Chapter 1

Probability

1.1 Basic Concepts

The discipline of statistics deals with the collection and analysis of data. In the study of statistics, we consider **(random) experiments** for which the outcome cannot be predicted with certainty. The collection of all outcomes is called the **outcome space/sample space/space** S.

Ex See Examples 1-8 on page 3 (scanned file).

The outcomes of an experiment could be numerical or non-numerical. We often associate each outcome with a number so that we can analyze the outcomes mathematically. In an experiment, a function $X : S \to \mathbf{R}$ (the set of real numbers) that associates each outcome with a number is called a **random variable (r.v.)** (often denoted by capital letters X, Y, Z, etc). When we repeat a random experiment n times, the fractions of times an experiment ends in the respective outcomes form a **distribution of the random variable**. We often **estimate** these fractions by performing the experiments many times. The collection of the outcomes obtained from such repeated trials is called a **sample**. The making of a conjecture about the distribution of a random variable based on the sample is called a **statistical inference**.

Ex Repeated flipping a fair coin 1000 times and record the outcomes.

Suppose we repeat an experiment n times. If an outcome A has occurred $N(A)$ times in these n trials, then $N(A)$ is called the **frequency**, and $N(A)/n$ is called the relative frequency, of the outcome A, respectively. As n increases, the relative frequency tends to approach the real **probability** of the outcome.

Frequency table is used to show the frequencies of outcomes. **Histograms** can be used to present graphically the frequencies/relative frequencies/probabilities of outcomes.

Ex Example 9 on page 5 (scanned file).

A **probability mass function (p.m.f.)** is a function that gives the probability that a discrete random variable is exactly equal to some value.

The value of x (in a r.v. X) that has a greatest probability is called the **mode** of the distribution.

Ex Examples 10-11 on pages 6-8 (scanned file).

Homework
§1.1 1, 5, 7, 9

Attachment: Scanned textbook pages of Section 1-1

1.2 Properties of Probability

The collection of all possible outcomes of a random experiment is called the outcome space S. Let $A \subset S$ be a part of the collections of outcomes in S. Then A is called an **event**. When the random experiment is performed and the outcome is in A, we say that **event** A **has occurred**.

The algebra of sets/events: ($A, B \subset S$. Show by **Venn diagrams** in scanned file)

ϕ	the null/empty set
$A \subset B$	A is a **subset** of B
$A \cup B$	the **union** of A and B
$A \cap B$	the **intersection** of A and B
A'	the **complement** of A

Special terminology associated with events:

1. $A_1, A_2, \cdots, A_k$ are **mutually exclusive events** if $A_i \cap A_j = \emptyset$ for any $i \neq j$.

2. $A_1, A_2, \cdots, A_k$ are **exhaustive events** if $A_1 \cup A_2 \cup \cdots \cup A_k = S$.

3. $A_1, A_2, \cdots, A_k$ are **mutually exclusive and exhastive events** if $A_i \cap A_j = \emptyset$ for any $i \neq j$ and $A_1 \cup A_2 \cup \cdots \cup A_k = S$. We say that $A_1, A_2, \cdots, A_k$ constitute a **partition** of the sample space S.

Repeat an experiment n times. Let $N(A)$ be the number of times that event A occurred in these n trials. Then $N(A)/n$ is called the **relative frequency** of event A. When n is large enough, $N(A)/n$ approaches the **probability** $P(A)$ of event A. A function such as $P(A)$ that is evaluated for a set A is called a **set function**

Ex 1 on p13 (See scanned file).

Ex 2 on p14

Def. Probability *is a real-valued set function P that assigns, to each event A in the sample space S, a number $P(A)$, subjected to the following requirements:*

1. $P(A) \geq 0$.

2. $P(S) = 1$.
3. If $A_1, A_2, A_3, \cdots$ are events and $A_i \cap A_j = \emptyset$, $i \neq j$, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

Thm 1.1.

1. $P(\emptyset) = 0$, $P(S) = 1$
2. $P(A) \leq 1$
3. $P(A) = 1 - P(A')$
4. If $A \subset B$, then $P(A) \leq P(B)$
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
6. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$

Most of the statements can be proved by using Venn diagrams.

Ex Examples 3 on p15, 4 on p16, 5 on p17, 6 on p18.

Homework
§1.2 1, 3, 5, 7, 9, 17

Attachment: Scanned textbook pages of Section 1-2

1.3 Methods of Enumeration

Multiplication Principle: Suppose that a procedure E_1 has n_1 outcomes and, for each of these possible outcomes, a procedure E_2 has n_2 possible outcomes. Then the composite procedure E_1E_2 that consists of performing first E_1 and then E_2 has n_1n_2 possible outcomes.

