

CONDITIONAL PROBABILITY DISTRIBUTIONS IN THE WIDE SENSE¹

LESTER E. DUBINS

Introduction. Doob [3, p. 29],² has shown that if $Q = [U, \mathfrak{u}, u]$ is a probability space (i.e. \mathfrak{u} is a σ -field of subsets of a set U and u is a countably additive non-negative measure defined on \mathfrak{u} , normalized by the condition $u(U) = 1$) and \mathfrak{u}_0 is a σ -subfield of \mathfrak{u} , and y is an n -dimensional random variable mapping U into an n -dimensional Euclidean space X , then y possesses a conditional distribution in the wide sense relative to \mathfrak{u}_0 . More specifically, he proved the existence of a function, p , of two variables Y and w , where Y ranges over the Baire subsets of X , and w ranges over U , which satisfies the following three conditions: (i) $p(Y, w)$ is, for each fixed w , a probability measure defined on the Baire subsets of X ; (ii) $p(Y, w)$ is, for each Y , a function measurable relative to \mathfrak{u}_0 ; and (iii) $\int_A p(Y, w) du(w) = u(A \cap y^{-1}(Y))$ for each $A \in \mathfrak{u}_0$ and each Baire subset Y of X . This result of Doob's is known to be false if one lets X be an arbitrary set with a distinguished σ -field of measurable subsets. It is our purpose (Theorem 1, below) to show that Doob's result does hold, at least if X is any locally compact, Hausdorff space whose one point compactification is metrizable.

The idea behind this generalization is the following observation. If $p(Y, w)$ exists with properties (i), (ii) and (iii) above then by (i) $p(Y, w)$ determines for each fixed w , a unique element $P(w)$ in the dual of $C(X)$ where $C(X)$ is the Banach space of all continuous real-valued functions on X which vanish at infinity. Furthermore from (ii) it follows that $(g, P(w))$ is for each g in $C(X)$ a function measurable relative to \mathfrak{u}_0 where $(g, P(w))$ is the value of the linear functional $P(w)$ at the vector g in $C(X)$. And from (iii) it follows that $\int_A (g, P(w)) du(w) = \int_A g(y(w)) du(w)$ for each $A \in \mathfrak{u}_0$ and each g in $C(X)$. This last is equivalent to $\int_A (g, P(w)) du(w) = \int_A (g, \bar{y}(w)) du(w)$ for each $A \in \mathfrak{u}_0$ and each g in $C(X)$ where for each fixed w in U , $\bar{y}(w)$ is the probability measure which lives at the one point $y(w)$. Thus P is nothing other than the conditional expectation of \bar{y} relative to \mathfrak{u}_0 . Here P and \bar{y} are Banach space valued random variables. Thus if y possesses a conditional distribution in the wide sense relative to \mathfrak{u}_0

Received by the editors March 26, 1957.

¹ This research was supported by the United States Air Force Office of Scientific Research.

² Bracketed numbers refer to the bibliography at the end of the paper.

then \bar{y} possesses a conditional expectation relative to \mathfrak{U}_0 , say P , such that $P(w)$ is a probability measure for almost all w . The Riesz-Von Neumann-Markoff theorem implies that the converse also holds. Thus the problem of the existence of a conditional distribution in the wide sense relative to \mathfrak{U}_0 for y is equivalent to the existence of a conditional expectation P for \bar{y} such that $P(w)$ is a probability measure for almost all w . Whenever one has a Radon-Nikodym type theorem one also has conditional expectations. Dieudonné [2] and Dubins [4], among others, have investigated Radon-Nikodym type theorems for random variables with values in a Banach space. It follows from a Radon-Nikodym type theorem in [4] that integrable random variables with values in the dual of a separable Banach space possess conditional expectations relative to any σ -subalgebra \mathfrak{U}_0 . Thus two further questions arise. The first is: what restriction must be imposed on X to insure that $C(X)$ is separable? The answer is both easy and well known: if X is a locally compact Hausdorff space then $C(X)$ is separable if and only if the one point compactification of X is metrizable. This latter condition is equivalent to X being a σ -compact, metrizable locally compact space. Equivalently X must be a second countable locally compact Hausdorff space. The second question is: Is it true that if $C(X)$ is separable then for almost all values of w , $P(w)$ is a probability measure (where P is the conditional expectation of \bar{y} relative to \mathfrak{U}_0)? In the event X is compact metric then an affirmative answer follows because of the Helly theorem, that the probability measures are then a compact convex set in the weak star topology, and the fact that a weak star compact convex set in the dual of a separable Banach space is the intersection of a countable number of half spaces which are open in the weak star topology. In the event that we merely assume that the one point compactification of X is metrizable, the probability measures form a convex set but not a compact set in the weak star topology. This difficulty can be overcome by showing that even in this case the probability measures are the intersection of a suitable countable collection of half spaces. The details now follow, but first we mention parenthetically that the theorem below follows easily from Doob's result alluded to above and the fact that the field of Borel subsets of a complete separable metric space is isomorphic to the field of Borel subsets of a Borel subset of the real line via an isomorphism which is induced by a point function (see [9]). We prefer to present the proof below for some of the ideas employed seem to be of independent interest.

