Lecture 2: Probability

MSU-STT-351-Sum-19A

In this lecture, we discuss

1. Random Experiments
2. Sample Space
3. Events
4. Tree Diagram
5. Venn Diagram

Our conclusions based on random samples can be extended to the population. Otherwise, they are summaries for that particular data set only.

Random Experiment

Probability methods help us to evaluate the reliability and confidence of the sample statistics.

Random (Chance) Experiment: An experiment (process, situation) whose outcomes are known, but cannot be predicted in advance. Random experiments arise out of natural phenomena or we perform them for inferential purposes.

Probability is a way modeling of outcomes of a random experiment. First we define the following concepts:

1. Sample Space
2. Events
3. Probability

Random Experiment

Sample space (S): The set of **all** possible outcomes (simple events) of a random experiment.

A Simple Event (e): An event consisting of only one possible outcome.

Compound Event: An event consisting of more than one outcome.

An event: Either a simple event or a compound event. That is, an event is any subset of S.

Note: We will always use the terminology "event" and will not differentiate between a simple and a compound event.

Example 1.

Experiment: A single die is rolled.
Sample space: $S = \{1, 2, 3, 4, 5, 6\}$.
Simple events: $\{1\}, \{2\}, \ldots, \{6\}$.
An event: $B = \{1, 3, 5\}$, an event of odd integers.

Note an event is a **set** of simple events or a **subset** of S.

Example 2

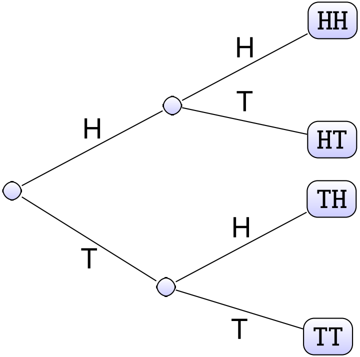
Experiment: Two dice are rolled.
Sample space: $S = \{(1, 1), (1, 2), \ldots, (6, 6)\}$.
Some simple events: $(3, 4)$ or $(5, 6)$ or $(6, 6)$.
An event: $C = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$

Note C denotes the event of getting the same outcome on both the dice.

Tree Diagram

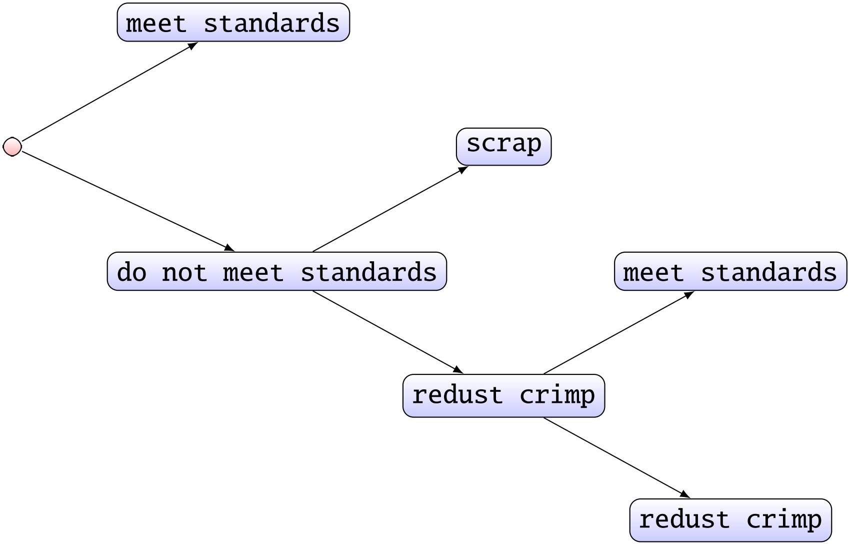
Tree Diagram: A sketch of steps performed/involved in a random experiment.

Example 3. The tree diagram for the outcomes of the experiment of flipping a coin **two** times:



Example 4. The following figure represents the tree diagram for the selection of fasteners (nuts and bolts) used in aircraft manufacturing:

All the branches need not be of the same length. That is, some branches may stop early, while others may extend through some additional branching points.


meet standards
scrap
do not meet standards
meet standards
redust crimp
redust crimp

Set Theoretic Relations

Definitions: Let A and B be two events. Then

(i) $A^c =$ Complement of the event A=the set of all simple points that are not in A. If A^c occurs, the event A does not occur.

(ii) Two events A and B are disjoint (mutually exclusive) if there is no common simple event. That is, $A \cap B = \phi$.

