Stat 609: Mathematical Statistics I (Fall Semester, 2016) Introduction

Course information

Instructor	Professor Jun Shao	TA	Mr. Han Chen
Office	1235A MSC		1335 MSC
Phone	608-262-7938		608-263-5948
Email	shao@stat.wisc.edu		hanchen@stat.wisc.edu
Office hours	???		???
	or by appointment		

Lecture	1:00-2:15 MW (SMI 331)
Discussion	1:00-2:15 F (CHAMBERLIN 2112)
	2:30-3:45pm F (VILAS 4014)
	Discussions taught by TA

Textbook

Book	Statistical Inference, 2nd edition
Authors	George Casella and Roger L. Berger
Publisher	Brooks/Cole Cengage Learning
Topics covered	Chapters 1-6
Syllabus	See the list of topics and schedule
Recommended reading books	Mathematical Statistics (1977, Bickel and Doksum)
	An Introduction to Probability Theory and Mathematical Statistics (1976, Rohatgi)

Homework Assignments

- In each lecture, there are 2-4 homework problems selected from the textbook.
- Problems are assigned in groups with a specified due date.
- Some or all problems are graded by the TA and a solution will be provided.

Exam Schedule (may be changed)

1st exam	1:00-2:15pm on Oct 12, 2016
2nd exam	1:00-2:15pm on Nov 16, 2016
Final exam	2:45-4:45pm on Dec 22, 2016?
Note	All exams are closed-book
	You may bring 1 sheet (2 pages) of notes

Course Grades

Assignments	20 points	Grades	A (90-100)
1st exam	20 points		AB (80-89)
2nd exam	20 points		B (70-79)
Final exam	40 points		BC (60-69)
Total	100 points		C(50-59) ······

Adjustments may be made

Course Material

Covered topics, schedules, homework assignments, and lecture slides can be found in my website http://www.stat.wisc.edu/˜shao/

Required Mathematics

Linear algebra

- Vectors $c = (c_1, ..., c_k)$, row or column, transpose c'
- Matrices A, its dimension, transpose A'
- Vector/matrix addition and multiplications: $c'A$, $A+B$, AB, ...
- Positive and non-negative definite matrices and their simple properties.

Calculus, advanced calculus or mathematical analysis

- Functions (univariate or multivariate), domain and range
- Limits (as $n \to \infty$ or as $x \to x_0$), ε-N or ε-δ arguments
- Continuity and differentiability of functions
- Right-continuous, left-continuous
- Right-derivative, left-derivative
- Integration with finite or infinite integral limits
- Existence of an integral: e.g., $\int_{-\infty}^{\infty} |g(x)| dx = \infty$ or a finite number?

Calculus, advanced calculus or mathematical analysis

- Series: when does $\sum_{n=1}^{\infty} a_n$ converge?
- Geometric series:

$$\frac{1}{1-x} = \sum_{t=0}^{\infty} x^t \qquad |x| < 1$$

- Power series:

$$e^x = \sum_{t=0}^{\infty} \frac{x^t}{t!} \qquad |x| < \infty$$

- Infinity
  - For any real number x, $\infty + x = \infty$, $x\infty = \infty$ if $x > 0$, $x\infty = -\infty$ if $x < 0$, and $0\infty = 0$;
  - $\infty + \infty = \infty$;
  - $\infty^a = \infty$ for any $a > 0$;
  - $x/0 = \infty$ if $x > 0$ and $x/0 = -\infty$ if $x < 0$;
  - $\infty - \infty$ or ∞/∞ or $0/0$ is not defined.

Mathematical induction

Prove a statement involving n is true for any fixed $n = 1, 2, 3, ...$
(but not $n \to \infty$)

Knowledge in probability and mathematical statistics

- It is not required, but is very helpful if you took a course in probability and statistics previously
- Sets, subsets and set operations

Purposes of Stat 609-610

- Learning statistical concepts and knowledge (together with Stat 601-602 or 671-672 sequence)
- Training at the level of the Mater of Science in Statistics
  - Stat 609: Basic probability tools useful in statistics (Chapters 1-5) and data reduction principles (Chapter 6)
  - Stat 610: Basic topics in statistical inference (Chapters 7-11)
- Training in derivations and simple proofs
- Preparation for a PhD level study in statistics

Suggestions for studying

- Take class notes wisely.
- Do not be afraid of asking questions to the instructor and TA, either during or after the class.
- Read the relevant part of the textbook carefully after each lecture.
- Do the related homework problems after each lecture. Do not delay the work until the due date or the day before the due date.
- Do homework problems independently. Group discussions are encouraged, but you should do your own work.
- If you have time, try to do more problems from the textbook, which contains many good problems but I don't want to assign all.
- Study the graded problems from the TA, and the solution. It is a good idea to study a different solution to a (complicated) problem.
- Review the material before each exam (there will be a review class prior to each exam) and carefully prepare the one sheet notes for each exam.
- Menage the time wisely in the exam.

