Spring 2005: EECS126 Take Home Pseudo-Final
Assigned: May 5 in class
Due: May 10 in class

- Aside from your pen/pencil, you are allowed to use only your textbook and blank sheets of paper. No notes, calculating devices, or Internet.
- You are obliged by the honor system to spend no longer than 4 hours from the time you first open the exam to the time that you finish working on it. (Eat something before, go to the bathroom, turn off your cellphone and stay away from your computer so you are not interrupted and thereby lose time.) Mark the start and stop time on your exam.
- If you wish to spend longer than that, you must clearly mark the sections that you have worked on once the time limit is up. You can receive some partial credit for that extra work if you want.

Problem 3.1 *(36pts) True or False. Prove or show a counterexample:*

a. *12pts. If a random variable has finite second moment, then it has finite first moment.*

True. For any random variable X we have $|X| \leq 1 + X^2$, which implies $E[|X|] \leq E[1 + X^2] = 1 + E[X^2]$. This is a consequence of $|x| \leq 1 + x^2$ for any real number x. Hence, if $E[X^2] < \infty$ we have $E[|X|] < \infty$ which implies that $E[X]$ exists and is finite.

b. *12pts. Let A_1, A_2, A_3 be events with $0 < P(A_3) < 1$. Suppose A_1 and A_2 are conditionally independent given A_3 and are also conditionally independent given A_3^c. Then A_1 and A_2 are independent.*

False. For example, suppose the sample space is the unit square, i.e. $\{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$, with the probability of an event (a subset of the unit square) equal to its area. Let A_3 be defined by

$$A_3 = \{(x, y) : 0 \leq x \leq \frac{1}{3}\}$$

Let

$$A_1 = A_3 \cup \{(x, y) : \frac{1}{3} \leq x \leq 1, \frac{1}{2} \leq y \leq 1\}$$

and let

$$A_2 = \{(x, y) : 0 \leq x \leq \frac{2}{3}\}$$

Then $A_1 \cap A_3 = A_2 \cap A_3 = A_3$. Hence

$$P(A_1 \cap A_2 | A_3) = 1 = P(A_1|A_3)P(A_2|A_3)$$

so A_1 and A_2 are conditionally independent given A_3. Also, $P(A_1 \cap A_3^c) = P(A_2 \cap A_3^c) = \frac{1}{3}$, $P(A_3^c) = \frac{2}{3}$ and $P(A_1 \cap A_2 \cap A_3^c) = \frac{1}{6}$. Hence,

$$P(A_1 \cap A_2 | A_3^c) = \frac{\frac{1}{6}}{\frac{2}{3}} = \frac{1}{4} = P(A_1|A_3^c)P(A_2|A_3^c)$$

so A_1 and A_2 are conditionally independent given A_3^c.

However, we have $P(A_1) = P(A_2) = \frac{2}{3}$ and $P(A_1 \cap A_2) = \frac{1}{2}$ so A_1 and A_2 are not independent.

c. 12pts. For any random variable X and any $a > 0$ we have

$$P(|X| < a) \leq a^2 E[\frac{1}{X^2}]$$

True. Let Y denote $\frac{1}{X}$. Then Chebyshev's inequality gives

$$P(|Y| > \frac{1}{a}) \leq a^2 E[Y^2]$$

Since $\{|Y| > \frac{1}{a}\} = \{|X| < a\}$ this is the same as

$$P(|X| < a) \leq a^2 E[\frac{1}{X^2}]$$

Problem 3.2 *"Pairing up"*

In a particularly simplified society, there are exactly n young males and exactly n young females. In their mythology, for every male i there is exactly 1 female who is destined to be his soulmate and amazingly he is also destined to be her soulmate.[1]

a. *20 pts. In one case, the strategy for pairing up individuals is as follows. At every time, two currently unattached males are randomly chosen and paired with a randomly chosen single unattached female. If either pairing turns out to be soulmates, then that pair drops out of the pool of unattached individuals, while the remaining male returns to the pool. Otherwise, all three enjoy themselves and all return to the pool. The process starts with all $2n$ individuals in the single pool and it will terminate when everyone is married.*

Write out a model for this and
1. find the probability that the first randomly drawn group will result in finding a pair of soulmates?
2. give an exact expression for the expected time till everyone is married.

