Assignment 2

Last date: February 19, 2025

- Log sum inequality: for any non negatives $a_1, \dots, a_n$ and non negatives $b_1, \dots, b_n$ we have $\sum_{i=1}^{n} a_i \ln \frac{a_i}{b_i} \geq (\sum_i a_i) \ln \frac{\sum_i a_i}{\sum_i b_i}$.
- In calculation of KL divergence, $0 \ln 0$ is considered as 0.

(1 marks each):

1. Using log sum inequality, show that $KL(P, Q) \geq 0$.
2. Consider an algorithm that always pulls arm 1 (throughout T rounds). Give an instance on which this algorithm achieves an expected regret of 0. Also give an instance on which it achieves an expected regret of T.
3. Which of the following statement(s) is/are correct (Justify briefly - no marks for just writing the answer).
   (a) The expected regret of Successive Elimination algorithm is $O(\sqrt{KT \log T})$ on all instances.
   (b) The expected regret of Successive Elimination algorithm can be $\leq (KT \log T)^{1/4}$ on some instances.
   (c) The expected regret of Successive Elimination algorithm can not be $\leq (KT \log T)^{1/4}$ on all instances
4. Assuming *Good* (as defined in UCB algorithm), show that $\mu_a \leq UCB_a(t)$ for UCB for any t.
5. Assuming *Good* (as defined in Successive Elimination algorithm), prove that the optimal arm will never be eliminated in the Succesive Elimination Algorithm.
6. In the class we show a lower bound of $\Omega(\sqrt{KT})$ on the expected regret of any algorithm. But we also see that the UCB can achieve an instance-dependent upper bound of $O(\log T) \sum_{a:\triangle_a > 0} \frac{1}{\triangle_a}$. Explain why they do not contradict each other.
7. Let $P = (p, 1-p)$ and $Q = (q, 1-q)$ be two distributions on two elements. Show
$$d_{tv}(P^{\otimes 2}, Q^{\otimes 2}) \leq 2\, d_{tv}(P, Q)$$

8. Show for any n,

$$KL(P^{\otimes n}, Q^{\otimes n}) = n \cdot KL(P, Q)$$

(2-mark)
Prove for distributions P and Q defined on $[n]$,

$$d_{tv}(P, Q) = \max_{S \subseteq [n]} |P(S) - Q(S)|$$

(1-mark)
In the first assignment, you showed that Testing Coin($\epsilon, 1/5$) problem can solved in $O(1/\epsilon^2)$ coin tosses. Show that Testing Coin(ϵ, δ) can be solved in $O(\frac{\log 1/\delta}{\epsilon^2})$ coin tosses for any $\delta > 0$.

(3-mark)
In the class, we saw a lower bound of $\frac{(1-2\delta)^2}{\epsilon^2}$ on coin tosses for deterministic algorithms for TestingCoin(ϵ, δ) problem. The above question shows, the upper bound is $O(\frac{\log 1/\delta}{\epsilon^2})$. Clearly, the upper and lower bounds are tight in terms of ϵ but not on δ. Show that the lower bound is also $\Omega(\frac{\log 1/\delta}{\epsilon^2})$ for deterministic algorithms. You can assume $\epsilon \leq 1/3$ (in fact, if you want, you can assume both $\epsilon, \delta \leq c$ for any constant $c < 1$ of your choice).

(Hint: Instead of Pinsker Inequality, use the following stronger bound

$$d_{tv}(P, Q) \leq \sqrt{1 - e^{-KL(P,Q)}}$$

)

(3+3 = 6 marks)
Consider bandit instances in which the them mean reward of all arms lie in $[1/2, 1/2 + \gamma]$ for some $\gamma > 0$. Modify Successive Elimination and UCB algorithm to achieve better regret bounds (both instance independent and instance dependent bounds) for these instances. Just to clarify, the algorithm knows the value of γ. Your answer should have both description of the modified algorithms as well as their regret analysis.

(3 marks)
Consider a following problem : input consists of K coins and $\epsilon > 0$. It is promised that either

- all coins are fair or
- there is exactly one coin which is biased and have $\Pr(H) = 1/2 + \epsilon$ and rest other coins are fair.

