Problem Set 10 Solutions

Due: Monday, May 2 at 9 PM

Problem 1. Justify your answers to the following questions about independence.

(a) Suppose that you roll a fair die that has six sides, numbered 1, 2, $\ldots$, 6. Is the event that the number on top is a multiple of 2 independent of the event that the number on top is a multiple of 3?

Solution. Let A be the event that the number on top is a multiple of 2, and let B be the event that the number on top is a multiple of 3. We have:

$$\Pr(A) \cdot \Pr(B) = \frac{3}{6} \cdot \frac{2}{6} = \frac{1}{6} = \Pr(A \cap B)$$

Therefore, these events are independent.

(b) Now suppose that you roll a fair die that has *four* sides, numbered 1, 2, 3, 4. Is the event that the number on top is a multiple of 2 independent of the event that the number on top is a multiple of 3?

Solution. As before, let A be the event that the number on top is a multiple of 2, and let B be the event that the number on top is a multiple of 3. Now, however, we have:

$$\Pr(A) \cdot \Pr(B) = \frac{2}{4} \cdot \frac{1}{4} = \frac{1}{8}$$

But:

$$\Pr(A \cap B) = 0$$

Since these results disagree, the events are not independent.

(c) Now suppose that you roll a fair die that has *eight* sides, numbered $1, 2, \ldots, 8$. Again, is the event that the number on top is a multiple of 2 independent of the event that the number on top is a multiple of 3?

Solution. As before, let A be the event that the number on top is a multiple of 2, and let B be the event that the number on top is a multiple of 3. This time, we have:

$$\Pr(A) \cdot \Pr(B) = \frac{4}{8} \cdot \frac{2}{8} = \frac{1}{8}$$

And:

$$\Pr(A \cap B) = 1/8$$

Therefore, these events *are* independent.

(d) Finally, suppose that you roll the fair, eight-sided die again. Let the random variable X be the remainder when the number on top is divided by 2, and let the random variable Y be the remainder when the number on top is divided by 3. Are the random variables X and Y independent?

Solution. First, let's tabulate the values of X and Y:

die roll	X	Y
1	1	1
2	0	2
3	1	0
4	0	1
5	1	2
6	0	0
7	1	1
8	0	2

Working from the table, we have:

$$\Pr(X = 1 \cap Y = 1) = \frac{2}{8}$$

But:

$$\begin{aligned}\Pr(X = 1) \cdot \Pr(Y = 1) &= \frac{4}{8} \cdot \frac{3}{8} \\ &= \frac{3}{16}\end{aligned}$$

Since these results conflict, the random variables are not independent.

Problem 2. Philo T. Megabrain, a noted parapsychology researcher, has discovered an amazing phenomenon! He puts a psychic on each side of an opaque, soundproof barrier. Each psychic rolls a fair die, looks at it, and tries to guess what number came up on the *other* die by telepathy. Since the dice are fair and independent, the psychics should guess correctly only 1 time in 6. However, after extensive testing, Philo has discovered that they actually do *slightly* better.

(a) Philo's somewhat-arbitrary policy is to run the test over and over each day until both psychics roll a 6 at the same time. Then he immediately halts testing for the day, before the psychics make guesses. Explain the flaw in Philo's experiment in qualitative terms.

Solution. If a psychic sees a 6 on her own die, she knows not to guess that the other die is a 6.

(b) If a psychic exploits this flaw optimally, with what probabilty can she guess the number on the opposite die?

Solution. If she sees a 1, 2, 3, 4, or 5, then her probability of guessing the other die is the normal $1/6$. However, if she sees a 6, then she knows that the other die is *not* a 6, and so her probability of guessing the other die is $1/5$. By the total probability law, her probability of guessing the other die correctly in general is:

$$\frac{5}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{5} = \frac{31}{180}$$

Problem 3. There is a set P consisting of 1000 people.

- The favorite color of 20% of the people is blue.
- The favorite color of 30% is green.
- The favorite color of 50% is red.

(a) Suppose we select a set of two people $\{p_1, p_2\} \subseteq P$ uniformly at random. Let the random variables C_1 and C_2 denote their favorite colors. Are C_1 and C_2 independent? Justify your answer.

Solution. No. For example, $\Pr(C_1 = \text{blue}) = 200/1000$. However,

$$\Pr(C_2 = \text{blue} \mid C_1 = \text{blue}) = 199/999$$

since 199 of the remaining 999 people like blue after one person who likes blue is selected.

(b) Suppose we select a sequence of two people $(p_1, p_2) \in P \times P$ uniformly at random. Let the random variables C_1 and C_2 denote their favorite colors. Now are C_1 and C_2 independent? Justify your answer.

Solution. Yes. Let $c(n)$ be the color that the n-th person likes. The random variables p_1 and p_2 are independent. Functions of independent random variables are independent, so $C_1 = c(p_1)$ and $C_2 = c(p_2)$ are independent.

