DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A COMMENT ON THE SUBJECTIVIST POSITION

Asad Zaman
California Institute of Technology
and University of Pennsylvania

SOCIAL SCIENCE WORKING PAPER 362

November 1980

ABSTRACT

It is shown that priors obtained by making coherent choices over lotteries are arbitrary, since they depend on the order in which the choices are made.

A COMMENT ON THE SUBJECTIVIST POSITION

Asad Zaman

An extremist Bayesian viewpoint, one which I have held in the past, maintains that in order to be able to make decision in a "reasonable" way, it is necessary to have a prior over all subsets of the space of states of nature. I shall show that this is not the case by giving a procedure which a frequentist may use to make "reasonable" decisions in presence of uncertainty. First I shall describe the Bayesian argument.

Let Ω be the space of states of nature, and $L_B = \{f : \Omega \to R, \sup_w f(w) < \infty\}$. L_B is a Banach space with norm $\|f\| = \sup_w f(w)$. Let L_B^* be the set of continuous linear functionals over L_B; this is identified with the set of finitely additive measures on Ω (equipped with the σ-field of all subsets). Every element $\mu \in L_B^*$ induces a (complete) ordering of L_B via the definition $f \underset{\mu}{>} g$ (f is μ-preferred to g) if $\mu(f) \geq \mu(g)$. For any partial order $>$ on L_B, define equivalence classes $\pi(f) = \{g \in L_B | \ f > g \text{ and } g > f\}$. For the preference $\underset{\mu}{>}$ we must have $\pi(0)$ is closed and $\pi(f) = f + \pi(0)$ for all $f \in L_B$. Furthermore, $\pi(0)$ must be of codimension 1; $\pi(0) = \mathrm{Ker}(\nu)$ if and only if $\nu = \alpha\mu$ for some α. What may be termed the Fundamental Theorem of Subjective Probability is the converse of this result.

<u>The Fundamental Theorem of Subjective Probability</u>: Suppose a partial order $>$ on L_B can be extended to a complete order on some closed subspace S of L_B, containing the constant functions, and the equivalence classes $\pi(s)$ of the extension satisfy the conditions

(i) $\pi(0)$ is closed in S (and therefore in L_B),

(ii) for any $f \in S$, $\pi(f) = f + \pi(0)$

(iii) $\pi(0)$ is a subspace of co-dimension 1 in S.

Furthermore,

(iv) Nonnegative functions are preferred to 0.

Then there exists a measure $\mu \in L_B^*$ such that $\underset{\mu}{>}$ coincides with the extension of $>$ on S.

<u>Proof</u>: This follows easily from Theorem 1.41, p. 29 Rudin [1973], according to which the quotient map $\pi : S \to S/\pi(0)$ is linear and continuous (by (i) and (ii)). Since $\pi(0)$ is of codimension 1, there exists a continuous linear bijection $\ell : S/\pi(0) \to R$. $\ell \circ \pi$ is then a continuous linear functional on S and it is easy to check that it induces the same order on S as $>$. $\ell \circ \pi$ must be a positive linear funtional by (iv). Since S contains at least one constant positive function, $\ell \circ \pi$ can be extended to a positive linear functional $\mu \in L_B^*$, by Proposition 6, Chapter II §3 No. 4, Bourbaki [1953]. Clearly the preference induced by μ is the same as that induced by $\alpha\mu$ for any $\alpha > 0$ so we can take μ to be of mass 1 without loss of generality. ∎

It is easy to check that if S is the set of all functions measurable with respect to some σ-field, S is closed and contains a

constant function. In Savage's [1954] treatment of this theorem, S is taken to be L_B. Assumption (ii) is referred to as the sure-thing principle. Assumption (i) and (iii) follow from an assumption about continuity of the ordering, which can be stated as follows: if $g < h$ and $g \notin \pi(h)$, there exist open neighborhoods N_h of h and N_g of g such that $x \in N_g$ $y \in N_h$ imply $x < y$, and $x \notin \pi(y)$. Similarly, other systems of axiomatization must be equivalent to the one given above, or stronger, since the conditions above are necessary and sufficient.

The assumption of the theorem above are said to describe "coherent" behavior. If $S = L_B$ the assumption (iii) implies that the probability measure μ is unique. Savage [1954] regards it as desirable on normative grounds to make coherent choices. Since, at least in principle, it is possible for anyone to evaluate introspectively choices one would make over lotteries, and thereby acquire a prior, it is a reasonable normative assertion that everyone should act as if he had a prior.

Since the Emersonian objection to this argument that "Coherency is the hobgoblin of little minds" is regarded as "unfair" or "unscientific," I will give a frequentist procedure for decision-making which is coherent but not Bayesian.

The frequentist position can be described as follows. Some subsets of Ω have been observed to occur (or can be deduced to occur) with stable relative frequencies over a sequence of trials. Let B_0 be the σ-algebra of such events, and for $B \in B_0$ let $\mu_0(B)$ be the frequency of occurence of B. (The interpretation of μ_0 and B_0 is unimportant for what follows.) The frequentist asserts that all decision must be based on the measure μ_0. If f, g ε L_B are B_0-measurable than the frequentist will choose f over g if $\int f d\mu_0 > \int g d\mu_0$. Quite often B_0 is the trivial σ-field so that only constant function are measurable. The response to the Bayesian question "How does a frequentist chose among non-measurable lotteries" is that he would rather not do so. However, if forced to choose he proceeds as follows. If $f \geq g$ with strict inequality for at least one $w \in \Omega$, then f is chosen over g. Otherwise, construct the smallest σ-field B_1 such that $B_0 \subseteq B_1$ and both f and g are measurable. Let M_0, M_1 be the subspaces of B_0 and B_1-measurable functions of L. There exists a positive linear functional separating f and g which is an extension of μ_0 to M_1. Normalize it to a probability measure and call it μ_1. Proceed to make subsequent choices by the same procedure. The Bayesian argument seems to be that one can go through such a procedure introspectively and come up with a unique prior at the end, which can then be used for purpose of inference. The frequentist response is contained in the following.

