CMSC651 Assignment 1

Yancy Liao, Fan Yang, Bowen Zhi

February 2018

Problem 1

Let A, B be independent events s.t. $0 < \Pr(A), \Pr(B) < 1$.

(a) Let $C = \neg(A \wedge B)$. Clearly, $\Pr(A \wedge B \mid C) = 0$. By the definition of conditional probability, we have:

$$\Pr(A \mid C)\Pr(B \mid C) = \frac{\Pr(A \wedge C)\Pr(B \wedge C)}{\Pr(C)^2} = \frac{\Pr(A \wedge \neg B)\Pr(B \wedge \neg A)}{\Pr(\neg(A \wedge B))^2}.$$

Note that since A and B are independent events, so are $\neg A$ and B, and $\neg B$ and A. Moreover, since each of these events have nonzero probability, we must have $\Pr(A \wedge \neg B), \Pr(B \wedge \neg A) > 0$, and thus their product is also greater than zero. Thus, in this case, we have $\Pr(A \wedge B \mid C) < \Pr(A \mid C)\Pr(B \mid C)$.

(b) Let $D = \neg(A \oplus B) = (A \wedge B) \vee \neg(A \vee B)$. We see that $\Pr(A \wedge B \mid D) = \Pr(A \wedge B)/\Pr(D)$, and:

$$\Pr(A \mid D)\Pr(B \mid D) = \frac{\Pr(A \wedge D)\Pr(B \wedge D)}{\Pr(D)^2} = \frac{\Pr(A \wedge B)\Pr(A \wedge B)}{\Pr(D)^2} = \Pr(A \wedge B \mid D)^2.$$

A and B are independent events each with nonzero probability, so $\Pr(A \wedge B) < \Pr(D)$, and thus $\Pr(A \wedge B \mid D) < 1 \implies \Pr(A \wedge B \mid D)^2 < \Pr(A \wedge B \mid D)$. Hence $\Pr(A \wedge B \mid D) > \Pr(A \mid D)\Pr(B \mid D)$.

Problem 2

(i) Let X be a variable for the number of edges cut from a graph with m edges. Note that $\mathrm{E}(X) = m/2$ (Lecture 4, Section 2.1). We have:

$$\Pr(X \geq 0.499m) = \Pr(m - X < 0.501m) = 1 - \Pr(m - X \geq 0.501m)$$

$$\geq 1 - \frac{\mathrm{E}(m - X)}{0.501m} = 1 - \frac{m - \frac{m}{2}}{0.501m} = 1 - \frac{0.5}{0.501} \approx 0.001996.$$

Note that the use of Markov's inequality is valid since $m - X$ is non-negative. Thus we have shown there is indeed a $c > 0$ s.t. $\Pr(X \geq 0.499m) \geq c$ (just set c to the value computed above).

(ii) As a consequence of the above, the probability that a random cut does not cut at least $0.499m$ edges is at most $\epsilon = 0.5/0.501$. Now consider k independent random cuts. The probability that all of them are of size less than $0.499m$ is ϵ^k. Hence, if we wish to construct a cut of size at least $0.499m$ with probability at least 0.9, simply make k random cuts and return the largest one, with k s.t. $\epsilon^k < 0.1$. Concretely, we have $k > \ln(0.1)/\ln(\epsilon) \approx 1152$.

Problem 3

Let X be as in Problem 2. To be exact, let X be the sum of m indicator variables X_i for the event that the i-th edge is cut. To find a lower bound for $\Pr(X \geq 0.499m)$, we can find an upper bound for its complement. Thus we have:

$$\Pr(X < 0.499m) \leq \Pr((X < 0.499m) \vee (X > 0.501m)) = \Pr(|X - 0.5m| > 0.001m) \leq \frac{\mathrm{E}((X - 0.5m)^2)}{(0.001m)^2}$$

$$= \frac{\mathrm{E}\left(\sum_i(X_i - 0.5)\right)^2}{(0.001m)^2} = \frac{\sum_i \mathrm{E}(X_i - 0.5)^2 + \sum_{i \neq j} \mathrm{E}((X_i - 0.5)(X_j - 0.5))}{(0.001m)^2}.$$

Consider the terms in the second sum: $\mathrm{E}\left((X_i - 0.5)(X_j - 0.5)\right) = \mathrm{E}\left(X_i X_j\right) - 0.5(\mathrm{E}\left(X_i\right) + \mathrm{E}\left(X_j\right)) + 0.25$. We can think of the graph cut as independently placing each vertex into one of two groups with equal probability (Lecture 4, Section 2.1). Hence $\mathrm{E}\left(X_i\right) = \mathrm{E}\left(X_j\right) = 1/2$. For $\mathrm{E}\left(X_i X_j\right) = \Pr\left(\text{edges i,j both cut}\right)$, there are two cases to consider: the edges share a vertex, and they do not share a vertex. In the first case, we see that both edges will be cut iff both of their unshared vertices are not in the same group as the shared vertex, which occurs with probability $1/4$. In the second case, both edges are cut iff they individually have vertices in both groups, which occurs with probability $1/4$. Thus $\mathrm{E}\left(X_i X_j\right) = 1/4$, and we see that the terms in the second sum become 0. Continuing, we have:

