ON DECOMPOSITION INTO STOCHASTICALLY INDEPENDENT COMPONENTS

S. P. LLOYD

1. Introduction. Let $\mathcal{G}$ be an arbitrary class of measurable sets in the probability space $(\Omega, \mathfrak{F}, P)$. We denote by $\mathcal{G}^s$ the class of sets which are stochastically independent of $\mathcal{G}$; that is,

$$\mathcal{G}^s = \{H \in \mathfrak{F}: P(H \cap G) = P(H)P(G) \text{ for all } G \in \mathcal{G}\}.$$

If $\mathcal{G} \vee \mathcal{G}^s$ denotes the smallest σ-field containing $\mathcal{G}$ and $\mathcal{G}^s$, we say that decomposition (of $\mathfrak{F}$ relative to $\mathcal{G}$) holds if $\mathcal{G} \vee \mathcal{G}^s = \mathfrak{F}$. In other words, when decomposition holds, every statistical experiment in the class under consideration (every $\mathfrak{F}$ set) can be expressed as a combination of (is in the σ-field generated by) experiments from the designated sublcass (sets in $\mathcal{G}$) and experiments independent of the subclass (sets in $\mathcal{G}^s$). Examples will be given where decomposition holds and where it fails, and a sufficient condition for decomposition to hold will be established.

2. Some properties of independence. We digress briefly to exhibit some elementary properties of the ^{s} operation. We denote by $\mathcal{O}$ the σ-field consisting of sets of probability 0 and their complements. The following relations hold for arbitrary subclasses $\mathcal{G}, \mathcal{H}, \cdots$, of $\mathfrak{F}$.

(i) $\mathcal{G}^s$ is closed under the formation of complements, proper differences, and (countable) disjoint unions (but not necessarily intersections),

(ii) $\mathcal{O}^s = \mathfrak{F}$, $\mathfrak{F}^s = \mathcal{O}$,

(iii) $\mathcal{G}^s \supset \mathcal{O}$,

(iv) If $\mathcal{G} \subset \mathcal{H}$ then $\mathcal{G}^s \supset \mathcal{H}^s$,

(1) (v) $(\mathcal{G} \cup \mathcal{H})^s = \mathcal{G}^s \cap \mathcal{H}^s$,

(vi) $(\mathcal{G} \cap \mathcal{H})^s \supset \mathcal{G}^s \cup \mathcal{H}^s$,

(vii) $\mathcal{G} \subset \mathcal{G}^{ss} (= (\mathcal{G}^s)^s)$,

(viii) $\mathcal{G}^{sss} = \mathcal{G}^s$,

(ix) $\mathcal{G} \cap \mathcal{G}^s \subset \mathcal{O}$,

(x) $(\mathcal{G} \cup \mathcal{G}^s)^s = \mathcal{O}$, $(\mathcal{G} \cup \mathcal{G}^s)^{ss} = \mathfrak{F}$,

(xi) $(\mathcal{G} \vee \mathcal{G}^s)^s = \mathcal{O}$, $(\mathcal{G} \vee \mathcal{G}^s)^{ss} = \mathfrak{F}$.

Proofs of the above properties are simple and straightforward, and are omitted. It is to be noted that the problem of when $\mathcal{G} \vee \mathcal{G}^s = \mathfrak{F}$ holds is subsumed by the problem of when $\mathcal{H}^{ss} = \mathcal{H}$ holds, from (xi).

Presented to the Society January 22, 1962 under the title *On decomposition into independent components*; received by the editors February 2, 1962.

3. **Conditional cumulative distributions.** In all that follows we assume that $\mathcal{G}$ is a sub-σ-field of $\mathcal{F}$.

THEOREM 1. *$H \in \mathcal{G}^s$ if and only if $P(H|\mathcal{G}) = P(H)$ with probability* 1.

