The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 970–978

CONDITIONAL PROBABILITY ON Σ-COMPLETE BOOLEAN ALGEBRAS[1]

BY ARDEL J. BOES

Purdue University and Colorado School of Mines

1. Summary. Probability as measure on a Boolean algebra was presented by Kappos [5], but a treatment of conditional probability relative to a subalgebra is missing. The Stone space of a σ-complete Boolean algebra (see [10], p. 24) enables one to apply the concepts of conditional probability for a σ-algebra of subsets of some space (see [2], pp. 15–28), but the problem deserves closer attention.

In this note we consider conditional probability with respect to a σ-subfield of the σ-field generated by the open-closed subsets of the Stone space of a Boolean σ-algebra. We show that there is always a regular conditional probability (see [4], p. 80) relative to a full σ-subalgebra of Baire sets. With a modified definition of probability on a Boolean algebra a treatment of conditional probability is possible without reference to the Stone space. For this a generalized integral is defined and the theory of integration is begun for it. A definition of conditional probability on a σ-complete Boolean algebra is given for which there is no regularity condition. We conclude the discussion with a study of the relationship of this theory with the conventional theory.

2. Introduction. Throughout this paper we will refer to a Boolean algebra (σ-complete Boolean algebra) whose elements are sets as a field (σ-field). If A is a σ-complete Boolean algebra, there is an isomorphism θ such that $\theta(A) = \mathcal{E}$ is the class of open-closed subsets of a perfect reduced field of subsets of a space Y. Y is called the Stone space of A. $\mathcal{E}$ forms a σ-complete Boolean algebra but only finite supremum and infimum correspond to set-theoretic union and intersection, respectively. Let $\mathcal{B}(\mathcal{E})$ be the σ-field generated by $\mathcal{E}$. $B \in \mathcal{B}(\mathcal{E})$ can be represented by

$$B = E \triangle I \tag{2.1}$$

where E is an open-closed set and I is a set of the first category in Y.

We will take λ to be a strictly positive probability on A. Define μ on $\mathcal{E}$ by

$$\mu(E) = \lambda(\theta^{-1}(E)), \qquad E \in \mathcal{E}.$$

In view of the isomorphism, μ is a finitely additive, strictly positive probability on $\mathcal{E}$. Define μ on $\mathcal{B}(\mathcal{E})$ by

$$\mu(B) = \mu(E), \qquad B \in \mathcal{B}(\mathcal{E}),$$

where E is given by (2.1). Then μ is a probability measure (σ-additive) on $\mathcal{B}(\mathcal{E})$.

Received 22 July 1968; revised 30 December 1968.
[1] This work was supported by the Air Force Office of Scientific Research under contract AFOSR 955–65, Glen Baxter, Project Director.

For details of some of the preceding remarks, the reader may refer to [5] and [10].

In order to study conditional probability on a Boolean algebra we assume the Boolean algebra of events is σ-complete even though some of theory can be carried out in greater generality (see [11]).

Let B be a σ-complete Boolean algebra and let R denote the set of real numbers. Bold faced letters such as **f**, **g** are used consistently in this paper to denote functions with values in B and domain R. **f** is a random variable if the following conditions are satisfied:

$$\text{(2.2)} \qquad \begin{aligned} &\text{(i)} \quad \mathbf{f}(\alpha)\downarrow \quad \text{as} \quad \alpha\uparrow, \\ &\text{(ii)} \quad \vee_\alpha \mathbf{f}(\alpha) = 1 \quad \text{and} \quad \wedge_\alpha \mathbf{f}(\alpha) = 0, \\ &\text{(iii)} \quad \vee_{\beta>\alpha} \mathbf{f}(\beta) = \mathbf{f}(\alpha) \quad \text{for every} \quad \alpha \text{ in } R. \end{aligned}$$

Notice that if B is a σ-field of subsets of a space X, a correspondence between random variables **f** and measurable real valued functions f is given by $\mathbf{f}(\alpha) = \{x : f(x) > \alpha\}$. This motivates the definitions of order, arithmetic operations, etc. for the class of random variables on B (see [8]).

3. Conditional probability on the Stone space. In this section the basic probability space is the space referred to in (2), namely $(Y, \mathcal{B}(\mathcal{E}), \mu)$. The sets of $\mathcal{B}(\mathcal{E})$ are called Baire sets and $\mathcal{B}(\mathcal{E})$-measurable functions on Y are called Baire functions. Let $\mathcal{A}$ be a σ-subfield of $\mathcal{B}(\mathcal{E})$ which is full, i.e. for each A in $\mathcal{A}$ the open-closed set of $\mathcal{E}$ which is μ-equivalent to A is also in $\mathcal{A}$. Assume that $(\mu \mid \mathcal{A})$ is complete, where $(\mu \mid \mathcal{A})$ denotes the restriction of μ to $\mathcal{A}$.

