Solutions to Problem Set 10

Problem 1. MIT students sometimes delay laundry for a few days. Assume all random values described below are mutually independent.

(a) A ***busy*** student must complete 3 problem sets before doing laundry. Each problem set requires 1 day with probability $2/3$ and 2 days with probability $1/3$. Let B be the number of days a busy student delays laundry. What is $\mathrm{E}[B]$?

Example: If the first problem set requires 1 day and the second and third problem sets each require 2 days, then the student delays for $B = 5$ days.

Solution. The expected time to complete a problem set is:

$$1 \cdot \frac{2}{3} + 2 \cdot \frac{1}{3} = \frac{4}{3}$$

Therefore, the expected time to complete all three problem sets is:

$$\begin{aligned}
\mathrm{E}[B] &= \mathrm{E}[\text{pset1}] + \mathrm{E}[\text{pset2}] + \mathrm{E}[\text{pset3}] \\
&= \frac{4}{3} + \frac{4}{3} + \frac{4}{3} \\
&= 4
\end{aligned}$$

■

(b) A ***relaxed*** student rolls a fair, 6-sided die in the morning. If he rolls a 1, then he does his laundry immediately (with zero days of delay). Otherwise, he delays for one day and repeats the experiment the following morning. Let R be the number of days a relaxed student delays laundry. What is $\mathrm{E}[R]$?

Example: If the student rolls a 2 the first morning, a 5 the second morning, and a 1 the third morning, then he delays for $R = 2$ days.

Solution. If we regard doing laundry as a failure, then the mean time to failure is $1/(1/6) = 6$. However, this counts the day laundry is done, so the number of days delay is $6 - 1 = 5$. Alternatively, we could derive the answer as follows:

$$\begin{aligned}
\mathrm{E}[R] &= \sum_{k=0}^{\infty} \Pr\{R > k\} \\
&= \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^3 + \dots \\
&= \frac{5}{6} \cdot \left(1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \dots\right) \\
&= \frac{5}{6} \cdot \frac{1}{1 - 5/6} \\
&= 5
\end{aligned}$$

■

(c) Before doing laundry, an ***unlucky*** student must recover from illness for a number of days equal to the product of the numbers rolled on two fair, 6-sided dice. Let U be the expected number of days an unlucky student delays laundry. What is $\mathrm{E}[U]$?

Example: If the rolls are 5 and 3, then the student delays for $U = 15$ days.

Solution. Let D_1 and D_2 be the two die rolls. Recall that a die roll has expectation $7/2$. Thus:

$$\begin{aligned}
\mathrm{E}[U] &= \mathrm{E}[D_1 \cdot D_2] \\
&= \mathrm{E}[D_1] \cdot \mathrm{E}[D_2] \\
&= \frac{7}{2} \cdot \frac{7}{2} \\
&= \frac{49}{4}
\end{aligned}$$

■

(d) A student is ***busy*** with probability $1/2$, ***relaxed*** with probability $1/3$, and ***unlucky*** with probability $1/6$. Let D be the number of days the student delays laundry. What is $\mathrm{E}[D]$?

Solution.

$$\mathrm{E}[D] = \frac{1}{2}\,\mathrm{E}[B] + \frac{1}{3}\,\mathrm{E}[R] + \frac{1}{6}\,\mathrm{E}[U]$$

■

Problem 2. There are about 250,000,000 people in the United States who might use a phone. Assume that each person is on the phone during each minute mutually independently with probability $p = 0.01$.

(To keep the problem simple, we are putting aside the fact that people are on the phone more often at certain times of day and on certain days of the year.)

(a) What is the expected number of people on the phone at a given moment?