Main result.

THEOREM. *Let $Q = [U, \mathfrak{u}, u]$ be a probability space and \mathfrak{U}_0 a σ -sub-*

field of \mathfrak{U} . Suppose X is a locally compact Hausdorff space whose one point compactification is metrizable, and let y be a measurable transformation of U into X (i.e. $y^{-1}(Y) \in \mathfrak{U}$ for each Baire subset Y of X). Then there exists a function p of Y and w for which (i), (ii) and (iii) above hold.

LEMMA 1. Let S be the dual of a real separable Banach space B , and let \bar{y} be a function mapping U into S such that, for each x in B , $(x, \bar{y}(w))$ is a \mathfrak{U} -measurable function of w , and $\int (x, \bar{y}(w)) du(w)$ exists. Then there exists a function, P , mapping U into S such that, for each x in B , $(x, P(w))$ is a \mathfrak{U}_0 -measurable function of w , and for each $A \in \mathfrak{U}_0$, $\int_A (x, P(w)) du(w)$ exists and is equal to $\int_A (x, \bar{y}(w)) du(w)$.

Any such function P is called the conditional expectation of \bar{y} relative to \mathfrak{U}_0 and is designated by $E[\bar{y} | \mathfrak{U}_0]$.

We omit the proof of Lemma 1 for it follows immediately from Corollaries 1 and 2 of Theorem 5 in [4] in much the same manner as the existence of the conditional expectation of a real valued integrable random variable follows from the usual Radon-Nikodym theorem (see [3]).

The following lemma is well-known and is stated here for ease of reference.

LEMMA 2. Let X be a locally compact Hausdorff space, and let $C(X)$ be the Banach space of all continuous real valued functions on X which vanish at infinity. Then $C(X)$ is separable if and only if the one point compactification of X is metrizable.

PROOF OF THEOREM. Let S be the dual of $C(X)$, and let \bar{y} be the mapping of U into S defined by the condition that for each w in U , $\bar{y}(w)$ is the probability measure on X which lives at the point $y(w)$. By Lemma 2, S is the dual of a real separable Banach space. It is easy to verify that the hypotheses of Lemma 1 hold. Therefore there exists a function P which maps U into S such that for each g in $C(X)$, $(g, P(w))$ is a \mathfrak{U}_0 -measurable function of w , and for each $A \in \mathfrak{U}_0$

$$\int_A (g, P(w)) du(w) = \int_A g(y(w)) du(w).$$

We will now show that for almost all w , $P(w)$ is a probability measure. Let K be the set of positive measures of total variation less than or equal to 1. Since K is a bounded weak-star closed subset of S it is weak star compact. It can be shown, (see [4]) that, since S is the dual of a separable Banach space, K is the intersection of a countable

number of weak star open half spaces G_i where G_i is the set of all measures ν such that $(g_i, \nu) < 1$ for some g_i in $C(X)$. Let A_i be the set of w such that $(g_i, P(w)) \geq 1$. Clearly $A_i \in \mathfrak{u}_0$. Suppose $u(A_i) > 0$ for some i . Then

$$(1/u(A_i)) \int_{A_i} (g_i, P(w)) du(w) \geq 1.$$

Hence $(1/u(A_i)) \int_{A_i} (g_i, \bar{y}(w)) du(w) \geq 1$. But $\bar{y}(w) \in K \subset G_i$. Hence $(g_i, \bar{y}(w)) < 1$ for all w . Therefore $(1/u(A_i)) \int_{A_i} (g_i, \bar{y}(w)) du(w) < 1$. This is a contradiction. Thus A_i is a set of measure zero for all i . Therefore for almost all w , $(g_i, P(w)) < 1$ for all i . Hence for almost all w , $P(w) \in G_2$ for all i . Thus we've shown that for almost all w , $P(w) \in K$. We will next show that $P(w)$ is a probability measure for almost all w by showing that the total variation of $P(w)$ is 1 for almost all w . Let H_n be the set of ν in S such that $\int d\nu \leq 1 - 1/n$. Let B_n be the set of w such that $P(w) \in H_n$. We first observe that $B_n \in \mathfrak{u}_0$: for there exists a uniformly bounded sequence g_k in $C(X)$ which converges to 1 everywhere. Hence $\int g_k(Z) d\nu(Z) \rightarrow \int d\nu$ for each ν . Hence the mapping $\nu \rightarrow \int d\nu$ is a linear functional h on S with the property $(g_k, \nu) \rightarrow (h, \nu)$ for each ν in S . Therefore h is measurable relative to \mathfrak{B} , where \mathfrak{B} is the σ -algebra of subsets of S determined by the linear functionals g in $C(X)$. Since H_n is in the σ -algebra determined by h , it is in the σ -algebra \mathfrak{B} . But P^{-1} of each element of \mathfrak{B} is in \mathfrak{u}_0 . Hence $B_n \in \mathfrak{u}_0$. Suppose $u(B_n) > 0$ for some n . Then, for this n ,

