European Journal of Operational Research 206 (2010) 667–675

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

The selection efficiency of tournaments

Dmitry Ryvkin *

Department of Economics, Florida State University, Tallahassee, FL 32306-2180, USA

ARTICLE INFO

Article history:
Received 20 April 2009
Accepted 5 March 2010
Available online 15 March 2010

Keywords:
Human resources
Applied probability
Selection
Tournament
Simulation

ABSTRACT

We discuss tournaments in terms of their efficiency as probabilistic mechanisms that select high-quality alternatives (“players”) in a noisy environment. We characterize the selection efficiency of three such mechanisms – contests, binary elimination tournaments, and round-robin tournaments – depending on the shape of the distribution of players' quality, the number of players, and noise level. The results have implications as to how, and under what circumstances, the efficiency of tournament-based selection can be manipulated.

1. Introduction

Consider recruitment of employees by a firm. After the initial screening of applications, those who meet the minimum qualifications form a set from which further selection is made by comparing applicants with one another.

Recruitment is an example of a *selection problem*. Generally, a decision maker (DM) faces a selection problem when one or several “best” alternatives must be chosen from a set. Nontrivial instances of this problem are, for example, selecting a population with the highest unobserved mean (e.g., Gibbons et al., 1977), the secretary problem (e.g., Bearden et al., 2006; Bearden and Murphy, 2007), or a multi-attribute choice problem (e.g., Leskinen et al., 2004; Baucells et al., 2008).

In this paper, we study the following selection problem: The DM faces a random sample of N alternatives (“players”). Each alternative is characterized by a single attribute (“ability”) – an independent draw from the population with a known distribution. The DM cannot observe abilities, but can observe the results of (possibly multiple) ordinal comparisons of randomly perturbed abilities among the alternatives. The DM is interested in choosing the alternative with the highest ability, or, equivalently, the lowest rank order. Multi-stage recruitment is an example of such an environment: applicants' abilities are unknown, and their overall performance difficult to quantify on a numerical scale, but it may be relatively easy to compare one applicant with another. This selection problem is different from the ones discussed in Gibbons et al. (1977), where alternatives are *populations* with unknown (but deterministic) parameters (e.g., means), and multiple observations from each population are available to the DM.

Selection in such an environment can be made through a *tournament* – a scheme that uses (possibly multiple) comparisons of alternatives (“players”) to produce one player as the *winner*. In a noisy environment, the best player wins a tournament with some probability different from one. A DM, therefore, can be interested in the efficiency of different selection schemes, and how it depends on parameters such as the number of players, the noise level, and the distribution of players' abilities in the population.

In the economics and management literature, tournaments are mainly discussed in relation to incentive provision in firms (e.g., Prendergast, 1999; Orrison et al., 2004; Gerchak and He, 2003), sports (e.g., Szymanski, 2003), research competition (e.g., Taylor, 1995), and rent seeking (e.g., Lockard and Tullock, 2001). Under tournament incentives, players choose the supply of costly effort or other resources, and the principal's objective is, typically, the maximization of total output.

An alternative view of tournaments as selection mechanisms that help identify better players was introduced to the economics literature by Hvide and Kristiansen (2003). The authors consider a contest in which players can choose the level of risk pertaining to their output and find that *selection efficiency*, defined there as the probability of a high-ability type player winning the contest, may be a nonmonotonic function of the number of competitors and of the proportion of high-ability types in the population.

The concept of selection efficiency of tournaments as quality-enhancing selection mechanisms is also discussed in the statistical decision theory literature (see, e.g., Gibbons et al., 1977; Narayana, 1979; David, 1988). Here, unlike in the economics and management literature, it is typically assumed that tournament “players”

---

* Tel.: +1 850 644 7209; fax: +1 850 644 4535.
E-mail address: dryvkin@fsu.edu

0377-2217/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2010.03.016

do not make strategic choices, i.e. their performance is pre-determined by their intrinsic quality and, possibly, random factors beyond their control. A DM's objective then is to choose the appropriate selection scheme to most reliably identify players with better quality by "filtering out" the noise through multiple observations.

In this paper, similar to the statistical decision theory literature, we assume that players' effort levels are not a choice variable. This assumption restricts the applicability of our results in managerial settings to situations when the incentives in a tournament game are set in such a way that all players always choose to perform at their best. Albeit restrictive, this assumption is realistic for an important class of selection situations involving human subjects, such as final stages of recruitment tournaments and other environments where stakes are high, and significant prior investment has already been made by competitors. Other examples include formation of the Olympic team, innovation races, elections, or high level sports tournaments. Our results, of course, are also applicable to choice situations where alternatives are not humans, such as, for example, choice among different technologies.

Tournament selection schemes can be constructed using different matching and/or elimination rules, or *formats*. The simplest tournament format is the *contest* format, in which all players perform only once and the player with top performance is the winner. Recruitment, however, is often done in several stages, which suggests that a one-shot format, such as a contest, is too noisy to make reliable inference in the selection problem discussed here. A multistage elimination format, for example, the *binary elimination* (also known as knock-out) format, can be used instead. In the economics and management literature, such formats are discussed in the context of promotions in hierarchical organizations (see, e.g., Rosen, 1986; O'Flaherty and Siow, 1995; Devaro, 2006). In the statistical decision theory literature, the selection efficiency of knock-out tournaments is studied for various configurations of winning probabilities and seeding (see Hartigan, 1968; Knuth, 1987; Marchand, 2002; Israel, 1981; Hwang, 1982; Horen and Reizman, 1985, among others).

