Mathematical Social Sciences 25 (1992) 95-98
North-Holland

Note

An elementary proof of Blackwell's theorem

Moshe Leshno and Yishay Spector

School of Business Administration, The Hebrew University, Jerusalem 91 905, Israel

Communicated by D. Samet
Received 25 April 1991
Revised 16 March 1992

This paper presents a short and elementary proof of Blackwell's theorem. This theorem states the statistical conditions under which one information structure (or experiment) is *more informative* than another information structure. By *more informative* we mean that one information structure does not have less economic value to any decision-maker than another information structure, regardless of the payoff function, the decision-maker's utility function, and the a priori probability distribution.

Key words: Experiments; ordering of experiments; information structure; Blackwell's theorem.

1. Introduction

In 1951 Blackwell established the equivalence of two partial orderings on the set of information structures. The term 'information structure' is based on a probabilistic concept and on the economic criterion of the value of information to a decision-maker (Blackwell, 1951, 1953, Marschak and Miyasawa, 1968, Marschak, 1971). The original proof of Blackwell's theorem is long and complicated (see Blackwell, 1951, 1953, and Marschak and Miyasawa, 1968). In this paper we present a more compact and simpler proof of the theorem. A simpler proof was given also by Cremer (1982) and Ponssard (1975). We present here a proof of Blackwell's theorem which is shorter than all of the above-mentioned proofs.

2. The basic model of the decision problem

The basic model of the decision problem is as follows (for further detail see, for

Correspondence to: Y. Spector, School of Business Administration, The Hebrew University, Jerusalem 91 905, Israel.

0165-4896/92/$05.00

example, Marschak and Miyasawa, 1968): let S be a finite set of states of nature, $S=\{s_1,\dots,s_{n_s}\}$. Let p be a vector of a priori probabilities associated with the states in S, $p^t=(p_1,\dots,p_{n_s})$, where $\sum_{i=1}^{n_s} p_i=1$, $p_i \geq 0$ for $i=1,\dots,n_s$ (the superscript t stands for the transpose operator).

Let Y be a finite set of signals, $Y=\{y_1,\dots,y_{n_y}\}$. An information structure Q is defined as a Markovian (stochastic) matrix of conditional probabilities (dimension $n_s \times n_y$) in which signals of the set Y will be displayed at the occurrence of a state s. Thus Q_{ij} is the probability that for a given state s_i, signal y_j will be displayed. Let A be a finite set of actions that can be taken by the decision-maker (DM), $A=\{a_1,\dots,a_{n_a}\}$. A cardinal payoff function $U: A\times S \to R$ associates payoffs to each action and state pair. The function U can be depicted by an $(n_a \times n_s)$ matrix, denoted U, the element $u_{ij}=U(a_i,s_j)$ of which is the payoff gained when an action a_i is taken and the state turns out to be s_j.

The DM can only observe the signals, not the events, and chooses actions accordingly. The DM's strategy is delineated by a $(n_s \times n_a)$ Markov matrix D, the element D_{ij} of which determines the probability that the DM takes action a_j on observing signal y_i. Let us assume that the DM wishes to optimize D to obtain the maximum expected payoff. Let $\hat{\pi}$ be a square matrix containing the elements of p in its main diagonal and zero elsewhere. The expected payoff gained from Q, U, and a decision rule D is given by $\mathrm{tr}(QDU\hat{\pi})$, where 'tr' represents the trace operator. Maximization of $\mathrm{tr}(QDU\hat{\pi})$ is obtained by solving a linear programming problem for the elements of D constrained by the properties of a Markovian matrix (McGuire, 1986). Denote $F(Q,U,p)=\max_D \mathrm{tr}(QDU\hat{\pi})$. Let Q and P be two information structures operating on the same set of events S. From the economic point of view Q is defined to be generally more informative than P ($Q \supseteq P$) if the maximal expected payoff yielded by P is not larger than that yielded by Q for all payoff matrices U and all probability vectors p. Formally:

Definition 1. Q is more informative than P ($Q \supseteq P$) if $F(Q,U,p) \geq F(P,U,p)$ for all U and any p.

