PERFECT PROBABILITY MEASURES AND REGULAR CONDITIONAL PROBABILITIES

By Robert H. Rodine

Purdue University and State University of New York at Buffalo

1. Summary and introduction. In an attempt to refine the axiomatic model of a probability space introduced by Kolmogorov [11], Gnendenko and Kolmogorov [5] introduced the concept of a perfect probability measure. The desirability of some sort of refinement has been pointed out by several well known examples [2], [3], [8], which display a certain amount of pathology inherent in Kolmogorov's theory. It is known that if $(X, \mathcal{S}, \mu)$ is a probability space and μ is perfect, then each of the examples mentioned is ruled out. There have been attempts (e.g., [1], [10]) at characterizing all those measurable spaces $(X, \mathcal{S})$ having the property that every probability measure μ on $\mathcal{S}$ is perfect. There have also been investigations [12], [13], [14] of the measure-theoretic properties of perfect measures and their relationships to other concepts in measure theory and probability theory.

In this paper, we consider mixtures of perfect probability measures and their relationship to regular conditional probabilities. In Section 2, mixtures are defined and some interesting special cases are considered. Section 3 is a brief study of the imbedding of mixtures in a regular conditional probability space. Using known results and some results from Section 2, the perfectness of the underlying probability measure of a regular conditional probability space is characterized.

Unless explicit mention is made to the contrary, the notation and terminology of [6] will be used throughout; however, the word "function," used without qualification, will always mean "real-valued function." If $(X, \mathcal{S}, \mu)$ is a probability space, then μ is called perfect if for every $\mathcal{S}$-measurable function f on X and every set A on the real line for which $f^{-1}(A)$ belongs to $\mathcal{S}$, there is a linear Borel set B contained in A such that $\mu(f^{-1}(A)) = \mu(f^{-1}(B))$. It is known (see [13]): (1) that a measure is perfect if and only if its restriction to every countably-generated sub-sigma-algebra is perfect, and (2) that the restriction to any sub-sigma-algebra of a perfect measure is perfect.

Since the following characterization of perfectness will be used frequently in the sequel, we quote it here as a lemma. A proof may be found in [14].

Lemma 1. *A measure μ on a measurable space $(X, \mathcal{S})$ is perfect if and only if for every $\mathcal{S}$-measurable function f on X there is a linear Borel set $B(f)$ contained in $f(X)$ such that $\mu(f^{-1}(B(f))) = \mu(X)$.*

The following definitions are those of Jiřina [9]. Let $(X, \mathcal{S}, \mu)$ be a probability space and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be sub-sigma-algebras of $\mathcal{S}$. Any function $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$ defined on $\mathcal{S}_1 \times X$ will be called a conditional probability (c.p.) if it satisfies:

CP1. for fixed S in $\mathcal{S}_1$, $\mu(S, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$ is $\mathcal{S}_2$-measurable, and

Received 2 November 1964; revised 27 July 1965.

CP2. for every S in $\mathcal{S}_1$, and every T in $\mathcal{S}_2$,

$$\mu(S \cap T) = \int_T \mu(S_1, x \mid \mathcal{S}_1, \mathcal{S}_2)\, d(\mu \mid \mathcal{S}_2),$$

where $(\mu \mid \mathcal{S}_2)$ is the restriction of μ to $\mathcal{S}_2$. If the c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$ also satisfies:

CP3. for each fixed x in X, $\mu(\cdot, x \mid \mathcal{S}_1, \mathcal{S}_2)$ is a probability measure on $\mathcal{S}_1$, then it will be called a regular conditional probability (r.c.p.).

It is well known that if $(X, \mathcal{S}, \mu)$ is a probability space, $\mathcal{S}_1$ is any sub-algebra of $\mathcal{S}$, and $\mathcal{S}_2$ is any sub-sigma-algebra of $\mathcal{S}$, then there is a c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$. We shall denote by $\mu(\cdot, \cdot \mid \mathcal{S}, \mathcal{S}_2)$ any function satisfying CP1 and CP2, with $\mathcal{S}_1 = \mathcal{S}$.

