MATH 8000

Exam 1: Due TBD

Problem 1: Let X_1 and X_2 be independent random variables with common probability mass function

$$f_X(x|\theta) = \frac{-\theta^x}{x\log(1-\theta)}$$

where $x = 1, 2, 3, ...$ and $\theta \in (0, 1)$.

(a) Find the mean and variance of X_1.

(b) Define the new random variable Y such that $Y = -I(X_1 = 1)$, and derive the probability mass function of Y. Then calculate the $E(Y)$.

(c) Consider the random variable $T(X_1, X_2) = X_1 + X_2$, derive the probability mass function of T.

(d) Consider the function $U(t) = -P(X_1 = 1|T = t)$, show that the $E[U(T)] = E(Y)$.

Problem 2: Chase and Mike each have a box containing balls which are numbered $1, 2, 3, ..., N$. Each of them will draw N balls from their respective boxes in succession, and will say that there is a **match** if at any of the N draws, the numbers on their chosen balls coincide; e.g., if at the 10th draw both of them select a ball numbered 32, then there is a match.

(a) If the drawing of the balls is **with replacement**, what is the probability that a match occurs? Provide a general formula, and evaluate it when $N = 5$.

(b) What happens to the probability in (a) when you let $N \to \infty$?

(c) If the drawing of the balls is **without replacement**, what is the probability that a match occurs? Provide a general formula, and evaluate it when $N = 5$.

(d) What happens to the probability in (c) when you let $N \to \infty$?

(e) Compare the limiting probabilities in (b) and (d). Are they the same? Comment on your results.

Problem 3: Consider a counting process $N(t)$, where for each $t > 0$, the random variable $N(t)$ is the number of events that have occurred up to and including time t. Let S_n denote the time at which the nth event occurs.

(a) For a given integer $n \geq 1$ and time $t > 0$, argue that the nth arrival epoch S_n, and the counting random variable, $N(t)$, are related by

$$\{S_n \leq t\} = \{N(t) \geq n\}$$

For the remainder of this problem, we will assume that $S_n \sim$ Gamma(n, λ^{-1}); i.e., we are considering a Poisson process with rate parameter λ.

(b) Consider a time point t_0, for which $N(t_0) = 0$. Derive the distribution of the nonnegative random variable Z which denotes the length of the interval from t_0 until the first arrival.

(c) Show that for any $t > 0$ that $N(t) \sim \text{Poisson}(\lambda t)$.

Problem 4. Sampling at random from a finite/infinite population: In this problem we will consider a series of convenient distributional approximations.

(a) Finite population: Consider the problem of selecting, without replacement at random, K units from a finite population consisting of N total units. Further assume that the population can be divided into two distinct groups consisting of M Class I units and the other of $N - M$ Class II units. Define the random variable

$$X = \{\text{to be the number of Class I units selected}\}.$$

Derive the PMF of X and prove that it is valid. Use the PMF to find the $E(X)$ and $V(X)$. (**Note**: Your derivation of the the PMF should be accompanied by a short discussion; i.e., you need to do more than simply state it.)

(b) Infinite population: Consider the problem of selecting K units at random from an infinite population. Further assume that each unit has a probability p of being Class I. Define the random variable

$$Y = \{\text{to be the number of Class I units selected}\}.$$

Derive the PMF of Y and prove that it is valid. Use the PMF to find the $E(Y)$ and $V(Y)$. (**Note**: Your derivation of the the PMF should be accompanied by a short discussion; i.e., you need to do more than simply state it.)

(c) Calculating probabilities directly from the PMF of X could, and will, become tedious when M and N are large (**Why?**). Consequently, we will develop an approximation to these probabilities. In particular, for the random variable X consider letting $N \to \infty$ and $M \to \infty$, such that $M/N \to p$. In this situation, prove that the resulting PMF for X is in fact the PMF that you derived for Y. Notice the PMF of Y is free of N and M so for K reasonably small we can use the PMF of Y to approximate the PMF of X. Discuss how you would do this.

(d) In the approximation above, when K is large we can again run into trouble (**Why?**). So we will develop an approximation to the PMF of Y. The parameters controlling the distribution of Y are p and K, so to emphasize this we write $Y_{K,p}$. Find the limiting distribution of $Y_{K,p}$ if $p \to 0$ as $K \to \infty$, such that $Kp \to \alpha$, for some constant α. Once this is established, one should note that the distribution of the limiting random variable is free of K and depends entirely on α. Suggest how this result could be used to approximate the PMF of $Y_{K,p}$.

(e) Another approximation for the distribution of $Y_{K,p}$ can be developed by first defining the new random variable $Z_K = \sqrt{K}\left(\frac{Y_{K,p}}{K} - p\right)$. Now derive the limiting distribution of Z_K (i.e., as $K \to \infty$), and suggest how it could be used to approximate the distribution of $Y_{K,p}$. (Hint: the moment generating function for a normal random variable W whose mean is μ and variance is σ^2 is given by $M_W(t) = \exp\{\mu t + \sigma^2 t^2/2\}$.)

Problem 5. Consider a continuous random variable X whose PDF is given by

$$f_X(x) = \frac{1}{\beta} e^{-x/\beta} I(x > 0).$$

(a) Find the distribution of $Y = \lfloor X \rfloor$, where $\lfloor \cdot \rfloor$ is the usual floor function. The random variable Y follows a named distribution that we have encountered before, you must identify what this distribution and explicitly identify its parameters in terms of β. Derive the $E(Y)$, $V(Y)$, the CDF of Y, and $M_Y(t)$.

(b) Find the distribution of $Y = X - \lfloor X \rfloor$. Derive the $E(Y)$, $V(Y)$, the CDF of Y, and $M_Y(t)$.

Problem 6. Two real numbers X and Y are chosen at random in the interval $(0, 1)$. Define the random variable $Z = \{\text{the closest integer to } X/Y\}$.

(a) Derive the PMF of Z and show that it is valid. In order to show that the PMF is valid you will be working with an infinite series, you should derive the necessary results to complete the problem.

(b) Compute the probability that the closest integer to Z is even. Your solution should involve π and should be concisely presented; i.e., no infinite sum representations.

Problem 7. Suppose there are N different types of coupons available when buying cereal; each box contains one coupon and the collector is seeking to collect one of each in order to win a prize. After buying n boxes, what is the probability that the collector has at least one of each type?

Problem 8. Does a distribution exist for which $M_X(t) = t/(1-t)$, $|t| < 1$? If yes, find it. If not, prove it.

Problem 9. Let S denote a finite sample space consisting of N sample points. Now consider randomly and independently sampling from S two subsets A and B. Calculate $P(A \subset B)$.

Problem 10. Let X and Y be any two random variables and define $L = min(X;Y)$ and $U = max(X;Y)$, show that

$$E(U) = E(X) + E(Y) - E(L).$$