2 LP Meets Linear Regression

(a) $\frac{n-1}{n}$. Consider the case where the cost of the chosen expert is always 1, and the cost of each other expert is 0. Let k be the least-frequently chosen expert, and let m_k be the number of times that expert is chosen. This will result in a regret of $\frac{1}{T}(T - m_k)$.
Since the best expert is the one that is chosen least often, the best strategy will try to maximize the number of times we choose the expert that is chosen least often. This means we want to choose all the experts equally many times, so expert k is chosen in at most T/n of the rounds. Therefore, $m_k \leq \frac{T}{n}$, thus the regret is at least $\frac{1}{T}\left(T - \frac{T}{n}\right) = \frac{n-1}{n}$.
(Note here that even if our strategy is adaptive, i.e. it chooses an expert on day i based on the losses from days 1 to $i1$, rather than committing to an expert for day i before seeing the loss for day 1, it still can't achieve regret better than $\frac{n-1}{n}$.)

(b) $(1 - \min_i p_i)$. Like in part (a), the distribution is fixed across all days, so we know ahead of time which expert will be chosen least often in expectation. Let $k = \text{argmin}_i\, p_i$ be the expert with least cost. Let $c_k^t = 0$ for all t, and let $c_k^t = 1$ for all $i \neq k$ and for all t. This way,

$$\mathbb{E}\left[\sum_{t=1}^{T} c_{i(t)}^t\right] = T\left(\sum_{i\neq k} p_i \cdot 1 + p_k \cdot 0\right) = T(1 - p_k).$$

We also have

$$\min_i \sum_{t=1}^{T} c_i^t = 0$$

so we end up with an expected regret of $\frac{1}{T}(T(1-p_k) - 0) = 1 - p_k$.
To minimize the expectation of R is the same as maximizing $\min_i p_i$, which is achieved by the uniform distribution. This gives us regret $\frac{n-1}{n}$ (this is the same worst case regret as in part (b)).

3 Variants on the Experts Problem

(a) Every time we make a mistake, at least half of the experts who have been correct so far must also make a mistake, so the set of experts who have not made a mistake decreases by at least half. This set starts at size n , and can't decrease past size 1, so we must make at most $\log n$ mistakes.

(b) Every time we make a mistake, the number of experts we stop listening to decreases in size by at least 1/2. So we go from n experts to at most $2k-1$ experts in at most $\lceil \log(n/(2k-1)) \rceil$ mistakes. At this point, the experts who are always correct are the majority, so we never make a mistake.
(We will give full credit for $\log(n/k)$, since this is the best upper bound you can get without using floor/ceiling functions; this is tight in the case where n, k are both powers of two and half the experts make a mistake every round. Other off-by-one solutions or solutions that are correct up to rounding also get full credit.)

(c) $k(\log n + 1) + \log n$. Every time we make a set of mistakes, the set of experts who have made at most m mistakes must have decreased in size by at least 1/2. When $m < k$, this set goes from size at most n to 0 before m increases by 1, so with every increase in m we make at most $\log n + 1$ mistakes. When $m = k$, this set can go from size n to 1 which takes us at most $\log n$ mistakes.

4 Weighted Rock-Paper-Scissors

(a)

Your Friend:

You:	rock	paper	scissors
rock	0	-2	1
paper	2	0	-4
scissors	-1	4	0

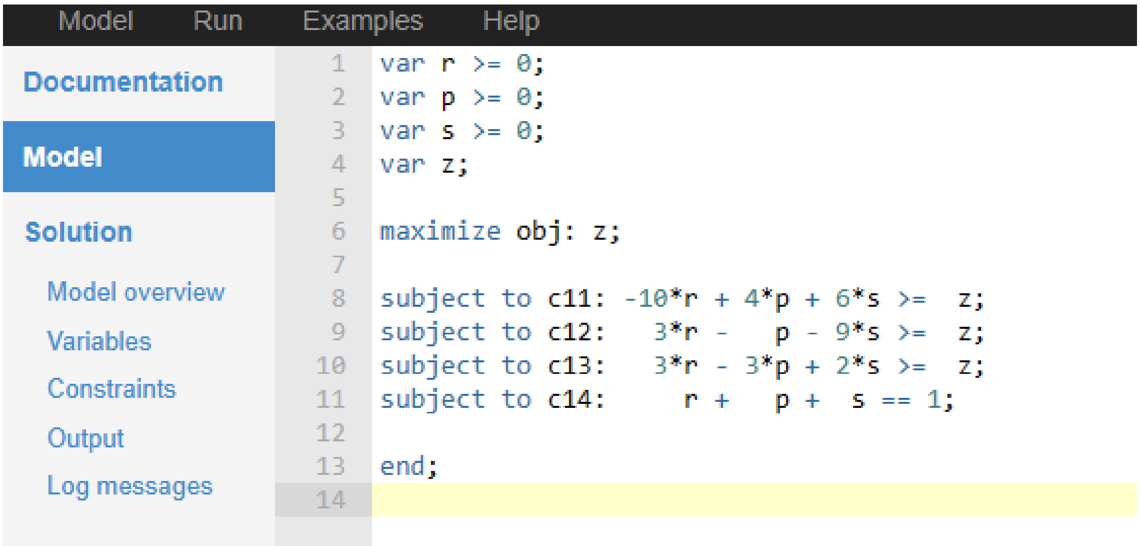
(b) Let r, p, s be the probabilities that you play rock, paper, scissors respectively. Let z stand for the expected payoff, if your opponent plays optimally as well.

$$\begin{aligned}
&\max z \\
&2p - s \geq z \quad \text{(Opponent chooses rock)} \\
&4s - 2r \geq z \quad \text{(Opponent chooses paper)} \\
&r - 4p \geq z \quad \text{(Opponent chooses scissors)} \\
&r + p + s = 1 \\
&r, p, s \geq 0
\end{aligned}$$

(c)

$$
\begin{aligned}
\max \quad & z \\
-10r + 4p + 6s &\geq z && \text{(Opponent chooses rock)} \\
3r - p - 9s &\geq z && \text{(Opponent chooses paper)} \\
3r - 3p + 2s &\geq z && \text{(Opponent chooses scissors)} \\
r + p + s &= 1 \\
r, p, s &\geq 0
\end{aligned}
$$

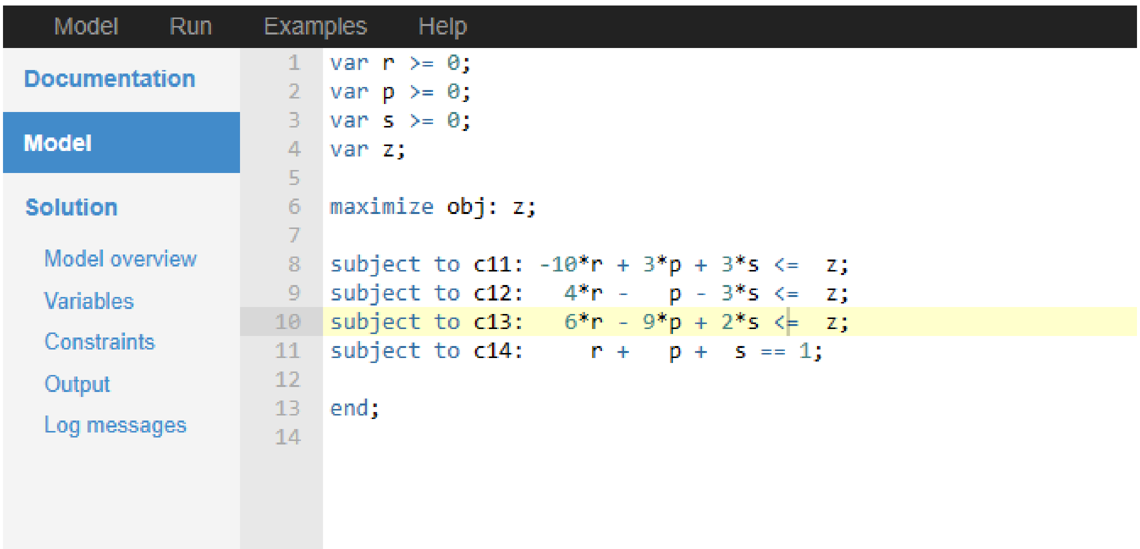
The optimal strategy is $r = 0.3346, p = 0.5630, s = 0.1024$ for an optimal payoff of -0.48.
If you are using the website suggested in this problem, here is what you should put in the model tab:



(d)

$$
\begin{aligned}
& \min z \\
& -10r + 3p + 3s \leq z \quad \text{(You choose rock)} \\
& 4r - p - 3s \leq z \quad \text{(You choose paper)} \\
& 6r - 9p + 2s \leq z \\
& r + p + s = 1 \\
& r, p, s \geq 0
\end{aligned}
$$

Your friend's optimal strategy is $r = 0.2677, p = 0.3228, s = 0.4094$. The value for this is -0.48, which is the payoff for you. The payoff for your friend is the negative of your payoff, i.e. 0.48, since the game is zero-sum.
If you are using the website suggested in this problem, here is what you should put in the model tab:



5 Domination

(a) 0. Regardless of what option the column player chooses, the row player always gets a higher payoff picking E than D, so any strategy that involves a non-zero probability of picking D can be improved by instead picking E.

(b) 0. We know that the row player is never going to pick D, i.e. will always pick either E or F. But in this case, picking B is always better for the column player than picking A (A is only better if the row player picks D). That is, conditioned on the row player playing optimally, B dominates A.

(c) Based on the previous two parts, we only have to consider the probabilities the row player picks E or F and the column player picks B or C. Looking at the 2-by-2 submatrix corresponding to these options, it follows that the optimal strategy for the row player is to pick E and F with probability 1/2, and similarly the column player should pick B, C with probability 1/2.

6 (Challenge) Follow the regularized leader

(a) Let the payoffs be at $t = 1$, $(0, 1 - \epsilon)$ and then for every other odd t be $(0, 1)$ and for each even t be $(1, 0)$.
Prior to every even numbered round, the strategy with the higher average payoff is strategy 2. Similarly, prior to ever odd numbered round, the strategy with the higher average payoff is strategy 1. But, by construction, this will yield an overall payoff of 0 as the strategies alternate in success in dissonance with the alterantion of average payoff. If you stuck to strategy 1, you would obtain a payoff of 50 and if you stuck to strategy 2, you would obtain a payoff of $50 - \epsilon$.

(b) A randomized solution is a linear combination of deterministic solutions. By the convexity of linear combinations, it can do no better than a deterministic solution.

(c)

$$\begin{aligned}
(1) &= \eta \sum_{i=1}^{n} p_t(i) \cdot \left(\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta - \ln p_t(i) \right) \\
&= \eta \sum_{i=1}^{n} p_t(i) \cdot \ln \left(e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} / p_t(i) \right) \\
&\leq \eta \cdot \ln \left(\sum_{i=1}^{n} \left(p_t(i) \cdot e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} / p_t(i) \right) \right) \\
&= \eta \cdot \ln \left(\sum_{i=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} \right).
\end{aligned}$$

(d) When using the MWU algorithm, we have:

$$\begin{aligned}
p_t(i) &= \frac{w_t(i)}{\sum_{j=1}^{n} w_t(j)} \\
&= \frac{(1-\epsilon)^{-\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]}}{\sum_{j=1}^{n} (1-\epsilon)^{-\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,j)]}}
\end{aligned}$$

If we set ϵ such that $(1 - \epsilon) = e^{-1/\eta}$, and plugin into the last equation, we have:

$$p_t(i) = \frac{e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta}}{\sum_{j=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,j)]/\eta}}$$

Therefore,

$$\begin{aligned}
(1) &= \eta \sum_{i=1}^{n} p_t(i) \cdot \ln \left(e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} / p_t(i) \right) \\
&= \eta \sum_{i=1}^{n} p_t(i) \cdot \ln \left(\sum_{i=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} \right) \\
&= \eta \cdot \ln \left(\sum_{i=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} \right).
\end{aligned}$$