THEOREM 3.1. *There is a regular conditional probability* $c(\cdot, \cdot \; \mathcal{B}(\mathcal{E}), \mathcal{A}) = c(\cdot, \cdot)$ *on* $\mathcal{B}(\mathcal{E}) \times Y$.

PROOF. Choose $p(E, \cdot \mid \mathcal{B}(\mathcal{E}), \mathcal{A}) = p(E, \cdot)$ so that $p(E, \cdot)$ is an $\mathcal{A}$-measurable function for which $0 \leqq p(E, \cdot) \leqq 1$ and for every A in $\mathcal{A}$,

$$\text{(3.1)} \qquad \mu(E \cap A) = \int_A p(E, \cdot)\, d(\mu \mid \mathcal{A}).$$

For each n, consider

$$\sum (k-1)/2^n \chi_{A_{n,k}}(y) \qquad (k = 1, \cdots, 2^n + 1),$$

where

$$A_{n,k} = \{y : (k-1)/2^n \leqq p(E, y) < k/2^n\}.$$

There is $E_{n,k}$ in $\mathcal{E}(2.1)$ such that $\mu(A_{n,k} \triangle E_{n,k}) = 0$. If we define

$$c_n(E, y) = \sum (k-1)/2^n \chi_{E_{n,k}}(y) \qquad (k = 1, \cdots, 2^n + 1),$$

then $c_n(E, \cdot)$ converges uniformly to a function $c(E, \cdot)$. $c(E, \cdot)$ is a continuous version of the functions defined by (3.1) because each $c_n(E, \cdot)$ is continuous and $\mathcal{A}$ is full.

There is N in $\mathcal{A}$, $\mu(N) = 0$, for which $c(E_1, y) + c(E_2, y) = c(E_1 \cup E_2, y)$ for y in N and E_1, E_2 disjoint. But these functions are continuous on Y, so $c(\cdot, y)$

is a finitely additive probability on $\mathcal{E}$ for fixed y. We can extend $c(\cdot, y)$ to $\mathcal{B}(\mathcal{E})$ because $\mathcal{E}$ is perfect. There will be no confusion if we denote this extension by $c(\cdot, y)$.

Let $\mathcal{B}$ denote the class of all B in $\mathcal{B}(\mathcal{E})$ for which $c(B, \cdot)$ satisfies (3.1). Then $\mathcal{B} = \mathcal{B}(\mathcal{E})$ and the proof is completed.

For any integrable function f on Y, let $E(f, \cdot)$ be an $\mathcal{A}$-measurable function for which $\int_A f\, d\mu = \int_A E(f, \cdot)\, d(\mu \mid \mathcal{A})$ for all A in $\mathcal{A}$. If $f = \chi_B$, we may take a version $E(\chi_B, \cdot) = c(B, \cdot)$. Then $E(\chi_B, \cdot) = \int_B c(dy, \cdot) = \int_Y \chi_B c(dy, \cdot)$. Proceeding in a standard manner, we obtain the following theorem.

THEOREM 3.2. *If f is integrable on Y, then there is a version of $E(f, \cdot)$ for which* $E(f, \cdot) = \int_Y fc(dy, \cdot)$. (cf. [2], p. 27).

4. Integration. The notion of probability on a Boolean algebra is modified here in order to investigate conditional probability without reference to the Stone space.

Let B be a σ-complete Boolean algebra and let $\Omega[B]$ be the set of random variables (2.2) with values in B. A mapping $\mathbf{\Phi}$ with domain B and range a subset of

$$\text{(4.1)}\quad \begin{aligned} &\text{(i)} && \mathbf{0} \leqq \mathbf{\Phi}(a) \leqq \mathbf{1} \quad \text{for every } a \text{ in } B,\\ &\text{(ii)} && \mathbf{\Phi}(a) = \mathbf{0} \quad \text{if and only if } a = 0,\\ &\text{(iii)} && \mathbf{\Phi}(1) = \mathbf{1}\\ &\text{(iv)} && \mathbf{\Phi}(a \vee b) = \mathbf{\Phi}(a) + \mathbf{\Phi}(b), \quad \text{if } a \wedge b = 0, \quad \text{and}\\ &\text{(v)} && \mathbf{\Phi}(a_n) \downarrow \mathbf{0} \quad \text{if } a_n \downarrow 0. \end{aligned}$$

In order to define the integral of a random variable $\mathbf{f}$ with respect to $\mathbf{\Phi}$, we require a little notation. A characteristic function $\mathbf{a}$ for an element a of B is given by

$$\begin{aligned} \mathbf{a}(\alpha) &= 1, && \alpha < 0\\ &= a, && 0 \leqq \alpha < 1\\ &= 0, && 1 \leqq \alpha. \end{aligned}$$

If α is a real number, let $\boldsymbol{\alpha}$ denote the random variable given by

$$\begin{aligned} \boldsymbol{\alpha}(\beta) &= 1, && \beta < \alpha\\ &= 0, && \alpha \leqq \beta. \end{aligned}$$

Generally, $\boldsymbol{\varepsilon}$ and $\mathbf{N}$ denote random variables whose "jump" occurs at a small positive and a large real number, respectively.

