GML Fall 23 - HW 1: Concentration, uniform and margin bounds

1 Optional Warm-up: Optimality of polynomial Markov

Chernoff's bound is obtained via Markov's inequality. In this question we show that Markov's inequality is actually tight. Furthermore, the k-th moment Markov bounds are in fact never worse than the Chernoff bound based on the moment generating function. a) **Find a** non-negative random variable X for which Markov's inequality is met with equality at a point $a > 0$. b) Suppose that $X \geq 0$ and that $\mathbb{E}e^{\lambda X}$ exists in an interval around zero. Given some $\delta > 0$ and integer $k = 1, 2, \ldots$ **show that**

$$\inf_{k=0,1,\ldots} \frac{\mathbb{E}|X|^k}{\delta^k} \leq \inf_{\lambda>0} \frac{\mathbb{E}e^{\lambda X}}{e^{\lambda\delta}}$$

1.1 Solution

a) Consider a random variable X with the pdf $p_X(x) = \delta_a(x)$, i.e. $X = a$ with probability 1. Then $E[X] = a$, $P(X \geq a) = 1$ and thus we have

$$1 = \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a} = \frac{a}{a} = 1.$$

b) We suppose that for $\lambda \in (-\Delta, \Delta)$, the expectation $\mathbb{E}\left[e^{\lambda X}\right]$ exists. Given a $\lambda \in (-\Delta, \Delta)$, we write

$$\mathbb{E}\left[e^{\lambda X}\right] = \mathbb{E}\left[\sum_{k\geq 0} \frac{\lambda^k X^k}{k!}\right] = \sum_{k\geq 0} \frac{\lambda^k \mathbb{E}\left[X^k\right]}{k!},$$

where we have used the Fubini-Tonelli theorem in the case of nonnegative measurable functions. To relate $\mathbb{E}\left[X^k\right]$ to the LHS, we rewrite

$$\mathbb{E}\left[X^k\right] = \frac{\delta^k \mathbb{E}\left[X^k\right]}{\delta^k} \geq \delta^k \inf_{k'\geq 0} \frac{\mathbb{E}\left[X^{k'}\right]}{\delta^{k'}}.$$

Thus,

$$\mathbb{E}\left[e^{\lambda X}\right] = \frac{\lambda^k \mathbb{E}\left[X^k\right]}{k!} \geq \inf_{k'\geq 0} \frac{\mathbb{E}\left[X^{k'}\right]}{\delta^{k'}} \sum_{k\geq 0} \frac{\lambda^k \delta^k}{k!} = \inf_{k'\geq 0} \frac{\mathbb{E}\left[X^{k'}\right]}{\delta^{k'}} e^{\lambda\delta}.$$

Dividing by $e^{\lambda\delta}$ and taking the infimum over λ yields the inequality.

2 Concentration and kernel density estimation

Let $\{X_i\}_{i=1}^n$ be an i.i.d. sequence of random variables drawn from a density f on the real line. A standard estimate of f is the kernel density estimate:

$$f_n(x) := \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

where $K : \mathbb{R} \to [0, \infty)$ is a kernel function satisfying $\int_{-\infty}^{\infty} K(t)\,dt = 1$, and $h > 0$ is a bandwidth parameter. Suppose that we assess the quality of f_n using the L_1-norm:

$$\|f_n - f\|_1 := \int_{-\infty}^{\infty} |f_n(t) - f(t)|\,dt.$$

Prove that:

$$P\left[\|f_n - f\|_1 \geq \mathbb{E}[\|f_n - f\|_1] + \delta\right] \leq e^{-\frac{n\delta^2}{18}}.$$

