

HOMEWORK 2 SOLUTIONS

Brendan Meeder¹

1 Privacy and Signed Cards

1.1

Let p be the probability of yes. The probability of YES given that we see a + can be calculated with Bayes' rule once we know a few quantities. For the 2/1 split, we have that

$$Pr[+] = 2p/3 + (1-p)/3 = (p+1)/3, \quad Pr[YES] = p, \quad Pr[+|YES] = 2/3,$$

which gives us $Pr[YES|+] = 2p/(1+p)$.

In the case of the 11/10 split we have

$$Pr[+] = 11p/21 + 10(1-p)/21 = (p+10)/21, \quad Pr[YES] = p, \quad Pr[+|YES] = 11/21,$$

giving us

$$Pr[YES|+] = Pr[+|YES]Pr[YES]/Pr[+] = \frac{11p}{10+p}.$$

1.2

The maximum absolute gain in the 2/1 split occurs at $p = \sqrt{2} - 1$ and is a gain of $3 - 2\sqrt{2}$ (approximately 17.16%). The maximum relative gains occur for small values of p with a limiting value of 2. The relative gain is decreasing on $[0, 1]$. The absolute gain starts at 0, increases to its maximum value, and decrease back to 0 at $p = 1$.

For the 11/10 split, we have that the maximum absolute gain is approximately 2.38% ($21 - 2\sqrt{110}$, exactly) at $p = .488$ ($\sqrt{110} - 10$, exactly). The maximum relative gain is 10% at small values of p . The relative gain is decreasing on $[0, 1]$. The absolute gain starts at 0, increases to its maximum value, and decrease back to 0 at $p = 1$.

1.3

Let E be the event that we see 75+ cards and 25- cards. We want to pick the value of $p \in [0, 1]$ such that $Pr[Pr[Y] = p|E]$ is maximized. The trick here is to use Bayes' rule:

$$Pr[Pr[Yes] = p|E] = \frac{Pr[E|Pr[Yes] = p]Pr[Pr[Yes] = p]}{Pr[E]}.$$

Since $Pr[E]$ and $Pr[Pr[Yes] = p]$ are constants, we simply want to find the value of p such that $Pr[E|Pr[Yes] = p]$ is maximized. The event E is equivalent to seeing 75 of 100 coin flips come up heads, where heads occurs with probability $q(p) = 2p/3 + (1-p)/3$. The probability of seeing 75 heads is

$$\binom{100}{75} q^{75} (1-q)^{25}.$$

¹**Collaboration notice:** I did not collaborate with anyone on this homework.

We want to maximize this value, so we can take logarithms (as log is monotonic) and maximize $\log\binom{100}{75} + 75 \log q + 25 \log(1 - q)$. Taking derivatives and setting equal to zero we get:

$$75/q - 25/(1 - q) = 0 \iff q = 3/4.$$

We can graph or take second derivatives to confirm that this is where the maximum value occurs. Great, we now found the value of q , but what about p ? Unfortunately, the maximum value that $q(p)$ can take on when p is restricted to $[0, 1]$ is $2/3$, which is less than the value of q which maximizes the probability that E happens. Since $q(p)$ is an increasing function of p , and the probability of E is an increasing function on $q \in [0, 3/4]$, the choice of $p = 1$ maximizes the probability that E occurs.

1.4

The 2/1 split clearly gives away more information (both relative and absolute) about an individual compared to the 11/10 split, but it also allows us to make more accurate estimates about what fraction of the population is YES. The 11/10 split barely gives away information about an individual when the result of the card is drawn. I personally wouldn't mind doing the 11/10 split method, and we could use it with a large number of students to make accurate predictions of the number of people who do X.

