in the end...
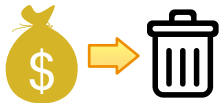
(1) forfeit all rewards
(2) output **1** arm

goal: find the single **best arm**
(largest expected reward)