$$\Pr\left(X < 0.499m\right) \leq \frac{\sum_i \mathrm{E}\left(X_i - 0.5\right)^2}{(0.001m)^2} = \frac{\sum_i(\mathrm{E}\left(X_i^2\right) - \mathrm{E}\left(X_i\right) + 0.25)}{(0.001m)^2} = \frac{\sum_i(\mathrm{E}\left(X_i\right) - \mathrm{E}\left(X_i\right) + 0.25)}{(0.001m)^2}$$
$$= \frac{0.25m}{(0.001m)^2} \sim \mathcal{O}\left(\frac{1}{m}\right).$$

Hence, we have shown that $\Pr\left(X \geq 0.499m\right) = 1 - \Pr\left(X < 0.499m\right) \geq 1 - \mathcal{O}\left(1/m\right)$.

Problem 4

(i) We have $\Pr\left(X \geq \mu + a\right) = \Pr\left(X - \mu + b \geq a + b\right)$, for some $b > 0$. The idea here is to shift the value on the LHS in the probability so that it is more likely to be positive. This way, once we square both sides of the inequality, the outcomes of X in the range $(\mu - 2b - a, \mu - b)$ no longer contribute to the resulting probability. We have:

$$\Pr\left(X \geq \mu + a\right) \leq \Pr\left((X - \mu + b)^2 \geq (a + b)^2\right) \leq \frac{\mathrm{E}\left((X - \mu + b)^2\right)}{(a + b)^2}$$
$$= \frac{\mathrm{E}\left((X - \mu)^2\right) + 2b(\mathrm{E}\left(X\right) - \mu) + b^2}{(a + b)^2} = \frac{v + b^2}{(a + b)^2}.$$

Let us minimize the above with respect to b. Note that $\frac{\mathrm{d}}{\mathrm{d}b}\left[(v + b^2)/(a + b)^2\right] = (2ab - 2v)/(a + b)^3$ is 0 for $b = v/a$. Moreover, $\frac{\mathrm{d}^2}{\mathrm{d}b^2}\left[(v + b^2)/(a + b)^2\right] = 2(a^2 - 2ab + 3v)/(a + b)^4$ is positive at $b = v/a$, thus we have a minimum at $b = v/a$. Setting b to this value, the above bound becomes:

$$\Pr\left(X \geq \mu + a\right) \leq \frac{v + (v/a)^2}{(a + v/a)^2} = \frac{a^2 v + v^2}{(a^2 + v)^2} = \frac{v(a^2 + v)}{(a^2 + v)^2} = \frac{v}{v + a^2}.$$

Similarly, for some $b > 0$:

$$\Pr\left(X \leq \mu - a\right) = \Pr\left(\mu - X + b \geq a + b\right) \leq \Pr\left((\mu - X + b)^2 \geq (a + b)^2\right)$$
$$\leq \frac{\mathrm{E}\left((\mu - X)^2\right) + b^2}{(a + b)^2} = \frac{v + b^2}{(a + b)^2}.$$

This bound is exactly the same as for $\Pr\left(X \geq \mu + a\right)$, and thus also becomes $v/(v + a^2)$.

Combining the two bounds, we get that $\Pr\left(|X - \mu| \geq a\right) = \Pr\left(X \geq \mu + a\right) + \Pr\left(X \leq \mu - a\right) \leq 2v/(v + a^2)$. Note that this bound is only less than the Chebyshev bound when a is smaller than the standard deviation. But in that case, both bounds are above 1, and hence useless. Thus it never makes sense to use this bound.

(ii) We have $\Pr\left(X \leq u\right) \geq 1/2$. Note that $\Pr\left(X \leq u\right) = \Pr\left(X \leq \mu - (\mu - u)\right) \leq v/(v + (\mu - u)^2)$. Thus we must have $v/(v + (\mu - u)^2) \geq 1/2 \implies (\mu - u)^2 \leq v \implies |\mu - u| \leq \sigma$. So $-\sigma \leq \mu - u \leq \sigma$, and hence $\mu - \sigma \leq u \leq \mu + \sigma$. Note that we need not check if the equivalent conditions for $\Pr\left(X \geq u\right) \geq 1/2$ would further restrict the range of u, as the bound we would use is the same, and thus should yield the same requirement.

