Acta Applicandae Mathematicae **69:** 141–162, 2001.

Optimality of the Expert Rule Under Partial Information

DANIEL BEREND[1] and LUBA SAPIR[2]*
[1] *Departments of Mathematics and of Computer Science, Ben-Gurion University, Beer-Sheva, 84105, Israel*
[2] *Department of Mathematics, Ben-Gurion University, Beer-Sheva, 84105, Israel*

(Received: 17 October 2000; in final form: 26 July 2001)

Abstract. We study the uncertain dichotomous choice model. In this model a group of decision makers is required to select one of two alternatives. The applications of this model are relevant to a wide variety of areas, such as medicine, management and banking. The decision rule may be the simple majority rule; however, it is also possible to assign more weight to the opinion of members known to be more qualified. The extreme example of such a rule is the expert decision rule. We are concerned with the probability of the expert rule to be optimal. Our purpose is to investigate the behaviour of this probability as a function of the group size for several rather general types of distributions. One such family of distributions is that where the density function of the correctness probability is a polynomial (on the interval $[1/2, 1]$). Our main result is an explicit formula for the probability in question. This contains formerly known results as very special cases.

Mathematics Subject Classifications (2000): 91B06, 91B12, 90B50.

Key words: dichotomous, choice, model, experts, decision rule, expert rule, optimality, probability, expertise, partial information.

1. Introduction

1.1. BACKGROUND

In various areas, such as politics, company management, medicine, and business, decisions are taken by voting. A group composed of several experts in the specific fields wishes to make correct decisions. We assume that there is an objectively correct decision, and all decision makers share the goal of identifying it. The algorithm for passing from all individual opinions to a group decision may be very distinct in various systems; for example, it may consist of a multi-stage process (Boland, 1989; Berg, 1997; Berg and Paroush, 1998).

We regard the dichotomous choice model, which goes back as far as Condorcet (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

* Research partially supported by an Eshkol Fellowship from the Israeli Science Ministry.

becomes large. Moreover, the probability of a correct decision increases to 1 as the number of individuals in the group tends to infinity. ‘Condorcet Jury Theorem’ usually refers to 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 of the members to make the right choice are equal and exceed $1/2$. Researchers have extended and completed Condorcet’s work in two directions: by relaxing the assumption of independence (Shapley and Grofman, 1984; Boland, 1989; Berg, 1993; Ladha, 1992, 1995) on the one hand, and by relaxing the assumption that individuals are homogeneous with respect to their decisional abilities (Grofman *et al*., 1983; Miller, 1986; Young, 1989; Paroush, 1998; Berend and Paroush, 1998) on the other hand.

In another direction, one may ask about identifying the optimal decision rule. Here a *decision rule* is a rule for translating the opinions of the various experts into a group decision. Such a rule is *optimal* if and only if it maximizes the probability that the group will reach the correct decision for all possible combinations of opinions. 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. In the case of symmetric alternatives (i.e. equi-probable *apriori* 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 (Nitzan and Paroush, 1982, 1984a, 1985; Grofman *et al*., 1983).

However, the assumption of full information regarding the decision makers competences 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, there exists some information 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 (1985), who obtained for this case an explicit expression for the probability $P_e(n)$ of the expert rule to be optimal. Nitzan and Paroush (1984b, 1985), considered the case of uniform distribution on $[1/2, 1]$. One of the primary goals was to compare the probability of the expert rule to be optimal with that of the majority rule. Nitzan and Paroush (1984b, 1985), using a Monte Carlo method, estimated the probability $P_e(n)$, as well as the probability $P_m(n)$ of the majority rule to be optimal, for small n’s. As one would expect, this probability decreases as a function of n. This line of the research was continued by Berend and Harmse (1993), Sapir (1998), Berend and Sapir (2001). Berend and Harmse (1993) obtained an asymptotic formula for $P_e(n)$ and an upper bound for $P_m(n)$ in case where the probabilities of being right for each expert are uniformly distributed on $[1/2, 1]$. These two results combined imply that the latter probability decays to 0 much more quickly than the former. They also obtain some generalizations of those results, which showed that 'slight deviations' from the two extreme rules are also unlikely to produce an optimal rule. Berend and Sapir (2001) generalized some of the results of Berend and Harmse (1993) under the assumption of generalized uniform distribution, which contains the uniform distribution as a particular case. In (Sapir, 1998) the assumption of the correctness probabilities being uniformly distributed in $[1/2, 1]$ was changed in such a way that the logarithmic expertise levels became distributed exponentially. In this case both $P_{\mathrm{e}}(n)$ and $P_{\mathrm{m}}(n)$ were calculated explicitly. The comparison of $P_{\mathrm{e}}(n)$ and $P_{\mathrm{m}}(n)$ in each of the situations, regarded by Berend and Harmse (1993), Sapir (1998), Berend and Sapir (2001).

The motivation for this paper starts with Sapir (1998). It was found there that

$$P_{\mathrm{e}}(n) = n \cdot \left(\frac{1}{2}\right)^{n-1}$$

under the assumption of exponential distribution of logarithmic expertise. It is natural to ask in which other cases can $P_{\mathrm{e}}(n)$ be expressed in such a simple form, as a single term or a finite sum. In this paper we are concerned with a general family of distributions, for which the probability of the expert rule to be optimal can be expanded explicitly into a finite or an infinite sum. Note that this family contains a wide variety of distribution families, and in particular some distributions related to well known ones, such as the Beta distribution. One such example is a distribution family where the density function of the correctness probability is a polynomial (on the interval $[1/2, 1]$), which contains the uniform distribution as a special case. Another example is the family, where the distribution function of the correctness probability is a polynomial in $(1-x)/x$ (again, on the interval $[1/2, 1]$). It contains the cases of exponentially distributed logarithmic expertise level and Beta-distributed inverse expertise as very particular instances. For this family (of polynomials in $(1-x)/x$) a simple expression is obtained for $P_{\mathrm{e}}(n)$. In addition to the above, several other distributions are included as examples of the general family. We also investigate the asymptotic behaviour of $P_{\mathrm{e}}(n)$ as a function of the group size and explore the influence of the distribution parameters on this probability.

It is worthwhile emphasizing that the main question addressed in this paper is not what the probability of the expert rule is to provide the right decision. This probability is certainly quite high in most cases, and in fact converges to 1 as the number of experts increases. It is what could be called 'the average case'. $P_{\mathrm{e}}(n)$, our object of interest through this paper, is rather 'the worst case'. Namely, each decision rule has some cases in which it is natural to doubt its performance. For the expert rule, the most questionable situation is where the top expert is opposed by all the others. To claim that, in a specific case, the expert rule is optimal, is tantamount to asserting that we should indeed favor the opinion of the top expert. Consequently, knowing that $P_{\mathrm{e}}(n)$ is, say, small does not mean that the expert rule performs badly in general. Rather, it assists us in modifying the rule in borderline cases.

Subsection 1.2 is devoted to a more accurate description of our model. In Section 2 we present the main results and show a wide variety of examples, and in Section 3 – their proofs. Section 4 provides some concluding remarks and questions.

