Unit 5

Probability

Overview

Like most sciences, the birth of probability can be traced to the desire to solve problems for which no reasonable method existed at the time. For centuries, scholars had been intrigued by profound issues of uncertainty, such as the existence of life on planets. Before the seventeenth century, however, relatively little had been accomplished in measuring uncertainty and the likelihood of things happening. In 1654, modern probability theory began, not as a complex issue such as the uncertainty of life, but from investigations of games of chance. The period around 1660 resulted in several contributions of probability in a variety of disciplines. (Bleasdell, 1993).

Today, the terms 'probability' and 'probably' and similar words are used quite liberally in our conversations and writings:
"He will probably drop by" or "She will probably attend the graduation ceremony";
"I'll probably go to the cinema tomorrow evening";
"The odds are slim that the Leewards Team will lift the top prize in the Caribbean's premier cricket tournament".
Such statements provide us with some idea about the likelihood that there will be a particular outcome. If we were to ask the person who makes one of these statements, or more precisely, "how probable, on a scale of 0 to 10?", the response might vary. The closer the value is to 10, the more likely the possible outcome and, conversely, the closer the value is to 0, the less likely the outcome. Of course, the scale chosen would influence the answer given – if the scale were from '0 to 100', for example, the corresponding response would be a number between 0 and 100.

Conventionally, probabilities are expressed by fractions between 0 and 1 or as percentages between 0 to 100%. In this unit, we shall define the notion of probability, demonstrate its importance and its use in statistical contexts.

Unit 5 Learning Objectives

After completing this unit, you should be able to:

- define the sample space for an experiment;
- identify events within the sample space;
- discuss the concept of probability and make use of its axioms;
- compare and contrast the three approaches to probability: the classical approach, the relative frequency approach and the subjective approach;

- identify the probability of an event from the results of an experiment;
- identify the conditional probability of an event from the results of an experiment;
- distinguish between mutually exclusive and mutually independent events;
- apply the addition and multiplication laws of probability.

The sessions in this Unit are:

Session 1: Defining the concept of the probability of an event.

Session 2: Approaches to determining the probability of an event.

Session 3: The laws of probability.

Note to students

This unit contains several activities and one practice assignment at the end of the unit. You are to work on the activities on your own. If you have any questions or concerns please post a message in the unit discussion forum so that your E-tutor can provide assistance to you. The assignment is to be uploaded in the practice assignment area.

Session 1

Defining the Concept of the Probability of an Event

Introduction

You will recall the concepts of experiment and outcomes from Unit 1. In statistics, an experiment is any activity that generates data; the results of the experiment are called outcomes. Before we discuss probability in any detailed way, we need to introduce two new concepts, namely, sample space and events. In this session we discuss these two concepts and define the concept of the probability of an event.

Objectives

After completing this session you should be able to:

- define the sample space resulting from an experiment;
- identify all possible events that belong to the sample space of an experiment;
- distinguish between simple events and compound events;
- discuss the concept of the probability of an event;
- list the axioms of probability.

Defining the Sample Space

Imagine an experiment that consisted of flipping a fair coin. The flip could only result in either the coin landing head up or landing tail up. In short, we say that the possible outcomes of this experiment are *head* and *tail* respectively. Thus the sample space will comprise the two outcomes of head and tail.

DEFINITION 5.1.1

The collection of all possible outcomes of an experiment is called a ***sample space.***

It is customary to write the outcomes as a set in which the designation for sample space is the capital letter S. Hence we write S = { Head , Tail }.

Suppose the experiment consisted of interviewing first year distance education students to measure their level of satisfaction with the registration process. To do so, we may ask some or all the students the following question:
"Which of the following statements best describes your level of satisfaction with the recently concluded registration process? :

Very Dissatisfied   Dissatisfied   Neutral   Satisfied   Very Satisfied

The possible outcomes of this experiment will be responses in the form of Very Dissatisfied, Dissatisfied, Neutral, Satisfied, and Very Satisfied. These will therefore constitute the sample space. Hence we write:
S = {Very Dissatisfied, Dissatisfied, Neutral, Satisfied, Very Satisfied}.

If instead, the experiment consisted of rolling a fair die, the possible outcomes would be die turned up Face 1, die turned up Face 2, die turned up Face 3, die turned up Face 4, die turned up Face 5, and die turned up Face 6. We abbreviate this by saying that the possible outcomes are the scores 1, 2, 3, 4, 5 and 6. These six scores comprise the sample space. Hence we write S = {1, 2, 3, 4, 5 6}.

So far we have been using the approach of a list to present a sample space. For example: S = {Very Dissatisfied, Dissatisfied, Neutral, Satisfied, Very Satisfied}; S = {Head, Tail }; and S = {1, 2, 3, 4, 5 6}

Alternatively, we can also use a *Venn Diagram* or a *Tree Diagram* to present the sample space.

