Review Problems 1

Professor Bryan Graham

Ec143 – Spring 2026

In preparing for midterm exams review your lecture notes, assigned course readings, problem sets and prepare answers to the questions on the review sheet(s). You may bring to the exam a single 8.5×11 inch sheet of paper with notes on it. No calculation aides are allowed in the exam (e.g., calculators, phones, laptops etc.). You may work in pencil or pen, however I advise the use of pencil. Please be sure to bring sufficient blue books with you to the exam if possible. Scratch paper will be provided.

[1] [Exploration vs. Exploitation] In each of $T = 4$ rounds an agent needs to pull either Arm 0 or Arm 1. Let $D_t = 0$ if Arm 0 is pulled in round t and $D_t = 1$ if Arm 1 is pulled instead. Let $Y_t(d)$ denote the (potential) outcome associated with pulling arm $d = 0, 1$ in round t. This outcome is either a success, in which case $Y_t(d) = 1$, or a failure, in which case $Y_t(d) = 0$. The observe round t rewards is

$$Y_t = (1 - D_t) Y_t(0) + D_t Y_t(1).$$

The agent's goal is to maximize the total expected number of successes across all $T = 4$ rounds (i.e., the expectation of the sum $Y_1 + Y_2 + Y_3 + Y_4$). The agent has the following priors over the properties of the two arms. Let θ_d denote the unknown success rate for arms $d = 0, 1$.

1. Arm 0: The agent knows, with certainty, that $\theta_0 \in \Theta_0 = \{0.25, 0.75\}$. *A priori* she believes that $\Pr(\theta_0 = 0.25) = \Pr(\theta_0 = 0.75) = \frac{1}{2}$. Arm 0 is either "good" ($\theta_0 = 0.75$) or "bad" ($\theta_0 = 0.25$).
2. Arm 1: The agent knows, with certainty, that $\theta_1 \in \Theta_1 = \{0, 1\}$. *A priori* she believes that $\Pr(\theta_1 = 0) = \Pr(\theta_1 = 1) = \frac{1}{2}$. Arm 1 is either "perfect" ($\theta_1 = 1$) or "broken" "($\theta_1 = 0$).

The two arm prior distributions are independent such that $\Pr(\theta_0 = t_0, \theta_1 = t_1) = \Pr(\theta_0 = t_0)\Pr(\theta_1 = t_1)$. Let $\mathcal{I}_t$ denote the begining-of-period t information set of the agent. At $t = 1$ the agent begins with nothing more that their prior, π. At $t = 2$ the information set additionally includes the choices and outcomes from round 1 (i.e, $\mathcal{I}_2 = (\pi, D_1, Y_1)$) and so on.

[a] Explain, in words, why

$$\mathbb{E}[Y_1 | D_1 = 0, \mathcal{I}_1] = \mathbb{E}[Y_1(0) | \mathcal{I}_1], \ \mathbb{E}[Y_1 | D_1 = 1, \mathcal{I}_1] = \mathbb{E}[Y_1(1) | \mathcal{I}_1].$$

[b] Calculate $\mathbb{E}[Y_1 | D_1 = 0, \mathcal{I}_1]$ and $\mathbb{E}[Y_1 | D_1 = 1, \mathcal{I}_1]$. Based on the expected reward alone, is there a preferred arm to start with?

[c] Suppose the agent pulls Arm 0 in Round $t = 1$ and observes a success (i.e., $(D_1, Y_1) = (0, 1)$). Calculate the agent's posterior belief that Arm 0 is "good"; that is,

$$\Pr(\theta_0 = 0.75 | \mathcal{I}_2) = ?$$

for $\mathcal{I}_2 = (\pi, D_1 = 0, Y_1 = 1)$. Further calculate the agent's (new) expected reward for Arm 0 were she to pull is again in Round $t = 2$.

[c] [Exploitation] Suppose the agent decides to stay with Arm 0 for the remaining three rounds ($t =$

$2, 3, 4)$ regardless of any future outcomes. Calculate the total expected payoff associated with this strategy

$$\mathbb{E}\left[Y_2 + Y_3 + Y_4 \,|\, D_2 = D_3 = D_4 = 0, \mathcal{I}_2\right]$$

[d] [Exploration] Suppose that after pulling Arm 0 in round 1 and observing a success, the agent nevertheless decides to switch and pull Arm 1 in round t =2. What are the agent's beliefs about the probability that this pull will be successful? What is the corresponding period t=2 expected payoff? How does this payoff compare with the one associated with Arm 0?

[e] [Exploration] If the pull of Arm 1 is successful in round $t = 2$, what is the agent's optimal strategy for remaining rounds $t = 3, 4$? What is the expected payoff associated with this strategy? If, instead, the pull of Arm 1 is unsuccessful in round $t = 2$, what is the agent's optimal strategy for remaining rounds $t = 3, 4$? What is the expected payoff associated with this strategy? Using your analysis calculate the expected payoff of "trying Arm 1 in round $t = 2$ and choosing optimally thereafter". Compare this expected payoff with exploitation one calculated in part [c].

[e] Explain, in words, why a rational agent would choose an arm with a lower immediate expected reward in Round 2. What is the "Information Premium" in this context?

[f] Suppose the game ends after $T = 2$ arounds. As earlier, the agent pulls Arm 0 in round 1 and observes a success. Will only one round remaining is it ever optimal for her to switch arms?

[g] [Harder] What is the expected payoff from pulling Arm 0 in round $t = 1$ and then optimally choosing arms thereafter? What is the expected payoff from pulling Arm 1 in round $t = 1$ and then optimally choosing arms thereafter? Should the agent have chosen Arm 0 in round $t = 1$? Explain.

[h] [Harder] Can you construct a perturbation of the problem faced by the agent such that she is indifferent between Arms 0 and 1 in round $t = 1$.