ISyE 3770, Spring 2024
Statistics and Applications

Introduction

Instructor: Jie Wang
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

jwang3163@gatech.edu
Office: ISyE Main 445

Course Outline

- **Course Pages**: https://gatech.instructure.com/courses/370378
- **Office Hours**: **Wednesday 03:30 PM - 04:45 PM.**

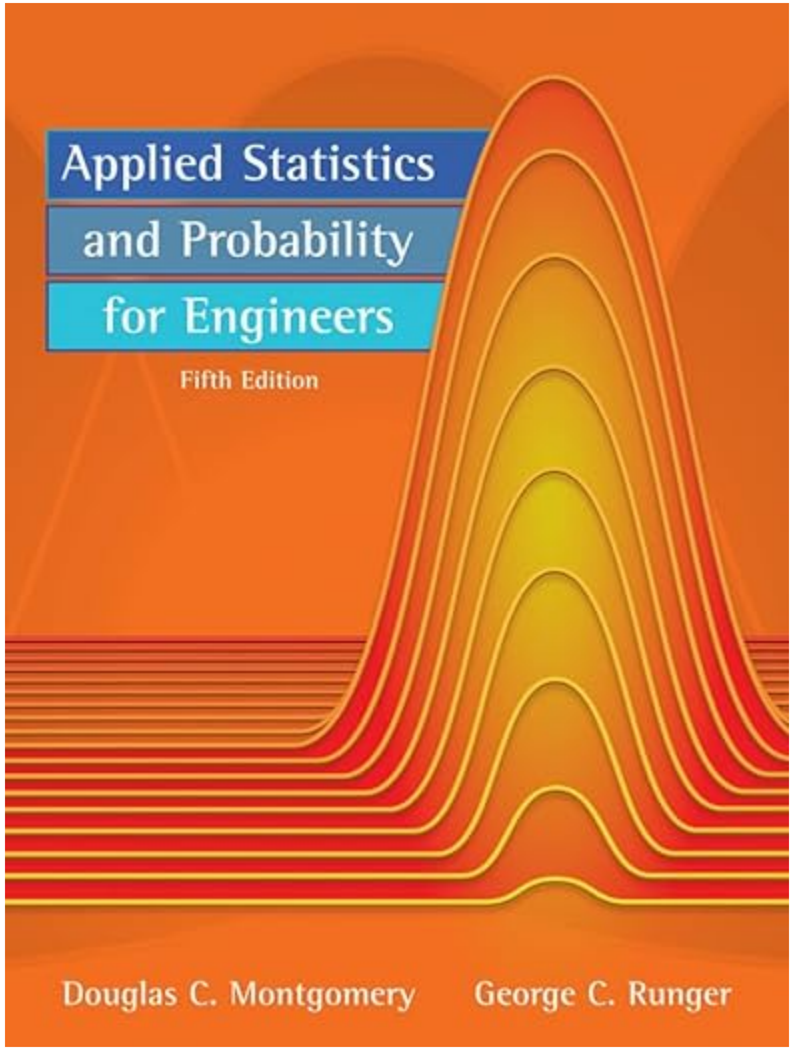
  Friday 02:00 PM - 04:45 PM.

