

## Задание 1.

Сравните эффективность  $\epsilon$ -жадных алгоритмов, с алгоритмом, использующим распределение Гиббса (мягкий максимум).

Сравнение проводите на наборе тестов следующего вида:

- каждому алгоритму предлагается 2000 случайно сгенерированных задач  $n$ -руких бандитов, где  $n=10$ .

- в каждой задаче вознаграждение за выбор действия  $a$  ( $a \in \overline{1:10}$ ) определяется случайным числом с нормальным распределением со средним значением  $Q^*(a)$  и дисперсией 1.

- для генерации каждой из 2000 задач значения  $Q^*(a)$  выбираются из нормального распределения со средним значением 0 и дисперсией 1.

Постройте графики изменения среднего вознаграждения на 1 шаг и процента выбора оптимального решения на протяжении 1000 раундов игры разными алгоритмами (усреднено по 2000 задачам).

Изучите влияние аргумента температура в распределении Гиббса на эффективность алгоритма.

Изучите влияние дисперсии при выборе вознаграждения на эффективность разных алгоритмов.