# Solutions to the Final Examination

**Problem 1 (20 points).** Professor Plum, Mr. Green, and Miss Scarlet are all plotting to shoot Colonel Mustard. If one of these three has both an *opporunity* and the *revolver*, then that person shoots Colonel Mustard. Otherwise, Colonel Mustard escapes.

**Exactly one** of the three has an *opportunity* with the following probabilities:

$$
\begin{aligned}
\Pr\{\text{Plum has opportunity}\} &= 1/6 \\
\Pr\{\text{Green has opportunity}\} &= 2/6 \\
\Pr\{\text{Scarlet has opportunity}\} &= 3/6
\end{aligned}
$$

**Exactly one** has the *revolver* with the following probabilities, regardless of who has an opporuntity:

$$
\begin{aligned}
\Pr\{\text{Plum has revolver}\} &= 4/8 \\
\Pr\{\text{Green has revolver}\} &= 3/8 \\
\Pr\{\text{Scarlet has revolver}\} &= 1/8
\end{aligned}
$$

**(a) (5 points)** Draw a tree diagram for this problem. Indicate edge and outcome probabilities.

**Solution.** ∎

**(b) (5 points)** What is the probability that Colonel Mustard is shot?

**Solution.** $13/48$ ∎

**(c) (5 points)** What is the probability that Colonel Mustard is shot, given that Scarlet does not have the revolver?

**Solution.** $5/21$ ■

**(d) (5 points)** What is the probability that Mr. Green had an opportunity, given that Colonel Mustard was shot?

**Solution.** $6/13$ ■

**Problem 2 (20 points).** For the following problems, you do not need to simplify your answers.

*Note:* The grading policy for the following four problems was:

$5 =$ all correct
$4 =$ minor error, e.g. arithmetic
$3 =$ significant error, e.g. one term wrong
$2 =$ more wrong than right
$1 =$ something vaguely relevant
$0 =$ all wrong

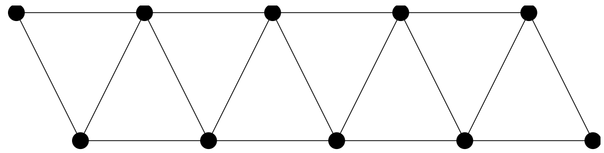
**(a) (5 points)** Next Christmas, the Grumpersons plan to divide among their three children not only 11 pieces of coal, but also 13 pieces of scrap metal and 3 shards of glass. In how many different ways can this be done?

**Solution.** By the product rule, the number of ways to divide up all the presents is equal to the number of ways to divide the coal times the number of ways to divide the scrap metal times the number of ways to divide the glass shard.

$$\binom{13}{2}\binom{15}{2}\binom{5}{2}$$

■

**(b) (5 points)** Each vertex of the graph below must be colored red, orange, yellow, or green such that adjacent vertices are colored differently. In how many different ways can this be done?



**Solution.** There are 4 ways to color the leftmost vertex, 3 ways to color the next vertex, and 2 ways to color each of the remaining vertices. By the general product rule, the total number of colorings is:

$$4 \cdot 3 \cdot 2^8$$

■

**(c) (5 points)** How many different sequences of natural numbers $(x_1, \ldots, x_7)$ satisfy *both* of the following equations:

$$x_1 + x_2 + x_3 = 50$$
$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 80$$

**Solution.** Subtracting the first equation from the second gives:

$$x_4 + x_5 + x_6 + x_7 = 30$$

The total number of solutions is equal to the number of ways to satisfy the first equation multiplies by the number of ways to satisfy this new equation, which is:

$$\binom{52}{2} \cdot \binom{33}{3}$$

■

**(d) (5 points)** Each day, an MIT student selects a breakfast from among 8 possibilities:

cereal, fruit, pancakes, ..., Doritos

The student selects lunch from among 10 possibilities:

sandwich from a truck, pizza from a truck, falafel from a truck, ..., Doritos

The student selects dinner from among 7 possibilities:

pasta, hamburger, tacos, pizza, ..., Doritos

In how many different ways can the student select a breakfast, a lunch, and a dinner if at most one meal can be Doritos?

