

Outline of the Course

1. Introduction
2. Discrete Distributions
3. Continuous Distributions
4. Bivariate Distributions
5. Distributions of Functions of Random Variables
6. Point Estimation
7. Interval Estimation
8. Test of Statistical Hypotheses
9. More Tests

Chapter 1

Introduction

- Basic Concepts
- Mean, Variance, Standard Deviation
- **Axioms** and Properties of Probability
- Methods of Enumeration
- Conditional Probability
- Independent Events
- Bayes' Theorem

1.1 Basic Concepts

- Statistics
 - data \rightarrow information \rightarrow decision
 - need **probability** (mathematical background)
- *Variation is almost everywhere*
 - probabilistic models for the variation
- Probability and Statistics are extremely useful
 - highly desirable for graduate work in many areas

Example 1



Table 1.1-1 Number of children per family

2	2	5	3	4	4	3	3	6	4	3	4	4	4	4	2	5	9	2	3
1	3	5	2	4	4	4	3	3	2	2	4	2	2	6	6	1	3	3	3
3	2	3	4	7	3	3	3	2	2	2	2	3	2	3	2	3	2	5	2
3	2	2	2	4	3	3	2	3	2	4	3	3	3	4	2	4	1	2	2
2	4	3	3	3	5	2	3	3	2	2	3	3	4	2	2	2	7	2	3