Problem 5

(i) Let $X = \sum_{i=1}^{n-4} X_i$, where each X_i is an indicator variable for "proof" appearing in the string starting at character i. Note then that:

$$\operatorname{Var}(X) = \operatorname{E}(X^2) - \operatorname{E}(X)^2 = \sum_i \operatorname{E}(X_i^2) + 2\sum_{i<j} \operatorname{E}(X_i X_j) - \frac{(n-4)^2}{26^{10}}$$

$$= \sum_i \operatorname{E}(X_i) + 2\sum_{i<j} \operatorname{E}(X_i X_j) - \frac{(n-4)^2}{26^{10}} = \frac{n-4}{26^5} + 2\sum_{i<j} \operatorname{E}(X_i X_j) - \frac{(n-4)^2}{26^{10}}.$$

Observing that the characters typed are independent, and that "proof" appearing at character i also implies it cannot appear again at the next 4 characters, we see that $\operatorname{E}(X_i X_j) = \Pr(\text{"proof" at i,j})$ is 26^{-10} if $j - 4 > i$, and 0 otherwise. Thus in the sum above, for each $i \in \{1, 2, \ldots, n-5\}$, there are $\max(n-4-i-4, 0)$ non-zero terms. Therefore the sum has, in total, $n - 9 + n - 10 + \cdots + 1$ or $(n-9)(n-8)/2$ non-zero terms. Hence we have $2\sum_{i<j} \operatorname{E}(X_i X_j) = (n-9)(n-8)/(26^{10})$, and thus:

$$\operatorname{Var}(X) = \frac{n-4}{26^5} + \frac{(n-9)(n-8) - (n-4)^2}{26^{10}} = \frac{n-4}{26^5} - \frac{9n-56}{26^{10}} \sim \mathcal{O}(n).$$

(ii) As above, let $X = \sum_{i=1}^{n-d+1} X_i$, and thus have:

$$\operatorname{Var}(X) = \frac{n-d+1}{26^d} + 2\sum_{i<j} \operatorname{E}(X_i X_j) - \frac{(n-d+1)^2}{26^{2d}}.$$

Let us break the sum above into two parts: one part has all terms where $j - d + 1 > i$, and the other has the remaining terms. By the same reasoning as in the previous part, the first part of the sum has value $(n-2d+1)(n-2d+2)/(2*26^{2d})$. As for the second part, we first observe that each term is $1/(26^{d+j-i})$ if it is possible for the string to occur at both positions, and 0 otherwise. In the worst case, the string consists entirely of one character, and thus all terms are nonzero. Thus, each $i \in \{1, 2, \ldots, n-d\}$ in the sum has terms $1/(26^{d+1}), 1/(26^{d+2}), \ldots, 1/(26^{2d-1})$ (some values of i do not have all these terms, but we ignore this since we will be bounding them from above anyways). Adding these terms up, we see that they are bounded by a geometric series whose sum is $1/((1 - 1/26)26^{d+1}) = 1/(25 * 26^d)$. Moreover, there are at most $n - d$ values of i in this sum, thus this part of the sum is no larger than $(n-d)/(25 * 26^d)$. Thus:

$$\operatorname{Var}(X) \leq \frac{n-d+1}{26^d} + \frac{2(n-d)}{25 * 26^d} + \frac{(n-2d+1)(n-2d+2) - (n-d+1)^2}{26^{2d}}$$

$$\leq \frac{n-d+1}{26^d} + \frac{2(n-d)}{25 * 26^d} + \frac{(n-d+1)(n-d+2) - (n-d+1)^2}{26^{2d}}$$

$$= \frac{27(n-d) + 25}{25 * 26^d} + \frac{n-d+1}{26^{2d}} \leq \frac{27n+25}{25 * 26^d} + \frac{n+1}{26^{2d}} \sim \mathcal{O}\left(\frac{n}{26^d}\right).$$

Problem 6

Let $Y = \sum_{i=1}^{n} Y_i$, and let each $Y_i = X_i - 1/2$. Clearly, we have:

$$Y = X - n/2,$$

and for each i:

$$Y_i = \begin{cases} -\dfrac{1}{2} & \text{with probability} = 0.5 \\ +\dfrac{1}{2} & \text{with probability} = 0.5. \end{cases}$$

It's easy to see that $\forall i\ E[Y_i] = E[Y_i^3] = 0, E[Y_i^2] = 1/4, E[Y_i^4] = 1/16$, thus:

$$
\begin{aligned}
E[(X - \frac{n}{2})^4] = E[Y^4] &= E\left[\left(\sum_{i=1}^{n} Y_i\right)^4\right] \\
&= E\left[\sum_i Y_i^4 + 4\sum_{i\neq j} Y_i Y_j^3 + 3\sum_{i\neq j} Y_i^2 Y_j^2 + 6\sum_{i\neq j\neq k} Y_i^2 Y_j Y_k\right] \\
&= \sum_i E[Y_i^4] + 4\sum_{i\neq j} E[Y_i]E[Y_j^3] + 3\sum_{i\neq j} E[Y_i^2]E[Y_j^2] + 6\sum_{i\neq j\neq k} E[Y_i^2]E[Y_j]E[Y_k] \\
&= \sum_i \frac{1}{16} + 4\sum_{i\neq j} 0\cdot 0 + 3\sum_{i\neq j} \frac{1}{4}\cdot\frac{1}{4} + 6\sum_{i\neq j\neq k} \frac{1}{4}\cdot 0\cdot 0 \\
&= \frac{n}{16} + 3\cdot n(n-1)\cdot\frac{1}{16} \\
&= \frac{3}{16}n^2 - \frac{1}{8}n \\
&= \mathcal{O}(n^2).
\end{aligned}
$$