Blackwell's theorem. *Q is more informative than P ($Q \supseteq P$) if and only if there exists a Markov matrix[1] M with appropriate dimension such that $QM=P$.*

3. A new proof of Blackwell's theorem

We use in the proof the following notation and propositions. For any matrices A and B we denote the inner product by (A,B) i.e.

$$(A,B)=\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij}.$$

[1] M is called the garbling matrix.

Proposition 1. *For any matrices* A, B *and* C *we have*: $(AB, C) = (B, A^{t}C) = (A, CB^{t})$.

Proof.

$$(AB, C) = \sum_{i,j} (AB)_{ij} C_{ij} = \sum_{i,j,k} A_{ik} B_{kj} C_{ij} = \sum_{k,j} B_{kj} \sum_{i} A^{t}_{ki} C_{ij} = (B, A^{t}C),$$

and by the first part,

$$(AB, C) = (B^{t}, A^{t}, C^{t}) = (A^{t}, BC^{t}) = (A, CB^{t}). \qquad \square$$

Proposition 2.

$$\mathrm{tr}(PDU\hat{\pi}) = (PDU, \hat{\pi}).$$

Proof.

$$\mathrm{tr}(PDU\hat{\pi}) = \sum_{i} (PDU)_{ii} \hat{\pi}_{ii} = \sum_{i,j} (PDU)_{ij} \hat{\pi}_{ij} = (PDU, \hat{\pi}). \qquad \square$$

The proof of Blackwell's theorem. Let P and Q be information structures with a common finite state of the world set S, and finite sets of signals Y and Z, respectively. Without loss of generality we can assume that the signals sets Y and Z are the same, and that all the matrices are quadratic.

(a) If $Q = PM$, then it is trivial that $Q \subseteq P$, because $\mathrm{tr}(QDU\hat{\pi}) = \mathrm{tr}(PMDU\hat{\pi})$. Since the matrix MD is also Markovian, we have

$$\max_{D} \mathrm{tr}(QDU\hat{\pi}) = \max_{D} \mathrm{tr}(PMDU\hat{\pi}) \leq \max_{D} \mathrm{tr}(PDU\hat{\pi}),$$

i.e. $Q \subseteq P$.

(b) Suppose that for every Markov matrix D, $Q \neq PD$. Then $Q \notin \mathscr{C}$, where $\mathscr{C} = \{A \mid \exists D \text{ Markov}, A = PD\}$. But $\mathscr{C}$ is a convex and closed set in $R^{n \times n}$, and therefore, by the separation theorem (see, for example, Rudin, 1973) there exists a matrix $\hat{U}$ such that for every $A \in \mathscr{C}$:

$$(A, \hat{U}) < (Q, \hat{U}),$$

or for every Markov matrix D:

$$(PD, \hat{U}) < (Q, \hat{U}).$$

Let the payoff matrix U be

$$U^{t} = \hat{\pi}^{-1} \hat{U},$$

then by Propositions 1 and 2 for every Markov matrix D we have

$$\mathrm{tr}(PDU\hat{\pi}) < \mathrm{tr}(QU\hat{\pi}).$$

Therefore

$$\max_{D} \mathrm{tr}(PDU\hat{\pi}) < \mathrm{tr}(QIU\hat{\pi}) \leq \max_{D} \mathrm{tr}(QDU\hat{\pi}),$$

that is $Q \not\subseteq P$. $\square$

This completes the proof of Blackwell's theorem.

References

D. Blackwell, The comparison of experiments, in: J. Neyman, ed., Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (Univ. of California Press, Berkeley, 1951), 93-102.

D. Blackwell, Equivalent comparison of experiments, Ann. Math. Statist. 24 (1953) 265-272.

J. Cremer, A simple proof of Blackwell's comparison of experiments theorem, J. Econ. Theory 27 (1982) 443-493.

J. Marschak, Economics of information systems, J. Amer. Statist. Assoc. 66 (1971) 192-291.

J. Marschak and K. Miyasawa, Economic comparability of information systems, Int. Econ. Rev. 9(2) (1968) 137-174.

B.C. McGuire, Comparison of information structures, in: C.B. McGuire and R. Rander, eds., Decision and Organization (Univ. of Minnesota Press, Minneapolis, 1986).

P.J. Ponssard, A note on information value theory for experiments defined in extensive form, Management Sci. 22(4) (1975) 449-454.

W. Rudin, Functional Analysis (McGraw-Hill, New York, 1973).