AD-A052 829 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF STATISTICS F/G 12/1

OPTIMAL CHALLENGES FOR SELECTION.(U)

JAN 78 M H DEGROOT, J B KADANE N00014-75-C-0516

UNCLASSIFIED TR-143 NL

1 OF 1

ADA

052829

END

DATE

FILMED

5-78

DDC

NR-042-309

AD A052829

AD No.
DDC FILE COPY

DEPARTMENT

OF

STATISTICS

Carnegie-Mellon University

PITTSBURGH, PENNSYLVANIA 15213

(1)

(6) OPTIMAL CHALLENGES FOR SELECTION*

by

(10) Morris H. DeGroot and

Joseph B. Kadane

Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

(12) 30p.

(11) January 1978

(9) Technical Report, No. 143

(14) TR-143

* This research was supported in part by the National Science Foundation under Grant SOC77-07548 and in part by the Office of Naval Research under Contract N00014-75-C-0516.

(15) NSF-SOC77-07548

Abstract

This paper generalizes the problems of optimal selection considered by the authors in an earlier report Roth, Kadane and DeGroot (1977) by allowing a set of J items to be chosen by two decision makers, the first of whom has A challenges and the second has B challenges. The two decision-makers each have an opportunity to challenge each item before it is accepted, in some fixed order we leave arbitrary. We assume that the decision-makers know the utility function of the other side as well as their own over sets of J items, and that they know the subjective distribution assigned by the other side of characteristics of potential items that will be observed, as well as their own. Under these conditions the other side's response to each potential time can be predicted with certainty, and backward induction defines an optimal strategy. The authors We study an important special case we call regular, and show that it is never disadvantageous to go first in the regular case. The use of peremptory challenges in jury trials motivates our model.

Key words and phrases:

Challenges; Househunting Problem; Juries; Optimal Jury Selection; Peremptory Challenges; Secretary Problem; Sequential Decisionmaking.

American Mathematical Society 1970 Subject Classifications
Primary 49C99; Secondary 90D35.

1. Introduction

In this paper we consider a legal case which is to be tried by a jury, and shall study optimal strategies to be followed by the lawyers for the prosecution (or plaintiff) and the lawyers for the defense in their use of the peremptory challenges that are available to them. We shall begin by describing our model of the process by which the jury is chosen.

We assume that jurors are to be chosen from a venire or panel of prospective jurors that have been selected for possible jury duty from the community in which the trial is being held. Prospective jurors are examined by the judge one at a time to make sure that they are qualified and then asked questions which could result in their being dismissed for cause. If a prospective juror is not dismissed for cause, the lawyers for either side then have the opportunity to have him dismissed by exercising a peremptory challenge. If neither side does so, he becomes a member of the jury.

We shall assume throughout this paper that all the prospective jurors who can be dismissed for cause have been eliminated from the panel, and shall restrict our attention to the population of prospective jurors who cannot be dismissed for cause. Thus, prospective jurors from this population are interviewed sequentially; after each person is interviewed, the lawyers for each side must specify whether they wish to exercise a peremptory challenge or to accept the juror. If he is accepted, he joins the jury and cannot later be dismissed. If he is challenged, he is dismissed and cannot later be recalled.

We assume that for each prospective juror, first one side must state its decision, and then the other side. Thus, the two sides never make decisions simultaneously and the two sides will not both exercise peremptory challenges on the same prospective juror. The question of which side is to decide first for each prospective juror is left arbitrary in our mathematics; it need not be the same side for each prospective juror. However, we assume that the rule is specified in advance and does not depend in any way on how peremptory challenges have been used on previous prospective jurors.

We assume that J jurors are to be chosen for the jury. Of course, J will usually be 12, but it could be 6 or some other number. We shall shortly describe some interesting applications of our theory in which $J = 1$. We assume also that the prosecution is limited to A peremptory challenges and the defense to B, where A and B are given positive integers.

We assume that each prospective juror is represented by a vector X of some arbitrary dimension that summarizes all the information available to the prosecution and the defense in regard to that person's possible behavior as a juror. In brief, each prospective juror is characterized by a vector X that summarizes any demographic, physical, or behavioral variables that might be relevant to his performance as a juror.

Since J jurors are to be seated on the jury and a total of $A + B$ peremptory challenges are available to the two sides, the jury must be fully formed after at most $J + A + B$ prospective jurors have been interviewed. We assume that the lawyers for the prosecution and the lawyers for the defense can each represent their own view of the process generating the sequence of prospective jurors by assigning a joint probability distribution to $X_1, \ldots, X_{J+A+B}$. This distribution need not be the same for the two sides, but it is assumed that each side knows the distribution that is assigned by the other side. The following four models for these joint distributions are of particular interest:

(1) The random vectors $X_1, \ldots, X_{J+A+B}$ might be independent and identically distributed with a particular known distribution that is agreed on by both sides. This might be the case if the distribution of the vector X in the population of prospective jurors was known quite precisely to both sides through a careful survey or poll of the community, and all that was known to the lawyers about $X_1, \ldots, X_{J+A+B}$ was that they were drawn as a random sample from this community.

