LUBA SAPIR

THE OPTIMALITY OF THE EXPERT AND MAJORITY RULES UNDER EXPONENTIALLY DISTRIBUTED COMPETENCE

ABSTRACT. We study the uncertain dichotomous choice model. In this model a set of decision makers is required to select one of two alternatives, say 'support' or 'reject' a certain proposal. Applications of this model are relevant to many areas, such as political science, economics, business and management. The purpose of this paper is to estimate and compare the probabilities that different decision rules may be optimal. We consider the expert rule, the majority rule and a few in-between rules. The information on the decisional skills is incomplete, and these skills arise from an exponential distribution. It turns out that the probability that the expert rule is optimal far exceeds the probability that the majority rule is optimal, especially as the number of the decision makers becomes large.

KEY WORDS: Decision rule, Expert rule, Majority rule, Optimal rule, Partial information, Logarithmic expertise

1. INTRODUCTION

Condorcet first regarded the dichotomous model in his *Essai* (1785). He considered the case of a group of decision makers reaching a decision on some issue using the simple majority rule. He made the statement that the group would be likely to make the correct choice as the size of the group becomes large. Moreover, the probability of a correct decision increases to 1 as the number of individuals in the group tends to infinity. In the following, by 'Condorcet Jury Theorem' (henceforward CJT) we mean a formulation of conditions under which Condorcet's statement is in fact valid. One of the most popular among such conditions is the condition that the members of the group vote independently and the probabilities that the members will make the right choice are equal and exceed 1/2.

Researchers have extended and completed Condorcet's work in different directions. A generalization of CJT has been presented by Grofman, Owen and Feld (1983), who relaxed the assumption

Theory and Decision **45:** 19–35, 1998.

that the correctness probability of each member must exceed 1/2. Instead, they only required the average competence of the group to exceed 1/2. In another direction, Sven Berg (1993) and Ladha (1995) relaxed the independence assumption, allowing correlated votes. Paroush (1996) and Berend and Paroush (1996) obtained exactly necessary and sufficient conditions for CJT in the case of independence. Other approaches have been tried by other authors as well.

Many works exploring the dichotomous choice model assume a known 'expertise level', i.e. a known probability for each group member to make the right choice. This assumption makes it easy to find the optimal decision rule – the one maximizing the probability for the group to obtain the correct alternative. In the case of symmetric alternatives (i.e., equiprobable *a priori* with equal penalties for each type of error), the optimal rule is a weighted majority rule, the weights being given by the logarithms of the members' odds of making the right choices (see Nitzan and Paroush 1982; Shapley and Grofman 1984). However, the assumption of full information regarding the decision makers' competence is very restrictive and often far from being fulfilled. thus, the case of incomplete information on decisional skills seems to better approximate practical situations. While no direct information on the 'expertise level' is available, some information does exist regarding the distribution of decisional competence in the population from which the individual members were drawn. The case of log-normal distribution of individual odds of choosing correctly was tackled by Nitzan and Paroush (1984, 1985). The case where the probabilities of being right for each expert are uniformly distributed in [1/2, 1] was discussed by Berend and Harmse (1993). This paper continues the exploration in the same direction under the assumption that the 'logarithmic expertise level' is exponentially distributed. The main accomplishment is the derivation of exact formulas for the probability of optimality of the expert and majority rules in this case (see Theorem 1 and Theorem 3 in Section 3). These probabilities are independent of the value of the parameter of the distribution. Section 2 is devoted to a more accurate description of our model. In Section 3 we present the main results, and their proofs in Section 4.

2. THE MODEL

The group consists of n members, and each advocated one of two alternatives. The probability that the ith member will make the right choice is denoted by p_i. A decision rule is a rule for translating the individual opinions into a group decision. Such a rule is *optimal* if it maximizes the probability that the group will make the correct decision for all possible combinations of opinions. If the members indexed by some subset $A \subseteq \{1, 2, \ldots, n\}$ of the group recommend the first alternative, while those indexed by $B = \{1, \ldots, n\} \setminus A$ recommend the second, then the first alternative should be chosen if and only if

$$\sum_{i \in A} \ln\left(\frac{p_i}{1 - p_i}\right) > \sum_{i \in B} \ln\left(\frac{p_i}{1 - p_i}\right) \tag{2.1}$$

