Midterm 1 Study Guide

William Sun

Probability

- $P(E)$: Probability of an event, which is a *set* of sample points.
- P: Subsets of sample space $\to [0, 1]$
  - Axiom 1: For any event E, $0 \leq P(E) \leq 1$
  - Axiom 2: $P(\Omega) = 1$
  - Axiom 3: (Countable additivity) If $E_1 \dots E_n$ disjoint, $P(E_1 \cup \dots E_n) = P(E_1) + \dots P(E_n)$
- Consequences of the Axioms
  - Complementary Rule
    * $E \cup E^C = \Omega \implies 1 = P(E \cup E^C) = P(E) + P(E^C)$
  - Simplest Probability Law
    * If all points in a sample space are equally likely to occur, $P(w_i) \forall w_i \in \Omega = \frac{1}{|\Omega|}$
  - Monotonicity
    * $E_1 \subseteq E_2 \implies P(E_1) \leq P(E_2)$
- Inclusion-Exclusion Theorem
  - $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$
  - Extends to a countable number of events
    * $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1) \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cup A_2 \cup A_3)$
- *DeMorgan's Law*
  - $P((A_1 \cup A_2)^C) = P(A^C \cap B^C)$
  - $P((A_1 \cap A_2)^C) = P(A^C \cup B^C)$
  - Also extends to a countable number of events
- Set operations - All still apply if you flip the $\cup$s and $\cap$s
  - Commutative Laws

        - $E_1 \cup E_2 = E_2 \cup E_1$
    - Associative Laws
        - $(E_1 \cup E_2) \cup E_3 = E_1 \cup (E_2 \cup E_3)$
    - Distributive Laws
        - $(E_1 \cup E_2) \cap E_3 = (E_1 \cap E_3) \cup (E_2 \cap E_3)$

Conditional Probability

- $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
- Law of Total Probability
    - If $B_1 \dots B_n$ partition the sample space:
    - $P(A) = \sum P(A|B_i)P(B_i)$
        - $P(A) = \sum P(A \cap B_i)$ as well, but the former is easier to calculate.
- Multiplicative Rules
    - $P(A_1 \cap A_2) = P(A_1)P(A_2|A_1)$
    - $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2$
    - Extends to a countable number of events
- Bayes' Rule: When $B_1 \dots B_m$ is a partition of Ω, and $P(A) > 0$:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{m} P(A|B_i)P(B_i)}$$

Independent Events

- $A, B \subseteq \Omega$ independent if $P(AB) = P(A)P(B)$
    - $P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$
    - (A, B) independent implies (A, B^C), (A^C, B), (A^C, B^C) independent

Counting

- Taking n objects at a time, $|\Omega| = n!$

- Matching
  - Let M_i be the event that there is a match at location i
  - $P(M_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$
    * $P(M_{i_1} \cap M_{i_2}) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$ if $i < j$
  - $\bigcup\limits_{i=1}^{n} M_i$ is the event that there is at least one match.
    * Also the inverse of there being no matches.
    * $P(\bigcap\limits_{i=1}^{n} M_i^C) \approx \frac{1}{e}$
  - P_n is the probability of no match at an index n
  - How can we compute $P(\text{exactly r matches})$?
- The Binomial Theorem
  - Suppose $n \geq 1$ is an integer and $x, y \in \mathbb{R}$. Then

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

- Partitions of Integers
  - Fix an $n \in \mathbb{Z}$ and let $r > 0$ be a fixed integer.
  - A list of r natural numbers $(x_1 \ldots x_r)$ is a partition of the integer n if $x_1 + \ldots x_r = n$
- Multinomial coefficient
  - If we have an n-element set and integers $n_1 \ldots n_r \geq 0$ that sum to n, the number of partitions of an n-element set into r disjoint subsets where the ith subset has n_i elements is

$$\binom{n}{n_1, n_2, \ldots n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}$$

  - Counts the number of *ordered subsets* with the ith subset having n_i elements.
    * If the sets were in a different order, it would be a different partition.

Random Variables

- $X : \Omega \to \mathbb{R}$ can be discrete or continuous
- If X is a discrete random variable, we will associate it with a *probability mass function (pmf)* P_X
  - $P_X(x) > 0 \ \forall \ x \in \{\text{values of } X\}$
  - $\sum_x P_X(x) = 1$, where the sum is over all the possible values of X
  - Used to calculate some probabilities: $P(X \in A) = \sum_{x \in A} P_X(x)$

