Due Monday August 1 at 1:59PM

1. **Beta Testing** (10 points: 4/6)

   Midterm 2 of CS70 consists of n questions. Before the test, m TAs are going to beta test the midterm. Each of them will solve a question correctly with probability p (independently of other TAs and independently of other questions). Let X be the number of distinct questions that no one solves correctly.

   (a) What is the expectation of X? What's the variance of X?

   (b) Now each TA is going to choose a question uniformly at random from the n questions to grade (independently of other TAs). Let Y be the number of distinct questions that no one chooses. What is the expectation of Y? What's the variance of Y?

2. **Distributions** (16 point: 2/2/2/2/2/2/2/2)

   For the following scenarios, what distribution would best model each one of them individually? You may choose from Poisson, geometric, exponential, uniform or normal distribution. Provide a brief justification of your answer.

   (a) Amount of time you need to wait until a fly enters through your window.

   (b) Number of shooting stars seen on a given night.

   (c) Average height of students in CS70.

   (d) Angle between two needles when we spin them at random about a common point.

   (e) Number of times you have to refresh a webpage until it loads.

   (f) Average grade in a CS70 exam.

   (g) Number of times a web server is accessed per minute.

   (h) Amount of time until the next time your telephone rings.

3. **Poisson**(12 points: 6/6)

   (a) It is fairly reasonable to model the number of customers entering a shop during a particular hour as a Poisson random variable. Assume that this Poisson random variable X has mean λ. Suppose that whenever a customer enters the shop they leave the shop without buying anything with probability p. Assume that customers act independently, i.e. you can assume that they each simply flip a biased coin to decide whether to buy anything at all. Let us denote the number of customers that buy something as Y and the number of them that do not buy anything as Z (so $X = Y + Z$). What is the probability that $Y = k$ for a given k? How about $\Pr[Z = k]$? Prove that Y and Z are Poisson random variables themselves.
   Hint: you can use the identity $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

(b) Assume that you were given two independent Poisson random variables X_1, X_2. Assume that the first has mean λ_1 and the second has mean λ_2. Prove that $X_1 + X_2$ is a Poisson random variable with mean $\lambda_1 + \lambda_2$.

4. **Bound It** (6 points)

   A random variable X is always strictly larger than -100. You know that $E[X] = -60$. Give the best upper bound you can on $P(X \geq -20)$. (Hint: think of the information you are given and the information you require to compute certain bounds)

5. **LLN** (18 points: 3/5/5/5)

   The Law of Large Numbers holds if, for all $\varepsilon > 0$,

   $$\lim_{n\to\infty} \Pr[\left|\tfrac{1}{n}S_n - \mathbb{E}(\tfrac{1}{n}S_n)\right| > \varepsilon] = 0.$$

   In class, we see the proof of the Law of Large Numbers for $S_n = X_1 + \cdots + X_n$, where the X_i's are i.i.d. random variables. This problem explores if the Law of Large Numbers holds under other circumstances.

   You are spinning a roulette wheel to collect n coupons. On each spin, you have a winning probability of $1 - p$ and the outcomes of different spins are independent. If you win, you get a certain amount of coupons; if you lose, you get nothing.

   For each of the following coupon rewarding schemes, determine whether the Law of Large Numbers holds when S_n is defined as the total number of coupons that you win out of the n coupons. Answer YES if the Law of Large Number holds, or NO if not, and give a brief justification of your answer. (Whenever convenient, you can assume that n is even.)

   (a) YES or NO: You spin the wheel n times. One each spin, if you win, you will get 1 coupon.

   (b) YES or NO: You spin the wheel $\frac{n}{2}$ times. Each time you win, you get 2 coupons.

   (c) YES or NO: You spin the wheel 2 times. Each time you win, you get $\frac{n}{2}$ coupons.

   (d) YES or NO: You spin the wheel once. If you win, you will get n coupons.

6. **Packets Sending** (8 points: 2/6)

   Assume Alice is trying to send m packets across a noisy channel to her friend Bob. The channel has probability $1 - p$ of dropping each packet. To account for the loss, Alice sends $n > m$ packets. Alice can send at most n packets, and she needs to ensure that Bob can receive at least m packets with probability at least r. She wants to figure out how big can m be.

   (a) Modeling each successfully sent packet as a coin toss with probability p, what is the probability that Bob receive at least m packets?

   (b) Assume $n = 100$, $r = 0.9$, and $p = 0.9$. What is the upperbound for m using the Chernoff bound? (use the Chernoff bound: $\Pr[X \leq (1-\delta)\mu] \leq e^{-\frac{\mu\delta^2}{2}}$)

7. **Midterm Review** (10 points)

   Alice and Bob are going to study for the upcoming midterm together. They agree to meet at time t this afternoon. Alice will show up X hours after t, where $X \in \text{Uniform}[0,2]$. Bob's arrival time is more unpredictable. He will be distracted by Pokemon Go and will show up Y hours after t, where $Y \in \text{Expo}(1)$. The person who shows up later is late for T hours. What is $E[T]$? (Hint: some useful integrals $\int x^2 e^{-x}dx = -e^{-x}(x^2+2x+2)+c$ and $\int -xe^{-x}dx = xe^{-x}+e^{-x}+c$)

8. **CIA** (8 points)

   Jason Bourne has been held captive in a prison from which there are three possible routes to escape: an air duct, a sewer pipe and the door (which happens to be unlocked). The air duct leads him on a three hour trip whereupon he falls through a trap door onto his head. The sewer pipe is similar but takes two hours to traverse. Each fall produces amnesia and he is returned to the cell immediately after each fall. Assume that he always immediately chooses one of the three exits from the cell with probability $\frac{1}{3}$. On average, how long does it take before he opens the unlocked door and escapes?

9. **Markov Chain** (12 points: 4/3/5)

   Consider the Markov chain $X(n)$ with the state diagram shown below, where $a, b \in (0, 1)$.

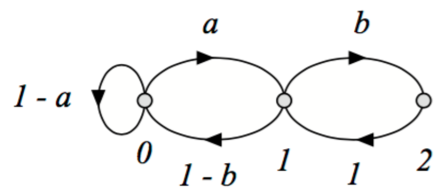

Figure 1: State diagram

   (a) Is this Markov chain irreducible? Is it aperiodic? Briefly justify your answers.

   (b) Calculate $\Pr[X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 1 | X_0 = 0]$.

   (c) Calculate the invariant distribution.