How does this time scale with increasing n?

We model this as a Markov chain where states represent number of unmarried males. There are n+1 states: $\{0, 1, 2, ...n\}$. The initial state is n. The state 0 is an absorbing state which denotes the situation when every male is married. Note that there is a only positive probability to transition from state i to itself and state $i-1$, because the number of unmarried males at every drawing can either remain the same (no match) or one of the males get married.

The transition probabilities depend only on the number of males currently left unmarried. If there are K such males (and females), then there are K^2 possible parings. Of them exactly K represent soulmate parings. Probability of getting a soulmate for one chosen male is $\frac{K}{K^2} = \frac{1}{K}$. Since the two males are drawn at one period the probability of marriage is $\frac{2}{K}$. As such, the number of drawings before the Markov chain enters the state $i-1$ if it started at state i denotes as T_i has a geometric distribution with $p_i = \min(1, \frac{2}{i})$. Note that the last two rounds are deterministic since there are only two males left.

1. Probability that the first randomly drawn group will result in finding a pair of soulmates is $p_1 = \frac{2}{n}$.

2. So the time when everyone is married is $T = \sum_{i=2}^{n} T_i + 1$. Each of the T_i for $i > 1$ is geometric and hence $E[T_i] = \frac{1}{p_i} = \frac{i}{2}$. For $i = 1$ it takes 1 drawing.

[1] As you can see, this model has almost no connection to reality as in the real world there are a host of good pairing possibilities for any individual.

Extra Space:

$$
\begin{aligned}
E[T] &= E[\sum_{i=2}^{n} T_i + 1] \\
&= \sum_{i=2}^{n} E[T_i] + 1 \\
&= \sum_{i=2}^{n} \frac{i}{2} + 1 \\
&= \frac{1}{2}(\sum_{j=1}^{n} j - 1) + 1 \\
&= \frac{1}{2}(\frac{n(n+1)}{2} - 1) + 1 \\
&= \frac{n(n+1)}{4} + \frac{1}{2}
\end{aligned}
$$

This time scales quadratically with increasing n.

b. *20 pts. In another case, the society decides to randomly deploy matchmakers to help get people paired with their soulmates. For every* $2n$ *singles, there are only about* $\lceil K \ln n \rceil$ *match-making elders. Every single person* i *has an independent probability* p *of knowing any particular matchmaker* k*. If any matchmaker knows both sides of a pair of soulmates, then that pair will be brought together and married. A pair cannot be brought together if they know no matchmakers in common. If there is any pair of soulmates that can not be brought together, it is considered a tragedy for the society.*

1. *What is the probability that the pair* (i, j) *do not have matchmaker* k *in common?*

For any individual matchmaker, there is a probability p^2 that both i and j have a connection to the matchmaker. So the probability of at least one of them not having a connection to the matchmaker is $(1 - p^2)$.

2. *Give an expression for the probability that a specified pair of soulmates* (i, j) *will not be brought together.*

The soulmates (i, j) will not be brought together exactly when they do not have any matchmakers in common. This is the same as saying that for all the matchmakers, they do not both have connections to them. Since the connections to different matchmakers are all independent and they are $\lceil K \ln n \rceil$ matchmakers, the probability that (i, j) share no common matchmakers is

$$(1 - p^2)^{\lceil K \ln n \rceil}$$

3. *What is the best bound you can give on the probability of a tragedy occurring?*

Let A_i denote the event that there is a tragedy involving male i. The event of a tragedy occurring with any male is therefore $\bigcup_{i=1}^{n} A_i$. We then apply the union bound:

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) &\leq \sum_{i=1}^{n} P(A_i) \\ &= \sum_{i=1}^{n} (1 - p^2)^{\lceil K \ln n \rceil} \\ &= n(1 - p^2)^{\lceil K \ln n \rceil} \\ &\leq n(1 - p^2)^{K \ln n} \end{aligned}$$

