Midterm Solutions

1. Basic counting problems. Give brief answers to each of the following questions. Unsimplified expressions (in terms of factorials, sums, etc.) are acceptable.

   (a) Suppose we have an unlimited supply of red balls, blue balls, and green balls. Suppose we start picking balls of random color from our supply, and putting them into a box. How many balls must be placed in the box to *guarantee* that the box contains at least 7 red balls or at least 11 blue balls or at least 2 green balls? (5 points)

   Solution: By pigeonhole principle, if we have $6 + 10 + 1 + 1 = 18$ balls, the box must contain at least 7 red balls or 11 blue balls or 2 green balls. ∎

   (b) Consider all of the permutations of the symbols from multiset $\{a, a, a, b, b, b, b, c, c\}$. How many distinct permutations are there if we do not allow permutations that have both $c's$ together? (5 points)

   Solution: Two common approaches to the problem.

   i. Line all the a's and b's in a row and place an empty space between each neighboring two letters and at the two ends. With 7 a's and b's, we have 8 such empty postions. Now select two empty positions and place c's at these positions. There are $\binom{8}{2}$ ways to select. And for each selection, there are $\frac{7!}{3!4!}$ ways to permutate the a's and b's. Hence, the answer is $\binom{8}{2}\frac{7!}{3!4!} = 980$.

   ii. Another way to see it is to bundle the two c's. Now there are $\frac{8!}{3!4!1!}$ permutations with both c's together. The total number of permutations of these 9 symbols is $\frac{9!}{3!4!2!}$. Hence, the answer is $\frac{9!}{3!4!2!} - \frac{8!}{3!4!1!} = 980$.

   ∎

   (c) How many distinct solutions are there to the following equation

   $$x_1 + x_2 + \cdots + x_k = n, x_i \geq 1, n \geq k$$

   (5 points)

   Solution: The problem is exactly the same as throwing n balls into k bins such that each bin contains at least one ball. First, we place a ball in every bin and then there are $\frac{(n-k+k-1)!}{(k-1)!(n-k)!}$ distinct ways to place the remaining balls in the k bins. The answer is $\frac{(n-1)!}{(k-1)!(n-k)!}$. ∎

2. Solve the following recurrences.

   (a) State the general form of the solution and then determine the coefficients. (10 points) $T(1) = 1, T(2) = 6$, and for all $n \geq 3$,

   $$T(n) = T(n-2) + 3n + 4$$

   Solution: Generally, you can solve the problem by either applying annihilators, which entails solving linear equation system, or expanding the recurrence.

   i. The homogeneous annihilator for $T(n) - T(n-2) = 0$ is $E^2 - 1$ and that for $3n+4$ is $(E-1)^2$. The complete annihilator is thus $(E-1)^3(E+1)$. Since in general, $(E-a)^n$ annihilates $p(i)a^i$ where $p(i)$ is any polynomial in i of degree $n-1$. We have

   $$T(n) = C_1n^2 + C_2n + C_3 + C_4(-1)^n$$

   We are left to calculate the constants C_i. Using the initial conditions and the recurrence, we have $T(1) = 1, T(2) = 6, T(3) = 14, T(4) = 22$. Now plugging each of these numbers into the left side of the above equation and solve the resulting linear equation systems, we get $C_1 = 3/4, C_2 = 7/2, C_3 = -29/8, C_4 = 3/8$. Hence our solution is

$$T(n) = \frac{3}{4}n^2 + \frac{7}{2}n - \frac{29}{8} + \frac{3}{8}(-1)^n$$

ii. By expanding the recurrence several times, it's easy to read the recurrence as $T(n) = T(n - 2i) + 3i(n-(i-1)) + 4i$. When n is even, $i = n/2 - 1$; when n is odd, $i = (n-1)/2$. Plugging i into the equation, we get

$$\begin{aligned} T(n) &= T(2) + 3(n/2-1)(n-(n/2-1-1)) + 4(n/2-1) \\ &= \frac{3}{4}n^2 + \frac{7}{2}n - \frac{32}{8} \text{ when } n \text{ is even} \end{aligned}$$

and

$$\begin{aligned} T(n) &= T(2) + 3(n/2-1)(n-(n/2-1-1)) + 4(n/2-1) \\ &= \frac{3}{4}n^2 + \frac{7}{2}n - \frac{26}{8} \text{ when } n \text{ is odd} \end{aligned}$$