2.1 Solution

We write the i.i.d. random variables $(X_1, ..., X_n)$ as a random vector and define the function

$$g(X_1, ..., X_n) = \|f - f_n(X_1, ..., X_n)\|_1.$$

We show that g satisfies the bounded differences property with $L = \frac{2}{n}$: For $x = (x_1, ..., x_n) \in \mathbb{R}^n$ and $k \in [n]$, define x^k by $x_i^k = x_i$ if $i \neq k$, and $x_k^k = y$, where $y \in \mathbb{R}$. We calculate

$$\begin{aligned}
|g(x) - g(x^k)| &= |\|f - f_n(x)\|_1 - \|f - f_n(x^k)\|_1| \leq \|f_n(x) - f_n(x^k)\|_1 \\
&= \frac{1}{nh} \int \left| K\left(\frac{t - x_k}{h}\right) - K\left(\frac{t - y}{h}\right) \right| dt \\
&\leq \frac{1}{n} \left(\int K(u - x_k/h) du + \int K(u' - y/h) du' \right) \leq \frac{2}{n}.
\end{aligned}$$

Thus, by the (one-sided) bounded differences inequality (Corollary 2.21 in MW), we obtain

$$\mathbb{P}\left(\|f - f_n\|_1 \geq \mathbb{E}\left[\|f - f_n\|_1\right] + \delta\right) \leq e^{\frac{-2\delta^2}{n\frac{4}{n^2}}} = e^{-\frac{n\delta^2}{2}}.$$

.... **10 points**

3 Sub-Gaussian maxima

In this exercise we prove an inequality used repeatedly in later lectures. Let $\{X_i\}_{i=1}^n$ be a sequence of zero-mean random variables, each subgaussian with parameter σ. The random variables X_i are *not* assumed to be independent. a) **Prove that** for all $n \geq 1$ we have

$$\mathbb{E} \max_{i=1,\dots,n} X_i \leq \sqrt{2\sigma^2 \log n}.$$

Hint: the exponential is a convex function. b) **Prove that** for all $n \geq 2$ we have

$$\mathbb{E} \max_{i=1,\dots,n} |X_i| \leq \sqrt{2\sigma^2 \log(2n)} \leq 2\sqrt{\sigma^2 \log n}.$$

3.1 Solution

We consider the moment generating function of $\max_{i \in [n]} X_i$. Since $\exp(\lambda \cdot)$ is a convex function, we utilize Jensen's inequality to obtain

$$\exp\left(\lambda \mathbb{E}\left[\max_{i \in [n]} X_i\right]\right) \leq \mathbb{E}\left[\exp\left(\lambda \max_{i \in [n]} X_i\right)\right].$$

We can interchange exp and max to obtain

$$\mathbb{E}\left[\exp\left(\lambda \max_{i\in[n]} X_i\right)\right] = \mathbb{E}\left[\max_{i\in[n]} e^{\lambda X_i}\right] \le \mathbb{E}\left[\sum_{i=1}^n e^{\lambda X_i}\right] = \sum_{i=1}^n \mathbb{E}\left[e^{\lambda X_i}\right] \le ne^{\lambda^2\sigma^2/2},$$

using the subgaussianity of the i.i.d. variables. In total, we have

$$\exp\left(\lambda \mathbb{E}\left[\max_{i\in[n]} X_i\right]\right) \le ne^{\lambda^2\sigma^2/2}.$$

Solving for $\mathbb{E}\left[\max_{i\in[n]} X_i\right]$, we get

$$\mathbb{E}\left[\max_{i\in[n]} X_i\right] \le \frac{1}{\lambda}(\log(n) + \lambda^2\sigma^2/2) = \frac{\log(n)}{\lambda} + \lambda\sigma^2/2.$$

This expression is minimized for $\lambda^* = \frac{\sqrt{2\log n}}{\sigma}$, where it achieves the value $\sqrt{2\sigma^2 \log n}$.

. . . . **5 points**

b) We have

$$\max_{i\in[n]} |X_i| = \max_{i\in[n]} \max\{-X_i, X_i\} = \max\{-X_1, X_1, ..., -X_n, X_n\},$$

which is a maximum over $2n$ subgaussian random variables. Thus, we have by (a)

$$\mathbb{E}\left[\max_{i\in[n]} |X_i|\right] \le \sqrt{2\sigma^2 \log 2n} = \sqrt{2\sigma^2(\log n + \log 2)} \le \sqrt{2\sigma^2 2\log n} = 2\sqrt{\sigma^2 \log n},$$

where we have used $n \ge 2$.