\mathcal{S} : Outcome space, **Sample space**, Space

$$\mathcal{S} = \{1, 2, 3, \dots\}$$

f : frequency of the outcome

n : number of trials

f/n : **relative frequency** of the outcome

f : frequency of the outcome

n : number of trials

f/n : **relative frequency** of the outcome

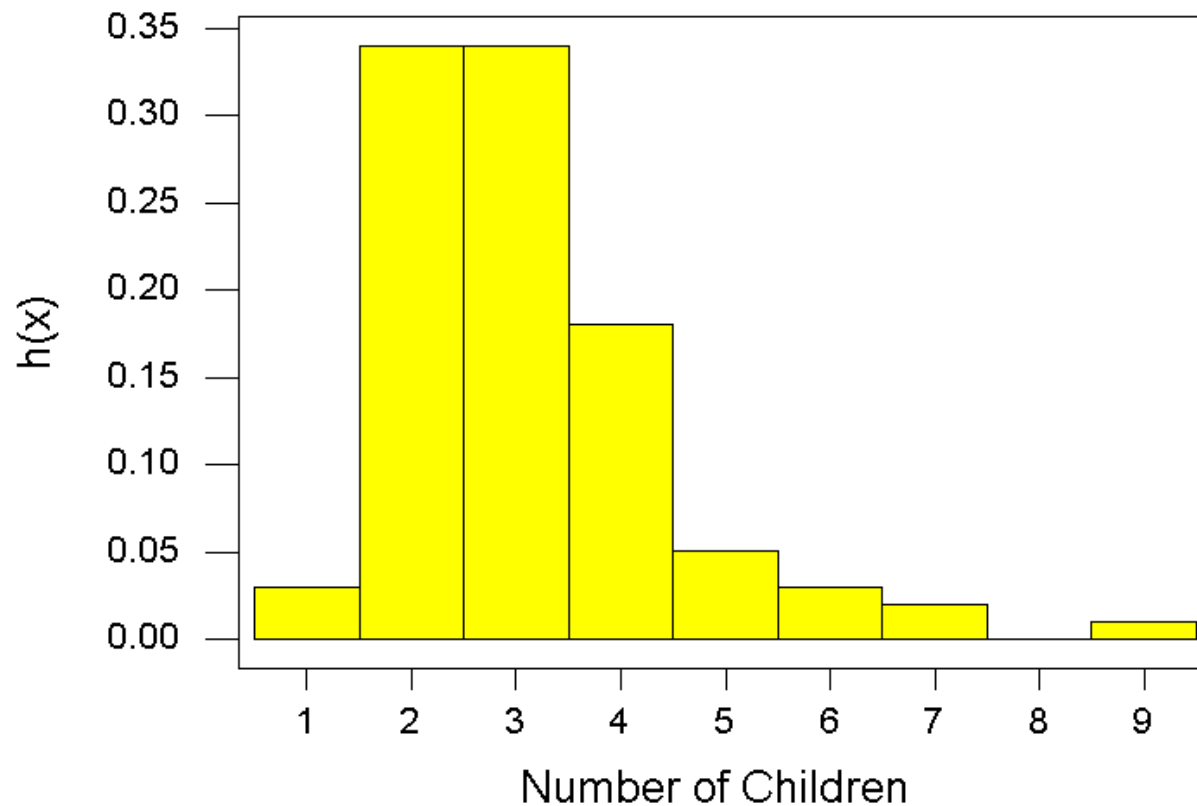
Frequency table

Table 1.1-2 Frequency table of number of children per family

Number of Children	Tabulation	Frequency	Relative Frequency
1		3	0.03
2	 	34	0.34
3	 	34	0.34
4	 	18	0.18
5	 	5	0.05
6		3	0.03
7		2	0.02
8		0	0.00
9		1	0.01
Totals		100	1.00

Histogram: *Relative Frequency Histogram*

Figure 1.1-1 Relative frequency histogram of
number of children per family



Example 2

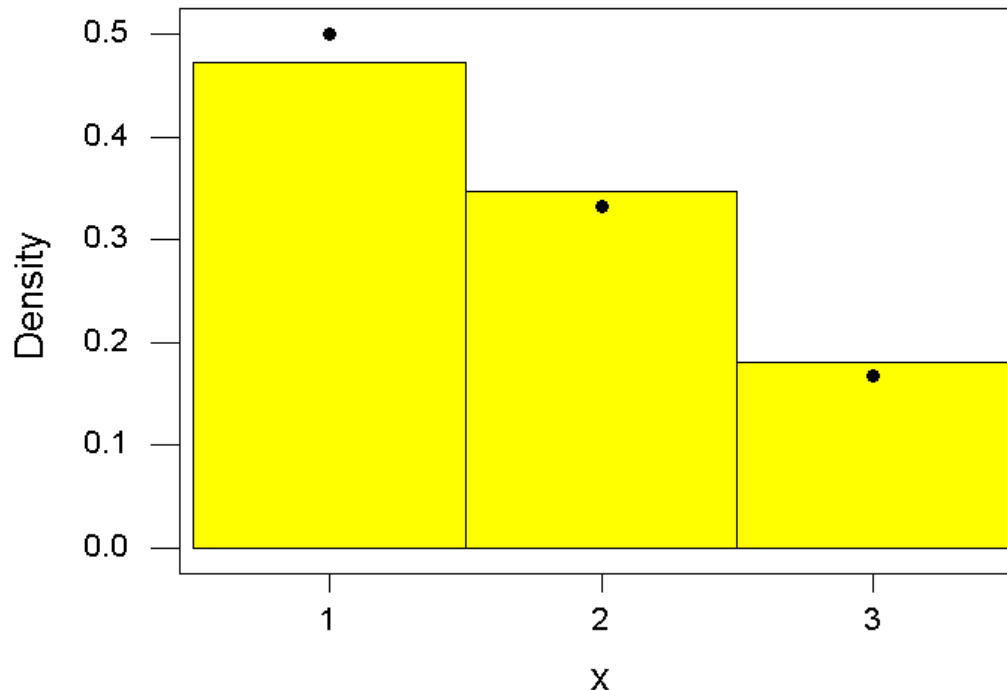
Six chips of the same size



$f(x)$: probability mass function, p.m.f.

$$S = \{1, 2, 3\} \quad \begin{cases} f(1) = 3/6, & f(2) = 2/6, & f(3) = 1/6, \\ h(1) = 284/600, & h(2) = 208/600, & h(3) = 108/600, \end{cases}$$

Figure 1.1-2 Relative Frequency Histogram (shaded) and Theoretical Probabilities (dots)



Chapter 1

Introduction

- Basic Concepts
- Mean, Variance, Standard Deviation
- Axioms and Properties of Probability
- Methods of Enumeration
- Conditional Probability
- Independent Events
- Bayes' Theorem

1.2 Mean, Variance, Standard Deviation

- **Mean** of the random variable X (or of its distribution)

