Assignment 1

Rohan (241110057)

Prove the following statements (1 mark each):

1. **For any two events A and B, $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.**

   Solution:

   Starting with the union of A and B:

   $$P(A \cup B)$$

   Breaking down the union into disjoint events (A only, B only, $A \cap B$):

   $$P(A \cup B) = P((A \setminus B) \cup (B \setminus A) \cup (A \cap B))$$

   Applying the rule for disjoint events (sum of probabilities):

   $$P(A \cup B) = P(A \setminus B) + P(B \setminus A) + P(A \cap B)$$

   Expressing probabilities of A and B:

   $$P(A) = P(A \setminus B) + P(A \cap B)$$

   $$P(B) = P(B \setminus A) + P(A \cap B)$$

   Substituting into the formula:

   $$P(A \cup B) = (P(A) - P(A \cap B)) + (P(B) - P(A \cap B)) + P(A \cap B)$$

   Simplifying:

   $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

   Hence proved.

2. **For any two random variables X and Y, $E[X + Y] = E[X] + E[Y]$.**

   Solution:

   By definition of expectation:

   $$E[X + Y] = \int_{-\infty}^{\infty} (x + y) f_{X,Y}(x, y) \, dx \, dy$$

   where $f_{X,Y}(x, y)$ is the joint probability density function of X and Y.

   Distributing the sum inside the integral:

   $$E[X + Y] = \int_{-\infty}^{\infty} x f_{X,Y}(x, y) \, dx \, dy + \int_{-\infty}^{\infty} y f_{X,Y}(x, y) \, dx \, dy$$

Recognize the marginal distributions:

The first term is the expectation of X:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

where $f_X(x)$ is the marginal probability density function of X.

The second term is the expectation of Y:

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy$$

where $f_Y(y)$ is the marginal probability density function of Y.

Combining the results:

$$E[X+Y] = E[X] + E[Y]$$

Hence Proved.

3. $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$**, where** c **is a constant and** X **is a random variable.**

   Solution:

   By the definition of variance:

   $$\mathrm{Var}(X) = E[X^2] - (E[X])^2$$

   For the random variable cX:

   $$\mathrm{Var}(cX) = E[(cX)^2] - (E[cX])^2$$

   Simplifying the expressions:

   $$E[(cX)^2] = E[c^2X^2] = c^2E[X^2]$$

   $$E[cX] = cE[X]$$

   Thus,

   $$(E[cX])^2 = c^2(E[X])^2$$

   Substituting these into the variance formula:

   $$\mathrm{Var}(cX) = c^2E[X^2] - c^2(E[X])^2$$

   Taking c^2 as common:

   $$\mathrm{Var}(cX) = c^2\left(E[X^2] - (E[X])^2\right)$$

   We know:

   $$E[X^2] - (E[X])^2 = \mathrm{Var}(X)$$

   So:

   $$\mathrm{Var}(cX) = c^2\,\mathrm{Var}(X)$$

   Hence Proved.

4. **Prove that with probability at least $1 - 1/k$, a uniformly random permutation $\sigma : [n] \to [n]$ has at most k fixed points.**

   Solution:

   Let σ be a random permutation of $[n]$, i.e., $\sigma : [n] \to [n]$.

   Define a fixed point i of σ as $\sigma(i) = i$.

   Let X_i be an indicator random variable for each element $i \in [n]$, where $X_i = 1$ if i is a fixed point and $X_i = 0$ otherwise.

   The probability that i is a fixed point is $\Pr(X_i = 1) = \frac{1}{n}$, since the permutation σ is chosen uniformly at random.

   Let $X = \sum_{i=1}^{n} X_i$ be the total number of fixed points in the permutation.

   The expected number of fixed points is:

$$E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} = 1$$

   By Markov's inequality, the probability that X exceeds k is:

$$\Pr(X \geq k) \leq \frac{E[X]}{k} = \frac{1}{k}$$

   Therefore, the probability that X is less than or equal to k is:

$$\Pr(X \leq k) \geq 1 - \frac{1}{k}$$

   Thus, with probability at least $1 - \frac{1}{k}$, a uniformly random permutation σ has at most k fixed points.