2. Mixtures of probability measures. Suppose there is given a measurable space $(X, \mathcal{S})$ and a probability space $(Y, \mathfrak{I}, \nu)$ such that for every y in Y there is a probability measure μ_y on $\mathcal{S}$ and such that $\mu_{(\cdot)}(S)$ is $\mathfrak{I}$-measurable for each S in $\mathcal{S}$. The set function μ defined on $\mathcal{S}$ by $\mu(S) = \int_Y \mu_y(S)\, d\nu$ is readily verified to be a probability measure on $\mathcal{S}$. It is usually called a mixture measure; more precisely, it is a ν-mixture of the μ_y's. The properties of μ depend, in general, upon properties of ν and of the μ_y's. In particular, the following questions arise. If the μ_y's are all perfect, is μ perfect? Conversely, if μ is perfect, does it follow that the μ_y's are perfect, except possibly for y's in a ν-null set? The answer to the first question will be shown to be negative, but the answer to the second question is unknown to us.

Before considering the general case, we discuss a few special cases of interest. If the mixing measure ν is discrete (i.e., if $(Y, \mathfrak{I}, \nu)$ is a discrete probability space), then the mixture measure μ is perfect if and only if μ_y is perfect for every y in Y. The easy proof of this fact can be based on Lemma 1.

If $(X, \mathcal{S})$ is a measurable space and μ_1 and μ_2 are probability measures on $\mathcal{S}$ such that $\mu_1 \ll \mu_2$, then the perfectness of μ_2 implies that of μ_1. Indeed, if $\{\mu_\gamma : \gamma \in \Gamma\}$ is a family of probability measures on $\mathcal{S}$ which is dominated by a probability measure λ, then the perfectness of λ implies that of μ_γ, for each γ in Γ. From this, it follows that the domination of the family $\{\mu_y : y \in Y\}$ by a perfect probability measure is sufficient to insure the perfectness of the mixture measure μ, whatever be the mixing measure, ν. For, if $\{\mu_n : n = 1, 2, \cdots\}$ is an equivalent countable subset of the $\{\mu_y\text{'s}\}$ (see [7], p. 232), then certainly μ_n is perfect, for each n. And if $\bar{\mu}(S) = \sum_{n=1}^{\infty} 2^{-n} \mu_n(S)$, then $\bar{\mu}$ is perfect, since it is a discrete mixture. It is clear that $\mu \ll \bar{\mu}$, and the result follows from the opening sentence of this paragraph.

Of course, the perfectness of μ does not imply the existence of a perfect probability measure which dominates the family $\{\mu_y : y \in Y\}$. As a counter example, we may take $X = Y = [0, 1]$ and $\mathcal{S} = \mathfrak{I} =$ the sigma-algebra of Borel sets of X. Let ν be Lebesgue measure on $\mathfrak{I}$ and for each y in Y, let $\mu_y(S) = I_S(y)$, the indicator function of the set S. Then μ is also Lebesgue measure on the Borel sets of $[0, 1]$. Moreover, μ, ν, and each μ_y are perfect probability measures on $\mathcal{S}$. But, if there were a measure λ on $\mathcal{S}$ such that $\{\lambda\} \gg \{\mu_y : y \in Y\}$, then $\lambda(\{y\}) > 0$, for each y in Y, which is impossible. However, the following is true: if the family

$\{\mu_y : y \varepsilon Y\}$ is dominated by a probability measure λ and $\mu \gg \lambda$, where μ is perfect, then μ_y is perfect, for every y in Y.