Solution. Let I_i be an indicator for the event that the i-th person is one the phone. The number of people on the phone is then:

$$\sum_{i=1}^{250{,}000{,}000} I_i.$$

The expectation of this sum is $250,000,000 \cdot 0.01 = 2,500,00$ by linearity of expectation. ∎

(b) Suppose that we construct a phone network whose capacity is a mere **one percent** above the expectation. Upper bound the probability that the network is overloaded in a given minute. (Use the approximation formula given in the notes. You may need to evaluate this expression in a clever way because of the size of numbers involved. For example, you could first evaluate the *logarithm* of the given expression.)

Solution. The network is overloaded if the fraction of people calling is greater than $1.01 \cdot 0.01 = 0.0101$. In complementary terms, the network is overloaded if the fraction of people *not* calling is less than $1 - 0.0101 = 0.9899$. Define the following variables:

$$\begin{aligned} n &::= 250,000,000 \quad \text{people in the US} \\ p &::= 0.99 \quad \text{prob. not on phone} \\ \alpha &::= 0.9899 \quad \text{min. allowable fraction not on phone} \end{aligned}$$

In these terms, the solution to the problem is $F_{n,p}(\alpha n)$. We can upper bound this approximately using the formula from the notes:

$$F_{n,p}(\alpha n) \approx \frac{1-\alpha}{1-\alpha/p} \cdot \frac{2^{nH(\alpha)}}{\sqrt{2\pi\alpha(1-\alpha)n}} \cdot p^{\alpha n}(1-p)^{(1-\alpha)n}.$$

By first evaluating the logarithm of this expression, we find that this is about e^{-120}. ∎

(c) What is the expected number of minutes (approximately) until the system is overloaded for the first time?

Solution. Applying the "expected time to failure" formula with probability $p = e^{-120}$ gives $1/p = e^{120}$. ∎

Problem 3. We are given a set of n distinct positive integers. We then determine the maximum of these numbers by the following procedure:

Randomly arrange the numbers in a sequence.

Let the "current maximum" initially be the first number in the sequence and the "current element" be the second element of the sequence. If the current element is greater than the current maximum, perform an "update": that is, change the current maximum to be the current element. Either way, change the current element to be the next element of the sequence. Repeat this process until there is no next element.

Prove that the expected number of updates is $\sim \ln n$.

Hint: Let M_i be the indicator variable for the event that the ith element of the sequence is bigger than all the previous elements in the sequence.

Solution. Note that the number of times we update the current maximum is precisely $M = M_1 + \ldots + M_n$. Since expectation is a linear operator, we can compute $\mathrm{E}[M]$ by finding $E(M_i)$ for all i and summing them up. Note also that since M_i is an "indicator" variable, we only have to find $\Pr\{M_i = 1\}$. In a random permutation, this happens with probability $1/i$. Why? Because all permutations of the first i numbers in the sequence are equally likely, and the largest among them occurs as the last element of the permutation in $1/i$ of the cases. Thus

$$\begin{aligned} \mathrm{E}[M] &= \sum_i \Pr\{M_i = 1\} \\ &= \sum_{i=1}^{n} 1/i \\ &= H_n \sim \ln n, \end{aligned}$$

where H_n is the nth Harmonic number. ■

Problem 4. In a certain card game, each card has a point value.

- Numbered cards in the range 2 to 9 are worth five points each.
- The card numbered 10 and the face cards (jack, queen, king) are worth ten points each.
- Aces are worth fifteen points each.

(a) Suppose that you thoroughly shuffle a 52-card deck. What is the expected total point value of the three cards on the top of the deck after the shuffle?

