Homework 11

Instructions: Standard course policies about typesetting, file size, and submission apply. You must show your work to receive credit. Your work must be your own, though you are permitted to get assistance from classmates or instructional staff. Your responses to the submission problems must be uploaded to Submitty by **8:59pm on Thursday, April 3**.

Attribution: these problems were chosen from the DMC text.

Note: you can employ Monte Carlo simulations to check some of your answers. But of course your arguments must be sound and not reference any simulations.

Recitation Problems

I. Let $\mathbf{X}$ be the waiting time with success probability p. Compute $f(n) = \mathbb{P}[\mathbf{X} \geq n]$ and $g(n)$, the Chebyshev upper bound on $\mathbb{P}[\mathbf{X} \geq n]$. Show that $f(n) = o(g(n))$. (Chebyshev's bound can be asymptotically worse than reality).

II. A couple has kids until five boys. Estimate $\mathbb{P}[7 \leq \text{number of children} \leq 13]$ using Chebyshev's inequality and compare with the true value.

III. 100 people toss their hats up. The hats land randomly on heads. Let the random variable $\mathbf{X}$ be the number of people who get their hats back.

   (a) Compute $\mathbb{E}[\mathbf{X}]$ and $\text{var}[\mathbf{X}]$.
   (b) Give a (non-trivial) upper bound on the probability that more than half the people get their hats back.

IV. An aggressive drunk takes 10 steps $\mathbf{X}_1, \ldots, \mathbf{X}_{10}$. Each step is independent, and moves left or right with equal probability. The size of the step increases with time, $|\mathbf{X}_i| = i$. Use linearity to compute the expected value and standard deviation of the drunk's position after 10 steps.

V. Prove that the unweighted finite graphs are countable. What about weighted graphs with integer weights? What if the weights are real?

VI. The positive rationals are $\mathbb{Q}_+ = \left\{ \frac{x}{y} \mid x, y \in \mathbb{N} \right\}$. We will use $\mathbb{Q}_+$ to prove that $\mathbb{Q}$ is countable.

   (a) Argue that $|\mathbb{Q}_+| \leq |\mathbb{N}|$ by establishing that $f(x, y) = 2^x 3^y$ is an injection from $\mathbb{Q}_+$ to $\mathbb{N}$.
   (b) Show that $\mathbb{Q}_+$ being countable means that $\mathbb{Q}$ is countable.

VII. Prove that eventually constant infinite sequences on $\mathbb{N}$ are countable.

VIII. Prove that increasing sequences on $\mathbb{N}$ have the same cardinality as $\mathbb{R}$.

Submission Problems

(1) An urn has 10 white and 20 black balls. Let $\mathbf{X}_1$ be the waiting time (number of balls sampled) until a white ball appears and $\mathbf{X}_2$ the waiting time until two white balls

appear. Use indicator random variables to find the mean and variance of $\mathbf{X}_1$ and $\mathbf{X}_2$ in each setting: (a) you sample with replacement, (b) you sample without replacement.

(2) Prove $\sum_{i=1}^n x_i^2 \geq \frac{1}{n}\left(\sum_{i=1}^n x_i\right)^2$.

(3) You are building a circuit board and need 50 transistors. About 2% of transistors are defective. Use Chebyshev's inequality to estimate the number of transistors you should order to ensure at least a 99% chance that you will have enough non-defective transistors. Compare with the correct number you should order.

