1. (a) (10 marks) Basics

Among one million coins one erroneously shows "heads" on both sides; all other coins are fine. Suppose you pick one of those one million coins uniformly at random and toss it twenty times. Further suppose that it shows "heads" every single time. What is the probability that it's one of the good coins? Show your work and provide a numerical approximation.

Solution: First, we introduce a needed result.

Theorem (Bayes' theorem)**.** *Let* $(\Omega, \Pr)$ *be a finite discrete probability space and* $A, B \subset \Omega$ *events with* $\Pr(B) > 0$*. Then*

$$\Pr[A \mid B] = \frac{\Pr[B \mid A] \cdot \Pr[A]}{\Pr[B]}.$$

Proof. Consider the definition of conditional probability:

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}. \tag{1}$$

We similarly have

$$\Pr[B \mid A] = \frac{\Pr[A \cap B]}{\Pr[A]}. \tag{2}$$

Rearranging (2) yields

$$\Pr[A \cap B] = \Pr[B \mid A] \Pr[A]$$

and substituting this into (1) yields

$$\Pr[A \mid B] = \frac{\Pr[B \mid A] \cdot \Pr[A]}{\Pr[B]}$$

as required. □

Let A be the event that we pick a good coin, and B be the event that our picked coin shows "heads" every time. Note $\overline{A}$ is the event we pick the erroneous coin. We are looking for $\Pr[A \mid B]$, which we can find using Bayes' theorem. We have

$$\begin{aligned} \Pr[A] &= 1 - 10^{-6}, \\ \Pr[B \mid A] &= 2^{-20}. \end{aligned}$$

We have left to find $\Pr[B]$, but we have

$$\Pr[B] = \Pr[B \mid A]\Pr[A] + \Pr[B \mid \overline{A}]\Pr[\overline{A}]$$

by the law of total probabilities (as $\{A, \overline{A}\}$ form a partition of our sample set). We have

$$\begin{aligned} \Pr[\overline{A}] &= 10^{-6}, \\ \Pr[B \mid \overline{A}] &= 1. \end{aligned}$$

thus

$$\Pr[B] = 2^{-20}(1 - 10^{-6}) + 10^{-6}.$$

We can now apply Bayes' theorem:

$$\begin{aligned} \Pr[A \mid B] &= \frac{\Pr[B \mid A] \cdot \Pr[A]}{\Pr[B]} \\ &= \frac{(2^{-20}) \cdot (1 - 10^{-6})}{2^{-20}(1 - 10^{-6}) + 10^{-6}} \\ &= 0.488 \end{aligned}$$

correct to 3 significant figures.

(b) Tail bounds

Consider the digits in your CIS username as an integer number C, e.g., if your username were "kqlz36" then that would be $C = 36$. If you find $C < 10$ then instead let $C = 17$. Let $p = 1/C$. Consider the independent random variables $X_1, X_2, \ldots, X_{10}$ with $P(X_i = 1) = p$. Let $X = \sum_{i=1}^{10} X_i$.

Derive numerical bounds, rounded to the first non-zero digit, for $P(X \geq 10)$ using

i. (3 marks) Markov's inequality,

Solution: We have $p = 1/55$ and $\mathbb{E}[X] = 10p$. By Markov's bound,

$$\Pr[X \geq 10] \leq \frac{\mathbb{E}[X]}{10} = p.$$

ii. (5 marks) Chebyshev's inequality,

Solution: We have

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = p - p^2 = p(1-p).$$

By Chebyshev's inequality, we have

$$\Pr[|X - 10p| \geq \alpha] \leq \frac{1}{\alpha^2}\,\mathrm{Var}[X].$$

We pick $\alpha = 10(1-p)$.