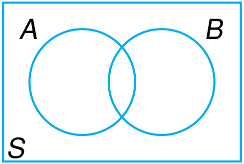
(iii) A and $B = A \cap B$= consists of all simple events common to both A and B. In this case, both A and B occur.

(iv) A or $B = A \bigcup B$= consists of all simple events that are either in A or in B or in both. (At least one of the events A or B occurs, that is, A occurs, or B occurs, or both occur).

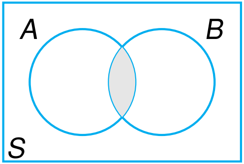
(v) Two events A and B are **independent** if the occurrence of one does not affect the occurrence of the other.

Venn diagram: Venn diagrams are used to show the relationships between events. They are a two dimensional figures (circles or rectangles) whose enclosed regions represent an event.

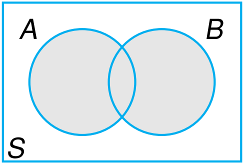
A rectangle denotes the sample space, and the circles inside denote the events. Common areas also represent common parts of the events. Some Venn diagrams for two events A and B are given below:



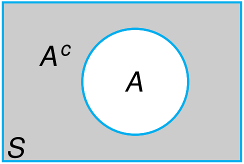
(a) Venn diagram of events A and B



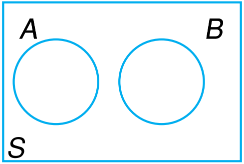
(b) Shaded region is $A \cap B$



(c) Shaded region is $A \cup B$



(d) Shaded region is A^c



(e) Mututally exclusive events

Set Theoretic Relations

Some simple excercises:

Exercise 2. Sketch a Venn diagram for the following two events in a single die experiment: $A = \{$a number less than 5$\}$; $B = \{$an even number$\}$.

Exercise 3. Observe the weather on two consecutive days. Sketch a Venn diagram for two events:

$A = \{$bad weather on 1st day$\}$; $B = \{$bad weather on 2nd day$\}$.

Then find on the Venn diagram the event: nice weather on 1st day or nice weather on 2nd day.

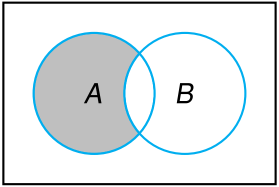
De Morgan Laws: Let A and B be two events. Then

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

Example 5. Draw the Venn Diagram for showing two events A and B which are not disjoint. Also, depict the event which corresponds to A but not B.

Solution: Any two events A and B for which (i) A and B are disjoint and (ii) the event A or B does not coincide with the entire sample space S and satisfy $P(A) + P(B) \neq 1$.

Two non-overlapping circles representing A and B, whose combined area is less than 1 will serve the purpose. The diagram is given below:

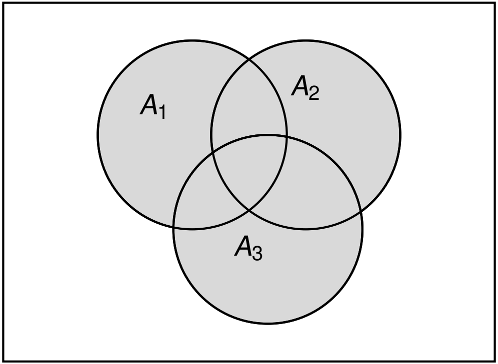


Example 6. An engineering firm is constructing power plant at three different sites. Define $A_i =$ the event that the plant at site i is completed by contract date, $i = 1, 2, 3$.

Draw the Venn diagram for the event and shade the region corresponding to the event that

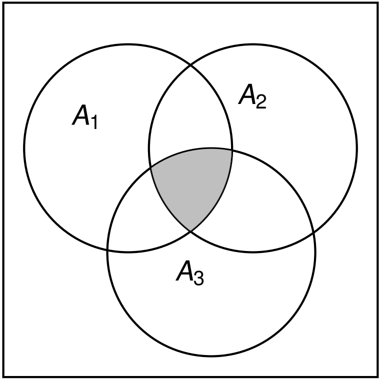
(a) At least one plant is completed by contract date.
(b) All plants are completed by the contract date.
(c) Only plant at site 1 is completed by the contract date.
(d) Exactly one plant is completed by the contract date.
(e) Either the plant at site 1 or both of the other two plants are completed by the contract date.