5. **Consider a particle that does an unbiased random walk on real line. It starts at position 0. For any i, if the particle is at i, it moves to position $i + 1$ with probability $1/2$ and to position $i - 1$ with probability $1/2$. Prove that after n steps, with at least $1 - 10/\sqrt{n}$ probability, the distance of the particle from start, i.e., 0 is at most $\sqrt{n}$ ln n. Prove that after n steps, with at least $1-10/\sqrt{n}$ probability, the distance of a particle from the start is at most $\sqrt{n \ln n}$.**

   Solution:

   Let X_n denote the position of the particle after n steps. The particle performs a random walk. The distance from the starting point after n steps is given by $|X_n|$.

   By the Central Limit Theorem (CLT), for large n, the random walk can be approximated by a normal distribution:

$$X_n \sim \mathcal{N}(0, n)$$

   where 0 is the mean and n is the variance of the distribution.

   Bounding the Probability of Large Deviations:

   The probability that the particle moves from the origin by more than $\sqrt{n \ln n}$ is:

$$P(|X_n| \geq \sqrt{n \ln n})$$

   Since $X_n \sim \mathcal{N}(0, n)$, we can apply concentration inequalities for normal distributions:

$$P(|X_n| \geq t) \leq 2 \exp\left(-\frac{t^2}{2n}\right)$$

where t is the deviation.
Substituting $t = \sqrt{n \ln n}$:

$$P\left(|X_n| \geq \sqrt{n \ln n}\right) \leq 2 \exp\left(-\frac{n \ln n}{2n}\right) = 2 \exp\left(-\frac{\ln n}{2}\right)$$

This simplifies to:

$$P\left(|X_n| \geq \sqrt{n \ln n}\right) \leq \frac{2}{\sqrt{n}}$$

Final Probability Bound:
The probability that the distance from the start is at most $\sqrt{n \ln n}$ is:

$$P\left(|X_n| \leq \sqrt{n \ln n}\right) = 1 - P\left(|X_n| \geq \sqrt{n \ln n}\right)$$

Substituting the previous result:

$$P\left(|X_n| \leq \sqrt{n \ln n}\right) \geq 1 - \frac{2}{\sqrt{n}}$$

For sufficiently large n:

$$1 - \frac{2}{\sqrt{n}} \geq 1 - \frac{10}{\sqrt{n}}$$

Thus, after n steps, with at least $1 - \frac{10}{\sqrt{n}}$ probability, the distance of the particle from the starting point is at most $\sqrt{n \ln n}$.

Solve/Prove the following statements (2 mark each):

1. **If X and Y are independent random variables then $E[XY] = E[X]E[Y]$.**
   Solution:
   By definition of expectation, the joint expectation of X and Y is:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x, y)\, dx\, dy$$

   where $f_{X,Y}(x, y)$ is the joint probability density function of X and Y.
   Since X and Y are independent, the joint probability density function:

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

   Thus, the joint expectation becomes:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y)\, dx\, dy$$

   Separating the integrals:

$$E[XY] = \left(\int_{-\infty}^{\infty} x f_X(x)\, dx\right) \left(\int_{-\infty}^{\infty} y f_Y(y)\, dy\right)$$

   The first integral is $E[X]$ and the second is $E[Y]$, therefore:

$$E[XY] = E[X]E[Y]$$

2. $E[X] = \Pr(A)E[X|A] + \Pr(A^c)E[X|A^c]$ **where** X **is a random variable and** A **is an event.**

   Solution:

   The law of total expectation is:

$$E[X] = E[X|A]P(A) + E[X|A^c]P(A^c)$$

   From the definition of conditional expectation:

$$E[X|A] = \frac{\sum_{x \in A} x \Pr(X = x|A)}{P(A)}$$

   Similarly for $E[X|A^c]$:

$$E[X|A^c] = \frac{\sum_{x \in A^c} x \Pr(X = x \mid A^c)}{P(A^c)}$$

   Thus:

$$E[X] = \Pr(A)E[X|A] + \Pr(A^c)E[X|A^c]$$

3. **If** $X_1, \ldots, X_n$ **are independent random variable then** $\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{Var}(X_i)$**.**