PROOF. $P(H|\mathcal{G})$ is the unique element of $L_\infty(\Omega, \mathcal{G}, P)$ satisfying

$$P(H \cap G) = \int_G P(H \mid \mathcal{G})_\omega dP(\omega) \qquad \text{for all } G \in \mathcal{G}. \tag{2}$$

When $H \in \mathcal{G}^s$ the left-hand side is equal to $P(H)P(G)$, and a solution of (2) is clearly $P(H|\mathcal{G}) = \text{const} = P(H)$ with probability 1. Conversely, if $P(H|\mathcal{G})$ is constant with probability 1 then $P(H \cap G) = P(H)P(G)$ holds for every $G \in \mathcal{G}$, whence $H \in \mathcal{G}^s$.

If x is a real valued random variable we denote by $[x]$ the smallest σ-field with respect to which x is Borel measurable, that is, $[x]$ is the class of sets $\{\omega: x(\omega) \in B\}$ for all linear Borel sets B. The conditional cumulative distribution of x relative to $\mathcal{G}$ will be denoted by

$$F(\lambda, \omega) = P(x \leqq \lambda \mid \mathcal{G})_\omega, \qquad -\infty < \lambda < \infty, \ \omega \in \Omega.$$

We can and do assume that the conditional probability on the right-hand side is determined for each λ in such a way that $F(\cdot, \omega)$ is a bona fide cumulative distribution function for each $\omega \in \Omega$ [1, p. 29].

THEOREM 2. *Suppose $F(\lambda, \omega)$, $-\infty < \lambda < \infty$, $\omega \in \Omega$, is continuous on $-\infty < \lambda < \infty$ for each ω in the complement of a set of probability* 0. *Then the random variable ξ defined by*

$$\xi(\omega) = F(x(\omega), \omega), \qquad \omega \in \Omega$$

has the properties

(i) $[\xi] \subset [x] \vee \mathcal{G}$,

(ii) $[x] \subset [\xi] \vee \mathcal{G}$,

(iii) $[\xi] \subset \mathcal{G}^s$.

REMARK. P. Lévy is the first to use the independence property of random conditional cumulatives; Lévy gives no proof [2]. Since there seems to be no published proof, and since proof is not as simple as would seem off hand, we give it here.

We need first a measure theoretic lemma. Let $\mathcal{R}$ denote the σ-field of linear Borel sets, and let $(\Omega \times \Lambda, \mathcal{G} \times \mathcal{R})$ denote the product measurable space of $(\Omega, \mathcal{G})$ with $(\Lambda, \mathcal{R})$, where Λ is the real line. If $\Gamma \subset \Omega \times \Lambda$ then Γ_ω denotes the ω section $\Gamma_\omega = \{\lambda \in \Lambda: (\omega, \lambda) \in \Gamma\}$. For any Borel set B, define

$$F(B, \omega) = \int_B F(d\lambda, \omega).$$

LEMMA 1. *If* $\Gamma \in \mathcal{G} \times \mathcal{R}$ *then*

$$P(x(\omega) \in \Gamma_\omega) = \int F(\Gamma_\omega, \omega) dP(\omega), \qquad \Gamma \in \mathcal{G} \times \mathcal{R}. \tag{3}$$

PROOF. The properties of $F(\cdot, \cdot)$ insure that $\Delta_1(\Gamma) = \int F(\Gamma_\omega, \omega) dP(\omega)$ is a measure on $\mathcal{G} \times \mathcal{R}$. With $\delta: \Omega \to \Omega \times \Lambda$ defined by $\delta(\omega) = (\omega, x(\omega))$, let Δ_2 be the measure $\Delta_2(\Gamma) = P(\delta^{-1}(\Gamma))$, $\delta^{-1}(\Gamma) \in \mathfrak{F}$, noting that $\Delta_2(\Gamma) = P(x(\omega) \in \Gamma_\omega)$. When Γ is a $\mathcal{G} \times \mathcal{R}$ rectangle, $\Gamma = G \times B$, $G \in \mathcal{G}$, $B \in \mathcal{R}$, we have

$$\begin{aligned} \Delta_2(\Gamma) &= P(G \cap \{x \in B\}) \\ &= \int_G P(x \in B \mid \mathcal{G})_\omega dP(\omega) \\ &= \int F(\Gamma_\omega, \omega) dP(\omega) = \Delta_1(\Gamma). \end{aligned}$$

Since measures Δ_1 and Δ_2 coincide on $\mathcal{G} \times \mathcal{R}$ rectangles, they coincide on $\mathcal{G} \times \mathcal{R}$, proving the lemma.