If $\mathbf{s} = \alpha_1\mathbf{a}_1 + \cdots + \alpha_n\mathbf{a}_n$ is a simple function, define

$$\text{(4.2)}\qquad \int \mathbf{s}\, d\mathbf{\Phi} = \alpha_1\mathbf{\Phi}(a_1) + \cdots + \alpha_n\mathbf{\Phi}(a_n).$$

We see that the integral of a simple function is a linear combination of random variables.

By $\mathbf{f}_n \to \mathbf{f}$ we mean $\wedge_n(\vee_{m \geqq n}\mathbf{f}_m) = \vee_n(\wedge_{m \geqq n}\mathbf{f}_m) = \mathbf{f}$, where $(\wedge_n\mathbf{f}_n)(\alpha) = \vee_m \wedge_n\mathbf{f}_n(\alpha + 1/m)$ and $(\vee_n\mathbf{f}_n)(\alpha) = \vee_n\mathbf{f}_n(\alpha)$. If $\mathbf{f}_n \leqq \mathbf{g}$, $\vee_n\mathbf{f}_n$ is a random variable, and if $\mathbf{f}_n \geqq \mathbf{g}$, $\wedge_n\mathbf{f}_n$ is a random variable.

If $\mathbf{f} \geqq \mathbf{0}$, $\mathbf{s}_n$ is a sequence of simple functions such that $\mathbf{0} \leqq \mathbf{s}_n \leqq \mathbf{f}$, and $\mathbf{s}_n \uparrow \mathbf{f}$, define

$$\int \mathbf{f}\, d\mathbf{\Phi} = \vee_n \int \mathbf{s}_n\, d\mathbf{\Phi}. \tag{4.3}$$

A random variable $\mathbf{f}$ is bounded above (below) if there is α such that $\mathbf{f}(\alpha) = 0$ ($\mathbf{f}(\alpha) = 1$). If we extend this definition to apply to integrals of non-negative $\mathbf{f}$, we see that if the integral in (4.3) is bounded, it is a random variable.

Remark. If $a \leqq b$, $b = a \vee (b - a)$. Then $\mathbf{\Phi}(a) \leqq \mathbf{\Phi}(b)$ by (4.1).

Theorem 4.1. *If* $\mathbf{s}$ *and* $\mathbf{t}$ *are simple functions with* $\mathbf{0} \leqq \mathbf{s} \leqq \mathbf{t}$, *then* $\int \mathbf{s}\, d\mathbf{\Phi} \leqq \int \mathbf{t}\, d\mathbf{\Phi}$.

Proof. If $\mathbf{s} = \mathbf{a}$ and $\mathbf{t} = \mathbf{b}$ for some a, b in B, $\int \mathbf{s}\, d\mathbf{\Phi} = \mathbf{\Phi}(a)$ and $\int \mathbf{t}\, d\mathbf{\Phi} = \mathbf{\Phi}(b)$ by (4.2). If $a \leqq b$, we see that $\int \mathbf{a}\, d\mathbf{\Phi} \leqq \int \mathbf{b}\, d\mathbf{\Phi}$ by the remark. In order to complete the proof we point out that if $\mathbf{s}$ and $\mathbf{t}$ are simple functions, there is a common decomposition.

In definition (4.3) we use $\mathbf{s}_n \uparrow \mathbf{f}$ so in view of Theorem 4.1 we write $\lim_n \int \mathbf{s}_n\, d\mathbf{\Phi} = \vee_n \int \mathbf{s}_n\, d\mathbf{\Phi}$.

Using a common decomposition and the distributive laws (see [8]), the following theorem is apparent.

Theorem 4.3. *If* $\mathbf{0} \leqq \mathbf{s}$ *and* $\mathbf{0} \leqq \mathbf{t}$ *are simple functions and* α *is a positive real number, then* $\int \alpha\mathbf{s}\, d\mathbf{\Phi} = \alpha \int \mathbf{s}\, d\mathbf{\Phi}$ *and* $\int (\mathbf{s} + \mathbf{t})\, d\mathbf{\Phi} = \int \mathbf{s}\, d\mathbf{\Phi} + \int \mathbf{t}\, d\mathbf{\Phi}$.

Theorem 4.4. *If* $\mathbf{r}_n$ *and* $\mathbf{s}_n$ *are sequences of simple functions with* $\mathbf{0} \leqq \mathbf{r}_n \uparrow \mathbf{f}$ *and* $\mathbf{0} \leqq \mathbf{s}_n \uparrow \mathbf{f}$, *then* $\lim_n \int \mathbf{s}_n\, d\mathbf{\Phi} = \lim_n \int \mathbf{r}_n\, d\mathbf{\Phi}$.