   Solution:

   By the definition of variance:

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = E\left[\left(\sum_{i=1}^n X_i\right)^2\right] - \left(E\left[\sum_{i=1}^n X_i\right]\right)^2$$

   Expanding the square:

$$E\left[\left(\sum_{i=1}^n X_i\right)^2\right] = E\left[\sum_{i=1}^n X_i^2 + 2\sum_{1 \le i < j \le n} X_i X_j\right]$$

   Using the linearity of expectation and the independence of $X_1, X_2, \ldots, X_n$, we get:

$$E\left[\sum_{i=1}^n X_i^2 + 2\sum_{1 \le i < j \le n} X_i X_j\right] = \sum_{i=1}^n E[X_i^2] + 2\sum_{1 \le i < j \le n} E[X_i]E[X_j]$$

   Since the X_i's are independent and the cross terms $E[X_iX_j]$ for $i \neq j$ are equal to $E[X_i]E[X_j]$.

   We have:

$$E\left[\left(\sum_{i=1}^n X_i\right)^2\right] = \sum_{i=1}^n E[X_i^2] + 2\sum_{1 \le i < j \le n} E[X_i]E[X_j]$$

   Calculating $\left(E\left[\sum_{i=1}^n X_i\right]\right)^2$:

$$\left(E\left[\sum_{i=1}^n X_i\right]\right)^2 = \left(\sum_{i=1}^n E[X_i]\right)^2$$

   Finally, the variance is:

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{Var}(X_i)$$

4. **Prove Markov's inequality, i.e., if X is any random variable that takes only non negative values then $\Pr(X \geq cE[X]) \leq \frac{1}{c}$.**

   Solution:

   Define the indicator random variable $I = 1_{\{X \geq cE[X]\}}$, which is 1 if $X \geq cE[X]$ and 0 otherwise.

   $$X \geq cE[X] \implies I = 1$$

   Using the expectation of the indicator variable:

   $$E[I] = \Pr(X \geq cE[X])$$

   Applying the following inequality:

   $$X \geq cE[X] \implies X \geq c \cdot 1_{\{X \geq cE[X]\}} \cdot E[X]$$

   This follows because when $X \geq cE[X]$, we have:

   $$X \geq c \cdot E[X]$$

   Therefore, the inequality:

   $$E[X] \geq E[X 1_{\{X \geq cE[X]\}}]$$

   We will apply the property of expectation because $1_{\{X \geq cE[X]\}}$ is always between 0 and 1:

   $$E[X 1_{\{X \geq cE[X]\}}] \leq E[X]$$

   Finally:

   $$\Pr(X \geq cE[X]) = E[1_{\{X \geq cE[X]\}}] \leq \frac{1}{c}$$

   Thus, it is proved Markov's inequality:

   $$\Pr(X \geq cE[X]) \leq \frac{1}{c}$$

5. **Prove Chebyshev's inequality, i.e., $\Pr(|X - E[X]| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}$ (Hint: use Markov inequality).**

   Solution:

   Let $Y = |X - E[X]|$.

   Define a new random variable $Z = (X - E[X])^2$.

   Clearly, $Z \geq 0$ because Z is a squared term.

   Applying Markov's inequality to Z:

   $$\Pr\left(Z \geq t^2\right) \leq \frac{E[Z]}{t^2}$$

   Since $Z = (X - E[X])^2$:

   $$E[Z] = E[(X - E[X])^2] = \mathrm{Var}(X)$$

   Thus, the probability will become:

   $$\Pr\left(|X - E[X]| \geq t\right) = \Pr\left((X - E[X])^2 \geq t^2\right) \leq \frac{\mathrm{Var}(X)}{t^2}$$