平均數 $\mu = \sum_{x \in S} xf(x) = u_1f(u_1) + u_2f(u_2) + \cdots + u_kf(u_k).$

- **Variance** of the random variable X (or of its distribution)
- *second moment about the mean*

變異數 $\sigma^2 = \sum_{x \in S} (x - \mu)^2 f(x) = \sum_{x \in S} (x^2 - 2\mu x + \mu^2) f(x)$

Standard deviation $= \sum_{x \in S} x^2 f(x) - 2\mu \sum_{x \in S} xf(x) + \mu^2 \sum_{x \in S} f(x)$

標準差 $= \sum_{x \in S} x^2 f(x) - 2\mu \cdot \mu + \mu^2 \cdot 1 = \sum_{x \in S} x^2 f(x) - \mu^2.$

- Perform a certain random experiment n times obtaining n observed values of the random variable, say x_1, x_2, \dots, x_n . \rightarrow *samples*

- *Empirical distribution*: determined by the data

$$x_1, x_2, \dots, x_n.$$

- *mean of the empirical distribution*

$$\sum_{i=1}^n x_i \left(\frac{1}{n} \right) = \frac{1}{n} \sum_{i=1}^n x_i,$$

- *sample mean* 樣本平均數

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

- *Variance of the empirical distribution*

$$v = \sum_{i=1}^n (x_i - \bar{x})^2 \left(\frac{1}{n} \right) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

- *Sample Variance* 樣本變異數

$$s^2 = \left[\frac{n}{n-1} \right] v = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Chapter 1

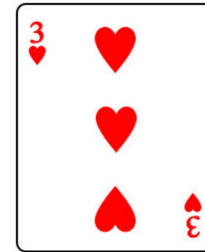
Introduction

- Basic Concepts
- Mean, Variance, Standard Deviation
- **Axioms and Properties of Probability**
- Methods of Enumeration
- Conditional Probability
- Independent Events
- Bayes' Theorem

1.3 Axioms and Properties of Probability

- Outcome space, Sample space
 - The set of all possible outcomes of an experiment
 - denoted by \mathcal{S}
- Event
 - a subset A of an outcome space \mathcal{S}
- The probability that event A occurs
 - denoted by $P(A)$.

- For example
 - randomly pick up a bridge card
 - then there are 52 possible outcomes in the outcome space of the experiment.



e.g., $X = H3$

– Followings are some events:

- {Black cards}.
- {Heart cards}.
- {Cards with number = 2 or 3}.



Definition 1.1-1 (Probability Axioms 公理)

- A **probability measure** $P(\cdot)$ is a function
 - that maps **events** in the sample space to **real numbers** such that :

Axiom 1. For any event $A \subseteq S$, $P(A) \geq 0$,

Axiom 2. $P(S) = 1$

Axiom 3. For any countable collection A_1, A_2, \dots
of mutually exclusive events ($A_i \cap A_j = \emptyset$ for all $i \neq j$),

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

(axiom of countable additivity)

Basic Theorems

- **Theorem 1.1-1** For each event A ,
 $P(A) = 1 - P(A')$, where $A' = S - A$.
- **Theorem 1.1-2** $P(\emptyset) = 0$.
- **Theorem 1.1-3** A and B are two events.
If $A \subseteq B$, then $P(A) \leq P(B)$.
- **Theorem 1.1-4** For each event A , $P(A) \leq 1$.
Therefore, $0 \leq P(A) \leq 1$
- **Theorem 1.1-5** If A and B are any two events.
 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- **Theorem 1.1-6**
 $P(A \cup B \cup C) = P(A) + P(B) + P(C) - \dots$

Basic Theorems

- **Theorem 1.1-1** For each event A ,
$$P(A) = 1 - P(A'), \quad \text{where } A' = S - A.$$

