STSCI 5080
Probability Models and Inference

Lecture 1: Introduction

August 23, 2018

Instructor, TA, and grader

- Instructor: Kengo Kato
  Associate Professor, Department of Statistical Science
  E-mail: kk976@cornell.edu
  Office: Comstock Hall 1194

- Teaching Assistant: Yichen Zhou

- Grader: Huijie Feng

Enrollment PIN

- To enroll in STSCI 5080, you need your own enrollment PIN.
- If you have not received your PIN, please contact me or Ms. Beatrix Johnson (bj11@cornell.edu) as soon as possible.

STSCI 5080

This course is about Probability and Statistics. In particular:

- We will study important concepts and results in probability theory, such as random variables, standard distributions, the law of large numbers, and the central limit theorem.
- We will study core statistical ideas: estimation, testing, and confidence intervals. Special emphasis is on the maximal likelihood method.

Logistics

- Lectures: Tu/Th 2:55pm – 4:10pm, Olin Hall 255
- Office Hours (instructor): Tu 4:30pm – 6:30pm
- Discussion: Fr 2:55pm – 4:10pm, Upson 216
- Office Hours (TA): W 10am – 12pm?
- Course materials are posted on Blackboard.

Grade (tentative!)

- Homework: 5 homework assignments (16%). Lowest score dropped.
- Midterms: 9/25 (Tu) and 10/30 (Tu), in-class, 1 hour (20% each). Closed-books and closed-notes. Cheatsheet (1 page letter size paper).
- Final: TBD (44%). Open-books and open-notes.

Homework

- Homework assignments have to be submitted on paper by each individual student in-class.
- No late homework is accepted.
- You may work on homework with your classmates, but write up your solutions individually.

Laptop use

- You may use your laptop or tablet for note taking.
- No sounds. No Youtube.
- No electronic devices in the exams.

Miscellaneous

- Prerequisite: BTRY 3010, Calculus II, or the equivalent
- Textbook:
  Rice, J.A. (2013). *Mathematical Statistics and Data Analysis*, 3rd Edition, Brooks/Cole.
- Other references:
  DeGroot, M.H. and Schervish, M.J. (2012). *Probability and Statistics*, 4th Edition, Pearson Education.
  Bertsekas, D.P. and Tsitsiklis, J.N. (2008). *Introduction to Probability*, 2nd Edition, Athena Scientific
- Lecture slides will be posted on Blackboard after each lecture.

Topics to be covered

- Probability (Chapter 1)
- Discrete and Continuous Random Variables (Chapter 2)
- Joint Distributions (Chapter 3)
- Expected Values (Chapter 4)
- Limit Theorems (Chapter 5)
- Sampling Distributions (Chapter 6)
- Parametric Estimation (Chapter 8)
- Hypothesis Testing (Chapter 9)

Why Probability and Statistics?

Probability and Statistics

- Statistics can help you make decisions based on data.
- Statistics thinks that the data are random, and tries to extract information from the data by controlling the randomoness.
- Probability theory is about randomness, and that's why we should learn (some) probability before statistics.

Example: Postponement of death

- There is a theory that people can postpone their death until after an important event.
- To test the theory, Phillips and Smith[1] (1990) collected data on deaths around some (important!) festival for a certain group of people.
- Of 103 deaths, 33 died the week before the festival and 70 died the week after.

[1] D.P. Phillips and D.G. Smith. (1990). “Postponement of death until symbolically meaningful occasions”. *JAMA* **263** 1947-1951.

- So,

$$\frac{\text{number of persons who died after the festival}}{\text{total sample size}}$$
$$= \frac{70}{103} = 0.68...$$

- Do you think the theory is true?

Potential thoughts

- Thought 1: Yes! Because about 70% of them died after the festival.

Potential thoughts

- Thought 1: Yes! Because about 70% of them died after the festival.

- Thought 2: Wait a minute! Compare this with coin tossing. Suppose you toss a coin 21 times. You know the probability of Head is $1/2$. But we may only have 7 heads in 21 trials, and

$$\frac{\text{number of heads}}{\text{number of trials}} = \frac{7}{21} = 0.33...$$

  So the number 68% might be just a matter of chance.

Potential thoughts

- Thought 1: Yes! Because about 70% of them died after the festival.

- Thought 2: Wait a minute! Compare this with coin tossing. Suppose you toss a coin 21 times. You know the probability of Head is $1/2$. But we may only have 7 heads in 21 trials, and

  $$\frac{\text{number of heads}}{\text{number of trials}} = \frac{7}{21} = 0.33...$$

  So the number 68% might be just a matter of chance.

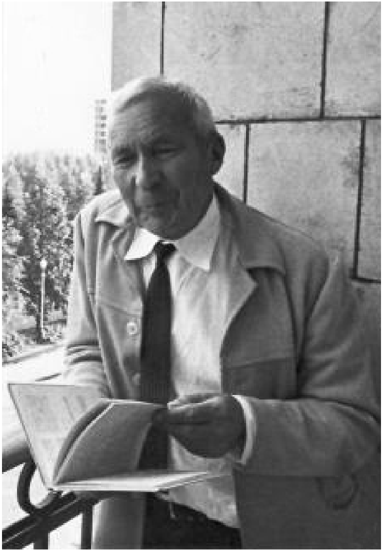
- You can answer this question at the end of the course!