The goal is to correctly determine the one of the above possibility with at least 4/5 probability. That is, if the input coins satisfy the first item, algorithm should return 'Fair' with at least 4/5 probability. On the other hand, if the input coins satisfy the second item, algorithm should return 'Biased' with at least 4/5 probability.

Modify the lower bound proof for Biased Coin Identification problem to show the same lower bound (of $\Omega(K/\epsilon^2)$) for the above problem.

Prove the Pinsker Inequality:

- (2-mark)
  Let $P = (p, 1-p)$ and $Q = (q, 1-q)$ be two binary distributions (distributions on two elements). Show

  $$d_{tv}(P,Q) \leq \sqrt{\frac{1}{2} KL(P,Q)}$$

  (Hint: use elementary calculus)

- (1-mark)
  Consider two distributions $P = (p_1, \dots, p_n)$ and $Q = (q_1, \dots, q_n)$ on $[n]$. Consider any set $S \subseteq [n]$. Let $P' = (P(S), 1-P(S))$ and $Q' = (Q(S), 1-Q(S))$ be binary distributions where recall that $P(S) = \sum_{i \in S} p_i$ and $Q(S) = \sum_{i \in S} q_i$. Show

  $$KL(P', Q') \leq KL(P,Q)$$

  (Hint: use log sum inequality)

- (2-mark)
  Consider two distributions $P = (p_1, \dots, p_n)$ and $Q = (q_1, \dots, q_n)$ on $[n]$. Show that

  $$d_{tv}(P,Q) \leq \sqrt{\frac{1}{2} KL(P,Q)}$$

(Proof of $E[\delta] \leq O(\sqrt{K/T})$ for the MOSS algorithm)

Let $r_1, \dots, r_T$ be T independent samples from a distribution D with mean μ (D is supported on $[0,1]$ so all r_i and μ is in $[0,1]$). For any $1 \leq x \leq T$, let $\hat{\mu}_x = \frac{\sum_{i=1}^{x} r_i}{x}$ and

$$I_x = \hat{\mu}_x + \sqrt{\frac{\log^+(\frac{T}{Kx})}{x}}$$

where $\log^+(z) = \max(z, 0)$. Let $\delta = \max(0, \mu - \min_{1 \leq x \leq T} I_x)$

- (3-marks)
  For any y, prove that

  $$\Pr(\delta \geq y) \leq \frac{K}{T} \cdot O(\frac{1}{y^2})$$

(Hint: use Hoffding's maximal inequality - Suppose $X_1, X_2, \dots$ are iid such that each X_i is in $[0,1]$ and $E[X_i] = \mu$ then for any $t > 0$ and $m \geq 1$, we have

$$\Pr(\exists 1 \leq r \leq m : \sum_{i=1}^{r} (\mu - X_i) \geq t) \leq exp(-\frac{2t^2}{m}).$$

Note that the above inequality is stronger than the one we used throughout the class. To apply the above inequality , t should not depend on r. However, you might come up with an expression where t may depend on r. In such case, use the following trick:

$$\Pr(\exists 1 \leq r \leq m : \sum_{i=1}^{r} (\mu - X_i) \geq t(r)) \leq \sum_{j} \Pr(\exists 2^j \leq r \leq 2^{j+1} : \sum_{i=1}^{r} (\mu - X_i) \geq f(j))$$

where $t(r) \geq f(j)$ for $r \in [2^j, 2^{j+1}]$. Now with some manipulation, you should be able to apply the Hoffding's maximal inequality.

Additionally, following inequality might be useful: $\sum_{j=1}^{\infty} 2^j exp(-2^j y^2) = O(\frac{1}{y^2}))$

- (2-marks)
  Prove that $E[\delta] = O(\sqrt{K/T})$.

  (Hint: note that δ is a continuous r.v.. Define a new r.v. β such that β is discrete and $\delta \leq \beta$ and so $E[\delta] \leq E[\beta]$. Obviously, you need to use the result of the first part, i.e., $\Pr(\delta \geq y) \leq \frac{K}{T} \cdot O(\frac{1}{y^2})$ which should be hint on how to define β.)