Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 1838–1850
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Asymptotic efficiency of majority rule relative to rank-sum method for selecting the best population

Samuel S. Wu[a, b, *], David H. Annis[c]

[a] College of Medicine, University of Florida, Gainesville, FL 32610, USA
[b] Malcom Randall VA Medical Center, Gainesville, FL 32608, USA
[c] Operations Research Department, Naval Postgraduate School, Monterey, CA 93943, USA

Received 27 May 2005; received in revised form 9 February 2006; accepted 28 March 2006
Available online 21 June 2006

Abstract

The ranking and selection problem has been well-studied in the case of continuous responses. In this paper, we address the situation in which continuous responses are replaced by discrete orderings. When individuals in the population provide exhaustive rank-orderings of the alternatives, two common decision rules are the majority rule and the rank-sum method. In the former case, the alternative receiving the most first-place votes is declared superior, while in the latter, the alternative with the smallest rank-sum is deemed the best. Both the Pitman efficiencies and the lower bounds on Bahadur efficiencies of the majority rule relative to the rank-sum method are derived, assuming that the rank data are generated from either the Plackett–Luce or the translative strengths models. In addition, finite sample properties of the two methods are compared with the maximum likelihood approach through simulation studies. Our results suggest two things. First, when it is substantially more difficult to obtain a complete rank-ordering than simply the top choice, the majority rule performs adequately and efforts would be better spent asking many voters to provide top choice rather than fewer voters to provide complete orderings. Second, the rank-sum rule compares favorably to, and is substantially more robust than, the maximum likelihood approach.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Asymptotic efficiency; Majority rule; Rank sum method; Plackett–Luce model; Translative strengths model

1. Introduction

The selection and ranking problem has been extensively studied since the 1950s, including pioneering works of Bechhofer (1954) and Gupta (1956). That these problems have been studied in such detail points to their wide-spread relevance to many applied problems. In many analyses, the observed responses (on which ranking and selection are based) are recorded on a continuous scale. In this paper, we address the situation in which continuous responses are replaced by their corresponding discrete rank-orders. In this case, suppose a committee of size n is charged with choosing the best among m potential alternatives. To this end, each voter submits an exhaustive rank-ordering of the alternatives, with ties prohibited (i.e., the ranking is strict). There are many potential ways to determine a winner in such a circumstance. For example, the alternative receiving the most first-place votes may be chosen. Such a system is commonly referred to as the majority (MJ) rule and is in wide use. On the other hand, when the number of alternatives,

* Corresponding author. College of Medicine, University of Florida, Gainesville, FL 32610, USA.
E-mail addresses: samwu@biostat.ufl.edu (S.S. Wu), annis@nps.edu (D.H. Annis).

0378-3758/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.03.012

m, is large, the rank-sum (RS) method may be preferable. This procedure selects the alternative that has the smallest rank sum over all votes. (Low ranks are considered preferable to higher ones—i.e., first-place corresponds to the best alternative, while mth-place corresponds to the worst).

Throughout, we assume that each of the alternatives possesses an inherent desirability or appeal, and that these determine how likely a particular alternative is to be preferred to another (or a group of others). In this paper, we investigate the asymptotic efficiency of the MJ rule relative to the RS method in determining the correct winner (i.e., the alternative whose desirability is largest), first under the Plackett–Luce model (Luce, 1959; Plackett, 1975) and then under the translative strengths model (Stern, 1990; Marden, 1995).

We derive two criteria to quantify asymptotic efficiency. The first is a Pitman (1948) type asymptotic efficiency which measures the asymptotic ratio of sample sizes required to achieve equal limiting power against the same sequence of alternative strengths whose differences tend to zero. The second measure is a Bahadur (1971) type efficiency, namely, the limiting ratio of sample sizes such that the probability of selecting the best alternative tends to 1 at the same rate, for any fixed alternative strengths.

