This and other notes are available from `http://www.cs.berkeley.edu/~yozo/cs170.fa05/`

Fun with Probability

- Random variable: $X = x_i$ with probability p_i.
- Expectation: $\mathrm{E}[X] = \sum_i x_i p_i$.
- Linearity of expectations: $\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y]$.
- Independence. Events $\mathcal{E}_1$ and $\mathcal{E}_2$ are independent if $\Pr[\mathcal{E}_1, \mathcal{E}_2] = \Pr[\mathcal{E}_1] \times \Pr[\mathcal{E}_2]$. Equivalently $\Pr[\mathcal{E}_2 \mid \mathcal{E}_1] = \Pr[\mathcal{E}_2]$. Intuitively, whether or not event $\mathcal{E}_1$ happens or not does not affect the occurrence of event $\mathcal{E}_2$.
- Indicator variables: a 0-1 random variable.
- Flip of a fair coin. Let $X_i = 1$ if the i-th flip is heads. What is $\mathrm{E}[X_i]$? What is the expected number of heads in the first n tosses?
- Consider n people at a hotel, each with his or her own room. After a party, they all get drunk and chooses a random room to sleep in.
  - What is the expected number of people that sleeps in his or her own room? Does things change when we restrict it so that only one person can sleep in any given room?
  - What is the probability that no one sleeps in his or her own room? What does this approach as $n \to \infty$?
  - What is the expected number of people that sleeps in his or her own room *alone*?
  - What is the expected number of people that sleeps alone, not necessarily in his or her own room?
  - If there are m rooms available, then what is the probability that there is two (or more) people sleeping in the same room?

Randomized Max-Cut Approximation
Random cut (assigning vertices randomly to S or T) gives expected $\frac{1}{2}|E|$ size.

- Prove it. Hint: consider an indicator variable X_{uv} for each edge (u, v). Are these X_{uv} independent?
- True or false: A random variable X satisfies $X \geq \mathrm{E}[X]$ with probability $\geq 1/2$.
- How can you convert this to a randomized algorithm that gives a cut of size $\geq \frac{1}{3}|E|$ with probability $1 - \epsilon$ for small ϵ? Give the running time.
- Now give an algorithm that gives a cut of size $\geq \frac{1}{2}|E|$ with probability $1-\epsilon$. Give the running time.
- Extend the result to weighted graph. Now the cut size is the sum of weights of the edges that cross the cut.

- Interesting application. Consider breaking a simple substitution cipher. The cipher uses some alphabet X. Use the fact that in many European languages, the consonants and vowels alternate quite frequently to try to separate X into two subsets: vowels V and consonants C.
- Markov inequality: if X is a nonnegative random variable, then

  $$\Pr\left[X \geq \lambda \mathrm{E}[X]\right] \leq \frac{1}{\lambda}$$

  for any $\lambda \geq 1$.

Randomized Min-Cut
This will be covered in lecture (I hope).

Questions

1. Does linearity of expectations hold even if the variables involved are not independent?
2. Birthday Paradox: if there are more than 23 persons in the class, then most likely there are two persons in the class with the same birthday (assuming uniform distribution of birthdays). Prove it.

   Another interesting case: if there are more 14 persons in the class, then most likely there are two persons in the class with birthdays within 1 day of each other.