CS70 Summer 2018 — Solutions to Homework 5

Sung Hyun Harvey Woo, SID 24190408, `CS70`

July 20, 2018

Collaborators: Dylan Hwang (dylanhwang@berkeley.edu), Allison Wang (awang24@berkeley.edu), Michael Hillsman (mhillsman@berkeley.edu)

Sundry

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up. — Sung Hyun Harvey Woo

Homework 5 Harvey Woo

1. (a) $2^{\binom{n}{2}}$.

Each edge can either exist or not exist, so each edge has 2 options and there are $\binom{n}{2}$ edges.

(b) $\left(\frac{1}{2}\right)^{\frac{k(k-1)}{2}}$ or $\left(\frac{1}{2}\right)^{\binom{k}{2}}$

To form a k-clique in a graph, each pair must be adjacent. In a k-clique, there are $\binom{k}{2}$ pairs and there is $\frac{1}{2}$ probability of an edge existing. Therefore, the probability is $\left(\frac{1}{2}\right)^{\binom{k}{2}}$

(c) $\binom{n}{k} \le n^k$.

$\frac{n!}{(n-k)!\,k!} \le n^k$

$\frac{n(n-1)(n-2)(n-3)\ldots(n-k+1)}{k(k-1)\ldots(1)} \le n^k$.

Comparing $n(n-1)(n-2)\ldots(n-k+1)$ and n^k.
we can see that $n^k = n \cdot n \cdot n \ldots n$ is the product of n, k times, which has to be bigger than $n(n-1)(n-2)\ldots(n-k+1)$ since they both have k terms and n, k are positive values. So, $n(n-1)\ldots(n-k+1) \le n^k$, and $k(k-1)\ldots(1) > 0$ since $k > 0$

therefore, $\frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots(1)} \le n^k$ since $n(n-1)\ldots(n-k+1)$ is smaller than n^k, and that term is again divided by a positive value.

(d) Lets assume event A_i is the probability of a graph containing a k-clique. Then, applying the union bound, we have

$$P(A) \leq P[A_1 \cup A_2 \cup \ldots A_n] \quad \text{for a positive } n.$$

$$P(A) \leq \sum_{n=1}^{\binom{n}{k}} \left(\frac{1}{2}\right)^{\binom{k}{2}}$$

$$P(A) \leq \binom{n}{k}\left(\frac{1}{2}\right)^{\binom{k}{2}}.$$

which we can get by using the probability from part 1.b. Now, we simplify the inequality using $k = 4\lceil \log n \rceil + 1$ to get at most $\frac{1}{n}$

$$\binom{n}{k}\left(\frac{1}{2}\right)^{\binom{k}{2}} \leq n^k \left(\frac{1}{2}\right)^{\binom{k}{2}} \quad \rightarrow \text{using part 1.c.}$$

$$\frac{n^k}{2^{\binom{k}{2}}} = \frac{n^k}{2^{\frac{k(k-1)}{2}}} = \frac{n^k}{\left(2^{\frac{k-1}{2}}\right)^k} = \frac{n^k}{\left(2^{\frac{4\lceil \log n \rceil + 1 - 1}{2}}\right)^k} = \frac{n^k}{\left(2^{2\lceil \log n \rceil}\right)^k}$$

$$= \frac{n^k}{n^{2k}} = \frac{1}{n} \qquad p(A) \leq \frac{1}{n}.$$

Therefore, we can see that using the union bound and part c, the probability that the graph contains a k-clique for $k = 4\lceil \log n \rceil + 1$ is at most $\frac{1}{n}$