Solution: (a) $A_1 \bigcup A_2 \bigcup A_3$



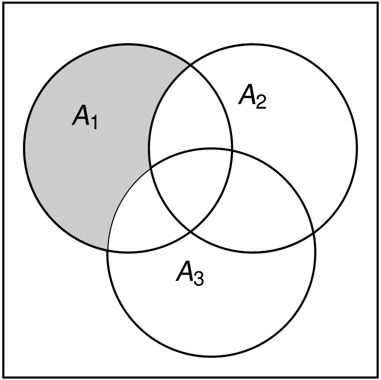
Set Theoretic Relations

(b) $A_1 \bigcap A_2 \bigcap A_3$



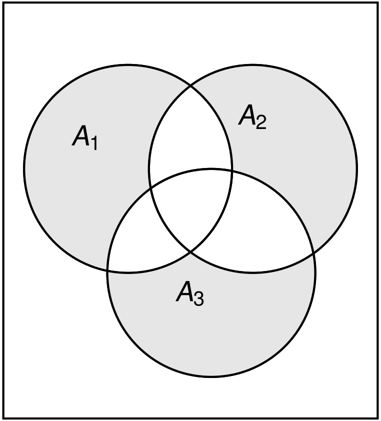
Set Theoretic Relations

(c) $A_1 \bigcap A_2^c \bigcap A_3^c$



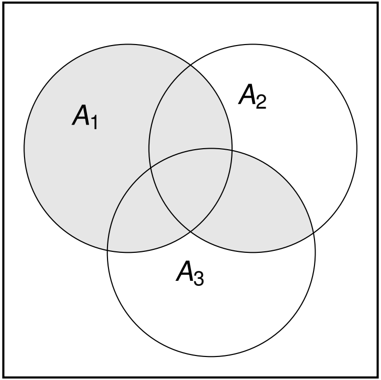
Set Theoretic Relations

(d) $(A_1 \bigcap A_2^c \bigcap A_3^c) \bigcup (A_1^c \bigcap A_2 \bigcap A_3^c) \bigcup (A_1^c \bigcap A_2^c \bigcap A_3)$



Set Theoretic Relations

(e) $A_1 \bigcup (A_2 \bigcap A_3)$



Probability: Definition

Probability of an event A is the quantification of the likelihood of A in a random experiment. A formal definition follows.

Definition 1

Probability is a function which assigns to each event a number P(A) with the properties:

Axiom 1: For any event A, $0 \leq P(A) \leq 1$.

Axiom 2: For the sample space S, $P(S) = 1$.

Axiom 3: If $A_1, A_2, \ldots$ is a sequence of **disjoint** events, then

$$P(A_1 \cup A_2 \cup \ldots) = \sum_{i=1}^{\infty} P(A_i).$$

Note: For any two **disjoint** events A and B,

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B).$$

Intuitively, P(A) represents the proportion of times an event A occurs in the long run.

Properties:

Proposition 1.
For any event A, $P(A) + P(A^c) = 1$ which implies $P(A^c) = 1 - P(A)$.
This implies $P(\phi) = 0$.

Proposition 2.
For any two events A and B, not necessarily disjoint ones,

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

Proposition 3:
For any three events A, B, and C,
$P(A \cup B \cup C) =$
$P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Excercise 4: Prove the above result using the result in Proposition 2.

Some Examples

Example 1. Suppose that the probability that it rains on Friday is 0.4, and that it rains on Saturday is 0.8, and that it rains on both days is 0.3. Then the probability of rain on Friday or Saturday is

$$P(A \cup B) = P(A) + P(B) - P(AB) = 0.4 + 0.8 - 0.3 = 0.9.$$

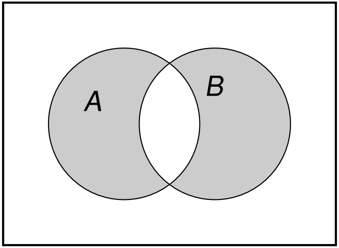
Example 2. (Ex 14) A utility company offers a lifeline rate to any household whose electricity usage falls below 240 kWh during a particular month. Let A and B respectively denote the event that a randomly selected household in a certain community does not exceed the lifeline usage during January, and during July. Suppose $P(A) = 0.8,\ P(B) = 0.7$, and $P(A \cup B) = 0.9$.
Compute the following:
(a) $P(A \cap B)$.
(b) The probability that the lifeline usage amount is exceeded in exactly one of the two months. Describe this event in terms of A and B.