Proof. For $\mathbf{r}_m$, $\mathbf{0} \leqq \mathbf{r}_m \leqq \mathbf{f}$ and $\lim_n \mathbf{s}_n \geqq \mathbf{r}_m$. Since $(\vee_n \wedge [\mathbf{r}_m, \mathbf{s}_n])(\alpha) = \vee_n (\mathbf{r}_m(\alpha) \wedge \mathbf{s}_n(\alpha)) = \mathbf{r}_m(\alpha)$, it follows that $\wedge[\mathbf{r}_m, \mathbf{s}_n] \uparrow \mathbf{r}_m$. Then $\mathbf{r}_m - \wedge[\mathbf{r}_m, \mathbf{s}_n] = \mathbf{t}_n \downarrow \mathbf{0}$. The $\mathbf{t}_n$ are simple functions so if

$$N = \sup\{\alpha : \mathbf{t}_1(\alpha) > 0\},$$

then $\mathbf{t}_n \leqq \mathbf{N}$ for all n. Let $\epsilon > 0$. Then $\mathbf{t}_n \leqq \boldsymbol{\varepsilon} + N\boldsymbol{\tau}_n$, where $\tau_n = \mathbf{t}_n(\epsilon)$. By the previous theorems, $\int \mathbf{t}_n\, d\mathbf{\Phi} \leqq \int \boldsymbol{\varepsilon}\, d\mathbf{\Phi} + \int N\boldsymbol{\tau}_n\, d\mathbf{\Phi} = \boldsymbol{\varepsilon} + N\mathbf{\Phi}(\mathbf{t}_n(\epsilon))$. But $\mathbf{t}_n(\epsilon) \downarrow 0$ so $\mathbf{\Phi}(\mathbf{t}_n(\epsilon)) \downarrow \mathbf{0}$. Thus $\lim_n \int (\mathbf{r}_m - \wedge[r_m, \mathbf{s}_n])\, d\mathbf{\Phi} = \mathbf{0}$ and $\int r_m\, d\mathbf{\Phi} = \lim_n \int \wedge[\mathbf{r}_m, \mathbf{s}_n]\, d\mathbf{\Phi} \leqq \lim_n \int \mathbf{s}_n\, d\mathbf{\Phi}$. Similarly, $\int \mathbf{s}_m\, d\mathbf{\Phi} \leqq \lim_n \int \mathbf{r}_n\, d\mathbf{\Phi}$.

A random variable $\mathbf{f}$ is integrable in case the integrals of $\mathbf{f}^+$ and $\mathbf{f}^-$ (see [8]) are bounded and

$$\int \mathbf{f}\, d\mathbf{\Phi} = \int \mathbf{f}^+\, d\mathbf{\Phi} - \int \mathbf{f}^-\, d\mathbf{\Phi}.$$

Linearity of the integral can be established using techniques analogous to those in standard measure theory. Theorems important to probability theory such as a monotone convergence theorem and dominated convergence theorem now can be easily obtained.

5. Conditional probability. Let B be a σ-complete Boolean algebra. Let $\Omega[0, 1]$ denote the class of random variables which assume only the 0 and 1 of B.

Thus if $\mathbf{f}$ is in $\Omega[0, 1]$ and $\mathbf{0} \leqq \mathbf{f} \leqq \mathbf{1}$, there is $0 \leqq \alpha \leqq 1$ such that $\mathbf{f}(\beta) = 1$ if $\beta < \alpha$ and $\mathbf{f}(\beta) = 0$ if $\beta \geqq \alpha$. If $\mathbf{\Phi}$ is as in (4.1) with range $\Omega[0, 1]$, $\mathbf{\Phi}$ is a probability and we denote it by $\mathfrak{u}$. If $\mathfrak{n}$ is a mapping of B into $\Omega[0, 1]$, $\mathfrak{n}$ is absolutely continuous with respect to $\mathfrak{u}$ in case, given any $\boldsymbol{\varepsilon} > \mathbf{0}$, there is $\boldsymbol{\delta} > \mathbf{0}$ such that $|\mathfrak{n}(a)| < \boldsymbol{\varepsilon}$ whenever $\mathfrak{u}(a) < \boldsymbol{\delta}$, for a in B.

If $\int_a \mathbf{f}\, d\mathfrak{u} = \int \mathbf{af}\, d\mathfrak{u}$ the following Radon-Nikodyn theorem can be proved using a method similar to that of [8], pp. 186–190.

