Module 6
Probability Basics

<u>Module Learning Outcomes</u>

- Interpretation of probabilities in general.
- Perform basic probability calculations by counting with the help of two-way tables.
- Perform probability calculations using formula, like complementary rule, addition rule, conditional probability, and multiplication rule.
- Apply the use of conditional probability and Bayes' Theorem in probability calculations.
- Show two events are mutually exclusive using probability calculations.
- Show two events are independent to each other using probability calculations.

6.1 Probability and Statistics

- You probably have heard that these two topics always go together.
- In fact, the use of statistics would be very limited without probability, which is where we are at this point of the course – getting samples, collecting data and reporting statistics etc.
- However, our main goal in statistics is to know something about the population.
  For example, knowing everything about my random sample of size 1,000 Canadians about how they would vote in the next federal election is not very useful because we want to know how all Canadians (the population) would vote instead.
- But with probability, we can link, with some certainty, what is happening in samples to what we want to know in the population.
- Again, we could only make a good link when the data collected are from random samples.

6.2 Basic Probability Terminology

- An **experiment** is a process that generate a set of non-overlapping but well-defined **outcomes** or **sample points**.
- Note: In this course and probability in general, we will only consider experiment that generates two or more outcomes. Otherwise, things are not too interesting to discuss. Does it sound like one type of variable that we know of?
- By definition, outcomes are always **mutually exclusive**. That means only one outcome will appear in one experiment.
- The **sample space** is a collection of all possible outcomes and it is usually denoted by S (S from sample Space) or U (U from Universe).
- An **event** is simply a collection of one or more outcomes, but it rarely contains all outcomes.
- Note: It is a probability convention that we write events in block letters, starting with A, B, C etc.
  Example: Consider an experiment of "rolling a fair die" (note that "die" is a singular form of "dice").

Q: List out all possible outcomes.	A: 1, 2, 3, 4, 5, 6
Q: Write down the sample space.	A: {1, 2, 3, 4, 5, 6}
Q: Define an event A as "getting an even number". List out all outcomes in event A.	A: 2, 4, 6
Q: Define an event B as "getting an odd number". List out all outcomes in event B.	A: 1, 3, 5
Q: Define an event C as "getting a multiples of three". List out all outcomes in event C.	A: 3, 6

6.3 Interpretation of Probability and the Probability Rules

- The probability interpretation is straight-forward once you know what to include.
- **Rule #1**: Probability is always between 0 and 1, inclusively.
- A probability of zero means that there is no way that thing is going to happen.
- From an interpretation point of view, it means that <u>when the experiment is repeated infinite number of times, 0 times or 0% of the time that this event will occur</u>.
- A probability of one means that something is certain to happen.
- From an interpretation point of view, it means that <u>when the experiment is repeated infinite number of times, 100% of the time that this event will occur</u>.
- In daily life, a 50% probability is like the "mid-point" between certainty and impossibility. In any election, a candidate would be thrilled to get 51% of the votes. When you participate in a random draw, you would be delighted if you were told that you had more than 50% chance of winning the jackpot.
- But once you have look at its interpretation "<u>when the experiment is repeated infinite number of times, 50% of the time that this event will occur</u>", it suddenly occurs to you that 50% of the time is not quite good enough. It is like students would not be happy if they only got 50% in a test. And your client would not be very happy to hear that there is only 50% chance your research results are correct.
- In statistics, we have a higher standard. How high? It really depends on the person. But for a proper research, we would not want anything under 90% probability.

- **Rule #2**: The sum of probability of all possible outcomes (from the same experiment) is always one.
- This rule is particularly handy in handling situation with complementary events (more later).

Example: Consider an experiment tossing three fair coins. The following table shows the number of heads and the corresponding probabilities. Find X.

Number of Heads	0	1	2	3
Probability	0.125	0.375	0.375	X

Answer: Using probability rule #1, we know that X must be positive and between 0 and 1. Using probability rule #2, we know that the sum of the above four probabilities must be 1, i.e. 0.125 + 0.375 + 0.375 + X = 1 → X = 0.125.

- MLO: Interpretation of probabilities in general.

6.4 Fundamental Probability Equation

- Most probability questions can be solved by using this simple but very powerful **fundamental probability equation**.
- Define A as an event, then the **probability of the event** A is calculated by

$$P(A) = \frac{number\ of\ items\ in\ A}{total\ number\ of\ items}$$

- Note: The "items" could be anything as long as it refers to the same thing in both the numerator and denominator. In this course, "items" could be "subjects", "outcomes" or even "samples" at times, depending on the situation.
  Example: Consider an experiment of "rolling a fair die". Define event A as "getting an even number"; event B as "getting an odd number"; event C as "getting multiples of three".

