Please provide complete and well-written solutions to the following exercises.

Due March 1, at the beginning of class.

Homework 6

Exercise 1 (Triangular Arrays)**.** For any $n \geq 1$, let $X_{n,1},\ldots,X_{n,n}\colon \Omega \to \mathbf{R}$ be a collection of independent random variables, and let $S_n = X_{n,1}+\cdots+X_{n,n}$. Let $\mu \in \mathbf{R}$.

- (i) (Weak law) If $\mathbf{E}X_{n,i}=\mu$ for all $1\leq i\leq n$ and $\sup_{i,n}\mathbf{E}\left|X_{n,i}\right|^2<\infty$, show that S_n/n converges in probability to μ.
- (ii) (Strong law) If $\mathbf{E}X_{n,i}=\mu$ for all $1\leq i\leq n$ and $\sup_{i,n}\mathbf{E}\left|X_{n,i}\right|^4<\infty$, show that S_n/n converges almost surely to μ.

Exercise 2. For any natural number n and a parameter $0<p<1$, define an Erdös-Renyi graph on n vertices with parameter p to be a random graph (V,E) on a (deterministic) vertex set V of n vertices (thus (V,E) is a random variable taking values in the discrete space of all $2^{\binom{n}{2}}$ possible undirected graphs one can place on V) such that the events $\{i,j\}\in E$ for unordered pairs with $i,j\in V$ are independent and each occur with probability p.

For each $n\geq 1$, let (V_n,E_n) be an Erdös-Renyi graph on n vertices with parameter $p=1/2$ (we do not require the graphs to be independent of each other).

- (i) Let $|E_n|$ be the number of edges in (V_n,E_n). Show that $|E_n|/\binom{n}{2}$ converges almost surely to $1/2$ (Hint: use Exercise 1.)
- (ii) Let $|T_n|$ be the number of triangles in (V_n,E_n) (i.e. the set of unordered triples $\{i,j,k\}$ with $i,j,k\in V_n$ such that $\{i,j\},\{i,k\},\{j,k\}\in E_n$), show that $|T_n|/\binom{n}{3}$ converges in probability to $1/8$. (Note: there is not quite enough independence here to directly apply the law of large numbers, so try using the second moment method directly.)
- (iii) Show in fact that $|T_n|/\binom{n}{3}$ converges almost surely to $1/8$. (Note: you don't need to compute the fourth moment here.)

Exercise 3. For each $n\geq 1$, let $A_n=(a_{ij,n})_{1\leq i,j\leq n}$ be a random $n\times n$ matrix (i.e. a random variable taking values in the space $\mathbf{R}^{n\times n}$ or $\mathbf{C}^{n\times n}$ of $n\times n$ matrices) such that the entries $a_{ij,n}$ of A_n are independent in i,j and take values in $\{-1,1\}$ with a probability of $1/2$ each. We do not assume any independence for the sequence $A_1,A_2,\ldots$.

- (i) Show that the random variables $\mathrm{Tr}A_nA_n^*/n^2$ are equal to the constant 1, where A_n^* denotes the matrix adjoint (which, in this case, is also the transpose) of A_n and Tr denotes the trace (or sum of the diagonal entries) of a matrix.

(ii) Show that for any natural number $k \geq 1$, the quantities $\mathbf{E}\mathrm{Tr}(A_n A_n^*)^k / n^{k+1}$ are bounded uniformly in $n \geq 1$ (i.e. they are bounded by a quantity C_k that can depend on k but not on n). (It may be helpful to first try $k = 2$ and $k = 3$.)

(iii) Let $\|A_n\|$ denote the operator norm of A_n, and let $\varepsilon > 0$. Show that $\|A_n\|/n^{1/2+\varepsilon}$ converges almost surely to zero, and that $\|A_n\|/n^{1/2-\varepsilon}$ diverges almost surely to infinity. (Hint: use the spectral theorem to relate $\|A_n\|$ with the quantities $\mathrm{Tr}(A_n A_n^*)^k$.