   Therefore, Chebyshev's inequality is:

   $$\Pr\left(|X - E[X]| \geq t\right) \leq \frac{\mathrm{Var}(X)}{t^2}$$

6. **Show that, for more than 2 events, pairwise independence does not imply independence.**

   Solution:

   Let the probability space be $\Omega = \{1, 2, 3, 4\}$, and define the probability distribution as follows:
   $$\Pr(\{1\}) = \Pr(\{2\}) = \Pr(\{3\}) = \Pr(\{4\}) = \frac{1}{4}$$

   Defining the events:

   - $A = \{1, 2\}$
   - $B = \{1, 3\}$
   - $C = \{1, 4\}$

   Checking pairwise independence:

   - $\Pr(A \cap B) = \Pr(\{1\}) = \frac{1}{4}$ and $\Pr(A)\Pr(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
     So A and B are independent.
   - $\Pr(A \cap C) = \Pr(\{1\}) = \frac{1}{4}$ and $\Pr(A)\Pr(C) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
     So A and C are independent.
   - $\Pr(B \cap C) = \Pr(\{1\}) = \frac{1}{4}$ and $\Pr(B)\Pr(C) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
     So B and C are independent.

   Thus, A, B, and C are pairwise independent.

   Now, checking if A, B, and C are mutually independent:

   - $\Pr(A \cap B \cap C) = \Pr(\{1\}) = \frac{1}{4}$
   - $\Pr(A)\Pr(B)\Pr(C) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

   Since $\Pr(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = \Pr(A)\Pr(B)\Pr(C)$

   A, B, and C are not mutually independent.

   Thus, it is shown that pairwise independence does not imply mutual independence for more than 2 events.

7. **We have a standard 6-sided dice. Let X be the number of times 6 appears over n throes of the dice. Let p be the probability of the event $\Pr(X \geq n/4)$. Compare the upper bounds on p that you can obtain using Markov's inequality, Chebyshev's inequality, and Hoffding's bounds.**

   Solution:

   1. Using Markov's Inequality

   Markov's inequality states that for a non-negative random variable Y and any $a > 0$, we have:
   $$P(Y \geq a) \leq \frac{E[Y]}{a}$$

   In question, X is the number of times 6 appears, so X is the sum of n independent Bernoulli random variables, each with success probability $\frac{1}{6}$.

   Thus, the expectation of X is:
   $$E[X] = n \times \frac{1}{6}$$

Now, applying Markov's inequality to get an upper bound on $p = P(X \geq n/4)$:

$$P(X \geq n/4) \leq \frac{E[X]}{n/4} = \frac{n/6}{n/4} = \frac{4}{6} = \frac{2}{3}$$

Thus, the upper bound using Markov's inequality is:

$$P(X \geq n/4) \leq \frac{2}{3}$$

2. Using Chebyshev's Inequality

Chebyshev's inequality states that for any random variable X with mean μ and variance σ^2, and for any $t > 0$, we have:

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$$

In question, X is the number of times 6 appears, so:

$$E[X] = n \times \frac{1}{6}$$

$$\text{Var}(X) = n \times \frac{1}{6} \times \frac{5}{6} = \frac{5n}{36}$$

Now, we want to bound $P(X \geq n/4)$. This can be written as:

$$P(X \geq n/4) = P(X - E[X] \geq n/4 - E[X])$$

Using Chebyshev's inequality, we get:

$$P(X \geq n/4) \leq \frac{\text{Var}(X)}{(n/4 - E[X])^2}$$

Substituting the values:

$$P(X \geq n/4) \leq \frac{5n/36}{(n/4 - n/6)^2}$$

Simplifying the expression:

$$P(X \geq n/4) \leq \frac{5n/36}{(n/12)^2} = \frac{5n/36}{n^2/144} = \frac{20}{n}$$

Thus, the upper bound using Chebyshev's inequality is:

$$P(X \geq n/4) \leq \frac{20}{n}$$

3. Using Hoeffding's Bounds

Hoeffding's inequality provides a bound on the probability that the sum of independent random variables deviates from its expectation. For independent random variables $X_1, X_2, \ldots, X_n$, each bounded between 0 and 1, Hoeffding's inequality states:

$$P\left(\sum_{i=1}^{n} X_i - E\left[\sum_{i=1}^{n} X_i\right] \geq t\right) \leq \exp\left(-\frac{2t^2}{n}\right)$$