  ☐ In person location: ISyE Main 445

  ☐ Zoom: https://gatech.zoom.us/my/jwang3163

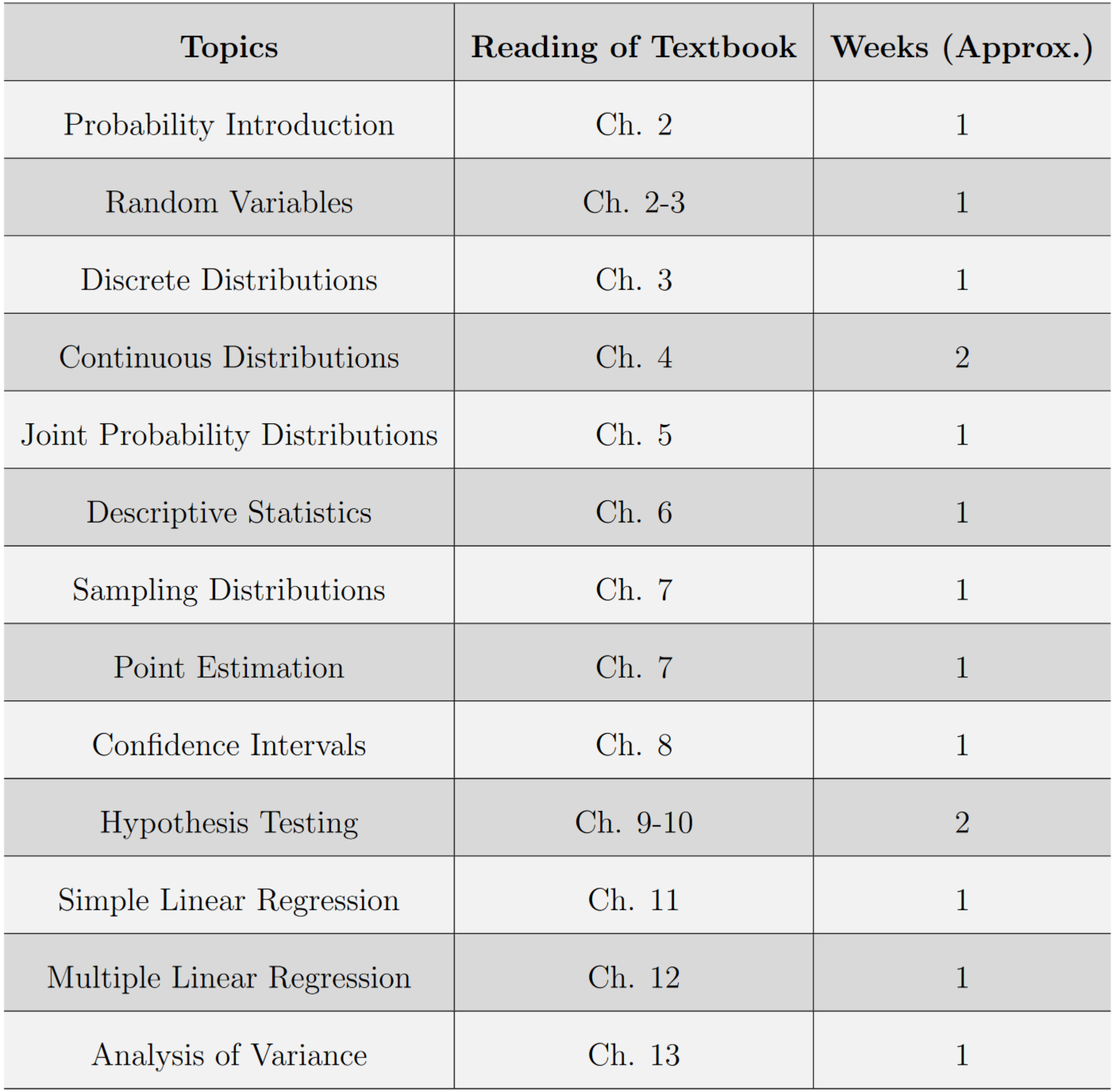
Course Outline

- **Textbook**: Applied Statistics and Probability for Engineers, 7th Edition



- **Course Description**: Introduction to probability, probability distributions, point estimation, confidence intervals, hypothesis testing, linear regression, and analysis of variance.
- **Prerequisites**: An undergraduate-level understanding of multivariate calculus.

Course Outline



Topics	**Reading of Textbook**	**Weeks (Approx.)**
Probability Introduction	Ch. 2	1
Random Variables	Ch. 2-3	1
Discrete Distributions	Ch. 3	1
Continuous Distributions	Ch. 4	2
Joint Probability Distributions	Ch. 5	1
Descriptive Statistics	Ch. 6	1
Sampling Distributions	Ch. 7	1
Point Estimation	Ch. 7	1
Confidence Intervals	Ch. 8	1
Hypothesis Testing	Ch. 9-10	2
Simple Linear Regression	Ch. 11	1
Multiple Linear Regression	Ch. 12	1
Analysis of Variance	Ch. 13	1

Grading Policy

- **Homework**: **30%**
- **Midterm 1**: 17.5%
- **Midterm 2**: 17.5%
- **Final**: 35%

Remark:

a) Exams will be closed book, but it is allowed to bring a one-page cheating sheet.

b) Any request regarding exams must be made within one or two weeks of getting the exams back. There will be no make-up exams for any reason. If you have an acceptable reason (e.g., illness with doctor statement of your inability to take the exam) for missing an exam, the weight associated with the exam will be transferred to the Final Exam.

Homework Policy

- **Homework**: **30%**
- **Midterm 1**: 17.5%
- **Midterm 2**: 17.5%
- **Final**: 35%

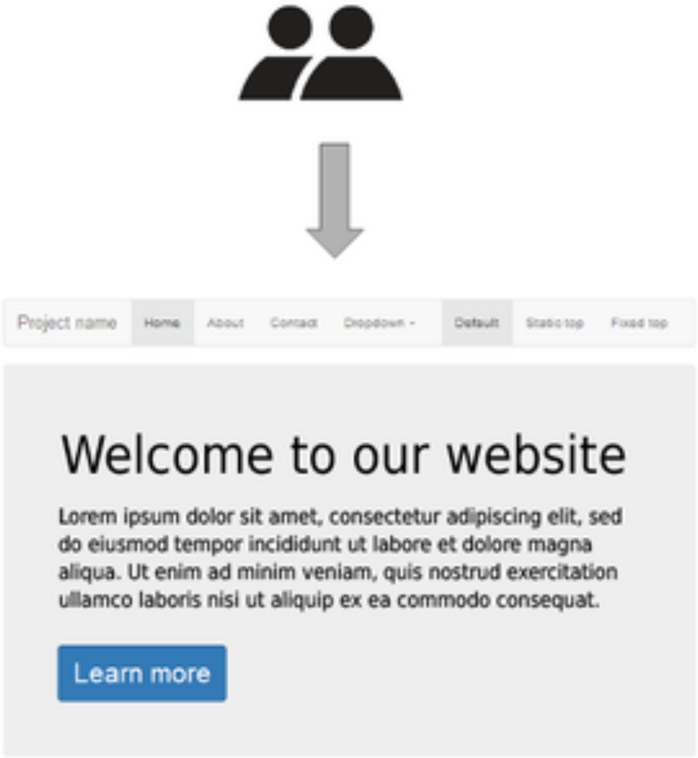
- **Late homework submission within 24 hours of the deadline incurs a 25% grade deduction, while late submission between 24 and 48 hours incurs a 50% deduction. Any work turned in more than two days past the deadline will not earn credit.**

Remark:

- Homework assignments will be posted approximately once for every 2 week. The due date is from Tuesday to the next Tuesday. Student collaboration is authorized and encouraged, but submitted homework must be worked out and written up on your own.
- Homework should be submitted electronically on Canvas as a single pdf file. Late homework should also be submitted electronically on Canvas. If Canvas happens not to accept late submissions, please email your homework solution directly to me (jwang3163@gatech.edu).