Claim. $[|X - 10p| \geq \alpha] = [X - 10p \geq \alpha]$.
Proof. Let ω be an outcome such that $|X(\omega) - 10p| \geq \alpha$. Then either $X(\omega) - 10p \geq \alpha$ or $X(\omega) - 10p \leq -\alpha$. For a contradiction, assume $X(\omega) - 10p \leq -\alpha$. Then

$$\begin{aligned} X(\omega) &\leq -\alpha + 10p \\ &\leq -10(1-p) + 10p \\ &\leq 10(2p-1) \\ &\leq 10\left(\tfrac{2}{55} - 1\right) \\ &\leq -\tfrac{106}{11}, \end{aligned}$$

but $X(\omega) \in \{0,1\}$, so we have a contradiction. □

Thus,

$$\begin{aligned} \Pr[X \geq 10] &= \Pr[|X - \mathbb{E}[X]| \geq 10] \\ &\leq \tfrac{1}{10^2}\,\mathrm{Var}[X] \\ &\leq \tfrac{1}{10^2} p(1-p). \end{aligned}$$

iii. (8 marks) the generic version of Chernoff bound.

Solution: By Chernoff bound,

$$\Pr[X \geq (1+\delta)\mathbb{E}[X]] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{\mathbb{E}[X]}.$$

Thus we pick δ such that $(1+\delta)\mathbb{E}[X] = 10$:

$$\begin{aligned}(1+\delta)10p &= 10\\ \delta &= \tfrac{1}{p} - 1 \geq 0.\end{aligned}$$

Thus

$$\begin{aligned}\Pr[X \geq 10] &= \Pr[X \geq (1+\delta)\mathbb{E}[X]]\\ &\leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mathbb{E}[X]}\\ &= \left(\frac{e^{\frac{1}{p}-1}}{\left(\frac{1}{p}\right)^{\frac{1}{p}}}\right)^{10p}.\end{aligned}$$

(c) (12 marks) Martingales

Consider the Markov chain $(X_n)_{n\geq 0}$ representing a two-dimensional random walk on $\mathbb{Z}^2$, starting at $X_0 = (0,0)$, with the transition probabilities

$$\Pr[X_{n+1} = (i+a, j+b) \mid X_n = (i,j)] = \tfrac{1}{8}$$

for all $a, b \in \{-1, 0, 1\}$ with $a = b = 0$ excluded. Prove that $Z_n = \|X_n\|^2 - \frac{3}{2}n$ is a martingale with respect to $X_0, X_1, \ldots, X_{n-1}$.

Solution: Z_n is clearly a function of $(X_i)_{i=0}^n$, and has finite expectation. We now prove that $Z_n = \mathbb{E}[Z_{n+1} \mid (X_i)_{i=0}^n]$.

$$\begin{aligned}\mathbb{E}[Z_{n+1} \mid (X_i)_{i=0}^n] &= \mathbb{E}[\|X_n\|^2 - \tfrac{3}{2}(n+1) \mid (X_i)_{i=0}^n]\\ &= \mathbb{E}[\|X_{n+1}\|^2 \mid (X_i)_{i=0}^n] - \tfrac{3}{2}(n+1).\end{aligned}$$

Now

$$\begin{aligned}\mathbb{E}[\|X_{n+1}\|^2 \mid (X_i)_{i=0}^n] &= \mathbb{E}[\|X_{n+1}\|^2 \mid X_n = (i,j)]\\ &= \frac{1}{8}\sum_{(a,b)\in A}\big((i+a)^2 + (j+b)^2\big)\end{aligned}$$

where $A = \{-1, 0, 1\}^2 \setminus \{0, 0\}$. Thus

$$\mathbb{E}[\|X_{n+1}\|^2 \mid (X_i)_{i=0}^n] = \|X_n\|^2 + \tfrac{3}{2}$$

giving the required result.