Proof [See Figure 1.1-1(a).] We have

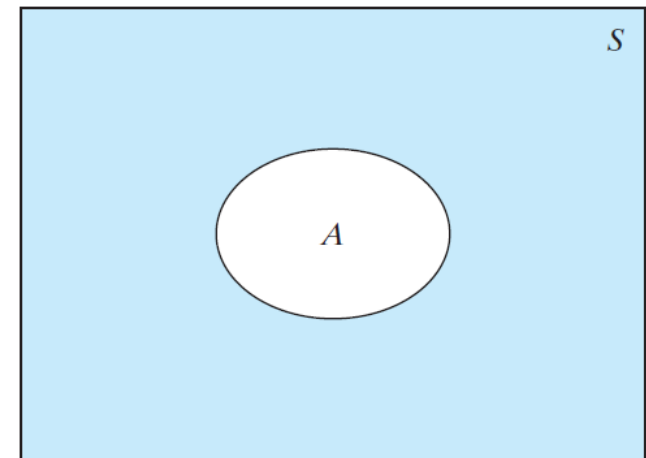
$$S = A \cup A' \quad \text{and} \quad A \cap A' = \emptyset.$$

Thus, from properties (b) and (c), it follows that

$$1 = P(A) + P(A').$$

Hence

$$P(A) = 1 - P(A').$$



(a) A'

Axiom 2. $P(S) = 1$

Axiom 3. For any countable collection A_1, A_2, \dots

of mutually exclusive events ($A_i \cap A_j = \emptyset$ for all $i \neq j$),

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

Basic Theorems

- **Theorem 1.1-2** $P(\emptyset) = 0$.

Proof In Theorem 1.1-1, take $A = \emptyset$ so that $A' = S$. Then

$$P(\emptyset) = 1 - P(S) = 1 - 1 = 0.$$

- **Theorem 1.1-1** For each event A ,
 $P(A) = 1 - P(A')$, where $A' = S - A$.

Basic Theorems

- **Theorem 1.1-1** For each event A ,
 - $P(A)=1 - P(A')$, where $A'=S - A$.
- **Theorem 1.1-2** $P(\emptyset) = 0$.
- **Theorem 1.1-3** A and B are two events.
If $A \subseteq B$, then $P(A) \leq P(B)$.
- **Theorem 1.1-4** For each event A , $P(A) \leq 1$.
Therefore, $0 \leq P(A) \leq 1$
- **Theorem 1.1-5** If A and B are any two events.
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
- **Theorem 1.1-6**
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - \dots\dots$$

Theorem 2.1-5 *If A and B are any two events, then*

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

Proof The event $A \cup B$ can be represented as a union of mutually exclusive events, namely,

$$A \cup B = A \cup (A' \cap B).$$

Hence, by property (c),

$$P(A \cup B) = P(A) + P(A' \cap B). \quad (2.1-1)$$

However,

$$B = (A \cap B) \cup (A' \cap B),$$

which is a union of mutually exclusive events. Thus

$$P(B) = P(A \cap B) + P(A' \cap B)$$

and

$$P(A' \cap B) = P(B) - P(A \cap B).$$

If this result is substituted in Equation 2.1-1 we obtain

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

which is the desired result. ■

Chapter 1

Probability

- Basic Concepts
- Mean, Variance, Standard Deviation
- Axioms and Properties of Probability
- **Methods of Enumeration**
- Conditional Probability
- Independent Events
- Bayes' Theorem

1.4 Methods of Enumeration 計數方法

排列組合

- Multiplication Principle
 - Assume that an experiment E_1 has n_1 possible outcomes
 - and for each of these possible outcomes, an experiment E_2 has n_2 possible outcomes.
 - Then, the composite experiment E_1E_2 that consists of performing E_1 first and then E_2 has n_1n_2 possible outcomes.

For example

- Let E_1 denote the selection of a rat from a cage containing one female rat and one male rat.
- Let E_2 denote the administering of either drug A, drug B, or drug **P** to the selected rat.
 - There are totally 6 outcomes as denoted in the following: (F, A), (F, B), (F, **P**), (M, A), (M, B), (M, **P**).