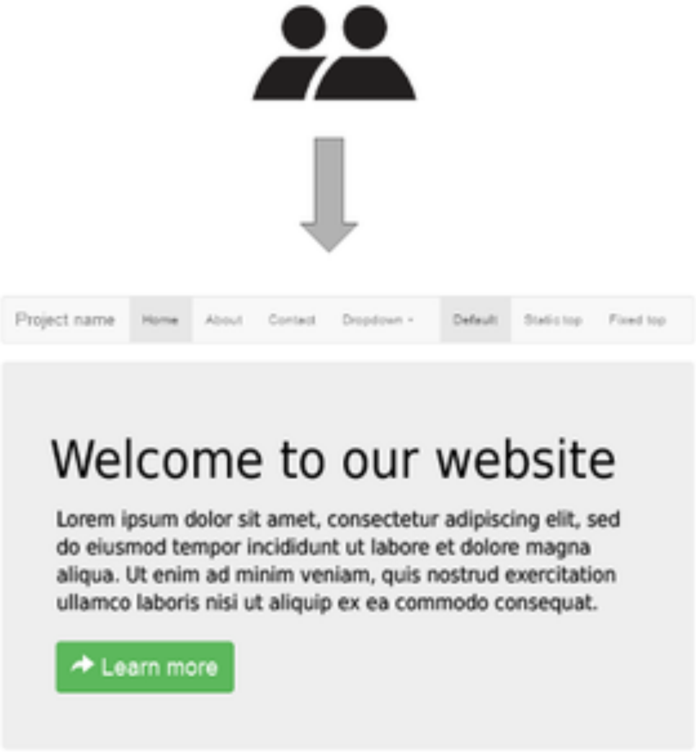
Motivation: *Statistics is the science of data*

Example: Hypothesis Testing



Welcome to our website

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.



Welcome to our website

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Click rate: 52 % 72 %

Motivation: *Statistics is the science of data*

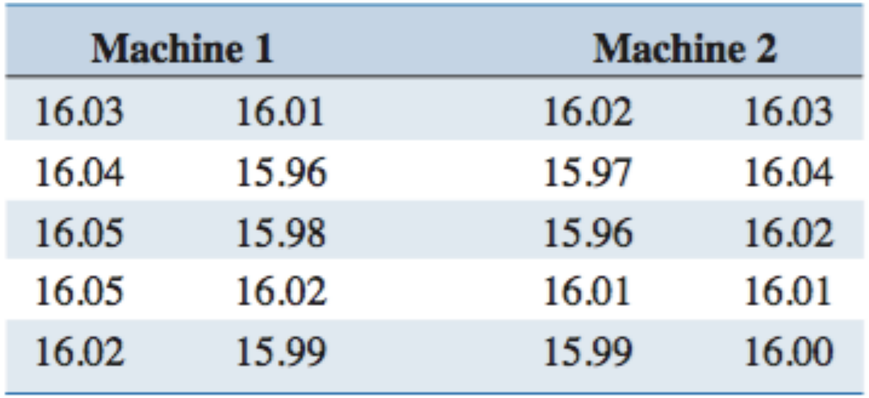
Example: Hypothesis Testing

The Coca-Cola Co. is releasing its new Georgia Peach and California Raspberry flavors.

THE COCA-COLA CO.

