IERG 2470 Tutorial for Midterm Review

Yanxiao Liu

23 March 2022

General Information

- Yanxiao Liu, 1155168298@link.cuhk.edu.hk
- Review: Chapter 1-3 sequentially with (some of) examples in lectures or questions from homework through the review.
- If you have particular questions on homeworks, please check:
  1. hw 1: ask Gongpu, cg019@ie.cuhk.edu.hk;
  2. hw 2: ask me;
  3. hw 3: ask Xiaohong, cx021@ie.cuhk.edu.hk;
  4. hw 4: ask Guodong, xg018@ie.cuhk.edu.hk;
- The slide and recording will be on blackboard today.
- My suggestion: although it's open-book, prepare a cheat sheet, it is the best way to review all the contents by yourselves.

Review for Sec 1.1: Preliminaries

- A **statement** can be True or False.
- Logic Symbols: AND/OR/NOT/Implication$\Rightarrow$/Equivalence$\Leftrightarrow$
  The only case that $X \Rightarrow Y$ is False is: X is True and Y is False.

Establish that the statement "$x \geq 2 \Rightarrow x \geq 0$" is True.

- Since x is not specified, interpret it as any real number.
- Verify the statement for all real numbers x.

for	$x \geq 2$	$\Rightarrow$	$x \geq 0$
$x < 0$	F		F
$0 \leq x < 2$	F		T
$x \geq 2$	T		T

- De Morgan's law:

$$\sim (X \vee Y) \Leftrightarrow (\sim X) \wedge (\sim Y)$$
$$\sim (X \wedge Y) \Leftrightarrow (\sim X) \vee (\sim Y)$$

- $\sim (\forall x, Y(x)) \Leftrightarrow \exists x, \sim Y(x)$

Review for Sec 1.2: Set Theory

- A set is a collection of objects. Ω, $\emptyset$
- Union, Intersection, Complement, Difference, Set inclusion
- Key to not make mistake: always draw Venn diagrams!
- De Morgan's law:

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

- Distribution laws:

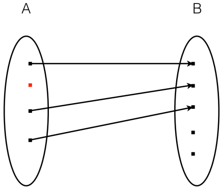
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Review for Sec 1.3: Relation and function

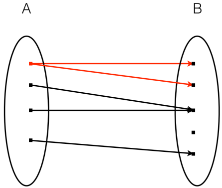
- A relation between two sets A and B is a subset of $A \times B$ s.t.

$$A \times B = \{(x, y) : x \in A, y \in B\}$$

- A function $f : A \to B$ is a relation between A and B such that every $x \in A$ is associated with a unique element in B, denoted by $f(x)$.



not a function



not a function

- injection, surjection, bijection.

Review for Sec 1.6: Different Proof Methods

- Direct proof;
- Prove by contradiction.
- Prove by induction(next page).

Review for Sec 1.7: Mathematical induction

- Mathematical induction:
  1. Base case, usually $P(n)$ for $n = 1$.
  2. Assume $P(n)$ is true for $n > 1$.
  3. Prove $P(n+1)$ is true.
- Inclusion-exclusion formula:

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$$

$$|A_1 \cup \cdots \cup A_n| = \sum_{1 \le i \le n} |A_i| - \sum_{1 \le i < j \le n} |A_i \cap A_j| + \cdots + (-1)^{n+1} |A_1 \cap A_2 \cap \cdots \cap A_n|$$

- Set-additive function: $\mu : B \mapsto \mathbb{R}$ for $B \subset \Omega$ s.t. $\forall B \cap B' = \emptyset$,

$$\mu(B \cup B') = \mu(B) + \mu(B').$$

Eg: cardinality is a set-additive function.

IERG 2470 Homework 2

- (c): Prove the binomial formula by mathematical induction.
- Binomial formula:

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}$$

- Base case, $n = 1 \Rightarrow RHS = \sum_{r=0}^{1} \binom{1}{r} a^r b^{1-r} = a + b = LHS$
- Suppose when $n = k$, $(a+b)^k = \sum_{r=0}^{k} \binom{k}{r} a^r b^{k-r}$. When $n = k+1$,

$$(a+b)^{k+1} = (a+b)^k (a+b)$$

$$= \sum_{r=0}^{k} \binom{k}{r} a^{r+1} b^{k-r} + \sum_{r=0}^{k} \binom{k}{r} a^r b^{k+1-r} = \cdots$$

$$= \binom{k+1}{0} a^0 b^{k+1-0} + \sum_{r=1}^{k} \binom{k+1}{r} a^r b^{k+1-r} + \binom{k+1}{k+1} a^{k+1} b^0 = \cdots$$

$$= \sum_{r=0}^{k+1} \binom{k+1}{r} a^r b^{k+1-r}$$

2. (Horses again!) Someone claims to have proved by mathematical induction that all the horses in the world have the same color. Here is his formulation. Let n be the number of horses in any group of horses. For all $n \geq 1$, the proposition is that all the horses in the group have the same color. Here is the proof:

   (a) The proposition is obviously true for $n = 1$.