$$\frac{1}{u(B_n)} \int_{B_n} (h, P(w)) du(w) \leq 1 - \frac{1}{n}.$$

Hence

$$\frac{1}{u(B_n)} \int_{B_n} \left[\lim_k (g_k, P(w)) \right] du(w) \leq 1 - \frac{1}{n}.$$

By the dominated convergence theorem of Lebesgue

$$\lim_k \frac{1}{u(B_n)} \int_{B_n} (g_k, P(w)) du(w) \leq 1 - \frac{1}{n}.$$

Hence

$$\lim_k \frac{1}{u(B_n)} \int_{B_n} (g_k, \bar{y}(w)) du(w) \leq 1 - \frac{1}{n}.$$

But $(g_k, \bar{y}(w)) = g_k(y(w))$ converges to 1 everywhere and is uniformly bounded. This is easily seen to be a contradiction. Hence $u(B_n) = 0$

for all n . Thus for almost all w , the total variation of $P(w)$ is $> 1 - 1/n$ for all n . This completes the proof that for all w not in a null set N , $P(w)$ is a probability measure. The remainder of the proof is easy, namely, choose a fixed probability measure and change P on this set N so that $P(w)$ is this fixed probability measure for all w in N . Then by the Riesz-Markoff theorem (see [7, p. 247]), for each w , $P(w)$ determines a unique probability measure (in the usual abstract sense) defined on all Baire subsets Y of X . Let $P(w)$ also designate this probability measure. We define p by $p(Y, w) = P(w)(Y)$. To confirm that p satisfies conditions (i), (ii), and (iii) one may use standard measure theoretic techniques. This completes the proof.

REMARK. If X were only assumed to be locally compact Hausdorff then $C(X)$ need not be separable. In this event the conditional expectation of \bar{y} with respect to \mathfrak{U}_0 need not (as far as is known) be a point function, but would certainly be a generalized random variable (see [4]). Thus with only technical modifications in the proof of the theorem above, it can be shown that y still possesses a generalized conditional distribution in the wide sense relative to \mathfrak{U}_0 . That is, it can be shown that there exists a σ -algebra homomorphism h defined on the weak star measurable subsets of the dual of $C(X)$ into the submeasure ring \mathfrak{M}_0 of Q determined by \mathfrak{U}_0 such that $\int_A h = \int_A \bar{y}$ for each $A \in \mathfrak{M}_0$.

REFERENCES

1. N. Bourbaki, *Integration*, Paris, Herman et Cie, 1952.
2. J. Dieudonné, *On the Theorem of Lebesgue-Nikodym* (III) (IV) and (V), *Annales de l'Université de Grenoble* vol. 23 (1948) pp. 25-53, *J. Indian Math. Soc.* vol. 15 (1951) pp. 77-86, and *Canadian J. Mathematics* (1951) p. 132, respectively.
3. J. L. Doob, *Stochastic processes*, New York, John Wiley and Sons, 1953.
4. L. E. Dubins, *Generalized random variable*, *Trans. Amer. Math. Soc.* vol. 87 (1957) pp. 273-309.
5. N. Dunford, *Uniformity in linear spaces*, *Trans. Amer. Math. Soc.* vol. 44 (1938) pp. 305-356.
6. N. Dunford and B. J. Pettis, *Linear operations on summable functions*, *Trans. Amer. Math. Soc.* vol. 47 (1940) pp. 323-392.
7. P. R. Halmos, *Measure theory*, New York, Van Nostrand, 1949.
8. J. L. Kelley, *General topology*, New York, Van Nostrand, 1955.
9. C. Kuratowski, *Topologie I*, Lwow, Warsaw, 1933.
10. B. J. Pettis, *On integration in vector spaces*, *Trans. Amer. Math. Soc.* vol. 44 (1938) pp. 277-304.
11. I. E. Segal, *Abstract probability spaces and a theorem of Kolmogoroff*, *Amer. J. Math.* vol. 76 (1954) pp. 721-732.