BNAD 276
Lecture 3
Probability Theory I

Phuong Ho

May 17, 2017

Outline

Outline

Introduction

- From today, we will take off from data for a while.
- We will discuss basic probability theory.
- Why?
- Since we need to build a better understanding of theories to develop tools for data analysis later.

Outline

Probability

Probability

Probability is a numerical measure of the likelihood that an **EVENT** will occur.

- What is an event?
- To answer to this question, we need to be more formal.
- We begin with defining experiments to define events.

Experiments

Experiment

An experiment is a process that generates well-defined outcomes.

Experiments	**Outcomes**
Toss a coin	Head, Tail
Select a part for inspection	Defective, nondefective
Roll a die	1,2,3,4,5,6
Election	Win, lose, tie

- Once we have an experiment, we can define a SAMPLE SPACE.

Sample Space

The sample space for an experiment is the set of all outcomes of the experiment.

- An experimental outcome, or an element of the sample space is also called a **sample point.**

e.g. Rolling a die

- The sample space of this experiment is

$$S = \{1, 2, 3, 4, 5, 6\}$$

- Once we have a sample space, we can define EVENTS.

Event

An event is a collection of outcomes in a sample space.
(i.e. an event is a subset of a sample space)

Ex 1 Rolling a die

$$S = \{1, 2, 3, 4, 5, 6\}$$

- Any subset of S will be an event.
- $\{1\}$, $\{2\}$, ..., $\{6\}$ are events.
- $\{1, 2, 3\}$, $\{1, 3, 5\}$, and $\{2, 4, 6\}$ are also events.

Example 1 cont'd. Interpret an event.

- The event, $\{1\}$, is the event that I get 1 when I roll a die.
- The event, $\{1, 3, 5\}$, is the event that I get odd numbers.
- The event, $\{1, 2, 3\}$, is the event that I get numbers less than 4.

Example 2

Let's consider another experiment, Election.

$$S = \{Win, Lose, Tie\}$$

- Possible events are

$$\{Win\},\ \{Lose\},\ \{Tie\},\ \{Win, Lose\},\ \{Win, Tie\},$$

  and so on.

- Recall the definition of probability, *a numerical measure of the likelihood that an event will occur*.
- Now that we defined what events are, we can assign probabilities on events.

Assigning Probabilities

Probability of an Event

The probability of an event is equal to the sum of the probabilities of the sample points in the event.

Example 1 cont'd (Rolling a die)

- Suppose we want to assign the probability on the event, $E = \{1, 2, 3\}$.
- Following the above statement, $P(E) = P(1) + P(2) + P(3)$.
- Thus, we need to assign a probability to each sample point.
- Why can we assign probabilities on each sample point?

Assigning Probability

Requirements for assigning probabilities

1. Let e_i denote i^{th} sample point from S from an experiment and $P(e_i)$ denote its probability, then

$$0 \leq P(e_i) \leq 1 \text{ for all } i$$

2. Let the sample space $S = \{e_1, e_2, ..., e_n\}$. Then,

$$P(e_1) + P(e_2) + \cdots + P(e_n) = 1$$

That is, sum of the probabilities for all sample points in the sample space must equal to 1.

Popular assigning methods: classical, subjective, relative frequency, ...

Classical Method: Equally Likely. Example 1

Consider tossing a coin.

$$S = \{H, T\}$$

- In this case, if the coin is "fair", it is reasonable to think H and T are equally likely to happen.
- Thus, we assign the following probabilities.

$$P(H) = \frac{1}{2}, \quad P(T) = \frac{1}{2}$$

Q. Do these probabilities satisfy the requirements for assigning probabilities?

Classical Method: Equally Likely. Example 2

Consider rolling a die.

$$S = \{1, 2, 3, 4, 5, 6\}$$

- In this case, if the die is "fair", it is reasonable to think $1, \ldots, 6$ are equally likely to happen.
- Thus, we assign the following probabilities.

