Midterm Review

CSE 312
Counting

• **Product Rule:** If there are \( n \) outcomes for some event \( A \), sequentially followed by \( m \) outcomes for event \( B \), then there are \( n \cdot m \) outcomes overall. General: \( n_1 \times n_2 \times \ldots \times n_k \)

• **Permutation:** an arrangement of objects in a definite order \( N!/(N-n)! \)

• **Combination:** a selection of objects with no regard to order \( N!/[n!(N-n)!] \)
Binomial Theorem

\[(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\]

\[\sum_{k=0}^{n} \binom{n}{k} = 2^n\]

**Proof:**

\[\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k} = (1 + 1)^n = 2^n\]
Inclusion-Exclusion

- for two sets or events $A$ and $B$, whether or not they are disjoint, $|A \cup B| = |A| + |B| - |A \cap B|$

- General: $|A \cup B \cup C| = |A| + |B| + |C| - |B \cap C| - |A \cap C| - |A \cap B| + |A \cap B \cap C|$
Pigeonhole Principle

• If there are \( n \) pigeons in \( k \) holes and \( n > k \), then some hole contains more than one pigeon. More precisely, some hole contains at least \( \lceil n/k \rceil \) pigeons.

• Problem: network problem on HW
Sample spaces / Events / Sets

- **Sample space:** $S$ is the set of all possible outcomes of an experiment (notation: $\Omega$)
- **Events:** $E \subseteq S$ is an arbitrary subset of the sample space
- **Set:**
  - subset: $A \subset B$
  - Union: $A \cup B = \{ x \mid x \in A \text{ or } x \in B \}$
  - Intersection: $A \cap B = \{ x \mid x \in A \text{ and } x \in B \}$
  - Complement: $A' = \{ x \mid x \notin A \} = A^c$
  - Mutually Exclusive / Disjoint: $A \cap B = \emptyset$
  - Any number of sets $A_1, A_2, A_3, \ldots$ are mutually exclusive if and only if $A_i \cap A_j = \emptyset$ for $i \neq j$
DeMorgan’s Laws

\[ \overline{E \cup F} = \bar{E} \cap \bar{F} \]

\[ \overline{E \cap F} = \bar{E} \cup \bar{F} \]
Axioms of Probability

- Axiom 1 (Non-negativity): \(0 \leq \Pr(E)\)
- Axiom 2 (Normalization): \(\Pr(S) = 1\)
- Axiom 3 (Additivity): If \(E\) and \(F\) are mutually exclusive (\(EF = \emptyset\)), then \(\Pr(E \cup F) = \Pr(E) + \Pr(F)\)

If events \(E_1, E_2, \ldots, E_n\) are mutually exclusive

\[
\Pr\left(\bigcup_{i=1}^{n} E_i\right) = \Pr(E_1) + \cdots + \Pr(E_n)
\]
Conditional Probability

- **Conditional probability** of $E$ given $F$: probability that $E$ occurs given that $F$ has occurred. $P(E|F)$

$$P(E \mid F) = \frac{|EF|}{|F|} = \frac{|EF|/|S|}{|F|/|S|} = \frac{P(EF)}{P(F)}$$
Chain Rule

\[ P(E \mid F) = \frac{P(EF)}{P(F)} \quad \text{where, } P(F) > 0 \]

• General definition of Chain Rule:

\[
P(E_1E_2\cdots E_n) = \\
P(E_1)P(E_2 \mid E_1)P(E_3 \mid E_1, E_2)\cdots P(E_n \mid E_1, E_2, \ldots, E_{n-1})
\]
Law of Total Probability

- \( E \) and \( F \) are events in the sample space \( S \):
  \[ E = EF \cup EF' \]

\[
P(E) = P(EF) + P(EF^c)
= P(E|F) P(F) + P(E|F^c) P(F^c)
= P(E|F) P(F) + P(E|F^c) (1-P(F))
\]

\[
P(E) = \sum_i P(E|F_i) P(F_i)
\]
Bayes Theorem

Most common form:

\[ P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)} \]

Expanded form (using law of total probability):

\[ P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)} \]

Proof:

\[ P(F \mid E) = \frac{P(EF)}{P(E)} = \frac{P(E \mid F)P(F)}{P(E)} \]
Independence

• Two events \(E\) and \(F\) are independent if 
\[ P(EF) = P(E)P(F). \]
If \(P(F) > 0\), \(P(E|F) = P(E)\)
Otherwise, they are dependent.

• Three events \(E\), \(F\), \(G\) are independent if 
\[ P(EF) = P(E)P(F) \quad P(EG) = P(E)P(G) \quad P(FG) = P(G)P(G) \]
and \(P(EFG) = P(E)P(F)P(G)\)

• Events \(E_1, E_2, \ldots, E_n\) are independent if for 
every subset \(S\) of \(\{1,2,\ldots, n\}\), we have 
\[
P \left( \bigcap_{i \in S} E_i \right) = \prod_{i \in S} P(E_i) \]
Independence

• Theorem: E, F independent $\Rightarrow$ E, F’ independent

• Theorem: if P(E)>0, P(F)>0, then E, F independent $\iff$ P(E|F)=P(E) $\iff$ P(F|E) = P(F)
Network Failure

