Probability Theory 2
Proposed solution of moed bet exam 2022

1. **(a)** Since $X \sim \mathrm{U}(-1,1)$, its density function is

$$f_X(x) = \begin{cases} 1/2 & -1 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Recall (see Proposition 1.1 in Lecture 10) that if X is a random variable with probability density function f and $g : \mathbb{R} \to \mathbb{R}$ is a function, then $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$. Hence, for every positive integer n it holds that

$$\mathbb{E}(X^n) = \int_{-\infty}^{\infty} x^n f_X(x)dx = \frac{1}{2}\int_{-1}^{1} x^n dx = \frac{x^{n+1}}{2(n+1)}\Big|_{-1}^{1} = \frac{1-(-1)^{n+1}}{2(n+1)}.$$

We conclude that

$$\mathbb{E}(X^n) = \begin{cases} 0 & n \text{ is odd} \\ \frac{1}{n+1} & n \text{ is even.} \end{cases}$$

(b) It follows by Part (a) of this question that $\mathbb{E}(X_i) = 0$ and $\mathrm{Var}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = 1/3$ for every positive integer i. Therefore

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left(-\sqrt{n} \le \sum_{i=1}^{n} X_i \le \sqrt{n}\right) &= \lim_{n\to\infty} \mathbb{P}\left(\frac{-\sqrt{n} - 0\cdot n}{\sqrt{1/3}\cdot\sqrt{n}} \le \frac{\sum_{i=1}^{n} X_i - 0\cdot n}{\sqrt{1/3}\cdot\sqrt{n}} \le \frac{\sqrt{n} - 0\cdot n}{\sqrt{1/3}\cdot\sqrt{n}}\right) \\ &= \lim_{n\to\infty} \mathbb{P}\left(-\sqrt{3} \le \frac{\sum_{i=1}^{n} X_i}{\sqrt{n/3}} \le \sqrt{3}\right) = \Phi(\sqrt{3}) - \Phi(-\sqrt{3}),\end{aligned}$$

where the last equality holds by the Central Limit Theorem. We conclude that $a = -\sqrt{3}$ and $b = \sqrt{3}$ are the desired real numbers.

2. We will present a randomized algorithm and then prove that it meets all the requirements of the question.

Algorithm: For every $1 \le i \le 100$ do the following

(i) Colour the elements of $\{1, \ldots, n\}$ uniformly at random, that is, for every $1 \le j \le n$ we colour j red with probability $1/2$ and blue with probability $1/2$, all colour choices being mutually independent.

(ii) Check all sets A_j to see if any of them in monochromatic. If none of them is monochromatic, then return this colouring.

First, we analyse the running time of the algorithm. Consider any one of the (at most) 100 rounds of the algorithm. Colouring any single element of $\{1, \ldots, n\}$ takes constant time, so colouring the entire set takes time $\Theta(n)$. Checking whether the resulting colouring is *good* (that is, A_i contains a red element and a blue element for every $1 \leq i \leq m$) or *bad* (that is, there exists some $1 \leq i \leq m$ such that A_i is monochromatic), takes time $O\left(\sum_{i=1}^m |A_i|\right) = O(mk)$. Since there are at most 100 rounds, we conclude that the running time of the algorithm is $O(n + mk)$.

Next, we prove the (randomized) correctness of the algorithm. Fix some $1 \leq i \leq m$. If the elements of $\{1, \ldots, n\}$ are coloured red and blue uniformly at random, then the probability that A_i is monochromatic is $2 \cdot 2^{-|A_i|} = 2^{1-k}$. A union bound over all sets A_i then shows that the probability of a random colouring being bad is at most $m2^{1-k} \leq 1/2$. The algorithm fails to produce a good colouring if and only if it produces 100 bad colourings. The probability of this occurring is at most 2^{-100}, as required.

3. Let $G \sim G(n,p)$, where $p = \frac{\ln n}{n}$.

   (a) A union bound argument shows that the probability of the event in question is at most

$$n\binom{n-1}{10\ln n} p^{10\ln n} \leq n\left(\frac{enp}{10\ln n}\right)^{10\ln n} = n\left(\frac{e}{10}\right)^{10\ln n} \leq ne^{-10\ln n} = n^{-9}.$$

   We conclude that

$$\lim_{n\to\infty} \mathbb{P}(\text{there exists a vertex in } G \text{ of degree at least } 10\ln n) = 0.$$

   (b) Fix some set $A \subseteq V(G)$ of size $1 \leq t \leq n^{0.9}$. The probability that A spans at least $2t$ edges is at most

$$\binom{\binom{t}{2}}{2t} p^{2t} \leq \binom{t^2/2}{2t} p^{2t} \leq \left(\frac{etp}{4}\right)^{2t} \leq \left(t^2p^2\right)^t.$$

   Therefore, a union bound over all values $1 \leq t \leq n^{0.9}$ and all sets $A \subseteq V(G)$ of size t shows that the probability that there exists a set $A \subseteq V(G)$ of size $1 \leq t \leq n^{0.9}$ which spans at least $2t$ edges is at most

$$\sum_{t=1}^{n^{0.9}} \binom{n}{t} \left(t^2p^2\right)^t \leq \sum_{t=1}^{n^{0.9}} \left(entp^2\right)^t \leq \sum_{t=1}^{n^{0.9}} \left(en^{-0.1}\ln^2 n\right)^t \leq \sum_{t=1}^{\infty} n^{-t/20} = \frac{n^{-1/20}}{1-n^{-1/20}}.$$

   We conclude that

$$\lim_{n\to\infty} \mathbb{P}(\exists A \subseteq V(G) \text{ of size } 1 \leq t \leq n^{0.9} \text{ which spans at least } 2t \text{ edges}) = 0.$$