(see Nitzan and Paroush 1985). It is therefore natural to define the *expertise* of an individual, whose probability of being correct is p, as $p/1 - p$, and his *logarithmic expertise* as $\ln(p/1 - p)$. It will be convenient to consider the functions F and f defined by:

$$F(p) = \frac{p}{1 - p},$$

$$f(p) = \ln\left(\frac{p}{1 - p}\right) = \ln(F(p)).$$

The *combined expertise* of a set of l experts with correctness probabilities $p_1, \ldots, p_l$ is

$$F(p_1, \ldots, p_l) = \prod_{j=1}^{l} F(p_j),$$

and the *combined logarithmic expertise* of the same experts is

$$f(p_1, \ldots, p_l) = \sum_{j=1}^{l} f(p_j).$$

3. THE MAIN RESULTS

In this paper we are concerned with the probabilities that various decision rules may be optimal. Specifically, we consider the expert rule (Theorem 1) and the majority rule (Theorem 3). We also estimate the probability that the optimal rule belongs to certain sets of rules (Theorem 2). Obviously, the results depend on n, and a comparison of the asymptotics of the answers is of primary importance. We will assume that the expertise levels $f(p_i)$, $i = 1, \dots, n$, are distributed $\text{Exp}(\lambda)$. In other words, the density function of each p_i is

$$g_p(t) = \begin{cases} \frac{\lambda(1-t)^{\lambda-1}}{t^{\lambda+1}}, & t \in [\frac{1}{2}, 1], \\ 0, & \text{otherwise}, \end{cases}$$

and its distribution function is

$$G_p(t) \begin{cases} 0, & t \leqslant \frac{1}{2}, \\ 1 - e^{-\lambda \ln \frac{t}{1-t}} = 1 - \left(\frac{1-t}{t}\right)^{\lambda}, & \frac{1}{2} < t \leqslant 1, \\ 1, & t > 1. \end{cases}$$

THEOREM 1. *The probability $P_e(n)$ that the expert rule is optimal is:*

$$P_e(n) = \frac{n}{2^{n-1}}.$$

Unlike most other results of the paper, one can provide a short proof of Theorem 1. In fact, as the top expert may be any of the n, the probability in question is exactly n times the probability that the expertise of a randomly selected expert exceeds the combined expertise of the rest. Let $f(p)$, $f(q_1) \dots f(q_{n-1})$ be independent random variables distributed $\text{Exp}(\lambda)$. Clearly:

$$\begin{aligned} P_e(n) &= n \cdot \text{Prob}\left\{ f(p) \geqslant \sum_{i=1}^{n-1} f(q_i) \right\} \\ &= n \cdot P\{f(p) \geqslant f(q_1)\} \cdot \prod_{k=2}^{n-1} \\ &\qquad P\left\{ f(p) \geqslant \sum_{i=1}^{k} f(q_i) | f(p) \geqslant \sum_{i=1}^{k-1} f(q_i) \right\}. \end{aligned}$$

Now, due to the lack-of-memory property of the exponential distribution, each of these probabilities is equal to 1/2. Hence:

$$P_e(n) = n \cdot \left(\frac{1}{2}\right)^{n-1}.$$

Table 1 illustrates the likelihood of optimality of the expert rule as obtained by Theorem 1, for $3 \leqslant n \leqslant 9$.

TABLE 1

Expert rule – optimality probability

Number of experts	3	4	5	6	7	8	9
$P_e(n)$	0.750	0.500	0.312	0.187	0.109	0.062	0.035

For two mutually exclusive and complementing subsets of experts A and B, we write $A \succeq B$ if the probability that the combined expertise of the experts indexed by A is greater than that of the experts indexed by B exceeds 1/2 (namely, (2.1) is satisfied with a probability greater than 1/2). Our next result is designed to check by how much a few of the 'bottom' group members add credibility to the top expert.

THEOREM 2. *For $n \geqslant 2(k+1)$:*

$$P(\{1, n-k+1, \ldots, n-1, n\} \succeq \{2, 3, \ldots, n-k\}) = \frac{n!}{(n-k-1)!(n-k-1)^k 2^{n-1}}.$$

Remark 1. The special case $k = 0$ of Theorem 2 amounts to Theorem 1.

Remark 2. In case $n < 2(k+1)$, we will get that half or more experts from the bottom join the top expert, $k \geqslant (n-1)/2$, so the probability that the top one and k experts from the bottom will have higher expertise than all the others can only increase in comparison with above case, i.e.

$$P(\{1, n-k+1, \dots, n-1, n\} \succeq \{2, 3, \dots, n-k\}) \geqslant \begin{cases} \dfrac{n!}{\left(\frac{n+1}{2}\right)!(n+1)^{\frac{n-3}{2}} 2^{\frac{n+1}{2}}}, & n \text{ is odd}, \\ \dfrac{n!}{(\frac{n}{2})! n^{\frac{n-2}{2}} 2^{\frac{n}{2}}}, & n \text{ is even}. \end{cases}$$

Table 2 illustrates the probabilities obtained by Theorem 2 for $4 \leqslant n \leqslant 13$ and $k \leqslant \frac{n}{2} - 1$. Monte Carlo simulations provided very similar data. (The $\approx$ sign in Table 2 provides data obtained only by Monte Carlo method. All blank entries are trivially 1.)