Diagram 5.1
Venn Diagram for Sample Space of the Experiment of Rolling a Fair Die

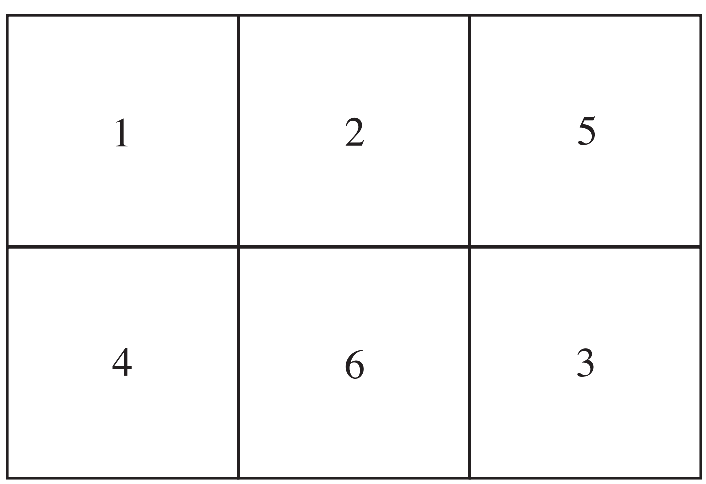
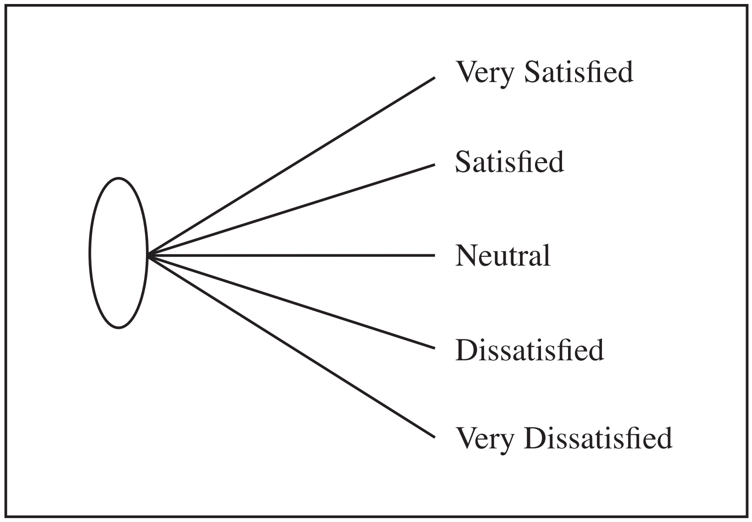
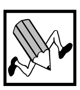


Diagram 5.2
A Tree Diagram for the Sample Space for a Satisfaction Feedback Experiment



Having defined a sample space, we are very often interested in only a subset of the sample space. Such subsets (or sub-collections) are called events.



ACTIVITY 5.1

Define the sample space from the experiment of recording the birth of two babies.

Defining an event for an experiment

Earlier we identified a sample space as all possible outcomes of an experiment. In this section we will discuss subjects of an event

DEFINITION 5.2

An ***event*** is any subset of the sample space.

Events may be simple or compound.

For example, in the experiment of rolling a fair die, the Sample Space is given by S = {1, 2, 3, 4, 5, 6}. Each outcome in this sample space is a simple event. Thus there are 6 simple events here.

> ***DEFINITION 5.3***
>
> A ***simple event*** comprises only one outcome from the sample space.

In our earlier sample space arising from the experiment of flipping a fair coin, there are two simple events, namely, the event of {head} and the event of {tail}.

For example in the experiment of rolling a fair die, the sample space is given by S = {1, 2, 3, 4, 5 6}. Each outcome in this sample space is a simple event. If we were to create events that comprise of more than one simple event, we would be creating compound events.

> ***DEFINITION 5.4***
>
> A ***compound event*** comprises more than one outcome from a sample space.

There are compound events comprised of two outcomes e.g. the event {1 , 2}; compound events comprised of three outcomes e.g. the event { 2, 4, 6}; compound events comprising four outcomes e.g. the event {1, 2, 3, 4}; and compound events comprising five outcomes e.g. the event {1, 2, 3, 4, 5}.

In our earlier sample space arising out of the satisfaction survey of the first year distance students, there were five simple events, namely the event of Very Dissatisfied, the event of Dissatisfied, the event of Neutral, the event of Satisfied, and the event of Very Satisfied. There are also compound events comprised of two of the outcomes e.g. the event of the Very Dissatisfied or Dissatisfied. There are also compound events comprising three outcomes and four outcomes.

> ***ACTIVITY 5.2***
>
> Define all the events within the sample space defined in Activity 5.1 for the experiment of recording the birth of two babies.

The common characteristic of the outcomes in an event gives the event its name. For example, the event { 2, 4, 6} from the toss of a fair die, will be named the event of getting an even score. Similarly, the event of {Satisfied, Very Satisfied} can be described as the event that the student is at least satisfied.

Now that we have defined and used sample spaces and events we can introduce probability.

Defining the Concept of the Probability of an Event

The probability of an event is a numerical measure of the likelihood that the event will occur.

Within the concepts of probability there are two extremes: (1) A probability of 0 is interpreted to mean that the event will never occur; (2) A probability of 1 is interpreted to mean that the event will always occur. Intermediate values, i.e. probabilities between 0 and 1 exclusive, indicate the existence of doubt or uncertainty about the occurrence of the particular event. These statements are subsumed in what is called the *axioms of probability.*

There are four axioms of probability:

1. The probability of an event A (given by P(A)) is a fraction or percentage between 0 and 1 inclusive. Put differently, $0 \leq P(A) \leq 1$.

2. $P(A) = 0$ if and only if the event A never occurs.