(2) Suppose that each side had conducted its own survey of the community and that the two surveys yielded different results. Then each side might be aware of the results of the survey conducted by the other side, but it might believe its own survey rather than the other one. In this case, each side might again assume that the random vectors $X_1, \ldots, X_{J+A+B}$ are independent and identically distributed, but the two sides would assign different distributions. Each side would believe that its own distribution was a more appropriate model than the distribution assigned by the other side.

(3) The exact sequence of values of $X_1, \ldots, X_{J+A+B}$ might be known in advance to both sides. This might be the case if the lawyers could interview or otherwise study the entire specific panel of prospective jurors from which the jury was to be chosen before beginning the sequential selection and challenge process.

(4) The random variables $X_1, \ldots, X_{J+A+B}$ might be exchangeable but dependent. This might be the case if the lawyers for both sides believe that $X_1, \ldots, X_{J+A+B}$ form a random sample from some distribution of prospective jurors that depends on a parameter θ whose value is unknown to both sides, and each side assigns a prior distribution to θ. The two sides might agree on the conditional distribution of X_i for any given value of θ, but they might assign different prior distributions to θ. In this case, it is assumed that each side is aware of the prior distribution assigned by the other side. The lawyers for each side will update their own distribution for θ after each prospective juror is interviewed, which in turn will affect their joint distribution for the remaining prospective jurors.

Of course, there are infinitely many other possibilities and our model can accommodate any specifications of the joint distribution of $X_1, \ldots, X_{J+A+B}$.

Now let $Y_1, \ldots, Y_J$ denote the X-vectors of characteristics for the J jurors actually chosen. Before the selection has actually been made, $Y_1, \ldots, Y_J$ will be random variables. For any possible values $y_1, \ldots, y_J$ of $Y_1, \ldots, Y_J$, we shall let $\psi_1(y_1, \ldots, y_J)$ denote the utility function of the prosecution and $\psi_2(y_1, \ldots, y_J)$ denote the utility function of the defense. Thus, we assume that the prosecution is trying to select a jury in order to maximize $E_1[\psi_1(Y_1, \ldots, Y_J)]$ and that the defense is trying to select a jury in order to maximize $E_2[\psi_2(Y_1, \ldots, Y_J)]$. Here, E_1 denotes expectation with respect to the joint distribution of $X_1, \ldots, X_{J+A+B}$ assigned by the prosecution and E_2 denotes expectation with respect to the distribution assigned by the defense.

The functions ψ_1 and ψ_2 could have various interpretations. For example, if both sides in a criminal trial were interested only in whether or not the juror voted for conviction, $\psi_1(y_1, \ldots, y_J)$ could be interpreted as the probability, in the view of the prosecution, that a jury comprising $y_1, \ldots, y_J$ would vote for conviction and $\psi_2(y_1, \ldots, y_J)$ could be interpreted as the probability, in the view of the defense, that the jury would not vote for conviction. In more general contexts, the functions ψ_1 and ψ_2 might take into the account the possibility that the jury does not reach any decision (i.e., that it becomes a hung jury), or the possibility that the jury might recommend a heavier or a lighter punishment or penalty.

Often the behavior of a jury depends only on the values of $y_1, \ldots, y_J$ and not on the order in which these J values were selected, so that the functions ψ_1 and ψ_2 can be taken to be symmetric. Sometimes, however, the first juror chosen, who has characteristics y_1, is automatically the foreman. Also, sometimes the last jurors chosen are alternates who participate in the jury decision only if one of the other jurors is incapacitated. Since the assumption of symmetry is not necessary for our mathematics and does not seem to simplify our analysis, we do not make that assumption. Regardless of whether or not some of the chosen jurors are to be alternates, it will be convenient throughout the paper to express our problems as though J jurors $Y_1, \ldots, Y_J$ are to be chosen.

Finally, it should be emphasized that although both sides observe the vector X for each prospective juror, the prosecution and defense may well disagree on what aspects of this vector are important or relevant. Thus, the utility function $\psi_1(y_1, \ldots, y_J)$ may depend on only certain components of each of the vectors y_i whereas $\psi_2(y_1, \ldots, y_J)$ may depend on only the remaining components of each y_i. On the other hand, the prosecution and defense may agree on the importance of the various components of each y_i, but may disagree on what these values imply about how the juror will behave. Thus, in general, the jury selection process will be a non-zero sum, two-player sequential game in which each side is trying to maximize its own expected utility.

We shall show that interesting results can be obtained in the important special case in which the prosecution and defense assign the same joint distribution to prospective jurors, have exactly the same perception of how each juror will behave, and also have diametrically opposed utilities for each possible outcome of the trial. In this case, the jury selection process will in effect be a zero-sum game; i.e., $\psi_1(y_1, \ldots, y_J) =$

$-\psi_2(y_1,\ldots, y_J)$ or, more generally, $a_1\psi_1(y_1,\ldots, y_J) + a_2\psi_2(y_1,\ldots, y_J) = c$ for constants a_1, a_2, and c such that $a_1a_2 > 0$ and all possible values of $y_1,\ldots, y_J$.