TABLE 2

Top and bottom versus middle

$n \backslash k$	1	2	3	4	5
4	0.750				
5	0.417	≈0.848			
6	0.234	0.417	≈0.917		
7	0.131	0.205	≈0.522	≈0.964	
8	0.073	0.105	0.205	≈0.604	≈0.978
9	0.040	0.055	0.095	≈0.249	≈0.723
10	0.022	0.029	0.046	0.095	≈0.291
11	0.012	0.015	0.023	0.042	≈0.120
12	0.006	0.008	0.011	0.019	0.042
13	0.001	0.001	0.002	0.003	0.006

Remark 3. Looking at Table 2, one immediately recognizes some patterns. For example, we always have

$$\begin{aligned} &P(\{1, k+4, k+5 \dots, 2k+3\} \succ \{2, 3, \dots, k+3\}) \\ &\quad = P(\{1, k+4, k+5, \dots, 2k+4\} \succ \{2, 3, \dots, k+3\}). \end{aligned}$$

While this result follows by a simple substitution in Theorem 2, it is instructive to note that it can be proved easily using the lack-of-memory property of the exponential distribution. More generally, if n is even and A and B are equinumerous with, say, $n \in A$, then $P(A \succ B) = P((A \backslash \{n\}) \succ B)$.

The case $k = 1$ of Theorem 2 gives

COROLLARY 1.

$$P(\{1, n\} \succ \{2, 3, \ldots, n-1\}) = \frac{n(n-1)}{2^{n-1}(n-2)}.$$

Employing Theorem 1 we also obtain

COROLLARY 2.

$$P(\{1, n\} \succ \{2, 3, \ldots, n-1\}, \{1\} \prec \{2, 3, \ldots, n\}) \\ = \frac{n(n-1)}{2^{n-1}(n-2)} - \frac{n}{2^{n-1}} = \frac{n}{(n-2)2^{n-1}}.$$

Note that the left hand side is the probability that the decision rule $(n-1, 1, \ldots, 1)$ is optimal. It is worthwhile mentioning that, perhaps somewhat surprisingly, even though this decision rule seems to be more 'balanced' than the expert rule, its probability of being optimal is quite a lot smaller than that of the expert rule.

THEOREM 3. *For odd $n = 2s + 1$ the probability that the majority rule is optimal is:*

$$P_m(n) = \frac{1}{2^s(s+1)^s} \left\{ 1 - \frac{\binom{2s+1}{s}(s+1)}{(2s)^s} \sum_{i=0}^{s-1} \frac{(s-i-1)^s(-1)^i}{s+i+1} \cdot \binom{s}{i} \right\}$$

In Table 3 we present the few initial values relating to Theorem 3. We added here also, for comparison, the corresponding probabilities for the expert rule. The table shows very distinctly that the expert rule is far more likely to be optimal than the majority rule for even quite small values of n.

It is interesting to compare, for the case $n = 5$, the probability of the optimality of each of the seven weighted majority rules with the

TABLE 3

Optimality probability – expert versus majority rule

n	3	5	7	9	11
$P_e(n)$	0.75	0.31	0.11	0.04	0.01
$P_m(n)$	0.25	0.01	$1.8 \cdot 10^{-4}$	$1.7 \cdot 10^{-6}$	$1.0 \cdot 10^{-8}$

TABLE 4

Optimality likelihoods of all weighted majority rules for $n = 5$

Rule	Distribution	
	$f(p_i) \sim \text{Exp}(\lambda)$	$p_i \sim U(\frac{1}{2}, 1)$
(1,0,0,0,0)	0.312	0.199
(1,1,1,1,1)	0.010	0.022
(1,1,1,0,0)	0.157	0.175
(3,1,1,1,1)	0.104	0.107
(2,1,1,1,0)	0.208	0.229
(3,2,2,1,1)	0.157	0.194
(2,2,1,1,1)	0.052	0.074

results of Nitzan and Paroush (1985) obtained for the uniformly distributed p_i. The point of the above comparison is that it substantiates the robustness of the conclusion that the expert rule is far more likely to be optimal than the majority rule, even for quite small values of n, for instance, for $n = 5$.