3. $P(A) = 1$ if and only if the event A always occurs.

4. If the event A is a compound event and comprised of a finite number (say *n*) of simple events $\{A_1, A_2, A_3, \ldots\ldots A_n\}$, then

$$P(A) = P(A_{1},) + P(A_2) + P(A_3) + \ldots\ldots + P(A_n).$$

Now that we have defined probability and its axioms, we must be curious about the question "How do we determine the probability of an event?" This question will be answered in the following session.

One immediate implication of this definition is that if we are given two events A and B and A is more likely to occur than B then $P(A) > P(B)$. The reverse is also correct, if given two events A and B and the statement that $P(A) > P(B)$, we can conclude that A is more likely to occur than B.

Summary

In this session we defined a sample space as the collection of all possible outcomes of an experiment and the event as a subset of a sample space. We pointed out that there are simple and compound events. It was important to define these concepts before moving on to the concept of probability. The probability of an event is a numerical measure of the likelihood that the event will occur. In looking at the concept of probability we discussed the four axioms of probability.

Session 2

Approaches to Determining the Probability of an Event

Introduction

Statistics provide us with an objective means of determining the outcome of an experiment. The concept of probability is an important element in arriving at an outcome. There are three approaches to determining the probability of an event, namely the *classical approach*, the *relative frequency approach*, the *subjective approach*. These approaches span the continuum from subjectivity to objectivity. In this session, we shall take a focused look at two examples. In each example, we shall attempt to identify sample space and simple events, assign probabilities by applying one or more of the approaches of probability, and demonstrate that the axioms of probability are satisfied.

Objectives

On completing this session students should be able to:

- select among the classical, relative frequency and subjective approaches for determining the probability of an event;
- apply each of the three approaches correctly and appropriately;
- interpret data from an experiment to determine the sample space, events and their corresponding probabilities;
- demonstrate that the axioms of probability are satisfied in a given circumstance.

The Classical Approach

This approach is applied to experiments in which all outcomes of the sample space are equally likely, e.g. the toss of a 'fair' coin, the roll of a 'fair' die and the sex of a newborn at birth. Under this approach, the probability of a simple event equals 1 divided by the total number of simple events in the sample space for the experiment.

DEFINITION 5.5

Thus if "A" is a simple event from a sample space S of a given experiment, the ***classical approach*** determines the probability of A by the formula.

$$P(A) = \frac{1}{\text{Number of Simple Events in the Sample Space S}}$$

Consider the toss of a 'fair' coin, the sample space is {head, tail}. By the 'fairness' of the coin, the outcomes of head and tail are equally likely to occur. Thus
P(head) = ½ and P(tail) = ½.

Consider the roll of a 'fair' die, the sample space is { 1, 2, 3, 4, 5, 6 }. By the 'fairness' of the die, the outcomes of 1, 2, 3, 4, 5, and 6 are equally likely to occur. Thus
P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = ⅙.

In both these examples, the probabilities were clear once we knew how many simple events comprised the sample space. Probabilities determined by the classical approach are therefore called *apriori* probabilities.

Using the fourth axiom of probability, we can immediately conclude that if B is an event comprised of say three simple events $\{A_1, A_2, A_3\}$ from a sample space S, the probability of event B is given by

$$P(B) = P(A_1,) + P(A_2) + P(A_3).$$

$$= \frac{3}{\text{Number of Simple Events in the Sample Space S}}$$

We can generalise this as follows:

$$P(B) = \frac{\text{number of simple events in event B}}{\text{number of simple events in S.}}$$

ACTIVITY 5.3

Using the classical approach:

(1) Determine the probability of getting an odd score when a fair die is rolled.

(2) Determine the probability of getting at least one head in the two tosses of a 'fair' coin.

There is one shortcoming of the classical approach and it is that few situations, if any, in the real world qualify as giving rise to equally likely outcomes. As such, it is not practical for several situations found in the real world. Consider for example the probability that a student passes a course such as *Mathematics for Social Sciences* ECON1003. The sample space here is comprised of **Pass** and **Fail**. However, the pass rate is not 50%; thus the events are not equally likely.

We therefore need an approach that can be applied to most situations in the real world. That approach is the relative frequency approach.

The Relative Frequency Approach

Suppose that we wanted to compute any of the following probabilities:

- the probability that a family selected at random owns two cars;
- the probability that an 80 year old person will live for at least one more year;
- the probability that the toss of an unbalanced coin will result in a 'head'.

The underlying experiments for these situations are as follows:

i. The experiment of selecting a family at random and recording whether the family owns 2 cars or not;
ii. The experiment of selecting an 80-year-old person today and checking in one year's time whether that person is still alive;
iii. The experiment of tossing an unbalanced coin and recording the result of the toss.

The classical approach will not be applicable in any of these experiments since the outcomes for these experiments are not equally likely.

Instead we recognize that

- each of these experiments can be performed repeatedly;
- after a large number of repetitions we can compute the frequency of occurrence of the given event and hence the relative frequency of the given event.

The relative frequency approach uses this relative frequency as an approximation to probability once the number of repetitions of the experiment is large enough. This approach defines probability as follows:

DEFINITION 5.6

The ***relative frequency approach to probability*** states as follows:

"if an experiment is repeated a sufficiently large number of times, the relative frequency of a particular event equals the probability of that event".