Solution. Let the random variable, X, be the total point value of the three cards on the top of the deck. Then we can write:

$$X = X_1 + X_2 + X_3$$

where the random variables X_1, X_2, and X_3 are the point values of the first, second, and third cards. By the definition of expectation:

$$\begin{aligned}
\mathrm{E}[X_i] &= \sum_{r \in X_i(S)} r \cdot \Pr\{X_i = r\} \\
&= 5 \cdot \frac{8}{13} + 10 \cdot \frac{4}{13} + 15 \cdot \frac{1}{13} \\
&= \frac{95}{13}
\end{aligned}$$

Now we can solve the problem by taking the expected value of both sides of our original equation and then using linearity of expectation:

$$\begin{aligned}
\mathrm{E}[X] &= \mathrm{E}[X_1 + X_2 + X_3] \\
&= \mathrm{E}[X_1] + \mathrm{E}[X_2] + \mathrm{E}[X_3] \\
&= \frac{95}{13} + \frac{95}{13} + \frac{95}{13} \\
&= \frac{285}{13}
\end{aligned}$$

■

(b) Suppose that you throw out all the red cards and shuffle the remaining 26-card, all-black deck. Now what is the expected total point value of the top three cards? (Note that drawing three aces, for example, is now impossible!)

Solution. The expected point value is the same as before, since expected point value of a single card is unchanged. Nothing in our solution assumed a 52 card deck. ■

Problem 5. A true story from World War II:

The army needs to identify soldiers with a disease called "klep". There is a way to test blood to determine whether it came from someone with klep. The straightforward approach is to test each soldier individually. This requires n tests, where n is the number of soldiers. A better approach is the following: group the soldiers into groups of k. Blend the blood samples of each group and apply the test once to each blended sample. If the group-blend doesn't have klep, we are done with that group after one test. If the group-blend fails the test, then someone in the group has klep, and we individually test all the soldiers in the group.

Assume each soldier has klep with probability, p, independently of all the other soldiers.

(a) What is the expected number of tests as a function of n, p, and k? (Assume for simplicity that n is divisible by k.)

Solution. There are n/k groups of size k each. Let X_i be a random variable that denotes the number of tests performed in group i. X_i takes value 1 with probability $(1-p)^k$ and value $k+1$ with probability $1-(1-p)^k$. Hence the expected number of tests is

$$\mathrm{E}\left[\sum_{i=1}^{n/k} X_i\right] = \sum_{i=1}^{n/k} E(X_i) = (\frac{n}{k})((1-p)^k + (k+1)(1-(1-p)^k)) = n(1-(1-p)^k + \frac{1}{k}). \quad (1)$$

■

(b) How should k be chosen to minimize the expected number of test performed, and what is the resulting expectation?

Solution. The k must be chosen so that the derivative w.r.t. k of the answer from part (a) is 0, namely,

$$(1-p)^k \ln(1-p) + \frac{1}{k^2} = 0.$$

Assuming that p is much smaller than $1/k$, we can approximate $(1-p)^k$ by 1 and $\ln(1-p)$ by $-p$, giving us

$$k \approx \sqrt{\frac{1}{p}}.$$

In particular, $p \approx 1/k^2$ comes out much smaller than $1/k$, so our approximations are justified. The resulting expectation is approximately $n\sqrt{p}$. ■

(c) What fraction of the work does the grouping method expect to save over the straightforward approach in a million-strong army where 1% have klep?

Solution. Using the approximation from the previous part, the expected fraction of work saved is $1-\sqrt{p}$, so for $p = 0.01$, we estimate a 90% savings. Using the exact formula (1), we find that the fraction of work saved is $(1-p)^k - 1/k$. So for $p = 0.01$ and $k = \sqrt{1/p} = 10$, the savings is $(1-0.01)^{10} - \sqrt{0.01} = 0.804$, that is, more than 80%. ■

Problem 6. The hat-check staff has had a long day, and at the end of the party they decide to return people's hats at random. Suppose that n people have their hats returned at random. We previously showed that the expected number of people who get their own hat back is 1, irrespective of the total number of people. In this problem we will calculate the variance in the number of people who get their hat back.

Let $X_i = 1$ if the ith person gets his or her own hat back and 0 otherwise. Let $S_n ::= \sum_{i=1}^{n} X_i$, so S_n is the total number of people who get their own hat back. Show that

(a) $\mathrm{E}\left[X_i^2\right] = 1/n$.