4. THE PROOFS

Proof of Theorem 1. As the top expert may be any of the n, the probability in question is exactly n times the probability that the expertise of a randomly selected expert exceeds the combined expertise of the rest. Let $f(p)$, $f(q_1) \ldots f(q_{n-1})$ be independent random variables distributed Exp(λ). Putting $t = \sum_{i=1}^{n-1} f(q_i)$ we have $t \sim$

Gamma($\lambda, n-1$), i.e. the density function of t is given by:

$$\rho_t(x) = \begin{cases} \frac{\lambda^{n-1} x^{n-2} e^{-\lambda x}}{\Gamma(n-1)}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Thus:

$$\begin{aligned} P_e(n) &= n \cdot \text{Prob}\left\{ f(p) \geqslant \sum_{i=1}^{n-1} f(q_i) \right\} \\ &= n \int_0^\infty \frac{\lambda^{n-1} t^{n-2} e^{-\lambda t}}{\Gamma(n-1)} \int_t^\infty \lambda e^{-\lambda s} \,\mathrm{d}s \,\mathrm{d}t \\ &= n \int_0^\infty \frac{\lambda (^{n-1} t^{n-2} e^{-\lambda t}}{\Gamma(n-1)} \cdot e^{-\lambda t} \,\mathrm{d}t \\ &= \frac{n}{2^{n-1}} \int_0^\infty \frac{(2\lambda)^{n-1} t^{n-2} e^{-2\lambda t}}{\Gamma(n-1)} \,\mathrm{d}t = \frac{n}{2^{n-1}} . \end{aligned}$$

$\square$

Proof of Theorem 2. Let Y_i, $i = 1, 2, \ldots, n$, be the order statistics of $f(p_i)$, namely:

$$Y_1 = \max f(p_i), \ldots, Y_n = \min f(p_i),$$

$$Y_n \leqslant Y_{n-1} \leqslant \cdots \leqslant Y_1.$$

Let a_i be fixed values of Y_i for $i = n-k+1, \ldots, n$. The random variables Y_i, $i = n-k+1, \ldots, n$, represent the k bottom experts and their densities are:

$$\rho_{Y_i | Y_{i+1} = a_{i+1}}(t) = \begin{cases} i\lambda e^{-i\lambda(t - a_{i+1})}, & t \geqslant a_{i+1} \\ 0, & \text{otherwise.} \end{cases}$$

Here we took $Y_{n+1} = a_{n+1} = 0$.

Fix Y_i, $i = n-k+1, \ldots, n$. Clearly, if we add to the bottom k experts one of the top $n-k$, but not the top expert, their combined expertise will be less than that of the others. Hence,

$$\begin{aligned} &P(\{1, n-k+1, \ldots, n-1, n\} \succeq \{2, 3, \ldots, n-k\}) \\ &\quad = (n-k) P(\{l, n-k+1, \ldots, n-1, n\} \\ &\quad \succeq \{1, \ldots, n-k\} \backslash \{l\}), \end{aligned}$$

where l is chosen uniformly among the numbers $1, 2, \ldots, n-k$.

Let $f^*(p_i)$, $i = 1, \ldots, n-k$, be independent, identically distributed random variables, representing the top $n-k$ experts, after the k bottom ones have been fixed. Then:

$$\begin{aligned}
&P(D|Y_{n-k+1} = a_{n-k+1}, \ldots, Y_n = a_n) \\
&\quad = P(Y_1 + a_{n-k+1} + \cdots + a_n \geqslant Y_2 + \cdots + Y_{n-k}) \\
&\quad = (n-k)P\left(f^*(p_1) + a_{n-k+1} + \cdots + a_n \geqslant \sum_{i=2}^{n-k} f^*(p_i)\right)
\end{aligned}$$

Due to the lack-of-memory property of the exponential distribution it is possible to write $f^*(p_i) = a_{n-k+1} + Z_i$, $i = 1, \ldots, n-k$, where $Z_1, Z_2, \ldots, Z_{n-k}$ are independent, $Z_i \sim \mathrm{Exp}(\lambda)$. Hence

$$\begin{aligned}
&P(D|Y_{n-k+1} = a_{n-k+1}, \ldots, Y_n = a_n) \\
&\quad = (n-k)P\left\{Z_1 \geqslant (n-k-3)a_{n-k+1} - \sum_{i=n-k+2}^{n} a_i + \sum_{i=2}^{n-k} Z_i\right\} \\
&\quad = (n-k)P\left\{Z_1 \geqslant X + (n-k-3)a_{n-k+1} - \sum_{i=n-k+2}^{n} a_i\right\},
\end{aligned}$$

where

$$X = \sum_{i=2}^{n-k} Z_i \sim \mathrm{Gamma}(\lambda, n-k-1).$$

Consequently:

$$\begin{aligned}
&P(D|Y_{n-k+1} = a_{n-k+1}, \ldots, Y_n = a_n) \\
&\quad = (n-k)\int_0^\infty \frac{\lambda^{n-k-1}x^{n-k-2}e^{-\lambda x}}{\Gamma(n-k-1)} \int_{x+(n-k-3)a_{n-k+1}-\sum_{i=n-k+2}^{n} a_i}^{\infty} \lambda e^{-\lambda t}\,\mathrm{d}t\,\mathrm{d}x \\
&\quad = (n-k)\int_0^\infty \frac{\lambda^{n-k-1}x^{n-k-2}e^{-2\lambda x}e^{-\lambda((n-k-3)a_{n-k+1}-\sum_{i=n-k+2}^{n} a_i)}}{\Gamma(n-k-1)}\,\mathrm{d}x \\
&\quad = \frac{(n-k)e^{-\lambda((n-k-3)a_{n-k+1}-\sum_{i=n-k+2}^{n} a_i)}}{2^{n-k-1}} \int_0^\infty \frac{(2\lambda)^{n-1-k}x^{n-k-2}e^{-2\lambda x}}{\Gamma(n-k-1)}\,\mathrm{d}x.
\end{aligned}$$

The last integrand is the density function of a Gamma$(2\lambda, n-k-1)$ random variable, so that

$$\int_0^\infty \frac{(2\lambda)^{n-k-1} x^{n-k-2} e^{-2\lambda}}{\Gamma(n-k-1)}\,\mathrm{d}x = 1,$$

and therefore:

$$\begin{aligned} &P(D|Y_{n-k+1} = a_{n-k+1}, \ldots, Y_n = a_n) \\ &\quad = \frac{(n-k)e^{-\lambda((n-k-3)a_{n+1-k} - \sum_{i=n+2-k}^{n} a_i)}}{2^{n-k-1}}. \end{aligned}$$

Denote:

$$P_j = P(D|Y_{n-k+1+j} = a_{n-k+1+j}, \ldots, Y_n = a_n), \quad j = 0, \ldots, k.$$

Thus, we have computed

$$P_0 = P(D|Y_{n-k+1} = a_{n-k+1}, \ldots, Y_n = a_n),$$

and our aim is to compute $P_k = P(D)$. More generally, let us show by induction on j that

$$P_j = \frac{(n-k)\cdots(n-k+j)}{(n-k-1)^j 2^{n-k+j-1}} e^{-\lambda((n-k-j-3)a_{n-k+j+1} - \sum_{i=n-k+j+2}^{n} a_i)},$$
$$j = 1, 2, \ldots, k.$$

(Here, for $j = k-1$ and $j = k$ the sum in the exponent is empty.) In fact, this has been shown for $j = 0$, and for $1 \leqslant j \leqslant k-2$ we have:

$$\begin{aligned} P_j &= \int_{a_{n-k+j+1}}^{\infty} P_{j-1}\rho_{Y_{n-k+j}}(t)\,\mathrm{d}t \\ &= \frac{(n-k+j)!e^{\lambda((n-k+j)a_{n-k+j} + \sum_{i=n-k+j+2}^{n} a_i)}}{(n-k-1)! * (n-k-1)^{j-1} 2^{n-k+j-2}} \\ &\qquad \int_{a_{n-k+j+1}}^{\infty} \lambda e^{-\lambda x(n-k-j-2)} e^{-(n-k+j)\lambda x}\,\mathrm{d}x \\ &= \frac{(n-k)\cdots(n-k+j)}{(n-k-1)^j 2^{n-k+j+1}} e^{-\lambda((n-k-j-3)a_{n-k+j+1} - \sum_{i=n-k+j+2}^{n} a_i)}. \end{aligned}$$

In particular:

$$\begin{aligned}P(D) = P_k &= \frac{(n-k)\cdots(n-1)n}{(n-k-1)^k 2^{n-1}}\\ &= \frac{n!}{(n-k-1)!(n-k-1)^k 2^{n-1}}.\end{aligned} \qquad \square$$

Proof of Theorem 3. Let Y_i, $i = 1, \ldots, 2s+1$, be the order statistics of $f(p_i)$, as in the proof of Theorem 2. Put $M = \{Y_1 + Y_2 + \cdots + Y_s \leqslant Y_{s+1} + \cdots + Y_{2s+1}\}$. Obviously, the probability sought for in the theorem is $P(M)$. Let a_i be arbitrary fixed values of the variables Y_i. Denote $A_{s+j} = \sum_{i=s+j}^{n} a_i$, $j = 1, 2, \ldots, s+1$.
Then:

$$\begin{aligned}&P(M|Y_{s+1} = a_{s+1}, Y_{s+2} = a_{s+2}, \ldots, Y_n = a_n)\\ &\quad = P(Y_1 + Y_2 + \cdots + Y_s \leqslant a_{s+1} + \cdots + a_n)\\ &\quad = P(f^*(p_1) + \cdots + f^*(p_s) \leqslant a_{s+1} + \cdots + a_n),\end{aligned}$$

where $f^*(p_i)$ are independent variables representing each of the s nonranked experts, after we have also fixed the $(s+1)$ bottom ones.

