Problem 1

Let us first define some notations. We define $\widehat{L}_S(Q) = \mathbb{E}_Q[\widehat{L}_S(h)]$ and $L(Q) = \mathbb{E}_Q[L(h)]$.

(a) Let $g(h) = (m-1)D(\widehat{L}_S(h)||L(h))$. We use Donsker-Varadhan inequality on distributions P and Q and obtain:

$$\begin{aligned}
\mathbb{E}_Q[g(h)] &\leq D(Q||P) + \log\left(\mathbb{E}_P[\exp(g(h))]\right) \\
\implies \mathbb{E}_Q[(m-1)D(\widehat{L}_S(h)||L(h))] &\leq D(Q||P) + \log\left(\mathbb{E}_P[\exp((m-1)D(\widehat{L}_S(h)||L(h)))]\right) \\
\implies (m-1)D(\widehat{L}(Q)||L(Q)) &\leq D(Q||P) + \log\left(\mathbb{E}_P[\exp((m-1)D(\widehat{L}_S(h)||L(h)))]\right)
\end{aligned}$$

The last inequality holds because KL Divergence is convex in both arguments and using Jensen's inequality, we obtain $D(\widehat{L}_S(Q)||L(Q)) \leq \mathbb{E}_Q[D(\widehat{L}_S(h)||L(h))]$.

(b) Now, we proceed to prove the upper bound for the second term in the RHS of the expression.

I We have the $f(\cdot)$ is an non-negative non-increasing function. So,

$$P(e^{(m-1)f(X)} \geq e^{(m-1)f(\epsilon)}) = P(f(X) \geq f(\epsilon)) = P(X \leq \epsilon) \leq e^{-mf(\epsilon)} \tag{1}$$

Now, as $e^{(m-1)f(X)}$ is a non-negative random variable, we have

$$\mathbb{E}[e^{(m-1)f(X)}] = \int_0^\infty P(e^{(m-1)f(X)} \geq t)dt$$

Let $t = e^{(m-1)f(\epsilon)} \implies dt = (m-1)e^{(m-1)f(\epsilon)}f'(\epsilon)d\epsilon$. So,

$$\begin{aligned}
\mathbb{E}[e^{(m-1)f(X)}] &= \int_0^\infty P(e^{(m-1)f(X)} \geq t)dt \\
&= \int_0^\infty P(e^{(m-1)f(X)} \geq e^{(m-1)f(\epsilon)})(m-1)e^{(m-1)f(\epsilon)}f'(\epsilon)d\epsilon \\
&\overset{(1)}{\leq} (m-1)\int_0^\infty e^{-mf(\epsilon)}e^{(m-1)f(\epsilon)}f'(\epsilon)d\epsilon \\
&= (m-1)\int_0^\infty e^{-f(\epsilon)}f'(\epsilon)d\epsilon \\
&\leq m\left(e^{-f(0)} - e^{-f(\infty)}\right) \\
&\leq m
\end{aligned} \tag{2}$$

The last inequality follows as $e^{-f(0)} - e^{-f(\infty)} \leq e^{-f(0)}$ and $f(0) \in [0.\infty) \implies -f(0) \in (-\infty, 0] \implies e^{-f(0)} \in (0, 1]$.

II Let $f(\epsilon) = D^+(\epsilon||p)$ where $X_1, \ldots, X_m$ are i.i.d random variables in $[0,1]$ and $p = \mathbb{E}[X_1]$. By definition, $f(\epsilon) \geq 0$.

Now, we show that $f(\epsilon)$ is non-increasing in $[0,1]$. For, $\epsilon \geq p$, $f'(\epsilon) = 0$ by definition. As $f(\epsilon)$ differentiable for $\epsilon \leq p$, we show that $f'(\epsilon) \leq 0$ in that range.

Proof. By defintion, for $\epsilon \leq p, f(\epsilon) = D(\epsilon||p) = \epsilon \log\left(\frac{\epsilon}{p}\right) + (1-\epsilon)\log\left(\frac{1-\epsilon}{1-p}\right)$.

$$\begin{aligned} f'(\epsilon) &= 1 + \log\epsilon - \log p - 1 - \log(1-\epsilon) + \log(1-p) \\ &= \log\left(\frac{\epsilon}{1-\epsilon}\right) - \log\left(\frac{p}{1-p}\right) \\ &= \log\left(\frac{\epsilon - p\epsilon}{p - p\epsilon}\right) \\ &\leq \log\left(\frac{\epsilon}{p}\right) \leq 0 && \text{If } a \leq b, \frac{a-ab}{b-ab} \leq \frac{a}{b} \end{aligned}$$

So, $f(\epsilon) = D^+(\epsilon||\mathbb{E}[X_1])$ is a non-increasing and non-negative function.