   (b) Assume that the proposition is true for some $n \geq 1$. Now consider any group of $n+1$ horses. By the induction hypothesis, Horses 1 to n, which is a group of n horses, have the same color. Again by the induction hypothesis, Horses 2 to $n+1$ have the same color. Since Horses 2 to n are common to both groups of n horses, the two groups of n horses must have the same color. Therefore, we conclude that Horses 1 to $n+1$ all have the same color.

   Are you convinced?

- Hint: consider $n = 1$ to $n = 2$.

Review for Sec 1.8: Combinatorics

- Factorial: number of ways to order n balls is $n!$.
- Permutation: pick k balls in an ordered manner from n ball:

$$P(n,k) = \frac{n!}{(n-k)!}$$

- Combination: pick k balls in an unordered manner from n ball:

$$C(n,k) = \frac{P(n,k)}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

- binomial formula:

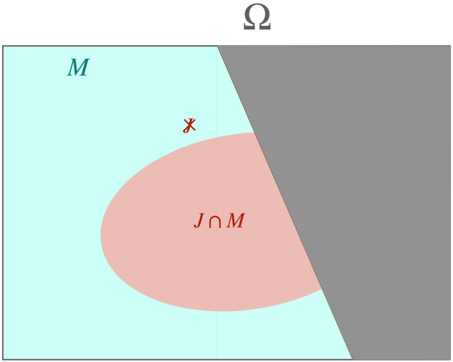
$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}$$

Review for Sec 2.1: Probability and Events

- sample space Ω: all possible outcomes.
- set function P: $0 \leq P(E) \leq 1$ for $\forall E \subset \Omega$; $P(\Omega) = 1$; $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$. axioms of probability.
- outcome: $\omega \in \Omega$.
- event: $E \subset \Omega$, occurs if $\omega \in E$.
- Probability: $P(E)$: probability that event E occurs.
- Corollary:
  1. $P(\emptyset) = 0$.
  2. $P(A^C) = 1 - P(A)$
  3. $P(A) \leq P(B)$ if $A \subset B$.

Review for Sec 2.2: Probability as a state of knowledge

- Conditioning:



- Let $f(\cdot)$ denote fraction,

$$f(J|M) = \frac{|J \cap M|}{|M|} = \frac{f(J \cap M)}{f(M)}$$

Review for Sec 2.3: Probability as a state of knowledge

- The probability of event A conditioning on event B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Review for Sec 2.4: The law of total probability and the Bayes theorem

- A collection of sets $\{B_i\}$ is a partition of Ω if

$$\cup_i B_i = \Omega$$
$$B_i \cap B_j = \emptyset \text{ if } i \neq j$$

- The law of total probability:

$$P(A) = \sum_i P(A|B_i)P(B_i)$$

- Example: $\Omega = \{\text{students}\}$, $M = \{\text{male}\}$, $F = \{\text{female}\}$. Check it is a partition! Let $J = \{\text{student: wears jacket}\}$, then

$$P(J) = P(J|M)P(M) + P(J|F)P(F)$$

Review for Sec 2.4: The law of total probability and the Bayes theorem

- Bayes theorem: $P(A), P(B) > 0$,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Proof:

$$P(A \cap B) = P(A)P(B|A) = P(A|B)P(B)$$

- Corollary: $\{B_i\}$ as a partition of Ω, $\forall A$,

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}$$

Review for Sec 2.4: The law of total probability and the Bayes theorem

- Example: $P(L) = 0.2$, $P(F|L) = 0.8$, $P(F|L^c) = 0.15$.

$$P(L) = 0.2$$
$$P(F|L) = 0.8$$
$$P(F|L^c) = 0.15$$

- By Corollary 2.18,

$$\begin{aligned} P(L|F) &= \frac{P(F|L)P(L)}{P(F|L)P(L) + P(F|L^c)P(L^c)} \\ &= \frac{P(F|L)P(L)}{P(F|L)P(L) + P(F|L^c)(1 - P(L))} \\ &= \frac{(0.8)(0.2)}{(0.8)(0.2) + (0.15)(1 - 0.2)} \\ &= \frac{4}{7} \end{aligned}$$

- Also we have $P(L^c|F) = 1 - P(L|F) = \dfrac{3}{7}$.

Review for Sec 2.5: Independent events

- Two events and are independent if $P(A \cap B) = P(A)P(B)$.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

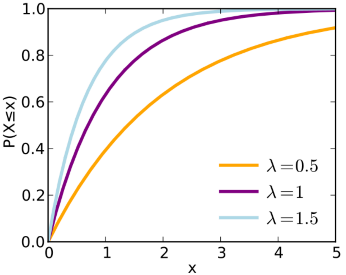
- Three events A, B, C are mutual independent if they are
  1. pairwise independent
  2. $P(A \cap B \cap C) = P(A)P(B)P(C)$

- A random variable X is a function of ω. $X : \Omega \to \mathbb{R}$
- A random variable X is called discrete if the set of all values taken by X is discrete.
- X is characterized by a pmf that gives the probability of occurrence of each possible value of $X(\omega)$. pmf $\{p_i\}$ satisfies: $p_i \leq 0$, $\sum_i p_i = 1$.
- Binomial distribution: $p_i = \binom{n}{i} p^i (1-p)^{n-i}$.
- Poisson distribution: $\lambda \geq 0$, $p_k = \frac{e^{-\lambda}\lambda^k}{k!}$ if $k \geq 0$ and $p_k = 0$ otherwise.