(d) Markov Chains

i. (20 marks) Consider a state space $S = \{s_1, \dots, s_m\}$, and two distributions σ, τ on S. Recall that the total variation distance $d_{\mathrm{TV}}(\sigma, \tau)$ is defined to be $d_{\mathrm{TV}}(\sigma, \tau) = \frac{1}{2}\sum_{i=1}^m |\sigma_i - \tau_i|$. Prove that

$$d_{\mathrm{TV}}(\sigma, \tau) = \max_{A \subset S} |\sigma(A) - \tau(A)|$$

where $\sigma(A) = \sum_{i \in A} \sigma_i$ and $\tau(A) = \sum_{i \in A} \tau_i$.

Solution: Let f be a real-valued function and define $f^+ = \max\{0, f\}$ and $f^- = -\min\{0, f\}$. Then $f = f^+ - f^-$ and $|f| = f^+ + f^-$.
We apply this to $\sum_{i=1}^m |\sigma_i - \tau_i|$.

$$\begin{aligned} 0 = \sum_{i=1}^m (\sigma_i - \tau_i) &= \sum_{i=1}^m (\sigma_i - \tau_i)^+ - \sum_{i=1}^m (\sigma_i - \tau_i)^- \\ \sum_{i=1}^m |\sigma_i - \tau_i| &= \sum_{i=1}^m (\sigma_i - \tau_i)^+ + \sum_{i=1}^m (\sigma_i - \tau_i)^-. \end{aligned}$$

Then

$$\sum_{i=1}^m (\sigma_i - \tau_i)^+ = \sum_{i=1}^m (\sigma_i - \tau_i)^-$$

and so

$$d_{\mathrm{TV}}(\sigma, \tau) = \sum_{i=1}^m (\sigma_i - \tau_i)^+ = \sum_{i=1}^m (\sigma_i - \tau_i)^-.$$

Claim. $|\sigma(A) - \tau(A)| \le d_{\mathrm{TV}}(\sigma, \tau)$ for all $A \subset S$.

Proof.

$$\begin{aligned}
\sigma(A) - \tau(A) &= \sum_{i \in A} (\sigma_i - \tau_i) \\
&= \sum_{i \in A} (\sigma_i - \tau_i)^+ - \sum_{i \in A} (\sigma_i - \tau_i)^- \\
&\leq \sum_{i \in A} (\sigma_i - \tau_i)^+ \\
&\leq \sum_{i=1}^{m} (\sigma_i - \tau_i)^+ \\
&= d_{\mathrm{TV}}(\sigma, \tau).
\end{aligned}$$

It can be similarly be shown for $\tau(A) - \sigma(A)$. $\square$

We have left to show that there is a $A \subset S$ which attains the upper bound. This is, by definition, the maximum value, thus we pick $A = \{i \in S : \sigma_i \geq \tau_i\}$ (such that $\sigma_i - \tau_i \geq 0$). Then

$$\begin{aligned}
\sigma(A) - \tau(A) &= \sum_{i \in A} (\sigma_i - \tau_i)^+ - \sum_{i \in A} (\sigma_i - \tau_i)^- \\
&= \sum_{i \in A} (\sigma_i - \tau_i)^+ \\
&= \sum_{i=1}^{m} (\sigma_i - \tau_i)^+ \\
&= d_{\mathrm{TV}}(\sigma, \tau),
\end{aligned}$$

as required.

ii. Let $(X_n)_{n \in \mathbb{N}}$ be a Markov chain with state space S and transition matrix $P = (p_{ij})_{i,j \in S}$. We say two states $i, j \in S$ are reachable from each other if there are $n, m, \in \mathbb{N}$ such that $p_{ij}^{(m)} > 0$ and $p_{ji}^{(n)} > 0$. Prove that two states that are reachable from each other must have the same period.

