The Annals of Statistics
1985, Vol. 13, No. 3, 1244–1248

COHERENT PREDICTIONS ARE STRATEGIC[1]

BY DAVID A. LANE AND WILLIAM D. SUDDERTH

University of Minnesota

Two random quantities x and y, taking values in sets X and Y, are to be observed sequentially. A predicter (bookie) posts odds on (x, y) and on y given x according to functions P and $q(x)$, respectively. The predicter is coherent (the bookie can avoid a sure loss) if and only if P is a finitely additive probability distribution on $X \times Y$ and q satisfies a general law of total probability:

$$P(A) = \int q(x)(Ax)P_0(dx)$$

for $A \subset X \times Y$, $Ax = \{y: (x, y) \in A\}$, P_0 = marginal of P on X.

1. Introduction. A general law of total probability can be written in the form

$$P(A) = \int q(x)(Ax)P_0(dx) \tag{1}$$

where P is a probability distribution on the product space $X \times Y$, P_0 is the marginal of P on X, $A \subset X \times Y$, $Ax = \{y: (x, y) \in A\}$, and q is a conditional probability distribution for y given x. (A measure P which admits such a disintegration is called *strategic* [1].) In the conventional theory of countably additive probability, q is required by definition to satisfy

$$P(C \times D) = \int_C q(x)(D)P_0(dx) \tag{2}$$

for $C \subset X$, $D \subset Y$ and (1) is then derived for measurable sets A using standard σ-field arguments. However, even (2) is difficult to justify when probabilities are interpreted as relative frequencies if $P_0\{x\} = 0$ for each $x \in X$. The object of this note is to derive (1) directly from an assumption of coherence which is formulated below.

Consider x and y to be random quantities that are to be observed in sequence. A predicter (bookie) expresses his predictions about the pair (x, y) by a function P defined on subsets of $X \times Y$ with values in $[0, 1]$, and his conditional predictions about y given x by a function $q(x)$ defined on subsets of Y and also having values in $[0, 1]$ for every $x \in X$. The functions are interpreted, following de Finetti [2] in this way: for $A \subset X \times Y$, $P(A)$ is the price in dollars at which the predicter is neutral between buying and selling a ticket worth \$1 if $(x, y) \in A$ and \$0

Received March 1984; revised December 1984.

[1] Research supported by National Science Foundation Grant MCS 8100789.

AMS 1980 *subject classifications.* 62A15, 60A05.

Key words and phrases. Coherence, strategic measures, conglomerable measures, finite additivity, prediction.

otherwise; for $x \in X$ and $D \subset Y$, $q(x)(D)$ is the price at which the predicter commits himself to buy or sell a ticket worth \$1 if $y \in D$, \$0 otherwise, should the firts quantity turn out to have value x.

Suppose the predictor contracts with a gambler to buy or sell tickets on (x, y) and on y after observing x at the prices determined by P and q. The gambler may sell or purchase partial shares so that a bet of c on $A \subset X \times Y$ would yield a net return of

$$\phi(x, y) = c[A(x, y) - P(A)] \tag{3}$$

and a bet of d on $Bx \subset Y$ conditional on $x \in S \subset X$ would yield

$$\psi(x, y) = dS(x)[B(x, y) - q(x)(Bx)]. \tag{4}$$

(Here A, B, and S are identified with their indicator functions and $B = \{(x, y): y \in Bx\}$.) The numbers c and d can be can be positive or negative, corresponding to a sale or a purchase, respectively.

The gambler is permitted a finite number of such transactions. The predicter is called coherent if the gambler cannot attain a uniformly positive return.

Here is a more formal statement.

DEFINITION. The pair (P, q) is *coherent* if there do not exist functions $\phi_1, \cdots, \phi_m$, of the form (3) and functions $\psi_1, \cdots, \psi_n$ of the form (4) such that

$$\inf\{\sum \phi_i + \sum \psi_j\} > 0. \tag{5}$$

This notion of coherence is quite similar to that of de Finetti [2, 3] in that incoherence corresponds to a *sure* loss under some betting system. Both definitions permit only a finite number of operative bets, i.e., bets that result in actual financial transactions. However, the formulation here does allow for a possibly infinite number of potential bets when the set X is infinite; only those bets corresponding to the actual observed x become operative. The formulation of de Finetti seems to allow only a finite number of potential bets.