. . . . **5 points**

4 Bonus: Sharper tail bounds for bounded variables: Bennett's inequality

Read MW Chapter 2 and learn about subexponential tail bounds and Bernstein's inequality, yielding some more tail bounds for empirical means of random variables satisfying conditions other than the subgaussian one. Bernstein's inequality is sometimes tighter for bounded variables than when applying the subgaussian bound. In this problem we prove an even tighter bound for bounded variables, known as Bennett's inequality a) Consider a zero-mean random variable such that $|X_i| \le b$ for some $b > 0$. **Prove that**

$$\log \mathbb{E}\mathrm{e}^{\lambda X_i} \le \sigma_i^2 \lambda^2 \frac{\mathrm{e}^{\lambda b} - 1 - \lambda b}{(\lambda b)^2}$$

for all $\lambda \ge 0$, where $\sigma_i^2 = \mathrm{Var}(X_i)$. b) Given independent random variables $X_1, \ldots, X_n$ satisfying the condition of part (a), let $\sigma^2 := \frac{1}{n}\sum_{i=1}^n \sigma_i^2$ be the average variance. **Prove *Bennett's inequality***

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \ge \delta\right) \le \mathrm{e}^{-\frac{n\sigma^2}{b^2} h\left(\frac{b\delta}{\sigma^2}\right)}$$

where $h(t) := (1+t)\log(1+t) - t$ for $t \ge 0$. c) *Bonus*: **Show that** Bennett's inequality is at least as good as Bernstein's inequality.

4.1 Solution

a) First, note that the function $f(x) = \frac{e^x - 1 - x}{x^2}$ is positive and monotonically increasing over $x \geq 0$, which is easy to verify by expanding e^x . Therefore, $f(\lambda X_i)$ is bounded by $f(\lambda b)$, where we use that $\lambda \geq 0$. We can write:

$$\mathbb{E}e^{\lambda X_i} = \mathbb{E}\sum_{k=0}^{\infty}\frac{(\lambda X_i)^k}{k!} = 1 + \lambda \underbrace{\mathbb{E}X_i}_{=0(\text{zero mean})} + \mathbb{E}\left(\lambda^2 X_i^2 \underbrace{\frac{e^{\lambda X_i} - 1 - \lambda X_i}{(\lambda X_i)^2}}_{=f(\lambda X_i)}\right)$$
$$\leq 1 + \lambda^2\sigma_i^2 f(\lambda b)$$
$$\Rightarrow \log \mathbb{E}e^{\lambda X_i} \leq \lambda^2\sigma_i^2\frac{e^{\lambda b} - 1 - \lambda b}{(\lambda b)^2} \leq \frac{\sigma_i^2}{b^2}(e^{\lambda b} - 1 - \lambda b)$$

where the last line follows from the definition of f and uses the numerical inequality $\log(1 + x) \leq x$ for $x \geq 0$. .... **5 points**

b) By monotonicity of the exponential function and Markov's inequality we can bound $\mathbb{P}(\frac{1}{n}\sum_i X_i \geq \delta) \leq \frac{\mathbb{E}\exp(\sum_i \frac{\lambda}{n}X_i)}{\exp(\lambda\delta)}$ like for Chernoff's bound. Since this holds for any $\lambda \geq 0$, we ultimately choose the one to achieve the best (lowest) probability. By independence we have by setting $\lambda \leftarrow \frac{\lambda}{n}$

$$\mathbb{E}\prod_{i=1}^{n}\exp(\frac{\lambda}{n}X_i) = \prod_{i=1}^{n}\mathbb{E}\exp(\frac{\lambda}{n}X_i) \leq \exp\left(\frac{n\sigma^2}{b^2}(\exp(\frac{b\lambda}{n}) - 1 - b\lambda/n)\right).$$

Finally, substituting into Markov's inequality we obtain

$$\mathbb{P}\left(\frac{1}{n}\sum_i X_i \geq \delta\right) \leq \exp\left(\frac{n\sigma^2}{b^2}(\mathrm{e}^{\frac{b\lambda}{n}} - 1 - \frac{b\lambda}{n} - \frac{b^2\lambda\delta}{n\sigma^2})\right) \quad (\star)$$