(4) Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be independent random variables with means $\mu_1, \ldots, \mu_n$ and variances $\sigma_1^2, \ldots, \sigma_n^2$. In a Monte Carlo simulation, you use a random variable generator to generate values for $\mathbf{X}_1, \ldots, \mathbf{X}_n$. Let the values generated by $z_1, \ldots, z_n$. The average of these values is the sample mean $\overline{z} = (z_1 + \cdots + z_n)/n$. Similarly, the sample variance is the average of the squared-deviations from the sample mean:

$$s^2 = \frac{1}{n}\sum_{i=1}^n (z_i - \overline{z})^2.$$

(a) Show that $\mathbb{E}[\overline{z}] = (\mu_1 + \cdots + \mu_n)/n$, the average of the means.
(b) Show that $s^2 = \frac{1}{n}\sum_{i=1}^n z_i^2 - \overline{z}^2$ (the average of the squares minus the squared average).
(c) Show that $\mathbb{E}[s^2] = \left(\frac{n-1}{n}\right) \cdot \frac{1}{n}\sum_{i=1}^n \sigma_i^2 + \frac{1}{n}\sum_{i=1}^n \mu_i^2 - \left(\frac{1}{n}\sum_{i=1}^n \mu_i\right)^2$. (A term slightly below average$(\sigma_1^2, \ldots, \sigma_n^2)$ plus variance$(\mu_1, \ldots, \mu_n)$, which measures the spread in the μ_i.)
(d) When each $\mathbf{X}_i$ has the same mean μ and variance σ^2, show that $\mathbb{E}[\overline{z}] = \mu$ and $\mathbb{E}[s^2] = \left(\frac{n-1}{n}\right)\sigma^2$.
(e) Justify the following Monte Carlo approach to estimating $\mathbb{E}[\mathbf{X}$ and $\sigma^2(\mathbf{X})$. Generate n values for $\mathbf{X}$ having sample mean $\overline{z}$ and sample variance s^2. Estimate $\mu = \overline{z}$ and $\sigma^2 = \left(\frac{n}{n-1}\right)s^2$.

(5) True or false. Explain your answers.

(a) The collection of all subsets of $\mathbb{N}$ is countable.
(b) The collection of all finite subset of $\mathbb{N}$ is countable.

(6) Use ideas from Problem VI in the recitation set to prove countable sets are closed under union and Cartesian product. Let $\{p_1, p_2, \ldots\}$ be primes. Let $\{A_1, A_2, \ldots\}$ be a countable set, where each $A_i = \{a_{i,1}, a_{i,2}, \ldots\}$ is itself countable.

(a) For $x \in \cup_i A_i$, let A_ℓ be the first set containing x, i.e. ℓ is such that $x \in A_\ell$ and $x \notin A_j$ for $j < \ell$. Let x be the kth element in A_ℓ. Use $f(x) = p_\ell^k$ to prove that $\cup_i A_i$ is countable. That is, the union of countably many countable sets is countable.
(b) Let $x = (x_1, x_2, \ldots, x_n)$ be an ordered n-tuple from $A_1, \ldots, A_n$, so $x \in A_1 \times A_2 \times \cdots \times A_n$.

(i) Let x_i be the k_i-th element in A_i. Use $f(x_1, \ldots, x_n) = p_1^{k_1} p_2^{k_2} \cdots p_n^{k_n}$ to prove that $A_1 \times A_2 \times \cdots \times A_n$ is countable. That is, the Cartesian product of finitely many countable sets is countable.

(ii) The argument in (i) fails for an infinite Cartesian product. In fact, prove that the Cartesian product of countably many sets is uncountable, if infinitely many of the sets have size at least 2.

(7) Prove or disprove that each intersection of the following collection of sets is always countable.

(a) Countably many countable sets.
(b) Uncountably many countable sets.
(c) Countably many uncountable sets.

(8) Algebraic numbers are solutions to integer-polynomial equations,

$$a_1x^{k_1} + a_2x^{k_2} + \cdots + a_\ell x^{k_\ell} = 0,$$

where a_i, k_i are integers and $k_1 < k_2 < \cdots < k_\ell$ are nonnegative. A number is transcendental if it is not algebraic. Prove that the transcendentals, e.g. π and e far outnumber the algebraics, e.g. 2 and $\sqrt{2}$, by showing that

$$|\{\text{algebraic numbers}\}| = \mathbb{N}$$
$$|\{\text{transcendental numbers}\}| > |\mathbb{N}|.$$