2. (a) P(no one gets own ID back) = 1 - P(at least one person gets own ID back)
using the inclusion/exclusion principle:

$$\sum_{n=1}^{n} \frac{1}{n} = n \cdot \frac{1}{n} = 1 = \frac{1}{1!} \quad [= P(A_1) + P(A_2) \ldots + P(A_n)]$$

$$\sum_{n=1}^{\binom{n}{2}} \frac{1}{n} \cdot \frac{1}{n-1} = \binom{n}{2} \frac{1}{n(n-1)} = \frac{n!}{(n-2)!\,2!\,n(n-1)} = \frac{n!}{n(n-1)(n-2)!\,2!} = \frac{1}{2!}$$

$$\sum_{n=1}^{\binom{n}{3}} \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2} = \binom{n}{3} \frac{1}{n(n-1)(n-2)} = \frac{n!}{(n-3)!\,3!\,n(n-1)(n-2)}$$

$$= \frac{n!}{n(n-1)(n-2)(n-3)!\,3!} = \frac{1}{3!}$$

The pattern for each portion of the inclusion/exclusion principle is evident since the calculation for each one is essentially following a pattern.
The inclusion/exclusion principle on 1 - P(at least one person gets own ID) is, given $A_1, A_2 \ldots A_n$ as n people:

$$\underbrace{P(A_1) + P(A_2) + \ldots + P(A_n)} - \underbrace{\sum_{n,m} P(A_n \cap A_m)} + \underbrace{\sum_{n,m,l} (A_n \cap A_m \cap A_l)} \ldots$$

The pattern for each term is $\frac{1}{1!}, \frac{1}{2!}, \frac{1}{3!} \ldots$
therefore,

$$1 - P(\text{At least one person} \ldots) = 1 - \left(\frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} \ldots\right)$$

$$= 1 - \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k!} \quad \text{or} \quad \sum_{k=0}^{n} \frac{(-1)^{k}}{k!}$$

(b) if $n \to \infty$

we can use $\sum_{k=0}^{n} \frac{(-1)^k}{k!}$ and $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ to get an answer.

$\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}$ is what we get if the probability goes from $n \to \infty$. It is in the exact form as the e^x equation where $x = -1$, therefore,

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1}$$

3 (a). you can use the total probability rule to get answer using each condition.

(1) All green skittles eaten. $(P(G \cap A) = P(A)P(G|A))$

$\frac{1}{2} \cdot \frac{0}{80} = 0$

(2) Half of green skittles eaten. $(P(G \cap B) = P(B)P(G|B))$

$\frac{1}{4} \cdot \frac{10}{90} = \frac{1}{36} \approx 0.0278$

(3) 5 green skittles eaten $(P(G \cap C) = P(C)P(G|C))$.

$\frac{1}{4} \cdot \frac{15}{95} = \frac{15}{95 \cdot 4} \approx 0.0395$

$Pr[\text{Green skittle}] = P(G \cap A) + P(G \cap B) + P(G \cap C)$

$Pr[\text{green skittle}] = 0.0 + 0.0278 + 0.0395 = 0.0673$

(b) You can use the total probability rule, just like part 3.a

$P[\text{at least 1 green}] := P[G] = 1 - P(\text{none are green})$; $P(N) := P(\text{none are green})$

$P[N] = P(N \cap A) + P(N \cap B) + P(N \cap C)$, where A, B, C are respective green skittle situation.

(1) All green skittles eaten $(P(N \cap A) = P(A)P(N|A))$

$\frac{1}{2} \cdot \frac{80}{80} \cdot \frac{79}{79} = \frac{1}{2}$

(2) Half of green skittles eaten $(P(N \cap B) = P(B)P(N|B))$

$\frac{1}{4} \cdot \frac{80}{90} \cdot \frac{79}{89} \approx 0.197$

(3) 5 green skittles eaten $(P(N \cap C) = P(C)P(N|C))$

$\frac{1}{4} \cdot \frac{80}{95} \cdot \frac{79}{94} \approx 0.177$

4 $P(G) = 1 - P(N) = 1 - (0.5 + 0.197 + 0.177) = 0.126$

(C) $P(\text{pick 3 skittles, three green}) := P(G)$

$P(G) = P(G \cap A) + P(G \cap B) + P(G \cap C)$ [total probability rule]