Solution. $X_i = 1$ with probability $1/n$ and 0 otherwise. Thus $X_i^2 = 1$ with probability $1/n$ and 0 otherwise. So $\mathrm{E}\left[X_i^2\right] = 1/n$. ■

(b) $\mathrm{E}\left[X_i X_j\right] = 1/n(n-1)$ for $i \neq j$.

Solution. The probability that X_i and X_j are both 1 is $1/n \cdot 1/(n-1) = 1/n(n-1)$. Thus $X_i X_j = 1$ with probability $1/n(n-1)$, and is zero otherwise. So $\mathrm{E}\left[X_i X_j\right] = 1/n(n-1)$. ■

(c) $\mathrm{E}\left[S_n^2\right] = 2$. *Hint:* Use (a) and (b).

Solution.

$$
\begin{aligned}
\mathrm{E}\left[S_n^2\right] &= \sum_i \mathrm{E}\left[X_i^2\right] + \sum_i \sum_{j \neq i} \mathrm{E}\left[X_i X_j\right] \\
&= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} \\
&= 2.
\end{aligned}
$$

■

(d) $\mathrm{Var}\left[S_n\right] = 1$.

Solution.

$$
\begin{aligned}
\mathrm{Var}\left[S_n\right] &= \mathrm{E}\left[S_n^2\right] - \mathrm{E}^2\left[S_n\right] \\
&= 2 - (n(1/n))^2 \\
&= 2 - 1 \\
&= 1.
\end{aligned}
$$

■

(e) Explain why you cannot use the variance of sums formula to calculate $\text{Var}[S_n]$.

Solution. The indicator random variables, X_i, are not even pairwise independent. This can be seen by comparing the marginal and conditional probability of a particular person, Alice, getting her hat back. The marginal probability, unconditioned on any other events, is $1/n$ as we've computed before. However, if compute this probability conditioned on the event that a second person, Bob, got his hat back, we find that the probability of Alice getting her hat back is $1/(n-1)$. ■

(f) Using Chebyshev's Inequality, show that $\Pr\{S_n \geq 11\} \leq .01$ for any $n \geq 11$.

Solution.

$$\begin{aligned}
\Pr\{S_n \geq 11\} &= \Pr\{S_n - \text{E}[S_n] \geq 11 - \text{E}[S_n]\} \\
&= \Pr\{S_n - \text{E}[S_n] \geq 10\} \\
&\leq \frac{\text{Var}[S_n]}{10^2} = .01
\end{aligned}$$

Note that the X_i's are Bernoulli variables but are *not* independent, so S_n does not have a binomial distribution and the binomial estimates from Lecture Notes do not apply. ■

Problem 7. Let R_1 and R_2 be independent random variables, and f_1 and f_2 be any functions such that $\text{domain}(f_i) = \text{codomain}(R_i)$ for $i = 1, 2$. Prove that $f_1(R_1)$ and $f_2(R_2)$ are independent random variables.

Solution. The event $[f_i(R_i) = a]$ is the disjoint union of all the events $[R_i = r]$ for $r \in f_i^{-1}(a)$, so

$$\Pr\{f_i(R_i) = a\} = \sum_{r \in f_i^{-1}(a)} \Pr\{R_i = r\}.$$

Also, the event $[f_1(R_1) = a \text{ and } f_2(R_2) = b]$ is the disjoint unon of the events $[R_1 =$