$$P(1) = \frac{1}{6}, \ P(2) = \frac{1}{6}, \ \cdots, \ P(6) = \frac{1}{6}$$

Q. do these probabilities satisfy the requirements for assigning probabilities?

Subjective Method: To reflect "subjective" belief

Consider rolling a die.

$$S = \{1, 2, 3, 4, 5, 6\}$$

- Suppose that I personally believe that 6 is more likely happen when I roll a die.
- We can reflect this "personal belief" when we assign probabilities on each outcome (sample point).
- Again, this is "sensible" probabilities just for me.

$$P(1) = P(2) = \cdots = P(5) = \frac{1}{10}, \; P(6) = \frac{1}{2}$$

Q. do these probabilities satisfy the requirements for assigning probabilities?

Relative Frequency Method: When data is available

Consider the following experiment. There is a hospital that has 3 waiting seats. We count the number of people who are on waiting seats.

$$S = \{0, 1, 2, 3\}$$

Suppose we could collect data for 20 days and construct the frequency distribution of data as follows.

Outcome	Frequency
# of people	# of days the outcome occured
0	7
1	6
2	4
3	3
Total	20

Relative Frequency Method. Cont'd

# of people	# of days the outcome occurred
0	7
1	6
2	4
3	3
Total	20

We may use the relative frequency of an outcome as the probability of that outcome.

$$P(0) = \frac{7}{20},\ P(1) = \frac{6}{20},\ P(2) = \frac{4}{20},\ P(3) = \frac{3}{20}$$

Outline

- As we saw, it is important to know how many outcomes in the sample space to assign probabilities.
- If we don't know how many outcomes we have in the sample space, we can't properly assign probabilities.
- Thus, it is important to know how to count all possible outcomes.
- Counting outcomes in the previous examples is straightforward.
- However, when we have a complicated experiment, it can be difficult.

Counting Rule for Multiple-step Experiments. Example 1

Consider the experiment of tossing a coin twice. Define the experimental outcomes by the pattern that heads and tails appear. For example, (H,T) is one outcome, (T,T) is another outcome. We want to know what is the sample space for this experiment so that we can assign probabilities. The sample space is as follows.

$$S = \{(H,H),(H,T),(T,H),(T,T)\}$$

Cont'd

$$S = \{(H,H),(H,T),(T,H),(T,T)\}$$

- If the coin is fair, we can assign the probability of $\frac{1}{4}$ to each outcome by using the classical method of assigning probabilities.
- (Again, notice that once we know the number of outcomes in a sample space it is easy to assign probabilities.)
- In this example, it is still straightforward to count the number of all possible outcomes.

Example 2

Suppose now, we toss a coin 5 times. An outcome is a pattern that heads and tails appear.

- What is an example of an outcome in this experiment?
- How many outcomes do we have?
- Following the classical method, what is the probability that we get for each outcome in this experiment?

Counting Rule for Multiple-step Experiments

Counting Rule for Multiple-step Experiments

Suppose an experiment has k steps. The number of all possible outcomes in the i^{th} step is denoted by n_i.
The total number of experimental outcomes is given by:

The number of outcomes of a k-step experiment $= n_1 \times n_2 \times \ldots \times n_k$

Combinations (not ordered, without replacement)

- Combination counting rule allows us to count the number of outcomes in an experiment that involves selecting k objects out of N objects at a time (i.e. choose k objects out of N objects without replacement and the order of chosen objects does not matter).

e.g. Examples of such an experiment are as follows

- Take 2 pens out of a box containing 5 pens at a time. (Select 2 out of 5 objects)
- Take out 4 cards from a deck of 52 cards at a time.

Examples cont'd

!NOTICE!

- Combination counting rule is valid only if the order does not matter and replacement is not allowed.

Again, draw 2 pens out of a box containing 5 pens without replacement. (Select 2 out of 5 objects.) Imagine the box contains 5 pens that has the labels A, B, C, D, E.

- An example of an outcome is AB.
- In this experiment, we assume that the order does not matter. That is, BA is considered as the same outcome as AB.
- Without replacement (taking out/drawing at a time) means that there is no outcome like AA.