In order to take the infimum over λ, we differentiate the term w.r.t λ and find that the derivative vanishes at $\lambda = \frac{n}{b}(\log(\frac{b\delta}{\sigma^2}) + 1)$. Plugging it into the right hand side of $(\star)$ concludes the proof. .... **5 points**

c) Denote $A := \frac{bt}{\sigma^2}$ and recall the inequalities for the concentration of a single random variable: *Bernstein's inequality*: $\mathbb{P}[X \geq t] \leq exp(-\frac{t^2}{2(\sigma^2+bt)})$
Bennett's inequality: $\mathbb{P}[X \geq t] \leq exp(-\frac{\sigma^2}{b^2}h(A))$
First, we show that for non-negative $A \geq 0$, $h(A) \geq \frac{A^2}{2(A+1)}$:

$$h(A) = (1 + A)\log(1 + A) - A \geq \frac{A^2}{2(A + 1)}$$
$$\Leftrightarrow g(A) := 2\log(1 + A) - \frac{2A}{A + 1} - \frac{A^2}{(A + 1)^2} \geq 0$$

Clearly, $g(0) = 0$. Hence, the claim follows when showing that $g'(A) \geq 0$ for any $A \geq 0$:

$$g'(A) = \frac{2}{1 + A} - \frac{2}{(1 + A)^2} - \frac{2A}{(1 + A)^3} \geq 0$$
$$\Leftrightarrow \frac{2(A + 1)^2 - 2(A + 1) - 2A}{(1 + A)^3} = \frac{2A^2}{(1 + A)^3} \geq 0$$

Rewriting the exponent of the RHS of the Bernstein's inequalit, one can show that it is an upper bound on the exponent in Bennett's inequality:
$-\frac{t^2}{2(\sigma^2+bt)} = -\frac{\sigma^2}{b^2}\frac{\frac{b^2t^2}{\sigma^2}}{2(\sigma^2+bt)} = -\frac{\sigma^2}{b^2}\frac{btA}{2(\sigma^2+bt)} = -\frac{\sigma^2}{b^2}\frac{A}{2(\frac{\sigma^2}{bt}+1)} = -\frac{\sigma^2}{b^2}\frac{A}{2(\frac{1}{A}+1)} = -\frac{\sigma^2}{b^2}\frac{A^2}{2+2A} \geq -\frac{\sigma^2}{b^2}h(A).$
.... **5 points**

5 Sharp upper bounds on binomial tails

Let $\{X_i\}_{i=1}^n$ be an i.i.d. sequence of Bernoulli variables with parameter $\alpha \in (0, \frac{1}{2}]$, and consider the binomial random variable $Z_n = \sum_{i=1}^n X_i$. The goal of this exercise is to prove, for any $\delta \in (0, \alpha)$, a sharp upper bound on the tail probability $P[Z_n \leq \delta n]$. a) **Show that**

$$P[Z_n \leq \delta n] \leq e^{-nD(\delta\|\alpha)},$$

where the quantity

$$D(\delta \parallel \alpha) := \delta \log \delta + (1-\delta)\log(1-\delta)$$

is the Kullback–Leibler divergence between the Bernoulli distributions with parameters δ and α, respectively.
b) **Show that** the bound from part (a) is strictly better than the Hoeffding bound for all $\delta \in (0, \alpha)$.

5.1 Solution

a) By Chernoff, we have

$$\mathbb{P}\left(Z_n \leq \delta n\right) \leq \frac{\mathbb{E}\left[e^{\lambda Z_n}\right]}{e^{\lambda \delta n}} = e^{-\lambda\delta n}(\alpha e^{\lambda} + (1-\alpha))^n,$$

where we have inserted the moment generating function of the binomial distribution. Taking the log of both sides and setting the derivative of the RHS w.r.t. λ to zero, we obtain

$$-\delta + \frac{\alpha e^{\lambda}}{\alpha e^{\lambda} + 1 - \alpha} = 0,$$

which yields

$$\lambda^* = \log\frac{\delta(1-\alpha)}{\alpha(1-\delta)} = \log\left(\frac{1-\alpha}{1-\delta}\right) - \log\left(\frac{\alpha}{\delta}\right).$$