Typically, in the real world, constants of time and cost, for example, sometimes make it impossible to repeat an experiment a very large number of times. Hence some refinement is necessary, namely, if an experiment is repeated N times and an event B occurs n of these times, then the probability of event B is given by:

$$P(B) = {}^{n}/_{N}$$

ACTIVITY 5.4

Discuss how you would use the relative frequency approach to determine the probability that an ECON1005 student selected at random from your class is male.

Post your response in the relevant discussion forum

Both the classical and relative frequency approaches are considered to be objective approaches.

There are however experiments in which the equally likely property of the outcomes in the sample space is not present and/or it is impractical to generate a large number of repetitions either, due to cost considerations, the destructive nature of the experiment, or little or no history of such experiments. In such situations, neither the classical nor the relative frequency approach will be applicable; we must look to the only remaining approach which is subjective by nature.

Subjective Probability

Suppose we wanted to compute the probability that

- scientists will find a cure for the common cold by 2010,
- an oil exploration company will find oil off Grenada.

Both these experiments have neither equally likely outcomes nor the potential of several repetitions. Accordingly, neither the classical nor the relative frequency approach will apply.

Instead the practice is to refer to experts for an opinion on the chance of occurrence of the stated event. The probability of the stated event will be defined as the modal response from the 'experts'. This approach is subjective and influenced primarily by the 'experts' selected.

DEFINITION 5.7

The ***subjective approach*** to probability is an approach in which the probability of an event is assigned on the basis of subjective judgement, belief, experience and/or information. The assignment is arbitrary and influenced by the biases, preferences and experience of the person(s) assigning the probability.

Identifying Available Probabilities from the Results of an Experiment

EXAMPLE 5.1

Consider the dataset data.mtw which provides some demographic data and examination marks for 137 full time first year students of the Faculty of Social Sciences in a range of Level I courses. This dataset reflects a gender split of 92 female and 45 male students.

Let us define our experiment as that of selecting one student at random and recording the student's sex.

Thus the sample space S = { male, female }

For a moment, let us focus on the event {female} and its probability.

If the classical approach were applicable, we could say that

$$P(\{\text{female}\}) = \frac{\text{No. of simple events in S that are female}}{\text{No. of simple events in the sample space S}} = {}^{92}\!/_{137} \approx 67\%$$

However this approach does not necessarily apply to this experiment.

Instead we can apply the relative frequency approach. In this approach,

- we repeat the experiment with replacement several times;
- as we repeat the experiment, we would check the relative frequency of a female outcome;
- as the number of repetitions gets sufficiently large, the relative frequency will settle down to 67%.

Hence we say by the relative frequency approach that P({Female}) = 67%.

In a similar way, we can show that the P({Male}) = 33%.

Looking back at the results we see that:

- the sample space S = { male, female}
- the probability attached to each outcome is a positive fraction less than 1

  $0 \le P(\{\text{male}\}) \le 1$

  $0 \le P(\{\text{female}\}) \le 1$

- the sum of the probabilities of the two simple events that exhaust the sample space is given by

  $P(\{\text{male}\}) + P(\{\text{female}\}) = 1$

EXAMPLE 5.2

Consider the file data.mtw again. This time we focus on the degree options pursued by the full time students. Table 5.1 summarises the raw data on degree options.

Table 5.1

Frequency Distribution of Degree Options of Full Time Social Science Students

Discipline	Outcome /Simple Event	Frequency
Accounting	A_1	16
Management	A_2	55
Economics	A_3	28
Soc/Mgmt	A_4	20
Other	A_5	18
		Total = 137

source : data.mtw

We can assign probabilities to the degree options on either the classical or relative frequency approach provided that the necessary conditions are present. Either way we should get probabilities as shown in the next table:

Table 5.2

Probabilities of Degree Options of Full Time Social Science Students

Discipline	Outcome/ Simple Event	Probability
Accounting	A_1	0.12
Management	A_2	0.40
Economics	A_3	0.20
Soc/Mgmt	A_4	0.15
Other	A_5	0.13
		Total = 1.00

Note the following:

- The sample space for this experiment is $S = \{ A_1, A_2, A_3, A_4, A_5)\}$
- For each outcome/simple event A_i we observe that

  $1 \geq P(A_i) \geq 0 \qquad i = 1,2,3,4,5$
- the sum of the probabilities of all outcomes/simple events that comprise the sample space is given by

$$P(A_1) + P(A_2) + P(A_3) + P(A_4) + P(A_5) = 1.$$

Once again, the axioms of probability hold true.

Summary

The classical, relative frequency and subjective approaches to determining the probability of an event were discussed in this session. The classical approach was defined as an approach applied to experiments in which all outcomes are equally likely. The relative frequency approach, unlike the classical approach, is more applicable to real world situations. It is used in cases where an experiment is repeated a sufficiently large number of times so the relative frequency of a particular event equals the probability of that event. Subjective probability was defined as the modal response of experts. Probability in this instance is assigned on the basis of selective judgement.

Session 3

Laws of Probability

Introduction

So far we can determine the probability of an event in situations where the event relates to one characteristic. Very often we are required to find probabilities of compound events that relate to two or more characteristics. Consider for example Table 5.3 below in which summary data is presented for the age of the suitcase traders by sex. If we select a trader randomly, we can focus on events like "the trader selected is either male or 41-50 years old" or "the trader selected is female and 31-40 years old".