□

Let $X_i = \ell(h, z_i)$, where z_i is the ith training sample, then $\bar{X} = \widehat{L}_S(h)$ and $\mathbb{E}[X] = L(h)$. So, using Chernoff-Hoeffding's inequality, we have:

$$P(\widehat{L}_S(h) \leq \epsilon) \leq \exp(-mD^+(\epsilon||L(h))) \tag{3}$$

Using result (2), (3) and considering $f(\widehat{L}_S(h)) = D^+(\widehat{L}_S(h)||L(h))$, we obtain:

$$\mathbb{E}_{S\sim\mathcal{D}^m}[e^{(m-1)D^+(\widehat{L}_S(h)||L(h))}] \leq m \tag{4}$$

III Consider the random variable $Y_S = \mathbb{E}_{h\sim P}[e^{(m-1)D^+(\widehat{L}_S(h)||L(h))}]$ which depends on the sample S from the distribution. Using Markov's inequality, we get:

$$P\left(Y_S \geq \frac{2\mathbb{E}_{S\sim\mathcal{D}^m}[Y_S]}{\delta}\right) \leq \frac{\delta}{2} \implies P\left(Y_S \leq \frac{2\mathbb{E}_{S\sim\mathcal{D}^m}[Y_S]}{\delta}\right) \geq 1 - \frac{\delta}{2} \tag{5}$$

Now,

$$\begin{aligned} \mathbb{E}_{S\sim\mathcal{D}^m}[Y_S] &= \mathbb{E}_{S\sim\mathcal{D}^m}[\mathbb{E}_{h\sim P}[e^{(m-1)D^+(\widehat{L}_S(h)||L(h))}]] && \text{By definition} \\ &= \mathbb{E}_{h\sim P}[\mathbb{E}_{S\sim\mathcal{D}^m}[e^{(m-1)D^+(\widehat{L}_S(h)||L(h))}]] && \text{Fubini's theorem} \\ &\overset{(4)}{\leq} \mathbb{E}_{h\sim P}[m] = m \end{aligned} \tag{6}$$

From (5) and (6), with probability at least $1 - \frac{\delta}{2}$, $\mathbb{E}_{h\sim P}[e^{(m-1)D^+(\widehat{L}_S(h)||L(h))}] \leq \frac{2m}{\delta}$.

■

Problem 2

We have a function class $S_{\alpha,\gamma}(C_{\max}, L)$ which is defined as:

$$S_{\alpha,\gamma}(C_{\max}, L) = \{f : [0,1] \to \mathbb{R} : |f^{(j)}|_\infty \leq C_{\max}, \forall 0 \leq j \leq \alpha, \text{and}$$
$$|f^\alpha(x) - f^\alpha(y)| \leq L|x-y|^\gamma, \forall x, y \in [0,1]\}$$

We have $y_i = f^*(x_i) + \epsilon_i, \forall i \in [n]$ where $\epsilon \sim \mathcal{N}(0,1)$ and $\widehat{f}$ is the minimizer of $\sum_{i=1}^n (y_i - f(x_i))^2$ for $f \in S_{\alpha,\gamma}(C_{\max}, L)$. Let $\left\|\widehat{f} - f^*\right\|_n = \sqrt{\frac{1}{n}\sum_{i=1}^n (\widehat{f}(x_i) - f^*(x_i))^2}$. So, $\texttt{MSE} = \mathbb{E}\left[\left\|\widehat{f} - f^*\right\|_n^2\right]$. Furthermore, for $f \in S_{\alpha,\gamma}(C_{\max}, L)$, we have $\|f\|_n = \sqrt{\frac{1}{n}\sum_{i=1}^n f(x_i)^2} \leq \sqrt{\frac{1}{n}\sum_{i=1}^n C_{\max}^2} = C_{\max}$. For ease of notation, let $S_{\alpha,\gamma}(C_{\max}, L$ be denoted by $\mathcal{F}$ hereafter.

Lemma 1. $\left\|\widehat{f} - f^*\right\|_n^2 \leq \frac{2}{n}\sum_{i=1}^n \epsilon_i(\widehat{f}(x_i) - f^*(x_i))$

Proof.

$$\begin{aligned}
\left\|\widehat{f} - f^*\right\|_n^2 &= \frac{1}{n}\sum_{i=1}^n (\widehat{f}(x_i) - f^*(x_i))^2 \\
&= \frac{1}{n}\sum_{i=1}^n (\widehat{f}(x_i) - y_i + y_i - f^*(x_i))^2 \\
&= \frac{1}{n}\sum_{i=1}^n (\widehat{f}(x_i) - y_i)^2 + (y_i - f^*(x_i))^2 + 2(\widehat{f}(x_i) - y_i)(f^*(x_i) - y_i) \\
&\leq \frac{2}{n}\sum_{i=1}^n (y_i - f^*(x_i))^2 + (\widehat{f}(x_i) - y_i)(f^*(x_i) - y_i) \\
&= \frac{2}{n}\sum_{i=1}^n \epsilon_i^2 + \epsilon_i(\widehat{f}(x_i) - f^*(x_i) + f^*(x_i) - y_i) \\
&= \frac{2}{n}\sum_{i=1}^n \epsilon_i^2 + \epsilon_i(\widehat{f}(x_i) - f^*(x_i) - \epsilon_i) \\
&= \frac{2}{n}\sum_{i=1}^n \epsilon_i(\widehat{f}(x_i) - f^*(x_i))
\end{aligned}$$

□

Now, using Lemma 1

$$\left\|\widehat{f} - f^*\right\|_n^2 \leq \frac{2}{n}\sum_{i=1}^n \epsilon_i(\widehat{f}(x_i) - f^*(x_i)) \leq \frac{4}{n}\sup_{f\in\mathcal{F}}\sum_{i=1}^n \epsilon_i f(x_i)$$
$$\implies \texttt{MSE} = \mathbb{E}\left[\left\|\widehat{f} - f^*\right\|_n^2\right] \leq 4\cdot\mathbb{E}\left[\frac{1}{n}\sup_{f\in\mathcal{F}}\sum_{i=1}^n \epsilon_i f(x_i)\right] \tag{7}$$

Theorem 2 (Dudley's Entropy Integral)**.** *The Gaussian complexity of as function class $\mathcal{F}$ can be upper bounded as the covering number of the function class in sup-norm as:*

$$\mathbb{E}\left[\frac{1}{n}\sup_{f\in\mathcal{F}}\langle\epsilon, f\rangle\right] \leq \frac{12}{\sqrt{n}}\int_0^D \sqrt{\log N(\delta, \mathcal{F}, \|\cdot\|_\infty)}d\delta$$

where D is the diameter of $\mathcal{F}$.