Exercise 4. The Cramér random model for the primes is a random subset $\mathcal{P}$ of the natural numbers such that $1 \notin \mathcal{P}$, $2 \in \mathcal{P}$, and the events $n \in \mathcal{P}$ for $n = 3, 4, \ldots$ are independent with $\mathbf{P}(n \in \mathcal{P}) := \frac{1}{\log n}$. Here we used the restriction $n \geq 3$ so that $\frac{1}{\log n} < 1$. This random set of integers $\mathcal{P}$ gives a reasonable way to model the primes $2, 3, 5, 7, \ldots$, since by the Prime Number Theorem, the number of primes less than n is approximately $n/\log n$, so the probability of n being a prime should be about $1/\log n$. The Cramér random model can provide heuristic confirmations for many conjectures in analytic number theory:

- (Probabilistic prime number theorem) Show that $\frac{1}{x/\log x}|\{n \leq x : n \in \mathcal{P}\}|$ converges almost surely to one as $x \to \infty$.
- (Probabilistic Riemann hypothesis) Let $\varepsilon > 0$. Show that
$$\frac{1}{x^{1/2+\varepsilon}}\Big(|\{n \leq x : n \in \mathcal{P}\}| - \int_2^x \frac{dt}{\log t}\Big)$$
converges almost surely to zero as $x \to \infty$.
- (Probabilistic twin prime conjecture) Show that almost surely, there are an infinite number of elements p of $\mathcal{P}$ such that $p + 2$ also lies in $\mathcal{P}$.
- (Probabilistic Goldbach conjecture) Show that almost surely, all but finitely many natural numbers n are expressible as the sum of two elements of $\mathcal{P}$.

Exercise 5. This exercise proves the Hardy-Ramanujan Theorem. This theorem, with probabilistic proof due to Turán, says that a typical large $n \in \mathbf{N}$ has about $\log\log n$ distinct prime factors. Unlike the previous exercise, the probabilistic proof here proves a rigorous result about primes.

Let $\mathcal{P} \subseteq \mathbf{N}$ denote the set of prime numbers (in this exercise $\mathcal{P}$ is deterministic, not random). When $p \in \mathcal{P}$ and $n \in \mathbf{N}$, we use the notation $p|n$ to denote "p divides n," i.e. n/p is a positive integer. Let $x \geq 100$ with $x \in \mathbf{N}$ (so that $\log\log x \geq 1$), and let N be a natural number that is uniformly distributed in $\{1, 2, \ldots, x\}$. Assume Mertens' theorem
$$\sum_{p \in \mathcal{P} \colon p \leq x} \frac{1}{p} = \log\log x + O(1).$$

- Show that the random variable $\sum_{p \in \mathcal{P} \colon p \leq x^{1/10}} 1_{p|N}$ has mean $\log\log x + O(1)$ and variance $O(\log\log x)$. (Hint: up to reasonable errors, compute the means, variances and covariances of the random variables $1_{p|N}$.)
- For any $n \in \mathbf{N}$, let $f(n)$ denote the number of distinct prime factors of n. Show that $\frac{f(N)}{\log\log N}$ converges to 1 in probability as $x \to \infty$. (Hint: first show that $f(N) = \sum_{p \in \mathcal{P} \colon p \leq x^{1/10}} 1_{p|N} + O(1)$.) More precisely, show that
$$\frac{f(N) - \log\log N}{g(N)\sqrt{\log\log N}}$$

converges in probability to zero as $x \to \infty$, whenever $g\colon \mathbf{N} \to \mathbf{R}$ is any function satisfying $\lim_{n\to\infty} g(n) = \infty$.

Exercise 6 (Kronecker's Lemma)**.** Let $y_1, y_2, \ldots$ be a sequence of real numbers. Let $0 < b_1 \leq b_2 \leq \cdots$ be a sequence of real numbers that goes to infinity. Assume that $\lim_{n\to\infty} \sum_{m=1}^{n} y_m$ exists. Then $\lim_{n\to\infty} \frac{1}{b_n} \sum_{m=1}^{n} b_m y_m = 0$. (Hint: if $s_n := \sum_{m=1}^{n} y_m$, then the summation by parts formula implies that $\frac{1}{b_n} \sum_{m=1}^{n} b_m y_m = s_n - \frac{1}{b_n} \sum_{m=1}^{n-1} (b_{m+1} - b_m) s_m$.)