Table 5.3
Age Distribution of Suitcase Traders by Sex

	Less than 21	21 - 30	31 – 40	41 – 50	51 - 60	Row Total	Row Percent
Male	1	11	5	0	0	17	27.4
Female	2	12	15	11	5	45	72.6
Column Total	3	23	20	11	5	62	100
Column Percent	4.8	37.1	32.3	17.7	8.1	100	

Source: Franklin, Hosein & Persaud (2007)

Finding the probabilities of these two events requires additional tools. These tools are called the laws of probability. In this session we introduce some definitions, notation and jargon, before discussing the laws of probability.

Objectives:

At the end of this session, students must be able to:

- identify the complement of an event;
- differentiate between the addition law and multiplication law of probability;
- select the appropriate law to be used in a given situation;
- recognize when events are mutually exclusive;
- recognize when events are mutually independent;
- compute conditional probabilities;
- apply the laws to compute probabilities.

Complement of an Event

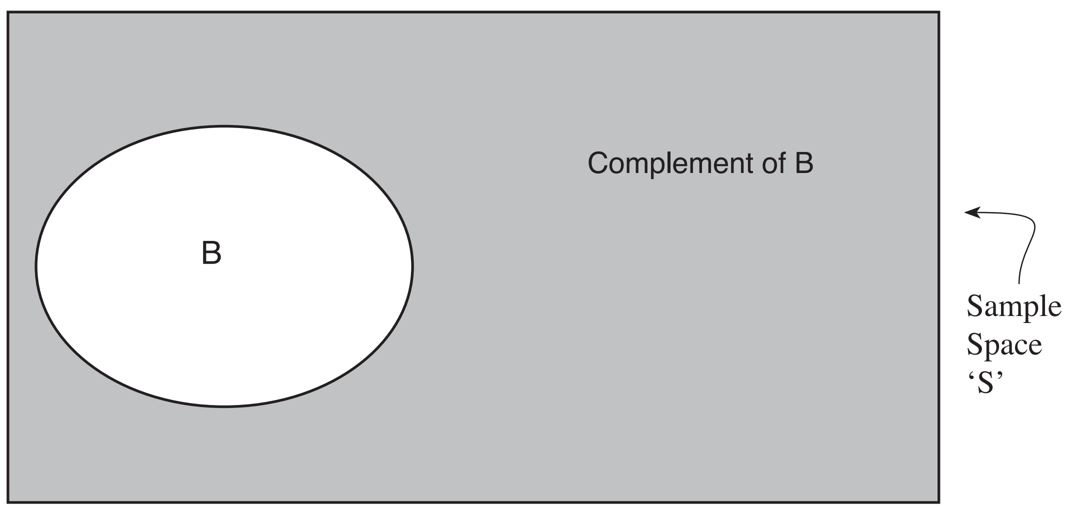
In preparation for the laws of probability we must utilise some concepts which you would have met at secondary school level in preparation for the Caribbean Examinations Council Exam (CXC) or in the Vacation Student Support Programme (VSSP) done by students in Trinidad and Tobago. The complement of a set, the null set, the intersection of a set, the union of sets and mutually exclusive sets. Further we will replace the word set with event.

DEFINITION 5.8

Given an event B in a sample space, the **complement of an event** B (written B^c) is the event that B does not occur. Hence $P(B^c)$ = P(event B does not occur).

The complement of an event B is represented by the shaded portion of Diagram 5.3

Diagram 5.3
Complement of an Event



For example, in the toss of a 'fair' coin, the sample space is {head, tail}.

The complement of the event {Head} is the event {tail}and vice versa.

Another example is the roll of a 'fair' die. Here the sample space is { 1, 2, 3, 4, 5, 6 }. The complement of the event {1, 3, 5} is the event { 2, 4, 6 }. The former is called the event that the score is odd; the latter is referred to the event that the score is even.

DEFINITION 5.9

The **null event** Ø is the event which comprises no outcomes from the sample space. Put differently, it is the event that never occurs. Hence P(Ø) = 0.

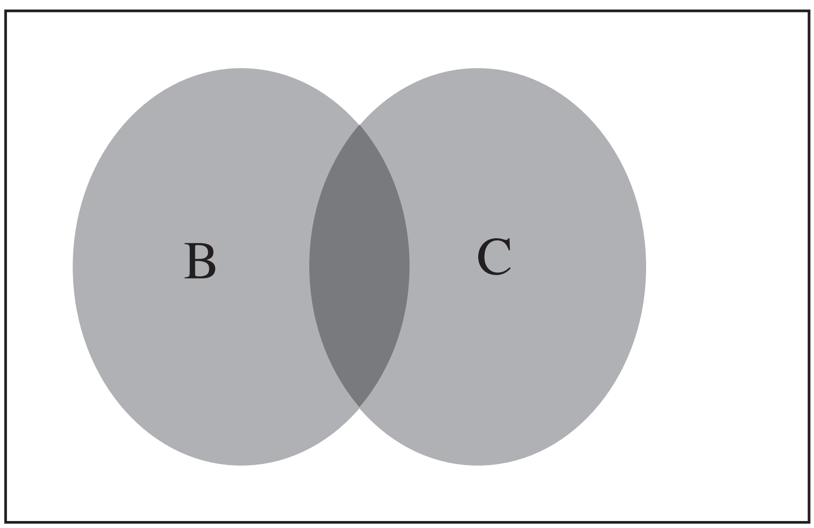
For example, let us consider the experiment of selecting a primary school student and measuring his/her age. Suppose we were interested in finding the probability that a primary school student is younger than 2 years old. The event of a primary school student younger than 2 years old does not occur in a primary school since a child must be at least 4 years to enroll in a primary school. Such an event is called a *null event*. Further, its probability is 0.