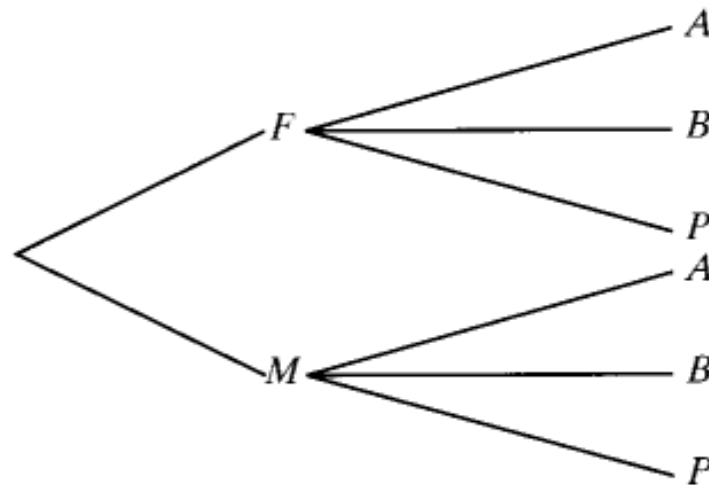


Figure 2.2-1 Tree diagram

- The **multiplication principle** can be extended to a sequence of more than two experiments.
 - Assume that the experiment E_i has n_i , $i = 1, 2, \dots, m$, possible outcomes after previous experiments have been performed.
 - The **composite experiment** $E_1E_2\dots E_m$
 - has **$n_1n_2\dots n_m$** possible outcomes.

Summary of Sampling

Taking k samples out of n samples

	With replacement	Without replacement
Ordered samples of size k	n^k	$\frac{n!}{(n-k)!}$
Unordered samples of size k	$\frac{(n+k-1)!}{k!(n-1)!}$	$\frac{n!}{k!(n-k)!}$

DEF 1: Permutation of n Objects

- A permutation of n objects
 - Assume that n positions are to be filled with n different objects.
 - There are n choices for filling the first position, $n-1$ for the second, ..., 1 choice for the last position.
 - So by the multiplication principle, there are $n(n-1)\dots(2)(1) = n!$ possible arrangements.
 - The symbol $n!$ is read n factorial.

Example 1.2-3

Example 1.2-3

- The number of **permutations** of the four letters a , b , c , and d is clearly $4!=24$.
 - *Ordered sampling without replacement*
- The number of **possible four-letter code words** using the four letters a , b , c , and d if letters may be repeated is $4^4=256$,
 - because in this case each selection can be performed in four ways.
 - *Ordered sampling with replacement*

DEF 2: permutation of n objects taken k at a time

- Each of the $n!$ arrangements of the n different objects is called a permutation of the n objects.
- If only k positions are to be filled with objects selected from n different objects, where $k \leq n$, then the number of possible arrangements is

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

- Each of the arrangements is called a permutation of n objects taken k at a time and is denoted by P_k^n or ${}_nP_k$.

Example 1.2-4, 1.2-5

Example 1.2-4

- The number of possible four-letter code words, selecting from the 26 letters in the alphabet, in which all four letters are different is

$${}_{26}P_4 = (26)(25)(24)(23) = \frac{26!}{22!} = 358,800$$

Example 1.2-5

- The number of ways of selecting a president, a vice president, a secretary, and a treasurer in a club consisting of 10 persons is

$${}_{10}P_4 = 10 \cdot 9 \cdot 8 \cdot 7 = \frac{10!}{6!} = 5040$$

Ordered Sampling

DEF 3: ordered sample of size k

- Assume that k objects are selected from a set of n different objects and the order of selection is noted.
- Then, the selected set of k objects is called an ordered sample of size k .
- Can be “with” or “without” replacement

DEF 4: Sampling **with Replacement**

- Sampling with replacement occurs when an object is selected and then is replaced before the next selection is made.
- By the multiplication principle, when ordered sampling with replacement is conducted, the number of possible ordered samples of size k taken from a set of n objects is n^k .

Example 1.2-6, 1.2-7

Example 1.2-6

- A die is rolled seven times. The number of possible ordered samples is $6^7=279,936$. Note that rolling a die is equivalent to sampling with replacement from the set $\{1,2,3,4,5,6\}$.

Example 1.2-7

- An urn contains 10 balls numbered 0, 1, 2, ..., 9. If four balls are selected, one at a time and with replacement, the number of possible ordered samples is $10^4 = 10,000$. Note that this is the number of four-digit integers between 0000 and 9999, inclusive.