Proof. Let V_j be a $2^{-j}D$-cover of $\mathcal{F}$ in ℓ_∞ norm. So, by definition, for any $f \in \mathcal{F}$, $\exists v_f^j \in V_j$, we have $\left\|f - v_f^j\right\|_n \leq 2^{-j}D$. Now, from the definition of Gaussian complexity, we have (ignoring the factor of $1/n$):

$$\begin{aligned}
\mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f\rangle\right] &= \mathbb{E}\left[\sup_{f\in\mathcal{F}}\left(\langle\epsilon, f - v_f^m\rangle + \sum_{j=1}^m \langle\epsilon, v_f^j - v_f^{j-1}\rangle\right)\right] \\
&\leq \mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f - v_f^m\rangle\right] + \sum_{j=1}^m \mathbb{E}\left[\sup_{v\in V_j, v'\in V_{j-1}}\langle\epsilon, v - v'\rangle\right] \\
&\overset{(a)}{\leq} \mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f - v_f^m\rangle\right] + \sum_{j=1}^m \sigma_j\sqrt{2\log|V_j||V_{j-1}|} \\
&\overset{(b)}{\leq} \mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f - v_f^m\rangle\right] + \sum_{j=1}^m 2\sigma_j\sqrt{\log|V_j|}
\end{aligned}$$

The statement (a) follows from the definition of Massart's Lemma
(b) holds because by construction, we have $|V_j| \geq |V_{j-1}|$.

Now, $\sigma_j = \sqrt{\mathrm{Var}(\langle\epsilon, v - v'\rangle)} = \sqrt{n}\|v - v'\|_n = \sqrt{n}\|v - f + f - v'\|_n \leq \sqrt{n}(\|f - v'\|_n + \|f - v\|_n)$. The last inequality follows from the triangle inequality for the $\|\cdot\|_n-$norm. Furthermore, by definition of v, v', we have $\sqrt{n}(\|f - v'\|_n + \|f - v\|_n) \leq \sqrt{n}(2^{-j}D + 2^{-(j-1)}D) = 3\cdot 2^{-j}\sqrt{n}D \implies \sigma_j \leq 3\cdot 2^{-j}\sqrt{n}D$. So,

$$\begin{aligned}
\mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f\rangle\right] &\leq \mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f - v_f^m\rangle\right] + 6\sqrt{n}D\sum_{j=1}^m 2^{-j}D\sqrt{\log N(2^{-j}, \mathcal{F}, \|\cdot\|_\infty)} \\
&\leq \mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f - v_f^m\rangle\right] + 12\sqrt{n}D\sum_{j=1}^m (2^{-j} - 2^{-(j-1)})\sqrt{\log N(2^{-j}D, \mathcal{F}, \|\cdot\|_\infty)} \\
&\leq \mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f - v_f^m\rangle\right] + 12\sqrt{n}\sum_{j=1}^m \int_{2^{-j}D}^{2^{-(j-1)D}} \sqrt{\log N(\delta, \mathcal{F}, \|\cdot\|_\infty)}d\delta \\
&\leq \mathbb{E}\left[\sup_{f\in\mathcal{F}}\langle\epsilon, f - v_f^m\rangle\right] + 12\sqrt{n}\int_{2^{-m}D}^{D} \sqrt{\log N(\delta, \mathcal{F}, \|\cdot\|_\infty)}d\delta
\end{aligned}$$

Now, if $m \to \infty, 2^{-m} \to 0$. Further, $\left\|f - v_f^m\right\|_n \leq 2^{-m}D \to 0$, so, $|f - v_f^m| \to 0 \implies$ the first term on RHS vanishes. Finally, dividing both sides by n, we get:

$$\mathbb{E}\left[\frac{1}{n}\sup_{f\in\mathcal{F}}\langle\epsilon, f\rangle\right] \leq \frac{12}{\sqrt{n}}\int_0^D \sqrt{\log N(\delta, \mathcal{F}, \|\cdot\|_\infty)}d\delta$$

□

It is known that:

$$\log N(\delta, \mathcal{F}, \|\cdot\|_\infty) \leq C\left(\frac{1}{\delta}\right)^{1/(\alpha+\gamma)} \tag{8}$$

Now, we compute the diameter of $\mathcal{F}$. $D = \sup_{f,f'\in\mathcal{F}}|f - f'| \leq 2|f|_\infty \leq 2C_{\max}$. Using (7), (8) and Theorem 2, we obtain:

$$\begin{aligned}
\texttt{MSE} &= \mathbb{E}\left[\left\|\widehat{f} - f^*\right\|_n^2\right] \\
&\leq 4\cdot\mathbb{E}\left[\frac{1}{n}\sup_{f\in\mathcal{F}}\sum_{i=1}^n \epsilon_i f(x_i)\right] \\
&\leq \frac{48}{\sqrt{n}}\int_0^D \sqrt{\log N(\delta, \mathcal{F}, \|\cdot\|_\infty)}d\delta \\
&= 48\sqrt{\frac{C}{n}}\int_0^{2C_{\max}} \delta^{-\frac{1}{2(\alpha+\gamma)}}d\delta \\
&= 48\sqrt{\frac{C}{n}}(2C_{\max})^{1-\frac{1}{2(\alpha+\gamma)}} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