Inserting λ^* in the logarithm of the RHS, we obtain

$$\log\mathbb{P}\left(Z_n \leq \delta n\right) \leq -n[\lambda^*\delta - \log(\alpha e^{\lambda^*} + (1-\alpha))] = -n[(\delta-1)\log\left(\frac{1-\alpha}{1-\delta}\right) - \delta\log\left(\frac{\alpha}{\delta}\right)] = -nD(\delta \parallel \alpha).$$

.... **5 points**

b) Any bounded random variable ($X \in [a, b]$) is sub-Gaussian with parameter at most $\frac{(b-a)}{2}$. Thus, X_i are sub-Gaussian with parameter 1/2. By Hoeffding, we have

$$\mathbb{P}\left(Z_n \leq \delta n\right) = \mathbb{P}\left((Z_n - \alpha n) \leq (\delta n - \alpha n)\right) \leq \exp\left(-n(\delta-\alpha)^2\right).$$

It remains to compare $D(\delta \parallel \alpha)$ and $(\delta-\alpha)^2$ for $\delta \in (0, \alpha)$. At $\delta = \alpha$, both functions are zero and their first derivatives are zero. The second derivative of $(\delta-\alpha)^2$ at $\delta = \alpha$ is 2, whereas the second derivative of $D(\delta \parallel \alpha)$ at $\delta = \alpha$ is $\frac{1}{\alpha(1-\alpha)}$, which is larger than 4 for $\alpha \in (0, 1/2)$. This yields the claim.

.... **5 points**

6 Robust estimation of the mean

Suppose we want to estimate the mean μ of a random variable X from a sample $X_1, \cdots, X_n$, drawn independently from the distribution of X. We want an ϵ-accurate estimate of the mean, i.e., one that falls with probability $\geq 1-\delta$ in the interval $[\mu-\epsilon, \mu+\epsilon]$. **Show that** a sample size of $N = O\left(\log(\delta^{-1})\frac{\sigma^2}{\epsilon^2}\right)$ suffices to compute an ϵ-accurate estimate of the mean with probability at least $1-\delta$. *Hint: Compute the median of* $\log(\delta^{-1})$ *weak estimates.*

6.1 Solution

We divide the proof into two steps, where we first construct weak learners which are with probability at least $p > \frac{1}{2}$ an ϵ-accurate estimate of the mean (for simplicity, we can simply choose $p = 3/4$). In a second step, we then show that the median of the weak learners is with proability at least $1-\delta$ an ϵ-accurate estimate of the mean.

Step 1: We begin with the construction of K weak learners μ_i. For this, we divide the dataset into K-parts equally large of size N_K and compute the mean μ_i for each of this subset. By Chebyshev's inequality, we get that

$$\mathbb{P}(|\mu - \mu_i| > \epsilon) \leq \frac{\sigma^2}{N_K \epsilon^2}.$$

In particular, when choosing $N_K \geq \frac{4\sigma^2}{\epsilon^2}$, we have that with probability at least $3/4$, μ_i is an ϵ-accurate estimate of the mean.

Step 2: Let $\tilde{\mu}$ be the median of the K estimates μ_i, which are by construction all independent. Furthermore, define the variables $\phi_i = 1\left[\mu_i \in [-\epsilon+\mu, \epsilon+\mu]\right]$ and $S = \sum b_i$. We can upper bound the probability that $\tilde{\mu}$ is not an ϵ-accurate estimate of the mean by:

$$\mathbb{P}(|\tilde{\mu} - \mu| > \epsilon) \leq \mathbb{P}(\sum b_i < \frac{K}{2}) = \mathbb{P}(\sum b_i - p < \frac{K}{2} - Kp).$$

We can now apply Hoeffdings inequality, which gives us

$$\mathbb{P}(\sum b_i - p < \frac{K}{2} - Kp) \leq \exp\left(-\frac{2(\frac{K}{2} - pK)^2}{\sum_{i=1}^{K}(1-0)^2}\right) = \exp\left(-\frac{1}{2}K(\frac{1}{2} - p)^2\right) = \delta$$

where we choose $p = \frac{1}{4}$ and $K = \lceil 32\log(\delta^{-1})\rceil$. Hence we can conclude the proof.