The example of the previous paragraph can be generalized so as to yield an example of a mixture of perfect measures which is not perfect. Let X be any set, and $\mathcal{S}$, any sigma-algebra of subsets of X. Take $Y = X$, $\mathfrak{J} = \mathcal{S}$ and let ν be any probability measure on $\mathfrak{J}$. For each y in Y, let $\mu_y(S) = I_S(y)$. By taking ν as the mixing measure, and the probability measures $\{I_{(\cdot)}(y): y \varepsilon Y\}$ as mixand measures, it is easy to see that for every S in $\mathcal{S}$, $\mu(S) = \nu(S)$. Thus, any nonperfect mixture of this family $\{\mu_y : y \varepsilon Y\}$ is itself nonperfect. That there is such a nonperfect mixing measure is well known (see, e.g. [14]). Hence not every mixture of perfect measures is perfect. A natural conjecture is that a perfect mixture of perfect measures is perfect, but whether or not this conjecture is true remains an open question.

3. Regular conditional probabilities and perfectness. In order to discuss mixtures and regular conditional probabilities, we introduce some convenient definitions. A family $\{\mu_\gamma : \gamma \varepsilon \Gamma\}$ of probability measures on a measurable space $(X, \mathcal{S})$ will be said to be an equiperfect family if for every $\mathcal{S}$-measurable function f on X there is a linear Borel set $B(f)$ contained in $f(X)$ such that $\mu_\gamma(f^{-1}(B(f))) = 1$ for every γ in Γ. If the index set Γ consists of a single point, then this definition reduces to that of perfectness (cf. Lemma 1). If the index set Γ has the structure of a probability space $(\Gamma, \mathfrak{F}, \alpha)$, and if for every $\mathcal{S}$-measurable function f on X there is a linear Borel set $B(f)$ contained in $f(X)$ and a set $N(f)$ in $\mathfrak{F}$ such that $\alpha(N(f)) = 0$ and $\mu_\gamma(f^{-1}(B(f))) = 1$ for every γ not in $N(f)$, then $\{\mu_\gamma : \gamma \varepsilon \Gamma\}$ will be called an equiperfect, almost everywhere $[\alpha]$ family (an e.p.a.e. $[\alpha]$ family).

If $(X, \mathcal{S}, \mu)$ is a probability space and $\mathcal{S}_1$ and $\mathcal{S}_2$ are sub-sigma-algebras of $\mathcal{S}$ for which there is a r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$, then $\mu \mid \mathcal{S}_1$ is a $\mu \mid \mathcal{S}_2$-mixture of the family $\{\mu(\cdot, x \mid \mathcal{S}_1, \mathcal{S}_2): x \varepsilon X\}$ of probability measures. What will be shown in the following is that given any mixture problem, it is possible to construct an associated probability space into which the mixture structure can be imbedded. We remark that the mixture structure cannot be completely recovered after imbedding into the r.c.p. space, since the imbedded structure is related to the original through isomorphisms of certain probability spaces.

Let $(X, \mathcal{S})$ be a measurable space and $(Y, \mathfrak{J}, \nu)$ a probability space such that for each y in Y, μ_y is a probability measure on $\mathcal{S}$ and such that $\mu_{(\cdot)}(S)$ is a $\mathfrak{J}$-measurable function for each S in $\mathcal{S}$. Let μ be a ν-mixture of the μ_y's; i.e., $\mu(S) = \int_Y \mu_y(S)\, d\nu$, S in $\mathcal{S}$. Consider the measurable space $(X \times Y, \mathcal{S} \times \mathfrak{J})$, where $X \times Y$ is the Cartesian product of X and Y and $\mathcal{S} \times \mathfrak{J}$ is the sigma-algebra generated by the class of sets $\{S \times T: S \varepsilon \mathcal{S}, T \varepsilon \mathfrak{J}\}$. For any set E in $\mathcal{S} \times \mathfrak{J}$, let E^y denote the y-section of E; i.e., $E^y = \{x: (x, y) \varepsilon E\}$. Define a function $\lambda(\cdot, \cdot)$ on $(\mathcal{S} \times \mathfrak{J}) \times (X \times Y)$ by $\lambda(E, (x, y)) = \mu_y(E^y)$ and a function $\lambda(\cdot)$ on $\mathcal{S} \times \mathfrak{J}$ by $\lambda(E) = \int_Y \lambda(E, (x, y))\, d\nu$. We observe that $\lambda(\cdot, \cdot)$ and $\lambda(\cdot)$ are well defined, since the class $\mathcal{R} = \{S \times T: S \varepsilon \mathcal{S}, T \varepsilon \mathfrak{J}\}$ is a semi-algebra of subsets of $X \times Y$ which is contained in the normal class $\mathcal{U}$ of subsets of $X \times Y$ for which U^y belongs to $\mathcal{S}$ for every y in Y and $\mu_y(U^y)$ is a $\mathcal{T}$-measurable function. A well known lemma (see [6], p. 28) insures that the normal class generated by $\mathcal{R}$ coincides with the sigma-algebra generated by $\mathcal{R}$, from which we conclude that $\mathcal{S} \times \mathcal{T} \subset \mathcal{U}$.