In question, X is the number of 6's, so it can be treted as the sum of n independent Bernoulli random variables. Each random variable has a success probability of $\frac{1}{6}$, so:

$$E[X] = n \times \frac{1}{6}$$

The deviation t is as follows:

$$t = \frac{n}{4} - E[X] = \frac{n}{4} - \frac{n}{6} = \frac{n}{12}$$

Now, applying Hoeffding's inequality:

$$P(X \geq n/4) \leq \exp\left(-\frac{2(n/12)^2}{n}\right) = \exp\left(-\frac{2n^2}{144n}\right) = \exp\left(-\frac{n}{72}\right)$$

Thus, the upper bound using Hoeffding's inequality is:

$$P(X \geq n/4) \leq \exp\left(-\frac{n}{72}\right)$$

Conclusion:

The upper bounds on $p = P(X \geq n/4)$ are:

- Using Markov's inequality: $P(X \geq n/4) \leq \frac{2}{3}$
- Using Chebyshev's inequality: $P(X \geq n/4) \leq \frac{20}{n}$
- Using Hoeffding's inequality: $P(X \geq n/4) \leq \exp\left(-\frac{n}{72}\right)$

3-marks each

1. **Let μ be some fixed quantity and $\varepsilon > 0$ be any fixed small number. Consider a randomized algorithm R which has the following guarantee: the value returned by the algorithm say $\hat{\mu}$ satisfies $\mu - \varepsilon < \hat{\mu} < \mu + \varepsilon$ with probability at least $2/3$. We wish an algorithm that achieves the same guarantee (that is value returned should be in $[\mu - \varepsilon, \mu + \varepsilon]$) but with very high probability say at least $1 - \delta$ for a small $\delta > 0$. Show that the following algorithm achieves the same. run the algorithm R independently $t = 100000\ log 1/\delta$ times. Return the median of these t outputs.**

   Note: This is called median trick. If we can estimate a quantity with probability c where $c > 1/2$ is any constant using at most s samples then we can also estimate that quantity with confidence $1-\delta$ for any $\delta > 0$, with only logarithmic blow up in the number of samples (i.e., using at most $O(slog1/\delta)$ samples).

   Solution:

   We are asked to apply the *median trick*, which is a method that allows us to amplify the probability of an event happening by taking the median of several independent trials.

   The given guarantee of randomized algorithm R is:

   $$P(\mu - \epsilon < \hat{\mu} < \mu + \epsilon) \geq \frac{2}{3}$$

   This means that for each independent execution of R, the probability that the output $\hat{\mu}$ lies within the interval $[\mu - \epsilon, \mu + \epsilon]$ is at least $\frac{2}{3}$ and the probability that the output lies outside this interval is at most $\frac{1}{3}$.

Running the algorithm t times:

We will run the algorithm R independently $t = 100000 \cdot \log \frac{1}{\delta}$ times.

Let the results of these t independent runs as $\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_t$.

Let X_i be an indicator variable for each trial, where $X_i = 1$ if $\hat{\mu}_i$ is within the interval $[\mu - \epsilon, \mu + \epsilon]$, and $X_i = 0$ otherwise.

We already know that:

$$P(X_i = 1) \geq \frac{2}{3}, \quad P(X_i = 0) \leq \frac{1}{3}$$

Applying Median Tric:k

We will compute the median of the t independent outputs of the randomized algorithm R. The median output will be the value that occurs in at least half of the t trials.

Let the median of these trials as $\hat{\mu}_{\text{median}}$.

The key observation here is that the median will lie within the desired interval $[\mu - \epsilon, \mu + \epsilon]$ with high probability if enough trials return values within the interval.

Specifically, since the probability of each trial returning a value within the interval is at least $\frac{2}{3}$, the number of trials where $\hat{\mu}_i$ is in $[\mu - \epsilon, \mu + \epsilon]$ follows a binomial distribution.