The remainder of the paper is organized as follows. Section 2 introduces the Plackett–Luce and translative strengths models. Sections 3 and 4 derive the Pitman efficiencies and the lower bounds on Bahadur efficiencies of the MJ rule relative to the RS method first under the Plackett–Luce model and then assuming a translative strengths formulation. Section 5 presents results of a series of simulations comparing the aforementioned approaches for various values of n (voting committee size) and m (alternatives under consideration). A discussion of main results is given in Section 6. Appendix A states a few properties of the Plackett–Luce model. Longer proofs are left for Appendix B, and Appendix C derives quantities related to Pitman efficiency for some common translative strengths models.

2. Modeling rank data

Let $\pi(\cdot)$ denote a permutation of the integers $\{1, 2, \ldots, m\}$ which corresponds to a rank-ordering of the alternatives. In particular, $\pi(r) = s$ denotes alternative s receiving rank r. The complete rank-ordering is expressed as $\pi(1) \to \pi(2) \to \cdots \to \pi(m)$, indicating that the alternative receiving the first-place vote, $\pi(1)$, is preferred to the alternative receiving the second-place vote, $\pi(2)$, and so forth. Implicit in this notation is the idea of transitivity of an individual's preference, namely that together $i \to j$ and $j \to k$ are sufficient to guarantee $i \to k$.

The n voters generate their rank-orderings as follows. Let $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_m)' \in \mathscr{R}_+^m$ denote the desirabilities of alternatives. Without loss of generality, we assume that the m alternatives are labeled according to their desirability ($\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_m$) throughout. Suppose $Y_j = (Y_{1,j}, Y_{2,j}, \ldots, Y_{m,j})'$, $j = 1, \ldots, n$ are i.i.d. random vectors taking values in $\mathscr{R}^m$ and consisting of independent components distributed according to $P_{\gamma_1}, P_{\gamma_2}, \ldots, P_{\gamma_m}$, respectively. The data we observe are orderings $\pi_j = \kappa(Y_j)$, where κ is a measurable mapping from $\mathscr{R}^m$ to Π, the set of all $m!$ permutations of alternatives.

If P_{γ_i} is the exponential distribution with rate γ_i (mean $1/\gamma_i$) and κ is given by the *increasing* order of the $Y_{i,j}$'s (that is, $\pi_j = (i_1, i_2, \ldots, i_m)$ if and only if $Y_{i_1,j} < Y_{i_2,j} < \cdots < Y_{i_m,j}$), then the ordering vectors follow the Plackett–Luce model (Luce, 1959; Plackett, 1975), which is perhaps the most common characterization of permutation probabilities. This model, denoted by PL($\boldsymbol{\gamma}$), assumes that each voter independently submits a ranking of all alternatives, and the probability assigned to a particular ordering π, is

$$P_{\boldsymbol{\gamma}}[\pi(1) \to \pi(2) \to \cdots \to \pi(m)] = \prod_{i=1}^{m} \frac{\gamma_{\pi(i)}}{\gamma_{\pi(i)} + \cdots + \gamma_{\pi(m)}}. \tag{1}$$

It is easy to verify that $P_{\boldsymbol{\gamma}}[\pi] = P_{c\boldsymbol{\gamma}}[\pi]$ for any constant $c > 0$. In other words, the probability of the ordering is solely dependent on the *relative* (rather than absolute) desirability of an alternative. To ensure a unique solution, we always assume $\sum_{i=1}^{m} \gamma_i = 1$ for this model. Our motivation of the PL($\boldsymbol{\gamma}$) model is not unique. Indeed, Stern (1990) shows that this model can be obtained in many different ways. In addition, the Plackett–Luce model possesses an appealing internal consistency property. Furthermore, it satisfies Luce's choice axiom and has a connection to the proportional hazard model (Cox, 1972). More details are given in Appendix A.

Another class of models for rank data involves a location-family with possibly different location parameters (Daniel, 1950; Marden, 1995). In this case, $P_{\gamma_1}, P_{\gamma_2}, \ldots, P_{\gamma_m}$ have cumulative distribution functions $F(y - \gamma_1), F(y - \gamma_2), \ldots, F(y - \gamma_m)$ and κ is a one-to-one mapping given by the *decreasing* order of the $Y_{i,j}$'s, that is,

Download English Version:

https://daneshyari.com/en/article/1150031

Download Persian Version:

https://daneshyari.com/article/1150031

Daneshyari.com