

## One-armed bandit process with a covariate

You Liang · Xikui Wang · Yanqing Yi

Received: 8 March 2012 / Revised: 7 February 2013 / Published online: 12 April 2013  
© The Institute of Statistical Mathematics, Tokyo 2013

**Abstract** We generalize the bandit process with a covariate introduced by Woodroffe in several significant directions: a linear regression model characterizing the unknown arm, an unknown variance for regression residuals and general discounting sequence for a non-stationary model. With the Bayesian regression approach, we assume a normal-gamma conjugate prior distribution of the unknown parameters. It is shown that the optimal strategy is determined by a sequence of index values which are monotonic and determined by the observed value of the covariate and updated posterior distributions. We further show that the myopic strategy is not optimal in general. Such structural properties help to understand the tradeoff between information gathering and immediate expected payoff and may provide certain insight for covariate adjusted response adaptive design of clinical trials.

**Keywords** Bandit process · Bayesian regression · Markov decision process · Optimal strategy