THEOREM 5.1. *If* $\mathfrak{n}$ *is absolutely continuous with respect to* $\mathfrak{u}$, *there is a unique random variable* $\mathbf{f}$ *such that* $\mathfrak{n}(a) = \int_a \mathbf{f}\, d\mathfrak{u}$. *Conversely, for a random variable* $\mathbf{f}$, *if* $\mathfrak{n}(a) = \int_a \mathbf{f}\, d\mathfrak{u}$, *then* $\mathfrak{n}$ *is absolutely continuous.*

Let A be a Boolean σ-subalgebra of B and let $(\mathfrak{u} \mid A)$ denote the restriction of $\mathfrak{u}$ to A. Then if $\mathbf{f}$ is integrable, there is a random variable $E_A(\mathbf{f})$ with values in A for which

$$\int_a E_A(\mathbf{f})\, d(\mathfrak{u} \mid A) = \int_a \mathbf{f}\, d\mathfrak{u} \tag{5.1}$$

for every a in A. $E_A(\mathbf{f})$ is the conditional expectation of $\mathbf{f}$ given A. The properties of conditional expectation are analogous to those in [2], p. 23.

If $\mathbf{f} = \mathbf{b}$ for some b in B, $E_A(\mathbf{b})$ is written $\mu_A(b)$ and is the conditional probability of b given A. Notice that (5.1) becomes

$$\int_a \mu_A(b)\, d(\mathfrak{u} \mid A) = \int_a \mathbf{b}\, d\mathfrak{u} = \mathfrak{u}(a \wedge b)$$

and let μ_A denote the mapping which takes b into $\mu_A(b)$.

Theorem 5.2. μ_A *takes values in* $\Omega[A]$ *and satisfies*:

$$\begin{array}{ll} \text{(i)} & \mathbf{0} \leqq \mu_A(b) \leqq \mathbf{1} \quad \textit{for every} \quad b \quad \textit{in} \quad B, \\ \text{(ii)} & \mu_A(b) = \mathbf{0} \quad \textit{if and only if} \quad b = 0, \\ \text{(iii)} & \mu_A(1) = \mathbf{1}, \\ \text{(iv)} & \mu_A(a \vee b) = \mu_A(a) + \mu_A(b), \quad \textit{if} \quad a \wedge b = 0, \quad \textit{and} \\ \text{(v)} & \mu_A(a_n) \downarrow \mathbf{0} \quad \textit{if} \quad a_n \downarrow 0. \end{array} \tag{5.2}$$

The reader will notice that (5.2) (cf. [2], p. 25) corresponds exactly with (4.1). In fact, conditional probability is the motivation for defining the integral as in Section 4. The idea of regularity does not apply to the conditional probabilities μ_A in view of Theorem 5.2. Although the theorem that follows is not true in general for the conventional theory, the proof can be done in the standard way by application of the previous results.

THEOREM 5.3. *If* $\mathbf{f}$ *is an integrable random variable*, $E_A(\mathbf{f}) = \int \mathbf{f}\, d\mu_A$.

The above completes this work to give a simply stated basis for the theory of probability. There is an apparent inability to correspond numerical values to events and random variables. This problem is considered in the next section.

6. Relationship with the conventional theory. If $(X, \mathfrak{A}, \lambda)$ is a probability space, then $\mathfrak{A} \mid \lambda$ is a complete Boolean algebra whose elements are denoted by

$a = [A]$ for A in $\mathcal{A}$, $1 = [X]$, and $0 = [\varnothing]$. Define $\boldsymbol{\mu}$ on $\mathcal{A} \mid \lambda$ as follows:

$$(6.1) \qquad (\boldsymbol{\mu}(a))(\alpha) = 1, \quad \alpha < \lambda(A)$$
$$= 0, \quad \lambda(A) \leqq \alpha.$$

We point out the following obvious results.

THEOREM 6.1. *If* $\boldsymbol{\mu}$ *is defined by* (6.1), $\boldsymbol{\mu}$ *is a probability* (*Section* 4) *on* $\mathcal{A} \mid \lambda$.

THEOREM 6.2. *If* $\boldsymbol{\mu}$ *is a probability* (*Section* 4) *on* $\mathcal{A} \mid \lambda$, *then there is a probability measure on the field of sets* $\mathcal{A}$ *which corresponds to* $\boldsymbol{\mu}$ *as in Theorem* 6.1.

The preceding theorems give us the numerical values for $\boldsymbol{\mu}$ sometimes necessary for computation. There is an advantage in considering probabilities with a positivity condition (see [5], pp. 25–32).

Let $\mathcal{B}$ be a σ-subfield of $\mathcal{A}$ and let $\mathcal{B} \mid \lambda$ be the set of residue classes of $\mathcal{A} \mid \lambda$ which contain an element of $\mathcal{B}$. Then $\mathcal{B} \mid \lambda \subset \mathcal{A} \mid \lambda$ and $\mathcal{B} \mid \lambda$ is a complete Boolean algebra.