The next section describes what we mean by an optimal strategy and the relation of our problem to the previous literature. Section 3 gives our notation, and, in Theorem 1, the expected utility of going first and of going second. Section 4 discusses the important concept of regularity, and shows that if the two sides have the same opinion about the characteristics of the unseen jurors, or if there is only one juror to be selected, then the problem is regular. An example is given which is not regular. Finally Section 5 discusses the advantage of going first. In a regular problem it is never disadvantageous to go first. If the two sides have the same opinion about the characteristics of unseen jurors, and have either diametrically opposed or exactly coincident utilities, then the order in which they exercise challenges is irrelevant to both of them. Finally an example is given in which it is strictly advantageous to go second.

2. Nature of the Model, Optimal Strategies, and Relation to Previous Literature

Roth, Kadane, and DeGroot (1977) studied problems of the type we have described under the following special assumptions: Both the prosecution and the defense are interested only in the probability that the jury will vote for conviction. Each prospective juror is characterized by a vector $X = (p_1, p_2)$, where p_1 is the prosecution's assessment of the probability that the prospective juror will vote for conviction and p_2 is the defense's assessment of that probability. It is assumed that these two-dimensional vectors for successive prospective jurors are independent and identically distributed drawings from some known bivariate distribution that is agreed on by both sides. Furthermore, each side assumes that the probability that the final jury will vote for conviction is equal to the product of the individual probabilities that each member of the jury will vote for conviction. Finally, each side assigns utilities only to the possibilities of a conviction or no conviction.

In this paper, these assumptions are relaxed to permit a more sophisticated view of the jury selection and jury decision process. We allow an arbitrary representation of the relevant characteristics of each prospective juror, arbitrary probability distributions, and arbitrary utility functions over possible juries. There are, however, two kinds of restrictive assumptions that remain in this paper.

First, we are considering only one of a wide variety of different sequential processes by which peremptory challenges are exercised. In another system, prospective jurors are examined for qualifications and possible excuses for cause until a full jury is found. The two sides are then invited to challenge peremptorily, the jury is refilled, and this process continues until no more peremptory challenges are exercised either because the jury is satisfactory or because the dissatisfied side has run out of peremptory challenges. A second system, called the struck jury system, has the judge examine prospective jurors until $J+A+B$ of them have been found qualified and not been excused for cause. Then the two sides use their peremptory challenges in a fixed order. Sometimes in the struck jury system, each side is allowed to challenge only prospective jurors lower on the list than the last juror it challenged. Each of these systems creates its own problems of optimal behavior. Some preliminary work along these lines has been done by Brams and Davis (1976a, 1976b).

The second sense in which the assumptions of this paper are restrictive is that we assume that each side is fully aware of the information and beliefs of the other side. Thus, both sides observe each vector X, and each side knows the probability distributions and utility function that have been assigned by the other side. Since full information of this sort is a special case of uncertainty about what the opponent is trying to do, and since such uncertainty greatly complicates an already complicated situation, we chose to examine this full-information case first.

We shall now define what we mean by optimal procedures in this problem. Since a jury of J persons must have been found after at most $A+B+J$ people have been interviewed, the number of decisions in the selection process is bounded. Consider the last possible decision that could arise in this process, when one side has one challenge remaining, the other side has none, and only one juror remains to be seated. If the sequential decision process has not terminated before this state is reached, then the side with the one available challenge has a well-defined optimal decision for each possible prospective juror that might be interviewed.

Under the assumption that this last possible choice will be made optimally, the consequences of the next-to-last possible decision are known, and hence it can also be made optimally. By backward induction, each decision can be made optimally under the assumption that both sides will act optimally in all possible subsequent decisions. The optimal procedure is taken to be the one resulting from all these optimal actions of both sides. Thus we assume that juror selection is a non-cooperative two-person game. If the sides could collude, they might under some circumstances both improve their expected utilities, but we assume that they do not. Non-cooperative two-person games with alternating choices are also used in studies of duopoly (Cyert and DeGroot, 1970, 1973).

The optimal use of peremptory challenges in trials by jury is a generalization of the problems of optimal selection discussed by Gilbert and Mosteller (1966) and DeGroot (1970, Sec. 13.4), among others, and much earlier by Cayley (1875). In those problems, a sequential random sample of independent and identically distributed variables $Z_1, Z_2, \ldots$ is to be observed without recall by a single decision maker. The decision maker must choose a stopping rule N in order to maximize $E(Z_N)$. There is a given upper bound n on the number of observations that can be taken; if the decision maker has not stopped before the value of Z_n has been observed, then he must stop and accept that value as his payoff.

This problem of optimum selection is a special case of the jury selection problem in which only one juror is to be selected ($J = 1$), the observed vectors $X_1, X_2, \ldots$ are independent and identically distributed, only one side has any challenges available (the prosecution, say, has $n - 1$ challenges and the defense has none), and $Z_i = \psi_1(X_i)$ for $i=1,2,\ldots$.

Gilbert and Mosteller (1966) also consider the problem in which r observations must be selected out of the total of n variables to be observed sequentially, without recall, and the payoff to the decision maker is the sum of the r selected observations. This problem also is essentially a jury selection problem in which $J = r$, the prosecution has $n - r$ challenges and the defense has none, and the function ψ_1 has the special form $\psi_1(y_1,\ldots, y_r) = \Sigma_{i=1}^{r} \psi(y_i)$.