**Solution.** There are $1 \cdot 9 \cdot 6$ menus with Doritos only for breakfast, $7 \cdot 1 \cdot 6$ menus with Doritos only for lunch, $7 \cdot 9 \cdot 1$ menus with Doritos only for dinner, and $7 \cdot 9 \cdot 6$ menus with no Doritos at all. Thus, the number of different menus with Doritos in at most one meal is:

$$9 \cdot 6 + 7 \cdot 6 + 7 \cdot 9 + 7 \cdot 9 \cdot 6$$

■

**Problem 3 (20 points).** Consider the following equation:

$$\binom{2n}{n+1} = \sum_{k=0}^{n-1} \binom{n}{k} \binom{n}{k+1} \tag{*}$$

**(a) (5 points)** Describe a set $S$ of binary sequences whose size is given by the expression on the left.

**Solution.** Let $S$ be all $2n$-bit sequences with exactly $n+1$ ones.

Note that the problem explicitly required $S$ to be a set of binary sequences. Solutions that correctly proved the equality above but did not have $S$ to be a set of binary sequences lost 1 point. ■

**(b) (10 points)** Describe a way of partitioning $S$ into disjoint subsets $T_0, \ldots, T_{n-1}$ such that:

$$|T_k| = \binom{n}{k} \binom{n}{k+1}$$

In particular, state clearly which elements of $S$ are in set $T_k$ and explain why $|T_k|$ satisfies this equation.

**Solution.** Let $T_k$ consist of $2n$-bit sequences with exactly $k$ zeros in the first $n$ positions. Each such sequence has $n-k$ ones in the first $n$ positions, and thus $k+1$ ones in the last $n$ positions. There are $\binom{n}{k}$ ways to select the first $n$ bits and $\binom{n}{k+1}$ ways to select the last $n$ bits, and so there are $\binom{n}{k}\binom{n}{k+1}$ elements of $T_k$ in all.

The most common mistake in this part was to describe $T_k$ as the set of strings that have $k$ ones in the first half of the sequence and $k+1$ ones in the second half of the sequence, which is wrong as this doesn't produce strings in $S$, in general. Solutions lost 4 points for this. ■

**(c) (5 points)** Assume that you answered the previous two parts correctly. Explain why equation (*) logically follows.

**Solution.** Since $S$ is equal to the disjoint union $T_0 \cup \ldots \cup T_{n-1}$, the sum rule implies:

$$|S| = |T_0| \cup \ldots \cup |T_{n-1}|$$

Substituting the results from the two preceding parts gives equation (*).

For this part, the most common mistake was to forget to mention the disjointness of the $T_k$'s as a justification for applying the sum rule. Solutions lost 1 point for this. ■

**Problem 4 (20 points).** *Two* roommates can decide who must scrub the bathtub by flipping a coin. But how can *three* roommates decide? Here is a procedure that works, even if the coin itself is biased!

1. Flip the coin three times.
2. If the coin lands the same way every time, then go to step 1.
3. Otherwise, one of the three coins lands differently from the other two; let the random variable $k$ be the number of this coin. For example, $k(HHT) = 3$ and $k(THT) = 2$.

Suppose that the coin lands heads-up with probability $p$, where $0 < p < 1$, and that the results on successive tosses are independent.

**(a) (6 points)** What is the probability that the branch back to step 1 is taken?

**Solution.** $p^3 + q^3$ where $q ::= 1 - p$.

Answering just "$p^3$," got only 2 points of credit.

If you assumed $p = \frac{1}{2}$, which allowed you to give a numeric answer, you lost 4 points on this part. You were not penalized again on the following two parts for this mistake *assuming* your reasoning was correct. However, you had to be explicit in your reasoning to receive credit. ■

**(b) (7 points)** If step 3 is reached, what is the probability that two of the last three flips were heads?

**Solution.** $p$.

Various equivalent formulas also received full credit:

$$\begin{aligned} 3p^2q/(1-(p^3+q^3)) &= 3p^2q/(3p^2q+3pq^2) \\ &= p/(p+q) = p. \end{aligned}$$

If you only provided the numerator, you lost 5 points because the purpose of the problem was to test your ability to do conditional probabilities. If you missed the coefficient 3 (in the first version of the solution) you lost 2 points. If you included the probability of the case HHH, you lost 2 points because the step 3. is not reached at that outcome. ■

**(c) (7 points)** What is the expected number of times that the coin is flipped?

**Solution.** $3/(1-(p^3+q^3))$ where $q ::= 1-p$.

Various equivalent formulas also received full credit:

$$\begin{aligned} 3/(1-p^3-q^3) &= 3/(3p^2q+3pq^2) \\ &= 1/(pq(p+q)) = 1/pq. \end{aligned}$$