r and $R_2 = t]$ for $(r,t) \in f_1^{-1}(a) \times f_2^{-1}(b)$. Hence,

$$\begin{aligned}
&\Pr\{f_1(R_1) = a \text{ and } f_2(R_2) = b\} \\
&= \sum_{(r,t)\in f_1^{-1}(a)\times f_2^{-1}(b)} \Pr\{R_1 = r \text{ and } R_2 = t\} \\
&= \sum_{(r,t)\in f_1^{-1}(a)\times f_2^{-1}(b)} \Pr\{R_1 = r\}\Pr\{R_2 = t\} && \text{[because } R, R_2 \text{ independent]} \\
&= \sum_{t\in f_2^{-1}(b)} \sum_{r\in f_1^{-1}(a)} \Pr\{R_1 = r\}\Pr\{R_2 = t\} \\
&= \sum_{t\in f_2^{-1}(b)} \Pr\{R_2 = t\} \left(\sum_{r\in f_1^{-1}(a)} \Pr\{R_1 = r\} \right) \\
&= \sum_{t\in f_2^{-1}(b)} \Pr\{R_2 = t\} (\Pr\{f(R) = a\}) \\
&= \Pr\{f(R) = a\} \sum_{t\in f_2^{-1}(b)} \Pr\{R_2 = t\} \\
&= \Pr\{f(R) = a\}\Pr\{f_2(R_2) = b\}
\end{aligned}$$

■

Problem 8. Let A, B, C be events, and let I_A, I_B, I_C be the corresponding indicator variables. Prove that A, B, C are mutually independent iff the random variables I_A, I_B, I_C are mutually independent.

Solution. ($\longrightarrow$):

$$\begin{aligned}
\Pr\{I_A = 1 \wedge I_B = 0 \wedge I_C = 1\} &= \Pr\left\{A \cap \overline{B} \cap C\right\} \\
&= \Pr\{A \cap C\} - \Pr\{A \cap B \cap C\} \\
&= \Pr\{A\} \cdot \Pr\{C\} - \Pr\{A\} \cdot \Pr\{B\} \cdot \Pr\{C\} \\
&= \Pr\{A\} \cdot \Pr\{C\} (1 - \Pr\{B\}) \\
&= \Pr\{A\} \cdot \Pr\{C\} \Pr\left\{\overline{B}\right\} \\
&= \Pr\{I_A = 1\} \cdot \Pr\{I_C = 1\} \Pr\{I_B = 0\} \\
&= \Pr\{I_A = 1\} \cdot \Pr\{I_B = 0\} \cdot \Pr\{I_C = 1\}
\end{aligned}$$

and similarly for any other three binary values in place of 101.

($\longleftarrow$):

$$
\begin{aligned}
\Pr\{A \cap B \cap C\} &= \Pr\{I_A = 1 \wedge I_B = 1 \wedge I_C = 1\} \\
&= \Pr\{I_A = 1\} \cdot \Pr\{I_B = 1\} \cdot \Pr\{I_C = 1\} \\
&= \Pr\{A\} \cdot \Pr\{B\} \cdot \Pr\{C\}
\end{aligned}
$$

Also,

$$
\begin{aligned}
\Pr\{A \cap B\} &= \Pr\{A \cap B \cap C\} + \Pr\left\{A \cap B \cap \overline{C}\right\} \\
&= \Pr\{I_A = 1 \wedge I_B = 1 \wedge I_C = 1\} + \Pr\{I_A = 1 \wedge I_B = 1 \wedge I_C = 0\} \\
&= \Pr\{I_A = 1\} \cdot \Pr\{I_B = 1\} \cdot \Pr\{I_C = 1\} + \Pr\{I_A = 1\} \cdot \Pr\{I_B = 1\} \cdot \Pr\{I_C = 0\} \\
&= \Pr\{A\} \cdot \Pr\{B\} \cdot \Pr\{C\} + \Pr\{A\} \cdot \Pr\{B\} \cdot \Pr\left\{\overline{C}\right\} \\
&= \Pr\{A\} \cdot \Pr\{B\} \left(\Pr\{C\} + \Pr\left\{\overline{C}\right\}\right) \\
&= \Pr\{A\} \cdot \Pr\{B\}
\end{aligned}
$$

and similarly for $B \cap C$ and $A \cap C$. ■