(1) All skittles eaten $(P(G \cap A) = P(A)P(G|A))$

$\frac{1}{2} \cdot 0 \cdot 0 \cdot 0 = 0$

(2) Half of green skittles eaten $(P(G \cap B) = P(B)P(G|B))$

$\frac{1}{4} \cdot \frac{10}{90} \cdot \frac{9}{89} \cdot \frac{8}{88} \approx 0.000255$

(3) 5 of green skittles eaten $(P(G \cap C) = P(C)P(G|C))$

$\frac{1}{4} \cdot \frac{15}{95} \cdot \frac{14}{94} \cdot \frac{13}{93} \approx 0.000822$

$P(G) = P(G \cap A) + P(G \cap B) + P(G \cap C) = 0 + 0.000255 + 0.000822$

$= 0.00108$

(d) We can solve this problem using Baye's rule.

(1) probability that all green skittles were eaten.
In this case P(A) would simply be 0
since it is impossible for you to pick 3 or
any green skittles if it was all eaten.

(2) P(B) → probability of half of green skittles eaten.

$$P(B|G) = \frac{P(B \cap G)}{P(G)} = \frac{P(G|B)P(B)}{P(G)} = \frac{\frac{0.000255}{0.25} \cdot 0.25}{0.00108}$$

$\approx 0.236.$

(3) P(C) → probability of 5 green skittles eaten

$$P(C|G) = \frac{P(C \cap G)}{P(G)} = \frac{P(G|C)P(C)}{P(G)} = \frac{\frac{0.000822}{0.25} \cdot 0.25}{0.00108}$$

≈ 0.762

(e) $P(C) \rightarrow$ probability of drawing 3 skittles of the same color.
for the color green we already know that the probability of picking 3 green skittles is 0.00108.

$P(C) = P(R) + P(O) + P(Y) + P(P) + P(G)$
where $P(R), P(O), P(Y), P(P), P(G)$ are the respective probabilities of picking the same Red, Orange, Yellow, Purple, and Green skittle 3 times in a row.

Since the roommate eating the green skittles affects the other colors the same way,
$P(R) = P(O) = P(Y) = P(P)$
$P(R) = P(R \cap A) + P(R \cap B) + P(R \cap C)$
(1) all green skittles eaten $(P(R \cap A))$
$\frac{1}{2} \cdot \frac{20}{80} \cdot \frac{19}{79} \cdot \frac{18}{78} \approx 0.00694$
(2) Half of all green skittles eaten $(P(R \cap B))$
$\frac{1}{4} \cdot \frac{20}{90} \cdot \frac{19}{89} \cdot \frac{18}{88} \approx 0.00243$
(3) 5 of green skittles eaten $(P(R \cap C))$
$\frac{1}{4} \cdot \frac{20}{95} \cdot \frac{19}{94} \cdot \frac{18}{93} \approx 0.00206$

$P(R) = P(O) = P(Y) = P(P) = 0.00694 + 0.00243 + 0.00206$
$= 0.01143$

$P(C) = (4 \cdot 0.01143) + 0.00108 = 0.0468$

$P(C) = 0.0468$

4 (a) $P(A \setminus B) \geq P(A) - P(B)$

$P(A \setminus B) = P(A \cup B) - P(B) = P(A) - P(A \cap B)$

$P(A) - P(A \cap B) \geq P(A) - P(B)$

all we need to do now is compare $P(B), P(A \cap B)$ since $P(A)$ is the same.

$(A \cap B)$ is a subset of B, therefore

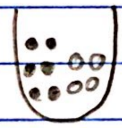
$P(A \cap B) \leq P(B)$, which indicates that

$P(A) - P(A \cap B) \geq P(A) - P(B)$ which means

$P(A \setminus B) \geq P(A) - P(B)$

(b) An event is any subset of the probability space. Therefore, the number of distinct events is: $\sum_{i=0}^{n} \binom{n}{i}$ or every possible way of choosing, i elements from n elements (unordered)

(C) Assume we have a bag with 6 black balls and 4 white balls. We pick one at random.



