DATA 37200 SUGGESTED PROBLEMS, Winter 2026

FR: Foster-Rakhlin notes. LS: bandit textbook by Lattimore and Svesepari.

Note: some problems are harder than others. Sometimes, you may need to look at the chapter of the notes/book to understand their notation. Reviewing the material from other references and looking at their problems. These problems do not 100% exhaustively cover all course content, but should be helpful for those looking to study.

1 Core math skills, esp. for using probability

1. Martingales: Exercise 3.2 of LS, parts (a,b,c).
2. Variance: Exercise 5.1 of LS.
3. MGFs: Exercise 5.9 of LS.
4. Taylor expansion: Exercise 5.11 of LS (following the hint).
5. Exercise 5.18 of LS.
6. Exercise 8.1 of LS (you do not need to read the context, just solve the problem following the hint).
7. Bayes rule: Exercise 34.1 of LS.
8. KL divergence: Exercise 14.7 of LS.
9. Give an example of three random variables X, Y, Z such that they are pairwise but not jointly independent.

2 Online learning and bandits

1. Exercise 6.6 of LS.
2. Exercise 34.8 of LS.
3. Exercise 4 of Chapter 2 of FR.
4. Exercise 5 of Chapter 3 of FR. (Was essentially covered in class, but still good practice.)
5. Exercise 6 of Chapter 3 of FR. (You can take the mentioned $O(\log |F|/\delta)$ guarantee for ERM for granted, we did not explicitly show this in class.)

3 Reinforcement learning and control

1. Exercise 38.3 of LS.
2. Exercise 38.8 of LS. (Was essentially covered in class.)
3. Give an example of an MDP where the optimal policy is not unique.
4. Consider the LQR problem with dynamics

   $$x_t = Ax_{t-1} + u_{t-1}$$

   with A is an arbitrary $d \times d$ matrix and u_t the control input, and cost function

   $$\sum_{t=1}^{\infty} \|x_t\|^2.$$

   (a) Is this system controllable, i.e. can we go from any initial state to any other state in a finite number of steps? (b) Explain (with math or words) the optimal policy for this problem.

4 Games

1. The game of Tic-Tac-Toe is played on a 3x3 grid where each square in the grid can either have an X, an O, or nothing. See `https://en.wikipedia.org/wiki/Tic-tac-toe` for needed background and information. Say that the X player receives 1 dollar from the O player if they win, and receives -1 dollars if they lose. What is the minimax value of this game, assuming X goes first?

2. Exercise 2.7 of https://homes.cs.washington.edu/~karlin/GameTheoryBook.pdf.

3. Exercise 2.11 of https://homes.cs.washington.edu/~karlin/GameTheoryBook.pdf. (You can solve for the optimal strategies using multivariable calculus.)