Session 1: An introduction to probability

Susan Thomas

http://www.igidr.ac.in/˜susant

susant@mayin.org

IGIDR

Bombay

The aim of this session

1. Why do we care for probability?
2. What is an outcome/event?
3. Categories of outcomes.
4. Defining probability.
5. Characteristics of probability.
6. Unconditional and conditional probability.
7. Independent events.

Motivating probability

Probability is the measure of the uncertainty of an outcome.

Outcome/Events

Outcomes/Events are observations of interest. Examples:

- an experiment: The rat dies after having been injected with the bird flu virus.
- financial markets: Nifty will go up by 10% at the end of April.
- life: By the end of the year, a million Indians will own an Amida.
- a game: Vishwanathan Anand will win the next match with Topalov in 25 moves.

Outcome/Events

Outcomes/Events are observations of interest. Examples:

- an experiment: The rat dies after having been injected with the bird flu virus.
- financial markets: Nifty will go up by 10% at the end of April.
- life: By the end of the year, a million Indians will own an Amida.
- a game: Vishwanathan Anand will win the next match with Topalov in 25 moves.

Each of these events is uncertain: probability tells us by how much they are uncertain.

Outcome/Events

Outcomes/Events are observations of interest. Examples:

- an experiment: The rat dies after having been injected with the bird flu virus. 12%
- financial markets: Nifty will go up by 10% at the end of April. 63%
- life: By the end of the year, a million Indians will own an Amida. 0.025%
- a game: Vishwanathan Anand will win the next match with Topalov in 25 moves. 27%

Each of these events is uncertain: probability tells us by how much they are uncertain.

Types of events

- Death of a rat: One out of two (binary) events.
- Nifty going up by 10%: One out of an infinite real-numbered events.
- A million Indians owning Amidas: One out of an infinite integer-numbered events.
- Vishy winning in 25 moves: Combination of two events – Vishy wins (binary events), and in 25 moves ("technically" ∞ integer events).

The event space

The first aspect of calculating the probability is defining the **event space**: this is the set of all possible events that could happen.

- The rat can either die or not.
- Nifty returns can take any value between -100% and ∞.
- Amida-owning Indians can be anywhere between 0 and 1.4 billion.
- Vishy can either win or not, and if he wins, he can do so in any number of moves between 10 and ∞.

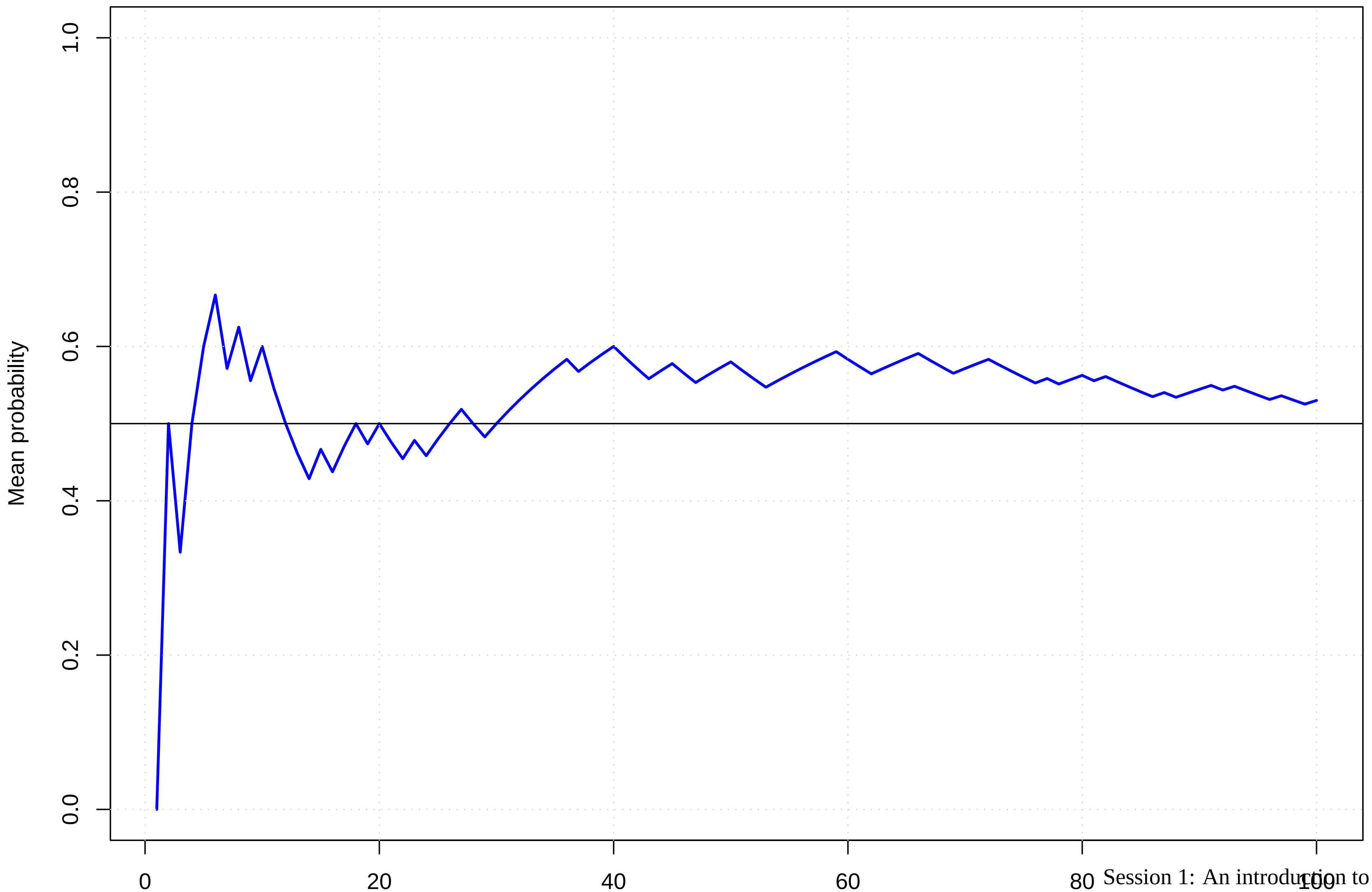
The events are either countable (rat death or Amida-owning Indians) or not (all possible values of changes in Nifty).