Probability: Properties

Solution: (a) Note $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Hence,

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.8 + 0.7 - 0.9 = 0.6$$

(b) $P(\text{shaded region}) = P(A \cup B) - P(A \cap B) = 0.9 - 0.6 = 0.3.$



Note the shaded region = event of interest = $(A \cap B^c) \cup (A^c \cap B)$

Example 3. The three major options on a certain type of new car are an automatic transmission (A), a sunroof (B), and a stereo with compact disc player (C). Suppose 70% of all purchases request A, 80% request B, 75% request C, 85% request A or B, 90% request A or C, 95% request B or C, and 98% request A or B or C.

Compute the probabilities of the following events.
(a) The next purchaser will request at least one of the three options.
(b) The next purchaser will select none of the three options.
(c) The next purchaser will request only an automatic transmission and not either of the other two options.
(d) The next purchaser will select exactly one of these three options.

Probability: Examples

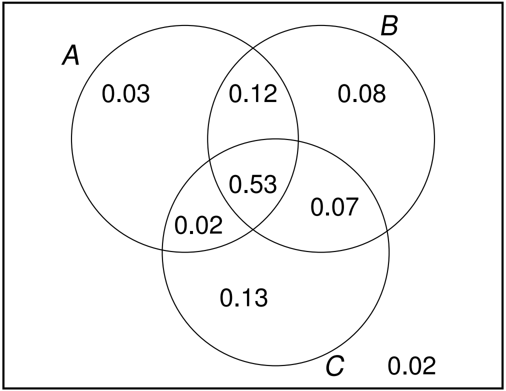
Solution: It is given $P(A) = 0.70,\ P(B) = 0.80,\ P(C) = 0.75$, Also, $P(A \cup B) = 0.85,\ P(A \cup C) = 0.90$, $P(B \cup C) = 0.95$ and $P(A \cup B \cup C) = 0.98$.

Then we compute

$$
\begin{aligned}
P(A \cap B) &= P(A) + P(B) - P(A \cup B) = 0.65 \\
P(A \cap C) &= 0.55, \quad P(B \cap C) = 0.60 \\
P(A \cap B \cap C) &= P(A \cup B \cup C) - P(A) - P(B) - P(C) \\
&\quad + P(A \cap B) + P(A \cap C) + P(B \cap C) \\
&= 0.98 - 0.7 - 0.8 - 0.75 + 0.65 + 0.55 + 0.60 \\
&= 0.53
\end{aligned}
$$

All the above information is given in the following Venn diagram.

Probability: Examples



Using the above quantities, we obtain

(a) $P(A \cup B \cup C) = 0.98$, as given

(b) $P(\text{none selected}) = 1 - P(A \cup B \cup C) = 1 - 0.98 = 0.02$

(c) P(only automatic transmission selected)=

$$
\begin{aligned}
P(A \cap B^c \cap C^c) &= P(A) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C) \\
&= .70 - .65 - .55 + .53 \\
&= 0.03 \quad \text{(from the Venn diagram).}
\end{aligned}
$$

(d) P(exactly one of the three)=0.03+0.08+0.13=0.24.

Alternatively,

$$
\begin{aligned}
(d) &= P(A \cup B \cup C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + 2P(A \cap B \cap C) \\
&= 0.24.
\end{aligned}
$$

Equally Likely Outcomes: In many random experiments of N (simple) outcomes, all the outcomes are equally likely. In such cases, the probability of each outcome is $\frac{1}{N}$. Also, for any event A, let $N(A)$ denote the number of outcomes contained in A. Then the probability of A is $P(A) = \frac{N(A)}{N}$.

2.3 Counting Techniques

When the outcomes of the experiment are equally likely, the task of computing probabilities of an event requires counting the number of elements in that event. Then the probability of an event A is

$$P(A) = \frac{|A|}{|S|},$$

where $|A| =$ cardinality of A.

The concept of permutations and combinations play an important role in computing the cardinality of A.

A Counting Technique:

The product rule: If the first element of an ordered pair can be selected in n_1 ways, and for each of these n_1 ways, the second element of the pair can be selected in n_2 ways, then the number of possible pairs is $n_1 n_2$.

Example: Playing cards have 13 face values and 4 suits. There are thus $4 \times 13 = 52$ face value/suit combinations.

Example 1. An 8-bit binary word is a sequence of 8 digits, each of which may be either a 0 or 1. How many different 8-bit words are there?

Answer: $2^8 = 256$. Here $n_1 = n_2 = \ldots = n_8 = 2$.