Problem 4. Secret documents are disappearing from CIA headquarters. Some documents are simply misplaced. But the Security Chief suspects that others are begin stolen by Agent X and passed to the government of Liechtenstein to further its relentless pursuit of global domination. Two inspectors are assigned to investigate the matter:

- Inspector AM determines that the event that a document disappears during a given day is independent of the event that Agent X is in headquarters that day.
- Similarly, inspector PM determines that the event that a document disappears during a given night is independent of the event that Agent X is around that night.

The Security Chief concludes that the event that a document disappears is independent of the event that Agent X is present. Therefore, Agent X is probably innocent.

(a) Construct a probability model of the situation. State the inspectors' determinations and the Security Chief's conclusion as probabilities.

Solution. Let the sample space S be a set of days and nights. Define the following three events:

$$\begin{aligned} D &= \text{A secret document disappears} \\ X &= \text{Agent X is at headquarters} \\ A &= \text{It is daytime.} \end{aligned}$$

In these terms, Inspector AM says:

$$\Pr(D \cap X \mid A) = \Pr(D \mid A) \cdot \Pr(X \mid A)$$

Inspect PM says:

$$\Pr\left(D \cap X \mid \overline{A}\right) = \Pr\left(D \mid \overline{A}\right) \cdot \Pr\left(X \mid \overline{A}\right)$$

And the Security Chief concludes:

$$\Pr(D \cap X) = \Pr(D) \cdot \Pr(X)$$

(b) Is the Security Chief's reasoning correct? Justify your answer.

Solution. The security chief is wrong. For example, suppose that S consists of a single day and a single night:

$$S = \{\text{day}, \text{night}\}$$

Assign night and day each probability $1/2$. Now suppose that Agent X is around during the night and a document disappears only at night:

$$\begin{aligned} D &= \{\text{night}\} \\ X &= \{\text{night}\} \\ A &= \{\text{day}\} \end{aligned}$$

Furthermore, suppose $\Pr(\text{day}) = \Pr(\text{night}) = 1/2$. These suppositions are consistent with the inspectors' determinations:

$$\begin{aligned} \Pr(D \cap X \mid A) &= \frac{\Pr(D \cap X \cap A)}{\Pr(A)} = 0 \\ \Pr(D \mid A) \cdot \Pr(X \mid A) &= \frac{\Pr(D \cap A)}{\Pr(A)} \cdot \frac{\Pr(X \cap A)}{\Pr(A)} = 0 \\ \Pr\left(D \cap X \mid \overline{A}\right) &= \frac{\Pr\left(D \cap X \cap \overline{A}\right)}{\Pr\left(\overline{A}\right)} = 1 \\ \Pr\left(D \mid \overline{A}\right) \cdot \Pr\left(X \mid \overline{A}\right) &= \frac{\Pr\left(D \cap \overline{A}\right)}{\Pr\left(\overline{A}\right)} \cdot \frac{\Pr\left(X \cap \overline{A}\right)}{\Pr\left(\overline{A}\right)} = 1 \end{aligned}$$

However, the Security Chief's conclusion is wrong, because:

$$\Pr(D \cap X) = \Pr(\text{night}) = 1/2$$

But:

$$\Pr(D) \cdot \Pr(X) = (1/2) \cdot (1/2) = 1/4$$

So Agent X may be guilty after all!

Problem 5. Suppose you flip n fair, independent coins. Let the random variable X be the number of heads that come up.

(a) What is the exact value of $\Pr(X \leq k)$, the probability of flipping k or fewer heads? Your answer need not be in closed form.

Solution.

$$\frac{\binom{n}{k} + \binom{n}{k-1} + \ldots + \binom{n}{0}}{2^n}$$

(b) Suppose $k < n/2$. Prove that:

$$\Pr(X \leq k) \leq \frac{n-k+1}{n-2k+1} \cdot \Pr(X = k)$$

(Upper bound your previous answer with an infinite geometric sum and then evaluate the sum.)

Solution. We can upper bound the numerator in the preceding answer as follows:

$$\begin{aligned}
&\binom{n}{k} + \binom{n}{k-1} + \ldots + \binom{n}{0} \\
&\quad = \binom{n}{k} + \tfrac{k}{n-k+1}\binom{n}{k} + \tfrac{k(k-1)}{(n-k+1)(n-k+2)}\binom{n}{k} + \tfrac{k(k-1)(k-2)}{(n-k+1)(n-k+2)(n-k+3)}\binom{n}{k} + \ldots \\
&\quad \leq \binom{n}{k} \cdot \left(1 + \frac{k}{n-k+1} + \frac{k^2}{(n-k+1)^2} + \frac{k^3}{(n-k+1)^3} + \ldots\right) \\
&\quad = \binom{n}{k} \cdot \frac{1}{1 - \frac{k}{n-k+1}} \\
&\quad = \binom{n}{k} \cdot \frac{n-k+1}{n-2k+1}
\end{aligned}$$

(Note that the geometric sum converges only if $k < n/2$.) Therefore:

$$\Pr(X \leq k) \leq \frac{n-k+1}{n-2k+1} \cdot \Pr(X = k)$$

(c) If you flip a coin 100 times, the probability of flipping exactly 30 heads is approximately 23 out of a million. Give an upper bound on the probability of flipping *30 or fewer* heads.