■

(b) Give a tight asymptotic solution to the following recurrence. $T(1) = 4$, and for all $n \geq 2$ a power of 2, (5 points)

$$T(n) = 3T(n/2) + n^2 - 2n + 1$$

Solution: You can solve it by applying Master theorem or by expanding the recurrence.

i. By expanding the recurrence, we get

$$\begin{aligned} T(n) &= 3T(n/2) + n^2 - 2n + 1 \\ &= 3(3T(n/4) + (n/2)^2 - 2(n/2) + 1) + n^2 - 2n + 1 \\ &\dots \\ &= 3^iT(n/2^i) + n^2(1 + (\frac{1}{2})^2 + \cdots + (\frac{1}{2^{i-1}})^2) - 2n(1 + \frac{1}{2} + \cdots + \frac{1}{2^{i-1}}) + i \end{aligned}$$

Since n is a power of 2, $i = \log n$,

$$\begin{aligned} T(n) &= 3^{\log n}T(1) + n^2\left(\frac{1-(\frac{1}{4})^{\log n}}{1-\frac{1}{4}}\right) - 2n\left(\frac{1-(\frac{1}{2})^{\log n}}{1-\frac{1}{2}}\right) + \log n \\ &= 4n^{\log_3 2} + \frac{4}{3}n^2 - 4n + \log n + \frac{8}{3} \\ &= \Theta(n^2) \end{aligned}$$

ii. By Applying Master theorem, it's easy to verify that the recurrence falls in category 3, in which $T(n) = \Theta(f(n)) = \Theta(n^2)$.

■

3. Toss a perfect six-sided die 12 times. Find the probability that every face of the die appears at least once. Unsimplified expressions (in terms of factorials, sums, etc.) are acceptable. Show your reasoning. (15 points)

Solution: We take the outcome of the 12 tosses as a sample point. For a perfect six-sided die, it's easy to verify that every sample point has the same probability associated with it. The size of the sample space is 6^{12}.

Let $P_i, 1 \leq i \leq 6$ be the property that face i doesn't appear in the 12 tosses and A_i be the set of sample points with property P_i. If follows that $\overline{P_i}$ is the property that face i appears at least once during the 12 tosses and $\overline{A_i}$ is the corresponding set. The set of all the sample points such that each face appears at least once can be expressed as $\overline{A_1} \bigcap \overline{A_2} \bigcap \ldots \bigcap \overline{A_6}$. Let S be the sample space.

$$
\begin{aligned}
|\overline{A_1} \bigcap \overline{A_2} \bigcap \cdots \bigcap \overline{A_6}| &= \overline{|A_1 \bigcup \mathcal{A}_2 \bigcup \cdots \bigcup A_6|} \\
&= |S| - |A_1 \bigcup \mathcal{A}_2 \bigcup \cdots \bigcup A_6| \\
&= |S| - \sum_{i=1}^{6} |A_i| + \sum_{1 \leq i < j \leq 6} |A_i \bigcap A_j| - \sum_{1 \leq i < j < k \leq 6} |A_i \bigcap A_j \bigcap A_k| + \cdots - \\
&\quad |A_1 \bigcap A_2 \bigcap A_3 \bigcap \cdots A_6| \\
&\quad /* \\
&\quad \text{since property } P_i \text{ is symmetric,which means} \\
&\quad |A_1| = |A_2| = \cdots = |A_6| = 5^{12} \\
&\quad |A_1 \bigcap A_2| = |A_1 \bigcap A_3| = \cdots = |A_5 \bigcap A_6| = 4^{12} \\
&\quad \ldots \\
&\quad |A_1 \bigcap A_2 \bigcap \cdots \bigcap A_5| = \cdots = |A_2 \bigcap A_3 \bigcap \cdots \bigcap A_6| = 1^{12} \\
&\quad */ \\
&= 6^{12} - \binom{6}{1} 5^{12} + \binom{6}{2} 4^{12} - \binom{6}{3} 3^{12} + \binom{6}{4} 2^{12} - \binom{6}{5} 1^{12}
\end{aligned}
$$

