Introduction to Statistical Modelling

STAT2507A

LECTURE 4 -1

Probability and Probability Distributions

REVIEW

- In Chapters 2 and 3, we used graphs and numerical measures to describe data sets which were usually **samples**

PREVIEW

- In Chapters 4-7, we will learn many different ways to calculate probabilities. We will assume that the population is known and calculate the probability of observing various sample outcomes

What is Probability?

- We measured "how often" using:
  - Relative frequency = frequency/sample size = f/n
- As n gets larger
  - Sample → Population
  - Relative frequency → probability

EVENTS AND THE SAMPLE SPACE

- Data obtained by observing uncontrolled events in nature or controlled situations in a laboratory.
- Experiment: the process by which an observation (or measurement) is obtained:
  - Experiment: record an age
  - Experiment: toss a die
  - Experiment record an opinion (yes, no)
  - Experiment: toss two coins

EVENTS AND THE SAMPLE SPACE

- Simple Event: It is the outcome that is observed on a single repetition of the experiment
  - The basic element to which probability is applied
  - One an only one simple event can occur when the experiment is performed
  - A simple event is denoted by E with a subscript

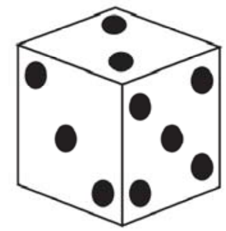
EVENTS AND THE SAMPLE SPACE

- If, when one event occurs, the other cannot, and vice versa, then these two events are **mutually exclusive.**
- E.g. If we obtain “head” in one toss, then we can not obtain “tail” in the same toss. So, these two simple events are mutually exculsive.

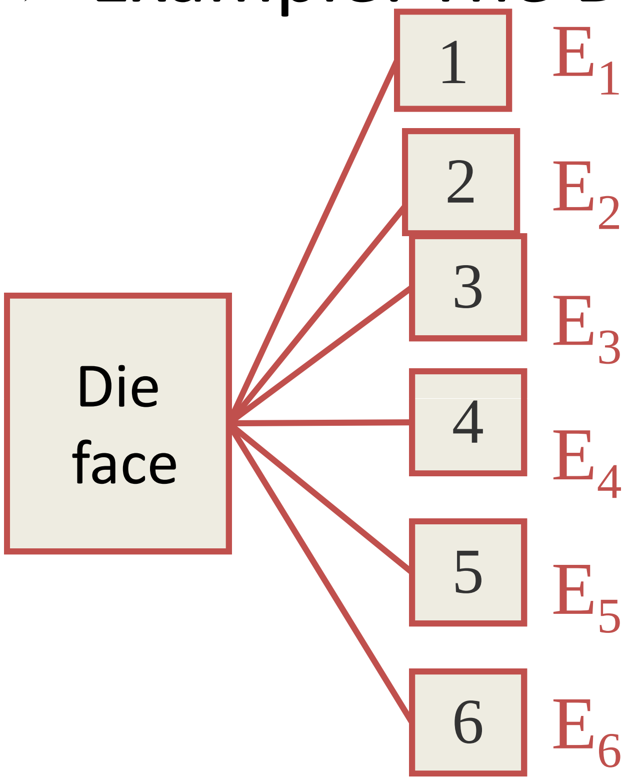
EXAMPLE

- Each simple event will be assigned a probability, measuring "how often" it occurs
- The set of all simple events of an experiment is called sample space, S.
- When a die is tossed, there are six possible outcomes. Simple events are:
  - Event E_1 – observe a 1 Event E_2 – observe a 2
  - Event E_3 – observe a 3 Event E_4 – observe 4
  - Event E_5 –observe a 5 Event E_5 – observe 5