Motivation: *Statistics is the science of data*

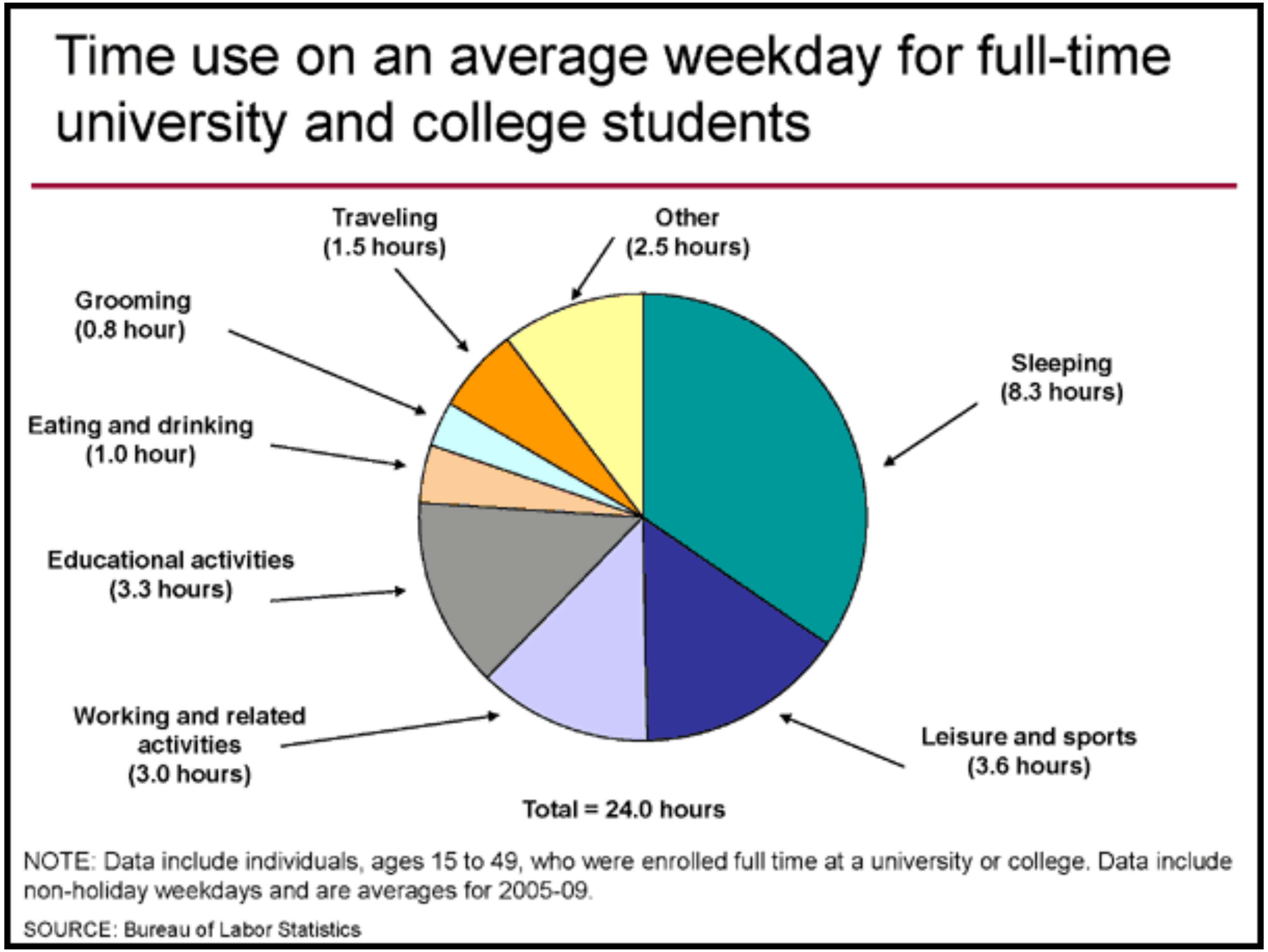
Example: Hypothesis Testing



Machine 1		Machine 2	
16.03	16.01	16.02	16.03
16.04	15.96	15.97	16.04
16.05	15.98	15.96	16.02
16.05	16.02	16.01	16.01
16.02	15.99	15.99	16.00