Defining probabilities

Once we know the event space,

- The probability of a particular event is how frequently the event appears as a fraction of how often all the other events in the event space occurs.
- The problem: we never know how frequently all the events occur – we have to simulate it, or model it.
- For example, the fact that the probability of a 50% chance of observing a head in a toss of a "fair" coin has been ascertained empirically by tossing a coin a very large number of times.

Results from a simulation of 100 tosses of a coin



Assumptions about probabilities

In the rest of this session,

- we will either assume that the events in the event space have equal probability, or
- that the probability of the events will be given exogenously.

Simplest set of events: Toss of a (fair) coin

Event: One coin-toss.

- Events: either H or T (equally likely)
- Event-space: 2 elements
- Then the probability of getting a head is 1 out of 2, or 50%.
- Similarly, the probability of a tail is 1 out of 2, or 50%.

Mutually exclusive events

Concept: – when the probability of two mutually exclusive events happening at the same time is zero.

For example, getting a head and a tail on a single toss of a coin has a probability of zero. (We say that the event space is a null set.)

Simple coin toss, complicated

Event: Two coin-tosses.

- Events: (H,H), (H,T), (T,T), (T,H)
- Event-space: four elements
- Probability of a head followed by a head: 1 out of 4 (25%)
- Probability of a head and a tail : 2 out of 4 (50%)
- Probability of a tail followed by a head: 1 out of 4 (25%)

Testing concepts: two coin tosses

I define events as follows:

E1 The event is head followed by a head

E2 The event is head followed by a tail

E3 The event is one head and one tail

Questions:

1. Is E1 and E2 mutually exclusive?
2. Is E1 and E3 mutually exclusive?
3. Is E2 and E3 mutually exclusive?

Testing concepts: two coin tosses

I define events as follows:

E1 The event is head followed by a head

E2 The event is head followed by a tail

E3 The event is one head and one tail

Questions:

1. Is E1 and E2 mutually exclusive? Yes
2. Is E1 and E3 mutually exclusive? Yes
3. Is E2 and E3 mutually exclusive? No

Basic Properties of probabilities

- Probability of all events taken together is 1.
  For example, the probability that the rat will either die or not is 1.
- The probability of a single event is greater than 0. A probability 0 event does not exist in the space.
  For example, the sun rising in the west.
- The highest probability of a single event is 1.
  For example, the sun rising in the east.

Manipulating probabilities

- The probability of two mutually exclusive events together is zero.
  For example, the rat cannot both live and die – ie, probability of 0.
- If two events are mutually exclusive, then the probability that either event will happen is the sum of their probabilities.
  For example, the probability that in two tosses of a coin, both will be heads or both will be tails is 50%.