1.2. SETUP AND NOTATIONS

The group consists of n members, each of whom advocates one of two alternatives. The probability that the ith member will make the right choice is denoted by p_i. We always assume that $p_i > 1/2$ for all i. (Note that, at least in the case where the p_i's are known, this implies no loss of generality.) We also assume that the members' choices are independent. 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

$$\prod_{i \in A} \frac{p_i}{1 - p_i} > \prod_{i \in B} \frac{p_i}{1 - p_i} \tag{1.1}$$

(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 $\ln p/(1-p)$.

Now we assume p_i, $i = 1, 2, \ldots, n$, to be (independent) random variables distributed on $[1/2, 1]$ according to the same distribution function $G_p(x)$. The corresponding density function is denoted by $g_p(x)$. Then we can consider the probability of a decision rule to be optimal. Moreover, we assume the ranking of the members of the team is (at least partly) known. Thus, one can follow rules based on this ranking. The extremes are the expert rule (i.e., always following the most competent member of the group) and the simple majority rule. In this paper we deal almost exclusively with the expert rule.

Sometimes it will be more convenient to describe the situation in terms not of the p_i's but of other random variables, such as the *inverse expertises* $Y_i = (1 - p_i)/p_i$, or the logarithmic expertises. Denote by $G_Y(y)$ and $G_{\ln \frac{p}{1-p}}(x)$ the respective distribution functions, and by $g_Y(y)$ and $g_{\ln \frac{p}{1-p}}(x)$ the density functions. Later we use the following connections between the distribution functions

$$G_Y(y) = 1 - G_p\left(\frac{1}{1+y}\right), \quad y \in [0, 1) \tag{1.2}$$

and

$$G_Y(y) = 1 - G_{\ln \frac{p}{1-p}}(-\ln y), \quad y \in [0, 1), \tag{1.3}$$

where both functions vanish on the negative axis and are identically 1 for $x \geqslant 1$.

2. Main Results

We are concerned with the probability $P_e(n)$ of the expert rule to be optimal. One verifies easily that $P_e(n)$ is the probability that, in the case where the top expert disagrees with all the others, the top expert is more likely to be correct than the rest. In this paper we regard a general family of distributions, for which the probability of the expert rule to be optimal can be expanded explicitly into a finite or infinite sum. The results obtained here contain some previous results as very particular instances.

2.1. EXPANDABLE $P_e(n)$

Suppose the inverse expertise levels of the committee members are i.i.d. random variables, distributed on [0, 1] according to the same distribution function, given in the form

$$G_Y(y) = \begin{cases} 0, & y < 0, \\ \sum_{j=1}^{\infty} c_j y^j, & 0 \leqslant y < 1, \\ 1, & y \geqslant 1, \end{cases} \tag{2.1}$$

for suitable $c_1, c_2, \ldots$. (Note that we do not require the series $\sum_{j=1}^{\infty} c_j y^j$ to converge at 1, but we do require $G_Y(y)$ to be continuous at 1; see Example 6 infra.)

Obviously, for $G_Y(y)$ to be a distribution function on [0, 1] it is necessary that $c_1 \geqslant 0$. (It follows from the equality $G_Y^{*'}(0) = c_1$.) An exact determination of the conditions on the coefficients c_j, under which $G_Y(y)$ is indeed a distribution function, is not trivial. Suffice it to say that, if all these coefficients are nonnegative, then $G_Y(y)$ is a (continuous) distribution function if and only if $\sum_{j=1}^{\infty} c_j = 1$. In the following, we usually assume only that G_Y is a distribution function, the c_j's being of arbitrary signs.

For the next theorem it will be convenient to use the moments of the distribution:

$$I_j = E[Y^j] = \int_0^1 g_Y(y) y^j \, dy, \quad j = 1, 2, \ldots.$$

Obviously $1 > I_1 > I_2 > \cdots$.

THEOREM 1. *Suppose the inverse expertise levels are distributed according to G_Y. If one of the conditions*

(i) $\sum_{j=1}^{\infty} c_j = 1,$

or

(ii) $$\sum_{j=1}^{\infty} |c_j| I_j^{n-1} < \infty,$$

holds, then the probability of the expert rule to be optimal is

$$P_{\mathrm{e}}(n) = n \sum_{j=1}^{\infty} c_j I_j^{n-1}. \tag{2.2}$$

It may be worthwhile mentioning already at this point the reason for the moments' occurrence in the theorem (for more details see the proof). One finds easily (conditioning on the identity of the top expert) that $P_{\mathrm{e}}(n) = nE_{G_Y^{n-1}}(G(\prod_{i=1}^{n-1} y_j))$, the expectation being taken with respect to $n-1$ independent variables, each with distribution G_Y. Theorem 1 implies the following corollary, which gives the asymptotic behavior of $P_{\mathrm{e}}(n)$.

COROLLARY 1. *Under the assumptions of Theorem* 1

$$P_{\mathrm{e}}(n) = nc_{j_0} I_{j_0}^{n-1} \cdot (1 + \mathrm{o}(1)),$$

as $n \to \infty$, *where* c_{j_0} *is the first nonzero term in the sequence* $(c_j)_{j=1}^{\infty}$.

Note that, while some of c_j's may be negative, c_{j_0} must be positive for $G_Y(y)$ to be a distribution function. Also, if condition (ii) in the theorem holds for some n, it clearly holds for all $n > n_0$.

Remark 1. Under condition (i) of Theorem 1, the moments may be expressed explicitly in terms of the coefficients of G_Y:

$$I_j = \sum_{s=1}^{\infty} \frac{sc_s}{j+s}.$$

2.2. FINITELY EXPANDED $P_{\mathrm{e}}(n)$

The results of Section 2.1 assume a particularly simple form if $G_Y(y)$ (restricted to [0, 1]) is a polynomial:

$$G_Y(y) = \sum_{j=1}^{l} c_j y^j, \quad y \in [0, 1]. \tag{2.3}$$

Note that $G_Y(y)$, given by (2.3), is a distribution function if and only if $\sum_{j=1}^{l} c_j = 1$ and the corresponding density $g_Y^*(y) = G_Y^{*'}(y)$ is nonnegative on [0, 1]. The last condition can be checked by Sturm's algorithm (cf. Isaacson and Keller, 1966). Also note that the condition $\sum_{j=1}^{l} c_j = 1$ coincides with condition (i) of Theorem 1. Hence, by Remark 1:

$$I_j = \sum_{s=1}^{l} \frac{sc_s}{j+s}.$$

Theorem 1 reduces then to the following:

THEOREM 1$'$. *Suppose $G_Y(y)$, given by (2.3), is a distribution function. Then the probability of the expert rule to be optimal is:*

$$P_{\mathrm{e}}(n) = n \sum_{j=1}^{l} c_j \left(\sum_{s=1}^{l} \frac{sc_s}{j+s} \right)^{n-1}.$$

In particular, as $n \to \infty$,

$$P_{\mathrm{e}}(n) = nc_{j_0} \left(\sum_{s=1}^{l} \frac{sc_s}{j_0+s} \right)^{n-1} \cdot (1 + \mathrm{o}(1)),$$

where c_{j_0} is the first nonzero coefficient.