$P(A) = P(\text{picking black ball}) = \frac{6}{10} = 0.6$

$P(B) = P(\text{picking white ball}) = \frac{4}{10} = 0.4$

$P(C) = P(\text{Someone takes out all the black balls}) = 0.9$.

Probability space is all the ways we can pick a ball.

$P(A) > P(B)$, $P(A|C) = 0$ $P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{0.9 \cdot 1}{0.9} = 1$.

$P(A|C) < P(B|C)$

(d) If C, D are disjoint, $P(C \cap D) = 0$ and $P(C) > 0, P(D) > 0$ then C, D cannot be independent.

For C, D to be independent,

$P(C \cap D) = P(C) \cdot P(D)$, where $P(C) > 0, P(D) > 0$.

but, $P(C \cap D) = 0$. Therefore $P(C) \cdot P(D)$ has to be both 0, and bigger than 0 → Contradiction.

$\therefore$ C, D cannot be independent.

7

(e) $P(D|C) = P(D|\bar{C})$

$$\rightarrow \frac{P(C \wedge D)}{P(C)} = \frac{P(D \wedge \bar{C})}{P(\bar{C})}$$

$$\frac{P(C \wedge D)}{P(C)} = \frac{P(D) - P(D \cap C)}{1 - P(C)}$$

$$P(C \cap D)(1 - P(C)) = P(C)(P(D) - P(D \wedge C))$$

$$P(C \wedge D) - P(C)P(C \wedge D) = P(C)P(D) - P(C)P(D \wedge C)$$

$$P(C \wedge D) = P(C)P(D)$$

$\hookrightarrow$ C and D are independent

(f) Assume we roll 2 4-sided Die.

A: Getting an even number on the first die

B: Getting an even number on the second die

C: Chances of getting 1 odd, 1 even number for both Dice

(1) $P(A \wedge B) = \{(2,2), (2,4), (4,2)(2,2) = 4/16 = \frac{1}{4}$

$P(A) = 1/2 \quad P(B) = 1/2.$

$P(A \wedge B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B).$

(2) $P(A \cap C) = \{(2,1), (2,3) \ (4,1), (4,3)\} = \frac{1}{4}$

$P(A) = 1/2 \quad P(C) = (\frac{1}{2} \cdot \frac{1}{2}) + (\frac{1}{2} \cdot \frac{1}{2}) = \frac{1}{2}$

$P(A \wedge B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B)$

(3) $(B \cap C) = \{(1,2), (3,2), (1,4), (3,4)\} = \frac{1}{4}$

$P(B) = 1/2, \quad P(C) = (\frac{1}{2} \cdot \frac{1}{2}) + (\frac{1}{2} \cdot \frac{1}{2}) = \frac{1}{2}$

$P(B \wedge C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B) \cdot P(C)$

pairwise independent

(4) $P(A \wedge B \wedge C) = \{\} = 0$

$P(A) = 1/2 \quad P(B) = 1/2 \quad P(C) = 1/2$

$$[P(A \wedge B \wedge C) = 0] \neq [P(A) \cdot P(B) \cdot P(C) = \frac{1}{8}]$$

$\therefore$ A, B, C are pairwise independent, but not mutually independent

5 (a) Box 1: 1000 total, 100 defective, 10%
Box 2: 2000 total, 100 defective, 5%
$P(A)$: Defective lightbulb
$P(B)$: Selecting box 1 $= \frac{1}{2}$
$P(C)$: selecting box 2 $= \frac{1}{2}$
Here we can use Baye's rule.

$$pr(B|A) = \frac{P(B)\cdot P(A|B)}{P(B)\cdot P(A|B) + P(B^c)\cdot P(A|B^c)} = \frac{P(B)\cdot P(A|B)}{P(B)\cdot P(A|B) + P(C)\cdot P(A|C)}$$

$$= \frac{\frac{1}{2}\cdot\frac{1}{10}}{\frac{1}{2}\cdot\frac{1}{10}+\frac{1}{2}\cdot\frac{1}{20}} = \frac{2}{3}$$

