Homework 4

1. Expectation, expectation, expectation

(a) In class, we gave two definitions for the expectation of a r.v. $X$:

$$E[X] := \sum_{\omega \in \Omega} X(\omega)p(\omega), \quad (1)$$

and

$$E[X] := \sum_a a \cdot p_X(a), \quad (2)$$

where $p_X$ is the p.m.f. of $X$ and the summation is over all $a$ in the range of $X$. Show formally that the two definitions are equivalent.

(b) Now suppose we have another r.v. $Y$ which is a function of the r.v. $X$, i.e. $Y = f(X)$. (For example, $Y = X^2$.) The expectation of $Y$ can be computed in one of several ways. First, one can appeal to definition (1):

$$E[Y] = \sum_{\omega \in \Omega} Y(\omega)p(\omega). \quad (3)$$

Second, one can first compute the p.m.f. $p_Y$ of $Y$ and use the second definition (2):

$$E[Y] = \sum_b b \cdot p_Y(b), \quad (4)$$

where the summation is over all $b$ in the range of $Y$.

Show that there is a third equivalent way of computing the expectation of $Y$:

$$E[Y] = \sum_a f(a) \cdot p_X(a), \quad (5)$$

where the summation is over the range of $X$. (Hint: show that (5) follows from one of (3) or (4).)

(c) Suppose $X$ is uniformly distributed over the integers from $-n$ to $n$, and $Y = X^2$. Compute the expectation of $Y$ using (4) and (5) and check that they give you the same answer.

2. Linearity of Expectations

Solve each of the following problems using linearity of expectation. Clearly explain the steps. (Hint: for each problem, think about what the appropriate random variables should be and define them explicitly.)

(a) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence “bonbon” appears?

(b) A coin with Heads probability $p$ is flipped $n$ times. A “run” is a maximal sequence of consecutive flips that are all the same. (Thus, for example, the sequence HTTHHHTTH with $n = 8$ has five runs.) Show that the expected number of runs is $1 + 2(n - 1)p(1 - p)$. Justify your answer carefully.
3. Packets again
In Q. 4 of Homework 3, we have computed the expectation of the number of packets lost under three different routing protocols, and they turn out to be all the same. Compute and compare the variances. Under which protocol is the variance the largest? the smallest? Which protocol do you prefer? Why?

4. Random homeworks again
In the problem where \( n \) homeworks are randomly returned to \( n \) students, compute the variance of the number of students that gets her own homework back.

5. More Family Planning
Mr and Mrs Brown are really desperate for girls and decide to continue having children until they have their first girl, no matter how long they have to wait. 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. Let \( T \) be the total number of children they have.

(a) List the outcomes of the sample space and assign probabilities to the outcomes.
(b) Write down the distributions of the random variables \( B, G \) and \( T \).
(c) Write down the expectations of \( B, G \) and \( T \).

Suppose instead the Browns decide to have children until they have two girls.

(d) List the outcomes in the new sample space and assign probabilities to the outcomes.
(e) Compute the distribution and the expectation of the total number of boys that the Browns have under this new strategy.

6. Machine Failures
Two faulty machines, \( M_1 \) and \( M_2 \), are repeatedly run synchronously in parallel (i.e., both machines execute one run, then both execute a second run, and so on). On each run, \( M_1 \) fails with probability \( p_1 \) and \( M_2 \) with probability \( p_2 \), all failure events being independent. Let the random variable \( X_1 \) denote the number of runs until the first failure of \( M_1 \), and \( X_2 \) denote the number of runs until the first failure of \( M_2 \).

Let \( X = \min(X_1, X_2) \) denote the number of runs until the first failure of either machine. Compute the distribution of \( X \). What is its expectation?