In fact, $P_{\mathrm{e}}(n)$ can be expressed in a simple form under the assumption of a more general distribution family than given by (2.3):

$$G_Y(y) = \begin{cases} 0, & y < 0, \\ \sum_{j=1}^{l} c_j y^{\alpha_j}, & 0 \leqslant y < 1, 0 < \alpha_1 < \alpha_2 < \cdots < \alpha_l, \\ 1, & y \geqslant 1. \end{cases} \tag{2.4}$$

Again, for $G_Y(y)$ to be a distribution function it is necessary that $c_1 \geqslant 0$ and $\sum_{j=1}^{l} c_j = 1$.

THEOREM 1$''$. *Suppose the function $G_Y(y)$, defined by (2.4), is a distribution function. Then the probability of the expert rule to be optimal is:*

$$P_{\mathrm{e}}(n) = n \sum_{j=1}^{l} c_j \left(\sum_{s=1}^{l} \frac{\alpha_s c_s}{\alpha_j + \alpha_s} \right)^{n-1}.$$

EXAMPLE 1. The case of $\mathrm{Exp}(\alpha)$ distributed logarithmic expertise is a particular instance of (2.4), obtained by taking $l = 1, c_1 = 1, \alpha_1 = \alpha$ (see (1.3)). Theorem 1$''$ yields $P_{\mathrm{e}}(n) = n \cdot (\frac{1}{2})^{n-1}$, which agrees with Theorem 1 of Sapir (1998).

Table I. Lower bound for $P_e(n)$.

l	1	3	10	100	1000
$P_e(n) \geqslant$	$n \cdot 0.5^{n-1}$	$\frac{n}{3} \cdot 0.64^{n-1}$	$\frac{n}{10} \cdot 0.8^{n-1}$	$\frac{n}{100} \cdot 0.96^{n-1}$	$\frac{n}{1000} \cdot 0.99^{n-1}$

EXAMPLE 2. If $G_Y(y)$ is as in 2.3 for $l = 2$, it turns out that $G_Y(y)$ forms a distribution function if and only if $0 \leqslant c_1 \leqslant 2$. Then $c_2 = 1 - c_1$, and Theorem 1′ gives:

$$P_e(n) = n\left(c_1\left(\frac{4-c_1}{6}\right)^{n-1} + (1-c_1)\left(\frac{3-c_1}{6}\right)^{n-1}\right), \quad c_1 \in [0, 2].$$

EXAMPLE 3. Let $c_1 = c_2 = \cdots = c_l = 1/l$. Then:

$$P_e(n) = \frac{n}{l^n}\sum_{j=1}^{l}\left(\sum_{s=1}^{l}\frac{s}{j+s}\right)^{n-1}. \tag{2.6}$$

Table I provides the lower bound for $P_e(n)$ obtained by taking just the first term (corresponding to $j = 1$) on the right-hand side of (2.6).

According to (2.6) (and as follows in particular from the last part of Theorem 1′), $P_e(n)$ (considered as a function of n) behaves asymptotically as $\frac{n}{l}\theta^{n-1}$ for some constant $\theta = \theta_l$. The table indicates that $\theta = \theta_l \underset{l\to\infty}{\longrightarrow} 1$. Indeed, this is the interesting point about this example, formalized by

PROPOSITION 1. *Under the assumptions of Example* 3*,*

$$P_e(n) \geqslant \frac{n}{l}\theta_l^{n-1}$$

for appropriate constants θ_l *satisfying* $\theta_l \underset{l\to\infty}{\longrightarrow} 1$.

EXAMPLE 4. Let $c_j = (-1)^{j-1}\binom{l}{j}$ for $1 \leqslant j \leqslant l$. Note that this sequence arises from the following simple density function:

$$g_Y(y) = \begin{cases} l(1-y)^{l-1}, & 0 \leqslant y \leqslant 1, \\ 0, & \text{otherwise.} \end{cases}$$

PROPOSITION 2. *Under the assumptions of Example* 4*,*

$$P_e(n) = n\sum_{j=1}^{l}(-1)^{j-1}\binom{l}{j}\frac{1}{\binom{l+j}{l}^{n-1}}. \tag{2.7}$$

In particular, $P_e(n) \leqslant nl\theta_l^{n-1}$ *for appropriate constants* θ_l *satisfying* $\theta_l \underset{l\to\infty}{\longrightarrow} 0$.

Propositions 1 and 2, taken together, show that the spectrum of possible types of asymptotic behavior of $P_{\mathrm{e}}(n)$ is quite wide.

Note that inverse expertise levels in Example 4 arise from the special case of Beta distribution with parameters 1 and l. Recall that, in general, X is Beta-distributed, say $X \sim \beta(v, w)$ (with $v,\ w > 0$), if its density function is

$$g_X(x) = \begin{cases} \dfrac{x^{v-1}(1-x)^{w-1}}{B(v, w)}, & 0 \leqslant x \leqslant 1, \\ 0, & \text{otherwise}, \end{cases}$$

where $B(v, w) = \frac{\Gamma(v)\Gamma(w)}{\Gamma(v+w)}$ is the beta function. The next example generalizes Example 4 as follows:

EXAMPLE 5. Suppose the inverse expertise levels are distributed $\beta(v, w)$ with integer parameters $v,\ w$. The distribution function can be expanded in the form (2.3) with

$$c_1 = c_2 = \cdots = c_{v-1} = 0,$$
$$c_j = (-1)^{v+j}\binom{v+w-1}{j}\binom{j-1}{v-1}, \quad v \leqslant j \leqslant v+w-1.$$

Similarly to Proposition 2, one can show that

PROPOSITION 3. *Under the assumptions of Example* 5*,*

$$P_{\mathrm{e}}(n) = nw\binom{v+w-1}{w}^n \cdot \sum_{j=0}^{w-1} \frac{(-1)^j\binom{w-1}{j}}{(j+v)\binom{2v+w+j-1}{w}^{n-1}},$$

and in particular, as $n \to \infty$ we have

$$P_{\mathrm{e}}(n) = \frac{n}{vB(v, w)}\left(\frac{B(2v, w)}{B(v, w)}\right)^{n-1}(1 + \mathrm{o}(1)).$$

2.3. INFINITELY EXPANDED $P_{\mathrm{e}}(n)$

The motivation for this section starts with Berend and Harmse (1993), who considered the case of uniform distribution of correctness probabilities. They obtained an expression for $P_{\mathrm{e}}(n)$ in terms of an infinite series, which yields arbitrarily good estimates for $P_{\mathrm{e}}(n)$. This section illustrates examples of families of distributions, given by (2.1). The case of a uniform distribution is one of the special cases presented in Example 6.