In summary, the problems of jury selection that we are studying can be described in the following general context: A group of J items with characteristics $y_1,\ldots, y_J$ is to be selected jointly by two persons. If $J = 1$, the item might be a house or a car to be selected jointly by a husband and wife, or an employee to be hired jointly by two executives in a particular firm. If $J \geq 2$, the items might be thought of as J persons that will form a team or committee, or a staff of J employees, to be selected jointly by the two decision makers. The interests and evaluations of the two decision makers need not coincide, and $\psi_1(y_1,\ldots, y_J)$ and $\psi_2(y_1,\ldots, y_J)$ represent the utilities to each of them of any possible selection of J items with characteristics $y_1,\ldots, y_J$. Each of the two decision makers can either accept or reject each item as it is observed, sequentially and without recall, and each has a fixed number of available challenges. The observations $X_1, X_2, \ldots$ have a specified joint distribution that is known to both decision makers, and each of them tries to maximize his own expected utility.

While we have limited our considerations in this paper to two sides, we believe that the problems, the point of view and our results generalize to more than two sides.

3. Notation and Preliminary Results

Suppose that k prospective jurors with characteristics $X_1, \ldots, X_k$ have been observed, some of whom may have been challenged and the others seated on the jury. The history comprising these values and the decisions that were made by each side about them is denoted h_k. The history h_k may affect the future decisions by the two sides in the following three ways: (i) If J' jurors have been seated then the first J' components of the utility functions ψ_1 and ψ_2 have been determined and the utility of remaining groups of jurors may be affected; (ii) The conditional joint distribution of X_{k+1}, $X_{k+2}, \ldots$ given $X_1, \ldots, X_k$ now becomes relevant for each side; (iii) The number of peremptory challenges remaining to each side and the number of jurors remaining to be seated may be reduced.

We require a notation for the expected utility to each side of the optimal strategies given the history h_k. This might be denoted $E_i(\psi_i | h_k)$, for $i=1,2$, but on some occasions it will be useful to emphasize particular aspects of the history h_k. Thus, for example, we write $E_i(\psi_i | h_k, a, b, j)$ to indicate that h_k includes the use of exactly $A - a$ peremptory challenges by the prosecution, leaving a; the use of exactly $B - b$ challenges by the defense, leaving b; and the seating of $J - j$ jurors, leaving j to be chosen.

Suppose now that some history h_k has already happened, and let $X_{k+1} = x$ be the characteristics of the next prospective juror observed. Let

$$F_1 = E_1(\psi_1 | h_k, x, a-1, b, j), \tag{3.1}$$

so that F_1 is the worth to the prosecution of the situation where j jurors are left to be chosen, the prosecution has a challenges, the defense has b challenges, and the prosecution uses one of its challenges. Similarly, let

$$G_1 = E_1(\psi_1 | h_k, x, a, b-1, j), \tag{3.2}$$

so that G_1 is the worth to the prosecution of the same situation except that the defense, instead of the prosecution, exercises the peremptory challenge. Finally, let

$$C_1 = E_1(\psi_1 | h_k, x, a, b, j-1), \tag{3.3}$$

so that C_1 is the worth of the above situation to the prosecution if the juror with characteristics x is not challenged by either side and, hence, is seated on the jury.

Similarly, let

$$F_2 = E_2(\psi_2 | h_k, x, a, b-1, j), \tag{3.4}$$

$$G_2 = E_2(\psi_2 | h_k, x, a-1, b, j), \tag{3.5}$$

$$C_2 = E_2(\psi_2 | h_k, x, a, b, j-1), \tag{3.6}$$

which are respectively the expected utilities to the defense if the defense challenges, the prosecution challenges, or the prospective juror is seated. Since one of these three outcomes must occur, the F's, G's, and C's fully describe the expected utility of the situation. We now describe the optimal strategies in these terms.

We impose the convention $F_1 = -\infty$ if $a = 0$ and $F_2 = -\infty$ if $b = 0$ so that neither side wishes to use a challenge it does not have. When $j = 0$ the process ends.

Suppose that the prosecution goes first. If the prosecution does not challenge the juror, then the defense will challenge the juror if $F_2 > C_2$, and will accept the juror if $F_2 < C_2$. A special problem occurs if $F_2 = C_2$, because in this case the defense could optimally take either action. In order not to burden our analysis and notation excessively, we will break ties by supposing that challenges will be exercised only when they are strictly necessary. Thus by assumption the defense will challenge the juror if $F_2 > C_2$, and will accept the juror otherwise. Therefore, the worth to the prosecution of the choice not to challenge is G_1 if $F_2 > C_2$ and C_1 otherwise. Consequently, when $F_2 > C_2$, the prosecution will challenge if $F_1 > G_1$. When $F_2 \le C_2$, the prosecution will challenge if $F_1 > C_1$. Hence, the expected utility to the prosecution of going first is

$$\begin{aligned} &F_1 \quad \text{if} \quad (F_2 > C_2 \ \text{and} \ F_1 > G_1) \ \text{or} \ (C_2 \ge F_2 \ \text{and} \ F_1 > C_1), \\ &G_1 \quad \text{if} \quad F_2 > C_2 \ \text{and} \ G_1 \ge F_1, \qquad (3.7) \\ &C_1 \quad \text{if} \quad C_2 \ge F_2 \ \text{and} \ C_1 \ge F_1. \end{aligned}$$