Combinations

Counting rule for combinations

Suppose we have N objects and select k objects out of N at a time. The number of outcomes in this experiment is

$$C_k^N = \left(\begin{array}{c} N \\ k \end{array} \right) = \frac{N!}{k!(N-k)!},$$

where

$$\begin{aligned} N! &= N(N-1)(N-2)\cdots(2)(1), \\ k! &= k(k-1)(k-2)\cdots(2)(1), \\ 0! &= 1. \end{aligned}$$

Example cont'd

Draw 2 pens out of a box containing 5 pens at a time.

- We choose 2 out of 5 objects. Thus, $N = 5$ and $k = 2$.
- By using the formula of counting rule for combinations,

$$\begin{aligned} C_2^5 &= \frac{5!}{2!(5-2)!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)(3)(2)(1)} \\ &= \frac{120}{10} = 10. \end{aligned}$$

- Thus, we have 10 possible outcomes in this experiment.

Permutations (ordered, without replacement)

- Recall that the combination counting rule is used when the order does not matter.
- However, sometimes, the order DOES mater.
- Permutations allow us to count the number of outcomes when we choose k out of N objects AND the order matters.

Note We still assume "without replacement" in an experiment.

Example: Permutation

Consider the following experiment. We will draw 2 pens from a box containing 3 pens that have different colors, Red, Blue, and Green. We may have two different experiments depending on what we are interested in.

1. We only care about the combination of 2 pens' colors.
   - i.e. (Red, Blue) is the same outcome as (Blue, Red). As far as we have Blue and Red combination, the order does not matter.
2. We care about both the combination of colors of 2 pens and their order.
   - i.e. Now, (Red, Blue) is a different outcome than (Blue, Red) since now we do care about the order.

Example Cont'd

1. In the first experiment, the order does not matter. Thus, sample space will be

$$S = \{(R, B), (R, G), (B, G)\}$$

2. In the second experiment, the order matters. Thus, sample space will be

$$S = \{(R, B), (B, R), (R, G), (G, R), (B, G), (G, B)\}$$

Counting rule for Permutations

Permutations

The number of k-permutations of N objects (take k objects sequentially out of N objects) is given by

$$P_k^N = k!\left(\begin{array}{c} N \\ k \end{array}\right) = \frac{N!}{(N-k)!}$$

Example

Consider the previous example again. We draw 2 pens from the box containing 3 pens, R,B, and G. In the second experiment, when we care about the order of colors, the number of possible outcomes will be

$$P_2^3 = 2! \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 2! \times C_2^3 = \frac{3!}{1!} = \frac{(3)(2)(1)}{1} = 6$$

Outline

Addition Law. Introduction

- Suppose we have two events, A and B.
- We are interested in the probability that either A or B occurs.
- The Addition Law helps us to calculate this probability.
- However, before we start, things will be easy if we define a new event: The event that either A or B occurs, i.e. the event that at least one of the two events occurs.

[Addition Law] Union vs. Intersection of two events

Union of two events

The **union** of A and B is the event containing all sample points belonging to A ***or*** B ***or*** both.
The union is denoted by $A \cup B$.

Intersection of two events

The **intersection** of A and B is the event containing the sample points belonging to both A ***and*** B.
The intersection is denoted by $A \cap B$.

Example: Roll a die

Let $A = \{1, 2, 3\}$ and $B = \{2, 3, 5\}$

- What is $A \cup B$?
- What is $A \cap B$?
- Let's define new event C as $A \cup B$. i.e. $C = A \cup B$.

Addition Law

Addition Law

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

- Why do we need to subtract $P(A \cap B)$ instead of just summing $P(A)$ and $P(B)$?
- Venn diagram

Example: Roll a die

Let $A = \{1, 2, 3\}$ and $B = \{2, 3, 5\}$

- What is $A \cup B$? What is the $P(A \cup B)$?
- What is $A \cap B$? What is the $P(A \cap B)$?
- What is the $P(A) + P(B)$?