2 ICM and Submodular Functions

2.1

The cardinality function is clearly nonnegative and monotone increasing by definition of the cardinality of a finite set. It is also submodular since for any finite sets we have that

$$|A \cup B| = |A| + |B| - |A \cap B| \Rightarrow |A \cup B| + |A \cap B| \leq |A| + |B|.$$

The function $f(A) = \max(A)$ is nonnegative as the ground set consists of nonnegative integers, and we defined it to be zero in the case that $A = \emptyset$. It's also monotone since whenever $A \subseteq B$, $x = \max(A)$ is also in B . Therefore, $\max(A) = x \leq \max(B)$. To show that it is submodular, let $x = \max(A \cup B)$ and without loss of generality assume $x \in A$. As $\max(A) = x$, it is sufficient to show that $\max(A \cap B) \leq \max(B)$. This is true by the monotonicity of max: $A \cap B \subseteq B \Rightarrow \max(A \cap B) \leq \max(B)$.

The function $f(A) = \text{avg}(A)$ is not monotone. Let $A = \{10\}$ and $B = \{1, 10\}$. Although $A \subset B$, $f(A) = 10 > f(B) = 5.5$ which shows that f is not monotone increasing (it's also not monotone decreasing by letting $A = \{1\}$ in the previous example).

Finally, the function $f(A) = |A|^2$ is not submodular (although it is clearly nonnegative and monotone increasing). Consider the case of $A = \{1, 2\}$ and $B = \{3, 4\}$. Then

$$f(A \cup B) + f(A \cap B) = 16 + 0 > f(A) + f(B) = 4 + 4.$$

2.2

Let F and G be arbitrary NN-M-SM-SFs and $\alpha, \beta \in \mathbb{R}^+$. For any $A \subseteq X$ we have that $F(A) \geq 0$ and $G(A) \geq 0$, so it is also the case that $\alpha F(A) + \beta G(A) = H(A) \geq 0$, and therefore H is nonnegative. Furthermore, let $A \subseteq B \subseteq X$. By the monotonicity of F, G we have that

$F(A) \leq F(B), G(A) \leq G(B)$, and as $\alpha, \beta > 0$, scaling F and G by these positive constants maintains these inequalities. Adding the scaled inequalities together we get

$$H(A) = \alpha F(A) + \beta G(A) \leq \alpha F(B) + \beta G(B) = H(B).$$

Applying the same reasoning that scaling inequalities by positive constants maintains the inequality to the inequality that defines submodular set functions, we get that

$$\begin{aligned} H(A \cup B) + H(A \cap B) &= \alpha F(A \cup B) + \beta G(A \cup B) + \alpha F(A \cap B) + \beta G(A \cap B) \\ &\leq \alpha F(A) + \beta G(A) + \alpha F(B) + \beta G(B) \\ &= H(A) + H(B) \end{aligned}$$

2.3

Let $A \subseteq B \subseteq X$ and $e \in X \setminus B$ be arbitrary. Define $X = A \cup \{e\}$ and $Y = B$ so that $X \cup Y = B \cup \{e\}$ and $X \cap Y = A$. Applying the definition of submodularity of F to the sets X and Y , we get

$$\begin{aligned} F(X \cup Y) + F(X \cap Y) &\leq F(X) + F(Y) \iff \\ F(B \cup \{e\}) + F(A) &\leq F(A \cup \{e\}) + F(B) \iff \\ F(B \cup \{e\}) - F(B) &\leq F(A \cup \{e\}) - F(A). \end{aligned}$$

2.4

The key insight here is that algorithm 1 would run identically to algorithm 1' in which all of the edges are determined (according to the same probabilities) before the algorithm runs and it looks up the values of the coin flips when it needs to. The distribution of the output for algorithm 1 and algorithm 1' are clearly the same since if algorithm 1 looks at a set of edges E , the probability of algorithm 1' observing the same outcome over the edges in E is the same. We will now argue that algorithm 1' and algorithm 2 have the same distribution of outputs by showing that for every outcome of coin flips, the algorithms output the same infected set. Arbitrarily fix an outcome E of edges that remain (an outcome of the coin flips). Algorithm 1' does something akin to breadth first search if you imagine the algorithm placing newly infected nodes at the end of a queue of nodes that will attempt to infect their neighbors. Thus, any vertex which is reachable via edges in E from vertices in A will be output as being infected, which is identical to the description of what algorithm 2 does.