In this case the defense, going second, achieves

$$F_2 \text{ if } F_2 > C_2 \text{ and } G_1 > F_1, \tag{3.8}$$
$$G_2 \text{ if } (F_2 > C_2 \text{ and } F_1 > G_1) \text{ or } (C_2 \geq F_2 \text{ and } F_1 > C_1),$$
$$C_2 \text{ if } C_2 \geq F_2 \text{ and } C_1 \geq F_1.$$

Similarly the defense, going first, achieves

$$F_2 \text{ if } (F_1 > C_1 \text{ and } F_2 > G_2) \text{ or } (C_1 \geq F_1 \text{ and } F_2 > C_2),$$
$$G_2 \text{ if } F_1 > C_1 \text{ and } G_2 \geq F_2, \tag{3.9}$$
$$C_2 \text{ if } C_2 \geq F_2 \text{ and } C_1 \geq F_1,$$

and in this case the prosecution, going second, achieves

$$F_1 \text{ if } F_1 > C_1 \text{ and } G_2 \geq F_2, \tag{3.10}$$
$$G_1 \text{ if } (F_1 > C_1 \text{ and } F_2 > G_2) \text{ or } (C_1 \geq F_1 \text{ and } F_2 > C_2),$$
$$C_1 \text{ if } C_2 \geq F_2 \text{ and } C_1 \geq F_1.$$

Hence we have the following result:

<u>Theorem 1</u>

For each side i, $(i = 1,2)$, the expected utility of going first is

$$F_i \text{ if } (F_{3-i} > C_{3-i} \text{ and } F_i > G_i) \text{ or } (C_{3-i} \geq F_{3-i} \text{ and } F_i > C_i),$$
$$G_i \text{ if } F_{3-i} > C_{3-i} \text{ and } G_i \geq F_i, \tag{3.11}$$
$$C_i \text{ if } C_i \geq F_i \text{ and } C_{3-i} \geq F_{3-i},$$

and the expected utility of going second is

$$F_i \text{ if } F_i > C_i \text{ and } G_{3-i} \geq F_{3-i}, \tag{3.12}$$
$$G_i \text{ if } (F_i > C_i \text{ and } F_{3-i} > G_{3-i}) \text{ or } (C_i \geq F_i \text{ and } F_{3-i} > C_{3-i}),$$
$$C_i \text{ if } C_i \geq F_i \text{ and } C_{3-i} \geq F_{3-i}.$$

4. Regularity

It is natural to suppose that $G_i \geq F_i$ $(i=1,2)$ for any possible values of h_k, x, a, b, and j that might arise during the jury selection process. These inequalities compare a situation in which the prosecution has a challenges remaining and the defense has $b-1$ remaining, with a situation in which the prosecution has $a-1$ remaining and the defense has b remaining. The inequalities simply state that the prosecution would prefer the first situation and that the defense would prefer the second situation, all other conditions being equal. We shall say that a problem in which these relations are satisfied is regular.

At first thought, it might appear that every problem must be regular. Certainly, every problem that might arise in practice will be regular, because if a problem is not regular then a situation might arise in which the lawyers for one side would actually want to give one of their remaining challenges to the other side, a somewhat impractical move. However, the following example shows that there do exist problems that are not regular.

Example 1. Suppose that two jurors are to be chosen. Suppose also that there are three kinds of jurors denoted H, M, and L, and the sequence of the appearance of the first four prospective jurors is known by both sides with certainty to be H, M, L, L. Suppose finally that the preferences of each side are given by the relations $\psi_1(HL) > \psi_1(HM) > \psi_1(ML) > \psi_1(LL)$ and $\psi_2(HM) > \psi_2(LL) > \psi_2(ML) > \psi_2(HL)$.

Now suppose that $a=1$ and $b=1$. If the defense goes first and accepts the first prospective juror, with characteristics H, then the prosecution will also accept him and will use its challenge on the second juror, with characteristics M, to obtain an optimal jury of type HL, which the defense does not like. Consequently, the defense will challenge the first juror, and the outcome will be a jury of type ML.

Next suppose that $a=0$ and $b=2$. In this case, the defense will not use either of its challenges, and the outcome will be a jury of type HM. Hence, the outcome for the prosecution is better with $a=0$ and $b=2$ than with $a=1$ and $b=1$. In this bizarre circumstance then, the prosecution would, if it could, give its challenge to the defense, contradicting regularity. ∎

We do not have a good characterization of when regularity obtains. Even in the special case studied in Roth, Kadane and DeGroot (1977), we do not know if regularity must always hold. However when the utility functions of the two sides are diametrically opposed, and they agree on the joint probability distribution of the sequence of prospective jurors, we have the following theorem:

<u>Theorem 2</u>. Suppose that both sides assign the same joint distribution to the sequence $X_1, \ldots, X_{J+A+B}$. Suppose also that there exist constants a_1, a_2, and c, with $a_1 a_2 > 0$, such that

$$a\psi_1(y_1, \ldots, y_J) + a_2\psi_2(y_1, \ldots, y_J) = c \qquad (4.1)$$

for all possible values of $y_1, \ldots, y_J$. Then the problem is regular.