PROOF OF THEOREM 2. Define the pseudoinverses of $F(\cdot, \omega)$ as

$$\begin{aligned} \overline{G}(u, \omega) &= \sup \lambda: F(\lambda, \omega) \leqq u \\ \underline{G}(u, \omega) &= \min \lambda: F(\lambda, \omega) \geqq u, \\ & \qquad 0 < u < 1, \ \omega \in \Omega. \end{aligned}$$

We have, for all values of the arguments,

$$\begin{aligned} &F(\lambda, \omega) < u \quad \text{if and only if} \quad \lambda < \underline{G}(u, \omega), \\ &0 < u < 1, \quad -\infty < \lambda < \infty, \quad \omega \in \Omega. \end{aligned} \tag{4}$$

The assumed continuity of $F(\cdot, \omega)$ gives

$$F(\underline{G}(u, \omega), \omega) = F(\overline{G}(u, \omega), \omega) = u, \qquad 0 < u < 1,$$

with probability 1.

We will show that the conditional distribution of ξ given $\mathcal{G}$ is the uniform-on-the-unit-interval distribution with probability 1. For each $G \in \mathcal{G}$ and each u, $0 < u < 1$, we have

$$\begin{aligned} P(G \cap \{\xi < u\}) &= P(G \cap \{F(x(\omega), \omega) < u\}) \\ &= P(G \cap \{x(\omega) < \underline{G}(u, \omega)\}) \\ &= P(x(\omega) \in \Gamma_\omega) \end{aligned}$$

where $\Gamma \subset \Omega \times \Lambda$ is the set

$$\Gamma = (G \times \Lambda) \cap \Delta,$$
$$\Delta = \{(\omega, \lambda): \lambda < \underline{G}(u, \omega)\}.$$

It follows from (4) that $\underline{G}(u, \cdot)$ is $\mathcal{G}$ measurable for each fixed u, $0<u<1$; to see that Δ is $\mathcal{G}\times\mathcal{R}$ measurable, observe that $\Delta=\lim_n \Delta_n$, with

$$\Delta_n = \bigcup_{j=-\infty}^{\infty} D_{n,j} \times (-\infty, j/2^n)$$
$$D_{n,j} = \{\omega: j/2^n < \underline{G}(u, \omega) \leqq (j+1)/2^n\},$$

and the approximating sequence is manifestly $\mathcal{G}\times\mathcal{R}$ measurable. Applying Lemma 1, we have, for each $G\in\mathcal{G}$ and each u, $0<u<1$,

$$\begin{aligned}\int_G P(\xi < u \mid \mathcal{G})_\omega dP(\omega) &= P(G \cap \{\xi < u\}) = P(x(\omega) \in \Gamma_\omega)\\ &= \int F(\Gamma_\omega, \omega) dP(\omega)\\ &= \int_G F(\underline{G}(u, \omega) - 0, \omega) dP(\omega)\\ &= uP(G), \qquad G \in \mathcal{G},\ 0 < u < 1,\end{aligned}$$

whence $P(\xi<u|\mathcal{G})=u$ with probability 1, and this for each u, $0<u<1$.

That ξ is $[x]\vee\mathcal{G}$ measurable follows easily from (4). It remains to show that x is $[\xi]\vee\mathcal{G}$ measurable. Define random variables $\underline{x}$ and $\bar{x}$ by

$$\begin{aligned}\underline{x}(\omega) &= \underline{G}(\xi(\omega), \omega)\\ \bar{x}(\omega) &= \overline{G}(\xi(\omega), \omega),\end{aligned} \qquad (\text{where } \omega \in \Omega),$$