Mutually Exclusive Events

Consider rolling a die again. We are interested in the following two events.

$$A = \{1, 3\},\ B = \{2, 4\}$$

- What is $A \cup B$? What is the $P(A \cup B)$?
- What is $A \cap B$? What is the $P(A \cap B)$?
- What is the $P(A) + P(B)$?

Note that $A \cap B$ does not contain any thing. Thus, $P(A \cap B) = 0$. Finally, we have

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

Mutually Exclusive Events

Mutually Exclusive Events

Two events are said to be mutually exclusive if the intersection of two events is empty. In other words, if the two events have no sample points in common.

Addition Law for Mutually Exclusive Events

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

Complement of an Event

Complement of an Event

Given an event A, the **Complement of** A is defined to be the event consisting of all sample point that are NOT in A. And the complement of A is denoted by A^c.

e.g. Suppose we have $S = \{1, 2, 3, 4\}$.

- Let's say we have an event $A = \{1, 2\}$.
- What is A^c?
- $A^c = \{3, 4\}$

What's special about A and A^c

- Note that these two events A and A^c are mutually exclusive.

e.g. Suppose we have $S = \{1, 2, 3, 4\}$.

- Let's say we have an event $A = \{1, 2\}$.
- $A^c = \{3, 4\}$.
- Then, what is $A^c \cap A$?
- What is $A^c \cup A$?

What's special about A and A^c

- According to the definition, clearly, either event A or event A^c MUST occur.
- The above statement directly suggests the following equation.

$$P(A \cup A^c) = P(S) = 1$$

- Since A and A^c are mutually exclusive, addition law also works:

$$P(A \cup A^c) = P(A) + P(A^c) = 1$$

e.g Roll a die. $A = \{1, 3, 5\}$, $A^c = \{2, 4, 6\}$.

Computing Probability using the complement

We know the following is always true.

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

Then, the following result is immediate.

Computing Probability using the Complement

$$P(A) = 1 - P(A^c)$$

Example

Consider the example of rolling a die again. What is the probability that we get the number greater than 1?

- It will be $P(2) + P(3) + P(4) + P(5) + P(6) = \frac{5}{6}$.
- Alternatively, we can calculate by applying the concept of complement?

Outline

1. Introduction
2. Experiments, Sample Space, and Probability
3. Counting Rules
4. Basic Probability Laws
5. Exercises

Exercise 1

A decision maker subjectively assigned the following probabilities to the four outcomes of an experiment: $P(E_1) = .10$, $P(E_2) = .15$, $P(E_3) = .40$, and $P(E_4) = .20$. Are these probabilities valid? Explain.

Exercise 2

The data for US adult age is as follows (*The World Almanac*, 2009). The data are in millions of people.

Age	**Number**
18 to 24	29
25 to 34	40
35 to 44	43
45 to 54	43
55 to 64	32
65 and over	33

Assume that a person will be randomly chosen from this population.

a) Which method is appropriate to assign probability to each class?

b) What is the probability that the person is 18 to 24 years old?

c) What is the probability that the person is 18 to 34 years old?

d) What is the probability that the person is 45 or older?

Exercise 3

We have a fair coin, a fair die, and a box containing 5 balls with different letters A, B, C, D, and E. Consider the following multiple step experiment. We toss the coin and roll the die and finally draw 3 balls out of the box. An example of outcome is $(H, 2, (B, D, E))$.

a) We don't care about the order of drawn balls but only care about combination of balls.

   a-1) What is the number of all possible outcomes in this experiment?
   a-2) Following the classical method of assigning probability, what is the probability that each outcome happens?

b) Now, we care about the order of drawn balls.

   b-1) What is the number of all possible outcomes in this experiment?
   b-2) Following the classical method of assigning probability, what is the probability that each outcome happens?

Exercise 4

A survey of magazine subscribers showed that 45.8% rented a car during the past 12 months for business reasons, 54% rented a car during the past 12 months for personal reasons, and 30% rented a car during the past 12 months for both business and personal reasons.

a) What is the probability that a subscriber rented a car during the past 12 months for both business and personal reasons?

b) What is the probability that a subscriber rented a car during the past 12 months for either business or personal reasons?