Testing Concepts: calculating probabilities

A survey of a large number of economists produced the following probability estimates for four aspects of the nation's economy next year:

	Event	Probability (%)
A	GNP will increase by more than 10%	0.05
B	GNP will be greater than 8.5% but less than 10%	0.20
C	GNP will be greater than 7% but less than 8.5%	0.30
D	GNP will be less than 7%	0.40

What is the economist profession's probability estimate that the GNP will increase by more than 8.5% next year?

Testing Concepts: calculating probabilities

Two prospective managers are selected by chance from among four candidates. If two candidates are really better than the others, what is the probability that at least one of these two will be selected?

Testing Concepts: calculating probabilities

Two prospective managers are selected by chance from among four candidates. If two candidates are really better than the others, what is the probability that at least one of these two will be selected?

- Four candidates: M1, M2, M3, M4 (M1, M2 are good ones).
- Event space: (M1,M2), (M1,M3) (M1,M4), (M2,M3), (M2,M4), (M3,M4) = 6.
- At least one of M1, M2 is subset (M1,M2), (M1,M3), (M1,M4), (M2,M3), (M2,M4) = 5.
- Prob(at least one good candidate) = 5/6.
- The complement is both bad candidates = (M3, M4) = 1/6
- Then, prob(at least one good) = 1 - Prob(both bad) = 5/6

Conditional probability

Draws out of box with black, green and white balls

Event: Two draws out of the box – what is the probability of drawing one green and one black ball?

- Without replacing balls
  - Events: (B,W), (B,G), (W,B), (W,G), (G,W), (G,B)
  - Event space: 6
- With replacing balls
  - Events: (B,B), (B,W), (B,G), (W,B), (W,W), (W,G), (G,B), (G,W), (G,G)
  - Event space: 9

Prob(B,G) depends upon whether the first ball was replaced or not.

Defining conditional probability

The probability of an event A given that another event B has already taken place is called the conditional probability of A given B.

It is typically written as Prob(A|B).

(The probability of an event A is called the unconditional probability of A.)

Testing concepts: conditional probability

A box contains three fuses, one good (G) and two bad (B1, B2). Two fuses are drawn, one after the other.

1. What is the probability that the second fuse drawn is bad?
2. What is the probability that the second fuse drawn is bad, given that the first fuse was bad?

Testing concepts: conditional probability

A box contains three fuses, one good (G) and two bad (B1, B2). Two fuses are drawn, one after the other.

1. What is the probability that the second fuse drawn is bad?
2. What is the probability that the second fuse drawn is bad, given that the first fuse was bad?

- Events: (G,B1), (G,B2), (B1,G), (B1,B2), (B2,G), (B2,B1)
- Prob(second fuse is bad) = 4/6 = 2/3
- Prob(second is bad|first is bad) = 2/6 = 1/3

Testing concepts: conditional probability

A (fair) coin is tossed 10 times with a result of getting 10 tails. If the coin is tossed one more time, what is the probability of observing a head?

Testing concepts: conditional probability

A (fair) coin is tossed 10 times with a result of getting 10 tails. If the coin is tossed one more time, what is the probability of observing a head?

- Prob(head on the 11th draw|ten tails) = 1/2

Testing concepts: conditional probability

A (fair) coin is tossed 10 times with a result of getting 10 tails. If the coin is tossed one more time, what is the probability of observing a head?

- Prob(head on the 11th draw|ten tails) = 1/2

The conditional probability of the event is equal to the unconditional probability of the event.

Independent events

1. If the probability of one event does not depend upon whether another event has taken place, the events are independent.
2. Mathematically, two events are independent if

   $$\text{Prob}(A|B) = \text{Prob}(A) \text{ or if } \text{Prob}(B|A) = \text{Prob}(B)$$

3. If two events A, B are independent with probabilities Prob(A) and Prob(B), the probabililty of both events taking place is

   $$\text{Prob}(A \text{ and } B) = \text{Prob}(A)\,\text{Prob}(B)$$

Testing concepts: calculating probabilities

What is the probability that you will observe two heads in two tosses of a coin?

Testing concepts: calculating probabilities

What is the probability that you will observe two heads in two tosses of a coin?

- We know that observing heads on any one draw is independent of the other for a coin toss.
- Prob(H) = 1/2.
- Prob(H and H) = Prob(H) Prob(H) = $\frac{1}{2} * \frac{1}{2}$ = 1/4

Testing concepts: calculating probabilities

Prior experience at a workshop shows that the assembly line produces one in ten defective units, on average. Suppose that during a given period, there are five defective units that emerge in sequence.

1. If previously, defective units occur at random, what is the probability of obsreving a sequence of five consecutive defective units?
2. If this actually happened, what would you conclude?

