Multi-Armed Bandits

Question 1. Read Section 2.3 of SB and generate the plots in Figure 2.2. Also, generate a plot that shows the average absolute error in the estimate for each action (arm) as a function of time steps. In addition to the ϵ in the book, generate the plots for when the ϵ changes with time in a matter such that the sequence of $\{\epsilon(t), t \geq 1\}$ satisfies Equation (2.7).

Question 2. Repeat Question 1 for when the variance corresponding to each arm is 4 instead of 1.

Question 3. Solve exercise 2.3. Do it for all choices of ϵ in Question 1.

Question 4. Show that the sample mean is not influenced by the initial choice of $Q_1(a), \forall a$, where as when using a constant step-size α (see Equation (2.5)) the estimate $Q_t(a)$ is a function of $Q_1(a)$. Also, show that the dependence is larger for a smaller α.

Propose a method such that we can have a constant step-size but no dependence of $Q_t(a)$ on $Q_1(a)$.

Question 5. Do Exercise 2.5.

Question 6. Generate Figure 2.4 and solve exercise 2.8. Repeat generation of the figure and solve the exercise for $c = 1$ and $c = 4$ too.

Question 7. Generate Figure 2.5.