DEF 5: Sampling **without Replacement**

- Sampling with replacement occurs when an object is selected and is not replaced before the next selection is made.
- By the multiplication principle, when ordered sampling without replacement is conducted, the number of possible ordered samples of size k taken from a set of n objects is P_k^n .
- \sim permutation of n objects taken k at a time

Example 1.2-8

Example 1.2-8

- The number of ordered samples of 5 cards that can be drawn without replacement from a standard deck of 52 playing cards is

$$(52)(51)(50)(49)(48) = \frac{52!}{47!} = 311,875,200$$



Unordered Sampling (without replacement) or **Combination**

- Sometimes, *the order of selection is not important.*
- We call a unordered subset of a set of n objects *a combination.*
- An example of unordered sampling
 - randomly selecting 10 students out of 50 students in a class.

DEF 6. Combinations of n objects taking k at a time

The total number of **combinations** resulting from selecting k objects from a set of n objects is

$$\frac{n!}{k!(n-k)!} \text{ and is denoted by } C_k^n \text{ or } \binom{n}{k}$$

- The formula of C_k^n can be derived from P_k^n based on observing that the same subset of objects appears in $k!$ permutations of n objects taken k at a time.

Example 1.2-9, 1.2-10

Example 1.2-9

- The number of possible five-card hands (hands in five-card poker) drawn from a deck of 52 playing cards is

$${}_{52}C_5 = \binom{52}{5} = \frac{52!}{5!47!} = 2,598,960$$

Example 1.2-10

- The number of possible 13-card hands (hands in bridge) that can be selected from a deck of 52 playing cards is

$${}_{52}C_{13} = \binom{52}{13} = \frac{52!}{13!39!} = 635,013,559,600$$

Binomial Coefficients

-- arise in the expansion of an binomial 二項式

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} b^k a^{n-k} \quad \text{Eq 1.2-1, p23}$$

$$(a + b)^n = (a + b) (a + b) \dots (a + b)$$

- Example 1.2-11

Example 1.2-11

Assume that each of the $\binom{52}{5} = 2,598,960$ five - card hands drawn from a deck of 52 playing cards has the same probability for being selected. The number of possible five - card hands that are all spades (event A) is

$$N(A) = \binom{13}{5} \binom{39}{0}$$

because the 5 spades can be selected from the 13 spades in $\binom{13}{5}$ ways after

which zero nonspades can be selected in $\binom{39}{0} = 1$ way. We have

$$\binom{13}{5} = \frac{13!}{5!8!} = 1287$$

Thus the probability of an all - spade five - card hand is

$$P(A) = \frac{N(A)}{N(S)} = \frac{1287}{2,598,960} = 0.000495$$

Example 1.2-11(cont.)

Suppose now that the event B is the set of outcomes in which exactly three cards are kings and two cards are queens. We can select the three kings in any one of $\binom{4}{3}$ ways and two queens in any one of $\binom{4}{2}$ ways. By the multiplication principle, the number of outcomes in B is

$$N(B) = \binom{4}{3} \binom{4}{2} \binom{44}{0},$$

where $\binom{44}{0}$ gives the number of ways in which 0 cards are selected out of the non-kings and non-queens and of course is equal to one. Thus

$$P(B) = \frac{N(B)}{N(S)} = \frac{\binom{4}{3} \binom{4}{2} \binom{44}{0}}{\binom{52}{5}} = \frac{24}{2,598,960} = 0.0000092.$$

Example 1.2-11(cont.)

Finally, let C be the set of outcomes in which there are exactly two kings, two queens, and one jack. Then

$$P(C) = \frac{N(C)}{N(S)} = \frac{\binom{4}{2} \binom{4}{2} \binom{4}{1} \binom{40}{0}}{\binom{52}{5}} = \frac{144}{2,598,960} = 0.000055$$

because the numerator of this fraction is the number of outcomes in C .