■

Problem 3

We define w'_{t+1} as follows:

$$\nabla\psi(w'_{t+1}) = \nabla\psi(w_t) - \eta\nabla f_t(w_t) \tag{9}$$

Regarding Bregman divergence, it is to be noted that $D_\psi(u, w) \geq 0$ and equality is attained iff $u = w$. This can be proved as follows: by definition, Bregman divergence is defined for strictly convex function, $\implies$ for $u \neq w$, $\psi(u) > \psi(w) + \langle\nabla\psi(w), u - w\rangle \implies D_\psi(u, w) > 0$. For $u = w, D_\psi(u, w) = 0$.

(a) By definition:

$$\begin{aligned}
w_{t+1} &= \operatorname*{argmin}_{w\in\Omega}\left[\langle w, \nabla f_t(w_t)\rangle + \frac{1}{\eta}D_\psi(w, w_t)\right] \\
&= \operatorname*{argmin}_{w\in\Omega}\left[\langle w, \eta\nabla f_t(w_t)\rangle + D_\psi(w, w_t)\right] \\
&\overset{(9)}{=} \operatorname*{argmin}_{w\in\Omega}\left[\langle w, \nabla\psi(w_t) - \nabla\psi(w'_{t+1})\rangle + D_\psi(w, w_t)\right] \\
&= \operatorname*{argmin}_{w\in\Omega}\left[\langle w, \nabla\psi(w_t) - \nabla\psi(w'_{t+1})\rangle + \psi(w) - \psi(w_t) - \langle\nabla\psi(w_t), w - w_t\rangle\right] \\
&= \operatorname*{argmin}_{w\in\Omega}\left[-\langle w, \nabla\psi(w'_{t+1})\rangle + \psi(w)\right] \\
&= \operatorname*{argmin}_{w\in\Omega}\left[\psi(w) - \psi(w'_{t+1}) - \langle\nabla\psi(w'_{t+1}), w - w'_{t+1}\rangle\right] \\
&= \operatorname*{argmin}_{w\in\Omega} D_\psi(w, w'_{t+1})
\end{aligned} \tag{10}$$

(b) We begin from RHS:

$$\begin{aligned}
D_\psi(u, w_t) - D_\psi(u, w'_{t+1}) + D_\psi(w_t, w'_{t+1}) &= \psi(u) - \psi(w_t) - \langle\nabla\psi(w_t), u - w_t\rangle \\
&\quad - \psi(u) + \psi(w'_{t+1}) + \langle\nabla\psi(w'_{t+1}), u - w'_{t+1}\rangle \\
&\quad + \psi(w_t) - \psi(w'_{t+1}) - \langle\nabla\psi(w'_{t+1}), w_t - w'_{t+1}\rangle \\
&= -\langle\nabla\psi(w_t), u - w_t\rangle + \langle\nabla\psi(w'_{t+1}), u - w'_{t+1}\rangle \\
&\quad - \langle\nabla\psi(w'_{t+1}), w_t - w'_{t+1}\rangle \\
&= -\langle\nabla\psi(w_t), u - w_t\rangle + \langle\nabla\psi(w'_{t+1}), u - w_t\rangle \\
&= \langle\nabla\psi(w'_{t+1}) - \nabla\psi(w_t), u - w_t\rangle \\
&\overset{(9)}{=} \langle\eta\nabla f_t(w_t), w_t - u\rangle
\end{aligned} \tag{11}$$

Dividing both sides by η, we get the desired result.

(c) From (10), we have $w_{t+1} = \operatorname{argmin}_{w\in\Omega} D_\psi(w, w'_{t+1})$. By definition, $w'_{t+1} \in \Omega \implies D_\psi(w, w'_{t+1})$ attains the global minima in Ω and hence $w_{t+1} = w'_{t+1}$. So,

$$D_\psi(u, w_{t+1}) \leq D_\psi(u, w'_{t+1}), \forall u \tag{12}$$

(d) As $f_t(\cdot)$ is convex,

$$\begin{aligned}
f_t(w_t) - f_t(u) &\le \langle \nabla f_t(w_t), w_t - u\rangle \\
&\overset{(11)}{=} \frac{1}{\eta}\left(D_\psi(u, w_t) - D_\psi(u, w'_{t+1}) + D_\psi(w_t, w'_{t+1})\right) \\
&\overset{(12)}{\le} \frac{1}{\eta}\left(D_\psi(u, w_t) - D_\psi(u, w_{t+1}) + D_\psi(w_t, w'_{t+1})\right)
\end{aligned}$$

Adding terms from $t = 1$ to $t = T$, we see that the first two terms on the RHS telescopes.

$$\begin{aligned}
\sum_{t=1}^{T}(f_t(w_t) - f_t(u)) &\le \sum_{t=1}^{T}\frac{1}{\eta}\left(D_\psi(u, w_t) - D_\psi(u, w_{t+1})\right) + \frac{1}{\eta}\sum_{t=1}^{T} D_\psi(w_t, w'_{t+1}) \\
&= \frac{1}{\eta}\left(D_\psi(u, w_1) - D_\psi(u, w_{T+1})\right) + \frac{1}{\eta}\sum_{t=1}^{T} D_\psi(w_t, w'_{t+1}) \\
&\le \frac{D_\psi(u, w_1)}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T} D_\psi(w_t, w'_{t+1}) \qquad (13)
\end{aligned}$$