In this problem as presented, the different pairs of soulmates are independent as far as their prospects for tragedy go. This is because soulmate pairs are disjoint and so the matchmaker connections of one soulmate pair are completely independent of the matchmaker connections of another soulmate pair. As a result, it might seem possible to avoid using the union bound here and instead computing directly:

$$\begin{aligned}
P(\bigcup_{i=1}^{n} A_i) &= 1 - P(\{\bigcup_{i=1}^{n} A_i\}^c) \\
&= 1 - P(\bigcap_{i=1}^{n} A_i^c) \\
&= 1 - \prod_{i=1}^{n} P(A_i^c) \\
&= 1 - (1 - P(A_i))^n \\
&= 1 - (1 - (1 - p^2)^{\lceil K \ln n \rceil})^n \\
&\leq 1 - (1 - (1 - p^2)^{K \ln n})^n
\end{aligned}$$

4. As a function of the parameter K, how does the probability of tragedy change in the limit of large n?

One way is to start from the union bound. Then, the limit of taking $n \to \infty$ is:

$$\begin{aligned}
\lim_{n\to\infty} P(\bigcup_{i=1}^{n} A_i) &\leq \lim_{n\to\infty} e^{\ln n}(1 - p^2)^{K \ln n} \\
&= \lim_{n\to\infty} \{e(1 - p^2)^K\}^{\ln n}
\end{aligned}$$

If $K > \frac{-1}{\ln(1-p^2)}$ then

$$\lim_{n\to\infty} P(\bigcup_{i=1}^{n} A_i) = 0$$

Since the union bound goes to zero we know that the true probability will also go to zero since the true probability has to be sandwiched between 0 below and the union bound above.

If $0 < K < \frac{-1}{\ln(1-p^2)}$ then union bound does not gives us any useful results since

$$\lim_{n\to\infty} P(\bigcap_{i=1}^{n} A_i) \leq \lim_{n\to\infty} (e(1 - p^2)^K)^{\ln n} = \infty.$$

Here, we can use a tighter bound to see what is actually happening to the probability of no tragedies. Taking a ln of this probability (since ln is monotonic) we are interested in:

$$\lim_{n\to\infty} \ln (1 - (1 - p^2)^{K \ln n})^n = \lim_{n\to\infty} n \ln (1 - (1 - p^2)^{K \ln n})$$

We apply the Taylor series expansion for $\ln(1-x)$ for the case of small x. Note that since $x = (1-p^2)^{K\ln n} = e^{\ln(1-p^2)K\ln n} = n^{K\ln(1-p^2)}$ and n is large while $K\ln(1-p^2) < 0$, x is indeed very small. In this regime:

$$\ln(1-x) \leq -x$$

So then:

$$\begin{aligned} \lim_{n\to\infty} n\ln\left(1-(1-p^2)^{K\ln n}\right) &= \lim_{n\to\infty} n\ln\left(1-n^{\ln(1-p^2)K}\right) \\ &\leq -\lim_{n\to\infty} n(n^{\ln(1-p^2)K}) \\ &= -\lim_{n\to\infty} n^{1+\ln(1-p^2)K} \end{aligned}$$

If indeed $0 < K < \frac{-1}{\ln(1-p^2)}$, then $1+\ln(1-p^2)K > 0$ and thus the limit on the right diverges to $-\infty$. Since the log of the probability of no tragedies tends to $-\infty$, the probability of no tragedies must tend to zero and hence the probability of tragedy tends to 1.

In summary, if $0 < K < \frac{-1}{\ln(1-p^2)}$, the probability of tragedy tends to 1 with increasing n. If $K > \frac{-1}{\ln(1-p^2)}$, then the probability of tragedy tends to 0.

c. *20 pts. After a pair of supposed soulmates gets married, they find themselves in a Markov process with three states:*

Honeymoon: From this state, they continue in this state for the next period with probability $\frac{3}{4}$ but otherwise fall into the "fighting state."