EXAMPLE 6. We start with the case where the probabilities $p_i,\ \ i = 1, 2, \ldots, n$, are distributed with polynomial density (restricted to the interval [1/2, 1]) $g_p(x)$ of degree l. By (1.2) it is readily verified that an equivalent condition to p_i having the above distribution is for $Y_i = (1 - p_i)/p_i$ to have the following distribution function on [0, 1]:

$$G_Y(y) = \sum_{j=0}^{l} \frac{g_p^{(j)}(0)}{(j+1)!}\big(1 - (1+y)^{-(j+1)}\big), \quad 0 \leqslant y \leqslant 1. \tag{2.8}$$

As $(1 + y)^{-(j+1)}$ can be expanded into a *binomial series* for $|y| < 1$, $G_Y(y)$ can be expanded into a power series on [0, 1). (Note that the expansion is invalid at the point $y = 1$. This explains why the expansion in (2.1) was assumed only in the half-open interval [0, 1).) Since $G_Y(1) = 1$, the function $G_Y^*(y)$ is of the form (2.1), with

$$c_j = (-1)^{j-1} \cdot \sum_{i=0}^{l} \frac{g_p^{(i)}(0)}{(i+1)!} \binom{i+j}{j}, \quad j = 1,\ 2, \ldots .$$

Note that as $j \to \infty$ we have:

$$c_j = \mathrm{O}(j^l). \tag{2.9}$$

The moments may be bounded as follows:

$$\begin{aligned} I_j \ &= \int_0^1 y^j g_Y(y)\,\mathrm{d}y = \int_{\frac{1}{2}}^1 \left(\frac{1-x}{x}\right)^j g_p(x)\,\mathrm{d}x \\ &\leqslant \max_{x\in[\frac{1}{2},1]} g_p(x) \cdot \int_{\frac{1}{2}}^1 \left(\frac{1-x}{x}\right)^j \mathrm{d}x. \end{aligned}$$

Putting $C = \max_{x\in[\frac{1}{2},1]} g_p(x)$, and substituting $t = 1/x$, we obtain

$$I_j \leqslant C \int_1^2 \frac{(y-1)^j}{y^2}\,\mathrm{d}y \leqslant C \int_1^2 (y-1)^j\,\mathrm{d}y = \frac{C}{j+1}. \tag{2.10}$$

By (2.9) and (2.10) we have

$$|c_j| I_j^{n-1} = \mathrm{O}\left(\frac{1}{j^{n-1-l}}\right).$$

Thus, for $n \geqslant l+3$ the series $\sum_{j=1}^{\infty} c_j I_j^{n-1}$ converges absolutely, so that condition (ii) of Theorem 1 holds. In this case, Theorem 1 reduces to:

THEOREM 2. *Let the density $g_p(x)$ be a polynomial of degree l. Then for $n \geqslant l+3$ the probability $P_{\mathrm{e}}(n)$ of the expert rule to be optimal is:*

$$P_{\mathrm{e}}(n) = n \sum_{j=1}^{\infty} c_j I_j^{n-1}, \tag{2.11}$$

where

$$c_j = (-1)^{j-1} \sum_{i=0}^{l} \frac{g_p^{(i)}(0)}{(i+1)!} \binom{i+j}{j}, \quad j = 1, 2, \ldots,$$

$$I_j = \int_{1/2}^{1} \left(\frac{1-x}{x}\right)^j g_p(x)\, \mathrm{d}x, \quad j = 1, 2, \ldots.$$

Similarly to Remark 1 we obtain:

Remark 2. I_j can be calculated explicitly by the following formula:

$$I_j = \sum_{i=0}^{l} \frac{g_p^{(i)}(0)}{i!} \left(\sum_{s=0,\, s\neq j-i-1}^{j} \frac{(-1)^s}{s-j+i+1} \binom{j}{s} \left(1 - \frac{1}{2^{i-j+s+1}}\right)\right) +$$
$$+ \ln 2 \sum_{i=0}^{l} \frac{(-1)^{j-i-1} g_p^{(i)}(0)}{i!} \binom{j}{i+1}, \quad j = 1, 2, \ldots.$$

Note that, by (2.9), the series $\sum_{j=1}^{\infty} c_j$ diverges, which means that condition (i) of Theorem 1 is not fulfilled in this case.

Recall that an infinite series is a *Leibniz-type series* if its terms (weakly) decrease to 0 in absolute value and have alternating signs from some place on (cf. Fikhtengol'ts, 1965, Sec. 244).

PROPOSITION 4. *The series on the right-hand side of* (2.11) *in Theorem* 2 *is a Leibniz-type series.*

The proposition is of practical value; it enables us to obtain arbitrarily good estimates for $P_{\mathrm{e}}(n)$ by bounding it between consecutive partial sums of the infinite series.

Next we consider a few specific examples of polynomial densities, which demonstrate some fine points in Theorem 2 and Proposition 4.

(a) For uniform distribution, Theorem 2 reduces to the results of Berend and Harmse (1993), where $c_j = 2 \cdot (-1)^{j-1}$ for any j, and $P_{\mathrm{e}}(n)$ is expressed by means of an infinite Leibniz-type series, thus yielding arbitrarily good asymptotics for $P_{\mathrm{e}}(n)$ as $n \to \infty$. In particular

$$2n(2\ln 2 - 1)^{n-1} - 2n(3 - 4\ln 2)^{n-1} \leqslant P_{\mathrm{e}}(n) \leqslant 2n(2\ln 2 - 1)^{n-1}.$$

(b) Berend and Sapir (2001) considered the case of p_i being distributed according to the distribution function

$$G_p(x) = \begin{cases} 0, & x < \frac{1}{2}, \\ (2x-1)^{\alpha}, & \frac{1}{2} \leqslant x \leqslant 1, \\ 1, & \text{otherwise}, \end{cases} \tag{2.12}$$

where α is any positive parameter. For $\alpha = 1$ it reduces to the uniform distribution $U[1/2, 1]$. If $\alpha = m$ is a positive integer, (2.12) produces a polynomial density function of degree $m-1$. In this case Theorem 2 and Proposition 4 provide the expansion of $P_{\mathrm{e}}(n)$ into an infinite Leibniz-type series, with

$$c_j = (-1)^{j-1} \sum_{k=0}^{m} \binom{m}{k} \binom{j-1+m-k}{m-1}, \quad j = 1, 2, \ldots .$$

Note that in this case we have $|c_j| I_j^{n-1} = \mathrm{O}(1/j^{m(n-2)+1})$, and therefore condition (ii) of Theorem 1 holds for any n ($\geqslant 3$). Corollary 1 produces the asymptotic behaviour of $P_{\mathrm{e}}(n)$ as n increases:

$$P_{\mathrm{e}}(n) = 2mnI_1^{n-1}(1 + \mathrm{o}(1)).$$

(c) In general, the condition $n \geqslant l+3$ in Theorem 2 cannot be discarded; take, for example,

$$g_p(x) = \begin{cases} \dfrac{24}{7}x^2, & x \in [\frac{1}{2}, 1], \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$c_k = (-1)^{k-1} \cdot \tfrac{4}{7}(k+1)(k+2)$$

and

$$\begin{aligned} I_k &= \int_{1/2}^{1} \left(\frac{1-x}{x}\right)^k \cdot \frac{24}{7}x^2\,\mathrm{d}x \geqslant \frac{24}{7} \int_1^2 \frac{(t-1)^k}{t^4}\,\mathrm{d}t \\ &\geqslant \frac{3}{14} \int_1^2 (t-1)^k\,\mathrm{d}t \geqslant \frac{1}{10k}. \end{aligned}$$

Thus for $n = 3$ the general term $c_k I_k^2$ in the series on the right-hand side of (2.11) does not even converge to 0, so the series certainly diverges. By the way, an explicit calculation of $P_{\mathrm{e}}(3)$ in this case (using symbolic integration in Matlab) yields:

$$P_{\mathrm{e}}(3) = \frac{1152}{343} \ln 2 + \frac{135}{343}\pi^2 - \frac{9657}{1715} \approx 0.5816.$$

(d) The series in (2.11) is not necessarily alternating from the beginning. Taking

$$g_p(x) = \begin{cases} -4x+5, & x \in [\frac{1}{2}, 1], \\ 0, & \text{otherwise,} \end{cases}$$

we have $c_k = (-1)^k(2k-3)$. In particular, $c_1 = 1,\ c_2 = 1$, and the series in (2.11) alternates only from the second place on. Note that, as Theorem 2 stands, (2.11) is applicable in our case only for $n \geqslant 4$. However, one can check (again, using symbolic integration in Matlab) that in this case

$$P_{\mathrm{e}}(3) = -\frac{117}{2} \ln 2 - \frac{17}{16}\pi^2 + \frac{207}{4} \approx 0.7144$$

and numerical calculations hint that the series on the right-hand side of (2.11) converges (conditionally) to the same value.

