1. **Packets Over the Internet**

   $n$ packets are sent over the Internet ($n$ even). Consider the following probability models for the packet loss process:

   (a) Each packet is routed over a different path and is lost independently with probability $p$.

   (b) All $n$ packets are routed along the same path, and with probability $p$, one of the links along the path fails and all $n$ packets are lost. Otherwise all packets are received.

   (c) The $n$ packets are divided into 2 groups of $n/2$ packets, and each group is routed along a different path and lost with probability $p$. Losses of different groups are independent events.

   In each of the three models, compute the distribution, mean and variance of the number of packet losses. For $n = 6$ and $p = 0.3$, plot the distribution in each of the three cases. Does the distribution, mean and variance depend on the probability model? Which of the three routing protocols do you prefer?

2. **Coupon Collection**

   Consider the coupon collection problem with $n$ distinct baseball cards and $m$ cereal boxes bought. Compute the distribution, mean and variance of the number of Babe Ruth cards acquired.

3. **DNA Sequencing Again**

   In Q.5 of HW 2, compute the expectation of the number of positions on the genome sequence not covered by any read.

4. **Family Planning**

   Mr and Mrs Brown decide to continue having children until they either have their first girl or until they have five children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let $B$ and $G$ denote the numbers of boys and girls respectively that the Browns have.

   (a) Write down the sample space together with the probability of each outcome.

   (b) Compute and plot the distributions of the random variables $B$ and $G$.

   (c) Compute the mean and variance of $B$ and $G$ using a direct calculation.

5. **Function of a Random Variable**

   Let $X$ be a random variable with the pmf $P_X$, and $Y = g(X)$ be another random variable. Recall that $\mathbb{E}[Y]$ is defined to be $\sum_b b P_Y(b)$, where $P_Y$ is the pmf of $Y$. In this question we will verify the intuitive statement:

   $$
   \mathbb{E}[Y] = \sum_a g(a) P_X(a),
   $$

   i.e. we don’t need to compute the pmf of $Y$ to compute $\mathbb{E}[Y]$. 

---

**Homework 3**

Due Tue, Oct 24
(a) First consider the example where $X$ is uniformly distributed in $\{-5, -4, \ldots, 4, 5\}$.

(i) Show that eq. (1) holds if $Y = 2X$.

(ii) Show that eq. (1) holds if $Y = X^2$.

(Hint: draw the outcomes of the sample space like how we showed $E[X + Y] = E[X] + E[Y]$ in class.)

(b) Give a proof of eq. (1) in the general case.