. . . . **10 points**

7 Best-arm identification

We now look at an interesting application of concentration bounds. Assume that we have K newly developed drugs to cure a disease and denote with $\mu_k \in [0,1]$ the probability of getting cured by the k-th drug, which is assumed to be unknown. In order to determine the best drug $k^\star$ with the highest chance of a successful treatment $\mu^\star = \mu_{k^\star} = \max_k \mu_k$, we treat different volunteers in a clinical trial with one drug each and record the outcome. We model the observation of the outcome on one patient as sampling from a Bernoulli distribution with parameter μ_k. We denote with $X_{k,i} \in \{0,1\}$ the random variable indicating whether the i-th volunteer treated with the k-th drug was successful.

In a randomized control trial, all drugs would have the same probability of getting assigned to any patient throughout the trial. In this exercise, we want to study an adaptive algorithm that assigns treatment depending on the outcome of previous treatments. The goal is to assign the drugs in a way such that for some $\delta \in (0,1)$, with probability $\geq 1-\delta$, the algorithm finds the best drug $k^\star$ in *as few volunteers as possible.*

This is ethically more reasonable than assigning a "bad" drug to patients even when their results are clearly inferior to others in the trial.

Context: This problem is often referred to as a best-arm identification problem. In adaptive or online learning scenarios, where at each time step we sample from one of k distributions $\{\mathbb{P}_1, \cdots, \mathbb{P}_K\}$ is often called a multi-armed bandit. Pulling an arm k then corresponds to sampling from $\mathbb{P}_k$. In our case they are Bernoulli distributions with means $\{\mu_1, \cdots, \mu_K\}$.

In this exercise, we analyze a specific type of algorithm to solve the problem called the Successive Elimination algorithm.

Algorithm 1: Successive Elimination

```
S_0 = {1, ..., K} ;
for 1 ≤ t ≤ ∞ do
    Pull all arms in S_{t-1} to obtain samples X_{k,t} ~ D_k with k ∈ S_{t-1};
    Update S_t = S_{t-1} - {i ∈ S_{t-1} : ∃k ∈ S_{t-1} : μ̂_{k,t} - U(t, δ/K) > μ̂_{i,t} + U(t, δ/K)};
    Stop when |S_t| = 1;
end
```

Notation:

- S_t: The active set of arms.
- $\hat{\mu}_{k,t} := \frac{1}{t}\sum_{i=1}^{t} X_{k,i}$: Estimated mean of the reward μ_k for arm k after t pulls.
- $U(t, \delta)$: An **any-time confidence interval**, such that for any arm k,

$$\mathbb{P}\left(\bigcup_{t=1}^{\infty}\{|\hat{\mu}_{k,t} - \mu_k| \geq U(t, \delta)\}\right) \leq \delta.$$

The goal of this exercise is to prove Theorem 1 where we show that the Successive Elimination algorithm is correct and derive an upper bound on the maximum amount of steps needed to for the algorithm to terminate.

Theorem 1. *With probability* $\geq 1 - \delta$*:*

1. *For any* $t \geq 1$*, the best arm* $k^\star$ *is contained in the set* S_t*.*
2. *There exists an any-time confidence interval* U *such that the Successive Elimination algorithm terminates after* $O(\sum_{k \neq k^\star}^{K} \triangle_k^{-2} \log(K\triangle_k^{-1}))$ *samples with* $\triangle_k := \mu^\star - \mu_k$ *and the* O *notation is with respect to* K *and* $\triangle_k$ *for a constant* δ*.*

We first prove that with high probability the best arm stays in the active set S_t for all t until termination.

a) Define $\mathcal{E}$ as the event that for any $t \geq 1$, the estimated reward $\hat{\mu}_{k,t}$ of any arm k is not contained in the confidence interval $U(t, \delta/K)$ around the true mean μ_k, i.e.

$$\mathcal{E} := \bigcup_{k=1}^{K}\bigcup_{t=1}^{\infty}\{|\hat{\mu}_{k,t} - \mu_k| > U(t, \delta/K)\}.$$

Show that $\mathbb{P}(\mathcal{E}) \leq \delta$.

b) **Prove** statement 1 in Theorem 1.