$\therefore \frac{2}{3}$

(b) $P(A)$: Both lightbulbs are defective
$P(B)$: choose box 1 $= \frac{1}{2}$
$P(C) = P(B^c)$: choose box 2 $= \frac{1}{2}$
Also using Baye's rule,

$$P(B|A) = \frac{P(B)\cdot P(A|B)}{P(B)\cdot P(A|B) + P(C)\cdot P(A|C)} = \frac{\frac{1}{2}\cdot\frac{100}{1000}\cdot\frac{99}{999}}{\left(\frac{1}{2}\cdot\frac{100}{1000}\cdot\frac{99}{999}\right)+\left(\frac{1}{2}\cdot\frac{100}{2000}\cdot\frac{99}{1999}\right)}$$

$$= 0.8$$

$\therefore 0.8$

6 (a) $P(\bar{O}\cap F) = P(F) - P(F\cap O)$ $\rightarrow$ independent if $P(\bar{O}\cap F) = P(\bar{O})\cdot P(F)$
$= P(F) - (P(F)\cdot P(O))$ $\rightarrow$ since O, F are independent
$= P(F)(1-P(O))$
$= P(F)P(\bar{O})$ $\therefore \bar{O}, F$ are independent

(b) $P(O\cap \bar{F}) = P(O) - P(O\cap F)$
$= P(O) - P(O)P(F)$ $\rightarrow$ since O, F are independent
$= P(O)(1-P(F))$
$= P(O)P(\bar{F})$ $\therefore O, \bar{F}$ are independent

(c) $Pr(\bar{o} \cap \bar{F}) = P(\bar{o}) - P(\bar{o} \cap F)$

$= P(\bar{o}) - P(\bar{o})P(F)$ → $\bar{o}, F$ are independent from 6.a.

$= P(\bar{o})(1 - P(F))$

$= P(\bar{o})P(\bar{F})$ ∴ $\bar{o}, \bar{F}$ are independent.

(d) $O \cap (F \cap C)$ is independent if $P(O \cap (F \cap C)) = P(O) \cdot P(F \cap C)$

O, F, C are mutually independent, so

$P(O \cap F \cap C) = P(O) \cdot P(F) \cdot P(C)$

$= P(O) \cdot (P(F) \cdot P(C))$

$= P(O) \cdot P(F \cap C)$ → since, C, F are independent

$P(O \cap (F \cap C)) = P(O \cap F \cap C) = P(O) \cdot P(F \cap C)$

therefore, O, $(F \cap C)$ are independent

(e) O and $(F \Delta C)$ are independent if $P(O \cap (F \Delta C)) = P(O) \cdot P(F \Delta C)$

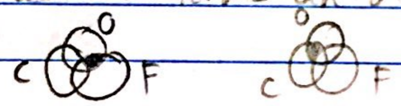
$F \Delta C = (F - C) \cup (C - F)$

$P(O \cap (F \Delta C)) = P(O \cap [(F - C) \cup (C - F)])$

$= P(O \cap (F - C)) \cup P(O \cap (C - F))$

$= P(O \cap F \cap \bar{C}) \cup P(O \cap C \cap \bar{F})$

These two terms are disjoint by nature.

→ 

Therefore, $P(O \cap F \cap \bar{C}) \cup P(O \cap C \cap \bar{F}) = P(O \cap F \cap \bar{C}) + P(O \cap C \cap \bar{F})$

$P(O \cap F \cap \bar{C})$

$= P((O \cap F) \cap \bar{C})$

$= P(O \cap F) - P(O \cap F \cap C)$

$= P(O) \cdot P(F) - P(O)P(F)P(C)$

$= P(O)P(F)(1 - P(C))$

$= P(O)P(F)P(\bar{C})$

$P(O \cap C \cap \bar{F})$

$= P((O \cap C) \cap \bar{F})$

$= P(O \cap C) - P(O \cap F \cap C)$

$= P(O)P(C) - P(O)P(F)P(C)$

$= P(O)P(C)(1 - P(F))$

$= P(O)P(C)P(\bar{F})$

$P(O \cap (F \Delta C)) = P(O)P(F)P(\bar{C}) + P(O)P(C)P(\bar{F})$

$= P(O)[P(F)P(\bar{C}) + P(C)P(\bar{F})]$

10 $= P(O)[P(F - C) + P(C - F)] = P(O)[P(F - C) \cup P(C - F)]$

$= P(O)P(F \Delta C)$ ∴ O and $(F \Delta C)$ are independent.