A real-valued mapping $\lambda(\cdot, \cdot) = \lambda(\cdot, \cdot \mid \mathcal{A}, \mathcal{B})$ defined on $\mathcal{A} \times X$ is a regular conditional probability on $\mathcal{A} \times X$ (see [4], p. 80) if

(6.2) for any A in $\mathcal{A}$, $\lambda(A, \cdot)$ is $\mathcal{B}$-measurable, and

$$\lambda(A \cap B) = \int_B \lambda(A, \cdot)\, d\lambda \quad \text{for all } B \text{ in } \mathcal{B}, \quad \text{and}$$

(6.3) for every x in X, $\lambda(\cdot, x)$ is a probability measure on $\mathcal{A}$.

If (6.2) is satisfied, $\lambda(\cdot, \cdot)$ is a conditional probability on $\mathcal{A} \times X$. The term "conditional probability on $\mathcal{A} \times X$" is used to distinguish a mapping $\lambda(\cdot, \cdot)$ described in (6.2) from a conditional probability on a Boolean algebra described in (5.2).

A conditional probability on $\mathcal{A} \times X$ satisfies the following conditions (cf. 5.2).

(6.4)

(i) for every A in $\mathcal{A}$, $0 \leqq \lambda(A, \cdot) \leqq 1$ with λ-probability one.

(ii) $\lambda(\varnothing, \cdot) = 0$ with λ-probability one,

(iii) $\lambda(X, \cdot) = 1$ with λ-probability one,

(iv) if $A \cap B = \varnothing$, $\lambda(A \cup B, \cdot) = \lambda(A, \cdot) + \lambda(B, \cdot)$ with λ-probability one, and

(v) if $A_n \downarrow \varnothing$, $\lambda(A_n, \cdot) \downarrow 0$ with λ-probability one.

Henceforth we denote μ_F by $\mu(F, \cdot)$ if $F = \mathcal{B} \mid \lambda$. Define $\mu(\mathcal{B} \mid \lambda, \cdot)$ on $\mathcal{A} \mid \lambda$ into $\Omega[\mathcal{B} \mid \lambda]$ by $(\mu(\mathcal{B} \mid \lambda, a))(\alpha) = [\{x : \lambda(A, x) > \alpha\}]$.

THEOREM 6.3. *If* $\lambda(\cdot, \cdot)$ *is a conditional probability on* $\mathcal{A} \times X$, *then* $\mu(\mathcal{B} \mid \lambda, \cdot)$ *is a conditional probability on the Boolean algebra* $\mathcal{A} \mid \lambda$.

PROOF. Only the continuity condition will be checked. If $a_n \downarrow 0$, choose B_n in a_n so that $B_n \downarrow \varnothing$. Given $\alpha > 0$, let $E_n = \{x : \lambda(B_n, x) > \alpha\}$. Then $\lambda(B_n \cap E_n) = \int_{E_n} \lambda(B_n, \cdot)\, d(\lambda \mid \mathcal{B}) > \alpha\lambda(E_n)$.

It follows that $\lambda(B_n) > \alpha\lambda(E_n)$ so that $\lambda(E_n) < (1/\alpha)\lambda(B_n) \downarrow 0$. But $(\mu(\mathcal{B} \mid \lambda, b_n))(\alpha) = e_n$ where $e_n \downarrow 0$ so $\bigwedge_n \mu(\mathcal{B} \mid \lambda, b_n))(\alpha) = 0$. Thus

$$\bigvee_m \bigwedge_n (\mu(\mathcal{B} \mid \lambda, b_n))(1/m) = 0$$

and finally $(\bigwedge_n \mu(\mathcal{B} \mid \lambda, b_n))(0) = 0$. In view of the fact that $\bigwedge_n \mu(\mathcal{B} \mid \lambda, b_n) \geqq 0$, the proof is complete.

Regularity does not enter into Theorem 6.3. The question of the existence of a regular conditional probability on $\mathcal{A} \times X$ defined from a conditional probability on the Boolean algebra $\mathcal{A} \mid \lambda$ arises for which only a partial solution is possible (see [1] or [3], p. 210). Perhaps an approach to the related problems using the ideas of this article could be helpful. We consider two important cases.

In what follows a set is chosen from each residue class of $\mathcal{B} \mid \lambda$. Let π denote a mapping that does this and define $\lambda(\cdot, x)$ on $\mathcal{A}$ by

$$\lambda(A, x) = \sup \{\alpha : x \in \pi((\mu(\mathcal{B} \mid \lambda, a))(\alpha))\} \tag{6.5}$$

where $\mu(\mathcal{B} \mid \lambda, \cdot)$ is a conditional probability on $\mathcal{A} \mid \lambda$. When $\lambda(\cdot, x)$ is a probability measure on $\mathcal{A}$ for each x in X, we refer to this class as a class of probability measures on $\mathcal{A}$ associated with $\mu(\mathcal{B} \mid \lambda, \cdot)$.