Another formulation of coherence for statistical models was introduced by Cornfield [1] and Freedman and Purves [6] and was studied in Heath and Sudderth [8] and Lane and Sudderth [10]. In that formulation, incoherence corresponds to an *expected* loss under some betting system which is uniformly positive when the expectation is taken under the possible values for the parameter or state of nature. As is explained in [10], this definition based on expected loss naturally reduces to the one here when no statistical model is posited.

Here is a characterization of our notion of coherence.

THEOREM 1. *The pair (P, q) is coherent if and only if P is a finitely additive probability measure and formula* (1) *holds for every $A \subset X \times Y$.*

The proof is in Section 2.

The theorem has several predecessors. In the case when X and Y are finite, it reduces to de Finetti's result for finite partitions. It is a close cousin of Dubins' result that the conglomerative measures with respect to a partition are the strategic measures (Theorem 1 of [4]). Dubins considers bounded functions while we consider sets, and he requires that the functions $q(x)$ be probability measures. In addition, Dubins does not treat the notion of coherence directly as is done here. Another related result was recently proved by Goldstein [7]. He discusses "posterior previsions" which may be based on an observation D for which there is no fixed set of possible outcomes. (Some mathematicians may find his definitions vague.) Goldstein's theorem is based on a quadratic loss function as in de Finetti's book [3]. Our definition is closer to de Finetti's earlier paper [2] and our proof is somewhat more direct: a violation of (1) results in a sure win for the gambler based on a system of just three transactions. A somewhat more general result appears in Lane and Sudderth [10], but the proof is nonconstructive using a separating hyperplane argument.

There is a titillating consequence of Theorem 1 (or a mild generalization of it stated for partitions) in a result of Schervish, Seidenfeld and Kadane [11]. (See also Hill and Lane [9].) These authors have proved the fundamental result that, for a finitely additive probability measure which is not countably additive, there is a countable partition of the underlying space on which the measure fails to be conglomerable. That is, there is a partition such that no conditional distribution exists satisfying the law of total probability. But Theorem 1 implies that coherent predictions must be conglomerable. Hence, coherent prediction with respect to every countable partition requires countable additivity.

There are two objections to this line of argument. The first is that it is unreasonable to require a predicter to make predictions which are not based on natural partitions such as that induced by an initial observation x. The second is that, if all countable partitions are allowed, there is no reason not to allow uncountable ones as well. But then even conventional countably additive measures are eliminated. (See, for example, Dubins [4].)

2. The proof of Theorem 1. If P is finitely additive and (P, q) satisfies (1) for every A, it is easily seen that

$$\int \phi \, dP = \int \psi \, dp = 0$$

for every ϕ as in (3) and every ψ as in (4). Hence, (5) is impossible.

Now assume (P, q) is coherent. P is then finitely additive by a result of de Finetti [2]. It remains to verify (1).

Suppose (1) fails for some $A \subset X \times Y$. To be specific, assume

$$P(A) < \int q(x)(Ax)P_0(dx). \tag{6}$$

(The argument would be similar if the opposite inequality were assumed to hold.) We will construct sets satisfying (5). The construction uses the following lemma.

LEMMA. *There is a set $S \subset X$ such that*

$$P(AS) < \{\inf_{x \in S} q(x)(Ax)\} P_0(S)$$

where

$$AS = \{(x, y): (x, y) \in A,\ x \in S\}.$$

PROOF. Let

$$\delta = \int q(x)(Ax)P_0(dx) - P(A).$$

For $n = 1, 2, \cdots$ and $i = 1, 2, \cdots, n + 1$, let

$$S_{in} = \{x: (i-1)/n \le q(x)(Ax) < (i/n)\}.$$

For each n, the sets $S_{1n}, \cdots, S_{n+1,n}$ form a partition of X. Hence,

$$\begin{aligned}\textstyle\sum_i P(AS_{in}) = P(A) &= \int q(x)(Ax)P_0(dx) - \delta \\ &= \textstyle\sum_i \int_{S_{in}} q(x)(Ax)P_0(dx) - \delta \\ &\le \textstyle\sum_i \frac{i}{n} P_0(S_{in}) - \delta \\ &= \textstyle\sum_i \frac{i-1}{n} P_0(S_{in}) + \frac{1}{n} - \delta.\end{aligned}$$