THEOREM 2. (i) $(X \times Y, \mathcal{S} \times \mathcal{T}, \lambda)$ *is a probability space.* (ii) $(X \times Y, \mathcal{S} \times Y, \lambda \mid (\mathcal{S} \times Y))$ *and* $(X, \mathcal{S}, \mu)$ *are set-isomorphic probability spaces, as are* $(X \times Y, X \times \mathcal{T}, \lambda \mid (X \times \mathcal{T}))$ *and* $(Y, \mathcal{T}, \nu)$. (iii) $\lambda(\cdot, \cdot)$ *is a* r.c.p., $\lambda(\cdot, \cdot \mid \mathcal{S} \times \mathcal{T}, X \times \mathcal{T})$.

PROOF. (i) is a special case of a result of Yushkevich ([15], Lemma 2, p. 218). To prove (ii), we observe that $\lambda(S \times Y, (x, y)) = \mu_y(S)$, $\lambda(S \times Y) = \mu(S)$, $\lambda(X \times T, (x, y)) = I_T(y)$, and $\lambda(X \times T) = \nu(T)$. The desired set isomorphisms are given by $\gamma_1(S) = S \times Y$ and $\gamma_2(T) = X \times T$. To prove (iii), we note that $\lambda(E; (x, y))$ is a constant function of the argument x and hence that $\lambda(E, \cdot)$ is, for fixed E in $\mathcal{S} \times \mathcal{T}$, an $X \times \mathcal{T}$-measurable function. For every E in $\mathcal{S} \times \mathcal{T}$ and every F in $X \times \mathcal{T}$, $\lambda(E \cap F) = \int_Y \mu_y(E \cap F)^y \, d\nu$, which may be written as $\int_F \lambda(E, (x, y)) \, d(\lambda \mid X \times \mathcal{T})$. Since μ_y is a probability measure on $\mathcal{S}$ for each y in Y, it follows from the properties of y-sections of measurable sets that $\lambda(\cdot, (x, y))$ is, for fixed (x, y) in $X \times Y$, a probability measure on $\mathcal{S} \times \mathcal{T}$.

If $(X, \mathcal{S}, \mu)$ is a probability space, and $\mathcal{S}_1$ and $\mathcal{S}_2$ are sub-sigma-algebras of $\mathcal{S}$ for which there is a r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$, then $\mu \mid \mathcal{S}_1$ is a $\mu \mid \mathcal{S}_2$-mixture of the probability measures $\{\mu(\cdot, x \mid \mathcal{S}_1, \mathcal{S}_2) : x \in X\}$. This fact, together with the following lemma will yield our next theorem.

LEMMA 3. *The mixture measure* μ *is perfect if and only if* $\{\mu_y : y \in Y\}$ *is an* e.p.a.e. $[\nu]$ *family.*

PROOF. The proof follows from Lemma 1 and standard integration techniques.