Note that the second condition of Theorem 1 does not imply the first condition, as seen by Example 6. In the other direction, if (i) is replaced by the stronger condition $\sum_{j=1}^{\infty} |c_j| < \infty$ or if we add the requirement that $g_Y(1) < \infty$, then condition (ii) holds as well. In general, though, (i) does not yield (ii), as the following example shows:

EXAMPLE 7. Let $m \geqslant 1$ be an integer, $c_1 = 1$ and

$$c_{2i} = -c_{2i+1} = -\prod_{k=1}^{i}\left(1 - \frac{1}{km}\right), \quad i = 1, 2, \ldots .$$

This sequence arises from the distribution function:

$$G_Y(y) = 1 - \left(\frac{1-y}{(1+y)^{m-1}}\right)^{1/m}, \quad 0 \leqslant y \leqslant 1 .$$

Using the estimate $1 - x \approx \mathrm{e}^{-x}$ for $x \approx 0$ we see that $|c_{2i}| = c_{2i+1}$ behaves (up to a multiplicative constant) as $j^{-1/m}$, and in particular $c_j \underset{j\to\infty}{\longrightarrow} 0$. Obviously, condition (i) of Theorem 1 holds. The theorem then gives:

$$P_{\mathrm{e}}(n) = n\left(I_1^{n-1} - \sum_{i=1}^{\infty} \frac{\prod_{k=1}^{i}(km-1)}{m^i i!} \cdot \left(I_{2i}^{n-1} - I_{2i+1}^{n-1}\right)\right).$$

Now we claim that condition (ii) of Theorem 1 is not fulfilled for any $n \leqslant m$. Indeed

$$\frac{1}{2m} B\left(\frac{1}{m}, i+1\right) \leqslant I_{2i+1} \leqslant \tfrac{1}{2} B\left(\frac{1}{m}, i+1\right),$$

so that I_j behaves as

$$\frac{\Gamma(1/m)\Gamma(i+1)}{\Gamma(i+1+1/m)} \approx C i^{-1/m}$$

for some constant $C = C(m)$ (cf. Abramowitz and Stegun, 1968). Hence the series $\sum_{j=1}^{\infty} c_j I_j^{n-1}$ behaves as $\sum_{j=1}^{\infty} j^{-n/m}$, which diverges for $n \leqslant m$. Thus, in general, condition (ii) does not follow from condition (i).

It is interesting that Theorem 1 also improves some of the results of Berend and Sapir (2001). In particular, we show in the next example that it provides an expansion of $P_{\mathrm{e}}(n)$ into an infinite series for generalized uniform distribution for $\alpha < 1$.

EXAMPLE 8. Let $G_p(x)$ be as in (refe2.12), with $\alpha < 1$. By (1.2):

$$G_Y(y) = 1 - (1-y)^{\alpha} \cdot (1+y)^{-\alpha}, \quad 0 \leqslant y \leqslant 1. \tag{2.13}$$

Since $\alpha < 1$, (2.13) can be expanded on [0, 1] in the form

$$G_Y(y) = \sum_{j=1}^{\infty} c_j y^j,$$

where

$$c_j = (-1)^{j-1} \sum_{i=0}^{j} \binom{\alpha}{i} \binom{\alpha + j - i - 1}{j - i}. \tag{2.14}$$

The expansion of $G_Y(y)$ is valid also at $y = 1$, so that $\sum_{j=1}^{\infty} c_j = 1$. Thus condition (i) of Theorem 1 takes place and we obtain the expansion

$$P_{\mathrm{e}}(n) = \sum_{j=1}^{\infty} c_j I_j^{n-1},$$

where

$$I_j = 2\alpha \int_0^1 \frac{t^j}{(1-t)^{1-\alpha}(1+t)^{\alpha+1}}\, \mathrm{d}t, \quad j = 1, 2, \ldots,$$

and c_j is given by (2.14).

We mention in passing that in the case $\alpha = 1/2$ the c_j's assume the simple form

$$c_1 = 1,$$
$$c_{2i+1} = -c_{2i} = \frac{(2i-1)!!}{(2i)!!}, \quad i = 1, 2, \ldots.$$

and the I_j's may be expressed using Wallis's formula (cf. Fikhtengol'ts, 1965, Sec. 188) in closed form:

$$I_j = \begin{cases} \dfrac{(2i)!!}{(2i-1)!!} - \dfrac{(2i-1)!!}{(2i-2)!!} \cdot \dfrac{\pi}{2}, & j = 2i, \\ -\dfrac{(2i-2)!!}{(2i-3)!!} + \dfrac{(2i-1)!!}{(2i-2)!!} \cdot \dfrac{\pi}{2}, & j = 2i-1. \end{cases} \quad i = 1, 2, \ldots,$$

(Recall that $(2i)!! = 2 \cdot 4 \cdot \cdots \cdot 2i$ and $(2i-1)!! = 1 \cdot 3 \cdot \cdots \cdot (2i-1)$.)

3. Proofs

Proof of Theorem 1. Consider the random variables

$$Y_i = \frac{1-p_i}{p_i}, \quad i = 1, 2, \ldots, n.$$

The density function of each Y_i is given by:

$$g_Y(y) = \begin{cases} \sum_{j=1}^{\infty} jc_j y^{j-1}, & 0 \leqslant y < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

Since the power series on the right-hand sides of (2.1) and (3.1) have the same the radius of convergence, which is at least 1, for any $\varepsilon > 0$ the two series converge uniformly to $G_Y^*(y)$ and $g_Y^*(y)$ on $[0, 1-\varepsilon]$. Therefore:

$$\begin{aligned} P_{\mathrm{e}}(n) &= \sum_{j=1}^{n} \mathrm{P}\left(Y_j < \prod_{i=1,\, i\neq j}^{n} Y_i\right) = n \cdot \mathrm{P}\left(Y_1 < \prod_{i=2}^{n} Y_i\right) \\ &= n \underbrace{\int \ldots \int}_{[0,1]^{n-1}} \prod_{i=1}^{n-1} g_Y(y_i) G_Y\left(\prod_{i=1}^{n-1} y_i\right) \mathrm{d}y_1\, \mathrm{d}y_2 \ldots \mathrm{d}y_{n-1} \\ &= n \underbrace{\int \ldots \int}_{[0,1]^{n-1}} \prod_{i=1}^{n-1} g_Y(y_i) \sum_{j=1}^{\infty} c_j \prod_{i=1}^{n-1} y_i^j\, \mathrm{d}y_1\, \mathrm{d}y_2 \ldots \mathrm{d}y_{n-1} \\ &= n \lim_{\varepsilon\to 0} \underbrace{\int \ldots \int}_{[0,1-\varepsilon]^{n-1}} \sum_{j=1}^{\infty} c_j \prod_{i=1}^{n-1} g_Y(y_i) y_i^j\, \mathrm{d}y_1\, \mathrm{d}y_2 \ldots \mathrm{d}y_{n-1}. \end{aligned}$$