Exercise 7 (Renewal Theory)**.** Let $t_1, t_2, \ldots$ be positive, independent identically distributed random variables. Let $\mu \in \mathbf{R}$. Assume $\mathbf{E}t_1 = \mu$. For any positive integer j, we interpret t_j as the lifetime of the j^{th} lightbulb (before burning out, at which point it is replaced by the $(j+1)^{st}$ lightbulb). For any $n \geq 1$, let $T_n := t_1 + \cdots + t_n$ be the total lifetime of the first n lightbulbs. For any positive integer t, let $N_t := \min\{n \geq 1 \colon T_n \geq t\}$ be the number of lightbulbs that have been used up until time t. Show that N_t/t converges almost surely to $1/\mu$ as $t \to \infty$. (Hint: if c, t are positive integers, then $\{N_t \leq ct\} = \{T_{ct} \geq t\}$. Apply the Strong Law to T_{ct}.)

Exercise 8 (Playing Monopoly Forever)**.** Let $t_1, t_2, \ldots$ be independent random variables, all of which are uniform on $\{1, 2, 3, 4, 5, 6\}$. For any positive integer j, we think of t_j as the result of rolling a single fair six-sided die. For any $n \geq 1$, let $T_n = t_1 + \cdots + t_n$ be the total number of spaces that have been moved after the n^{th} roll. (We think of each roll as the amount of moves forward of a game piece on a very large Monopoly game board.) For any positive integer t, let $N_t := \min\{n \geq 1 \colon T_n \geq t\}$ be the number of rolls needed to get t spaces away from the start. Using Exercise 7, show that N_t/t converges almost surely to $2/7$ as $t \to \infty$.

Exercise 9 (Random Numbers are Normal)**.** Let X be a uniformly distributed random variable on $(0, 1)$. Let X_1 be the first digit in the decimal expansion of X. Let X_2 be the second digit in the decimal expansion of X. And so on.

- Show that the random variables $X_1, X_2, \ldots$ are uniform on $\{0, 1, 2, \ldots, 9\}$ and independent.
- Fix $m \in \{0, 1, 2, \ldots, 9\}$. Using the Strong Law of Large Numbers, show that with probability one, the fraction of appearances of the number m in the first n digits of X converges to $1/10$ as $n \to \infty$.

(Optional): Show that for any ordered finite set of digits of length k, the fraction of appearances of this set of digits in the first n digits of X converges to 10^{-k} as $n \to \infty$. (You already proved the case $k = 1$ above.) That is, a randomly chosen number in $(0, 1)$ is normal. On the other hand, if we just pick some number such that $\sqrt{2} - 1$, then it may not be easy to say whether or not that number is normal.

(As an optional exercise, try to explicitly write down a normal number. This may not be so easy to do, even though a random number in $(0, 1)$ satisfies this property!)

Exercise 10 (Cheap Law of the Iterated Logarithm)**.** Let $X_1, X_2, \ldots \colon \Omega \to \mathbf{R}$ be independent random variables with mean zero and variance one. The Strong Law of Large Numbers says that $\frac{1}{n}(X_1 + \cdots + X_n)$ converges almost surely to zero (if the random variables are also identically distributed). The Central Limit Theorem says that $\frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)$ converges in distribution to a standard Gaussian random variable (if the random variables are also identically distributed). But what happens if we divide by some function of n in between $n^{1/2}$ and n ? This Exercise gives a partial answer to this question.

Let $\varepsilon > 0$. Show that

$$\frac{X_1 + \cdots + X_n}{n^{1/2}(\log n)^{(1/2)+\varepsilon}}$$

converges to zero almost surely as $n \to \infty$. (Hint: Re-do the proof of the Strong Law of Large Numbers, but divide by $n^{1/2}(\log n)^{(1/2)+\varepsilon}$ instead of n. You don't need to do any truncation.)