Important Discrete Random Variable Distributions

- Format: $X \sim$ *distribution* if X has the pmf $p(x)$

1) The Bernoulli(p)
   - $p(x) = \{p \text{ when } x = 1, 1 - p \text{ when } x = 0\}$
   - Where $0 \leq p \leq 1$
   - For infinite populations with proportion p of successes and selecting one at random
   - Interesting note: $\mathbb{E}(X^n) = \mathbb{E}(X) = p$ for the Bernoulli
   - Example: Experiment has $P(A) = p$. Then $I_A(\omega) = \{1 \text{ if } \omega \in A, 0 \text{ if } \omega \notin A\}$ is a Bernoulli(p) random variable
2) The Binomial(n, p)
   - $p(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, \ldots n$
   - One explanation
     - $n > 0, n \in \mathbb{Z}^+$ represents either the number of trials or the sample size.
     - p represents the probability of drawing a success
   - Another: If an experiment consists of n *independent* Bernoulli trials and X counts the number of successes
     - If $X_1 \ldots X_n$ are independent Bernoulli(p) random variables, $X := X_1 + \ldots X_n$ has a binomial distribution.
3) The Geometric(p)
   - $P_X(x) = P(X = x) = (1-p)^{x-1} p$ for $x = 1, 2, 3, \ldots$
   - An experiment consists of an infinite sequence of Bernoulli(p) trials. Let X be the random variable that gives the trial of the first success.
     - Example: A geometric sequence for the 1st head to appear on trial x starts with $x - 1$ tails.
4) The Poisson(λ)
   - $P_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for $x = 0, 1, 2, \ldots$
     - $\lambda \in \mathbb{R}$
   - First arose as an approximation to the Binomial(n, p) when n is large and p is small
     - Suppose $Y \sim \text{binomial}(n, p)$ where $p = \frac{\lambda}{n}$ and n is large. $\lambda = \text{rate} = \frac{\text{\# of events}}{\text{time}}$
     - Then for \$y = 0, 1, . . . \$ that are fixed
     - $P(Y = y) \to \frac{e^{-\lambda}\lambda^y}{y!}$ as $n \to \infty$
   - In practice: Poisson(λ) random variable is a random variable that counts the number of rare events that happen in an interval of time

- Example: Probability that someone else in a room has my birthday.

5) Negative binomial(r,p) / Pascal(r,p)
   - $P_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$
     - When $r \geq 1$ is a fixed integer
   - Interpretation: X is the trial of the rth success.

6) Hypergeometric

- $$P_X(x) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$
- Describes the probability of k successes in n draws without replacement, out of K total successes and N total draws.
  - In contrast, the binomial is just $k = K$ successes out of $n = N$ draws with replacement.
- Example: 5 green and 45 red marbles in an urn. Draw 10 without replacement; what is the probability that exactly 4 of the 10 are green?
  - $k = 4, K = 5, n = 6, N = 39$

Functions of Random variables

- With pmf of X and $Y = g(X)$:

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x)$$

Expected Value and Variance of a Discrete Random Variable

- $\mathbb{E}$: also the mean, weighted average, or center of mass

$$\mathbb{E}(X) = \sum_x x P_X(x)$$

- The variance of an random variable: $Var(x) = (x-\mu)^2$, where $\mu = \mathbb{E}(X)$
  - So $\mathbb{E}(\{X-\mu\}^2) := Var(X)$
- A form of the $Var(X)$ more amenable to calculations: $\mathbf{Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2}$

Cumulative Distribution Function (CDF)

- $F(X) = P(X \leq x)$
  1) $F : \mathbb{R} \to [0, 1]$
  2) If $x < y$, $F(x) \leq F(y)$
- Notation: Left-limit notation $F(c-) = lim_{x \to c-} F(x)$
- If we know the CDF,
  - $P(a < x \leq b) = F(b) - F(a)$
  - $P(a \leq x \leq b) = F(b) - F(a-)$
  - $P(a \leq x < b) = F(b-) - F(a-)$
  - $P(a < x < b) = F(b-) - F(a)$
  - General rule: If near “$<$”, $a \to -F(a)$, $b \to F(b-)$
    * “a $<$ b” ... so if “$<$” is near a, the “-” is on the left; “-” is on the right for b

Law of the Unconscious Statistician

- If X is discrete and $G : \mathbb{R} \to \mathbb{R}$, then
  - $\mathbb{E}(\mathbf{g}(\mathbf{X})) = \sum_{\mathbf{x}} \mathbf{g}(\mathbf{x})\mathbf{P}(\mathbf{X} = \mathbf{x})$ when the expectation exists
- Used when we know $G(X)$ and the distribution of X but not the distribution of $G(X)$
- Some consequences:
  - Linearity of Expectation #1
    * $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
  - Linearity of Expectation #2
    * $\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}X_n)$
    * Expectation of a sum is the sum of the individual expected values
    * For any random variables for which $\mathbb{E}(X_i)$ exists for all i
    * *To be proved later*

Miscellaneous

- Sum of a geometric sequence

$$\frac{\text{1st term in the series}}{1 - \text{geometric ratio}}$$