Solutions left in terms of an infinite summation were also accepted:

$$\sum_{i=0}^{\infty} 3(i+1)(p^3+q^3)^i(1-(p^3+q^3)).$$

Omitting the coefficient 3 for the expression cost 2 points. That was by far the most common mistake. ■

**Problem 5 (20 points).** Let $p_1, p_2, p_3, \ldots$ be the sequence of primes. Thus:

$$p_1 = 2, \quad p_2 = 3, \quad p_3 = 5, \quad p_4 = 7, \quad \ldots.$$

Let $k$ be an integer selected uniformly at random from the set $\{1, 2, 3, \ldots, p_n!\}$. (Note the factorial symbol.)

**(a) (4 points)** Suppose $1 \leq i \leq n$. What is the probability that $p_i \mid k$?

**Solution.**

$$\frac{1}{p_i}$$

■

**(b) (8 points)** What is the expected number of primes in the set $\{p_1, p_2, p_3, \ldots, p_n\}$ that divide $k$? You may leave your answer as a sum; a closed form is **not** required.

**Solution.**

$$\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n}$$

■

**(c) (8 points)** Suppose $1 \leq i < j \leq n$. Are the events $p_i \mid k$ and $p_j \mid k$ independent or not? Briefly justify your answer.

**Solution.** Yes, because

$$\begin{aligned}\Pr\{(p_i \mid k) \wedge (p_j \mid k)\} &= \Pr\{(p_i p_j \mid k)\} \\ &= \frac{1}{p_i p_j} \\ &= \Pr\{(p_i \mid k)\} \cdot \Pr\{(p_j \mid k)\}\end{aligned}$$

■

**Problem 6 (20 points).** Amy and Bill say sorry to Poor Pete for cheating him in the coin-flipping game. To make him feel better, they offer to play a new game:

1. Each player puts \$2 on the table.
2. Each player secretly writes a number between 1 and 4.
3. They roll a fair, four-sided die with faces numbered 1, 2, 3, and 4.
4. The money on the table is divided among the players that guessed correctly. If no one guessed correctly, then everyone gets their money back *and Poor Pete is paid* $\$0.25$ *in "apology money"*.

Suppose that, once again, Amy and Bill cheat by picking a pair of *distinct* numbers uniformly at random.

**(a) (12 points)** For each event listed below, indicate the probability of the event and Poor Pete's profit if that event occurs.

| | | |
|---|---|---|
| Pete guesses right AND either Amy or Bill guesses right | | |
| Pete guesses right AND both Amy and Bill guess wrong | | |
| Pete guesses wrong AND either Amy or Bill guesses right | | |
| Pete guesses wrong AND both Amy and Bill guess wrong | | |

**(b) (8 points)** What is Poor Pete's expected profit?

**Problem 7 (20 points).** We consider the unbiased Gambler's Ruin game: the gambler starts with $n$ dollars and has some fixed target of $T > 0$ dollars, where $n$ may vary between 0 and $T$. The gambler makes a series of 1 dollar bets. He wins each individual bet with probability 1/2; if he wins a bet, his wealth increases by 1 dollar, and if he loses a bet, his wealth decreases by 1 dollar. He continues betting until he has 0 dollars left (he is "ruined") or he has $T$ dollars (he "reaches his target").

Let

$$w_n ::= \Pr\{\text{the gambler reaches his target starting with } \$n\}.$$

**(a) (4 points)** We know that for $T \geq n \geq 2$, the probability $w_n$ satisfies the recurrence

$$w_n = aw_{n-1} + bw_{n-2}.$$

Fill in the values of $a$ 2, $b$ -1, $w_0$ 0, $w_T$ 1.

**Solution.** The recurrence equation can be derived from

$$w_{n-1} = \frac{1}{2}w_n + \frac{1}{2}w_{n-2}.$$

Most students gave the correct values of $w_0, w_T$. ■

**(b) (7 points)** Let $w(x)$ be the generating function for the sequence $w_0, w_1, w_2, \ldots$ where values of $w_n$ for $n > T$ are *defined* according to the recurrence above. Derive a closed form expression for $w(x)$ in terms of $x$ and $w_1$. **Note:** If you're unsure of your answers to part (a), you can also use the symbols "$a, b, w_0, w_T$" instead of their numerical values in your formula.