Random experiment

An experiment (in a general sense) that results in one of more than one outcomes with uncertainty in which outcome will be the realization

Examples of random experiments

- Tomorrow's weather: Sunny, Cloudy, Rainy
- Tomorrow's temperature: $(-M, M)$
- Number of car accidents in a city: 0, 1, 2, 3, ...
- The life time of a television set: $(0, \infty)$
- Toss a coin: Heads or Tails?

Statistical inference

- Treat data from the experiment as a sample from some population
- Make inference about the population
- Find things having regular patterns, not appear by chance

Probability theory

A basic tool for statistics, which explains a random experiment in terms of frequencies and their characteristics

For example, some past observations (data) may tell us that tomorrow's weather is sunny with 80%, cloudy with 15%, and rainy with 5%.

To study statistical inference, we must have a very good knowledge in probability theory.

Definition 1.1.1. Sample space (or outcome space)

The set S containing all possible outcomes is called the sample space

Example

$S = \{Heads, Tails\}, = \{0,1\}, = \{0,1,2,3,...\}, = (-M,M), = \mathscr{R}$ (all real numbers), $= (0,\infty)$

Discrete sample space: S contains integers (finitely or infinitely many)

Continuous sample space: e.g., $S = [0,1]$, $\mathscr{R}$, or $\mathscr{R}^2$.

Subsets or sets

In probability and statistics, we are interested in subsets of a given sample space S (sometime called sets for simplicity).

$A \subset S$: A is a subset of S.

Example: the life time of a television set is between 0 and 100 is a subset $A = (0, 100]$ of the sample space $S = (0, \infty)$.

Two very special subsets:

- Empty set: $\emptyset$ is the subset of S containing no element.
- The sample space S itself is a subset of S (the notation $A \subset S$ allows the special case of $A = S$).

Expressions of subsets

A subset is expressed in words or math formulas inside of $\{ \ \ \}$.

Example:
$A = \{x : x \text{ satisfies some condition}\}$
$A = \{x : x \in \mathscr{R} \text{ and } x \geq 0\} = [0, \infty)$

Relationship

- $A \subset B$: $x \in A \Rightarrow x \in B$
- $A = B$: $A \subset B$ and $B \subset A$
- Disjoint: A and B have no common elements

Example

$A = \{\text{all even integers}\}$
$B = [2, \infty)$

Obviously, for any $x \in A$, $x \geq 2$ and thus $x \in B$. Hence $A \subset B$.

A more complicated example

Let

$$A = \{x \in \mathscr{R}^k : x'x \leq c\}$$

where $c > 0$ is a given constant, x' is the transpose of x, and $a'b$ is the usual vector product, and

$$B = \left\{x \in \mathscr{R}^k : (y'x)^2 \leq cy'y \ \text{ for all } y \in \mathscr{R}^k\right\}$$

We want to show $A = B$.

A more complicated example

First, we show $A \subset B$.
For any $x \in A$, $x'x \leq c$ and by the Cauchy-Schwartz inequality,

$$(y'x)^2 \leq x'xy'y \leq c(y'y)$$

for all $y \in \mathscr{R}^k$.
This means $x \in B$ and thus $A \subset B$.

Next, we show $B \subset A$.
If $x \in B$,

$$(y'x)^2 \leq cy'y \quad \text{for all } y \in \mathscr{R}^k$$

In particular, $y = x$ should satisfy this inequality, which means

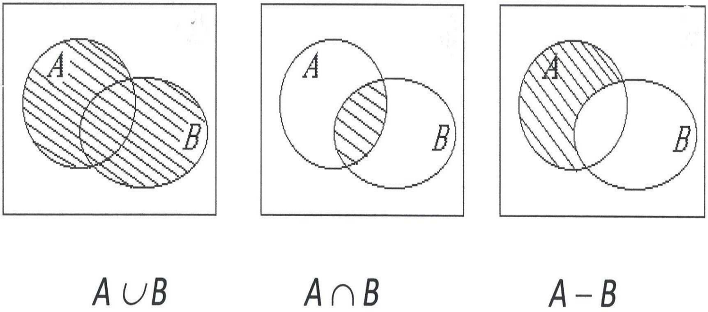
$$(x'x)^2 \leq cx'x \quad \Rightarrow \quad x'x \leq c$$

This means $x \in A$ and thus $B \subset A$.