Motivation: *Statistics is the science of data*

Example: Descriptive Statistics

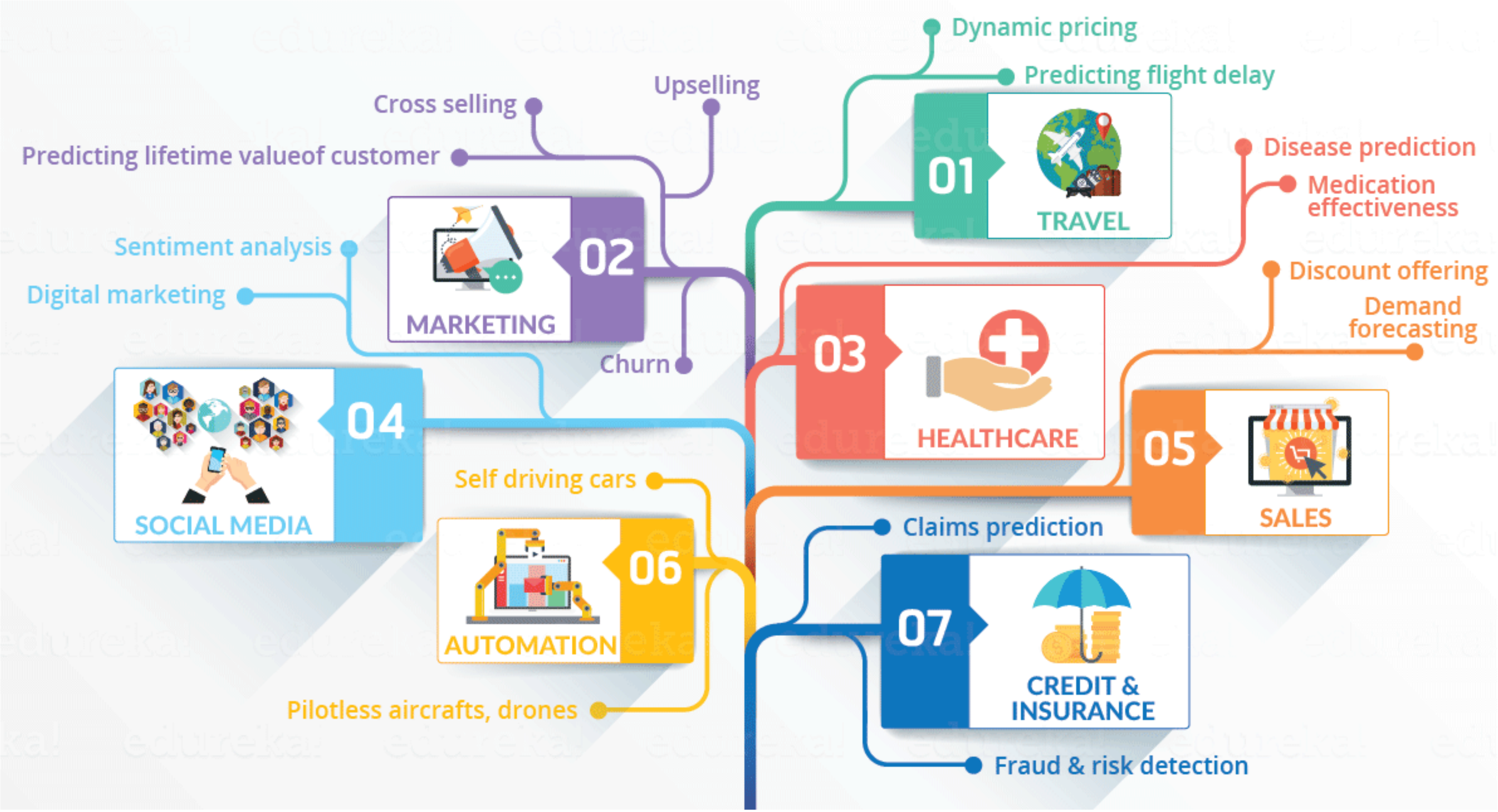


Time use on an average weekday for full-time university and college students

NOTE: Data include individuals, ages 15 to 49, who were enrolled full time at a university or college. Data include non-holiday weekdays and are averages for 2005-09.

SOURCE: Bureau of Labor Statistics

Motivation: *Statistics is the science of data*



Course Objectives

- ✓ Collect, summarize, and present data graphically
- ✓ Familiar with basic probability concepts
- ✓ Use statistical tests and confidence intervals in statistical decisions
- ✓ Select proper statistical techniques for practical applications
- ✓ Use statistical software to conduct data analysis and interpret output
- ✓ Draw statistical conclusions from data.

1.1.1 Fundamental concepts of Probability

Definition **1(Experiment)**: Any procedure that can be **infinitely** repeated and has a well-defined set of possible outcomes

Definition **2(Random Experiment)**: An experiment is said to be random if it has more than one possible outcome. In other words, its outcome cannot be **predicted with certainty**.

Definition **3(Sample Space or Outcome Space)**: The set, S, the collection of all possible outcomes of a particular experiment is called the **sample space** or **outcome space** for the experiment.

Definition **4(Event):**An event is **any collection** of possible outcomes of an experiment, that is, any **subset** of S (including S itself).

Let A be an event, a subset of S. We say the event A **occurs** if the outcome of the experiment is in the set A. When speaking of probabilities, we generally speak of the probability of an event, rather than a set. But we use the terms **interchangeably**.

Example 1:

Throwing a fair or perfectly manufactured 6-side die

1. This is a random experiment.
2. Sample space: $S = \{1, 2, 3, 4, 5, 6\}$.
3. Consider one event: $A = \{1, 2\}$.
4. Throw the die, if the outcome is either 1 or 2, the event A has occurred

Some terminology

- $\emptyset$ denotes the **null** or **empty** set;
- $A \subset B$ means A is a **subset** of B;
- $A \cup B$ is the **union** of A and B;
- $A \cap B$ is the **intersection** of A and B;
- A' is the **complement** of A (i.e., all elements in the entire set S that are not in A).

1.1.2 **Algebra of sets** (set theory --- a fundamental role)

Some terminology

- Set: a collection of distinct objects
- $\emptyset$: the null set or empty set

In the following, let A and B be two sets.

- $A \subset B$ means A is a **subset** of B; If $x \in A$, then $x \in B$
- $A \cup B$ is the **union** of A and B; If $x \in A \cup B$, then $x \in A$ or $x \in B$
- $A \cap B$ is the **intersection** of A and B; If $x \in A \cap B$, then $x \in A$ and $x \in B$
- A' is the **complement** of A (i.e., all elements in the entire set S that are not in A). If $x \in A'$, then $x \notin A$

Some terminology

$A_1, A_2, \ldots, A_k$ are

- **mutually exclusive events**: $A_i \cap A_j = \emptyset$, $i \neq j$; that is $A_1, \ldots, A_k$ are disjoint sets;
- **exhaustive events**: $\cup_{i=1}^{k} A_i = A_1 \cup A_2 \cup \cdots \cup A_k = S$.
- **mutually exclusive and exhaustive events**: $A_i \cap A_j = \emptyset$, $i \neq j$ and $\cup_{i=1}^{k} A_i = S$.

For any three events, A, B, and C, defined on a sample space S,

a.Commutativity $A \bigcup B = B \bigcup A,$

$A \bigcap B = B \bigcap A;$

b.Associativity $A \bigcup (B \bigcup C) = (A \bigcup B) \bigcup C,$

$A \bigcap (B \bigcap C) = (A \bigcap B) \bigcap C;$

c.Distributive Laws $A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C),$

$A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C);$

d.DeMorgan's Laws $(A \bigcup B)' = A' \bigcap B',$

$(A \bigcap B)' = A' \bigcup B'.$

Attention：You might be familiar with the use of Venn diagrams to “prove” these theorems in set theory. We caution that although Venn diagrams are sometimes helpful in visualizing a situation, they do not constitute a formal proof!

The proof of much of this theorem is left as your exercise. To illustrate the technique, however, we will prove the Distributive Law later:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Distributive Laws

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Prove: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(i)Sufficiency : Suppose $x \in (A \bigcap (B \bigcup C))$. *By the definition of inersetion,it must be that* $x \in (B \bigcup C)$, *that is, either* $x \in B$ *or* $x \in C$. *Since* x *also must be in* A,*we have that either* $x \in (A \bigcap B)$ *or* $x \in (A \bigcap C)$; *therefore,* $x \in ((A \bigcap B) \bigcup (A \bigcap C))$.