<u>Proof</u>. Since utility functions are defined only up to an arbitrary increasing linear transformation, we can assume without

loss of generality that $a_1 = a_2 = 1$ and $c = 0$. In this proof we shall be considering strategies for each side other than the optimal ones, so we shall extend our notation to indicate explicitly the strategies being used. Also, since both sides are using the same probability distribution, we shall delete the subscript on the expectation symbol. Thus, we write $E(\psi_i | h_k, a, b, j, s_1, s_2)$ to mean the expected utility of the situation to side i after history h_k is observed and before the next juror is drawn, when a challenges remain to the prosecution and b to the defense, j jurors remain to be chosen, and the prosecution follows strategy s_1 and the defense follows s_2.

Consider now the situation just described, and suppose that the defense adopts the following strategy s_d: At any stage of the process where the prosecution has a^* challenges remaining, the defense has b^*, and j^* jurors remain to be seated, the defense makes the decision that would be optimal if the stage were (a^*+1, b^*-1, j^*) until the first time, if ever, that a stage of the form $(0, b^*, j^*)$, with $b^* \geq 1$ is reached and the next prospective juror is such that, if the stage were $(1, b^*-1, j^*)$, it would be optimal for the prosecution to challenge him. (In other words if the prosecution went first it would be optimal for it to challenge the candidate, and if the defense went first it would be optimal for the defense to accept and the prosecution to challenge.) At such a stage, the defense challenges, and then simply follows its true optimal strategy for the remainder of the process.

Suppose that the prosecution's optimal strategy at (a^*, b^*, j^*) is s_p^*. Then since s_d is not necessarily optimal, we have

$$E(\psi_2|h_k,a,b,j,s_p^*,s_d^*) \geq E(\psi_2|h_k,a,b,s_p^*,s_d) \qquad (4.2)$$

where s_d^* is the defense's optimal strategy. Now let s_p be the optimal prosecution strategy against s_d, so that

$$\begin{aligned} E(\psi_1|h_k,a,b,j,s_p,s_d) &\geq E(\psi_1|h_k,a,b,j,s_p^*,s_d) \\ &\geq E(\psi_1|h_k,a,b,j,s_p^*,s_d^*). \end{aligned} \qquad (4.3)$$

But when the prosecution and defense use the strategies s_p and s_d, they will obtain exactly the same jury as if they had used their optimal strategies at $(a+1, b-1)$. Hence

$$\begin{aligned} E(\psi_1|h_k,a+1,b-1,j) &= E(\psi_1|h_k,a,b,j,s_p,s_d) \\ &\geq E(\psi_1|h_k,a,b,j) \end{aligned} \qquad (4.4)$$

and

$$\begin{aligned} E(\psi_2|h_k,a+1,b-1,j) &= E(\psi_2|h_k,a,b,j,s_p,s_d) \\ &\leq E(\psi_2|h_k,a,b,j). \blacksquare \end{aligned} \qquad (4.5)$$

Theorem 2 pertains only to problems in which the jury selection process is, in effect, a zero-sum game, and it says nothing about regularity when the interests of the two sides are not directly opposed. However, the next result applies to all utility functions ψ_i and all possible probability distributions that might be used by either side when only one juror is to be chosen. Since $j=1$ throughout this theorem, it is suppressed from our notation.

<u>Theorem 3</u>. When $J=1$, every problem is regular; i.e.,

$$E_1(\psi_1|a+1,b) \geq E_1(\psi_1|a,b+1)$$

and (4.6)

$$E_2(\psi_2|a+1,b) \leq E_2(\psi_2|a,b+1)$$

for all nonnegative integers a and b.

Proof. We proceed by induction on the value of $a+b$. Suppose first that $a+b=0$; i.e., $a=b=0$. We know that in this case $E_1(\psi_1|1,0) \geq E_1(\psi_1|0,1)$, since $E(\psi_1|1,0)$ is the maximum value of $E_1(\psi_1)$ that can be obtained when there is only one challenge available to the two sides. Similarly,

$$E_2(\psi_2|1,0) \leq E_2(\psi_2|0,1).$$

Now suppose that the relations (4.6) hold when $a+b=k$, where k is some nonnegative integer. We shall prove that (4.6) holds when $a+b=k+1$.

To be specific, suppose that the prosecution must decide first on the first prospective juror, and consider the following nine conceivable pairs of optimal decisions in the two problems $(a+1, b)$ and $(a, b+1)$, depending on the value x of the first prospective juror

	Optimal decisions in (a+1, b)	Optimal decisions in (a, b+1)
(1)	Accept by both sides	Accept by both sides
(2)	Accept by both sides	Challenge by prosecution.
(3)	Accept by both sides	Accept by prosecution, challenge by defense
(4)	Challenge by prosecution	Accept by both sides
(5)	Challenge by prosecution	Challenge by prosecution
(6)	Challenge by prosecution	Accept by prosecution, challenge by defense
(7)	Accept by prosecution, challenge by defense	Accept by both sides
(8)	Accept by prosecution, challenge by defense	Challenge by prosecution
(9)	Accept by prosecution, challenge by defense	Accept by prosecution, challenge by defense.

We shall now calculate the differences

$$\Delta_1 = E_1(\psi_1|a+1,b) - E_1(\psi_1|a,b+1)$$

and

$$\Delta_2 = E_2(\psi_2|a+1,b) - E_2(\psi_2|a,b+1)$$

in each of these nine cases; i.e., conditionally on x lying in each of these nine regions.