Fighting: From this state, they continue fighting in the next period with probability $\frac{1}{2}$ but otherwise make up and enter the "Civil" state.

Civil: From this state, they stay civil in the next period with probability $\frac{2}{3}$ but otherwise fall into the "fighting state."

They start out in the "honeymoon state" once they get married.

1. Which states are recurrent and which are transient? How many recurrent classes are there?

2. For arbitrary time t after they are married, give an expression for the probability that they are currently fighting.

3. As t gets very large, what does the probability of fighting tend towards?

To solve this problem without having to calculate any determinants at all, we notice that the honeymoon state is transient. The fighting and civil states are recurrent. There is only one recurrent class in this Markov chain. The transition matrix for this Markov chain is

$$A = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

where we use the row vector $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ to denote the honeymoon state with certainty, the row vector $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ to denote the fighting state with certainty, and the row vector $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ to denote the civil state with certainty.

The two of the eigenvalues and eigenvectors of this system are going to correspond purely to the embedded recurrent chain described by:

$$A' = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

For this chain, one eigenvalue is guaranteed to be 1. Moreover, its corresponding eigenvector is the steady state probabilities. The balance equations in this case tell us immediately that the steady state probability of fighting is $\frac{(\frac{1}{3})}{(\frac{1}{2})} = \frac{2}{3}$ times the probability of being civil and hence $\pi = v_1^T = \begin{bmatrix} 0 & \frac{2}{5} & \frac{3}{5} \end{bmatrix}$.

The other λ of the recurrent class can be found by realizing that the sum of the eigenvalues must be the trace of the matrix. And so $1 + \lambda = \frac{1}{2} + \frac{2}{3}$ giving us $\lambda = \frac{1}{6}$. To solve for the corresponding eigenvector, we just notice $\frac{1}{6} = \frac{1}{2} - \frac{1}{3}$ and hence $v_{\frac{1}{6}}^T = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$.

To now get the λ associated with staying in the transient states, we can apply the trace trick again to the big matrix A to get $\lambda_T = \frac{3}{4}$. To solve for the corresponding eigenvector $v_{\frac{3}{4}}^T$, we just set the first entry to 1 for convenience and then get the following system of equations (with f and c as the entries in the fighting and civil components of this eigenvector.):

$$\begin{aligned}\frac{3}{4}f &= \frac{1}{4} + \frac{1}{2}f + \frac{1}{3}c \\ \frac{3}{4}c &= \frac{1}{2}f + \frac{2}{3}c\end{aligned}$$

The second equation tells us that $c = 6f$ and so the first equation tells us that $f = -\frac{1}{7}$ and thus $c = -\frac{6}{7}$. So we have $v_{\frac{3}{4}}^T = \left[\begin{array}{ccc} 1 & -\frac{1}{7} & -\frac{6}{7}\end{array}\right]$.

All that remains is to decompose the vector representing our starting with certainty in the honeymoon state. We want $\alpha_{\frac{3}{4}}, \alpha_1, \alpha_{\frac{1}{6}}$ so that

$$\left[\begin{array}{ccc} 1 & 0 & 0\end{array}\right] = \alpha_{\frac{3}{4}} v_{\frac{3}{4}}^T + \alpha_1 v_1^T + \alpha_{\frac{1}{6}} v_{\frac{1}{6}}^T$$

because then we know that the probability of fighting at some arbitrary time t is:

$$\alpha_{\frac{3}{4}}(\frac{3}{4})^t(-\frac{1}{7}) + \alpha_1 1^t(\frac{2}{5}) + \alpha_{\frac{1}{6}}(\frac{1}{6})^t(1)$$

Now, it might appear that we have to solve a system of three equations to get the α terms, but the special structure gives us the answers almost immediately. We know that at $t = 0$, all the probability needs to be in the honeymoon state. And thus, $\alpha_{\frac{3}{4}} = 1$. We also know that at time infinity, we are going to be in the steady state which means that the probability of fighting needs to be $\frac{2}{5}$. And thus, $\alpha_1 = 1$. The only term left is $\alpha_{\frac{1}{6}}$ which we can get by looking at the probability of fighting at time 0:

$$0 = -\frac{1}{7} + \frac{2}{5} + \alpha_{\frac{1}{6}}$$

and thus $\alpha_{\frac{1}{6}} = \frac{1}{7} - \frac{2}{5} = -\frac{9}{35}$.