DEFINITION 5.10

The ***union of two events B and C*** is the event that either B alone occurs or C alone occurs or both B and C occur. It is usually referred to as B or C and comprises the events of the sample space that belongs to 'B' alone or 'C' alone or both 'B' and 'C'. (See diagram 5.4)

Diagram 5.4 below shows event 'B' as the circular set to the left, event 'C' as th circular set to the right and event 'B' or 'C' as the total shaded Set.

Diagram 5.4
Union of Two Events.

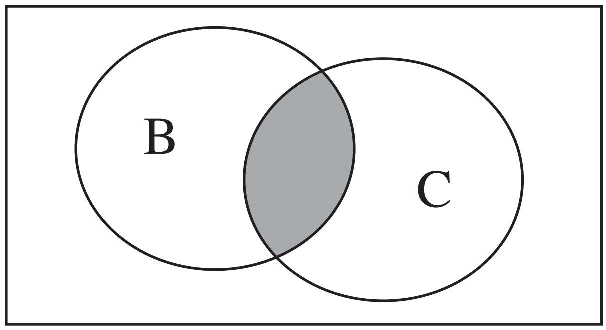


For example, if we return to the situation and the summary data in Table 5.3, we may want to focus on the event "the trader selected is either male or 31-40 years old". We can decompose this event so as to showcase two events viz. Event B = "the trader selected is male" and Event C = "the trader selected is 31-40 years old". Thus the event "the trader selected is either male or 31-40 years old" is the narrative version of the event B or C.

DEFINITION 5.11

The ***intersection of two events B and C*** is the event that both B and C occur. It is usually referred to as B and C and comprises the elements of the sample space that belong to both B and C.

Diagram 5.5
The Intersection of Two Events.



For example, if we return to the situation and the summary data in Table 5.3, we may want to focus on the event "the trader selected is female and 31-40 years old". We can decompose this event so as to showcase two events viz. Event B = "the trader selected is female" and Event C = "the trader selected is 31-40 years old". Thus the event "the trader selected is female and 31-40 years old" is the narrative version of the event B and C.

One immediate implication of the Diagrams 5.4 and 5.5 is that

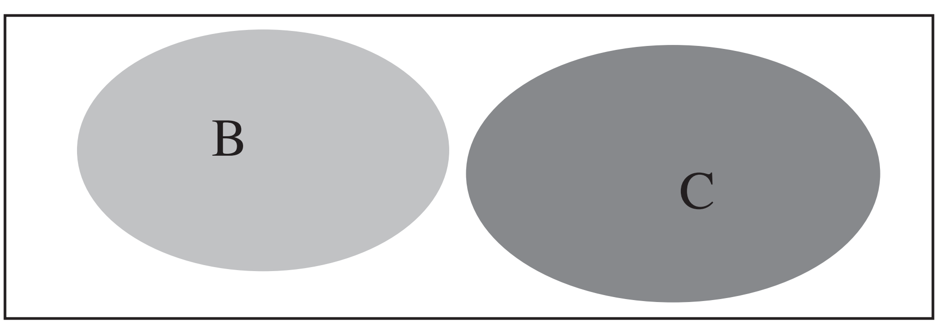
# of outcomes in B or C = # of outcomes in B + # of outcomes in C
– # of outcomes in B and C.

(This will be derived later in this Session)

DEFINITION 5.12

Two events B and C from the same sample space are said to be ***mutually exclusive events*** if they have no outcomes in common. In other words, the intersection of B and C is the null event. Hence the P(B and C) = P(Ø) = 0

Diagram 5.6
The Intersection of Two Mutually Exclusive Events.



Return for a moment to the example of the roll of a 'fair' die. Here the sample space is { 1, 2, 3, 4, 5, 6 }. The event {1, 3, 5} and the event { 2, 4, 6 } are mutually exclusive since these do not have any score in common. Put another way, you could not have an odd score and at the same time have an even score.

Yet another example, return to Table 5.3. The events {male} and {female} are mutually exclusive. So too are the events {trader selected is 21-30 years old} and {trader is 41-50 years old}.

Note that mutually exclusive events are also called *disjoint events.*

ACTIVITY 5.5

Identify at least three (3) more pairs of **mutually exclusive events** in Table 5.3
.

Addition Law of Probability

With these definitions, we can proceed to compute the probability that "the trader selected is either male or 31-40 years old". Recall that we can decompose this event so as to showcase two events, namely,

Event B = "the trader selected is male" and Event C = "the trader selected is 31-40 years old". Thus the event "the trader selected is either male or 31-40 years old" is the narrative version of the event B or C. We therefore wish to compute P(B or C).