Finally divide this number with 6^{12} to get the probability. ■

4. Find a tight asymptotic solution to the following recurrence:

$$T(n) = T(\lfloor n/4 \rfloor) + \log \log n + 4$$

in which log denotes the base two logarithm. Partial credit can be earned for upper and lower bounds, if you are unable to find a tight bound. You may safely ignore the floor operator. (15 points)

Solution: By expanding the recurrence, we get

$$
\begin{aligned}
T(n) &= T(n/4) + \log \log n + 4 \\
&= T(n/16) + \log \log (n/4) + 4 + \log \log n + 4 \\
&\quad \ldots \\
&= T(\frac{n}{4^i}) + \log \log n + \log \log \frac{n}{4} + \cdots + \log \log \frac{n}{4^{i-1}} + 4i \\
&= T(\frac{n}{4^i}) + \log \log n + \log(\log n - 2) + \log(\log n - 4) + \cdots + \log(\log n - 2(i-1)) + 4i \\
&\quad \ldots \\
&= T(1) + \log \log n + \log(\log n - 2) + \log(\log n - 4) + \cdots + \log(4) + 4(\frac{1}{2}) \log n \\
/* &\quad T(1) = \Theta(1),\ 4(\frac{1}{2}) \log n = \Theta(\log n) \\
&\quad \text{denoting } D(n) = \log \log n + \log(\log n - 2) + \log(\log n - 4) + \cdots + \log(4) */ \\
&= \Theta(\log n) + D(n)
\end{aligned}
$$

We now bound $D(n)$. Recall the approach used in HW 3. We partition the range $[1 \ldots n]$ into intervals $I_i = [2^i, \ldots, 2^{i+1} - 1]$, for $i = 0, \ldots, \lg n$. Initially, $\log n$ is the current value and it falls in interval

$I_{\log\log n}$. Every time we subtract 2 from the current value and add to the sum the binary bits of the current interval,which is i. This process continues until the current value equals 4. There are at most 2^i distinct values in interval I_i. Hence,

$$D(n) = \Theta(\sum_{i=2}^{\log\log n} \frac{2^i}{2} \cdot i) = \Theta(\log n \log\log n)$$

Put together,

$$T(n) = \Theta(\log n) + \Theta(\log n \log\log n) = \Theta(\log n \log\log n)$$

■

5. Proofs by combinatorial reasoning and by induction.

   (a) Prove by *combinatorial reasoning* the following theorem. (8 points)
   Theorem: For all integers $n \geq 0$ and all real numbers $x_1, x_2, \cdots, x_k$,

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{n_1 + n_2 + \cdots + n_k = n \\ n_i \geq 0}} \frac{n!}{n_1!n_2!\cdots n_k!} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}$$

   Solution: Denote $(x_1+x_2+\cdots+x_k)$ as $S(k)$. The LHS can be written as $(x_1+x_2+\cdots+x_k)(x_1+x_2+\cdots+x_k)\cdots(x_1+x_2+\cdots+x_k)$. Considering the unsimplified result of the product, each $S(k)$ contributes one x_i to each term. For any fixed $n_1, n_2, \cdots, n_k$ where $n_1 + n_2 + \cdots + n_k = n$ and $n_i \geq 0$, the coefficient of term $x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}$ is the number of all the distinct ways of contribution from these n $S(k)$ such that x_i is selected n_i times, which is $\frac{n!}{n_1!n_2!\cdots n_k!}$. And it's easy to see that all the combinations of n_i are possible, which justifies the summation over all non-negative n_i such that they sum to n. Hence, we get the RHS of the equation. ■

   (b) Prove the same theorem by induction on n. (12 points)

   Solution: i. The Basic Step:
   For $n = 1$, the LHS is $x_1 + x_2 + \cdots + x_k$. Since $x_1 + x_2 + \cdots + x_k = 1$ and $x_i \geq 0$, exactly one x_i could equal to 1 every time. The RHS simplifies to

$$\sum_{\substack{n_1 + n_2 + \cdots + n_k = n \\ n_i \geq 0}} \frac{n!}{n_1!n_2!\cdots n_k!} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k} = \sum_{i=1}^{k} x_i = x_1 + x_2 + \cdots + x_k$$

   LHS equals RHS. The equation holds for $n = 1$.

