Further Contributions to the "Two-Armed Bandit" Problem

Robert W. Keener

A version of the two-armed bandit with two states of nature and two repeatable experiments is studied. With an infinite horizon and with or without discounting, an optimal procedure is to perform one experiment whenever the posterior probability of one of the states of nature exceeds a constant ξ^*, and perform the other experiment whenever the posterior is less than ξ^* with indifference when the posterior equals ξ^*. This cutoff value ξ^* is expressed in terms involving expectations of ladder variables and can be calculated using Sptizer series.