Another prominent example of a nontrivial tournament selection scheme is the *round-robin* format, in which multiple binary comparisons determine the score of each player, and the player with the top score is the winner (see, e.g., Harary and Moser, 1966; Rubinstein, 1980; Mendonça and Raghavachari, 2000). A number of authors compare round-robin and knock-out tournaments in some special cases (David, 1959; Glenn, 1960; Searls, 1963; Appleton, 1995; McGarry and Schutz, 1997). In this literature, tournaments involving multiple binary comparisons (such as knock-out and round-robin) are typically parameterized by a matrix of deterministic winning probabilities. This setting does not allow one to explore the parametric dependence of tournament outcomes on the number of players, noise level, or the distribution of players' abilities, therefore in this paper we use a different approach.

The efficiency of a tournament selection scheme can be characterized by several criteria. In the context of the selection problem discussed here, Ryvkin and Ortmann (2008) explore the *predictive power* criterion – the probability of selecting the best player as the winner. One of the central findings of Ryvkin and Ortmann (2008) is that predictive power exhibits nonmonotonicity as a function of the number of players for fat-tailed distributions of players' abilities.

In this paper, we explore two alternative measures of selection efficiency of tournaments: the expected ability of the winner, and the expected rank of the winner. From a DM's perspective, these measures are more "balanced" than the predictive power measure in that they refer to the characteristics of the chosen alternative not requiring necessarily that it be the best one. For example, a recruiting committee interested in hiring the candidate with the highest possible ability, regardless of ranking, can choose to maximize the expected ability of the winner. At the same time, a recruiting committee whose goal is to get an edge in a race against other firms will tend to minimize the expected rank (see, e.g., Assaf and Samuel-Cahn, 1996).

Ryvkin and Ortmann (2008) provide the results of exploratory simulations for the expected ability and expected rank of the winner. Simulations suggest that, similarly to predictive power, the two efficiency criteria can exhibit nonmonotonic behavior. No theoretical foundation is provided, however. In this paper, we analyze the expected ability and expected rank of the winner for three tournament formats – contests, binary elimination and round-robin tournaments. We provide a theory, a comprehensive set of simulation results, and a detailed discussion for the two criteria. Our major contribution is in showing that (i) the expected ability of the winner always increases in the number of players; and, (ii) like predictive power, the expected rank of the winner exhibits nonmonotonicity as a function of the number of players for fat-tailed distributions of abilities. Also, both the expected ability and expected rank of the winner become nonmonotonic as a function of noise level for round-robin tournaments when the number of players is sufficiently large. Our results have important implications for DMs facing the selection problem in noisy environments, such as recruitment committees in organizations, or Olympic committees.

The rest of the paper is organized as follows. In Section 2, we present a general model. In Sections 3–5, the selection efficiency of three tournament formats – contests, binary elimination, and round-robin tournaments – is analyzed. Specifically, in Section 3 we provide a theory for the expected ability and expected rank of the winner in contests, and numerically illustrate our results. A theory for the expected ability of the winner in binary elimination tournaments is developed in Section 4. In Sections 4 and 5, we provide the results of numerical simulations for binary elimination and round-robin tournaments. Section 6 contains a discussion of our findings and concluding remarks.

2. The model

Let $\mathcal{N} = \{1, \ldots, N\}$ be a set of N alternatives ("players"). Each player $i = 1, \ldots, N$ is characterized by an attribute $X_i \in \mathbb{R}$ ("ability"). Abilities $\mathbf{X} = (X_1, \ldots, X_N)$ are independently and identically distributed (i.i.d.) with a probability density function (pdf) $f(x)$ and the corresponding cumulative density function (cdf) $F(x)$. It is assumed, for simplicity, that $f(x)$ is continuous on its support, and has a finite second moment.

The *selection problem* is to identify the "best" player, i.e. the player with the highest ability. Albeit straightforward when abilities X_i are directly observable, selection becomes nontrivial when the abilities are perturbed by noise. In this case, a *selection scheme* has to be employed that can only give the "right" answer with some probability less than one.

We consider a special class of selection schemes we call *tournaments*, which use ordinal comparisons of perturbed abilities to identify the best player as the "winner." Specifically, we consider three prominent tournament formats: contests, binary elimination tournaments, and round-robin tournaments.

A tournament selection scheme can involve one or several stages. At each stage t, player i's *output*, Y_{it}, is her perturbed ability: $Y_{it} = X_i + \epsilon_{it}$, where ϵ_{it} are zero-mean, i.i.d. across players and across stages, with a symmetric pdf $\phi(\epsilon)$, cdf $\Phi(\epsilon)$, and a finite second moment. The overall level of noise can be characterized by parameter σ^2, the variance of ϵ_{it}. According to the tournament scheme used, output levels Y_{it} are compared, and the winner of the tournament, player i_w, is determined. Let $X_w \equiv X_{i_w}$ denote the