Problem 0: Readings
Read notes 3 and 4 from the course web site.

Problem 1: What is random?
Consider 10 independent tosses of a fair coin. These are two possible outcomes:

HHTHTHTHTH
HHHHHHHHHH

These two outcomes have the exact same probability of \((1/2)^{10}\), yet, one seems to better conform to randomness than the other. Is this a dilemma?

(a) Ignore the actual sequencing, i.e. the first outcome consists of 5 heads and 5 tails; and the second outcome consists of 10 heads. Use a binomial random variable to settle the dilemma.

(b) Now consider a geometric random variable \(X\) with parameter \(p = 1/2\). What is \(P(X > 1)\) and \(P(X > 10)\)? Compare the two and explain how they can settle the dilemma.

(c) Mr. Bayes to the rescue: consider the possibility that the coin is biased with a probability \(p > 1/2\) for heads, and that \(P(\text{fair}) = P(\text{biased}) = 1/2\) (zero knowledge). Using a Bayesian approach, compute \(P(\text{fair} | \text{outcome})\) for each of the two outcomes. How does the Bayesian approach settle the dilemma?

Problem 2: Sample mean variance (Optional)
Let \(X\) be a random variable with unknown mean \(\mu\) and unknown variance \(\sigma^2\). Let \(x_1 \ldots x_n\) be \(n\) independent samples. Let \(\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}\). This is called the sample mean. Observe that, by the linearity of expectation, \(E[\bar{x}] = \mu\).

Define:

\[ S_n^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1} \]

Show that \(E[S_n^2] = \sigma^2\). \textit{Hint}: this can be easily computed using the definition and properties of expectation, but it might be messy. You will find useful:
• $E[ax_i] = aE[x_i]$ if $a$ is a constant
• linearity of expectation always holds (regardless of independence)
• $E[x_i x_j] = E[x_i]E[x_j]$ if $x_i$ and $x_j$ are independent, and this can be simply written as $E[x_i]^2$ because all $x_i$’s are identically distributed.

The implication of this mathematical result is that $S_n^2$, and not $\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$, should be taken as an estimate of the sample variance.

PS: This is well known. The purpose of the exercise is to simply manipulate expectations. I will not grade it.

Problem 3: Memoryless
Consider a geometric distribution with parameter $p$. Show the following memoryless property:

$$P(X = a + b | X > a) = (1 - p)^b - 1$$

for $b \in \mathbb{N}$.

Note: This is also a very well known result about the geometric distribution. We will see that this property also holds for the exponential distribution.

Problem 4: Limiting distribution
Let $n \in \mathbb{N}$. Let $r > 1$ and define $X_n$ to be a discrete random variable with

$$P(X_n = i/n) = \frac{r^{i/n} - 1}{r^{1/n} - 1}$$

for $i = 0, 1, 2, \ldots$.

(a) Show that $\sum_i P(X_n = i/n) = 1$.

(b) What kind of probability mass function do we get when we set $n = 1$?

(c) Find an approximation for $P(a \leq X_n \leq b)$ when $n$ is large, so for instance, you can assume that $a$ and $b$ are both integer multiples of $1/n$.

(d) What is the limit of $P(a \leq X_n \leq b)$ when $n$ goes to infinity? Find the density $f(x)$ such that:

$$\lim_{n \to \infty} P(a \leq X_n \leq b) = \int_a^b f(x)dx$$