The moral of this problem is that when sequential randomness is involved and you are trying to reason about the operation of an algorithm, it usually suffices to fix all of the random choices first and then consider the algorithm looking up outcomes as it is needed.

2.5

Let $R(A, E)$ be the set of all vertices reachable from vertices in A using only edges in E . Notice that this includes the vertices in A . We see that $\sigma(A)$ is the expected size of $R(A, E)$ where the randomness is taken over the distribution D of edge coin flip outcomes. An edge set $E \subset P$ (P is the set of potential edges in the graph) occurs with probability

$$Pr_D[E] = \left(\prod_{e \in E} p_e \right) \left(\prod_{e \in P \setminus E} (1 - p_e) \right),$$

where p_e is the probability that the message passes along edge e . Using the definition of the expectation of a random variable X over an outcome space,

$$\sigma(A) = \mathbb{E}_{E \sim D}[|R(A, E)|] = \sum_{E \subseteq P} |R(A, E)| \Pr_D[E].$$

2.6

Based on the previous part, we can show (by part 2) that σ is a NN-M-SM-SF if we show that $F_E(A) = |R(A, E)|$ is a NN-M-SM-SF for every subset of edges E . Let $E \subseteq P$ be arbitrary and define $F_E(A) = |R(A, E)|$. Clearly $F_E(A)$ is nonnegative since it is the cardinality of some finite set. If $A \subseteq B$ and $v \in R(A, E)$, then $v \in R(B, E)$ since for some $u \in A$ there is a u-v path using only edges in E and this same path is clearly a path from some vertex in B to v using only edges in E . Thus, $A \subseteq B \Rightarrow R(A, E) \subseteq R(B, E) \Rightarrow F_E(A) \leq F_E(B)$. Finally, to show that F_E is submodular

First, note that $R(A \cup B, E) = R(A, E) \cup R(B, E)$. Something which is a bit less obvious is that $R(A \cap B, E) \subseteq R(A, E) \cap R(B, E)$. If v is reachable via edges in E from some $u \in A \cap B$, then $v \in R(A, E), v \in R(B, E) \Rightarrow v \in R(A, E) \cap R(B, E)$. However,

$$\begin{aligned} F_E(A \cup B) + F_E(A \cap B) &= |R(A, E) \cup R(B, E)| + F_E(A \cap B) \\ &= |R(A, E)| + |R(B, E)| - |R(A, E) \cap R(B, E)| + F_E(A \cap B). \end{aligned}$$

However, notice that by the somewhat less obvious observation $|R(A, E) \cap R(B, E)| \geq F_E(A \cap B)$, and thus

$$F_E(A \cup B) + F_E(A \cap B) = F_E(A) + F_E(B) - |R(A, E) \cap R(B, E)| + F_E(A \cap B) \leq F_E(A) + F_E(B).$$

3 Erdos-Reyni

3.1

Yay, theory works, especially for large values of n . There's a small chance that for $p = \log n/n$ the graph is disconnected, but in that case the largest component has size $n - c$ for small values of c such as 1 or 2. You should have also gotten that the diameter is small (between 4 and 6 is what I saw).

3.2

Let's do the theoretical calculation, which as most of you observed closely matches what you get in the generated model. Let X_i^d be the indicator random variable that v_i has degree d , and $X^d = \sum_{i=1}^n X_i^d$ be the number of nodes with degree d . The degree of a particular vertex is distributed as a Binomial RV with parameters $n - 1$ and p , so that the probability that a particular vertex has degree d is $\binom{n-1}{d} p^d (1-p)^{n-1-d}$. Thus,

$$\mathbb{E}[X^d] = n \binom{n-1}{d} p^d (1-p)^{n-1-d}.$$

Notice that for $n = 1000$ and $p = 2 \log n/n$, when $d > 26$ the expected value is less than 1 and quickly decays.