Chapter 1 Probability

Some history

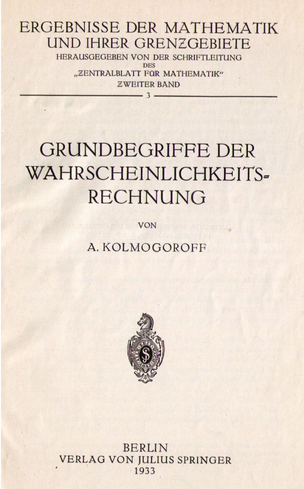
- History of mathematical theory of Probability goes back to Pascal and Fermat in the 17th century.
- Mathematical foundations of Probability were shaky until Kolmogorov in 1933.
- Kolmogorov "defines" Probability as a function of sets that satisfies three axioms.
- Kolmogorov's foundation of Probability is so general and so powerful. The "standard" theory of Probability.

A.N. Kolmogorov (1903-1987)



Taken from Wikipedia

Kolmogorov “Foundations of Probability Theory” (in German)



ERGEBNISSE DER MATHEMATIK
UND IHRER GRENZGEBIETE
HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK“
ZWEITER BAND
3

GRUNDBEGRIFFE DER
WAHRSCHEINLICHKEITS-
RECHNUNG

VON
A. KOLMOGOROFF

BERLIN
VERLAG VON JULIUS SPRINGER
1933

Taken from Wikipedia

Sample spaces

- Probability will be defined for subsets of a sample space.
- A sample space is the collection of all possible outcomes of some experiment.
- Ω: a sample space.
- $\omega \in \Omega$: an element of Ω.
- Subsets of Ω are called events.

Sample spaces: Examples

Example

- Experiment: tossing a coin three times.
- The sample space is $\Omega = \{hhh, hht, hth, thh, htt, tht, tth, ttt\}$.

Example

- Experiment: rolling a dice.
- The sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Example

- Experiment: measuring (in hours) the lifetime of a laptop.
- The sample space is $\Omega = \{t \mid t \geq 0\}$.

Review of set operations

- Let Ω be a sample space, and let A, B be events (subsets of Ω).
- Notation:

$$\omega \in A \leftrightarrow \omega \text{ is an element of } A$$
$$\omega \notin A \leftrightarrow \omega \text{ is not an element of } A$$

- Union: $A \cup B = \{\omega \mid \omega \in A \text{ or } \omega \in B\}$.
- Intersection: $A \cap B = \{\omega \mid \omega \in A \text{ and } \omega \in B\}$.
- Complement: $A^c = \{\omega \mid \omega \notin A\}$.
- Empty set: $\emptyset$ is the set with no elements.
- If $A \cap B = \emptyset$, we say that A and B are disjoint.

Subsets

- A is a subset of B if and only if any element of A is also an element of B.
- Notation for subsets: $A \subset B$ means that A is a subset of B.
- Useful trick to verify that $A = B$: check $A \subset B$ and $B \subset A$.

Example

Let $\Omega = \{1, 2, 3, 4, 5, 6\}$, and let $A = \{1, 2\}$ and $B = \{2, 4\}$.

- Union: $A \cup B = \{1, 2, 4\}$
- Intersection: $A \cap B = \{2\}$.
- Complement: $A^c = \{3, 4, 5, 6\}$.
- A and B are not disjoint, but A and $C = \{3, 4\}$ are disjoint:

$$A \cap C = \emptyset.$$

Union and intersection of n events

- Union:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup \cdots \cup A_n = \{\omega \mid \omega \in A_i \text{ for some } i = 1, \ldots, n\}.$$

- Intersection:

$$\bigcap_{i=1}^{n} A_i = A_1 \cap \cdots \cap A_n = \{\omega \mid \omega \in A_i \text{ for all } i = 1, \ldots, n\}.$$

Properties of sets and operations with sets

- $\emptyset \subset A$ for any event $A \subset \Omega$.
- $(A^c)^c = A$.
- $\Omega^c = \emptyset, \ \emptyset^c = \Omega$.
- Commutative laws:

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

- de Morgan's laws:

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

- If $A \subset B$, then
$$A \cap B = A, \ A \cup B = B.$$
In particular,
$$A \cap \Omega = A, \ A \cup \Omega = \Omega \quad \text{and} \quad A \cap \emptyset = \emptyset, \ A \cup \emptyset = A.$$
- Partition of sets:
$$B = (B \cap A) \cup (B \cap A^c),$$
where $B \cap A$ and $B \cap A^c$ are disjoint. In particular,
$$\Omega = A \cup A^c \quad \text{and} \quad A \cap A^c = \emptyset.$$

- Associative laws:

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

- Distributed laws:

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

Draw Venn's diagram.