SOME DESIGNS FOR MULTICRITERIA BANDITS

BY P. W. JONES, A. M. LEWIS AND R. HARTLEY

Keele University

Abstract

The Bernoulli two-armed bandit with finite horizon and independent beta priors is considered from a multicriteria perspective. Several sequential designs are suggested and their characteristics are derived.

1. Introduction. Consider the Bernoulli two armed bandit with probabilities p_1, p_2 of obtaining a reward and $(1-p_1)$, $(1-p_2)$ of obtaining nothing when pulling arms 1 and 2, respectively. It is assumed that the maximum number of pulls is N. The original two armed bandit problem concentrated on designs which maximize the expected return $E(R)$ [see, for example, Jones (1975)], this paper deals with two additional criteria, $P(CS)$, which is the probability of correctly selecting the superior arm at the termination of the process, and $E(N_{(1)})$, which is the expected number of pulls on the poorer arm. In the context of clinical trials where the arms are treatments and the number of patients is N, these additional criteria have ethical and statistical importance and need to be taken into account in evaluating designs. There are two ways of looking at this problem in this context, either as an optimization problem for this particular set of N patients, or as a decision problem where a recommendation on the superior treatment will be made for future patients. The two criteria, $E(R)$ and $E(N_{(1)})$ are of the first type and ethical considerations suggest that these should be