infty} (1 - \mu^*\{s^* \mid \|X - X_{j+k}\|_{L_1(\mu)}(s^*) \geqq \epsilon\}) \\ &\quad\leqq 1 - \textstyle\prod_{k=1}^{\infty} (1 - \epsilon^{-1}(n_{j+k} - 1)^{\frac{1}{2}}/m_{j+k}^{\frac{1}{2}}). \end{aligned} \tag{3.19}$$

Then, according to Theorem 7 [1], p. 96,

$$\lim_{j\to\infty} \mu^* \textstyle\bigcup_{k=1}^{\infty} \{s^* \mid \|X - X_{j+k}\|_{L_1(\mu)}(s^*) \geqq \epsilon\} = 0 \tag{3.20}$$

if $\sum_{j=1}^{\infty} (n_j - 1)^{\frac{1}{2}}/m_j^{\frac{1}{2}} < \infty$. Thus, for example, if $m_j = n_j^{3+\delta}$, $\delta > 0$, $\sum_{j=1}^{\infty} (n_j - 1)^{\frac{1}{2}}/m_j^{\frac{1}{2}} \leqq \sum_{j=1}^{\infty} n_j^{-1-\delta/2}$, a subseries of a convergent harmonic series with coefficient $1 + \delta/2$. Then,

$$\lim_{j\to\infty} \mu^* \bigcup_{k=1}^{\infty} \{s^* \mid \|X - X_{j+k}\|_{L_1(\mu)}(s^*) \geqq \epsilon\} = 0 \quad \text{for all} \quad \epsilon > 0,$$

and the convergence a.e. criterion [3]. p. 115, is satisfied, implying that $X_j \to X$ μ^* a.e.

DEFINITION 3.2. A sequence $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ of partitions, integers, and sample measures is said to converge μ^* a.e. to (Σ_1, μ), where Σ_1 is the smallest σ subalgebra containing $\bigcup_{j=1}^{\infty} \theta_j$ completed with respect to μ null sets in Σ, if $X_l \to X$ μ^* a.e.

LEMMA 3.4. *Let* $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ *converge* μ^* a.e. *to* (Σ_1, μ) *and* $\theta_j \leqq \theta_{j+1}$, $j = 1, 2, \cdots$. *Then* $C \varepsilon \Sigma_1$ *if and only if there exists a sequence of sets* $C_l \varepsilon \Theta_l'$, $l = 1, 2, \cdots$, *such that for each* $\epsilon > 0$

$$\lim_{l\to\infty} \|\chi_C(s^*, s) - \chi_{C_l}(s) X_l(s^*, s)\|_{L_1(\mu)} = 0 \tag{3.21}$$

for μ^* a.e. $s^* \varepsilon S^*$.

PROOF. Let $\{C_l\}_{l=1}^{\infty}$ be the sequence of sets constructed in Lemma 3.2. Then,

$$\begin{aligned} &\|\chi_C(s^*, s) - \chi_{C_l}(s) X_l(s^*, s)\|_{L_1(\mu)} \\ &\quad \leqq \|\chi_C(s) X(s^*, s) - \chi_{C_l}(s) X(s^*, s)\|_{L_1(\mu)} \\ &\qquad + \|\chi_{C_l}(s) X(s^*, s) - \chi_{C_l}(s) X_l(s^*, s)\|_{L_1(\mu)} \\ &\quad \leqq \int |\chi_{C_l}(s) - \chi_{C_l}(s)| \, \mu(ds) + \int |\chi_{C_l}(s)| |X(s^*, s) - X_l(s^*, s)| \, \mu(ds) \\ &\quad \leqq \int |\chi_C(s) - \chi_{C_l}(s)| \, \mu(ds) + \int |X(s^*, s) - X_l(s^*, s)| \, \mu(ds). \end{aligned} \tag{3.22}$$

Given $\epsilon > 0$ and s^* contained in the convergence set of $\{X_l(s^*, s)\}_{l=1}^{\infty}$, there exists $l(\epsilon, s^*)$ such that $\|\chi_C - \chi_{C_l}\|_{L_1(\mu)} < \epsilon/2$ and $\|X(s^*, s) - X_l(s^*, s)\|_{L_1(\mu)} < \epsilon/2$ for $l \geqq l(\epsilon, s^*)$. Since ϵ is arbitrary, equation (3.21) holds. The converse is immediate from Lemma 3.2, since $\chi_{C_l}(s) X_l(s^*, s) \to \chi_C(s^*, s)$ μ^* a.e. implies convergence in μ^* measure Q.E.D.

We will call Σ_1, as in Lemmas 3.2 and 3.4, the operationally defined domain of definition of the measure μ. It might be argued, however, that Σ_1 is not the maximal family of limits of sequences of sets in $\bigcup_{j=1}^{\infty} \Theta_j'$ and that Σ_1 was imposed by the two rather strong definitions of convergence. But even if a weak definition of convergence for the $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ was given so that the domain of definition of the limit measure was extended to all the set indicator functions in the weak closure in $L_1(\mu)$ of the linear manifold spanned by $\{\chi_E \mid E \varepsilon \bigcup_{j=1}^{\infty} \Theta_j'\}$, no new sets would be added to the domain: the weak closure of a convex set is equal to its strong closure. We can now summarize the preceding lemmas and discussion in our major theorem.