EVENTS AND SAMPLE SPACE



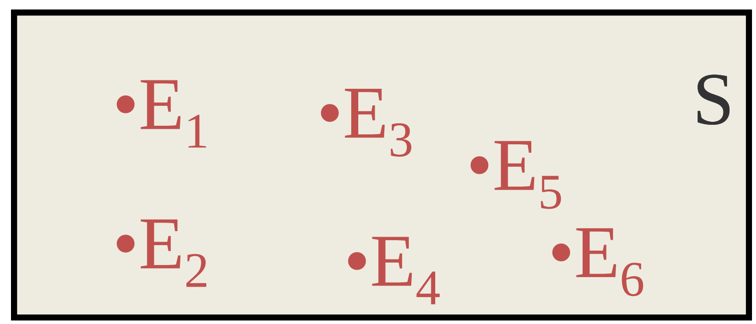
- Example: The Die Toss



Simple events

Sample space

$$S = \{E_1, E_2, E_3, E_4, E_5, E_6\}$$



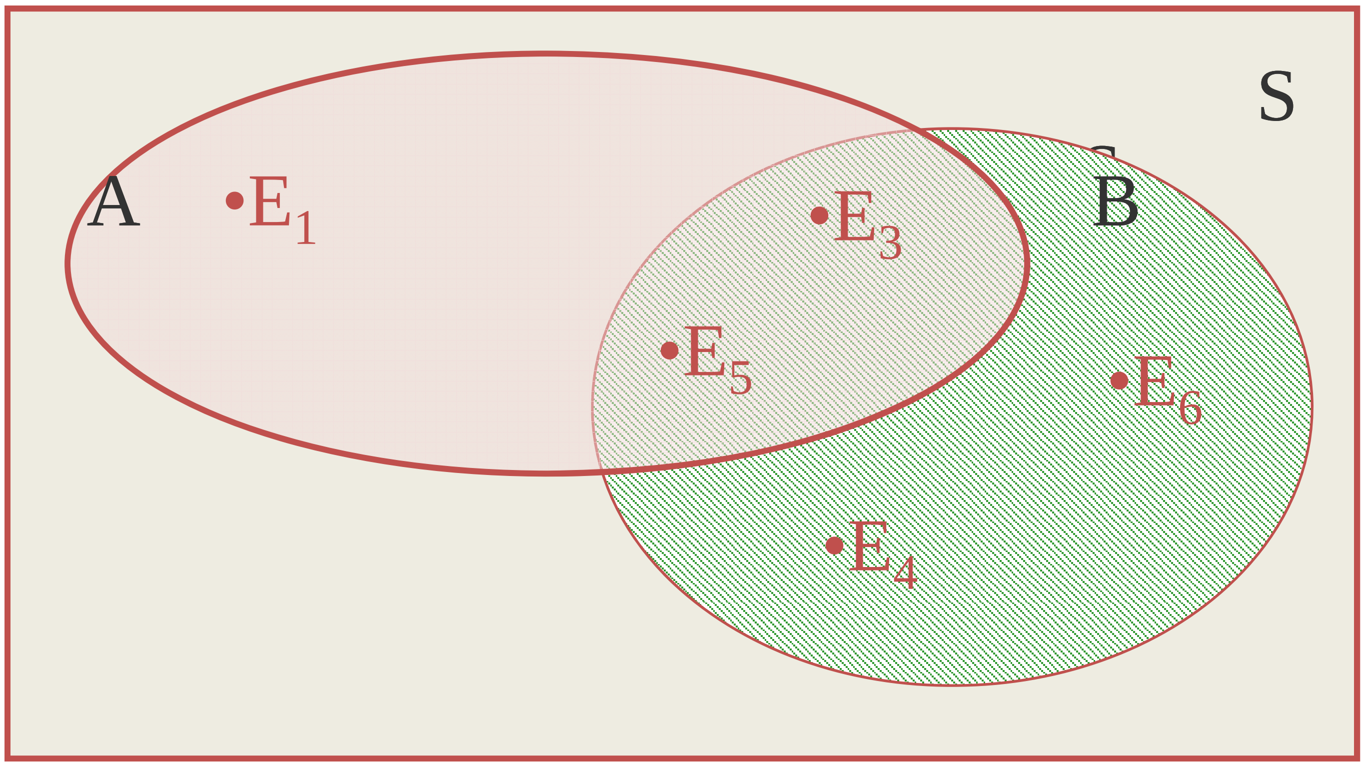
EVENTS AND THE SAMPLE SPACE

- Simple events are mutually exclusive: When the experiment is performed once, one and only one of these simple events can occur
- **Event**: It is the collection of one more simple events

EVENTS AND SAMPLE SPACE

- Experiment: Die toss
  - Event A: observe an odd number
  - Event B: observe a number greater than 2
- $A = \{E_1, E_3, E_5\}$
- $B = \{E_3, E_4, E_{5,} E_6\}$

