Homework #5 Solutions

1. Family planning.
   (a) The outcomes are
   \[ \{g, bg, bbg, bbbg, \ldots, bb\cdot \cdot bg, \ldots \} \]
   i.e., all patterns of b and b in which a string of 0 or more boys is followed by a single girl.
   Since the gender of each child is independent and each child is a boy with probability \( \frac{1}{2} \), the probability of any outcome \( bb\cdot \cdot bg \) with \( a \geq 0 \) boys is
   \[ \Pr(a \text{ boys then a girl}) = \Pr(\text{boy})^a \times \Pr(\text{girl}) = \left( \frac{1}{2} \right)^{a+1}. \]

   (b) \[ P(B = i) = 0.5^{i+1} \text{ if } i \geq 0 \text{ and 0 otherwise} \]
   \[ P(G = i) = 1 \text{ if } i = 1 \text{ and 0 otherwise} \]
   \[ P(T = i) = 0.5^i \text{ if } i \geq 1 \text{ and 0 otherwise} \]

   (c) Since \( T \) is a standard geometric distribution with success probability \( \frac{1}{2} \), i.e., \( T \sim \text{Geom}(\frac{1}{2}) \), we know
   \[ E[T] = \frac{1}{\frac{1}{2}} = 2. \]
   Moreover, since \( G \) is always 1, we know
   \[ E[G] = 1. \]
   Finally, we know that \( B + G = T \). Thus, by linearity of expectation we have
   \[ E[T] = E[B] + E[G], \]
   so
   \[ E[B] = E[T] - E[G] = 2 - 1 = 1. \]

   Suppose instead the Browns decide to have children until they have \textit{two} girls.

   (d) The outcomes are
   \[ \{gg, bgg, bgb, bbg, bbgg, bbgb, \ldots, bb\cdot \cdot bbg, \ldots \} \]
   i.e., all patterns of b and g containing two g's and ending with a g. Since the gender of each child is independent and each child is a girl with probability \( \frac{1}{2} \), the probability of an outcome \( bb\cdot \cdot bg\cdot \cdot bg \) with \( a_1 \geq 0 \) boys, then a girl, then \( a_2 \geq 0 \) boys, and finally another girl is
   \[ \Pr(a_1 \text{ boys then a girl, then } a_2 \text{ boys then a girl } ) = \Pr(\text{boy})^{a_1} \times \Pr(\text{girl}) \times \Pr(\text{boy})^{a_2} \times \Pr(\text{girl}) \]
   \[ = \left( \frac{1}{2} \right)^{a_1+a_2+2}. \]
(e) Compute the distribution and the expectation of the total number of boys that the Browns have under this new strategy.

Define \( X \) as above, obtaining the distribution

\[
\Pr(X = i) = \frac{(i + 1)}{2^{i+2}}
\]

via either answer 1 or answer 2. The expectation is given by the formula

\[
E[X] = \sum_{i=1}^{\infty} i \times \Pr(X = i) = \sum_{i=1}^{\infty} \frac{i(i + 1)}{2^{i+2}}.
\]

But how do we evaluate this summation? Well, it gets messy, but let’s do it. Let \( S \) denote this summation. If we write out \( S \) and \( 2S \) as a sum of terms, we find:

\[
2S = \frac{2}{4} + \frac{6}{8} + \frac{12}{16} + \frac{20}{32} + \frac{30}{64} + \cdots
\]

\[
S = \frac{2}{8} + \frac{6}{16} + \frac{12}{32} + \frac{20}{64} + \cdots
\]

With everything aligned so nicely, it seems natural to subtract these two equations. We get

\[
S = \frac{2}{4} + \frac{4}{8} + \frac{6}{16} + \frac{8}{32} + \frac{10}{64} + \cdots
\]

