Do not submit

1. **Continuous versus discrete**

   In class, a random variable X is specified by its *distribution* in the discrete case and by its *probability density function (pdf)* in the continuous case. To unify the two cases, we can define the *cumulative distribution function (cdf)* F, which is valid for both discrete and continuous random variables X, as follows:

   $$F(a) = \Pr[X \le a], \qquad a \in \mathbb{R}.$$

   (a) In the discrete case, show that the cdf of a random variable contains exactly the same information as its distribution, by expressing F in terms of the distribution and expressing the distribution in terms of F. For simplicity, you may assume that the discrete random variable only takes on integer values.

   (b) In the continuous case, show that the cdf of a random variable contains exactly the same information as its pdf, by expressing F in terms of the pdf and expressing the pdf in terms of F.

   (c) Identify two key properties that a cdf of any random variable has to satisfy.

2. Discrete and continuous random variables have a lot of similarities but some differences too.

   (a) Suppose X is a discrete random variable. Let $Y = cX$ for some constant c. Express the distribution of Y in terms of the distribution of X.

   (b) Suppose X is a continuous random variable. Let $Y = cX$ for some constant c. Express the pdf of Y in terms of the pdf of X. Is there any difference with the discrete case? (Hint: work with cdf's.)

   (c) If $X = N(\mu, \sigma^2)$, what is the density of $Y = cX$?

3. We begin by proving two very useful properties of the exponential distribution. We then use them to solve a problem in digital photography.

   (a) Let random variable X have exponential distribution with parameter λ. Show that, for any positive s, t, we have

   $$\Pr\left[X > s+t \;\middle|\; X > t\right] = \Pr[X > s].$$

   This is the "memoryless" property of the exponential distribution. We already saw the analogous memoryless property of the geometric distribution in the section 11 worksheet.

   (b) Let random variables X_1, X_2 be independent and exponentially distributed with parameters λ_1, λ_2. Show that the random variable $Y = \min\{X_1, X_2\}$ is exponentially distributed with parameter $\lambda_1 + \lambda_2$. (Hint: work with cdf's.)

(c) You have a digital camera that requires two batteries to operate. You purchase n batteries, labeled $1, 2, \ldots, n$, each of which has a lifetime that is exponentially distributed with parameter λ and is independent of all the other batteries. Initially you install batteries 1 and 2. Each time a battery fails, you replace it with the lowest-numbered unused battery. At the end of this process you will be left with just one working battery. What is the expected total time until the end of the process? Justify your answer.

(d) In the scenario of part c, what is the probability that battery i is the last remaining working battery, as a function of i?

4. Suppose a set of final grades for a course are approximately normally distributed with a mean of 64 and a standard deviation of 7.1. (You are free to use the facts that $\Pr[N(0,1) \leq 1.3] \approx 0.9$ and $\Pr[N(0,1) \leq 1.65] \approx 0.95$.)

   (a) Find the lowest passing grade if the bottom 5% of the students fail the class.

   (b) Find the highest B+ if the top 10% of the students are given A's or A-'s.

5. **Central Limit Theorem**

   Suppose you roll a standard die 2000 times and let X be the sum of the values you get. Using the Central Limit Theorem, for what value of a is $\Pr[X \geq a] \approx \Pr[N(0,1) \geq 2]$? Justify your answer.