One instructor and 72 students are in the classroom. Find:
(a) What is the probability that no student has the same birthday as the instructor?
(b) What is the probability that no students share a common birthday?

Answer: (a): For the first question, $\frac{(364)^{72}}{(365)^{72}} = 0.82.$

(b): For the second question,

$$\frac{365 \times 364 \times 363 \times ... \times (365 - 72 + 1)}{365^{72}} < 1\%.$$

This is called “The Birthday Paradox” because it seems counterintuitive that the probability would be low in such a small number of people.

Permutations

Any **ordered** sequence of k objects taken from a set of n **distinct** objects called a permutation of size k of the objects. The number of permutations of size k that can be constructed from the n objects is denoted by $P_{k,n}$. By the product rule,

$$P_{k,n} = n(n-1)(n-2)...(n-k+2)(n-k+1) = \frac{n!}{(n-k)!}.$$

Example 3. How many different ordered sequences is possible by arranging three different objects?

Answer: $3 \times 2 \times 1 = 3! = 6$.

Combinations

Given a set of n **distinct** objects, any **unordered** subset of size k of the objects is called a combination. The number of combinations of size k that can be formed from n distinct objects will be denoted by $\binom{n}{k}$. The number of combinations of size k from a particular set is smaller than the number of permutations because, when order is disregarded, a number of permutations correspond to the same combination. Also,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{P_{k,n}}{k!}$$

Note: For example, $0! = 1$ and

$$\binom{2}{1} = \frac{2!}{1!1!} = 2; \ \binom{4}{2} = \frac{4!}{2!2!} = 2.3 = 6.$$

Example 4. A bridge hand consists of any 13 cards selected from a 52-card deck without regard to order. There are $\binom{52}{13}$ different bridge hands, which is 635013559600 (roughly 635 billion).

Example 5. A player of the California state lottery could win the jackpot prize by choosing the 6 numbers from 1 to 53 that were subsequently chosen at random by the lottery officials. The probability of winning with one ticket is 1 in $\binom{53}{6}$ or about 23 million.

Example 6. (Ex. 34) Shortly after putting to service, some buses manufactured by a company have developed cracks on the underside of the main frame. Suppose a particular city has 25 of these buses and cracks of have actually appeared in 8 of them.
(a) In how many ways a sample of 5 buses contain exactly 4 with visible cracks.
(b) If a sample of 5 buses is chosen a random what is the probability that exactly 4 of the 5 will have visible cracks.
Solution: (a)

$$\binom{8}{4} \times \binom{17}{1} = 1190$$

(b) $P(\text{exactly 4 have cracks}) = \dfrac{\binom{8}{4} \times \binom{17}{1}}{\binom{25}{5}} = \dfrac{1190}{53130} = 0.022.$

Example 7. (Ex 39) Fifteen telephones have just been received at an authorized service centre. Five of these telephones are cellular, five are cordless, and the other five are corded phones. Suppose that these components are randomly allocated the numbers $1, 2, \ldots, 15$ to establish the order in which they will be serviced.

Find
(a) What is the probability that all the cordless phones are among the first ten to be serviced?
(b) What is the probability that after servicing ten of these phones, phones of only two of the three types remain to be serviced?
(c) What is the probability that two phones of each type are among the first six serviced?

Solution: (a) We want to choose all of the 5 cordless and 5 of the 10 others, to be among the first 10 services, so the desired probability is

$$\frac{\binom{5}{5} \times \binom{10}{5}}{\binom{15}{10}} = \frac{252}{3003} = 0.0839.$$

(b) Isolating one group, say the cordless phones, we want the other two groups represented in the last 5 serviced. So we choose 5 of the 10 others, except that we don't want to include the outcomes where the last five are all the same.

So, we have $\dfrac{\binom{10}{5}-2}{\binom{15}{5}}$. But we have three groups of phones, so the desired probability is

$$\frac{3[\binom{10}{5}-2]}{\binom{15}{5}} = \frac{3(250)}{3003} = 0.2498.$$

(c) We want to choose 2 of the 5 cordless, 2 of the 5 cellular, and 2 of the corded phones:

$$\frac{\binom{5}{2}\times\binom{5}{2}\times\binom{5}{2}}{\binom{15}{6}} = \frac{1000}{5005} = 0.1998$$

Homework

Sect 2.1: 1, 3, 6, 10

Sect 2.2: 11, 14, 19, 24, 27

Sect 2.3: 29, 34, 39 (a, b), 43.