(e) Let $\Omega = \Delta_N$ (N-dimensional simplex). Let $f_t(w_t) = \langle l_t, w_t\rangle$ where l_t is the loss vector for time t and $\psi(w) = \sum_i w_i \log w_i$. Clearly, $\nabla\psi(w) = [1 + \log w_1, \dots, 1 + \log w_N]^\top$ is invertible. So,

$$\begin{aligned}
D_\psi(u, w) &= \psi(u) - \psi(w) - \langle \nabla\psi(w), u - w\rangle \\
&= \sum_i u_i \log u_i - \sum_i w_i \log w_i - \sum_i (1 + \log w_i)(u_i - w_i) \\
&= \sum_i u_i \log \frac{u_i}{w_i} \\
&= D(u||w)
\end{aligned}$$

Let $u = e_j$ such that $j = \text{argmin}_{i\in[N]} \sum_{t=1}^{T}\langle l_t, e_i\rangle$ where e_i's are N-dimensional unit vectors. Now, we compute w'_{t+1}. Using (9), we get:

$$1 + \log w'_{t+1,i} = 1 + \log w_{t,i} - \eta l_{t,i} \implies w'_{t+1,i} \propto w_{t,i}e^{-\eta l_{t,i}} \implies w'_{t+1,i} = \frac{w_{t,i}e^{-\eta l_{t,i}}}{\sum_{j=1}^{N} w_{t,j}e^{-\eta l_{t,j}}}$$

Let $w_1 = \frac{1}{N}\mathbf{1}$. So, (13) becomes:

$$\begin{aligned}
R_T = \sum_{t=1}^{T}\langle l_t, w_t\rangle - \min_{i\in[N]}\sum_{t=1}^{T}\langle l_t, e_i\rangle &\le \frac{D(u||w_1)}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T} D(w_t||w'_{t+1}) \\
&= \frac{\log N}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T}\sum_{i=1}^{N} w_{t,i}\log\frac{w_{t,i}}{w'_{t+1}} \\
&= \frac{\log N}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T}\sum_{i=1}^{N} w_{t,i}\log\frac{w_{t,i}(\sum_{j=1}^{N} w_{t,j}e^{-\eta l_{t,j}})}{w_{t,i}e^{-\eta l_{t,i}}}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\log N}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T} \eta\langle l_t, w_t\rangle + \log\left(\sum_{j=1}^{N} w_{t,j} e^{-\eta l_{t,j}}\right) \\
&\overset{(a)}{\leq} \frac{\log N}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T} \eta\langle l_t, w_t\rangle + \log\left(\sum_{j=1}^{N} w_{t,j}(1 - \eta l_{t,j} + \eta^2 l_{t,j}^2)\right) \\
&= \frac{\log N}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T} \eta\langle l_t, w_t\rangle + \log\big(1 - \eta\langle l_t, w_t\rangle + \eta^2\langle w_t, l_t^2\rangle\big) \\
&\overset{(b)}{\leq} \frac{\log N}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T} \eta\langle l_t, w_t\rangle - \eta\langle l_t, w_t\rangle + \eta^2\langle w_t, l_t^2\rangle \\
&= \frac{\log N}{\eta} + \eta\sum_{t=1}^{T}\langle w_t, l_t^2\rangle \\
&\overset{(c)}{\leq} \frac{\log N}{\eta} + \eta T
\end{aligned}
$$

The inequality (a) follows because $e^{-x} \leq 1 - x + x^2, x \geq 0$,
(b) follows as $\ln(1 + x) \leq x, x \in \mathbb{R}$, and
(c) holds as $\langle w_t, l_t^2\rangle = \sum_{i=1}^{N} w_{t,i} l_{t,i}^2 \leq \sum_{i=1}^{N} w_{t,i} = 1$ as $0 \leq l_{t,i} \leq 1$.

Setting $\eta^* = \sqrt{\frac{\log N}{T}}$, we obtain $R_T \leq \mathcal{O}(\sqrt{T \log N})$, the regret bound for Hedge algorithm.

■

Problem 4

In this question, we compare the performance of three bandit algorithms namely: ϵ-greedy, UCB1 and Exp3. We have four Bernoulli bandits with expected rewards as: $\boldsymbol{p} = [0.5, 0.95, 0.2, 0.8]$. We have set $\epsilon = 0.05$ for the ϵ-greedy algorithm and $\gamma = 0.05$ for the Exp3 algorithm (γ is the random exploration probability).

(a) In this part, we plot the average fraction of times any arm a was selected by algorithm π by some time instant t in semi-log scale in Figure 1,2 and 3. As $T \to \infty$, we can expect that probability of choosing the arm with maximum reward approaches 1 and the choosing rest of the arms tends to 0.

Furthermore, from these plots we can observe that for $T \gg 1$, the order of $N_a^\pi(t)$ for arms is according to the expected reward of the arm, i.e, $N_{a_2}^\pi(t) > N_{a_4}^\pi(t) > N_{a_1}^\pi(t) > N_{a_3}^\pi(t)$ as $p(a_2) > p(a_4) > p(a_1) > p(a_3)$ where $p(a)$ denotes the average reward of arm a.