Theorem 6.4. *Let $\mathcal{B} \mid \lambda$ be such that there is a mapping $\pi : \mathcal{B} \mid \lambda \to \mathcal{B}$ so that the class of images under π is a σ-field, then the class $\{\lambda(\cdot, x) : x \in X\}$ is a class of probability measures on $\mathcal{A}$ associated with $\mu(\mathcal{B} \mid \lambda, \cdot)$.*

Proof. Denote the set $\pi((\mu(\mathcal{B} \mid \lambda, a))(\alpha))$ by $S(A, \alpha)$, then (6.5) becomes

$$\lambda(A, x) = \sup \{\alpha : x \in S(A, \alpha)\}. \tag{6.6}$$

It is evident that $0 \leqq \lambda(A, x) \leqq 1$ for all A and x, and that $\lambda(X, x) = 1$ for all x.

If $A \cap B = \emptyset$, $S(A \cup B, \alpha) = \bigcup_\beta (S(A, \beta) \cap S(B, \alpha - \beta))$. In order to show that $\lambda(A \cup B, x) = \lambda(A, x) + \lambda(B, x)$ we must show

$$\sup\{\alpha : x \in \textstyle\bigcup_\beta (S(A, \beta) \cap S(B, \alpha - \beta))\} = \sup\{\alpha : x \in S(A, \alpha)\} + \sup\{\alpha : x \in S(B, \alpha)\}. \tag{6.7}$$

To accomplish this let α_0 be in the set of numbers on the left of (6.7), then there is β_0 so that $x \in S(A, \beta_0)$ and $x \in S(B, \alpha_0 - \beta_0)$, i.e. β_0 is in the first set on the right of (6.7) and $\alpha_0 - \beta_0$ is in the second. Thus the sum of the suprema on the right of (6.7) is larger than α_0, hence larger than the supremum on the left. On the other hand let α_1 and α_2 be in the first and second sets on the right of (6.7), respectively. Suppose that one of them, say α_1, is not the supremum. Then there is, in the dense set of real numbers, a $\beta_0 > \alpha_1$ such that $x \in S(A, \beta_0)$. Moreover $x \in S(B, \alpha_1 + \alpha_2 - \beta_0)$ since $\alpha_2 + \alpha_1 - \beta_0 < \alpha_2$. Consequently

$$x \in \textstyle\bigcup_\beta (S(A, B) \cap S(B, \alpha_1 + \alpha_2 - \beta))$$

so that $\alpha_1 + \alpha_2$ is in the set on the left of (6.7).

If $B_n \uparrow B$, $S(B, \alpha) = \bigcup_n S(B_n, \alpha)$ so if α is such that $x \in S(B_n, \alpha)$ there is n such that $x \in S(B_n, \alpha)$. Then $\lambda(B, x) \leqq \lim \lambda(B_n, \alpha)$ by (6.6). The other inequality is obvious since $\lambda(B, x) \geqq \lambda(B_n, x)$ for all n.

If the probability space is discrete, the existence of a mapping π is apparent. For more general cases we use a theorem of Maharam ([6], p. 992): If $(X, \mathcal{B}, \lambda)$ is complete, there is $\pi: \mathcal{B} \mid \lambda \rightarrow \mathcal{B}$ such that $\pi(0) = \emptyset$, $\pi(1) = X$, $\pi(a \wedge b) = \pi(a) \cap \pi(b)$, $\pi(a \vee b) = \pi(a) \cup \pi(b)$, and $\pi(b) \in b$. Henceforth we assume $(X, \mathcal{B}, \lambda)$ is complete and let π be as above. If $\mu(\mathcal{B} \mid \lambda, \cdot)$ is a conditional probability on $\mathcal{A} \mid \lambda$, we define $\lambda(\cdot, x)$ as in (6.6).

The reader will notice that this π differs from the mapping of Theorem 6.4 in that the class of images might not form a σ-field. However, by making the appropriate changes in the proof, we obtain the following results.

LEMMA 6.5. *If $A \cap B = \emptyset$, there is $N \in \mathcal{B}$ with $\lambda(N) = 0$ such that $\lambda(A \cup B, x) = \lambda(A, x) + \lambda(B, x)$ for $x \notin N$.*

COROLLARY 6.6. *If $\mathcal{A}$ is a field which consists of a countable number of sets, there is $N \in \mathcal{B}$ with $\lambda(N) = 0$ such that $\lambda(\cdot, x)$ is finitely additive on $\mathcal{A}$ for $x \notin N$.*

LEMMA 6.7. *If $A_n \uparrow A$, then there is $N \in \mathcal{B}$ with $\lambda(N) = 0$ such that $\sup_n \lambda(A_n, x) = \lambda(A, x)$ for $x \notin N$.*