Q: Find P(A).	A: Because there are a total of six outcomes (6 items) when we roll a fair die (so "6" goes to the denominator) and event A has three possible outcomes (3 items in A), we have P(A) = 3/6 = 1/2 or 0.5. Note that "outcomes" happens in both the numerator and denominator.
Q: Find P(C).	A: Because event C has two outcomes (so "2" goes to the numerator), we have P(C) = 2/6 = 1/3 or 0.33.

- MLO: Perform basic probability calculations by counting with the help of two-way tables.

6.5 Complementary Events and its Probability

- A **complementary event** is an event containing all outcomes in the sample space that are not in the event of concern.
- There are a few notations: $\bar{A}$ (say "A bar"), A' (say "A prime") or even A^c (say "A complement").
- One way to find the probability associating with complementary event is to use the fundamental probability equation, by counting the number of items going to the numerator (top) of the equation.
- Or, if you know the probability of the event A, then the following formula gives you the probability of a complementary event (A^c) is
$$P(A^c) = 1 - P(A)$$
- Note: This rule makes use of probability rule #2, where the sum of all outcomes, like P(A) and P(A^c), is always one.
  Example: Consider an experiment of "rolling a fair die". Define event A as "getting an even number"; event B as "getting an odd number"; event C as "getting multiples of three".

Q: Use the complementary event formula to find P(B). A: Note the events A and B are complementary events. Therefore, P(B) = 1 – P(A) = 1 – 0.5 = 0.5.
Q: Use the complementary event formula to find $P(C^c)$. Answer: $P(C^c) = 1 - P(C) = 1 - 0.33 = 0.67$.

- MLO: Perform probability calculations using formula, like complementary rule, addition rule, conditional probability, and multiplication rule.

6.6 Intersection of Two Events and its Probability

- So far, we have looked at examples with one event only. But life would be boring if we only considered one event at a time. (Recall that looking at a single variable was boring too. That was why we went on with bivariate analysis. See the resemblance?)
- For any two events A and B, the **intersection of two events** is defined as an event that contains all outcomes in the sample space that are in both event A **and** event B.
- There are a few notations: $A \cap B$ (in mathematics set theory), $A \wedge B$ (mostly in computing science or logic course) or even $A\ and\ B$ (in statistics).
- Intersection has the notion that two events happening together or at the same time.
- To recognize intersection, look for the three main keywords: **AND**, **BUT**, **BOTH**.
- Sometimes, you may not be able to find any of these three words, in that case, you will have to think if the multiple events have to happen at the same time.
  Example: Define event A as being good at academic at high school and event B as being good in sports at high school. Suppose the requirement of one scholarship is to be good in schoolwork and sports. The keyword here is AND. In other words, we want "$A\ and\ B$".

6.7 Union of Two Events and its Probability

- For any two events A and B, the **union of two events** is defined as an event that contains all outcomes in the sample space that are in event A, in event B, or in both.
- There are a few notations: $A \cup B$ (in mathematics set theory), $A \vee B$ (mostly in computing science or logic) or even $A\ or\ B$ (in statistics).
- Union has the notion that either one of the two events, or both events happening. Another way of saying that is "at least one of the two events happening".
- To recognize union, look for the following combos: **EITHER ONE**, **AT LEAST ONE**.
- Note: In statistics, the combo "neither-nor" is a complementary event to "either-or".
  Example: Define event A as being good at academic at high school and event B as being good in sports at high school. Suppose the requirement of one scholarship is to be good in either schoolwork or sports. The keyword here is OR. In other words, we want "A or B".

6.8 Addition Rule

- Besides the fundamental probability equation, this general form of the **addition rule** (or law) is another way to find the probability of union, and it is given by:
$$P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$$
- Please note that if the any three of the four terms are given, we could also use this formula to find the remaining term.
- MLO: Perform probability calculations using formula, like complementary rule, addition rule, conditional probability, and multiplication rule.

6.9 Conditional Probabilities

- So far, all (general) probabilities make use of the sample space as the base (or denominator). **Conditional probability** allows you to focus on a sub-group of the sample space.
- The formula of conditional probability is:

$P(A \mid B) = \frac{P(A\ and\ B)}{P(B)}$ or $P(B \mid A) = \frac{P(A\ and\ B)}{P(A)}$, as long as the denominator is not zero.

- We say $P(A \mid B)$ as "the probability of A given B".
- In terms of meaning, given the event B has already occurred or observed, the conditional probability of $P(A \mid B)$ tells you how likely event A will occur.
- Another way of thinking of it is among all the subjects that are all in the event B, it represents the percentage of them are also in the event A.
- One important note about conditional probability: <u>Never add conditional probabilities directly</u>.