**Solution.**

$$\begin{array}{rcccccccccc}
w(x) & ::= & w_0 & + & w_1 x & + & w_2 x^2 & + & w_3 x^3 & + \dots \\
-2xw(x) & ::= & & - & 2w_0 x & - & 2w_1 x^2 & - & 2w_2 x^3 & + \dots \\
x^2 w(x) & ::= & & & & & w_0 x^2 & + & w_1 x^3 & + \dots \\
\hline
(1-2x+x^2)w(x) & = & 0 & + & w_1 x & + & 0 & + & 0 & + \dots
\end{array}$$

So

$$w(x) = \frac{w_1 x}{1-2x+x^2}.$$

5 points for using generating function but having incorrect closed-form expression for $w(x)$. ■

**(c) (1 point)** Briefly explain how to use the closed form in part (b) to conclude that

$$w(x) = \frac{A}{1-x} + \frac{B}{(1-x)^2}.$$

for some constants $A, B$.

**Solution.** $1 - 2x + x^2 = (1-x)^2$ so by partial fractions, there must be $A, B$ such that $w(x) = w_1 x/(1-x)^2 = A/(1-x) + B/(1-x)^2$. ■

**(d) (8 points)** Use part (c) and the known values of $w_0$ and $w_T$ to derive a closed form for $w_n$. *Note:* This part does *not* depend on the values of $a$ and $b$ in part (a).

**Solution.**

$$\begin{aligned}
0 = & w_0 = w(0) = A + B \\
1 = & w_T = \text{ coeff of } x^T \text{ in } A/(1-x) + \text{ coeff of } x^T \text{ in } B/(1-x)^2 \\
& = A + B(T+1).
\end{aligned}$$

So $A = -1/T$ and $B = 1/T$. Now

$$\begin{aligned}
w_n = & \text{ coeff of } x^n \text{ in } \frac{-1}{T(1-x)} + \text{ coeff of } x^n \text{ in } \frac{1}{T(1-x)^2} \\
& = \frac{-1}{T} + \frac{n+1}{T} = \frac{n}{T}.
\end{aligned}$$

4 points for a solution without finding the values of $A, B$. ■

**Problem 8 (10 points).** Suppose we perform $2n$ independent flips of a fair coin.

**(a) (3 points)** What is the number, $L_n$, of heads that is *most likely* to come up? $\underline{n}$

**(b) (7 points)** Derive an asymptotic ($\sim$) closed form for the probability of flipping *exactly* $L_n$ heads.

**Solution.** The closed form expression is

$$\frac{1}{\sqrt{\pi n}}.$$

We derive this closed form using Stirling's formula. We know the probability of flipping exactly $n$ heads is

$$\binom{2n}{n} 2^{-2n} = \left(\frac{2n!}{(n!)^2}\right) 2^{-2n}$$

Now

$$\begin{aligned}
\left(\frac{2n!}{(n!)^2}\right) 2^{-2n} &\sim \frac{\sqrt{2\pi 2n}\left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}\right)^2} \cdot 2^{-2n} \\
&= \frac{2\sqrt{\pi n}\left(\frac{2n}{e}\right)^{2n}}{2\pi n\left(\frac{n}{e}\right)^{2n}} \cdot 2^{-2n} \\
&= \frac{\sqrt{\pi n}\left(\frac{n}{e}\right)^{2n} \cdot 2^{2n}}{\pi n\left(\frac{n}{e}\right)^{2n}} \cdot 2^{-2n} \\
&= \frac{1}{\sqrt{\pi n}}.
\end{aligned}$$

3 points for the probability in binomial expression. 2 points for applying Stirling's formula. 2 points for the asymptotic closed form. ■

**Problem 9 (20 points).** We want to estimate the fraction, $d$, of defective silicon wafers in a long run of wafers from a fabrication facility. To do so, we make $n$ independent random choices of wafers from the run and test them for defects.

**(a) (10 points)**

Explain how to use the Binomial Sampling Theorem (given below) to calculate a near-minimal number, $n$, that will allow you to estimate $d$ within 0.005 with 98% confidence. Be sure to describe explicitly how to estimate $d$ and why there will be such an $n$.