Ex 1 on p20

Def. *Each of the $n!$ arrangement of n different objects is called a* **permutation** *of the n objects.*

Ex 3 on p21

If only r positions are to be filled with objects selected from n different objects, $r \leq n$, then the number of possible *ordered* arrangement is

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

Def. *Each of the $_nP_r$ arrangements is called a* **permutation of n objects taken r at a time**.

Ex 4 on p21. the probability that all students in a classroom have different birthday.

Def. *If r objects are selected from a set of n objects, and if the order of selection is noted, then the selected set of r objects is called an* **ordered sample of size r.** *There are two situations:*

1. **Sampling with replacement** *occurs when an object is selected and then replaced before the next object is selected.*
2. **Sampling without replacement** *occurs when an object is not replaced after it has been selected.*

Ex 6 & 8 on p22.

Often the order of selection is not important. The number of (unordered) subsets of size r that can be selected from n different objects is

$$\binom{n}{r} = {}_nC_r = \frac{_nP_r}{r!} = \frac{n!}{r!(n-r)!}$$

Def. *Each of the* $\binom{n}{r}$ *unordered subsets is called a* **combination of** n **objects taken** r **at a time**.

The **binomial coefficients** $\binom{n}{r}$ comes from the identity

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} b^r a^{n-r}$$

Ex 11 on p24, 12 on p25

Suppose that in a set of n objects, n_1 are similar, n_2 are similar, $\cdots$, n_s are similar, where $n_1+n_2+\cdots+n_2 = n$. Then the number of distinguishable permutations of the n objects is the **multinomial coefficient**

$$\binom{n}{n_1, n_2, \cdots, n_2} = \frac{n!}{n_1! n_2! \cdots n_s!}$$

Homework
§1.3 1, 5, 7, 11, 19

1.4 Conditional Probability

Ex 1 on p29 (scanned file)

Def. *The* **conditional probability** *of an event A, given that event B has occurred and $P(B) > 0$, is defined by*

$$P(A|B) := \frac{P(A \cap B)}{P(B)}$$

Def. *The probability that two event A and B both occur is given by the* **multiplication rule**

$$P(A \cap B) = P(A)P(B|A) \qquad or \qquad P(A \cap B) = P(B)P(A|B)$$

Ex Examples 2 & 3 (p30), 4 (p31), 6 (p33), 7 (p34), 9 (p34), 11 (p35) (scanned file)

Homework
§1.4 3, 5, 11, 17, 21

Attachment: Scanned textbook pages of Section 1-4

1.5 Independent Event

Def. *Events A and B are* **independent** *if and only if $P(A \cap B) = P(A)P(B)$. Otherwise A and B are* **dependent** *events.*

Thm 1.2. *If A and B are independent, then the following pairs of events are also independent:*

1. *A and B'*
2. *A' and B*
3. *A' and B'*

Ex 2 & 3 (p39) (scanned file)

Ex 4 (p40) (scanned file) shows that pairwise independent events A, B, C, may not be completely independent.

Def. *Events A, B, C are* **mutually independent** *if and only if*

1. $P(A \cap B) = P(A)P(B)$, $P(B \cap C) = P(B)P(C)$, $P(C \cap A) = P(C)P(A)$.
2. $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Def. *In general, events $A_1, A_2, A_3, \cdots$ are* **mutually independent** *if and only if for any subset I of the index set $\{1, 2, 3, \cdots\}$,*

$$P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$$

Ex 5 (p41), 6 (p42) (scanned file)

Homework
§1.5 1, 5, 9, 11, 13

Attachment: Scanned textbook pages of Section 1-5

1.6 Bayes's Theorem

It is often useful to use Venn diagram and tree diagram to analyze the probability involving several events.

Let $B_1, \cdots, B_m$ constitute a partition of the sample space S:

$$S = B_1 \cup B_2 \cup \cdots \cup B_m, \qquad B_i \cap B_j = \emptyset, \quad i \neq j.$$

Let A be any event. Then

$$\begin{aligned} A &= (B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_m \cap A) \\ P(A) &= \sum_{i=1}^{m} P(B_i \cap A) = \sum_{i=1}^{m} P(B_i) P(A|B_i). \end{aligned}$$

The **Bayes's theorem** states that:

$$P(B_k|A) = \frac{P(B_k \cap A)}{P(A)} = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{m} P(B_i)P(A|B_i)},$$

where $P(B_k|A)$ is called the **posterior probability** (shown by tree diagram).

Ex 1 (p45), 2 (p46), 3 (p47) (scanned file)

Homework
§1.6 1, 3, 7

Attachment: Scanned textbook pages of Section 1-6