Statistics

Review of Probability Model

Shiu-Sheng Chen

Department of Economics
National Taiwan University

Fall 2019

Section 1

Probability Theory

Probability Theory

Definition (Random Experiments)

The basic notion in probability is that of a random experiment: an experiment whose outcome cannot be determined in advance, but which is nevertheless subject to analysis.

Examples:

- Tossing a die and observing its face value.
- Choosing at random ten people and surveying their income level.

Probability Theory

Basic concepts:

- Sample spaces
- Events
- Probability measure

Probability Theory

Definition (Sample Space/State Space)

The set, Ω, of all possible outcomes of a particular experiment is called the sample space for the experiment.

(1) Roll a die (discrete and finite) $\Omega = \{1, 2, 3, 4, 5, 6\}$

(2) Flip a coin until a head appears (discrete and infinitely countable)

$$\Omega = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$$

(3) The height of a randomly selected student (continuous):[1]

$$\Omega = \mathbb{R}^+ = [0, \infty)$$

[1]Notice that for modeling purpose, it is often easier to take the sample space larger than is necessary.

Probability Theory

Definition (Event)

An event, denoted by E, is just a subset of Ω.

(1) Roll a die

$$E = \{\text{the event of odd numbers}\} = \{1, 3, 5\} \subset \Omega$$

(2) Flip a coin until a head appears

$$E = \{\text{the event of at most two tails}\} = \{H, TH, TTH\} \subset \Omega$$

(3) The height of a randomly selected student

$$\begin{aligned} E &= \{\text{the event that a student is shorter than 150cm}\} \\ &= \left\{x \middle| x \in [0, 150)\right\} \subset \Omega \end{aligned}$$

What is Probability?

- Probability is a mathematical language for quantifying uncertainty.
  - To answer the question "how likely is it...?"
- To put it loosely, probability is a number between 0 and 1, where:
  - a number close to 0 means not likely
  - a number close to 1 means quite likely
- How to assign probability?
  (A) The classical approach
  (B) The relative frequency approach
  (C) The subjective approach

(A) The Classical Approach

- Principle of Indifference: every outcome is equally likely to occur.

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$$

- Examples:
  - Roll a six-side die
- It is the interpretation identified with the works of Jacob Bernoulli and Pierre-Simon Laplace.

(B) The Relative Frequency Approach

- Some people argue that we need to further justify the assumption that "every outcome is equally likely to occur" by experience.[2]
- The relative frequency approach involves taking the follow three steps in order to determine $P(A)$, the probability of an event A:
  - Perform an experiment N times.
  - Count the number of times the event A of interest occurs, call the number $N(A)$.
  - Then, the probability of event A is:

$$P(A) = \lim_{N\to\infty} \frac{N(A)}{N}$$

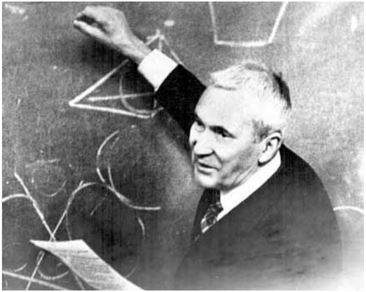
[2] Such as Richard von Mises.

(C) Subjective Approach

- The subjective approach is simply a personal opinion.
  - "I think there is an 80% chance of rain today."
  - "I think there is a 50% chance that I will get an A+ in this course"
- It is also called personal probability

Probability Models: Kolmogorov Axioms

- Now we present a probability model using the axioms of probability.
- This axiomatic approach to probability is developed by a Soviet mathematician, Andrey Kolmogorov (1903–1987).



Probability Model

Definition (Probability Measure)

A real-value function $P(\cdot)$ is a probability measure on the sample space Ω if all events $A \subseteq \Omega$ are assigned numbers $P(A)$ satisfying

(a) $P(\Omega) = 1$

(b) $P(A) \geq 0$ for all $A \subseteq \Omega$

(c) For all disjoint $A, B \subseteq \Omega$,

$$P(A \cup B) = P(A) + P(B)$$

- The pair (Ω, P) is called a probability model.
- Axiom (c) can be extended to a finite union or a countably infinite union of disjoint events.

Corollaries from Kolomogorov's Axioms

Corollary

(a) $P(A) + P(A^c) = 1$

(b) $P(\varnothing) = 0$

(c) $A \subseteq B$ *implies that* $P(A) \leq P(B)$

(d) $P(A) \leq 1$

(e) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- Rule (e) is known as the additive theorem of probability.

Conditional Probability

Definition

The conditional probability of an event A given that an event B has occurred is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ whenever } P(B) \neq 0.$$

- Die Roll

$$P(\{1\}|\text{Odd}) = P(\{1\}|\{1,3,5\}) = 1/3 > P(\{1\}) = 1/6$$

Section 3

Bayes' Theorem

Thomas Bayes

- British mathematician (1701–1761)
- He is credited with inventing Bayes Theorem

Bayes' Theorem: Motivation

- An iPhone was found to be defective (D). There are three factories (A, B, C) where such smartphones are manufactured. A Quality Control Manager (QCM) is responsible for investigating the source of found defects. Here are some information:

Factory	% of total production	Probability of defective product
A	$0.35 = P(A)$	$0.015 = P(D\|A)$
B	$0.35 = P(B)$	$0.010 = P(D\|B)$
C	$0.30 = P(C)$	$0.020 = P(D\|C)$

Q: If a randomly selected iPhone is defective, what is the probability that the iPhone was manufactured in factory C? That is, $P(C|D) =$?

Bayes' Theorem: A General Framework

- Let $A_1, A_2, \ldots, A_n \subseteq \Omega$ be a partition of Ω
- Suppose that we know
  - $P(A_i)$, which is called prior probability
  - $P(T|A_i)$, wich is called sample probability
- How to compute $P(A_i|T)$?

$$P(A_i|T) = \frac{P(A_i \cap T)}{P(T)}$$

- How to compute $P(A_i \cap T)$ and $P(T)$?
- It is easy to compute $P(A_i \cap T)$:

$$P(A_i \cap T) = P(T|A_i)P(A_i)$$

Law of Total Probability

Theorem (Law of Total Probability)

Let $A_1, A_2, \ldots, A_n \subseteq \Omega$ *be a partition of* Ω*, and* $\exists\ T \subseteq \Omega$ *with* $P(T) > 0$*. Then the probability of an event* T *can be calculated as*

$$P(T) = \sum_{j=1}^{n} P(T|A_j)P(A_j)$$

Bayes' Theorem

Theorem

$$P(A_i|T) = \frac{P(A_i \cap T)}{P(T)} = \frac{P(T|A_i)P(A_i)}{\sum_{j=1}^{n} P(T|A_j)P(A_j)},$$

where $P(A_i|T)$ is called posterior probability

- In the iPhone example,

$$\begin{aligned} P(C|D) &= \frac{P(D|C)P(C)}{P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C)} \\ &= \frac{0.020 \times 0.30}{0.015 \times 0.35 + 0.010 \times 0.35 + 0.020 \times 0.30} \\ &= 0.407 \end{aligned}$$

Section 4

Independence

Independent Events

Definition

Two events $A, B \subseteq \Omega$ are said to be independent if

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

- Hence, by the definition of conditional probability,

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