observing that $\underline{x}$ is $[\xi]\vee\mathcal{G}$ measurable, from (4), and that $\underline{x}(\omega)\leqq x(\omega)\leqq\bar{x}(\omega)$ holds for each $\omega\in\Omega$. We will show that $\underline{x}=\bar{x}$ with probability 1. For each $\omega\in\Omega$ let $I(\omega)$ denote the union of the disjoint nontrivial closed intervals on which $F(\cdot, \omega)$ is constant. Since $\{\omega: \underline{x}(\omega)<\bar{x}(\omega)\}\subset\{\omega: x(\omega)\in I(\omega)\}$, it will suffice to prove that $P(x(\omega)\in I(\omega))=0$. If Lemma 1 is applicable we have $P(x(\omega)\in I(\omega))=\int F(I(\omega), \omega)dP(\omega)=0$, under the assumption that the $\Omega\times\Lambda$ set Γ whose ω sections are $\Gamma_\omega=I(\omega)$ is $\mathcal{G}\times\mathcal{R}$ measurable. To see that this is the case, observe that $\Gamma=\lim_n \Gamma_n$, with

$$\Gamma_n = \bigcup_{-\infty<j<k<\infty} G_{n,j,k} \times [(j-1)/2^n, (k+1)/2^n],$$
$$\begin{aligned}G_{n,j,k} = \{\omega: &F((j-1)/2^n, \omega) < F(j/2^n, \omega)\\ &= F(k/2^n, \omega) < F((k+1)/2^n, \omega)\}.\end{aligned}$$

Our main result is the following, to the effect that if there exists one such random variable ξ then there exist many.

THEOREM 3. *Suppose that there exists a random variable independent of* $\mathcal{G}$ *whose distribution is continuous. Then decomposition holds; that is,* $\mathcal{G}\vee\mathcal{G}^s=\mathcal{F}$.

PROOF. From Theorem 2, we may assume given a random variable ξ with the property $P(\xi<u\mid\mathcal{G})=u$ with probability 1, $0<u<1$. For arbitrary $A\in\mathcal{F}$ consider the random variable $\eta_A=\chi_A+\xi$, where χ_A denotes the characteristic function of set A. The conditional distribution of η_A has the property

$$P(\lambda_1\leqq\eta_A<\lambda_2\mid\mathcal{G})=P(A\cap\{\lambda_1-1\leqq\xi<\lambda_2-1\}\mid\mathcal{G})$$
$$+P(A'\cap\{\lambda_1\leqq\xi<\lambda_2\}\mid\mathcal{G})\leqq 2(\lambda_2-\lambda_1)$$

with probability 1, implying that some version of $P(\eta_A<\lambda\mid\mathcal{G})$ is continuous in λ with probability 1. Applying Theorem 2 to η_A, we obtain a random variable ξ_A such that η_A is $[\xi_A]\vee\mathcal{G}$ measurable and such that $[\xi_A]\subset\mathcal{G}^s$. Since $0<\xi<1$ with probability 1, set A differs from the set $\{\eta_A>1\}$ by a set of probability 0, whence $A\in[\eta_A]\subset[\xi_A]\vee\mathcal{G}\subset\mathcal{G}^s\vee\mathcal{G}$ modulo set of probability 0. But A is an arbitrary element of $\mathcal{F}$, so that $\mathcal{F}\subset\mathcal{G}\vee\mathcal{G}^s$, or since the reverse inclusion is trivial, $\mathcal{F}=\mathcal{G}\vee\mathcal{G}^s$.