• **Parallel**: $n$ routers in parallel, $i$th has probability $p_i$ of failing, independently

$$P(\text{there is functional path}) = 1 - P(\text{all routers fail}) = 1 - p_1 p_2 \ldots p_n$$

• **Series**: $n$ routers, $i$th has probability $p_i$ of failing, independently

$$P(\text{there is functional path}) = P(\text{no routers fail}) = (1 - p_1)(1 - p_2) \ldots (1 - p_n)$$
Conditional Independence

- Two events $E$ and $F$ are called conditionally independent given $G$, if
- $P(EF|G) = P(E|G) \cdot P(F|G)$
- Or, $P(E|FG) = P(E|G)$, $(P(F)>0, P(G)>0)$
PMF / CDF

• PMF: probability mass function

\[ p(a) = \begin{cases} P(X = a) & \text{for } a \in T \\ 0 & \text{otherwise} \end{cases} \]

• CDF: cumulative distribution function:

\[ F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \leq a \end{cases} \]
Expectation

• For a discrete r.v. $X$ with p.m.f. $p(\bullet)$, the expectation of $X$ (expected value or mean), is $E[X] = \sum x \cdot p(x)$
Properties of Expectation

• Linearity:
• For any constants a, b: \( E[aX + b] = aE[X] + b \)

• Let \( X \) and \( Y \) be two random variables derived from outcomes of a single experiment. Then \( E[X+Y] = E[X] + E[Y] \)
Variance

- The variance of a random variable $X$ with mean $E[X] = \mu$ is $\text{Var}[X] = E[(X-\mu)^2]$, often denoted $\sigma^2$. 
Properties of Variance

1. \( \text{Var}(X) = E[X^2] - (E[X])^2 \)

2. \( \text{Var}[aX+b] = a^2 \times \text{Var}[X] \)

3. \( \text{Var}[X+Y] \neq \text{Var}[X] + \text{Var}[Y] \)
r.v.s Independence

• Defn: Random variable $X$ and event $E$ are independent if the event $E$ is independent of the event $\{X=x\}$ (for any fixed $x$), i.e. $\forall x \ P(\{X=x\} \ & \ E) = P(\{X=x\}) \cdot P(E)$

• Defn: Two random variables $X$ and $Y$ are independent if the events $\{X=x\}$ and $\{Y=y\}$ are independent (for any fixed $x$, $y$), i.e. $\forall x, y \ P(\{X=x\} \ & \ {Y=y}) = P(\{X=x\}) \cdot P(\{Y=y\})$
Joint Distributions

• Joint probability mass function:
  \[ f_{XY}(x, y) = P(\{X = x\} \& \{Y = y\}) \]

• Joint cumulative distribution function:
  \[ F_{XY}(x, y) = P(\{X \leq x\} \& \{Y \leq y\}) \]
Marginal Distributions

• Marginal PMF of one r.v.: sum over the other

\[ f_Y(y) = \sum_x f_{XY}(x,y) \]
\[ f_X(x) = \sum_y f_{XY}(x,y) \]
Discrete Random Variables
Bernoulli Distribution

**Definition:** value 1 with probability \( p \), 0 otherwise (prob. \( q = 1-p \))

**Example:** coin toss \( (p = \frac{1}{2} \text{ for fair coin}) \)

**Parameters:** \( p \)

**Properties:**

\[
E[X] = p \\
\text{Var}[X] = p(1-p) = pq
\]
Binomial Distribution

**Definition:** sum of \( n \) independent Bernoulli trials, each with parameter \( p \)

**Example:** number of heads in 10 independent coin tosses

**Parameters:** \( n, p \)

**Properties:**

\[
E[X] = np
\]

\[
\text{Var}(X) = np(1-p)
\]

pmf: \( \Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \)
Poisson Distribution

**Definition:** number of events that occur in a unit of time, if those events occur independently at an average rate $\lambda$ per unit time

**Example:** # of cars at traffic light in 1 minute, # of deaths in 1 year by horse kick in Prussian cavalry

**Parameters:** $\lambda$

**Properties:**
- $E[X] = \lambda$
- $\text{Var}[X] = \lambda$
- pmf: $Pr(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$
Geometric Distribution

**Definition:** number of independent Bernoulli trials with parameter $p$ until and including first success (so $X$ can take values 1, 2, 3, ...)

**Example:** # of coins flipped until first head

**Parameters:** $p$

**Properties:**

$$E[X] = \frac{1}{p}$$

$$\text{Var}[X] = \frac{1 - p}{p^2}$$

$$\text{pmf: } \Pr(X = k) = (1 - p)^{k-1}p$$
Hypergeometric Distribution

Definition: number of successes in $n$ draws (without replacement) from $N$ items that contain $K$ successes in total

Example: An urn has 10 red balls and 10 blue balls. What is the probability of drawing 2 red balls in 4 draws?

Parameters: $n$, $N$, $K$

Properties:

$$E[X] = n \frac{K}{N}$$

$$\text{Var}[X] = n \frac{K(N - K)}{N(N - 1)}$$

pmf: $$Pr(X = k) = \frac{\binom{K}{k} \binom{N - K}{n - k}}{\binom{N}{n}}$$

Think about the pmf; we've been doing it for weeks now: ways-to-choose-successes times ways-to-choose-failures over ways-to-choose-$n$

Also, consider that the binomial dist. is the with-replacement analog of this