1. Tickets to a lottery cost 1 euro. There are two possible prizes: a 10 euro payoff with probability 1/50 and a million euro payoff with probability 1/2,000,000. What is the expected monetary value of a lottery ticket? When (if ever) is it rational to buy a ticket? Be precise — show an equation involving utilities. You may assume current wealth of $k\text{ euros}$ and that $U(S_{k+10}) = 10 \times U(S_{k+1})$, but you may not make any assumption about $U(S_{k+1,000,000})$.

2. In 1738, J. Bernoulli investigated the *St. Petersburg paradox*, which works as follows. You have the opportunity to play a game in which a fair coin is tossed repeatedly until it comes up heads. If the first heads appears on the $n$th toss, you win $2^n$ euros.

   (a) Show that the expected monetary value of this game is infinite.

   (b) Bernoulli resolved the apparent paradox by suggesting that the utility of money is measured on a logarithmic scale: $U(S_n) = a \log_2 n + b$, where $S_n$ the state of having $n$ euros. What is the expected utility of the game under this assumption?

   (c) What is the maximum amount that it would be rational to pay to play the game, assuming that one’s initial wealth is $k\text{ euros}$?

3. For the $4 \times 3$ grid world from lectures, calculate which squares can be reached from $[1, 1]$ by the action sequence $[U, U, R, R, R]$ and with what probabilities?

4. Suppose that we define the utility of a state sequence to be the *maximum* reward obtained in any state in the sequence. Show that this utility function does not result in a stationary preferences between state sequences.

5. Consider an undiscounted MDP having three states (1, 2, and 3), with rewards $-1, -2, 0$, respectively. State 3 is a terminal state. In states 1 and 2 there are two possible actions: $a$ and $b$. The transition model is as follows:

   - In state 1, action $a$ moves the agent to state 2 with probability 0.8 and makes the agent stay put with probability 0.2.
   - In state 2, action $a$ moves the agent to state 1 with probability 0.8 and makes the agent stay put with probability 0.2.
   - In either state 1 or state 2, action $b$ moves the agent to state 3 with probability 0.1 and makes the agent stay put with probability 0.9.
Answer the following questions:

(a) Apply policy iteration, showing each step in full, to determine the optimal policy and the values of states 1 and 2. Assume that the initial policy has action $b$ in both states.

(b) What happens to policy iteration if the initial policy has action $a$ in both states? Does discounting help?