(1) In this case

$$\Delta_1 = \psi_1(x) - \psi_1(x) = 0$$

and

$$\Delta_2 = \psi_2(x) - \psi_2(x) = 0.$$

(2) The decisions under $(a,b+1)$ imply, by the induction hypothesis, that the defense would have accepted the juror with characteristics x in $(a,b+1)$ if the prosecution had accepted. Therefore, $E_1(\psi_1|a-1,b+1) > \psi_1(x)$. But these inequalities violate the induction hypothesis. Hence, this case is impossible.

(3) Here $\Delta_1 = \psi_1(x) - E_1(\psi_1|a,b) \geq 0$, since this is why it is optimal for prosecution to accept the juror with characteristics x under $(a+1,b)$. Also, $\Delta_2 = \psi_2(x) - E_2(\psi_2|a,b) \leq 0$, since this is why the defense challenges under $(a,b+1)$.

(4) Here $\Delta_1 = E_1(\psi_1|a,b) - \psi_1(x) \geq 0$, since this is why the prosecution challenges under $(a+1,b)$. Also, $\Delta_2 = E_2(\psi_2|a,b) - \psi_2(x) \leq 0$, since this is why the defense accepts under $(a,b+1)$.

(5) Here $\Delta_1 = E_1(\psi_1(a,b) - E_1(\psi_1|a-1,b+1) \geq 0$ and $\Delta_2 = E_2(\psi_2|a,b) - E_2(\psi_2|a-1,b+1) \leq 0$, by induction.

(6) Here $\Delta_1 = E_1(\psi_1|a,b) - E_1(\psi_1|a,b) = 0$ and $\Delta_2 = E_2(\psi_2|a,b) - E_2(\psi_2|a,b) = 0$.

(7) Since the defense challenges under $(a+1,b)$, then

$$E_2(\psi_2(a+1,b-1) > \psi_2(x).$$

But the defense accepts under $(a,b+1)$, which means that

$$E_2(\psi_2|a,b) \leq \psi_2(x).$$

These inequalities violate the induction hypothesis. Hence, this case is impossible.

(8) Here $\Delta_1 = E_1(\psi_1|a+1,b-1) - E_1(\psi_1|a-1,b+1) \geq 0$, by the induction hypothesis applied twice. Also, $\Delta_2 = E_2(\psi_2|a+1,b-1) - E_2(\psi_2|a-1,b+1) \leq 0$.

(9) Here $\Delta_1 = E_1(\psi_1|a+1,b-1) - E_1(\psi_1|a,b) \geq 0$ and $\Delta_2 = E_2(\psi_2|a+1,b-1) - E_2(\psi_2|a,b) \leq 0$, by induction.

A similar breakdown can be given if the defense must decide first. Thus, in all possible cases, the relations (4.6) are satisfied. ∎

5. The Advantage of Going First

In this section we show that in a regular problem it is never disadvantageous to go first and that in an important class of problems it is irrelevant which side goes first. We begin by establishing the special conditions that are needed in order for it to be strictly advantageous to go second.

Theorem 4. It is strictly advantageous for side i to go second if and only if

$$G_i > F_i > C_i \quad \text{and} \quad C_{3-i} \geq F_{3-i} \geq G_{3-i}. \tag{5.1}$$

Proof. Let Δ be the expected utility of going first minus the expected utility of going second. Using Theorem 1,

$$\Delta = \begin{cases} G_i - F_i & \text{if } G_i \geq F_i > C_i \text{ and } G_{3-i} \geq F_{3-i} > C_{3-i}, & (5.2a) \\ F_i - G_i & \text{if } F_i > C_i,\ F_i > G_i,\ F_{3-i} > C_{3-i}, \text{ and } F_{3-i} > G_{3-i}, & (5.2b) \\ F_i - G_i & \text{if } C_i \geq F_i > G_i \text{ and } F_{3-i} > C_{3-i}, & (5.2c) \\ F_i - G_i & \text{if } F_i > C_i \text{ and } C_{3-i} \geq F_{3-i} > G_{3-i}, & (5.2d) \\ 0 & \text{otherwise.} & (5.2e) \end{cases}$$

Only in (5.2d) might Δ be negative. Thus $\Delta < 0$ if and only if $G_i > F_i > C_i$ and $C_{3-i} \geq F_{3-i} > G_{3-i}$. ∎

Corollary. In a regular problem,

$$\Delta = \begin{cases} G_i - F_i & \text{if } F_i > C_i \text{ and } F_{3-i} > C_{3-i}, \\ 0 & \text{otherwise}, \end{cases} \qquad (5.3)$$

and it is never disadvantageous to go first.

Proof. Regularity (i.e., $G_i \geq F_i$ and $G_{3-i} \geq F_{3-i}$) eliminates (5.2b), (5.2c), and (5.2d), and reduces (5.2a) as shown. ∎

The expression for Δ in the Corollary has a natural interpretation. If $F_i > C_i$ and $F_{3-i} > C_{3-i}$, then each side prefers challenging the prospective juror to letting him be seated on the jury. Consequently, whichever side goes first can accept the juror, confident that the side going second must challenge. This yields a gain of $G_i - F_i$ to the side that goes first, which is how much that side prefers the other side to exercise a challenge compared with exercising a challenge itself. In all other cases, where at least one side would prefer seating the juror to challenging him, the order of challenges is irrelevant, and there is no gain in going first.

