

Data Management

MDM4U

Qinghao Hu

October 9, 2025

Contents

1	Unit 1	2
1.1	Lecture 1	2
1.1.1	The Fundamental or Multiplicative Counting Principle	2
1.1.2	Additive Counting Principle	2
1.2	Lecture 1.2	2
1.3	Like Term Permutations	2
1.4	Pascal Triangle	3
1.5	Venn Diagrams	3
1.6	Combination	3
2	Unit 2, Probability	4
2.1	Probability	4
2.1.1	Definitions of Probability	4
2.1.2	Definitions of other stupid stuffs	4
2.1.3	Formulas	5
2.1.4	Examples	5
2.2	Dependent Events	5
2.2.1	Definitions	5
2.3	Mutually Exclusive Events	5
2.3.1	Some boring Definitions	5
3	One Variable Statistics	7
3.1	Variables and Data	7
3.1.1	Definitions	7

Chapter 3

One Variable Statistics

3.1 Variables and Data

3.1.1 Definitions

Categorical variables represent data that are generally grouped into categories, and are also known as qualitative variables

Ordinary variables are categorical variables whose data has a natural order but the difference between values cannot be determined or is not meaningful

Nominal variables describe names, labels, or categories that have no natural order

Quantitative variables describe data values that are numerical, and are also known as numerical variables

Continuous variables are numerical variables which can assume an infinite number of values in a given interval

Discrete variables are numerical variables that only take on a finite number of possible values in a given interval

Primary data are data that are collected by the statisticians who are analyzing the data, from first-hand sources such as surveys or experiments

Secondary data are data that the statisticians who are analyzing the data did not participate in the first hand data collection process (ie Surveys or experiments)

Microdata contains records for each individual surveyed

Aggregate or summary data are data that are combined or summarized in such a way that the individual microdata can no longer be determined.

Data gathered from a **cross** sectional study considers individuals from different groups at the same time

Data gathered from a **longitudinal** study considers how the characteristics of a specific sample changes over time

An **index** is a continuous variable such that it is an arbitrarily defined number that provides a measure of scale. It is used to relate the values of a variable to a base level

The **consumer** price index, CPI, provides a broad picture of the cost of living in Canada by comparing the cost of a wide variety of consumer goods, such as food, clothing, fuel, heating cost, transportation, shelter, and recreation

Health officials use the **body** mass index to determine whether a person is overweight. The BMI is calculated by dividing a person's mass in kilograms by the square of their height in meters