Testing concepts: calculating probabilities

- The unconditional probability of observing defective units is 1/10.
- If they used to be produced at random, then the production of defective units is an independent process.
- Prob(five defective units in sequence) = $\frac{1}{10} * \frac{1}{10} * \frac{1}{10} * \frac{1}{10} * \frac{1}{10}$ = 1/10,000, a very low probability
- We'd conclude the production process has become unstable if there were five defective units in a row.

Non-independent events

- **Multiplicative law**
  If A and B are two events which are not independent of each other, the Prob(A and B) is calculated as

$$Prob(A)Prob(B|A)$$

- **Additive law**
  If A and B are independent, then Prob(either A or B or both) is calculated as:

$$Prob(A) + Prob(B) - Prob(both A and B)$$

Testing concepts: non-independent events

A box contains three fuses, one good (G) and two bad (B1, B2). Two fuses are drawn, one after the other.

1. What is the probability that the second fuse drawn is bad?
2. What is the probability that the second fuse drawn is bad, given that the first fuse was bad?

Testing concepts: non-independent events

A box contains three fuses, one good (G) and two bad (B1, B2). Two fuses are drawn, one after the other.

1. What is the probability that the second fuse drawn is bad?
2. What is the probability that the second fuse drawn is bad, given that the first fuse was bad?

- Event A: the second draw is a bad fuse.
- Event B: the first draw is a bad fuse. Prob(B) = 1/3
- Prob(A|B) = 1/2
- Prob(A and B) = Prob(B) Prob(A|B) = $\frac{1}{3} * \frac{1}{2}$ = 1/6.

Testing concepts: dependent events

A survey of 1000 small business ventures are categorised as to their profitability and their age (less than 2 years, between 2 and 5 years, greater than 5 years). The percentages of businesses in each category is given as:

	Time the business has been operational (Years)			Totals
	Less than 2	Between 2 and 5	Greater than 5	
Profitable	2	8	14	24
Unprofitable	16	35	25	76
Total	18	43	39	100

- Event A: The business is profitable.
- Event B: The business has been operational for more than 5

Testing concepts: Dependent events

1. Prob(A and B) = 14/100 = 14%
2. Prob(A|B) = 14/39 = 35.89%
3. Complement of Prob(A) = 76/100 = 76%
4. Prob(A or B or both) = Prob(A) + Prob(B) - P(A and B) = 0.24 + 0.39 - 0.14 = 0.49 = 49%

Problems to be solved

Problem 1

The event space is formed by three tosses of a coin.

- What is the event space?
- What is the probability that three tosses will give two heads and one tail?
- What is the probability of having two heads followed by a tail?

Problem 2

Consider: Two dice are tossed and the number on the upper faces are summed up.

1. What is the probability that the sum is equal to 2?
2. What is the probability that the sum is equal to 7?
3. What is the probability that the sum is equal to 14?

Problem 3

13 cards are dealt face up from a standard, well-mixed, 52-card deck and are observed. If an event A is defined as follows:
4 hearts, 4 spades, 3 clubs and 2 diamonds
it can be shown that the probability of A, ie, Prob(A) = 0.01796. What is the meaning of “Prob(A) = 0.01796”?

Problem 4

Assume that

- $\text{Prob}(A) = 0.6$
- $\text{Prob}(B) = 0.3$
- $\text{Prob}(C) = 0.5$
- $\text{Prob}(A|B) = 0.15$
- $\text{Prob}(A|C) = 0.5$
- $\text{Prob}(B|C) = 0.3$

1. Are A and B independent?
2. Are B and C independent?
3. Are A and C mutually exclusive?

Problem 5

Your broker has set up a deal that involves two separate investments in London gold and American silver. You are told that the gold investment has a 95% chance of being successful and silver has a 80% chance. However, the deal is worked so that if either the gold or silver investment is a success, your overall investment will also be a success.

If the success of the gold and the silver investments are *independent* of each other, what is the probability that your overall investment will be a success?

Problem 6

Astrologists claim that they can foretell the future based on knowledge of the zodiac because of the specific character traits that these signs assign. For example, persons born under the sign of Leo are said to possess outstanding leadership qualities. Suppose the probability of a randomly selected person being born a Leo is 0.085. The probability that a randomly selected person being a great leader given they were born a Leo is 0.004.

1. Find the prob(a randomly selected person is both an outstanding leader and a Leo).
2. The unconditional probability that a randomly selected person is an outstanding leader is 0.001. Find the probability that a randomly selected person is either an outstanding leader or a Leo or both.

References

- Chapter 1, Sheldon Ross. *Introduction to Probability Models*. Harcourt India Pvt. Ltd., 2001, 7th edition