Theorem 2: Let B', μ' be an extension of B_0, μ_0, so that $B_0 \subseteq B'$ and μ' is a finitely additive positive measure which coincides with μ_0 on B_0. There exists a well-ordered set T with minimal element 0, and for each $t \in T$, a pair of lotteries (f^t, g^t), a σ-field A_t with measure ν_t satisfying

(i) $A_0 = B_0 \quad \nu_0 = \mu_0$

(ii) Fix s; then for all $t < s$, f^t and g^t are measurable with respect to B_s, and $\int g^t d\nu_s \geq \int f^t d\nu_s$, and B_s and ν_s are extensions of B_t and ν_t for all t.

(iii) For all $t \in T$, f^t, g^t are B' measurable and $\int g^t d\mu' \geq \int f^t d\mu'$.

Discussion: The theorem states that by following the rule "pick the second element in a lottery when the two are not comparable" and picking the right sequence of lotteries, one can arrive at any extension of the measure μ_0. Thus the process of introspection does not lead to a determinate subjective probability; one might arrive at any probability measure depending on the sequence in which the choices are considered. (It might be of some interest to characterize the possible extensions of a measure μ_0 for a statistician following some fixed rule such as minimax, etc., for choosing among admissiable rules.) It is clear that the probability arrived at in this manner is not a reflection of the beliefs of the frequentist, nor does it contain information about the world, and hence cannot have any value in making inferences. When looked at in this light, the Bayesian Frequentist dispute revolves around how one should choose B_0 and μ_0. The Bayesian typically insist the B_0 should contain a large collection of events and the frequestists hold the B_0 is typically a small set. The resolution of this dispute is beyond the scope of theory.

Proof of Theorem 2: This is straight forward. Let $W : T \to B' \setminus B_0$ be a well-ordering of sets belonging to B' but not B_0. Define $g^t = \chi_{W(t)}(\omega)$

if $\int_{W(t)} \mu'(d\omega) \geq \int_{W(t)^c} \mu'(d\omega)$ and $g^t = \chi_{W(t)^c}$ otherwise (where χ_A is the characteristic function of A). Define $f^t = 1-g^t$. Fix $s \in T$ and assume A_t and γ_t have been defined for all $t < s$ to satisfy (i), (ii) and (iii). Let $A_s' = \bigcup_{t<s} A_t$. Let A_s be the smallest measure which makes f^t and g^t measurable, and includes A_s', and define γ_s to be the restriction of μ' to A_s. It is easy to verify that all the induction hypotheses are satisfied. This completes the proof. ■

Conclusions

We have seen that the argument that one can arrive at a prior distribution by introspective choice over lotteries is true but misleading since the prior chosen through such a process will in general depend on the order in which these choices are made. Another way of phrasing the argument is as follows. If one starts with a completely defined coherent prior over all subsets of the parameter space, then one can of course make coherent choices over lotteries. In this case there is no objection to the subjectivist position. However, if one has only vaguely defined prior beliefs, these cannot be extended to a completely specified prior in a unique way simply by the requirement of coherence.

It should be noted that a similar argument can be given to show that economic agents can make consistent choices without having completely defined utility functions.

REFERENCES

Anscombe, F. J. and Aumann, R. J. (1963) "A Definition of Subjective Probability." Annals of Mathematical Statistics 34:199-205.

Bourbaki, Nicholas (1953) Espaces Vectoriel Topologique. Elements de Mathematique, Livre V. Paris: Herman and CIE.

De Finetti, Bruno (1970) Teoria Delle Probabilita, Sintesi Introduttiva Con Appendice Critica. Torina: Guilio Einaudi.

Freedman, David A. and Purves, Roger A. (1979) "Bayes' Method for Bookies." Annals of Mathematical Statistics 40:1177-1186.

Lindley, D. V. (1972) Bayesian Statistics: A Review. Philadelphia: Siam.

Loeve, Michel, (1955) Probability Theory. New York: Van Nostrand Co, Inc.

Pratt, John W.; Raiffa, Howard; and Schlaifer, Robert. (1964) "The Foundations of Decision Under Uncertainty: An Elementary Exposition." Journal of the American Statistical Association 59:353-375.

Ramsey, Frank Plumpton. (1964) "Truth and Probability," Studies in Subjective Probability, edited by Henry E. Kyburg Jr. and Howard E. Smokler. New York: John Wiley, pp. 61-92; see also The Foundations of Mathematics and Other Essays. London: Kegan, Paul, Trench, Trubnet & Co., Ltd., 1931, pp. 159-198.

Rudin, Walter, (1973) Functional Analysis. New York: McGraw Hill.

Savage, Leonard J. (1954) The Foundations of Statistics. New York: John Wiley, pp. xv and 294.

Wald, Abraham. (1950) Statistical Decision Functions. New York: John Wiley, pp. ix and 179.