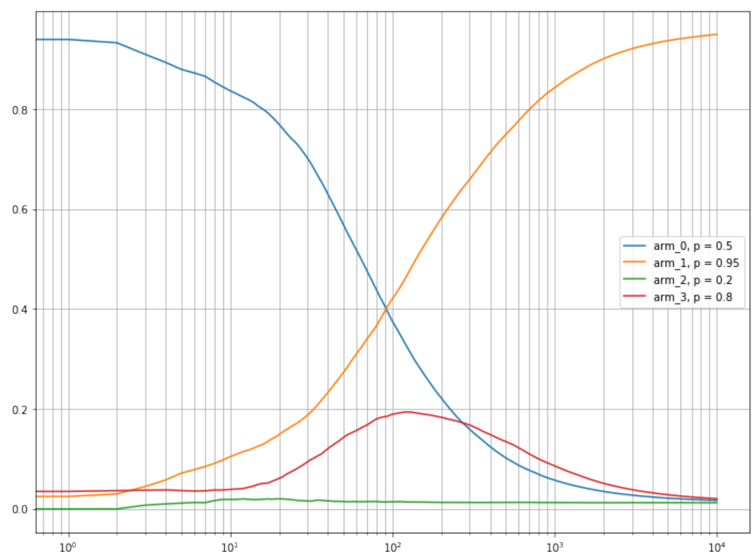

Figure 1: $N_a^\pi(t)$ for ϵ-greedy algorithm

(b) In this part, we compare the pseudo-regret, $R^\pi(t)$, of each algorithm. Concretely,

$$R^\pi(t) = t \max_a p(a) - \sum_{\tau=1}^{t} \bar{r}^\pi(\tau) \tag{14}$$

where $\bar{r}^\pi(t)$ is the average reward obtained by algorithm π at time t. The time-evolution of $R^\pi(t)$ is plotted in Figure 4. From the plot we can observe that all the algorithms achieve sublinear regret and UCB1 achieves the best regret bound.

(c) In this part, we check the sensitivity of the pseudo-regret with hyperparameters of the algorithms.

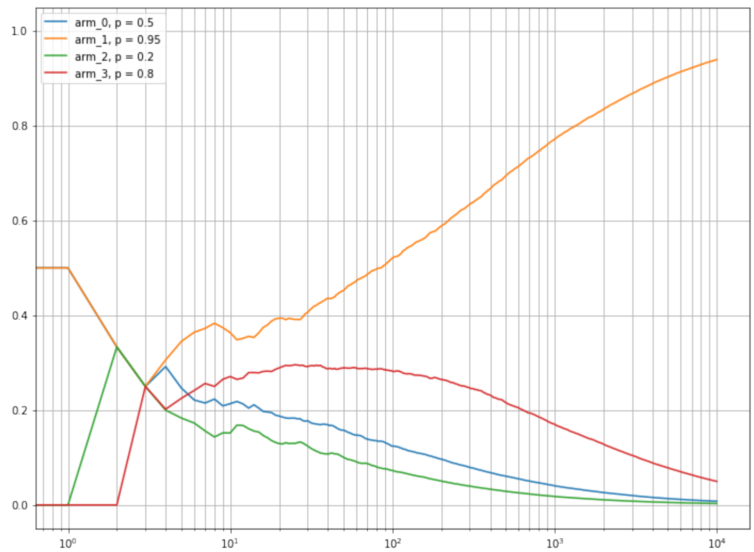

Figure 2: $N_a^\pi(t)$ for UCB1 algorithm

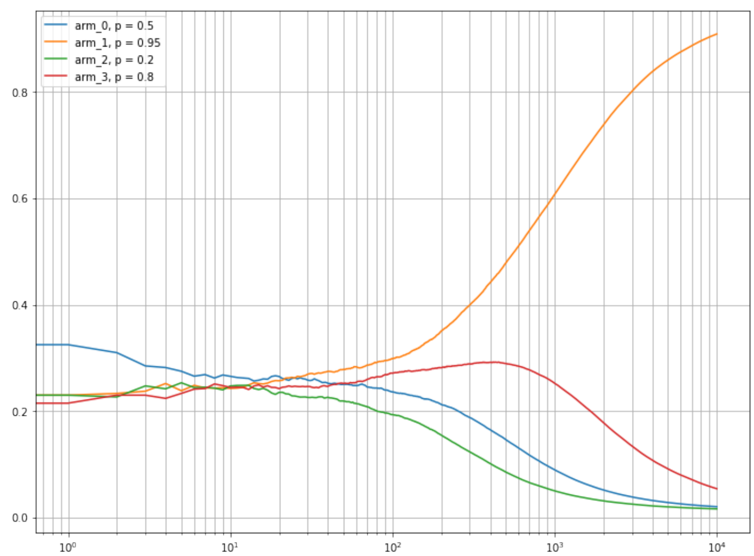

Figure 3: $N_a^\pi(t)$ for Exp3 algorithm

(i) Variation of ϵ in ϵ-greedy algorithm:
We plot the variation of pseudo-regret with time as ϵ is varied in Figure 5 for five values of $\epsilon = [0.02, 0.05, 0.1, 0.2, 0.5]$. From the plot, we observe that $\epsilon = 0.02$ achieves the best regret bound till $T = 10^4$ steps.