DEF7: Distinguishable Permutation

Example 1.2-12

Example 1.2-13

- Distinguishable Permutation \sim Combination
 - A **combination** of n objects taken k at a time can be regarded as a **binary partition** of a set containing n objects into two subsets containing k objects and $n - k$ objects, respectively.
- If a set of n objects is partitioned into subsets of sizes k_1, k_2, \dots, k_m , then the number of distinguishable permutations is

$$\binom{n}{k_1, k_2, \dots, k_m} = \frac{n!}{k_1! k_2! \cdots k_m!}, \text{ where } n = k_1 + k_2 + \cdots + k_m.$$

Example 1.2-12

- A coin is flipped 10 times and the sequence of heads and tails is observed. The number of possible 10-tuplets that result in four heads and six tails is
$$\binom{10}{4} = \frac{10!}{4!6!} = \frac{10!}{6!4!} = \binom{10}{6} = 210$$

Example 1.2-13

- Students on a boat send signals back to shore by arranging seven colored flags on a vertical flagpole. If they have four orange and three blue flags, they can send
$$\binom{7}{4} = \frac{7!}{4!3!} = 35$$

different signals. Note that if they had seven flags of different colors, they could send $7!=5040$ different signals.

Example 1.2-14

- If the students on the boat have three red flags, four yellow flags, and two blue flags to arrange on a vertical pole, the number of possible signals is

$$\binom{9}{3,4,2} = \frac{9!}{3!4!2!} = 1260.$$

Unordered Sampling with Replacement

- Assume that we **take k samples out of n objects** with replacement and **ignore the order** of samples.
- To figure out the number of possible outcomes, we can regard this problem as inserting $(n-1)$ bars into a list of k objects as follows: (e.g., $n=6$, $k=10$)
 - oo | ooo | | o | o | ooo

- Each distinguishable permutation of the string corresponds to an unordered sample.
- Therefore, the number of possible outcomes is

$$C_k^{(n+k-1)} = \frac{(n+k-1)!}{k!(n-1)!}$$

Summary of Sampling

Taking k samples out of n samples

	With replacement	Without replacement
Ordered samples of size k	n^k	$\frac{n!}{(n-k)!}$
Unordered samples of size k	$\frac{(n+k-1)!}{k!(n-1)!}$	$\frac{n!}{k!(n-k)!}$

Chapter 1

Probability

- Basic Concepts
- Mean, Variance, Standard Deviation
- Axioms and Properties of Probability
- Methods of Enumeration
- **Conditional Probability**
- Independent Events
- Bayes' Theorem

1.5 Conditional Probability 條件機率

- **Definition 1.4-1**

The conditional probability of an event A given that event B has occurred is **defined** by

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

provided that $P(B) > 0$.

Let A be the event $X_1 \geq 2$.

Let B denote the event $X_2 > X_1$.

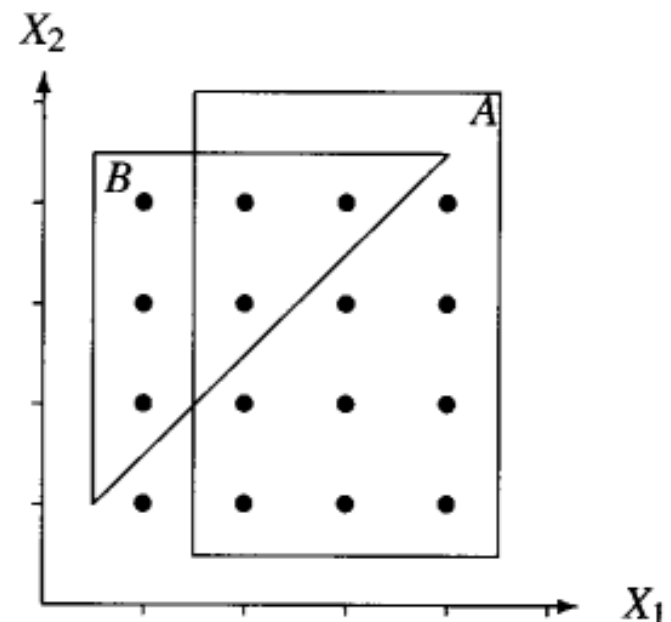


Example 1.18. Roll two four-sided dice. Let X_1 and X_2 denote the number of dots that appear on die 1 and die 2, respectively. Draw the 4 by 4 sample space. Let A be the event $X_1 \geq 2$. What is $P[A]$? Let B denote the event $X_2 > X_1$. What is $P[B]$? What is $P[A|B]$?