Theorem 4 leaves open the question of whether there could be a problem that was not regular in which it was actually advantageous to go second. The next example presents such a problem.

Example 2. Suppose as in Example 1 that $J=2$. Suppose now, however, that there are four kinds of jurors denoted S, H, M and L, and that the sequence of the appearance of the first five prospective jurors is known by both sides with certainty to be S, H, M, L, L. Suppose also that the preferences of each side are given by the relations

$$\psi_1(SH) = \psi_1(SM) = \psi_1(SL) > \psi_1(HL) > \psi_1(HM) > \psi_1(ML) > \psi_1(LL)$$

and

$$\psi_2(HM) > \psi_2(LL) > \psi_2(ML) > \psi_2(HL) > \psi_2(SH) = \psi_2(SM) = \psi_2(SL).$$

Now suppose that $a=1$ and $b=2$. If the defense goes first, and accepts, the prosecution will also accept, which is bad for the defense. Thus the defense will challenge, which leads to a situation with $a=1$ and $b=1$. By Example 1, the outcome will be a jury of type ML.

If the prosecution goes first, it sees that the defense will challenge . However, it was shown in Example 1 that the prosecution would prefer to use a challenge itself rather than have the defense use a challenge. Consequently, the prosecution will challenge, which leads to a situation with $a=0$, and $b=2$. By Example 1, the outcome will be a jury of type HM. Hence, the defense prefers to go second in this situation. ∎

It is interesting to note that, in Example 2, it is optimal for the prosecution, going first, to challenge S even though it prefers any jury containing S to any jury without S.

The next result shows that when the prosecution and defense use the same probability distributions and have either exactly the same utilities or diametrically opposed utilities, neither side cares who goes first. Thus, it is only when their interests are partially coincident and partially opposed, as must typically be the case, that it matters who goes first.

Theorem 5. Suppose that both sides assign the same joint distribution to $X_1, \ldots, X_{J+A+B}$. Suppose also that there exist constants a_1, a_2 and c, with $a_1 a_2 \neq 0$, such that (4.1) holds for every value of $y_1, \ldots, y_J$. Then the order in which the two sides exercise their challenges is irrelevant to both of them.

Proof. As in Theorem 2, we may, without loss of generality, take $\psi_1 = \pm \psi_2$.

If $\psi_1 = -\psi_2$, then by Theorem 2 the problem is regular. Consequently by Corollary 1, the only case to be concerned with is that in which $F_i > C_i$ and $F_{3-i} > C_{3-i}$. But since $\psi_1 = -\psi_2$, it follows that $F_i = -G_{3-i}$ and $C_i = -C_{3-i}$. Thus, by regularity, $F_i > C_i$ implies that $C_{3-i} > G_{3-i} \geq F_{3-i}$. Hence the only case in which $\Delta > 0$ in Corollary 1 cannot occur here, so $\Delta = 0$ and the order is irrelevant.

If $\psi_1 = \psi_2$, then $F_i = G_{3-i}$ and $C_i = C_{3-i}$. Substituting these values into the expressions given in Theorem 1 for the expected utility of going first and of going second, we find that they both reduce to $\mathrm{Max}(F_1, G_1, C_1)$. Consequently $\Delta = 0$ and again the order is irrelevant. ∎

6. Summary

We have shown that in a regular problem, it is never disadvantageous to go first. We have shown also, that if the two sides assign the same joint distribution to prospective jurors and have diametrically opposed utilities then the problem is regular. In this particular type of problem, the two sides never want to challenge the same prospective juror, so it does not matter which side goes first. We have also shown that any problem in which only one juror is to be chosen must be regular.

Examples were given to establish that there are problems that are not regular and in which it is advantageous to go second.

References

1. Brams, S.J. and M.D. Davis (1976a), "Optimal Jury Selection: A Game-Theoretic Model for the Exercise of Peremptory Challenges", unpublished manuscript.

2. Brams, S.J. and M.D. Davis (1976b) "A Game-Theory Approach to Jury Selection", Trial 12, 47-49.

3. Cayley, A. (1875), "Mathematical Questions and Their Solutions", Educational Times 22, 18-19.

4. Cyert, R.M. and M.H. DeGroot (1970), "Multiperiod Decision Models with Alternating Choice as a Solution to the Duopoly Problem", Quart. J. Econ., 84, 410-429.

5. Cyert, R.M. and M.H. DeGroot (1973), "An Analysis of Cooperation and Learning in a Duopoly Context", Am.Econ.Rev. 63, 24-37.

6. DeGroot, M.H. (1970), Optimal Statistical Decisions, New York: McGraw-Hill.

7. Gilbert, J.P. and F. Mosteller (1966), "Recognizing the Maximum of a Sequence", J.Am.Stat.Assn. 61, 35-73.

8. Roth, A., J.B. Kadane, and M.H. DeGroot (1977), "Optimal Peremptory Challenges in Trials by Juries: A Bilateral Sequential Process", Operations Research, 25, 901-919.