(ii) Variation of γ in Exp3 algorithm:
We plot the variation of pseudo-regret with time as γ is varied in Figure 6. The best

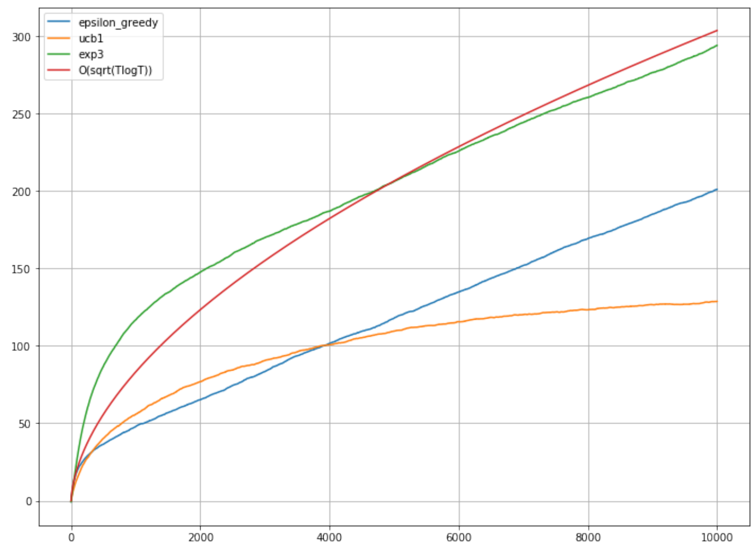

Figure 4: Time-evolution of $R^{\pi}(t)$ for each algorithm. For comparison, we also plot the theoretical regret bound for most bandit algorithms: $\mathcal{O}(\sqrt{T \log T})$

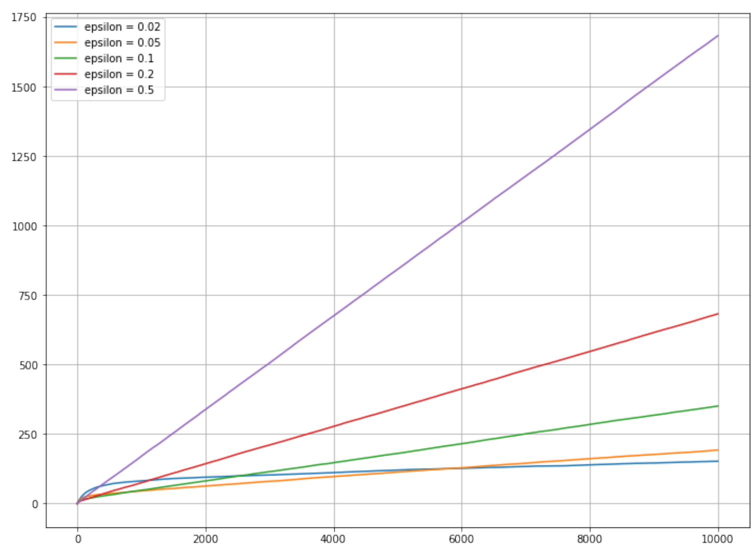

Figure 5: Time-evolution of pseudo-regret as ϵ is varied for ϵ-greedy algorithm.

performing value of γ in Exp3 algorithm is obtained according to[1]:

$$\gamma = \min \left\{ 1, \sqrt{\frac{N \log N}{(e-1)G}} \right\} \tag{15}$$

where G is the maximum achievable reward and N is the number of arms. Substituting the values, we obtain $\gamma^* = 0.018$. But, in Figure 6, we observe that $\gamma = 0.05$ attains the

[1] https://www.cs.princeton.edu/courses/archive/fall16/cos402/lectures/402-lec22.pdf

best regret till time $T = 10^4$. It is to be also noted that the growth rate for $\gamma = 0.02$ is slower than $\gamma = 0.05$. So, as $T \to \infty$, we can expect $\gamma = 0.02$ achieves the best regret.

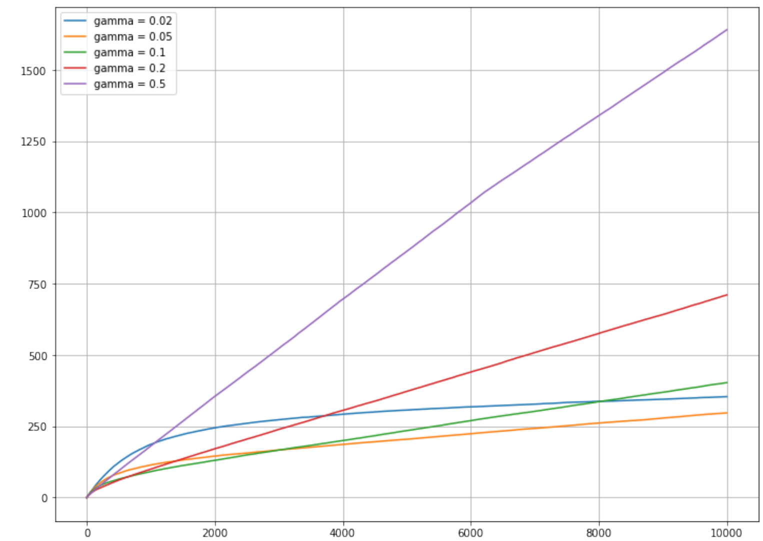

Figure 6: Time-evolution of pseudo-regret as γ is varied for Exp3 algorithm.

■

Problem 5

Let $C(a) = \texttt{cost}(a) = \sum_{t=1}^{T} c_t(a)$ and $C(a^*) = \min_a \texttt{cost}(a)$. Furthermore, let $\mathbb{E}[\texttt{cost}(a)] = Tc(a)$.