**Theorem (Binomial Sampling).** *Let $K_1, K_2, \ldots,$ be a sequence of mutually independent 0-1-valued random variables with the same expectation, $p$, and let*

$$S_n ::= \sum_{i=1}^{n} K_i.$$

*Then, for $1/2 > \epsilon > 0$,*

$$\Pr\left\{\left|\frac{S_n}{n} - p\right| \geq \epsilon\right\} \leq \frac{1-2\epsilon}{2\epsilon} \cdot \frac{2^{-n(1-H((1/2)-\epsilon))}}{\sqrt{2\pi(1/4-\epsilon^2)n}} \tag{1}$$

*where*

$$H(\alpha) ::= -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha).$$

**Solution.** The estimate of $d$ will be the fraction of defective wafers among the sample of $n$. To find $n$, let $\epsilon = 0.005$, and search for the smallest $n$ such that the righthand side of (1) is $\leq 0.02$. There must be such an $n$ because, for any fixed $\epsilon \geq 0$, the righthand side expression approaches 0 as $n$ approaches infinity.

Many students lost 2 points for the incorrect claim that the needed $n$ must exist because of the Weak Law of Large Numbers. The Weak Law of Large Numbers ensures that the *lefthand* side of (1) can be made as small as desired by choosing $n$ large. However, we *find* such an $n$ by evaluating the bound given by the *righthand* expression, and the reason we can find the desired $n$ is that the righthand expression goes to zero.

In fact, it's this limiting behavior of the righthand side that *implies* the Weak Law. But the Weak Law does not conversely imply that a large $n$ will make the righthand small. For example, as $n$ increases, the lefthand side will go to zero, but a (not so good) upper bound for it might not go to zero. ■

**(b) (10 points)** The calculations in part (a) depend on some facts about how the $n$ wafers are chosen from the run. Write T or F next to each of the following statements to indicate whether it is True or False; there will be a penalty for wrong answers, so do not guess randomly.

- F Given a particular wafer, the probability that it is defective is $d$.
- T Given a particular wafer, the probability that it is defective is 1 or 0.
- F The probability that all $n$ randomly chosen wafers will be different is $\Theta(1/\sqrt{n})$.
- T The probability that the first randomly chosen wafer will be chosen again at another time approaches one as $n$ increases.
- T All wafers in the run are equally likely to be selected as the third among the $n$ random choices (assuming $n \geq 3$).

- T The expectation of the indicator variable for the last wafer *chosen* being defective is $d$.
- F The expectation of the indicator variable for the last wafer in the *fabrication run* being defective is $d$.
- F Given that the first randomly chosen wafer was defective, the probability that the second one will be defective is greater than $d$.
- T It turns out that there are several different colors of wafer. Given that the first and second randomly chosen wafers are the same color, the probability that the first wafer is defective may be $< d$.

**Solution.** (1) is false and (2) is true, since any particular wafer is either defective or not. We justify these answers in the same way we justify (7) below. Many students got these two questions wrong, perhaps because it was not clear what was meant by being "given a particular wafer." A better phrasing would have been to ask instead about the "9th wafer in the fabrication run," for example.
(3) is false; our analyis of the Birthday "paradox" implies the probability is exponentially decreasing in $n$.

(4) is true. The probability that the next $n$ wafers chosen after the first all differ from the first one is, $(1 - 1/(\text{length of run}))^n$, which approaches zero as $n$ increases. So the probability that the first wafer will be chosen again on one of the next $n$ choices approaches one as $n$ increases. So if we keep choosing, we are certain to pick the first wafer again.

(5) is true by definition of independent choice.

(6) is true because the expectation of an indicator variable is the probability that it equals 1.

(7) is false because we're sampling from a fixed fabrication run. We have no information about the probability of different kinds of runs, and we are not treating the run as a random process (it's our sampling that is random). So the last wafer in the given run is not random: it is either defective or not, and so it's indicator variable is a constant 0 or 1.

(8) is false by definition of independent choice.

(9) is true because the probability of defects may not be independent of color. For example, suppose $d = 1/2$ because half the wafers are the same color and are not defective, while all the other wafers are of different colors and are defective . Now if the fabrication run is large, the probability that the same wafer will be chosen the first two times is very small. So given that the first two wafers are the same color, it's almost certain that they are different wafers of the same color and therefore are not defective. In other words, the probability the first one is defective, given that the first two are the same color, will be very small —much less than $1/2$. ■