According to the lack-of-memory property of the exponential distribution these variables can be represented as $f^*(p_i) = Z_i + a_{s+1}$, $Z_i \sim \mathrm{Exp}(\lambda)$, $i = 1, \ldots, s$, and $X = \sum_{i=1}^{s} Z_i \sim \mathrm{Gamma}(\lambda, s)$. Hence:

$$\begin{aligned}&P(M|Y_{s+1} = a_{s+1}, Y_{s+2} = a_{s+2}, \ldots, Y_n = a_n)\\ &\quad = P\left(X + sa_{s+1} \leqslant \sum_{i=s+1}^{n} a_i\right)\\ &\quad = P\left(X \leqslant \sum_{i=s+2}^{n} a_i - (s-1)a_{s+1}\right).\end{aligned}$$

If $a_{s+1} \geqslant (1/s - 1)A_{s+2}$ the last expression clearly vanishes. For $a_{s+1} < (1/s - 1)A_{s+2}$ we have

$$\begin{aligned}&P(M|Y_{s+1} = a_{s+1}, Y_{s+2} = a_{s+2}, \ldots, Y_n = a_n)\\ &\quad = \int^{A_{s+2}-(s-1)a_{s+1}} \frac{\lambda^s x^{s-1} e^{-\lambda x}}{\Gamma(s)}\, \mathrm{d}x\end{aligned}$$

$$= \int_0^\infty \frac{\lambda^s x^{s-1} e^{-\lambda x}}{\Gamma(s)}\,\mathrm{d}x - \int_{A_{s+2}-(s-1)a_{s+1}}^\infty \frac{\lambda^s x^{s-1} e^{-\lambda x}}{\Gamma(s)}\,\mathrm{d}x$$
$$= 1 - e^{-\lambda(A_{s+2}-(s-1)a_{s+1})} \cdot \sum_{i=0}^{s-1} \frac{(\lambda(A_{s+2}-(s-1)a_{s+1}))^i}{i!}.$$

Put:

$$P_{s+k} = P(M|Y_{s+k+1} = a_{s+k+1},\ Y_{s+k+2} = a_{s+k+2}, \dots, Y_n = a_n),$$
$$k = 0, \dots, s.$$

Thus, we have calculated P_{s+k} for $k = 0$, whereas the theorem concerns $k = s + 1$. We shall make the transition by increasing k gradually. For the next calculation let us introduce:

$$b_i = A_{s+i} - (s-i+1)a_{s+i-1}, \qquad i = 2, 3, \dots, s+1,$$
$$C_j = (s+1)(s+2)\dots(s+j) = \tfrac{(s+j)!}{s!}, \quad j = 1, 2, \dots, s+1,$$
$$C_0 = 0.$$

We have shown that (in the non-trivial case $a_{s+1} < \frac{1}{s-1}A_{s+2}$):

$$P_s = 1 - e^{-\lambda b_2} \sum_{i=0}^{s-1} \frac{(\lambda b_2)^i}{i!}.$$

Then

$$a_{s+2} \leqslant a_{s+1} \leqslant \frac{1}{s-1} A_{s+2}$$

and

$$P_{s+1} = (s+1)\lambda e^{\lambda(s+1)a_{s+2}} \cdot \left(\int_{a_{s+2}}^{\frac{1}{s-1}A_{s+2}} e^{-(s+1)\lambda t}\,\mathrm{d}t \right.$$
$$\left. - e^{-\lambda A_{s+2}} \sum_{i=0}^{s-1} \frac{1}{i!} \int_{a_{s+2}}^{\frac{1}{s-1}A_{s+2}} e^{-2\lambda t} \lambda (A_{s+2} - (s-1)t)^i\,\mathrm{d}t \right).$$