Set operation

- The union of A and B: $A \cup B = \{x \colon x \in A \text{ or } x \in B\}$
- The intersection of A and B: $A \cap B = \{x \colon x \in A \text{ and } x \in B\}$
  Disjoint A and B: $A \cap B = \emptyset$ (the empty set)
- The complement of A: $A^c = \{x : x \notin A\}$, $S^c = \emptyset$
- Difference of A and B: $A - B = A \cap B^c$

Venn Diagram



Properties

- $A \cap \emptyset = \emptyset$
- $A \cup \emptyset = A$
- $A \cap S = A$
- $A \cup S = S$
- $\emptyset \subset A \cap B \subset A$ or $B \subset A \cup B \subset S$
- $A \subset B \Rightarrow A \cup B = B$ and $A \cap B = A$
- $A - B = A - A \cap B$
- $A = (A \cap B) \cup (A \cap B^c)$
- $A \cup B = A \cup (B - A) = A \cup (B - A \cap B)$
- $\cup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \cdots = \{x : x \in A_i \text{ for some } i\}$
- $\cap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \cdots = \{x : x \in A_i \text{ for all } i\}$
- $B \cap \left(\cup_{i=1}^{\infty} A_i\right) = \cup_{i=1}^{\infty} (B \cap A_i)$
- $B \cup \left(\cap_{i=1}^{\infty} A_i\right) = \cap_{i=1}^{\infty} (B \cup A_i)$
- DeMorgan's laws: $\left(\cup_{i=1}^{\infty} A_i\right)^c = \cap_{i=1}^{\infty} A_i^c$ and $\left(\cap_{i=1}^{\infty} A_i\right)^c = \cup_{i=1}^{\infty} A_i^c$

Subset occurrence

If x in a subset A is the realization of an experiment, then we say that A occurs.

Let A, B, C be 3 subsets.
Their occurrence can be expressed by the following subsets:

- A occurs: A
- All subsets occur: $A \cap B \cap C$
- At least one subset occurs: $A \cup B \cup C$
- Only A occurs: $A \cap B^c \cap C^c$
- Exactly one subset occurs: $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$
- None occurs: $A^c \cap B^c \cap C^c$
- Not all subsets occur: $A^c \cup B^c \cup C^c$
- At least two subsets occur: $(A \cap B) \cup (A \cap C) \cup (B \cap C)$
- At most one subset occurs: $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C) \cup (A^c \cap B^c \cap C^c)$

Definition 1.2.1

A collection $\mathscr{F}$ of subsets of a sample space S is called a σ-field (or σ-algebra) iff it has the following properties:
(i) The empty set $\emptyset \in \mathscr{F}$;
(ii) If $A \in \mathscr{F}$, then the complement $A^c \in \mathscr{F}$;
(iii) If $A_i \in \mathscr{F}$, $i = 1, 2, \ldots$, then their union $\cup A_i \in \mathscr{F}$.

- If $A \in \mathscr{F}$, then A is called an event.
- Why do we need to consider σ-fields?
  When S is continuous, there are some “bad" subsets for which a logical definition of measure or probability is not possible.
- $\mathscr{F}$ contains good subsets we deal with, but they have to satisfy (i)-(iii).
- If S is discrete, we do not need this concept, i.e., $\mathscr{F} =$ all subsets of S.
- In $\mathscr{R}^k$, the k-dimensional Euclidean space ($\mathscr{R}^1 = \mathscr{R}$ is the real line), we consider $\mathscr{F}$ to be the smallest σ-field containing all open sets, and sets in $\mathscr{F}$ are called Borel sets.

How to define the probability of an event A?

- Intuitive definition: the rate or percentage of the occurrence of A
- Statistical definition: if we repeat the same experiment M times, the probability of $A =$(the number of times A occurs)$/M$ when M is sufficiently large.
- Mathematical definition: probability is a special measure.
- A measure is an abstract extension of length, area, volume, ...

Definition 1.2.4

A set function ν defined on a σ-field $\mathscr{F}$ is called a *measure* iff it has the following properties.
(i) $0 \leq \nu(A) \leq \infty$ for any $A \in \mathscr{F}$.
(ii) $\nu(\emptyset) = 0$.
(iii) If $A_i \in \mathscr{F}$, $i = 1, 2, \ldots$, and A_i's are disjoint, i.e., $A_i \cap A_j = \emptyset$ for any $i \neq j$, then

$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \nu(A_i).$$

If $\nu(S) = 1$, then ν is a probability and we use notation P instead of ν.