In the next theorem we change the setting so that it can be readily compared to [9], p. 241 and [4], pp. 80–83. Let $(X, \mathcal{A}, \lambda)$ be as before, $\mathcal{B}_1$ a subfield of $\mathcal{A}$, $\mathcal{B}_2$ a σ-subfield of $\mathcal{A}$ for which every subset of a λ-null set in $\mathcal{B}_2$ is in $\mathcal{B}_2$. There is a conditional probability $\mu(\mathcal{B}_2 \mid \lambda, \cdot)$ on $\mathcal{B}(\mathcal{B}_1) \mid \lambda$ into $\Omega[\mathcal{B}_2 \mid \lambda]$, where $\mathcal{B}(\mathcal{B}_1)$ is the σ-field generated by $\mathcal{B}_1$ (cf. sec. 5). A class $\mathcal{C}$ of subsets of a set X is compact, if for each sequence C_n in $\mathcal{C}$ the relation $\bigcap_1^n C_i \neq \emptyset$ for $n = 1, 2, \cdots$ implies $\bigcap_1^\infty C_i \neq \emptyset$. A finitely additive probability measure λ defined on a field $\mathcal{B}$ is compact if there is a compact class $\mathcal{C}$ which approximates $\mathcal{B}$ with respect to λ, i.e. for each A in $\mathcal{B}$ and $\eta > 0$ there is a set C in $\mathcal{C}$ and a set B in $\mathcal{B}$ such that $B \subset C \subset A$ and $\lambda(A - B) < \eta$.

Theorem 6.8. *If $\mu(\mathcal{B}_2 \mid \lambda, \cdot)$ is a conditional probability on $\mathcal{B}(\mathcal{B}_1) \mid \lambda$, $\mathcal{B}_1$ is a field with a countable number of elements, and $(\lambda \mid \mathcal{B}_1)$ is compact, then there is a class of probability measures on $\mathcal{B}(\mathcal{B}_1)$ associated with $\mu(\mathcal{B}_2 \mid \lambda, \cdot)$.*

PROOF. If π is defined on $\mathcal{B}_2 \mid \lambda$, define $\lambda(\cdot, x)$ on $\mathcal{B}_1$ as in (6.6). By Corollary 6.6 $\lambda(\cdot, x)$ is finitely additive for all x outside a λ-null set. There is a compact class $\mathcal{C}$ such that for every B in $\mathcal{B}_1$ there is C_n in $\mathcal{C}$ and B_n in $\mathcal{B}_1$ such that

$$B_n \subset C_n \subset B \quad \text{and} \quad \lambda(B) = \sup_n \lambda(B_n).$$

By Lemma 6.7 $\lambda(B, x) = \sup_n \lambda(B_n, x)$ for all x outside a λ-null set. $\mathcal{B}_1$ countable implies that $\lambda(\cdot, x)$ is also compact for x not in a λ-null set. Let $N \in \mathcal{B}_2$ be a λ-null set for which $\lambda(\cdot, x)$ is a compact, finitely additive measure on $\mathcal{B}_1$ for $x \notin N$. If $B \in \mathcal{B}_1$ define

$$\begin{aligned} \mu(B, x) &= \lambda(B, x), \qquad x \notin N \\ &= \lambda(B), \qquad x \in N. \end{aligned}$$

But $\mu(\cdot, x)$ can be extended to a probability on $\mathcal{B}(\mathcal{B}_1)$ for each x (see [7], p. 118).

7. Acknowledgment. The author wishes to thank Professor L. J. Cote under whose guidance this research was conducted.

REFERENCES

[1] DIEUDONNE, J. (1948). Sur le theoreme de Lebesgue-Nikodym, III. *Ann. Univ. Grenoble.* **23** 25–53.

[2] DOOB, J. (1953). *Stochastic Processes.* Wiley, New York.

[3] HALMOS, P. (1950). *Measure Theory.* Van Nostrand, New York.

[4] JIIŘNA, M. (1954). Conditional probabilities on σ-algebras with countable basis. *Select. Transl. Math. Statist. Prob.* **2** 79–85.

[5] KAPPOS, D. (1960). Strukturtheorie der Wahrscheinlichkeits-felder und Raume. *Ergebnisse der Mathematik.* Springer, Berlin.

[6] MAHARAM, D. (1958). On a theorem of von Neumann. *Proc. Amer. Math. Soc.* **9** 987–994.

[7] MARCZEWSKI, E. (1953). On campact measures. *Fund. Math.* **40** 111–124.

[8] OLMSTED, J. (1942). Lebesgue theory on a Boolean algebra. *Trans. Amer. Math. Soc.* **51** 164–193.

[9] SAZONOV, V. (1962). On perfect measures. *Amer. Math. Soc. Transl.* **48** 229–254.

[10] SIKORSKI, R. (1964). Boolean algebras. *Ergebnisse der Mathematik.* Springer, Berlin.

[11] VARADARAJAN, V. (1962). Probability in Physics and a theorem on simultaneous observability. *Comm. Pure Appl. Math.* **15** 189–217.