Since the series $\sum_{j=1}^{\infty} c_j y^j$ converges uniformly to $G_Y^*(y)$ on $[0, 1-\varepsilon]$, the series $\sum_{j=1}^{\infty} c_j (\prod_{i=1}^{n-1} y_i)^j$ converges uniformly to $G_Y(\prod_{i=1}^{n-1} y_i)$ on $[0, 1-\varepsilon]^{n-1}$. Also, $\prod_{i=1}^{n-1} g_Y(y_i)$ is bounded on $[0, 1-\ \varepsilon]^{n-1}$, and therefore:

$$\begin{aligned} P_{\mathrm{e}}(n) &= n \lim_{\varepsilon\to 0} \sum_{j=1}^{\infty} c_j \underbrace{\int \ldots \int}_{[0,1-\varepsilon]^{n-1}} \prod_{i=1}^{n-1} g_Y(y_i) y_i^j\, \mathrm{d}y_1\, \mathrm{d}y_2 \ldots \mathrm{d}y_{n-1} \\ &= n \lim_{\varepsilon\to 0} \sum_{j=1}^{\infty} c_j \left(\int_0^{1-\varepsilon} g_Y(y) y^j\, \mathrm{d}y\right)^{n-1}. \end{aligned} \tag{3.2}$$

We would like to interchange the order of the summation and taking the limit on the right-hand side of (3.2). To this end, we need one of the two conditions (i) and (ii) in the statement of the theorem. Define functions v_j and u_j by:

$$\left.\begin{aligned} v_j(\varepsilon) &= \left(\int_0^{1-\varepsilon} g_Y(y) y^j\, \mathrm{d}y\right)^{n-1}, \\ u_j(\varepsilon) &= c_j v_j(\varepsilon). \end{aligned}\right. \quad j = 1, 2, \ldots,\ 0 \leqslant \varepsilon \leqslant 1,$$

Assume first that condition (i) takes place, namely $\sum_{j=1}^{\infty} c_j = 1$. Since for each fixed ε we have $v_j(\varepsilon) \underset{j\to\infty}{\longrightarrow} 0$, and the convergence is monotonic, we may invoke Abel's test to prove that the series $\sum_{j=1}^{\infty} u_j(\varepsilon)$ converges uniformly on [0, 1]. If condition (ii) of the theorem takes place, namely $\sum_{j=1}^{\infty} |c_j| I_j^{n-1} < \infty$, then by Weierstrass M-test we obtain the uniform convergence of $\sum_{j=1}^{\infty} u_j(\varepsilon)$ on [0, 1].

Consequently, if at least one of the conditions holds, then:

$$\begin{aligned} P_{\mathrm{e}}(n) &= n \sum_{j=1}^{\infty} c_j \left(\lim_{\varepsilon\to 0} \int_0^{1-\varepsilon} g_Y(y) y^j \,\mathrm{d}y \right)^{n-1} \\ &= n \sum_{j=1}^{\infty} c_j \left(\int_0^1 g_Y(y) y^j \,\mathrm{d}y \right)^{n-1} = n \sum_{j=1}^{\infty} c_j I_j^{n-1}. \end{aligned}$$

Proof of Remark 1. Condition (i) of Theorem 1 ensures that the series on the right-hand side of (2.1) converges uniformly on [0, 1]. Using integration by parts we obtain

$$\begin{aligned} I_j &= \int_0^1 g_Y(y) y^j \,\mathrm{d}y = G_Y(y) y^j \big|_0^1 - j \int_0^1 G_Y^*(y) y^{j-1} \,\mathrm{d}y \\ &= 1 - j \int_0^1 G_Y(y) y^{j-1} \,\mathrm{d}y = 1 - \int_0^1 j \sum_{s=1}^{\infty} c_s y^{s+j-1} \,\mathrm{d}y \\ &= 1 - \sum_{s=1}^{\infty} \frac{j c_s}{j+s} = \sum_{s=1}^{\infty} \frac{s c_s}{j+s}, \end{aligned}$$

which is the required result.

Proof of Corollary 1. If condition (ii) of Theorem 1 holds, then the corollary is trivial. Assume that condition (i) is fulfilled. Put

$$B_m = \sum_{j=j_0}^{m} c_j, \qquad S_m = \sum_{j=j_0}^{m} c_j I_j^{n-1}, \qquad S = \sum_{j=j_0}^{\infty} c_j I_j^{n-1}.$$

As $c_j = B_j - B_{j-1}$, using Abel's transformation we have:

$$S_m = B_m I_m^{n-1} + \sum_{j=j_0}^{m-1} B_j \left(I_j^{n-1} - I_{j+1}^{n-1} \right).$$

Since $B_m \underset{m\to\infty}{\longrightarrow} 1$, we can find M such that $1/2 \leqslant B_m \leqslant 3/2$ for $m \geqslant M$. Then:

$$S_m = B_m I_m^{n-1} + \sum_{j=j_0}^{M-1} B_j \left(I_j^{n-1} - I_{j+1}^{n-1} \right) + \sum_{j=M}^{m-1} B_j \left(I_j^{n-1} - I_{j+1}^{n-1} \right).$$

Clearly, $B_m I_m^{n-1} > 0$ and $\sum_{j=M}^{m-1} B_j(I_j^{n-1} - I_{j+1}^{n-1}) > 0$, so that

$$\tfrac{1}{2} \cdot I_M^{n-1} + \sum_{j=j_0}^{M-1} B_j\big(I_j^{n-1} - I_{j+1}^{n-1}\big) \leqslant S_m \leqslant \tfrac{3}{2} \cdot I_M^{n-1} + \sum_{j=j_0}^{M-1} B_j\big(I_j^{n-1} - I_{j+1}^{n-1}\big).$$

Passing to the limit with respect to m, we obtain:

$$\tfrac{1}{2} \cdot I_M^{n-1} + \sum_{j=j_0}^{M-1} B_j\big(I_j^{n-1} - I_{j+1}^{n-1}\big) \leqslant S \leqslant \tfrac{3}{2} \cdot I_M^{n-1} + \sum_{j=j_0}^{M-1} B_j\big(I_j^{n-1} - I_{j+1}^{n-1}\big).$$

Rewriting these inequalities in terms of the c_j's, we obtain

$$\left(\frac{1}{2} - \sum_{j=j_0}^{M-1} c_j\right) I_M^{n-1} + \sum_{j=j_0}^{M-1} c_j I_j^{n-1} \leqslant S \leqslant \left(\frac{3}{2} - \sum_{j=j_0}^{M-1} c_j\right) I_M^{n-1} + \sum_{j=j_0}^{M-1} c_j I_j^{n-1},$$

which implies:

$$S = c_{j_0} I_{j_0}^{n-1} \cdot (1 + \mathrm{o}(1)) = c_{j_0} \left(\sum_{s=1}^{\infty} \frac{s c_s}{j_0 + s} \right)^{n-1} \cdot (1 + \mathrm{o}(1)).$$

This proves the corollary.