Using the change of variables $y = \lambda(A_{s+2} - (s-1)t)$ we find that

$$\begin{aligned}
P_{s+1} =&(s+1)\lambda e^{\lambda(s+1)a_{s+2}} \cdot \left[\frac{e^{-(s+1)a_{s+2}\lambda} - e^{-\frac{(s+1)}{(s-1)}A_{s+2}\lambda}}{(s+1)\lambda} \right. \\
&\left. - e^{-\lambda A_{s+2} - \frac{2\lambda}{s-1}A_{s+2}} \sum_{i=0}^{s-1} \frac{1}{i!} \int_0^{\lambda(A_{s+3}-(s-2)a_{s+2})} \frac{e^{\frac{2y}{s-1}}}{\lambda(s-1)} y^i \, \mathrm{d}y \right] \\
=&1 - e^{-\lambda\frac{s+1}{s-1}(A_{s+3}-(s-2)a_{s+2})} - (s+1)e^{-\lambda\frac{(s+1)}{(s-1)}(A_{s+3}-(s-2)a_{s+2})} \\
&\cdot \sum_{i=0}^{s-1} \frac{(s-1)^i}{i!2^{i+1}} \cdot \int_0^{\frac{2\lambda}{s-2}(A_{s+3}-(s-2)a_{s+2})} e^{\frac{2y}{s-1}} \left(\frac{2y}{s-1}\right)^i \mathrm{d}\left(\frac{2y}{s-1}\right).
\end{aligned}$$

Now, since

$$\int_0^h e^z z^i \, \mathrm{d}z = e^b \sum_{j=0}^{i} h^j \frac{i!}{j!}(-1)^{i-j} + (-1)^{i-1} i!$$

and $b_3 = A_{s+3} - (s-2)a_{s+2}$ we obtain a more convenient form for P_{s+1}:

$$\begin{aligned}
P_{s+1} =&1 - e^{-\lambda\frac{s+1}{s-1}b_3} - (s+1)e^{-\lambda\frac{s+1}{s-1}b_3} \\
&\cdot \sum_{i=0}^{s-1} \frac{(s-1)^i}{i!2^{i+1}} \left(e^{\frac{2\lambda}{s-1}b_3} \sum_{j=0}^{i} \left(\frac{2\lambda}{s-1}\right)^j b_3^j \frac{i!}{j!}(-1)^{i-j} + (-1)^{i-1} i! \right) \\
=&1 - e^{-\lambda\frac{s+1}{s-1}b_3} + \frac{(s+1)}{2} e^{-\lambda\frac{s+1}{s-1}b_3} \\
&\cdot \sum_{i=0}^{s-1} \left(\frac{s-1}{2}\right)^i (-1)^i - e^{-\lambda b_3} \sum_{i=0}^{s-1} \frac{(\lambda b_3)^i}{i!} \left(1 - \left(\frac{1-s}{2}\right)^{s-i}\right).
\end{aligned}$$

Thus:

$$\begin{aligned}
P_{s+1} =&1 - \left(\frac{1-s}{2}\right)^i e^{-\lambda\frac{s+1}{s-1}b_3} \\
&- e^{-\lambda b_3} \sum_{i=0}^{s-1} \frac{(\lambda b_3)^i}{i!} \left(1 - \left(\frac{1-s}{2}\right)^{s-i}\right).
\end{aligned}$$

In the same way we can continue the process of obtaining P_{s+k}:

$$(1) \quad P_{s+k} = \int_{a_{s+k+1}}^{\frac{1}{s-k}A_{s+k+1}} P_{s+k-1}(t) \cdot \rho_{y_{s+k}}(t) \, \mathrm{d}t, \qquad k = 1, 2, \ldots, s-1,$$

where $\rho_{y_{s+k}}(t)$ is the density function of the order statistics Y_{s+k}. By induction we can prove

$$
\begin{aligned}
P_{s+k} =& 1 + \frac{(-1)^k C_k}{s^{k-1}} \sum_{j=1}^{k} \frac{(-1)^{j-1} j^{k-1}}{(j-1)!(k-j)!(s+j)} \\
&\cdot \left(\frac{j-s}{2j}\right)^{s+k-1} e^{-\frac{s+j}{s-j}\lambda b_{k+2}} - e^{-\lambda b_{k+2}} \sum_{i=0}^{s-1} \\
&\cdot \left[1 + \frac{(-1)^k C_k}{s^{k-1}} \sum_{j=1}^{k} \frac{(-1)^{j-1} j^{k-1}}{(j-1)!(k-j)!(s+j)} \right. \\
&\cdot \left. \left(\frac{j-s}{2j}\right)^{s+k-1-i} \right] \frac{(\lambda b_{k+2})^i}{i!}, \quad 0 \leqslant k \leqslant s-1.
\end{aligned}
$$

Then the last probability calculated is

$$
\begin{aligned}
P_{2s-1} =& 1 + \frac{(-1)^s (2s-1)!}{s^{s-2} s!} \sum_{j=1}^{s-1} \frac{(-1)^j j^{s-2}}{(j-1)!(s-1-j)!(s+j)} \\
&\cdot \left(\frac{j-s}{2j}\right)^{2s-2} e^{-\frac{s+j}{s-j}\lambda a_{2s+1}} - e^{-\lambda a_{2s+1}} \\
&\cdot \sum_{i=0}^{s-1} \left[1 + \frac{(-1)^s (2s-1)!}{s^{s-2} s!} \sum_{j=1}^{s-1} \right. \\
&\cdot \left. \frac{(-1)^j j^{s-2}}{(j-1)!(s-1-j)!(s+j)} \left(\frac{j-2}{2j}\right)^{2s-2-i} \right] \frac{(\lambda a_{2s+1})^i}{i!}.
\end{aligned}
$$