THEOREM 4. *If* $(X, \mathcal{S}, \mu)$ *is a probability space and* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *are sub-sigma-algebras of* $\mathcal{S}$ *for which there is a* r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$, *then* $\mu \mid \mathcal{S}_1$ *is perfect if and only if* $\{\mu(\cdot, x \mid \mathcal{S}_1, \mathcal{S}_2) : x \in X\}$ *is an* e.p.a.e. $[\mu \mid \mathcal{S}_2]$ *family.*

Jiřina [9] has shown that if $(X, \mathcal{S}, \mu)$ is a probability space and $\mathcal{S}_1$ and $\mathcal{S}_2$ are sub-sigma-algebras of $\mathcal{S}$ with $\mathcal{S}_1$ countably generated, then the perfectness of μ implies the existence of a r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$. Our next lemma is a sharpening of this result which will allow us to improve Theorem 4.

LEMMA 5. *Let* $(X, \mathcal{S}, \mu)$ *be a probability space and* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *be sub-sigma-algebras of* $\mathcal{S}$, *with* $\mathcal{S}_1$ *countably generated. If* $\mu \mid \mathcal{S}_1$ *is perfect, then there is a* r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$.

PROOF. A proof for this lemma can be constructed along the lines suggested by the proof of a theorem of Doob (see [4], p. 31), after noting that the hypotheses of the lemma ensure the existence of an $\mathcal{S}_1$-measurable function h on X such that $\mathcal{S}_1$ coincides with the class of inverse images, under h, of the linear Borel sets. This, in turn, implies ([4], p. 29) the existence of a "wide-sense" r.c.p., and the proof now follows via Lemma 1.

THEOREM 6. *Let* $(X, \mathcal{S}, \mu)$ *be a probability space, and let* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *be sub-sigma-algebras of* $\mathcal{S}$, *with* $\mathcal{S}_1$ *countably generated. The probability measure* $\mu \mid \mathcal{S}_1$ *is perfect* *if and only if* (i) *there is a* r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$ *and* (ii) $\{\mu(\cdot, x \mid \mathcal{S}_1, \mathcal{S}_2): x \varepsilon X\}$ *is an* e.p.a.e. $[\mu \mid \mathcal{S}_2]$ *family.*

PROOF. (i) follows immediately from Lemma 5. Theorem 4 gives (ii), even without the countability restriction. The converse has been proved above.

THEOREM 7. *Let* $(X, \mathcal{S}, \mu)$ *be a probability space and* $\mathcal{S}_2$ *any sub-sigma-algebra of* $\mathcal{S}$. *The probability measure* μ *is perfect if and only if* (i) *there is a* r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$ *for every countably generated sub-sigma-algebra* $\mathcal{S}_1$ *of* $\mathcal{S}$, *and* (ii) $\{\mu(\cdot, x \mid \mathcal{S}, \mathcal{S}_2): x \varepsilon X\}$ *is an* e.p.a.e. $[\mu \mid \mathcal{S}_2]$ *family.*

PROOF. (Necessity). (i) follows from Lemma 5 and (2) of Section 1. Suppose that f is any $\mathcal{S}$-measurable function on X, and let $\mathcal{S}_1$ be the class of inverse images under f of the linear Borel sets. Since for any S in $\mathcal{S}$, $\mu(S, x \mid \mathcal{S}, \mathcal{S}_2) = \mu(S, x \mid \mathcal{S}_1, \mathcal{S}_2)$ except for x's in a set $N(S)$ in $\mathcal{S}_2$ with $(\mu \mid \mathcal{S}_2)(N(S)) = 0$, we see that there is a set $N_1(f)$ in $\mathcal{S}_2$ such that $(\mu \mid \mathcal{S}_2)(N_1(f)) = 0$ and if x is not in $N_1(f)$, then $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2) = (\mu \mid \mathcal{S}_1)(\cdot, \cdot \mid \mathcal{S}, \mathcal{S}_2)$. It follows from Theorem 6 that there is a linear Borel set $B(f)$ contained in $f(X)$ and a set $N_2(f)$ in $\mathcal{S}_2$ such that $(\mu \mid \mathcal{S}_2)(N_2(f)) = 0$ and $\mu(f^{-1}(B(f)), x \mid \mathcal{S}_1, \mathcal{S}_2) = 1$ for each x not belonging to $N_2(f)$. Hence for x not in $N_1(f) \cup N_2(f)$, $\mu(f^{-1}(B(f)), x \mid \mathcal{S}, \mathcal{S}_2) = 1$, and $(\mu \mid \mathcal{S}_2)(N_1(f) \cup N_2(f)) = 0$, which proves (ii).