Let Y be the number of trials where $\hat{\mu}_i$ is within the interval $[\mu - \epsilon, \mu + \epsilon]$, i.e.,

$$Y = \sum_{i=1}^{t} X_i$$

The expected value of Y is:

$$E[Y] = t \cdot \frac{2}{3}$$

The median of these outputs will lie within the interval $[\mu - \epsilon, \mu + \epsilon]$ if at least half of the trials return values in the interval.

To ensure this, we need $Y \geq \frac{t}{2}$, which is equivalent to:

$$P(Y \geq \frac{t}{2})$$

Using concentration inequalities (such as Chernoff bounds), we can show that this probability is at least $1 - \delta$ for large t.

Specifically, for $t = O(\log \frac{1}{\delta})$, the probability that the median lies within the interval $[\mu - \epsilon, \mu + \epsilon]$ is at least $1 - \delta$.

Conclusion:

By running the algorithm R, $t = 100000 \cdot \log \frac{1}{\delta}$ times and taking the median of these outputs, we can say that the final output $\hat{\mu}_{\text{median}}$ satisfies:

$$P(\mu - \epsilon < \hat{\mu}_{\text{median}} < \mu + \epsilon) \geq 1 - \delta$$

This guarantees that with high probability, the value returned by the algorithm will lie within the desired interval $[\mu - \epsilon, \mu + \epsilon]$.

2. **We show that the expected reward $E[R(T)]$ of the algorithm ETC satisfies $E[R(T)] = O(T^{2/3}K^{1/3}(logT)^{1/3})$. Note that we have assumed that the support of all reward distributions is in $[0, 1]$, i.e, all rewards are in $[0, 1]$. Suppose all rewards are in $[0, x]$ instead of $[0, 1]$. Now determine $E[R(T)]$ (now x should also come in the final expression of expected regret).**
(Hint: Carefully go through the proof of $E[R(T)] = O(T^{2/3}K^{1/3}(logT)^{1/3})$. Find exact places where it was used that rewards are in $[0, 1]$ and modify accordingly.)

Solution:

We are given that the expected reward of the ETC (Explore-Then-Commit) algorithm in the case where all rewards are in the range $[0, 1]$ satisfies the following bound:

$$E[R(T)] = O\left(T^{\frac{2}{3}}K^{\frac{1}{3}}(\log T)^{\frac{1}{3}}\right),$$

where:

- T is the total number of rounds,
- K is the number of arms in the bandit problem,
- The reward distributions are assumed to be supported in $[0, 1]$.

Scaling the reward range from $[0, 1]$ to $[0, x]$:

Suppose that instead of being in the range $[0, 1]$, the rewards are in the range $[0, x]$, where $x > 0$. We will apply a reward range transformation that maps the reward r in the range $[0, x]$ to the corresponding reward r' in the range $[0, 1]$. The transformation is given by:

$$r' = \frac{r}{x}, \quad \text{so that } r' \in [0, 1].$$

This transformation scales all rewards by a factor of x.

Effect on Expected Reward:

Since the reward range is transformed from $[0, x]$ to $[0, 1]$, all rewards are scaled by the factor x. As the expected reward is based on the distribution of rewards, scaling all rewards by x results in the expected reward being scaled by the same factor.

Thus, the expected reward after transforming the rewards is given by:

$$E[R(T)] = x \cdot E[R'(T)]$$

where $E[R'(T)]$ represents the expected reward computed for the normalized rewards in the range $[0, 1]$.

