(Fall 2022) 515502 Probability **Early Bird: 2022/10/31, 11am; Normal: 2022/11/1, 11am**

Homework 2: Random Variables, Expectation, and Moments

Submission Guidelines: Please combine all your write-ups (photos/scanned copies are acceptable; please make sure that the electronic files are of good quality and reader-friendly) into one single .pdf file and submit the file via E3.

Problem 1 (Special Random Variables) (10+10+10=30 points)

(a) Let $X_1, X_2, \cdots, X_n$ be n independent Geometric random variables with the same success probability $p \in (0, 1)$. Define $X = \max(X_1, \cdots, X_n)$ and $Y = \min(X_1, \cdots, X_n)$. What is the PMF of X? Moreover, what is the PMF of Y? What kind of random variable is Y?

(b) Suppose we are given an urn that initially contains 1 red ball and 1 blue ball. In every step, we choose from the urn a random ball which we replace, and then we also insert another ball of the same color (e.g., if the first choice was red, then in the second step we have 2 red balls and one blue ball in the urn). Define a random variable X to be the number of red balls in the urn after the n-th step. Then, show that X is a discrete uniform random variable with parameters $(1, n + 1)$. This classic problem is called the *Polya's urn*. (Hint: Show this by induction)

(c) Let $X \sim \mathcal{N}(\mu, \sigma^2)$ be a normal random variable. Consider another random variable $Y = a|X| + b$, where a, b are real numbers. Please write down the CDF and a PDF of Y. Under what condition is Y also a normal random variable? (Hint: Try to write down the CDF of Y and then derive the PDF by taking the derivative as discussed in Lecture 11)

Problem 2 (PMF and Entropy) (10+10=20 points)

Quantifying the amount of information carried by a random variable has been a central issue in various domains, including machine learning and wireless communication. In this problem, let us take a quick look at one fundamental and super useful concept in information theory, namely the *entropy* of the distribution of a discrete random variable.

(a) Consider a discrete random variable X with the set of possible values $\{1, 2, \cdots, n\}$. Define $p_i := P(X = i)$, for all $i = 1, \cdots, n$. Next, we define a metric called *entropy*:

$$H(X) := -\sum_{i=1}^{n} p_i \ln p_i.$$

(Note: "ln" means the base of the logarithm is e). What is the maximum possible value of $H(X)$? Find the PMF $\{p_i\}_{i=1}^n$ that achieves this maximum. (Hint: You may use the weighted inequality of arithmetic and geometric means, i.e., $\frac{w_1x_1+w_2x_2+\cdots+w_nx_n}{w} \geq (x_1^{w_1} \cdot x_2^{w_2} \cdots x_n^{w_n})^{\frac{1}{w}}$, where $\{w_i\}$ and $\{x_i\}$ are non-negative real numbers and $w = w_1 + \cdots + w_n$)

(b) Given the same setting of (a), what is the minimum possible value of $H(X)$? Please find out all the PMFs $\{p_i\}_{i=1}^n$ that achieve this minimum.

Problem 3 (Expectation and Moments) (10+10+10=30 points)

(a) Suppose $X \sim \text{Geometric}(p)$. Show that (i) $E[X] = 1/p$; (ii) $E[e^{tX}] = \frac{pe^t}{1-(1-p)e^t}$, for $t < -\ln(1 - p)$. (Hint: Regarding $E[X]$, write down the PMF and try to reuse the fact that the total probability is 1.)

(b) Let $n \geq 3$ be a positive integer. Let X and Y be two discrete random variables with the identical set of possible values $\{a_1, a_2, \cdots, a_n\}$, where $a_1, a_2, \cdots, a_n$ are n distinct real numbers. Show that $E[X^m] = E[Y^m]$ for all $m \in \{1, 2, \cdots, n - 1\}$ if and only if X and Y are identically distributed, i.e., $P(X = t) = P(Y = t)$, for all $t \in \{a_1, \cdots, a_n\}$. (Note: This property is typically called *moment matching*)

(c) Let $z_n = (-1)^n \sqrt{n}$, for $n = 1, 2, 3 \cdots$. Let Z be a discrete random variable with the set of possible values $\{z_n : n = 1, 2, 3 \cdots\}$. The PMF of Z is

$$p_Z(z_n) = P(Z = z_n) = \frac{6}{(\pi n)^2}, \ \forall n \in \mathbb{N}.$$

What is $\mathrm{Var}[Z]$? How about the value of $\sum_{n=1}^{\infty} z_n^3 \cdot p_Z(z_n)$? Does $E[Z^3]$ exist? Please carefully justify each step of your answer.

Problem 4 (Inverse Transform Sampling) (10+10=20 points)

In Lecture 12, we learn *inverse transform sampling*, a general method for generating random variables with customized distributions. In this problem, let us take a closer look at how we shall generate continuous and discrete random variables.

(a) Suppose we are given the PDF of a random variable X as follows: Let λ be a positive constant.

$$f(x) = \begin{cases} \lambda \exp(-\lambda x), & \text{if } x > 0. \\ 0, & \text{otherwise.} \end{cases}$$

How shall we generate such a random variable X with PDF $f(x)$ via inverse transform sampling? (Hint: You may first derive the CDF and then find its inverse)

(b) Similarly, suppose we are given the PMF of a random variable Y as follows:

$$P(Y = k) = \begin{cases} 0.1, & \text{if } k = 1, \\ 0.3, & \text{if } k = 3, \\ 0.4, & \text{if } k = 6, \\ 0.2, & \text{if } k = 10, \\ 0, & \text{otherwise.} \end{cases}$$

How shall we generate such a random variable Y via inverse transform sampling?