GROUND RULES:

- This exam contains 6 questions. The questions are of equal weight.
- Print your name at the top of this page in the upper right hand corner.
- This exam is closed book and closed notes.
- Show all of your work and explain all of your reasoning! **Translation:** No work, no credit. Insufficient explanation, no credit. If you are unsure about whether or not you should explain a result or step in your derivation/proof, then this means you probably should explain it.
- Do not talk with anyone else about this exam. You must work by yourself. No communication of any type with others.

1. A pinball machine has 7 holes through which a ball can drop. Five balls are played and we observe which hole each ball goes down. For example, the first ball could go down hole 1, hole 2, ..., or hole 7 (similarly for the other 4 balls).

(a) Provide a probability model $(S, \mathcal{B}, P)$ for this random experiment. Clearly describe what S, $\mathcal{B}$, and P are; e.g., describe what an outcome $\omega \in S$ "looks like."

(b) On each play, assume the ball is equally likely to go down any one of the 7 holes. Find the probability that more than one ball goes down at least one of the holes.

(c) Let X count the number of balls that go down hole 1. Pick two or three outcomes $\omega \in S$ and calculate $X(\omega)$ for them. Also, find the distribution of X.

2. Suppose $(S, \mathcal{B}, P)$ is a probability space. In the parts that follow, assume A, B, and C are events in $\mathcal{B}$. For parts (b) and (c), assume that $P(A) > 0$, $P(B) > 0$, and $P(C) > 0$. Do not use Venn Diagrams in your proofs.

(a) The *symmetric difference* of A and B, denoted by $A \triangle B$, is the set of all outcomes belonging to A or B but not both events. Prove that

$$P(A \triangle B) = P(A) + P(B) - 2P(A \cap B).$$

(b) Prove: If B and C are independent, then

$$P(A|B) = P(A|B \cap C)P(C) + P(A|B \cap C^c)P(C^c).$$

(c) The converse to (b) is not true. However, if $P(A|B) = P(A|B \cap C)P(C) + P(A|B \cap C^c)P(C^c)$ and $P(A|B \cap C) \neq P(A|B)$, then B and C are independent. Prove this.

3. A random variable X measures the time to failure for a transistor and has cumulative distribution function (cdf)

$$F_X(x) = \begin{cases} 1 - e^{-x^2}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

(a) Verify that $F_X(x)$ is a valid cdf.

(b) Find the probability density function $f_X(x)$.

(c) Find the median of X.

(d) The transistor will be replaced for free if it fails at or before time 1. Otherwise, the cost incurred from the transistor failing is $10(X-1)$. Therefore, the cost of replacement is $Y = 10(X-1)I(X > 1)$. Find the cdf of Y and graph it.

4. A random variable X is said to have a *logistic distribution* if its cumulative distribution function is given by

$$F_X(x) = \frac{1}{1+e^{-x}},$$

for all $x \in \mathbb{R}$.

(a) Show that

$$\ln\left(\frac{F_X(x)}{1-F_X(x)}\right)$$

is a linear function of x. This forms the basis for *logistic regression*.

(b) Show that X and $-X$ have the same distribution.

(c) What is the distribution of $Y = 1/(1+e^{-X})$?

(d) Derive the distribution of $Z = 1 + e^X$. Show that $E(Z)$ does not exist.

5. A discrete random variable X has probability mass function

$$f_X(x) = \begin{cases} \dfrac{c}{x2^x}, & x = 1, 2, 3, \ldots \\ 0, & \text{otherwise,} \end{cases}$$

where the constant $c = (\ln 2)^{-1}$.

(a) Sketch a graph of what the cumulative distribution function $F_X(x)$ looks like. Be precise.

(b) Show that the moment generating function of X is

$$M_X(t) = -c \ln\left(1 - \frac{e^t}{2}\right), \quad \text{for } t < \ln 2.$$

Hint: Write out $h(a) = -\ln(1-a)$ in its Maclaurin series expansion. Where does the condition $t < \ln 2$ come from?

(c) Find $E(X)$ and $\mathrm{var}(X)$.

6. A continuous random variable X has probability density function

$$f_X(x) = \begin{cases} 1 - |1 - x|, & 0 < x < 2 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Show that $f_X(x)$ can be written as follows:

$$f_X(x) = \begin{cases} x, & 0 < x \leq 1 \\ 2 - x, & 1 < x < 2 \\ 0, & \text{otherwise.} \end{cases}$$

Why do you think this density is called "a triangular density?"

(b) Derive $F_X(x)$, the cumulative distribution function of X.

(c) The moment generating function (mgf) for $t \neq 0$ is

$$M_X(t) = \frac{2e^t}{t^2}(\cosh t - 1).$$

How should the mgf be defined at $t = 0$ to make $M_X(t)$ a continuous function of t? *Hint:* Recall from calculus that $\cosh t = (e^t + e^{-t})/2$.

(d) Find $E[(X - 1)^2]$.