Substituting the known boundExpected Reward in the New Setting:

Using the original bound for the expected reward with rewards in the range $[0, 1]$, we can substitute the following:

$$E[R'(T)] = O\left(T^{\frac{2}{3}}K^{\frac{1}{3}}(\log T)^{\frac{1}{3}}\right)$$

Thus, the expected reward with rewards in the range $[0, x]$ is:

$$E[R(T)] = x \cdot O\left(T^{\frac{2}{3}}K^{\frac{1}{3}}(\log T)^{\frac{1}{3}}\right)$$

Final Expected Reward Expression:

Therefore, the expected reward for the ETC algorithm when the reward distributions are in the range $[0, x]$ is:

$$E[R(T)] = O\left(xT^{\frac{2}{3}}K^{\frac{1}{3}}(\log T)^{\frac{1}{3}}\right).$$

3. **Fix some large T (number of rounds). Give the best upper bound (that you can) on the expected regret of the ETC algorithm after T rounds on the following bandit instances consisting of only two arms:**

   (a) **both arm have mean of** $1/2$**.**

   (b) **one arm has mean of** $1/2$ **and another arm has mean of** $1/2 + 1000\frac{(logT)^{1/3}}{T^{1/3}}$.

   (c) **one arm has mean of** $1/2$ **and another arm has mean of** $1/2 + \frac{1}{\sqrt{T}}$

   Solution:

   General Expected Regret Bound:

   The expected regret $E[R(T)]$ for the ETC (Explore-Then-Commit) algorithm after T rounds with K arms is given by:

$$E[R(T)] = O\left(\sum_{i\neq\hat{i}} \frac{\Delta_i\sqrt{T\log T}}{n_i}\right)$$

   where:

   - $\hat{i}$ is the arm chosen most frequently,
   - Δ_i is the difference between the expected reward of arm i and the optimal arm (which has the maximum expected reward),
   - n_i is the number of times arm i is selected during the exploration phase.

   Case (a): Both arms have mean 1/2:

   Both arms have the same expected reward, so there is no difference between the arms. So:

$$\Delta = 0$$

   Since there is no difference in expected rewards, the expected regret is simply:

$$E[R(T)] = O(0) = 0$$

   Thus, the upper bound for the expected regret is:

$$E[R(T)] = 0$$

   Case (b): One arm has mean 1/2, the other has mean $1/2 + 1000\frac{(logT)^{1/3}}{T^{1/3}}$**:**

   The difference between the expected rewards of the two arms is:

$$\Delta = \frac{1000(\log T)^{1/3}}{T^{1/3}}$$

   Using the general expected regret formula and substituting this value of Δ, we get:

$$E[R(T)] = O\left(\frac{\Delta\sqrt{T\log T}}{n}\right)$$

   Substitute $\Delta = \frac{1000(\log T)^{1/3}}{T^{1/3}}$:

$$E[R(T)] = O\left(\frac{1000(\log T)^{1/3}\sqrt{T\log T}}{T^{1/3}}\right)$$

Simplifying:

$$E[R(T)] = O\left(\frac{1000(\log T)^{2/3}}{T^{1/6}}\right)$$

Thus, the expected regret for this case is:

$$E[R(T)] = O\left(\frac{1000(\log T)^{2/3}}{T^{1/6}}\right)$$

Case (c): One arm has mean 1/2, the other has mean $1/2 + \frac{1}{\sqrt{T}}$:

The difference between the expected rewards of the two arms is:

$$\Delta = \frac{1}{T}$$

Using the general expected regret formula and substituting $\Delta = \frac{1}{T}$, we get:

$$E[R(T)] = O\left(\frac{\Delta\sqrt{T \log T}}{n}\right)$$

Substitute $\Delta = \frac{1}{T}$:

$$E[R(T)] = O\left(\frac{\sqrt{T \log T}}{T}\right)$$

Simplifying:

$$E[R(T)] = O\left(\frac{\sqrt{\log T}}{T^{1/2}}\right)$$

Thus, the expected regret for this case is:

$$E[R(T)] = O\left(\frac{\sqrt{\log T}}{T^{1/2}}\right)$$

Final Results:

Summarizing the upper bounds for expected regret in each case:

- (a) $E[R(T)] = 0$
- (b) $E[R(T)] = O\left(\frac{1000(\log T)^{2/3}}{T^{1/6}}\right)$
- (c) $E[R(T)] = O\left(\frac{\sqrt{\log T}}{T^{1/2}}\right)$