EVENTS AND THE SAMPLE SPACE



EVENTS AND SAMPLE SPACE

4.1)Experiment: Toss a die

A: Observe an odd number

B: Observe a number greater than 3

C: Observe a 6

D: Observe a 3

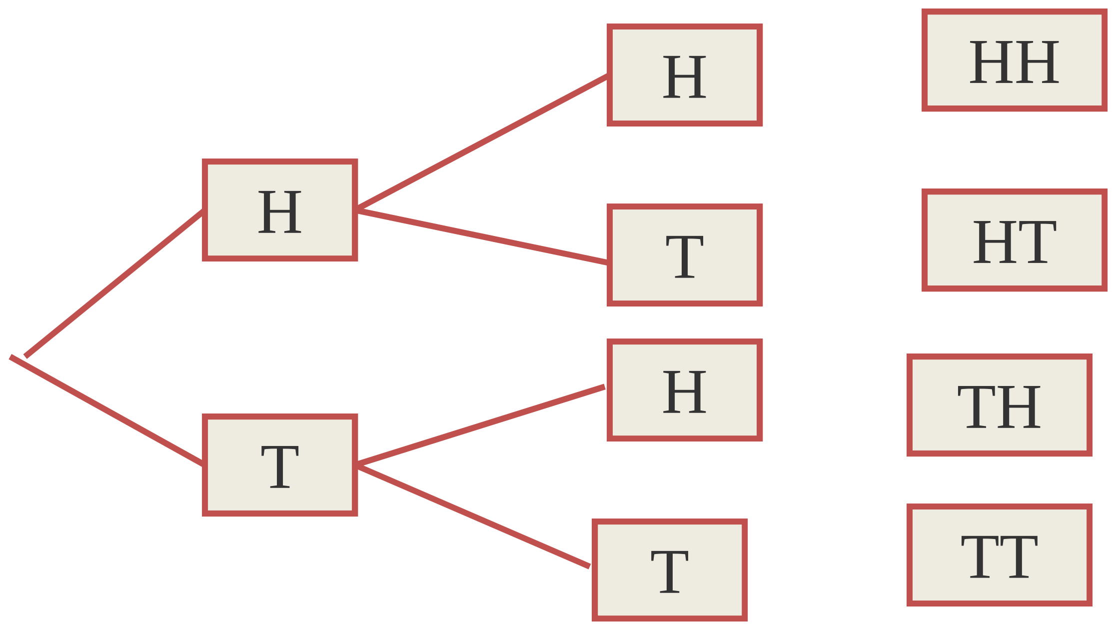
Are Events A and B, events C and D, events B and C, events B and D mutually exclusive?

EVENTS AND THE SAMPLE SPACE

- Sometimes, experiments can be generated in stages, the sample space can be displayed in a tree diagram. Each successive level of branching on the tree corresponds to a step required to generate the final outcome.