7. Let us assume that jar A is empty. we will need to pick out n cookies from jar A, and then choose jar A again, so that we know jar A is empty.
Let x be the number of cookies left in jar B once we find out jar A is empty. x can range from $1 \sim n$.
We also have $\frac{1}{2}$ chance of choosing jar A, and $\frac{1}{2}$ of choosing jar B.

The distribution of X is $P(X=k)$ for $1 \le k \le n$.

$$P(X=k) = \binom{2n-k}{n} 2\left(\frac{1}{2}\right)^{n+1}\left(\frac{1}{2}\right)^{n-k}$$

$\binom{2n-k}{n}$ is the number of ways we can choosing n cookies from all the cookies we picked when cookie jar A is empty.

This is multiplied by 2, because we have the case where jar A is empty, and the case where jar B is empty.

Then $\left(\frac{1}{2}\right)^{n+1}$ is the probability of choosing n cookies, plus 1 to discover that jar A is empty.

$\left(\frac{1}{2}\right)^{n-k}$ is the probability of choosing $(n-k)$ cookies from jar B, leaving k cookies at the end.

$$\therefore P_X(X=k) = \binom{2n-k}{n} 2\left(\frac{1}{2}\right)^{n+1}\left(\frac{1}{2}\right)^{n-k}$$

8 (a) $\sum_{i=0}^{k} \binom{k}{k-i}\binom{n-k}{i} = \binom{n}{k}$

The left side is showing that Alvin flips $(n-k)$ times and gets i heads, while Allen flips k times and gets $(k-i)$ heads. The summation gets all combinations of when $i=0$ all the way to when $i=k$ $(0 \leq i \leq k)$.

The sum of all possible ways Alvin and Allen gets $(i + (k-i) = k$ heads from $((n-k) + k = n)$ flips is equal to the right side, which is simply all possible ways that you can get k heads from n flips. Therefore, they are equal.

(b)

The probability of getting heads is $1/2$, tails is $1/2$. we can set n as the total number of flips so, the probability of getting k heads is simply $\binom{n}{k}\left(\frac{1}{2}\right)^n$ and the probability of Alvin and Allen getting the same number of heads is:

$$\sum_i \binom{k}{i}\left(\tfrac{1}{2}\right)^i\left(\tfrac{1}{2}\right)^{k-i}\binom{n-k}{i}\left(\tfrac{1}{2}\right)^i\left(\tfrac{1}{2}\right)^{n-k-i}$$

which is the sum of all i where $\binom{k}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{k-i}$ is the probability of Allen throwing k times and getting i heads, and $\binom{n-k}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{n-k-i}$ is the probability of Alvin throwing $n-k$ times and getting i heads

$$= \sum_i \binom{k}{i}\left(\tfrac{1}{2}\right)^k\binom{n-k}{i}\left(\tfrac{1}{2}\right)^{n-k} = \sum_i \left(\tfrac{1}{2}\right)^n\binom{k}{i}\binom{n-k}{i}$$

and since we know from part 8.a that

$$\sum_i^k \binom{k}{k-i}\binom{n-k}{i} = \binom{n}{k},$$

$$\sum_i \left(\tfrac{1}{2}\right)^n\binom{k}{i}\binom{n-k}{i} = \sum_i \left(\tfrac{1}{2}\right)^n\binom{k}{k-i}\binom{n-k}{i} = \left(\tfrac{1}{2}\right)^n\binom{n}{k}$$

12 Therefore, $\sum_i \binom{k}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{k-i}\binom{n-k}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{n-k-i} = \left(\frac{1}{2}\right)^n\binom{n}{k}$.