Solution. Applying the bound above gives:

$$\left(23 \cdot 10^{-6}\right) \cdot \frac{100 - 30 + 1}{100 - 2 \cdot 30 + 1} \approx 40 \cdot 10^{-6}$$

The actual value is about $39.25 \cdot 10^{-6}$.

Problem 6. Many of the best computer algorithms rely on randomization. However, generating uniform, mutually independent random bits is not so easy! (The mathematician John von Neumann said, "Anyone who considers arithmetic methods of producing random digits is, of course, in a state of sin.") Fortunately, some algorithms work equally well with *pairwise-independent* random bits, which are relatively "cheap". In particular, one can covert a set of mutually independent bits into an *exponentially larger* set of pairwise-independent random bits.

Let B be a set of n uniform, mutually-independent 0-1 random variables.

(a) Let S be a nonempty subset of the bits in B. Let the random variable s be the XOR of all the bits in S. Show that s is uniformly distributed on $\{0, 1\}$.

(Hint: Let b be one of the bits in S and let s' be the XOR of all other bits in S.)

Solution.

$$\begin{aligned}
\Pr(s = 0) &= \Pr(s' = 0 \cap b = 0) + \Pr(s' = 1 \cap b = 1) \\
&= \Pr(s' = 0)\Pr(b = 0) + \Pr(s' = 1)\Pr(b = 1) \\
&= \frac{1}{2}\Pr(s' = 0) + \frac{1}{2}\Pr(s' = 1) \\
&= \frac{1}{2}(\Pr(s' = 0) + \Pr(s' = 1)) \\
&= \frac{1}{2}
\end{aligned}$$

We first rewrite the event $s = 0$ and then use the independence of b and s'. The remaining steps use the facts that b is 0 or 1 with equal probability and that s' is either 0 or 1 (with unknown probabilities). Since $s = 0$ with probability $1/2$, we must have $s = 1$ with probability $1/2$ as well, so s is uniformly distributed on $\{0, 1\}$.

(b) Now let T be another nonempty subset of bits in B. Let the random variable t be the XOR of all the bits in T. Show that s and t are independent.

(Hint: Define s' to be the XOR of bits in $S - T$, t' to be the XOR of bits in $T - S$, and i to be the XOR of bits in $S \cap T$. Now consider three cases: (1) $S \cap T = \emptyset$, (2) $S \cap T = S$ or $S \cap T = T$, and (3) $S \cap T \neq \emptyset, S,$ or T.)

Solution. We must show that $\Pr(s = a \cap t = b) = \Pr(s = a)\Pr(t = b)$ for all a and b. By the preceding problem part, $\Pr(s = a) = \Pr(t = b) = 1/2$. So we really only need to show that for all a and b:

$$\Pr(s = a \cap t = b) = 1/4$$

Define random variables s', t', and i as described above. These random variables are mutually independent since they are functions of mutually independent bits. We can rewrite the quantity we're trying to analyze, $\Pr(s = a \cap t = b)$, in terms of these variables as follows:

$$\begin{aligned}\Pr(s = a \cap t = b) &= \Pr(s' = a \cap t' = b \cap i = 0) \\ &\quad + \Pr\left(s' = \overline{a} \cap t' = \overline{b} \cap i = 1\right) \\ &= \Pr(s' = a)\Pr(t' = b)\Pr(i = 0) \\ &\quad + \Pr(s' = \overline{a})\Pr\left(t' = \overline{b}\right)\Pr(i = 1) \qquad (*)\end{aligned}$$

Now we analyze the three cases:

1. If $S \cap T = \emptyset$, then $\Pr(i = 0) = 1$ and $\Pr(i = 1) = 0$. However, the sets $S - T$ and $T - S$ are nonempty, so $\Pr(s' = a) = \Pr(t' = b) = 1/2$ by the preceding part. Substituting into (*) gives:

$$\Pr(s = a \cap t = b) = \frac{1}{2} \cdot \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} \cdot 0 = \frac{1}{4}$$

2. If $S \cap T = S$, then $S - T = \emptyset$ and so $\Pr(s' = 0) = 1$ and $\Pr(s' = 1) = 0$. The sets $S \cap T$ and $T - S$ are nonempty, so i and t' are uniformly distributed by the preceding part. Substituting into (*) gives:

$$\Pr(s = a \cap t = b) = 0 \cdot \frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

   If $S \cap T = T$, then a symmetric argument applies.

3. If $S \cap T \neq \emptyset, S,$ or T, then the sets $S - T$, $T - S$, and $S \cap T$ are all nonempty. Therefore, s', t', and i are all uniformly-distributed. Substituting into (*) gives:

$$\Pr(s = a \cap t = b) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Therefore, s and t are independent.

(c) Explain how to construct a set of $2^n - 1$ uniform, pairwise-independent 0-1 random variables from a set of n uniform, mutually-independent 0-1 random variables.

Solution. Take the sums of all nonempty subsets modulo 2. In the two preceding parts, we proved that these random variables are uniform and pairwise independent.

(The quantity a_1 XOR a_2 XOR $\ldots$ XOR a_n is equal to $(a_1 + a_2 + \ldots + a_n) \text{ rem } 2$.)