EXAMPLE

- Two stage experiment: Toss a coin twice

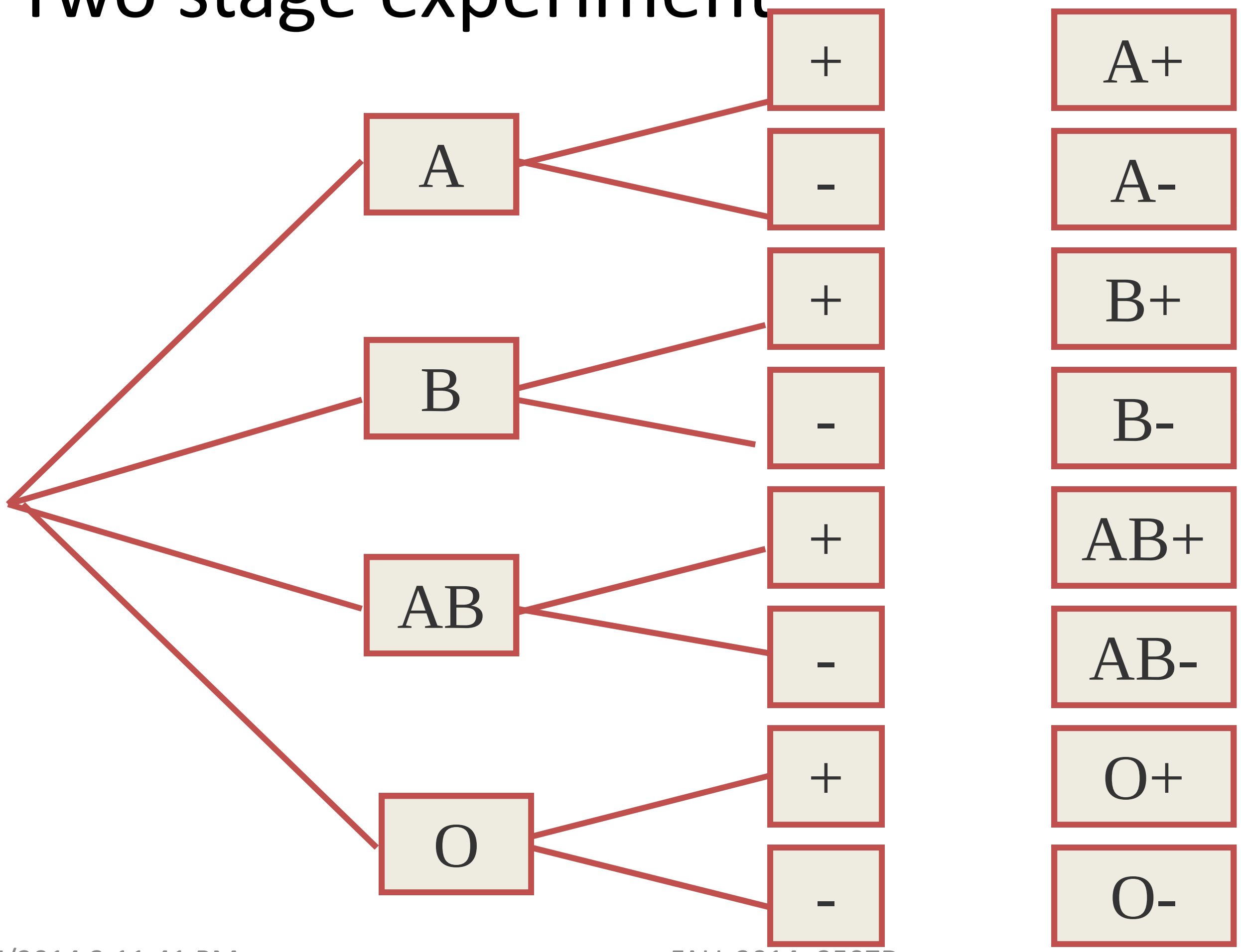


EXAMPLE

- A medical technician records a person's blood type and Rh factor. List the simple events in the experiment
  - For each person, a two-stage procedure is needed to record the two variable of interest.

TREE DIAGRAM

- Two stage experiment



CALCULATING PROBABILITIES USING SIMPLE EVENTS

- The probability of an event A measures "how often" we think A will occur; we write **P(A)**
- Suppose that an experiment is performed n times. The relative frequency for an event A is,

$$\frac{\text{Number of times A occurs}}{\text{n}} = \frac{f}{n}$$

- If n is infinitely large

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

CALCULATING PROBABILITIES USING SIMPLE EVENTS

- Requirements for simple event probabilities
  - If event A can never occur, $P(A) = 0$
  - If event A always occurs when the experiment is performed, $P(A) = 1$
- The sum of the probabilities for all simple events in S equals 1
- The **probability of an event A** is found by adding the probabilities of all the simple events contained in A

CALCULATING PROBABILITIES USING SIMPLE EVENTS

- Example for sum of the probabilities for all simple events in S equals 1
- Experiment: Coin toss
  - Simple events: S = {H, T}
  - P(H) + P(T) =1
- Experiment: Die toss
  - Simple events: S={1, 2, 3, 4, 5, 6}
  - P(1)+P(2)+P(3)+P(4)+P(5)+P(6)=1

FINDING PROBABILITIES

- Probabilities can be found using
  - Estimates from empirical studies: 77% of Canadian are identified as Christians (C). Select a person at random => P(C) = 0.77
  - Common sense estimates based on equally likely events: Tossing a coin has two simple events. If these events are equally likely (unbiased coin), then P(Head)=1/2 as these two events sum to 1

EXERCISE

4.2)Toss two fair coins. What is the probability of observing at least one head?

EXERCISE

4.3) A bowl contains three M&Ms: one red, one blue, and one green. A child selects two M&Ms at random. What is the probability that at least one is red?