Using the classical approach to probability we can say that

$$P(B \text{ or } C) = \frac{\#.\text{ of traders who are B or C}}{\text{Total }\#.\text{ of traders in the Sample Space S}}$$

We may be tempted to say that the # of traders who are either B or C , i.e. the # of traders who are either male between 31-40 years old equals the sum of the # of traders who are male and the # of traders who are 31-40 years old.

If we do so, we have double counted the # of traders who are both Male and 31-40 years old. Hence we must subtract the # of traders who are Male and 31-40 years old from the sum above.

Thus we say that:

# of traders who are either male between 31-40 yrs old = # of traders who are male add # of traders who are 31-40 yrs old minus # of traders who are male and 31-40 yrs old.

Since Event B = "the trader selected is male" and Event C = "the trader selected is 31-40 years old" we can rewrite the above statement as follows:

# of traders in B or C = # of traders in B + # of traders in C – # of traders in B and C.

In Table 5.3 the # of traders in B or C = 17 + 20 -5 = 32. Also the total of traders in the sample space is 62.

$$\text{Hence } P(B \text{ or } C) = \frac{17 + 20 - 5}{62} = \frac{17}{62} + \frac{20}{62} - \frac{5}{62} = \frac{32}{62}$$

The first fraction $^{17}/_{62}$ is P(B) ; the second fraction $^{20}/_{62}$ is P(C) and the third fraction $^{5}/_{62}$ is P(Band C).

Thus P(B or C) = P(B) + P(C) - P(B and C)

We can generalise this result by the definition.

DEFINITION 5.13

The ***addition law of probability*** states that given any two events B and C from a sample space,

P(B or C) = P(B) + P(C) - P(B and C)

Note that the 'operative' word here is 'or'.

In the discussion that preceded the definition above, P(B) = $^{17}/_{62}$; P(C) = $^{20}/_{62}$ and

P(Band C) = $^{5}/_{62}$. Thus we computed

P(B or C) = P(B) + P(C) - P(Band C) = $^{17}/_{62}$ + $^{20}/_{62}$ - $^{5}/_{62}$ = $^{32}/_{62}$.

There is a special case of the addition law of probability and this applies when the events B and C are mutually exclusive. When B and C are mutually exclusive, P(B and C) = 0 by Definition 5.12. When P(Band C) = 0 in Definition 5.13, the addition law of probability reduces to a simple addition of P(B) and P(C).

DEFINITION 5.14

The **special case of the addition law of probability** states that given any two mutually exclusive events B and C from a sample space,

P(B or C) = P(B) + P(C)

For example, return to Table 5.3 and focus on the event "the trader selected is either male or 41-50 years old". We can decompose this event so as to showcase two events viz. Event B = "the trader selected is male" and Event C = "the trader selected is 41-50 years old". Thus the event "the trader selected is either male or 41-50 years old" is the narrative version of the event B or C.

Here P(B) = $^{17}/_{62}$; P(C) = $^{11}/_{62}$ and P(Band C) = 0.

By Definition 5.12, the fact that P(B and C) = 0 is just an alternative way of saying that B and C are mutually exclusive.

Thus we compute P(B or C) = P(B) + P(C) = $^{17}/_{62}$ + $^{11}/_{62}$ 2 = $^{28}/_{62}$.

ACTIVITY 5.6

Recall Table 5.3. A trader is selected at random, find the probability that the trader is female or 51-60 years old.

Multiplication Law of Probability

So far all our probabilities have been unconditional. But if we return to Table 5.3, we can change our focus to events such as "the trader is 21-30 years old given that the trader is male". The probability of such an event is called a conditional probability. In other words, if we let Event B = "the trader is 21-30 years old" and event C = "trader is male", then we seek P(B given C). Conditional probabilities are always recognized by phrases such as 'given that', 'in light of', 'in view of' etc.

Conditional probability P(B/C) restricts our attention to the event C that has already occurred. As such, we treat the event C as a 'restricted' sample space. The probability that event B will occur in the 'restricted' sample space is the conditional probability.

In Table 5.3, the restricted sample space comprises only 17 traders and these are found in the first row of the table. This row is reproduced below as Table 5.4

Table 5.4
Restricted Sample Space of Male Traders

	Less than 21	21 - 30	31 – 40	41 – 50	51 - 60	Row Total
Male	1	11	5	0	0	17

From Table 5.4, it is clear that there are 11 traders who are of age 21-30 years. Hence the probability of 21-30 years in this restricted sample space is $^{11}/_{17}$.
We therefore conclude that P(B given C) = $^{11}/_{17}$.

The fraction $^{11}/_{17}$ is the same as the result of dividing $^{11}/_{62}$ by $^{17}/_{62}$; the former is the P(B and C) while the latter is P(C). Thus we can conclude that

$$P(B \text{ given } C) = \frac{P(B \text{ and } C)}{P(C)}$$

We can now generalise to the next definition.

DEFINITION 5.15

Given two events B and C of the same sample space, the ***conditional probability*** of B given C, interpreted i.e. the probability that an event B will occur given that another event C has already occurred, and written P(B I C), is given by the formula

$$P(B \mid C) = \frac{P(B \text{ and } C)}{P(C)}$$

In the discussion before the definition above, P(B and C) = $^{11}/_{62}$ and P(C) = $^{17}/_{62}$; we computed P(B | C) as $^{11}/_{17}$.