(Sufficiency). If $\mathcal{S}_1$ is any countably generated sub-sigma-algebra of $\mathcal{S}$, then by (i), there is a r.c.p. $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$. If f is any $\mathcal{S}_1$-measurable function, then, by (ii), there is a set $N_1(f)$ in $\mathcal{S}_2$ and a linear Borel set $B(f)$ contained in $f(X)$ such that if x is not in $N_1(f)$, then $\mu(f^{-1}(B(f)), x \mid \mathcal{S}, \mathcal{S}_2) = (\mu \mid \mathcal{S}_1)(f^{-1}(B(f)), x \mid \mathcal{S}, \mathcal{S}_2) = 1$. But $\mu(\cdot, \cdot \mid \mathcal{S}_1, \mathcal{S}_2)$ and $(\mu \mid \mathcal{S}_1)(\cdot, \cdot \mid \mathcal{S}, \mathcal{S}_2)$ differ on (at most) a set $N_2(f)$ in $\mathcal{S}_2$ with $(\mu \mid \mathcal{S}_2)(N_2(f)) = 0$. Thus for x not in $N_1(f) \cup N_2(f)$, $\mu(f^{-1}(B(f)), x \mid \mathcal{S}_1, \mathcal{S}_2) = 1$, and according to Theorem 6, $\mu \mid \mathcal{S}_1$ is perfect. It now follows from (1), Section 1, that μ is perfect.

4. Acknowledgments. This paper is based on a portion of the author's doctoral dissertation written at Purdue University. I wish to thank Professor L. J. Cote for his guidance and the referee for his helpful comments.

REFERENCES

[1] BLACKWELL, D. (1956). On a class of probability spaces. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **2** 1–6.

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

[3] DOOB, J. (1948). On a problem of Marczewski. *Colloq. Math.* **1** 216–217.

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

[5] GNEDENKO, B. and A. KOLMOGOROV (1954). *Limit Distributions of Sums of Independent Random Variables.* Addison-Wesley, Cambridge.

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

[7] HALMOS, P. and L. SAVAGE (1949). Application of the Radon-Nikodym theorem to the theory of sufficient statistics. *Ann. Math. Statist.* **20** 225–241.

[8] JESSEN, B. (1948). On two notions of independent functions. *Colloq. Math.* **1** 214–215.

[9] JIŘINA, M. (1954). Conditional probabilities on strictly separable sigma-algebras. *Czechoslovak Math. J.* **4** 372–380.

[10] KALLIANPUR, G. (1959). A note on perfect probability. *Ann. Math. Statist.* **30** 169–172.
[11] KOLMOGOROV, A. (1933). *Grundebegriffe der Wahrscheinlichkeitrechnung.* Springer, Berlin.
[12] MARCZEWSKI, E. (1953). On compact measures. *Fund. Math.* **40** 111–124.
[13] RYLL-NARDZEWSKI, C. (1953). On quasi-compact measures. *Fund. Math.* **40** 125–130.
[14] SAZONOV, V. (1962). On perfect measures. *Izvestia Acad. Sci. S.S.S.R. Seria Math.* **26** 391–414 (In Russian).
[15] YUSHKEVICH, A. (1960). On the definition of a strong Markov process. *Theor. Prob. Appl.* **5** 216–220.