So the answer to the problem is that the probability of fighting at some time t is

$$\frac{2}{5} - (\frac{3}{4})^t(\frac{1}{7}) - \frac{9}{35}(\frac{1}{6})^t$$

There is another way of working this problem out by conditioning on the amount of time L it takes to leave the honeymoon state. This is a geometric random variable. Once we have left the honeymoon state, we can ask the question of what is the probability of fighting at time t given that we were fighting at time L. This will express the desired probability as a sum.

As t gets very large, what does the probability of fighting tend towards?

This is just the steady state probability of fighting: $\frac{2}{5}$

Problem 3.3 *"Transmitters With Messages"*

Transmitters A and B independently send messages to a single receiver in a Poisson manner, with rates of λ_A and λ_B respectively. All messages are so brief that we may safely assume that they occupy single points in time. The number of words in the i-th message, regardless of the source that is transmitting it, is a random variable W_i with PMF:

$$P_W(w) = \begin{cases} \frac{1}{2} & \text{if } w = 1 \\ \frac{1}{3} & \text{if } w = 2 \\ \frac{1}{6} & \text{if } w = 3 \\ 0 & \text{otherwise} \end{cases}$$

and is independent of anything else.

a. *5 pts. In an interval of duration t, what is the probability that exactly 9 messages are received?*

This problem is the same as Prob.5.10 in the textbook with some numeric changes.

Let R denote the total number of messages received during the interval of duration t. Note that R is a Poisson random variable with parameter $(\lambda_A + \lambda_B)t$. Therefore, the probability that exactly 9 messages are received is

$$P(R = 9) = \frac{((\lambda_A + \lambda_B)t)^9 e^{(-(\lambda_A+\lambda_B)t)}}{9!}$$

b. *5 pts. Let N be the total number of words received in an interval of duration t. What is E[N]?*

The total number of words received in an interval of duration t can be expressed as

$$N = N_1 + 2N_2 + 3N_3$$

where N_i is the number of arrivals that have i words in them. Then each word is coming from independent Poisson processes by the fact that they are coming from a random splitting of a single poisson process of rate $(\lambda_A + \lambda_B)t$.

$$\begin{aligned} E[N] &= E[N_1] + 2 * E[N_2] + 3 * E[N_3] \\ &= (1\frac{1}{2} + 2\frac{1}{3} + 3\frac{1}{6})(\lambda_A + \lambda_B)t \\ &= \frac{5}{3}(\lambda_A + \lambda_B)t \end{aligned}$$

c. *10 pts. Determine the PDF of the length of time from* 0 *until the receiver receives the third message that is exactly 2 words long from transmitter A.*

Two-word messages arrive from transmitter A in a Poisson manner, with rate $\lambda_A P_W(2) = \frac{\lambda_A}{3}$. Therefore , the random variable Y of interest is Erlang of order 3, and its PDF is

$$f_Y(y) = \frac{\left(\frac{\lambda_A}{3}\right)^3 y^2 e^{-\frac{\lambda_A y}{3}}}{2!}$$

d. *10 pts. What is the probability that exactly* 2 *out of the next* 5 *messages received will be from transmitter A.*

Every message originates from either transmitter A or B, and can be viewed as an independent Bernoulli trial. Each message has probability $\frac{\lambda_A}{\lambda_A+\lambda_B}$ of originating from transmitter A. Thus the the number of messages from transmitter A is a binomial random variable, and the desired probability is equal to

$$\binom{5}{2}(\frac{\lambda_A}{\lambda_A+\lambda_B})^2(\frac{\lambda_B}{\lambda_A+\lambda_B})^3$$