(a) We do an "clean event" analysis by using Hoeffding's inequality on the cost sequence of each arm. For any arm a, using Hoeffding's inequality, we have:

$$P\left(\left|\frac{1}{T}\sum_{t=1}^{T} c_t(a) - c(a)\right| \geq \sqrt{\frac{1}{2T}\log\frac{2K}{\delta}}\right) \leq \frac{\delta}{K}$$
$$\implies P\left(\left|\sum_{t=1}^{T} c_t(a) - \mathbb{E}[\texttt{cost}(a)]\right| \geq \sqrt{\frac{T}{2}\log\frac{2K}{\delta}}\right) \leq \frac{\delta}{K}$$

So, *for all* the K arms, we get the following tail bound using union bound:

$$P\left(\left|\sum_{t=1}^{T} c_t(a) - \mathbb{E}[\texttt{cost}(a)]\right| \geq \sqrt{\frac{T}{2}\log\frac{2K}{\delta}}\right) \leq \delta$$

Choosing $\delta = 2/T^4$, we obtain:

$$P\left(\left|\sum_{t=1}^{T} c_t(a) - \mathbb{E}[\texttt{cost}(a)]\right| \leq \sqrt{2T\log KT}\right) \geq 1 - \frac{2}{T^4}$$

So, w.h.p we have:

$$\sum_{t=1}^{T} c_t(a) \geq \mathbb{E}[\texttt{cost}(a)] - \sqrt{2T\log KT} \tag{16}$$
$$\sum_{t=1}^{T} c_t(a^*) \leq \mathbb{E}[\texttt{cost}(a^*)] + \sqrt{2T\log KT} \tag{17}$$

From (16) and (17), we obtain:

$$\mathbb{E}[\texttt{cost}(a)] \leq \mathbb{E}[\texttt{cost}(a^*)] + 2\sqrt{2T\log KT}$$
$$\implies \min_a \mathbb{E}[\texttt{cost}(a)] \leq \mathbb{E}[\texttt{cost}(a^*)] + \mathcal{O}(\sqrt{T\log KT})$$

(b) Consider a scenario in which all the arms have 0-1 costs with mean 0.5. In this case,

$$\mathbb{E}[\texttt{cost(ALG)}] = \mathbb{E}\left[\sum_{t=1}^{T} c_t(a_t)\right]$$
$$= \sum_{t=1}^{T} \mathbb{E}[c_t(a_t)] = \frac{T}{2} \tag{18}$$

Now, we consider the random variable $\texttt{cost}(a)$. It is to be noted that we are dealing the case of deterministic adversary and i.i.d. costs, so, the costs in each step are independent. So,

$$\texttt{cost}(a) = X_a = X_{a,i} + \cdots + X_{a,T}$$

where $X_{a,i}, i \in [T], a \in [K]$ is Bernoulli r.v. with mean 0.5. So, $X_a \sim B(T, 1/2).\forall a \in [K]$. Now, we need to estimate $\mathbb{E}[\min_a \texttt{cost}(a)] = \mathbb{E}[\min_a X_a]$. For X_a, using Hoeffding's inequality, we have:

$$P\left(\left|\frac{1}{T}\sum_{t=1}^{T} X_{a,t} - \mathbb{E}[X_{a,1}]\right| \geq \sqrt{\frac{1}{2T}\log\frac{2K}{\delta}}\right) \leq \frac{\delta}{K}$$
$$\implies P\left(\left|X_a - \frac{T}{2}\right| \geq \sqrt{\frac{T}{2}\log\frac{2K}{\delta}}\right) \leq \frac{\delta}{K}$$

So, *for all* the K arms (random variables), we get the following tail bound using union bound:

$$P\left(\left|X_a - \frac{T}{2}\right| \geq \sqrt{\frac{T}{2}\log\frac{2K}{\delta}}\right) \leq \delta$$

So, with probability $1-\delta$, we have:

$$X_a \in \left[\frac{T}{2} - \sqrt{\frac{T}{2}\log\frac{2K}{\delta}}, \frac{T}{2} + \sqrt{\frac{T}{2}\log\frac{2K}{\delta}}\right]$$

Let $Y = \min_a X_a$. Clearly, with probability $1-\delta$, $Y \geq \frac{T}{2} - \sqrt{\frac{T}{2}\log\frac{2K}{\delta}} \implies \mathbb{E}[Y] \geq \frac{T}{2} - \sqrt{\frac{T}{2}\log\frac{2K}{\delta}}$. Now, we can choose δ such that:

$$\frac{T}{2} - \sqrt{\frac{T}{2}\log\frac{2K}{\delta}} \leq \mathbb{E}[Y] \leq \frac{T}{2} - \sqrt{\frac{T}{2}\log\frac{K}{\delta}} \tag{19}$$

As $K \geq 2$, let $\delta = 1/K^{2\gamma-1}$ for some $\gamma \geq 1$. From (19), we get

$$\begin{aligned}
\mathbb{E}[\min_a \texttt{cost}(a)] = \mathbb{E}[Y] &\leq \frac{T}{2} - \sqrt{\frac{T}{2}\log\frac{K}{\delta}} \\
&= \frac{T}{2} - \sqrt{\gamma T \log K} \\
&= \frac{T}{2} - \Omega(\sqrt{T\log K}) \\
&\overset{(18)}{=} \mathbb{E}[\texttt{cost(ALG)}] - \Omega(\sqrt{T\log K})
\end{aligned}$$
$$\implies \mathbb{E}[\texttt{cost(ALG)} - \min_a \texttt{cost}(a)] \geq \Omega(\sqrt{T\log K})$$

■