Review for Sec 3.2: Cumulative distribution function

- The CDF of a random variable X is defined by $F_X(x) = P(X \leq x) = P(-\infty < X \leq x)$
- It gives weight of the left part of the wire up to and including point x.
- $F_X(x)$ is non-decreasing and right-continuous
- For any interval $(a, b]$, $P(X \in (a, b]) = F(b) - F(a)$

- $\mathscr{E}(\lambda)$: exponential distribution with parameter λ, where $\lambda \geq 0$
- $F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$



Review for Sec 3.3: Probability density function

- We want $f(x)$ (not necessarily unique) to satisfy

$$F(x) = \int_{-\infty}^{x} f(u)du$$

- Thus let $f(x) = F'(x)$
- For a random variable X with a pdf, we have $P(X = x) = 0, \forall x$.

Review for Sec 3.4: Function of a random variable

- Consider a random variable X, let $Y = g(X)$, Y is also a random variable.
- Linear transformation: let $g(x) = ax + b$, $Y = g(X)$, $a > 0$, we find

$$F_Y(y) = F_X(\frac{y-b}{a})$$
$$f_Y(y) = \frac{1}{a} f_X(\frac{y-b}{a})$$

Review for Sec 3.5: Expectation

- If X is discrete, expectation of X is

$$E[X] = \sum_{x \in \mathcal{X}} xp(x)$$

- If X is continuous, expectation of X is

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

- $E(X)$ is the value taken by X on the average, also called *mean*.
- Binomial distribution:

$$E[X] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np$$

- Exponential distribution:

$$E[X] = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

Review for Sec 3.5: Expectation

- The expectation of a function g of a random variable X, if X is discrete:

$$Eg(X) = \sum_{x \in \mathcal{X}} g(x)p(x)$$

If X is continuous:

$$Eg(X) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

Review for Sec 3.6: Moments

- n-th moment: $m_n = E[X^n]$
- n-th central moment: $\mu_n = E[(X - EX)^n]$.
- m_1 is mean or expectation
- μ_2 is called the variance: $\text{Var}(X) = E[(X - EX)^2] = E[X^2] - [EX]^2$
- The Gaussian distribution $\mathcal{G}(m, \sigma^2)$ is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

has mean m and variance σ^2

Review for Sec 3.7: Jensen's inequality

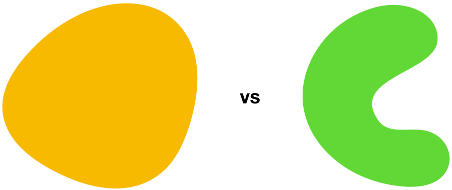
- a continuous random variable X:

$$Eg(X) = \int g(x)f(x)dx$$

- For linear combination $y = \sum_i \alpha_i x_i$, it is called a convex combination of $\{x_i\}$ if
  1. $\alpha_i \geq 0, \forall i$
  2. $\sum_i \alpha_i = 1$
- A set S s convex if $\forall x, y \in S$,

$$(\alpha x + (1 - \alpha)y) \in S,$$

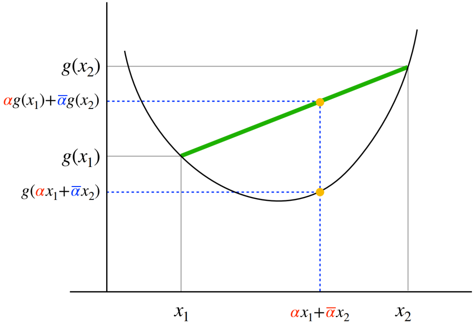
for any $0 \leq \alpha \leq 1$.



Review for Sec 3.7: Jensen's inequality

- A function g defined on a convex set S is convex if $\forall x, y \in S$, $\forall \alpha \in [0, 1]$,

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha)y$$

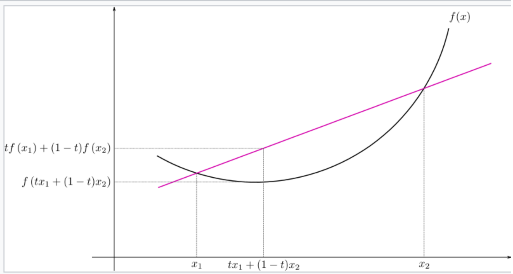


$$g(\alpha x_1+\bar{\alpha} x_2) \leq \alpha g(x_1)+\bar{\alpha} g(x_2)$$

Review for Sec 3.7: Jensen's inequality

- A function g defined on a convex set S is convex if $\forall x, y \in S$, $\forall \alpha \in [0, 1]$,

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha)y$$

- 

- Jensen's inequality: let g convex, then

$$E[g(X)] \geq g(EX)$$

Thank you and good luck to your midterm!