It is not yet shown whether and after how many steps the algorithm terminates. To do so, we derive a sufficiently tight any-time confidence interval U based on the concentration inequalities discussed in the lecture.

c) Let $\{Z_t\}_{t=1}^{\infty}$ be i.i.d bounded random variables with $Z_t \in [a,b]$ with $a \leq b$. **Show that**

$$U = \sqrt{\frac{(b-a)^2 \log(4t^2/\delta)}{2t}}$$

is a valid any-time confidence interval for the random variable Z_t. *Hint*: Use Hoeffding's bound and union bound.

d) *Bonus*: **Prove** statement 2 in Theorem 1.

7.1 Solution

a) First, by the Union bound,

$$\mathbb{P}(\mathcal{E}) \leq \sum_{k=1}^{K} \mathbb{P}(\bigcup_{t=1}^{\infty}\{|\hat{\mu}_{k,t} - \mu_k| > U(t,\delta/K)\}).$$

Next, we already know from the definition of the any-time confidence interval that

$$\mathbb{P}\left(\bigcup_{t=1}^{\infty}\{|\hat{\mu}_{k,t} - \mu_k| > U(t,\delta/K)\}\right) \leq \delta/K.$$

Hence, combining these results, we get $\mathbb{P}(\mathcal{E}) \leq \sum_{k=1}^{K} \delta/K = \delta$. .... **5 points**

b) We assume that $\mathcal{E}^{\mathsf{c}}$ holds and show that the best arm $k^\star$ will never be dropped. The proof then the follows trivially from a). Any arm k, with $1 \leq k \leq K$, will only be dropped by the algorithm if there exists $t \geq 1$ and $i \in S_{t-1}$ such that $k \in S_{t-1}$ and

$$\hat{\mu}_{i,t} - U(t,\delta/K) > \hat{\mu}_{k,t} + U(t,\delta/K).$$

Now consider k^*. $\mathcal{E}^{\mathsf{c}}$ holds by assumption and for any $t \geq 1$ we have:

$$\hat{\mu}_{k^\star,t} \geq \mu_{k^\star} - U(t,\delta/K)$$

Furthermore, for any $t \geq 1$ and $1 \leq i \leq K$ that $\mu_i \geq \hat{\mu}_{i,t} - U(t,\delta/K)$. By definition $\mu_{k^\star} \geq \mu_i$, and we get:

$$\hat{\mu}_{k^\star,t} + U(t,\delta/K) > \mu_{k^\star} \geq \mu_i > \hat{\mu}_{i,t} - U(t,\delta/K)$$

Meaning that $k^\star$ is never dropped, hence completing the proof. .... **5 points**

c) First, we take the Union bound to obtain

$$\mathbb{P}\left(\bigcup_{t=1}^{\infty}\{\left|\hat{\mu}_{k,t} - \mu_k\right| > U(t,\delta)\}\right) \leq \sum_{t=1}^{\infty} \mathbb{P}\left(\left|\hat{\mu}_{k,t} - \mu_k\right| \geq U(t,\delta)\right) \quad \leq \sum_{t=1}^{\infty} \mathbb{P}\left(\frac{1}{t}\left|\sum_{s=1}^{t} Z_s - \mathbb{E}Z_s\right| \geq U(t,\delta)\right)$$

Next, note that the case where $a = b$ follows trivially. Hence we can assume $a < b$ and observe that the random variable Z_i is a σ-subgaussian random variable with parameter $\sigma = \frac{b-a}{2}$. Therefore, we can apply Hoeffindgs inequality:

$$\mathbb{P}\left(\frac{1}{t}\left|\sum_{s=1}^{t} Z_s - \mathbb{E}Z_s\right| \geq U(t,\delta)\right) \leq 2\exp\left(-\frac{tU(t,\delta)}{2\sigma^2}\right) = 2\exp\left(-\frac{t(b-a)^2 \log(4t^2/\delta)}{2t\frac{b-a}{4}}\right)$$

$$= 2\exp(-\log(4t^2/\delta)) = 2\frac{\delta}{4t^2}$$

where the factor 2 in front of the exponential comes from the fact that we take a two sided bound, i.e.