ACTIVITY 5.7

Recall Table 5.3. Let Event B = “the trader is 21-30 years old” and event C = “trader is male”, compute the P(C given B) and show that

$$P(B \mid C) = \frac{P(B \text{ and } C)}{P(B)}$$

If we cross multiply the formula in Definition 5.15, we get the following:

$$P(B \text{ and } C) = P(B \mid C) \times P(C).$$

This result is formally known as the *multiplication law of probability.*

ACTIVITY 5.8

Refer to the result in Activity 5.7. Use cross multiplication to show that

$$P(B \text{ and } C) = P(C \mid B) \times P(B).$$

DEFINITION 5.16

The ***multiplication law of probability*** states that given any two events B and C from the same sample space,

$$P(B \text{ and } C) = P(B \mid C) \times P(C).$$

Alternatively

$$P(B \text{ and } C) = P(C \mid B) \times P(B).$$

Note here that the ‘operative’ word is ‘and’.

DEFINITION 5.17

Two events B and C from the same sample space are ***mutually independent*** if the occurrence of one event does not affect the probability of occurrence of the other. That is,

$$P(B \mid C) = P(B) \quad \text{and} \quad P(C \mid B) = P(C).$$

Whenever the events B and C are mutually independent we can replace $P(B \mid C)$ in Definition 5.16 by P(B) and $P(C \mid B)$ by P(C). Hence we get the following special case of the multiplication law of probability.

DEFINITION 5.18

The ***special case of the multiplication law of probability*** states that given any two events mutually independent B and C from the same sample space,

$$P(B \text{ and } C) = P(B) \times P(C).$$

A word on the logic for application of the two laws of probability

1. Read the question.
2. Identify the operative word 'or' or 'and'.
3. If 'or', focus on the addition law; if 'and', focus on the multiplication law.
4. If the addition law, check whether the events are mutually exclusive.
5. If the events are mutually exclusive then use the special form of the addition law; if not, then use the general form of the addition law and go to 8.
6. If the multiplication law, check whether the events are mutually independent.
7. If the events are mutually independent, then use the special form of the multiplication law; if not, then use the general form of the multiplication law and go to 8.
8. State the law.
9. Identify the event B and the event C.
10. Identify the probabilities for B alone, C alone, B and C, B given C, and C given B as demanded by the law.
11. Insert the required probabilities into the law to compute the required probability.

EXAMPLE 5.3

In 2002, McDonald's had 31,108 restaurants systemwide. Of these, 17,864 were operated by franchises, 9,000 by the company, and 4,244 by affiliates. What is the probability that a randomly selected restaurant is either a franchise or an affiliate? Source: McDonald's Corporation 2002 Annual Report. pg 1.

This example presents three events; these are as follows:
Event A – restaurant is operated by the company
Event B – restaurant is a franchise
Event C – restaurant is a franchise.

The probability that we are required to compute is P(B or C).

The conjunction 'or' points us to the Addition Law.

Which form of the Addition Law do we use here? The answer lies in whether the events B and C are mutually exclusive. The very definitions of 'franchise' and 'affiliate' tell us that a restaurant cannot be both. Thus Events B and C are mutually exclusive. This is further corroborated by the fact that if we add the franchises to the owned operated and then add the affiliates we get 17864 + 9000 + 4244 = 31,108.

Accordingly, we use the short form of the Addition Law i.e.
P(B or C) = P(B) + P(C).

We need values for P(B) and P(C). From the McDonald's 2002 Annual Report we discern that P(B) = 17864/31108 = 0.574 and P(C) = 4244/31108 = 0.136

Thus P(B or C) = 0.574 + 0.136 = 0.71

ACTIVITY 5.9

Two thousand randomly selected adults were asked if they were financially better off than their parents.

The table below gives a summary of the responses based on the education levels of the persons included in the survey and whether they are financially better off, the same, or worse off than their parents.

	Less than High School Education	High School Education	More than High School Education
Better off	140	450	420
Same	60	250	110
Worse off	200	300	70

Suppose one adult is selected at random from these two thousand adults.

cont'd

ACTIVITY 5.9 cont'd

1. Identify the experiment.
2. Identify the number of repetitions of the experiment.
3. Identify the simple events.
4. Identify the mutually exclusive events.
5. Find the following probabilities:
    a. P(better off or high school)
    b. P(more than high school or worse off)
    c. P(better off or worse off)
    d. P(better off)
    e. P(high school)
    f. P(better off and high school)
    g. P(high schooll better off)
6. Are the events {high school} and {better off} mutually independent?

Summary

In this session we introduced the concepts of the complement of an event, a null event, the union of two events and the intersection of two events. We also discussed mutually exclusive events and introduced the addition and multiplication rules of probability and discussed their applications in simple and complex cases.

Wrap Up

In this unit, we defined the notion of probability and stated its axioms. Since probability exists for events, we focused on simple events and compound events and the three approaches to assigning probabilities to events. We further defined mutually exclusive events and mutually independent events.

The unit ended with the introduction of the addition law of probability and the multiplication law of probability and a logic for application of these laws.

References

Bleasdell, E.A. 1993. *Statistics in Practice*, Philadelphia, PA. Saunders Collage Publishing, Harcourt Bruce College Publishers, 1993.

Franklin, M , Hosein, R. & Persaud, A. 2007. ICI in the Caribbean: Evidence from Trinidad and Tobago and Guyana. (unpublished paper).