   ii. The Induction Step:
   Suppose that the equation holds for $n = n^* \geq 1$. When $n = n^* + 1$,LHS equals

$$\begin{aligned} & (x_1 + x_2 + \cdots + x_k)^{n^*+1} \\ = \; & (x_1 + x_2 + \cdots + x_k)^{n^*} (x_1 + x_2 + \cdots + x_k) \\ = \; & \left(\sum_{\substack{n_1 + n_2 + \cdots + n_k = n^* \\ n_i \geq 0}} \frac{n^*!}{n_1!n_2!\cdots n_k!} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k} \right) (x_1 + x_2 + \cdots + x_k) \end{aligned}$$

For any term $x_{i_1}^{n_{i_1}} x_{i_2}^{n_{i_2}} \cdots x_{i_k}^{n_{i_m}}$ in the final product where $n_{i_j} > 0$, $1 \leq j \leq m \leq k$ and $\sum_{j=1}^{m} n_{i_j} = n^* + 1$, its coefficient comes from exactly m coefficients of those terms in case $n = n^*$. Namely, its coefficient is

$$\begin{aligned}
& \frac{n^*!}{(n_{i_1}-1)!n_{i_2}!\cdots n_{i_m}!} + \frac{n^*!}{n_{i_1}!(n_{i_2}-1)!\cdots n_{i_m}!} + \cdots + \frac{n^*!}{n_{i_1}!n_{i_2}!\cdots (n_{i_m}-1)!} \\
= \; & n^*! \left(\frac{n_{i_1}}{n_{i_1}!n_{i_2}!\cdots n_{i_m}!} + \frac{n_{i_2}}{n_{i_1}!n_{i_2}!\cdots n_{i_m}!} + \cdots + \frac{n_{i_m}}{n_{i_1}!n_{i_2}!\cdots n_{i_m}!} \right) \\
= \; & n^*! \frac{n_{i_1} + n_{i_2} + n_{i_3} + \cdots + n_{i_m}}{n_{i_1}!n_{i_2}!\cdots n_{i_m}!} \\
= \; & n^*! \frac{n^*+1}{n_{i_1}!n_{i_2}!\cdots n_{i_m}!} \\
= \; & \frac{(n^*+1)!}{n_{i_1}!n_{i_2}!\cdots n_{i_m}!} \text{ where } \sum_{j=1}^{m} n_{i_j} = n^* + 1
\end{aligned}$$

Since this is true for every term, the final product could be written as

$$\sum_{\substack{n_1 + n_2 + \cdots + n_k = n^* + 1 \\ n_i \geq 0}} \frac{(n^*+1)!}{n_1!n_2!\cdots n_k!} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}$$

This means the equation holds for $n = n^* + 1$. This ends the induction step, which completes the proof.

■

6. A box contains 10 pairs of socks, and each pair is a distinct color (e.g., there is a pair of yellow socks, a pair of red socks, etc.). Unsimplified expressions (in terms of factorials, sums, etc.) are acceptable. Show your reasoning.

   (a) Draw 2 socks at random from the box (without replacement). What is the probability that they are a pair of the same color? (5 points)

   Solution: After picking the first sock, there are 19 socks left. Among these socks, only one sock can form a pair with the chosen sock. Therefore the probability that they are a pair of the same color is $\frac{1}{19}$. ■

   (b) Draw 11 socks at random from the box of 20 socks (without replacement). Find the expected number of matching pairs in your sample. (15 points)

   Solution: Define indicator random variable $X_i, 1 \leq i \leq 11$ such that $X_i = 1$ if the i-th sock drawn forms a pair with another sock in the 11 socks and $X_i = 0$ otherwise. Let $X = \frac{1}{2}\sum_{i=1}^{11} X_i$. It's clear that X is the number of matching pairs in the sample. By part (i), we know that $Pr[X_i = 1] = \binom{10}{1}\frac{1}{19}$, we have

$$E[X] = 1/2E[\sum_{i=1}^{11} X_i] = 1/2\sum_{i=1}^{11} E[X_i] = 1/2\sum_{i=1}^{11} \Pr[X_i = 1] = 1/2\sum_{i=1}^{11} \binom{10}{1}\frac{1}{19} = 11/2 \cdot 10 \cdot \frac{1}{19} = \frac{55}{19}$$

■