We have found a somewhat simpler expression for \( S \), but it’s not simple enough yet to compute \( S \) directly. So let’s do the same trick again with this summation. We find:

\[
2S = \frac{2}{2} + \frac{4}{4} + \frac{6}{8} + \frac{8}{16} + \frac{10}{32} + \frac{12}{64} + \cdots
\]

\[
S = \frac{2}{4} + \frac{4}{8} + \frac{6}{16} + \frac{8}{32} + \frac{10}{64} + \cdots
\]

Subtracting these two equations, we obtain:

\[
S = \frac{2}{2} + \frac{2}{4} + \frac{2}{8} + \frac{2}{16} + \frac{2}{32} + \frac{2}{64} + \cdots
\]

\[
= 2 \times \left( \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots \right)
\]

\[
= 2 \times 1 = 2.
\]

In other words, \( S = 2. \) But \( S \) was exactly the expectation \( E[X] \) we were trying to compute. In summary,

\[
E[X] = S = 2.
\]

2. **Machine failures.** Let’s compute the probability that neither machine fails, on any particular run. Since failures of the two machines are independent,

\[
\Pr(M_1 \text{ doesn’t fail and } M_2 \text{ doesn’t fail}) = \Pr(M_1 \text{ doesn’t fail}) \times \Pr(M_2 \text{ doesn’t fail}) = (1-p_1)(1-p_2).
\]

Therefore, the probability that at least one machine fails, on any particular run, is

\[
\Pr(\text{either } M_1 \text{ or } M_2 \text{ fails (or both)}) = 1 - (1-p_1)(1-p_2) = p_1 + p_2 - p_1p_2.
\]
We repeatedly perform runs until one of the machines fail. Since failures of both machines at different runs are independent events, the number of runs until one of the machines fail is a geometric distribution with parameter $p_1 + p_2 - p_1p_2$: $X \sim \text{Geom}(p_1 + p_2 - p_1p_2)$. By the formula in Lecture Notes, $E[X] = \frac{1}{p_1 + p_2 - p_1p_2}$.

3. **Packets over the Internet.** Let $X$ be the random variable which denotes the number of packets lost.

(a) Since each of the $n$ packets is independently lost with probability $p$, we know that $X \sim \text{Bin}(n, p)$. Hence, for $k \in \{0, 1, \ldots, n\}$,

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$E[X] = np$$

As can be seen from the probability distribution tables and plots, the distributions depend on the probability models of packet loss, even though the expected number of lost packets is the same in all three models.

Which protocol is most preferable is quite dependent on the context, and it’s possible to construct situations where each of the protocols performs best:

- If the packets were being transmitted without using an error-correcting code, then all of them are needed at the recipient’s end to reconstruct the whole message. In this case, protocol (a) succeeds with probability $(1 - p)^n$, protocol (b) with probability $1 - p$ and protocol (c) with probability $(1 - p)^2$. Therefore, protocol (b), which routes all packets on the same path, would work best in this situation.

- Suppose you used an error-correcting code that can tolerate the loss of up to 35% of packets, and suppose $n = 100$ and $p = 0.3$. Then protocol (a) succeeds with probability $\approx 0.884$, protocol (b) with probability 0.7, and protocol (c) with probability 0.49. Therefore, protocol (a) works best in this situation.

- Suppose you used an error-correcting code that can tolerate the loss of up to 50% of packets, and suppose $n = 3$ and $p = 0.5$. Then protocol (a) and (b) succeed with probability 0.5, while protocol (c) succeeds with probability 0.75. In this case, protocol (c) works best.

- Finally, suppose that each packet that makes it to the recipient is worth the same amount; the total value of the communication is proportional to the number of packets received. Then it might not matter which protocol you choose: in the long run, they will deliver the same value, on average.

Any of these sorts of answers would be acceptable, as long as you give a coherent justification for which protocol you said is preferable.

4. **Coupon collector’s problem.** Suppose $T_i$ is the time needed for getting the $i$-th new coupon.
Then \( T = \sum T_i \).

\[
E[T] = \sum E[T_i] = \sum \frac{1}{p_i} = \sum \frac{n}{n-i} \sim n \log n
\]

5. *Simulating coupon collector's problem.*