That is, $A \bigcap (B \bigcup C) \subset (A \bigcap B) \bigcup (A \bigcap C)$.

(ii)Now assume $x \in ((A \bigcap B) \bigcup (A \bigcap C))$. *This implies that* $x \in (A \bigcap B)$ *or* $x \in (A \bigcap C)$.

If $x \in (A \bigcap B)$, *then* x *is both in* A *and* B. *Since* $x \in B$, $x \in (B \bigcup C)$ *and thus* $x \in (A \bigcap (B \bigcup C))$.

If, on the other hand,$x \in (A \bigcap C)$,*the argument is similar, and we again conclude that* $x \in (A \bigcap (B \bigcup C))$.

Thus, we have established $(A \bigcap B) \bigcup (A \bigcap C) \subset A \bigcap (B \bigcup C)$, *showing containment in the other direction.*

Combining (i) and (ii), we can prove the Distributive Law.

1.1.3 Probability and its properties

- **Goal**: To define the probability of event A,(the chance of A occurring)

- An intuitive idea:

  - Step 1: Repeat the experiment a number of times, say n times

  - Step 2: Count the number of times that the event A actually occurs.

- Relative frequency of event A in n repetition of the experiment : $\frac{\#(A)}{n}$

Example 1(c.n.t.):

Throwing a fair or perfectly manufactured 6-side die

- $S = \{1, 2, 3, 4, 5, 6\}$ and $A = \{1, 2\}$
- Outcome is either 1 or 2 means A has occurred
- $\frac{\#(A)}{n} \to \frac{1}{3}$ as $n \to \infty$.
- Define $p := \lim_{n\to\infty} \frac{\#(A)}{n} \to \frac{1}{3}$, called the probability of event A. It is denoted by $P(A)$

Definition: Probability

Probability is a real-valued set function P that assigns, to each event A in the sample space S, a number $P(A)$, called the probability of the event A, such that the following properties are satisfied:

(a) $P(A) \geq 0$;

(b) $P(S) = 1$;

(c) If $A_1, A_2, \ldots$ are events and $A_i \cap A_j = \emptyset$, $i \neq j$, then

$$P(\cup_{i=1}^{k} A_i) = P(A_1 \cup \cdots \cup A_k) = P(A_1) + \cdots + P(A_k) = \sum_{i=1}^{k} P(A_i)$$

for each positive integer k, and

$$P(\cup_{i=1}^{\infty} A_i) = P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots = \sum_{i=1}^{\infty} P(A_i)$$

for an infinite, but countable, number of events.

$Property 1: For \quad each \quad event \quad A,$

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

$Proof: S = A \cup A', \qquad A \cap A' = \phi$

$$\Rightarrow P(S) = P(A \cup A') = P(A) + P(A') = 1 \Rightarrow P(A) = 1 - P(A').$$

$Property 2: P(\phi) = 0.$

$Proof: P(S) = 1 \overset{by \quad Property \quad 1}{\Rightarrow} P(\phi) = 1 - P(S) = 0.$

$Property 3: For \quad each \quad event \quad A \quad and \quad B,$

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

$Proof: Note \quad that \quad for \quad any \quad events \quad A \quad and \quad B \quad we \quad have$

$$B = \{B \cap A\} \cup \{B \cap A'\},$$

$and \quad therefore$

$$P(B) = P(\{B \cap A\} \cup \{B \cap A'\}) = P(B \cap A) + P(B \cap A')$$

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

Property4: For each event A and B,

$$P(A\bigcup B)=P(A)+P(B)-P(A\bigcap B)$$

Proof: To establish Property4, we use the identity

$$A\bigcup B=A\bigcup\{B\bigcap A'\},\qquad\qquad (\Psi)$$

A Venn diagram will show why (Ψ) holds, although a formal proof is not difficult.

Using (Ψ) and the fact that A and $B\bigcap A'$ are disjoint (Since A and A' are), we have

$$P(A\bigcup B)=P(A\bigcup\{B\bigcap A'\})=P(A)+P(B\bigcap A')\overset{by\ \ Property3}{=}P(A)+P(B)-P(A\bigcap B)$$

Property5: For each event A and B, if $A\subset B$, then

$$P(A)\le P(B)$$

Proof: if $A\subset B$, then $A\bigcap B=A$. Therefore, Using property 3 we have

$0\le P(B\bigcap A')=P(B)-P(A)$, establishing Property 5.

Property6: For each event $A, P(A)\le 1$

Proof: $P(S)=1=P(A\bigcup A')=P(A)+P(A')\ge P(A)$.

1.2 Method of Enumeration

- Let the sample space S contains m possible outcomes $e_1, e_2, \dots, e_m$. In other words, $S = \{e_1, e_2, \dots, e_m\}$.
- Moreover, those m outcomes are equally likely, i.e.,

$$P(\{e_i\}) = \frac{1}{m}, \quad i = 1, \dots, m.$$

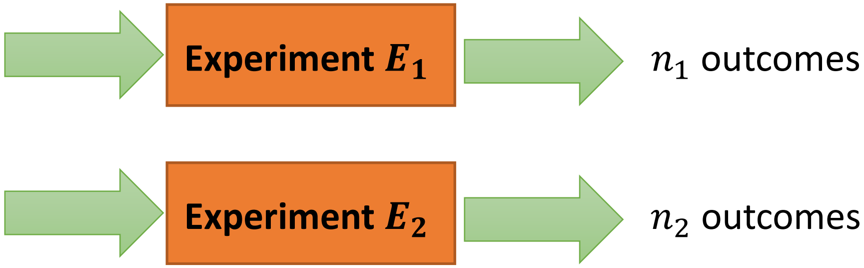
- Then the probability of an event A is

$$P(A) = \frac{N(A)}{N(S)},$$

where $N(X)$ denotes the number of outcomes contains in the set X.