$$\mathbb{P}\left(\frac{1}{t}\left|\sum_{s=1}^{t} Z_s - \mathbb{E}Z_s\right| \geq U(t,\delta)\right) = \mathbb{P}\left(\frac{1}{t}\sum_{s=1}^{t} Z_s - \mathbb{E}Z_s \geq U(t,\delta)\right) + \mathbb{P}\left(\frac{1}{t}\sum_{s=1}^{t} Z_s - \mathbb{E}Z_s \geq -U(t,\delta)\right).$$

Plugging this equation into the previous equation gives the desired solution:

$$\mathbb{P}\left(\bigcup_{t=1}^{\infty}\{\left|\hat{\mu}_{k,t} - \mu_k\right| > U(t,\delta)\}\right) \leq \sum_{t=1}^{\infty} \frac{\delta}{2t^2} \leq \delta$$

.... **5 points**

d) We can again assume that the event $\mathcal{E}^{\mathsf{c}}$ holds. Clearly, for any $k \neq k^\star$, we know that the algorithm removes the k-th when

$$\hat{\mu}_{k^\star,t} - U(t,\delta/K) > \hat{\mu}_{k,t} + U(t,\delta/K). \tag{1}$$

While the arm can also drop earlier, we note that we are only interested in an upper bound for the total amount of samples. Next, because $\mathcal{E}^{\mathsf{c}}$ holds by assumption, we have that $\hat{\mu}_{k^\star,t} \geq \mu^\star - U(t,\delta/K)$ and $\mu_k + U(t,\delta/K) \geq \hat{\mu}_{k,t}$. Therefore, Equation 1 is guaranteed to hold as long as:

$$\mu^\star - 2U(t,\delta/K) > \mu_k + 2U(t,\delta/K).$$

As a result, we obtain that the k-th arm must drop if

$$\triangle_k > 4U(t,\delta/K).$$

Next, the goal is to show that we can find a constant $c > 0$ independent of $0 < \triangle_k \leq 1$ and $K \geq 1$, such that for $T_k = c\triangle_k^{-2}\log(K\triangle_k^{-1})$, we have that $\triangle_k > 4U(T_k,\delta/K)$.

As a result, and because $U(t,\delta/K)$ is monotonically decreasing with respect to t, we can conclude that the k-th arm will be removed by the algorithm at least after $\lceil T_k \rceil$ steps. Plugging the expression for U from c) into the above equation, we get that

$$\triangle_k \geq 4\sqrt{\frac{\log(\frac{4K}{\delta}(c\triangle_k^{-2}\log(K\triangle_k^{-1}))^2)}{2c\triangle_k^{-2}\log(K\triangle_k^{-1})}} \tag{2}$$

$$\Leftrightarrow 1 \geq \frac{16\log(\frac{4n}{\delta}(c\triangle_k^{-2}\log(K\triangle_k^{-1}))^2)}{2c\log(K\triangle_k^{-1})} \tag{3}$$

Clearly, for any fixed $1 \geq \triangle_k > 0$ and $n \geq 1$, we can find c such that the inequality holds. Hence, the only thing we need to show is that we do not require $c \to \infty$ as $K \to \infty$ or $\triangle_k \to 0$. However, this follows trivially from the fact that $a\log(b) = \log(b^a)$. We can conclude that there exists $c > 0$ such that the inequality holds for all $\triangle_k$ and n. As a result, we can see that the total amount of samples for the algorithm needed to terminate is at most

$$\sum_{k\neq k^\star}\lceil T_k \rceil = \sum_{k\neq k^\star}\lceil c\triangle_k^{-2}\log(K\triangle_k^{-1})\rceil = O(\sum_{k\neq k^\star}\triangle_k^{-2}\log(K\triangle_k^{-1})).$$

.... **5 points**