For n sufficiently large, $(1/n) - \delta < 0$. Hence, one of the terms on the left must be smaller than the corresponding term on the right. That is, for some i,

$$P(AS_{in}) < ((i-1)/n)P_0(S_{in}).$$

Take $S = S_{in}$. □

Now let

$$q_* = \inf_{x \in S} q(x)(Ax)$$

and consider the following function

$$\lambda(x, y) = [(AS)(x, y) - P(AS)] - q_*[S(x) - P_0(S)] - S(x)[A(x, y) - q(x)(Ax)].$$

This function is of the form considered in the definition of coherence; just take $A_1 = AS$, $A_2 = S \times Y$ and $B_1 = A$, $S_1 = S$. We will have a contradiction once it is verified that $\inf \lambda > 0$. For the verification, consider three cases.

CASE 1. $(x, y) \in AS$. In this case,

$$\begin{aligned}\lambda(x, y) &= 1 - P(AS) - q_*(1 - P_0(S)) - 1 + q(x)(Ax) \\ &= (q_* P_0(S) - P(AS)] + [q(x)(Ax) - q_*] \\ &\ge q_* P_0(S) - P(AS) > 0.\end{aligned}$$

CASE 2. $(x, y) \in A^c S$. Similar to case 1.

CASE 3. $x \in S^c$. In this case,

$$\lambda(x, y) = -P(AS) + q_* P_0(S) > 0.$$

This completes the proof. □

For simplicity, Theorem 1 was presented without measurability assumptions. However, the theorem and its proof remain true under conventional measurability conditions. Likewise the theorem could be proved for general partitions as in Dubins [4]. Finally, the proof can be modified to cover the case in which predictions are made separately (not jointly) for x and y, as well as for y given x.

REFERENCES

[1] CORNFIELD, J. (1969). The Bayesian outlook and its applications. *Biometrics* **25** 617–658.
[2] DE FINETTI, B. (1937). La Prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7** 1–68. (An English translation appears in *Stud. Subjective Probab.* (Kyberg, S. ed.) Wiley, New York.)
[3] DE FINETTI, B. (1975). *Theory of Probability* 1. Wiley, New York.
[4] DUBINS, L. E. (1975). Finitely additive conditional probabilities, conglomerability, and disintegrations. *Ann. Probab.* **3** 89–99.
[5] DUBINS, L. E. (1977). Measurable, tail disintegrations of the Haar integral are purely finitely additive. *Proc. Amer. Math. Soc.* **62** 34–36.
[6] FREEDMAN, D. and PURVES, R. (1969). Bayes method for bookies. *Ann. Math. Statist.* **40** 1177–1186.
[7] GOLDSTEIN, M. (1983). The prevision of a prevision. *J. Amer. Statist. Assoc.* **78** 817–819.
[8] HEATH, D. AND SUDDERTH, W. (1978). On finitely additive priors, coherence, and extended admissibility. *Ann. Statist.* **6** 333–345.
[9] HILL, B. AND LANE, D. (1985). Conglomerability and countable additivity. In *Bayesian Inference and Decision Techniques*: Essays in Honor of Bruno de Finetti (P. K. Goel and A. Zellner, editors). North-Holland, Amsterdam.
[10] LANE, D. AND SUDDERTH, W. (1984). Coherent predictive inference. *Sankhyā*, Series A, **46** 166–185.
[11] SCHERVISH, M. SEIDENFELD, T. AND KADANE J. (1984). The extent of nonconglomerability of finitely additive probabilities. *Z. Wahrsch. verw. Gebiete* **66** 205–226.

SCHOOL OF STATISTICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA 55455