Solution: Define $\mathrm{per} : S \to \mathbb{N}$ such that $\mathrm{per}(s)$ denotes the period of state $s \in S$. We have

$$\mathrm{per}(i) = \{n \in \mathbb{N} : p_{ii}^{(n)} > 0\},$$
$$\mathrm{per}(j) = \{n \in \mathbb{N} : p_{jj}^{(n)} > 0\}.$$

We trivially have

$$p_{jj}^{(n+m)} \geq p_{ji}^{(n)} p_{ij}^{(m)} > 0,$$

and so $\mathrm{per}(j) \mid (n+m)$. Define $A_i = \{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}$. Then for all $k \in A_i$,

$$p_{jj}^{(n+k+m)} \geq p_{ji}^{(n)} p_{ii}^{(k)} p_{ij}^{(m)} > 0,$$

thus $\mathrm{per}(j) \mid n+k+m$. Thus, $\mathrm{per}(j) \mid k$. As $\mathrm{per}(i) = \gcd A_i$, $\mathrm{per}(j) \leq \mathrm{per}(i)$. We can use a similar method to show that $\mathrm{per}(i) \leq \mathrm{per}(j)$, and thus $\mathrm{per}(i) = \mathrm{per}(j)$.

(e) Probabilistic Method

For $k \in \mathbb{N}$ we say that a graph $G = (V, E)$ is *k-strong-and-stable* if for every pair $X, Y \subset V$ with $X \cap Y = \varnothing$ and $|X| = |Y| = k$ we can find a vertex $u \in V$ outside of X and Y with the following properties: $\forall v \in X : (u, v) \in E$ and $\forall v \in Y : (u, v \notin E)$.

i. (4 marks) Provide an explicit 1-strong-and-stable graph. Justify your claim.

Solution: We have the trivial empty graph, which is k-strong-and-stable for all $k \in \mathbb{N}$. For a non-trivial example, we may consider C_5.

ii. (15 marks) Prove that, for each $k \in \mathbb{N}$, $k \geq 1$, there exists a k-strong-and-stable graph.

Solution:

Definition. Let $G = (V, E)$ be a graph. A *k-vertex-pair* is a pair of vertex subsets (X, Y), $X, Y \subset V$, such that $|X| = |Y| = k$ and $X \cap Y = \varnothing$.

Definition. Let $G = (V, E)$ be a graph and $k \in \mathbb{N}$. A k-vertex-pair $X, Y \subset V$ is a *bad pair* if for all $u \in V \setminus (X \cup Y)$ either

- there is $v \in X$ such that $(u, v) \notin E$; or
- there is $v \in Y$ such that $(u, v) \in E$.

Randomly construct a (undirected) graph $G = (V, E)$ with $n > 2k$ vertices such that for all disjoint $u, v \in V$, the probability of an edge between u and v is $1/2$. Let be $(X_i, Y_i)_{i=1}^m$ be an enumeration of all possible k-vertex pairs, and for each $i \in \{1, \ldots, m\}$ define the event $A_i = [(X_i, Y_i)$ is a bad pair$]$. A_i depends on every other event, thus $|\Gamma(A_i)| \leq m$. Furthermore, $|V \setminus (X_i \cup Y_i)| = n - 2k$ for all $i \in \{1, \ldots, m\}$. Fix a k-vertex-pair (X_i, Y_i) and let $u \in V \setminus (X_i \cup Y_i)$. Let

$$B = [\exists v \in X : (u, v) \notin E],$$
$$B' = [\exists v \in Y : (u, v) \in E].$$

Then

$$\Pr[B] = \Pr[B'] = 1 - \left(\tfrac{1}{2}\right)^k = 1 - 2^{-k}$$

and

$$\Pr[B \wedge B'] = \left(1 - 2^{-k}\right)^2.$$

Thus

$$\Pr[B \vee B'] = \Pr[B] + \Pr[B'] - \Pr[B \cap B'] = 1 - 2^{-2k}.$$

So

$$\Pr[A_i] = (1 - 2^{-2k})^{(n-2k)} = p.$$

Therefore, by the Lovász local lemma, we generate a k-strong-and-stable graph with non-zero probability if

$$ep(m + 1) < 1,$$

which is true for sufficiently large n.