Proceeding to the next state, b_{s+2} is undefined, but it is easy to check that $P_{2s-1} = P_{2s}$, because P_{2s-1} is independent of a_{2s} and the integration region in (1) is the support of Y_{2s}.

$$
\begin{aligned}
P_{2s+1} =& P\{M\} = P_m(n) = (2s+1) \int_0^{\infty} e^{-(2s+1)\lambda t} \, \mathrm{d}t \\
&+ \frac{(2s+1)(2s-1)!}{s! s^{s-2}} \sum_{j=1}^{s-1} \frac{(-1)^{s+j} j^{s-2}}{(j-1)!(s-1-j)!(s+j)}
\end{aligned}
$$

$$\times \left(\frac{j-s}{2j}\right)^{2s-2} \cdot \mathcal{I}_1 - (2s+1) \sum_{i=0}^{s-1} \left[1 + \frac{(2s-1)!}{s^{s-2}s!} \sum_{j=1}^{s-1} \right.$$
$$\left. \times \frac{(-1)^{s+j} j^{s-2}}{(j-1)!(s-1-j)!(s+j)} \cdot \left(\frac{j-s}{2j}\right)^{2s-2-i} \right] \mathcal{I}_2,$$

where

$$\mathcal{I}_1 = \int_0^\infty e^{-\lambda t \left(2s+1+\frac{s+j}{s-j}\right)} \, \mathrm{d}t = \frac{s-j}{2s(s-j+1)\lambda},$$
$$\mathcal{I}_2 = \int_0^\infty e^{-\lambda t(2s+2)} \frac{(\lambda t)^i}{i!} \, \mathrm{d}t = \frac{1}{(2(s+1))^{i+1}}.$$

Then we obtain:

$$P_m(n) = \frac{1}{2^s(s+1)^s} \left[1 + \frac{(2s+1)!}{2^s s^s s!} \right.$$
$$\left. \cdot \sum_{j=1}^{s-1} \frac{(-1)^j (s-j)^s}{(j-1)!(s-j+1)!(s+j)} \right].$$

The last expression can finally be rewritten as:

$$P_m(n)$$
$$= \frac{1}{2^s(s+1)^s} \left[1 - \frac{\binom{2s+1}{s}(s+1)}{(2s)^s} \sum_{i=0}^{s-1} \frac{(s-i-1)^s(-1)^i}{s+i+1} \binom{s}{i} \right].$$

□

REFERENCES

Berend, D. and Harmse, J.E. (1993), ‘Expert rule versus majority rule under partial information’, *Theory and Decision* 35: 179–197.

Berend, D. and Paroush, J. (1996), ‘Sufficient conditions for Condorcet’s Jury Theorem’, *Social Choice and Welfare* (forthcoming).

Berg, S. (1993), ‘Condorcet’s Jury Theorem revisited’, *European Journal of Political Economy* 9: 437–446.

Condorcet, N.C. de (1785), *Essai sur l’Application de l’Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, Paris.

Grofman, B., Owen, G. and Feld, S. (1983), ‘Thirteen theorems in search of the truth’, *Theory and Decision* 15: 261–278.
Ladha, K. (1995), ‘Information pooling through majority-rule voting Condorcet’s jury theorem with correlated votes’, *Journal of Economics Behavior and Organization* 26: 358–372.
Nitzan, S. and Paroush, J. (1982), ‘Optimal decision rules in uncertain dichotomous choice situations’, *International Economic Review* 23: 289–297.
Nitzan, S. and Paroush, J. (1984), ‘Partial information on decisional competence and the desirability of the expert rule in uncertain dichotomous choice situations’, *Theory and Decision* 17: 275–286.
Nitzan, S. and Paroush, J. (1985, *Collective Decision Making*. Cambridge: Cambridge University Press.
Paroush, J. (1996), ‘Stay away from fair coins or the correct Condorcet’s Jury Theorem’, *Social Choice and Welfare*, forthcoming.
Shapley, L. and Grofman, B. (1984), ‘Optimizing group judgmental accuracy in the presence of interdependencies’, *Public Choice* 43: 329–343.

Address for correspondence: L. Sapir, Department of Mathematics and Computer Science, Ben-Gurion University, P.O. Box 653, ISR-84105 Beer-Sheva, Israel
Phone: +972 7 441-883; Fax: +972 7 472-910