Proof of Proposition 1. Denote:

$$\theta_l = I_1(l) = \frac{1}{l} \sum_{s=1}^{l} \frac{s}{1+s} = \left(\frac{1}{2} + \frac{2}{3} + \cdots + \frac{l}{l+1}\right) \frac{1}{l}.$$

Thus, $(\theta_l)_{l=1}^{\infty}$ is the sequence of arithmetic means of the sequence $(s/(s+1))_{s=1}^{\infty}$. Since the latter converges to 1, so does the former, so that $\theta = \theta_l \underset{l\to\infty}{\longrightarrow} 1$. By (2.6)

$$P_{\mathrm{e}}(n) \geqslant \frac{n}{l} \left(\frac{1}{l} \sum_{s=1}^{l} \frac{s}{1+s} \right)^{n-1} = \frac{n}{l} \theta_l^{n-1},$$

which proves the proposition.

Proof of Proposition 2. Recall that the jth moment of a $\beta(v, w)$ distributed random variable is $\prod_{k=0}^{j-1}(v+k)/(v+w+k)$ (cf. Hastings and Peacock, 1975). In particular, since $Y_i \sim \beta(1, l)$,

$$I_j = \prod_{k=0}^{j-1} \frac{1+k}{1+l+k} = \frac{j!\,l!}{(j+l)!} = \frac{1}{\binom{l+j}{l}}.$$

By Theorem 1, this yields (2.7).

It is easy to verify that the series on the right-hand side of (2.7) is Leibniz type. Moreover, its terms decrease in absolute value already from the beginning. In particular, using the first two terms, we obtain

$$\frac{nl}{(l+1)^{n-1}}\left(1-\frac{(l-1)}{2}\left(\frac{2}{l+2}\right)^{n-1}\right) \leqslant P_{\mathrm{e}}(n) \leqslant \frac{nl}{(l+1)^{n-1}}.$$

Proof of Remark 2. A routine calculation gives:

$$\begin{aligned}
I_j &= \int_{\frac{1}{2}}^{1}\left(\frac{1-x}{x}\right)^j \sum_{i=0}^{l}\frac{g_p^{(i)}(0)}{i!}\cdot x^i\,\mathrm{d}x \\
&= \int_{\frac{1}{2}}^{1}\sum_{s=0}^{j}(-1)^s\binom{j}{s}x^{-(j-s)}\sum_{i=0}^{l}\frac{g_p^{(i)}(0)}{i!}\cdot x^i\,\mathrm{d}x \\
&= \sum_{i=0}^{l}\frac{g_p^{(i)}(0)}{i!}\sum_{s=0}^{j}(-1)^s\binom{j}{s}\int_{\frac{1}{2}}^{1}x^{-j+s+i}\,\mathrm{d}x \\
&= \sum_{i=0}^{l}\frac{g_p^{(i)}(0)}{i!}\left(\sum_{s=0,\ s\neq j-i-1}^{j}\frac{(-1)^s}{s-j+i+1}\binom{j}{s}\left(1-\frac{1}{2^{i-j+s+1}}\right)\right)+ \\
&\quad +\ln 2\sum_{i=0}^{l}\frac{(-1)^{j-i-1}g_p^{(i)}(0)}{i!}\binom{j}{i+1}, \qquad j=1,2,\ldots.
\end{aligned}$$

Proof of Proposition 4. The fact that the sequence $c_k I_k^{n-1}$ converges to 0 is contained in Theorem 2. The sequence c_k, considered as a function of l, is (except for sign) some polynomial of degree l, say $c_k = (-1)^{k-1}(b_0 + b_1 k + \cdots + b_l k^l)$. Hence for sufficiently large k the terms in the sum on the right-hand side of (2.11) are of alternating signs. It remains to show that $|c_k| I_k^{n-1}$ decreases from some place on. First, note that

$$\begin{aligned}
\frac{|c_{k+1}|}{|c_k|} &= \frac{b_0 + b_1(k+1) + \cdots + b_l(k+1)^l}{b_0 + b_1 k + \cdots + b_l k^l} \\
&= \frac{b_l k^l + (l b_l + b_{l-1})k^{l-1} + \mathrm{O}(k^{l-2})}{b_l k^l + b_{l-1}k^{l-1} + \mathrm{O}(k^{l-2})} \\
&= \frac{b_l + (l b_l + b_{l-1})\frac{1}{k} + \mathrm{O}\left(\frac{1}{k^2}\right)}{b_l + b_{l-1}\frac{1}{k} + \mathrm{O}\left(\frac{1}{k^2}\right)} \\
&= 1 + \frac{l}{k} + \mathrm{O}\left(\frac{1}{k^2}\right).
\end{aligned} \tag{3.3}$$

Next, it will be convenient to introduce the constants:

$$J_{k,s} = \int_{\frac{1}{2}}^{1} \left(\frac{1-x}{x}\right)^{k} x^{s}\,\mathrm{d}x, \quad k = 1, 2, \ldots, s = 0, 1, \ldots, l.$$

The substitution $t = (1-x)/x$ yields:

$$J_{k,s} = \int_{0}^{1} \frac{t^{k}}{(1+t)^{s+2}}\,\mathrm{d}t.$$

Integration by parts gives:

$$\begin{aligned} J_{k,s} &= \frac{1}{2^{s+2}(k+1)} + \frac{s+2}{k+1} J_{k+1,s+1} \\ &= \frac{1}{2^{s+2}(k+1)} + \frac{s+2}{k+1}\left(\frac{1}{2^{s+3}(k+2)} + \frac{s+3}{k+2} J_{k+2,s+2}\right). \end{aligned}$$

Since $J_{k,s} \leqslant \int_0^1 t^k\,\mathrm{d}t = \frac{1}{k+1}$:

$$J_{k,s} = \frac{1}{2^{s+2}(k+1)}\left(1 + \frac{s+2}{2(k+2)}\right) + \mathrm{O}\left(\frac{1}{k^3}\right).$$

Using the last equality for large k, we get

$$\begin{aligned} \frac{J_{k,s}}{J_{k+1,s}} &= \frac{\frac{1}{k+1}\left(1 + \frac{s+2}{2(k+2)}\right) + \mathrm{O}\left(\frac{1}{k^3}\right)}{\frac{1}{k+2}\left(1 + \frac{s+2}{2(k+3)}\right) + \mathrm{O}\left(\frac{1}{k^3}\right)} \\ &= \frac{\frac{k+2}{k+1}\left(1 + \frac{s+2}{2}\frac{1}{k+3} + \frac{s+2}{2}\left(\frac{1}{k+2} - \frac{1}{k+3}\right)\right) + \mathrm{O}\left(\frac{1}{k^2}\right)}{\left(1 + \frac{s+2}{2}\frac{1}{k+3}\right) + \mathrm{O}\left(\frac{1}{k^2}\right)} \\ &= \frac{\frac{k+2}{k+1}\left(1 + \frac{s+2}{2}\frac{1}{k+3}\right) + \mathrm{O}\left(\frac{1}{k^2}\right)}{1 + \frac{s+2}{2}\frac{1}{k+3} + \mathrm{O}\left(\frac{1}{k^2}\right)} \\ &= \frac{k+2}{k+1} + \mathrm{O}\left(\frac{1}{k^2}\right) = 1 + \frac{1}{k} + \mathrm{O}\left(\frac{1}{k^2}\right). \end{aligned}$$