  Example: Suppose Tony is running for the class representative (against another student) in his Grade 5 class. After surveying all students in his class, he finds out that 20% (4 out of 20) of the boys will vote for him, where 40% (4 out of 10) of the girls will vote for him. Then, he believes that he is going to win the class representative position because "he has 60% of the class voting for him"! Upon further investigation, the whole story is summarized in the following two-way table.

		Gender		
		Boys (A)	Girls (A^c)	
Supporting Tony	Yes (B)	**4**	**4**	**8**
	No (B^c)	**16**	**6**	**22**
		20	**10**	**30**

  In other words, Tony only has 8 out of 30 students supporting him!
- MLO: Perform probability calculations using formula, like complementary rule, addition rule, conditional probability, and multiplication rule.

6.10 Multiplication Rule

- Besides the fundamental probability equation (and addition rule), this <u>general form</u> of the **multiplication rule** (or law) is another equation involving the probability of intersection, and it is given by:

$$P(A\ and\ B) = P(A) \times P(B \mid A) \text{ or } P(A\ and\ B) = P(B) \times P(A \mid B)$$

- You would notice that both expressions above are merely rearrangement of the conditional probability formula.
- Note: The probability of intersection is often called **joint probability**.
- MLO: Perform probability calculations using formula, like complementary rule, addition rule, conditional probability, and multiplication rule.

6.11 Bayes' Theorem

- **Bayes' Theorem** is an application of the conditional probability.
- It allows us to find probability that would otherwise seem impossible with the given information or probability.
- For example, we are interested in knowing if someone indeed has COVID-19 when she is tested positive, when the only information we have is how likely a COVID-19 is (or its prevalence) among all test takers.
- Other times, we want to know if a woman is (indeed) pregnant when she gets a positive pregnancy test, when the box only says 99.9% accuracy (it actually refers to the probability of a pregnant woman getting a positive test result).
- In this course, no formula will be used but we will use the two-way table method.
- MLO: Apply the use of conditional probability and Bayes' Theorem in probability calculations.

6.12 Mutually Exclusive Events

- Two events are said to be **mutually exclusive** or **disjoint** when (1) they do not share common outcomes, or (2) they cannot happen at the same time.
- Because the two events cannot happen at the same time, the event "A and B" has nothing in it. Hence, $P(A\ and\ B) = 0$.
- As a result, the general form of the addition rule is simplified to this special form:
  $P(A\ or\ B) = P(A) + P(B)$ when A and B are mutually exclusive.
- Another application of this formula is to test if two events are disjoint algebraically, we could test if $P(A\ and\ B)$ is indeed zero.
- Here is the three-step process.
  1) Assume that the two events are mutually exclusive, or assume that $P(A\ and\ B) = 0$.
  2) Calculate the $P(A\ and\ B)$.
  3) If $P(A\ and\ B)$ is indeed zero, then the two events are mutually exclusive. Otherwise, they are not.
- MLO: Show two events are mutually exclusive using probability calculations.

6.13 Independent Events

- Two events are said to be **independent** when the outcome of one event does not affect the outcome of another event, or vice versa.
- There are three identities that you can use to show independence between the two events.
- 1) $P(A\ and\ B) = P(A) \times P(B)$,
- 2) $P(A \mid B) = P(A)$,
- 3) $P(B \mid A) = P(B)$.

- The first identity above is the general form of the multiplication rule simplified to this <u>special form</u>:
  $P(A \; and \; B) = P(A) \times P(B)$ when A and B are independent to each other.
- Another application of this formula is to test if two events are independent. Another one of the three equations or identities above could be used to do that.
- Note: <u>In probability, we cannot assume events are independent unless it is stated clearly in the question</u>.
- To show two events are independent, we could show if the above identity is true or not by following a three-step process.
  1) Assume that the two events are independent first and write down one of the three identities.
  2) Calculate the LHS (left hand side) and RHS (right hand side) of the stated identity **<u>separately</u>**.
  3) If LHS is exactly the same as RHS, then the two events are independent to each other.
- Note: Even if LHS and RHS are off by the smallest margin, the two events are still not independent to each other.
- MLO: Show two events are independent to each other using probability calculations.

6.14 Mutually Exclusive Events vs. Independent Events

- Students generally are confused with these two relations or conditions.
- These terms are like how an ESL student (English as a Second Language) students see the words "handsome" and "pretty". One might be confused with these two terms because they are both describing the appearance of people. Also, they are both on the positive side of things. At the beginning, you just memorize how and when they are used. But after a while, you would get the difference – "handsome" is for men while "pretty" is for women.
- By the same token, once we know the proper meaning or the use of them, we do not need to fuzz about their difference. All we need to know is that "mutually exclusive-ness" is a condition of addition rule (hence, about union) and "independence" is a condition of multiplication rule (hence, about intersection).