Coherent Predictions are Strategic

by

David A. Lane and William D. Sudderth[1]
University of Minnesota
School of Statistics

Technical Report No. 429

Abstract

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 x 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

$$(1) \quad P(A) = \int q(x)(Ax)P_o(dx)$$

where P is a probability distribution on the product space $X \times Y$, P_o 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. In the conventional theory, q is required by definition to satisfy

$$(2) \quad P(C \times D) = \int_C q(x)(D)P_o(dx)$$

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 $P_o\{x\} = 0$ 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 which 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 [3], 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 otherwise; for $x \in X$ and $B \subset Y$, q(x)(B) is the price at which, the predicter commits himself to buy or sell a ticket worth \$1 if $y \in B$, \$0 otherwise, should the first quantity turn out to have value x.

Suppose the predicter contracts with a gambler to buy or sell tickets

-1-

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

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

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

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

(Here A and B 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.

<u>Definition</u>. The pair (P,q) is <u>coherent</u> if there do not exist functions $\phi_1, \ldots, \phi_m$, of the form (3) and functions $\psi_1, \ldots, \psi_n$ of the form (4) such that
(5) inf $\{ \Sigma \phi_i + \Sigma \psi_j \} > 0$.

If (P,q) is coherent in this sense, then P is certainly coherent in the sense of de Finetti [3] and, by his famous result, must be a finitely additive probability measure.

<u>THEOREM 1</u>. 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 [1]). Dubins consi-
ders 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 similar
result was recently proved by Goldstein [5]. 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 [4]. Our definition is closer to de Finetti's earlier paper
[3] 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 [6], but the proof
is non-constructive using a separating hyperplane argument.

There is a titillating consequence of Theorem 1 (or a mild generalization
of it stated for partitions) and a result of Seidenfeld, Shervish, and Kadane
[8]. (See also Hill and Lane [6].) 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 [2].)

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 additively by a result of de Finetti [3]. It remains to verify (1).

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

$$(6) \quad P(A) < \int q(x)(Ax)P_o(dx).$$

(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

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

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

Proof: Let

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

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

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

-4-

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

$$\sum_i P(AS_{in}) = P(A)$$

$$= \int q(x)(Ax)P_o(dx) - \delta$$

$$= \sum_i \int_{S_{in}} q(x)(Ax)P_o(dx) - \delta$$

$$\leq \sum_i \frac{i}{n} P_o(S_{in}) - \delta$$

$$= \sum_i \frac{i-1}{n} P_o(S_{in}) + \frac{1}{n} - \delta$$

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

$$P(AS_{in}) < \frac{i-1}{n} P_o(S_{in}).$$

Take $S = S_{in}$.   □

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

and consider the following function

$$\lambda(x,y) = [(AS)(x,y) - P(AS)] - q_*[S(x) - P_o(S)]$$

$$- S(x)[A(x,y) - q(x)(Ax)]$$

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.

1. $(x,y) \in AS$.

In this case,

$$\lambda(x,y) = 1 - P(AS) - q_*(1 - P_o(S)) - 1 + q(x)(Ax)$$

$$= [q_* P_o(S) - P(AS)] + [q(x)(Ax) - q_*]$$

$$> q_* P_o(S) - P(AS)$$

$$> 0.$$

2. $(x,y) \in A^c S$

Similar to case 1

3. $x \in S^c$

In this case,

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

This completes the proof.

Remarks.

1. This proof of Theorem 1 was suggested by our study of
de Finetti's proof in [3] that coherence implies

$$P(AB) = P(A)P(B/A).$$

2. For simplicity, Theorem 1 was presented without measura-
bility 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
[1]. 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.

3. Note that we only allow a finite number of operative bets,
but each function ψ (as in (4)) specifies a possibly infinite
number of potential bets - only the one determined by the realized
x becomes operative. Thus our result only applies to the situ-
ation in which one can specify in advance that at the time at
which x is observed, no other information will become available
relevant to y.

References

1. Dubins, Lester E. (1975). Finitely additive conditional probabilities, conglomerability, and disintegrations. <u>Annals of Probability</u> 3 89-99.

2. Dubins, Lester E. (1977). Measurable, tail disintegrations of the Haar integral are purely finitely additive. <u>Proceedings of the American Mathematical Society</u> 62 34-36.

3. de Finetti, Bruno (1937). La Prevision: ses lois logiques, ses sources subjectives, Annals de l'Institut Henri Poincare´ 7 1-68. (An English translation appears in <u>Studies in Subjective Probability</u> ed. Kyberg, Smokler. Wiley, New York.)

4. de Finetti, Bruno (1975). <u>Theory of Probability</u> 1. Wiley, New York.

5. Goldstein, Michael (1983). The prevision of a prevision. <u>Journal of the American Statistical Association</u>. 78 817-819.

6. Hill, Bruce and Lane, David (1983). Conglomerability and countable additivity. (to appear).

7. Lane, David and Sudderth, William (1983). Coherent predictive inference. <u>Sankhya</u>. (to appear).

8. Schervish, M. Seidenfeld, T. and Kadane J. (1981). The extent of non-conglomerability of finitely additive probabilities. Carnegie-Mellon University Report.