Write the density function $g_p(x)$ as a polynomial:

$$g_p(x) = \sum_{s=0}^{l} a_s x^s, \quad \tfrac{1}{2} \leqslant x \leqslant 1.$$

Then

$$\begin{aligned} I_k &= \int_{\frac{1}{2}}^{1} \frac{1}{F^k(x)} g_p(x)\,\mathrm{d}x = \sum_{s=0}^{l} a_s J_{k,s} \\ &= \sum_{s=0}^{l} a_s \left(1 + \frac{1}{k} + \mathrm{O}\left(\frac{1}{k^2}\right)\right) J_{k+1,s} = \left(1 + \frac{1}{k} + \mathrm{O}\left(\frac{1}{k^2}\right)\right) I_{k+1}, \end{aligned}$$

so that

$$\frac{I_k^{n-1}}{I_{k+1}^{n-1}} = 1 + \frac{n-1}{k} + \mathrm{O}\left(\frac{1}{k^2}\right). \tag{3.4}$$

By (3.3) and (3.4), for sufficiently large k we have

$$\frac{|c_{k+1}|}{|c_k|} < \frac{I_k^{n-1}}{I_{k+1}^{n-1}},$$

which completes the proof.

4. Concluding Remarks

The paper continues the analysis of the behavior of the probability of the expert rule to be optimal. Whereas the distributions of expertise levels for which this has been carried out in the past belong to several very restricted families, we have dealt here with a wide (infinite-dimensional) family. This family consists of all distributions defined by a power series, provided they satisfy one of two rather mild conditions (Theorem 1). The cases discussed in other works, such as uniformly distributed correctness probabilities and exponentially distributed logarithmic expertise levels, are very particular instances of our family. Moreover, the family of distributions G_Y contains in particular all polynomials. Since continuous functions on closed intervals may be approximated arbitrarily well by polynomials, and for polynomials Theorem $1'$ provides an explicit formula for $P_{\mathrm{e}}(n)$, this means that in principle we have a way of estimating $P_{\mathrm{e}}(n)$ for any continuous distribution G_Y. Of course, this would be useful mostly for theoretical purposes, as finding polynomials which approximate an arbitrary continuous function is not practical. To estimate $P_{\mathrm{e}}(n)$ in practice, it would probably be most convenient to use Monte Carlo. We emphasize also that the above procedure for estimating $P_{\mathrm{e}}(n)$ refers to any specific n. It cannot assist us in finding the asymptotic behavior of $P_{\mathrm{e}}(n)$ as n grows for an arbitrary continuous distribution G_Y.

Another natural family of distributions to consider is where the correctness probability is distributed polynomially. (Notice the difference between G_p being a polynomial and the case discussed in the preceding paragraph, where G_Y is a polynomial.) We have seen in Example 6 that here again G_Y is an infinite power series. However, the assumptions under which we obtained our formulas were not satisfied unless the group size n is at least $\deg(G_Y)+2$. Moreover, for $n \leqslant \deg(G_Y)$ the formula does not even make sense, as the expression it yields is a divergent series. This raises two interesting questions:

QUESTION 1. *Is Theorem 2 valid for* $n = \deg(G_Y) + 1$*?*

QUESTION 2. *Find an expression for* $P_{\mathrm{e}}(n)$ *for* $n \leqslant \deg(G_Y) + 1$.

Finally, we note that the family of distributions to which our results pertain may be further broadened. In fact, just as Theorem 1″ generalizes Theorem 1′ by replacing the polynomial in (2.3) by a combination of fractional powers in (2.4), we can replace the power series of (2.1) by a series consisting of fractional powers. One can easily prove the corresponding generalization of Theorem 1.

Acknowledgements

The authors wish to express their gratitude to the referees of the paper, whose comments contributed a lot to the presentation of this paper and provide interesting questions for further research.

References

Abramowitz, M. and Stegun, I. A. (1968) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 7th printing, National Bureau of Standards Appl. Math. Ser. 55.

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

Berend, D. and Paroush, J. (1998) When is Condorcet's jury theorem valid?, *Social Choice and Welfare* **15**, 481–488.

Berend, D. and Sapir, L. Expert rule versus majority rule under partial information, II, Preprint.

Berg, S. (1993) Condorcet's jury theorem, dependence among jurors, *Social Choice and Welfare* **10**, 87–95.

Berg, S. (1997) Indirect voting systems: banzhaf numbers, majority functions and collective competence, *Europ. J. Political Econ.* **13**, 557–573.

Berg, S. and Paroush, J. (1998) Collective decision making in hierarchies, *Math. Soc. Sci.* **35**, 233–244.

Boland, P. G. (1989) Majority systems and the Condorcet jury theorem, *The Statistician* **38**, 181–189.

Boland, P. G., Proschan, F. and Tong, Y. (1989) Modelling dependence in simple and indirect majority systems, *J. Appl. Probab.* **26**, 8–88.

de Condorcet, N. C. (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris.

Fikhtengol'ts, G. M. (1965) *The Fundamentals of Mathematical Analysis*, 1st edn, Vol. II, Pergamon Press, Oxford.

Grofman, B., Owen, G. and Feld, S. (1983) Thirteen theorems in search of the truth, *Theory and Decision* **15**, 261–278.

Hastings, N. A. J. and Peacock, J. B. (1975) *Statistical Distributions*, Wiley, New York.

Isaacson, E. and Keller, H. B. (1966) *Analysis of Numerical Methods*, Courant Institute of Math. Sci., New York University, Wiley, New York.

Ladha, K. (1992) The Condorcet jury theorem. Free speech and correlated votes, *Amer. J. Political Sci.* **36**, 617–634.

Ladha, K. (1995) Information polling through majority rule voting: Condorcet's jury theorem with correlated votes, *J. Econ. Behavior & Organization* **26**, 353–372.

Miller, N. R. (1986) Information, electorates, and democracy: some extensions and interpretations of Condorcet jury theorem, In: (B. Grofman and G. Owen (eds)), *Information Pooling and Group Decision Making*, Vol. 2, CT:JAI Press, Greenwich, pp. 173–192.

Nitzan, S. and Paroush, J. (1982) Optimal decision rules in uncertain dichotomous choice situations, *Internat. Econ. Rev.* **23**, 289–297.

Nitzan, S. and Paroush, J. (1984) A general theorem and eight corollaries in search of a correct decision, *Theory and Decision* **17**, 211–220.

Nitzan, S. and Paroush, J. (1984) Partial information on decisional competences 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 Univ. Press, Cambridge.

Paroush, J. (1998) Stay away from fair coins: A correct Condorcet's jury theorem, *Social Choice and Welfare* **15**, 15–20.

Sapir, L. (1998) The optimality of the expert and majority rules under exponentially distributed competence, *Theory and Decision* **45**, 19–35.

Shapley, L. and Grofman, B. (1984) Optimizing group judgmental accuracy in the presence of interdependencies, *Public Choice* **43**, 329–343.

Young, H. P. (1989) Condorcet's theory of voting, *Amer. Political Sci. Rev.* **82**, 1231–1244.