4. Examples.

Example 1. Decomposition holds. Let $(\Omega,\mathcal{F},P)$ be atomless, and let $\mathcal{G}$ be purely atomic. That is, there is a finite or countable partition of Ω into disjoint sets G_i, $P(G_i)>0$, $i=1,2,\cdots$, and $\mathcal{G}$ is the σ-field generated by the G_i, $i=1,2,\cdots$. From Theorem 1, $H\in\mathcal{G}^s$ if and only if $P(H\cap G_i)/P(G_i)$ is independent of i. To see that decomposition holds, let $A\in\mathcal{F}$ be an arbitrary measurable set, and define $H_{ii}=A\cap G_i$, $i=1,2,\cdots$. Since $(\Omega,\mathcal{F},P)$ is atomless, for each i there exist for all j sets $H_{ij}\subset G_j$ such that $P(H_{ij})/P(G_j)=P(H_{ii})/P(G_i)$. With $H_i=\bigcup_j H_{ij}$, we have $H_i\in\mathcal{G}^s$, $i=1,2,\cdots$, and finally $A=\bigcup_i H_{ii}=\bigcup_i(G_i\cap H_i)\in\mathcal{G}\vee\mathcal{G}^s$. We will prove also that $\mathcal{G}=\mathcal{G}^{ss}$ modulo sets of probability 0. Suppose $A\in\mathcal{F}$ does not differ from some $\mathcal{G}$ set by a set of probability 0. Then both $P(A\cap G_i)>0$ and $P(A'\cap G_i)>0$ hold for at least one i, say $i=j$. Choose $H_j\subset A\cap G_j$ and $K_j\subset A'\cap G_j$ so that $P(H_j)=P(K_j)>0$. For all $k\neq j$ choose sets $H_k=K_k\subset G_k$ so that $P(H_k)/P(G_k)=P(H_j)/P(G_j)$. Defining $H=\bigcup_l H_l$, $K=\bigcup_l K_l$, we have H, $K\in\mathcal{G}^s$ and $P(H)=P(K)$. Since $P(A\cap H)-P(A\cap K)=2P(H_j)>0$, at least one of H, K is not independent of A, implying $A\notin\mathcal{G}^{ss}$. Thus $\mathcal{G}^{ss}\subset\mathcal{G}$ modulo sets of probability 0, whence $\mathcal{G}^{ss}=\mathcal{G}$ modulo sets of probability 0, from (1.vii).

Example 2. Decomposition fails. Let $\Omega = \{(x, i): 0 \leqq x \leqq 1, i = 1, 2\}$, and let probability have density $p_i(x)$ at (x, i) relative to linear Borel measure on the two intervals which comprise Ω. Assume that $p_1(x) + p_2(x) = 1$, $0 \leqq x \leqq 1$. If B_1, B_2 are Borel subsets of the unit interval we denote by $[B_1, B_2]$ the ω set $\{(x, 1): x \in B_1\} \cup \{(x, 2): x \in B_2\}$. Let $\mathcal{G}$ be the sub-σ-field of sets of the form $[B, B]$ for all Borel B. It is readily verified that conditional probabilities relative to $\mathcal{G}$ have the form

$$P([B_1, B_2] \mid \mathcal{G})_{(x,i)} = \chi_{B_1}(x)p_1(x) + \chi_{B_2}(x)p_2(x), \qquad 0 \leqq x \leqq 1,\ i = 1, 2. \tag{5}$$

For each ω the measure $P(\cdot \mid \mathcal{G})_\omega$ is purely atomic, the measures of the two atoms being $p_1(x)$, $p_2(x)$. It is easy to choose p_1, p_2 in such a way that the only constant functions of the form (5) are 0 and 1, which is to say, $\mathcal{G}^s = \mathcal{O}$. Thus decomposition fails. Note also that $\mathcal{G}^{ss} = \mathcal{O}^s = \mathcal{F} \neq \mathcal{G}$.

Example 2 suggests the following conjecture: for decomposition to hold it is sufficient that the conditional probability $P(\cdot \mid \mathcal{G})$ be an atomless measure with probability 1. (This statement makes sense in the Stone representation space; see [3].)

5. **Acknowledgment.** Example 2 was suggested by Professor Murray Rosenblatt, and the author wishes to express his thanks to Professor Rosenblatt for a number of interesting conversations on this problem.

References

1. J. L. Doob, *Stochastic processes*, John Wiley, New York, 1953.
2. P. Lévy, *Theorie de l'addition variables aléatoires*, Gauthier-Villars, Paris, 1937.
3. P. R. Halmos, *On a theorem of Dieudonné*, Proc. Nat. Acad. Sci. U.S.A. **35** (1949), 38–42.

Bell Telephone Laboratories,
Murray Hill, New Jersey