Steps to calculate the probability of an Event

1. List all the simple events in the sample space
2. Assign appropriate probability to each simple event
3. Determine which simple events result in the event of interest
4. Sum the probabilities of the simple events that result in the event of interest

EXERCISE

4.4) An experiment involves tossing a single die. These are some events:

A: Observe a 2

B: Observe an even number

C: Observe a number greater then 2

D: Observe both A and B

E: Observe A or B or both

F: Observe both A and C

EXERCISE

a) List the simple events in the sample space
b) List the simple events in each of the events A through F
c) What probabilities should you assign to the simple events?
d) Calculate the probabilities of the six events A through F by adding the appropriate simple-event probabilities.

EXERCISE

4.5)A sample space S consists of five simple events with these probabilities:

$P(E_1)=P(E_2)=0.15$, $P(E_3)=0.4$, $P(E_4)=2P(E_5)$

a) Find the probabilities for simple events E4 and E5
b) Find the probabilities for these two events: A={E1,E2,E3}, B={E2,E3}
c) List the simple events that are either in event A or event B or both
d) List the simple events that are in both event A and event B

EXERCISE

4.6)Three people are randomly selected from voter registration and driving records to report for jury duty. The gender of each person is noted by the country clerk

a) Define the experiment and list all simple events
b) If each person is just as likely to be a man as a woman, what probability do you assign to each simple event?

EXERCISE

d) What is the probability that only one of the three is a man?

e) What is the probability that all three are woman?

Steps to calculate the probability of an Event

- It is important to find all simple events in the sample space
- One way to determine the required number of simple events is to use the counting rules.
  - If the simple events in an experiment are **equally likely**, you can calculate

$$P(A) = \frac{n_A}{n} = \frac{\text{Number of simple events in event A}}{\text{Total number of simple events}}$$

  - You can use **counting rules** to find n_A and N

USEFUL COUNTING RULES

- If an experiment is performed in two stages, with ***m*** ways to accomplish the first stage and ***n*** ways to accomplish the second stage, then there are ***mn*** ways to accomplish the experiment
  - This rule is easily extended to ***k*** stages, with the number of ways equal to $n_1 n_2 n_3 \ldots n_k$

USEFUL COUNTING RULES

(4.7)Find the total number of simple events for experiments in (a)-(d)

a) Toss two coins
b) Toss three coins
c) Toss two dice
d) Two M&Ms® are drawn from a dish containing two red and two blue candies

EXERCISE

4.8) There are three books that needs to be put on the shelf. But, there are space for only two books on the shelf. How many ways can you select and arrange the two books from three books?

FACTORIAL

- $5! = (5)(4)(3)(2)\ (1) = 120$
- $10! = (10)(9)\ \ldots(2)(1) = 3628800$
- $N! = (N)(N-1)\ldots(3)(2)(1)$
- $0! = 1$

Permutations

- The number of ways you can arrange *n* distinct objects, taking them *r* at a time is

$$P_r^n = \frac{n!}{(n-r)!}$$

Where n! = n(n-1)(n-2).......(3)(2)(1) and 0! = 1

PERMUTATIONS

- Since r objects are chosen, this is an r-stage experiment.
  - The 1st object can be chosen in n ways
  - 2nd in (n-1) ways
  - 3rd in (n-2) ways
  - rth in (n-r+1) ways.
  - total number of ways is n(n-1)(n-2)……(n-r+1).

$$n(n-1)(n-2)\cdots(n-r+1)=\frac{n(n-1)\cdots(n-r+1)(n-r)\cdots(2)(1)}{(n-r)\cdots(2)(1)}=\frac{n!}{(n-r)!}\,P_r^n$$

EXERCISE

4.9) How many 3-digit lock combinations can we make from the numbers 1, 2, 3, and 4?

EXERCISE

4.10) A piece of equipment is composed of five parts that can be assembled in any order. A test is to be conducted to determine the time necessary for each order of assembly. If each order is to be tested once, how many tests must be conducted?

COMBINATIONS

- When ordering or arrangement of the objects is not important, but the only the objects that are chosen , then the counting rule for combinations can be used.
- The number of distinct combinations of *n* distinct objects that can be formed, taking them *r* at a time is

$$C_r^n = \frac{n!}{r!(n-r)!}$$

COMBINATIONS

- The number of combinations and the number of permutations are related

$$C_r^n = \frac{P_r^n}{r!}$$

- Where r! is the number of ways of rearranging each distinct group of r objects chosen from the total n

EXERCISE

4.11) A printed circuit board may be purchased from five suppliers. In how many ways can three suppliers be chosen from the five?

4.12) A box contains six M&Ms®, four red and two green. A child selects two M&Ms at random. What is the probability that exactly one is red?