THEOREM 3.1. *Let* (S, Σ, μ) *be a probability space, and* $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ *be a sequence of finite measurable partitions, integers, and sample measures such that* $\theta_1 \leqq \theta_2 \leqq \cdots$. *Then,*

(1) *if* $(n_j - 1)^{?}/m_j^{\frac{1}{2}} \to 0$ *as* $j \to \infty$, $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ *converges* $L_1(\mu^*, L_1(\mu))$ to (Σ_1, μ);

(2) *if* $\sum_{j=1}^{\infty} (n_j - 1)^{\frac{1}{2}}/m_j^{\frac{1}{2}} < \infty$, $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ *converges* μ^* a.e. *to* (Σ_1, μ). *In both* (1) *and* (2), Σ_1 *is the* σ *subalgebra generated by* $\bigcup_{j=1}^{\infty} \Theta_j$ *and completed with respect to* μ*-null sets.*

Further, in (1) $C \varepsilon \Sigma_1$ *if and only if there exists a sequence of sets* $C_l \varepsilon \Theta_l'$, *the smallest* σ *subalgebra over the partition* θ_l, *such that for each* $\epsilon > 0$,

$$\lim_{l\to\infty} \mu^*\{s^* \mid \|\chi_C(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)} \geqq \epsilon\} = 0.$$

In (2), $C \varepsilon \Sigma_1$ *if and only if there exists a similar sequence such that for* μ^* a.e. $s^* \varepsilon S^*$,

$$\lim_{l\to\infty} \|\chi_C(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)} = 0. \tag{3.23}$$

Moreover (3.23) *implies*

$$\lim_{l\to\infty} |\mu(C) - \mu_l(s^*, C_l)| = 0. \tag{3.24}$$

Proof. All assertions have been proved except that (3.23) implies (3.24). But,

$$\begin{aligned} |\mu(C) - \mu_l(s^*, C_l)| &= |\textstyle\int \chi_C\mu(ds) - \int_{C_l} X_l(s^*, s)\mu(ds)| \\ &\leqq \textstyle\int |\chi_C(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)|\mu(ds) \\ &\leqq \|\chi_C(s^*, s) - \chi_{C_l}(s)X_l(s^*, s)\|_{L_1(\mu)}. \end{aligned} \tag{3.25}$$

Q.E.D.

Equation (3.24) is exactly what an experimenter would want: the sample measures on some sets at his disposal converge almost everywhere to the measure on the limit set.

Corollary 3.1. *A probability measure* μ *is operationally defined on its whole domain of definition* Σ *if and only if each* $C \varepsilon \Sigma$ *differs by a* μ *null set from* $A' \varepsilon \Sigma'$, *where* Σ' *is a countably generated* σ *subalgebra whose completion with respect to* μ *null sets is* Σ.

Proof. This is an immediate consequence of preceding results.

Corollary 3.2. *Let* X_0 *be a real-valued random variable on a probability space* (S_0, Σ_0, μ_0). X_0 *induces a probability measure* μ *on* (S, Σ), *the real line with the Borel sets. Let* $\{(\theta_j, m_j, \mu_j)\}_{j=1}^{\infty}$ *converge* $L_1(\mu^*, L_1(\mu))$ *to* (Σ, μ), *and* $\{F_{X_j}(s^*, s)\}_{j=1}^{\infty}$ *be the distribution functions of the sample measures. Then, given* $\epsilon > 0$,

$$\lim_{j\to\infty} \mu^*\{s^* \mid s^* \epsilon S^*, \sup_{s\epsilon S} | F_{X_0}(s) - F_{X_j}(s^*,s)| \geq \epsilon\} = 0. \tag{3.26}$$

Proof.

$$\begin{aligned} &\sup_{s\varepsilon S} |F_{X_0}(s) - F_{X_j}(s^*, s)| \\ &\quad\leqq \sup_{E\varepsilon\Sigma} |\mu(E) - \textstyle\int_E X_j(s^*, s)\mu(ds)| \\ &\quad\leqq \sup_{E\varepsilon\Sigma} \textstyle\int |\chi_E(s)X(s^*, s) - \chi_E(s)X_j(s^*, s)|\mu(ds) \\ &\quad\leqq \|X(s^*, s) - X_j(s^*, s)\|_{L_1(\mu)}. \end{aligned} \tag{3.27}$$

Therefore,

$$(3.28)\quad \{s^* \mid \sup |F_{X_0}(s) - F_{X_j}(s^*, s)| \geqq \epsilon\} \subset \{s^* \mid \|X(s^*, s) - X_j(s^*, s)\|_{L_1(\mu)} \geqq \epsilon\}.$$

Lemma 3.2 implies the desired conclusion. Q.E.D.

REFERENCES

[1] KNOPP, K. (1956). *Infinite Sequences and Series*. New York.
[2] KOLMOGOROV, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Ergebnisse der Mathematik und Ihrer Grenzgebiete, Berlin.
[3] LOÈVE, M. (1933). *Probability Theory* (3rd edition). Princeton Univ. Press.