.....
Each outcome is a pair (X_1, X_2) . To find $P[A]$, we add up the probabilities of the sample points in A .

From the sample space, we see that A has 12 points, each with probability $1/16$, so $P[A] = 12/16 = 3/4$. To find $P[B]$, we observe that B has 6 points and $P[B] = 6/16 = 3/8$. The compound event AB has exactly three points, $(2,3), (2,4), (3,4)$, so $P[AB] = 3/16$. From the definition of conditional probability, we write

$$P[A|B] = \frac{P[AB]}{P[B]} = 1/2$$



Another example of Conditional Probability

- Assume that we roll a dice three times and want to figure out the probability that **the number coming out from the first roll is 2**, given that **the summation of the three rolls is 5**.
- There are 6 outcomes of which the summation is 5: (1,1,3), (1,3,1), (3,1,1), (1,2,2), (2,1,2), (2,2,1).
- Among the six possible outcomes, two of them have the first number equal to 2.
- Therefore, the probability of concern is **$2/6$**

Properties of Conditional Probability

- Conditional probability **satisfies the axioms for a probability function.**

That is, with $P(B) > 0$,

(a) $P(A|B) \geq 0$

(b) $P(B|B) = 1$

(c) If A_1, A_2, \dots are mutually exclusive events,
then $P(A_1 \cup A_2 \cup \dots | B) = P(A_1|B) + P(A_2|B) + \dots$

- Proof of (a)

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

Since $P(B) > 0$ and $P(A \cap B) \geq 0$, $P(A|B) \geq 0$.

- Proof of (b)

$$P(B | B) = \frac{P(B \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

- Proof of (c)

$$\begin{aligned} & P(A_1 \cup A_2 \cup \dots | B) \\ &= \frac{P((A_1 \cup A_2 \cup \dots) \cap B)}{P(B)} \\ &= \frac{P((A_1 \cap B) \cup (A_2 \cap B) \cup \dots)}{P(B)} \\ &= \frac{P(A_1 \cap B) + P(A_2 \cap B) + \dots}{P(B)} \\ &= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} + \dots \\ &= P(A_1 | B) + P(A_2 | B) + \dots \end{aligned}$$

Multiplication Rule of Two Events

- **Definition 1.3-2** (from Multiplication Rule of Two Events)
The probability that two events, A and B, both occur is given by

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

provided that $P(A) > 0$ and $P(B) > 0$

Example 1.3-6

- A bowl contains seven blue chips and three red chips. Two chips are to be drawn successively at random and without replacement.
- We want to compute the probability that the first draw results in a red chip (A) and the second draw results in a blue chip (B).
- It is reasonable to assign the following probabilities:

$$P(A) = \frac{3}{10} \text{ and } P(B|A) = \frac{7}{9}$$

- The probability of red on the first draw and blue on the second draw is

$$P(A \cap B) = \frac{3}{10} \cdot \frac{7}{9} = \frac{7}{30}$$

- It should be noted that in many instances, it is possible to compute a probability by two seemingly different methods. For illustration, consider Example 1.3-6 but find the probability of drawing a red chip on each of the two draws.

Following that example, it is $\frac{3}{10} \cdot \frac{2}{9} = \frac{1}{15}$

- However, we can also find this probability using combinations as follows:

$$\frac{\binom{3}{2}\binom{7}{0}}{\binom{10}{2}} = \frac{\frac{3 \cdot 2}{1 \cdot 2}}{\frac{10 \cdot 9}{1 \cdot 2}} = \frac{1}{15}$$

- Thus we obtain the **same answer**, as we should, provided that our reasoning is consistent with the underlying assumptions.

Example 1.3-9

- Four cards are to be dealt successively at random and without replacement from an ordinary deck of playing cards. The probability of receiving in order a spade, a heart, a diamond, and a club is

$$\frac{13}{52} \cdot \frac{13}{51} \cdot \frac{13}{50} \cdot \frac{13}{49},$$

a result that follows from the extension of the multiplication rule and reasonable assignments to the probabilities involved.