Mathematics > Statistics Theory

Stochastic Bandit Based on Empirical Moments

Junya Honda, Akimichi Takemura

(Submitted on 14 May 2011)

In the multiarmed bandit problem a gambler chooses an arm of a slot machine to pull considering a tradeoff between exploration and exploitation. We study the stochastic bandit problem where each arm has a reward distribution supported in a known bounded interval, e.g. [0,1]. For this model, policies which take into account the empirical variances (i.e. second moments) of the arms are known to perform effectively. In this paper, we generalize this idea and we propose a policy which exploits the first d empirical moments for arbitrary d fixed in advance. The asymptotic upper bound of the regret of the policy approaches the theoretical bound by Burnetas and Katehakis as d increases. By choosing appropriate d, the proposed policy realizes a tradeoff between the computational complexity and the expected regret.

Subjects:	**Statistics Theory (math.ST)**
MSC classes:	62L05, 60G40
Cite as:	**arXiv:1105.2879 [math.ST]**
	(or **arXiv:1105.2879v1 [math.ST]** for this version)

Submission history

From: Akimichi Takemura [view email]
[v1] Sat, 14 May 2011 09:51:25 GMT (30kb)

Which authors of this paper are endorsers?