Goal: To develop counting techniques for determining the number of outcomes associated with the events of random experiments.

1.2.1 Multiplication Principle:



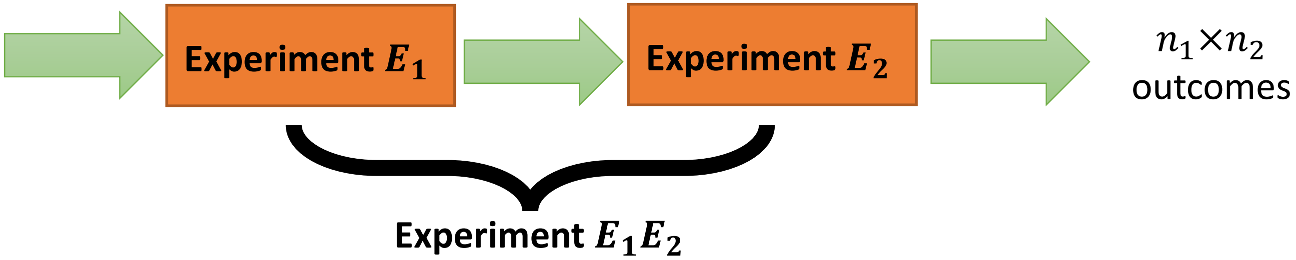
Experiment E_1

n_1 outcomes

Experiment E_2

n_2 outcomes

Sequential order:



Experiment E_1

Experiment E_2

$n_1 \times n_2$
outcomes

Experiment $E_1 E_2$

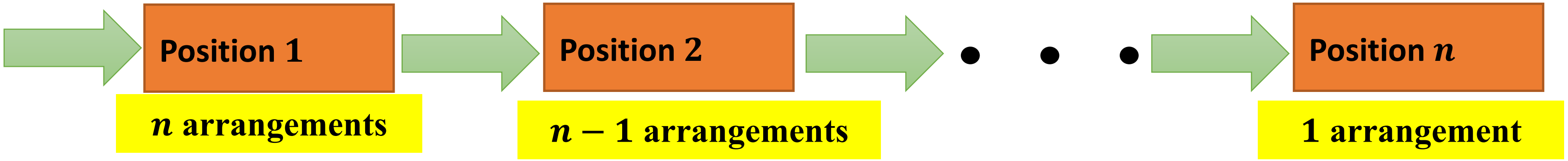
Example 1

- E_1: Select a rat from the cage containing either male or female
- E_2: for each selected rate, either drug A, drug B, or a placebo (P) is applied.
- The outcomes for the composite experiment are denoted by the ordered pair:

$$\begin{gathered}(F, A), (F, B), (F, P), \\ (M, A), (M, B), (M, P).\end{gathered}$$

1.2.2 permutation and Combination:

- Consider that n positions are to be filled with n different objects.
- This task can be handled by multiplication principle:



Position 1

n arrangements

Position 2

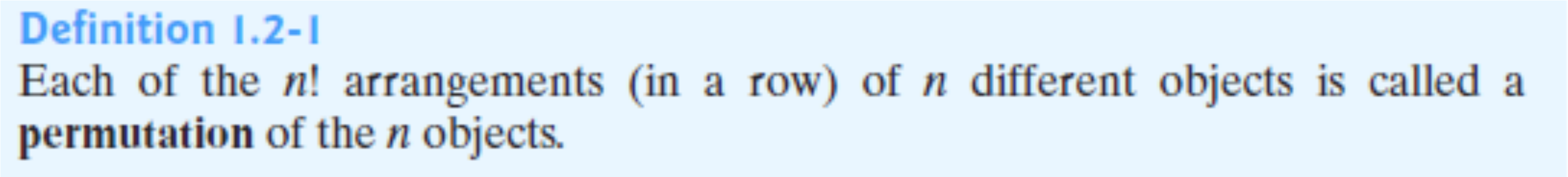
$n-1$ arrangements

• • •

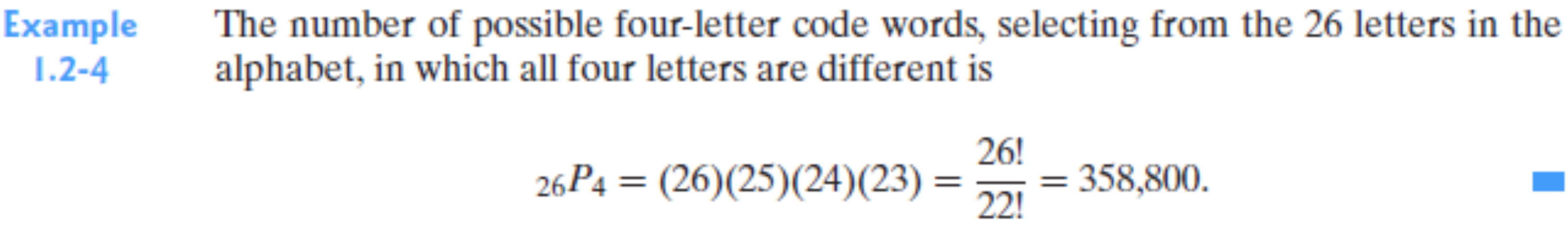
Position n

1 arrangement

In total ${}_nP_r \triangleq n(n-1)\cdots(n-r+1)$ possible arrangements



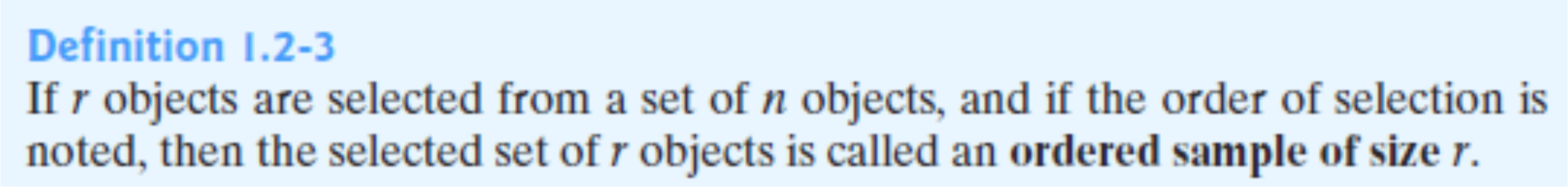
Definition 1.2-1
Each of the $n!$ arrangements (in a row) of n different objects is called a **permutation** of the n objects.



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

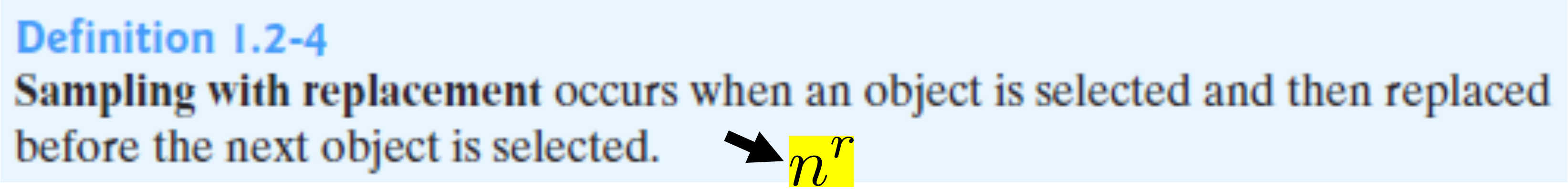
$$ {}_{26}P_4 = (26)(25)(24)(23) = \frac{26!}{22!} = 358{,}800. $$

■



Definition 1.2-3

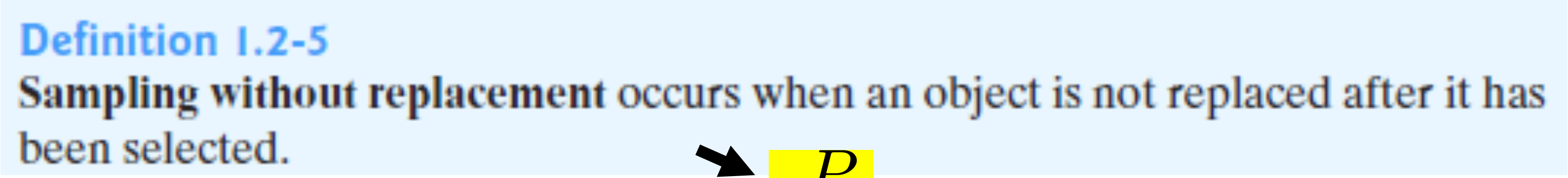
If r objects are selected from a set of n objects, and if the order of selection is noted, then the selected set of r objects is called an **ordered sample of size** r.



Definition 1.2-4

Sampling with replacement occurs when an object is selected and then replaced before the next object is selected.

n^r



Definition 1.2-5

Sampling without replacement occurs when an object is not replaced after it has been selected.

${}_nP_r$

Example 1-2-4 (Revised)

- The number of 4-letter word with different letters

$${}_{26}P_4$$

Sampling without replacement

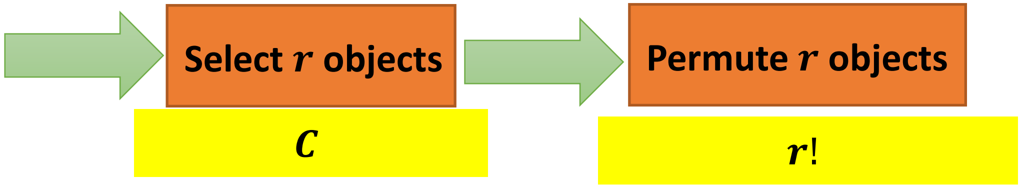
- The number of 4-letter word which may have the same letters

$$26^4$$

Sampling with replacement

- Sometimes, the order of selection is not important. We are interested in **the number of subsets of size r** taken from a set of n different objects.
  1. Recall permutation of n objects taken r at a time: ${}_nP_r$
  2. Let C denote the number of (unordered) subsets of size r that can be selected from n different objects.

     Then:



- $${}_nP_r = C \cdot r! \Rightarrow C = \frac{{}_nP_r}{r!} = \boxed{\frac{n!}{r!(n-r)!}}$$

Denoted as $\binom{n}{r}$ or ${}_nC_r$.
Named as ”choose r from n”

Definition 1.2-6

Each of the ${}_nC_r$ unordered subsets is called a **combination of n objects taken r at a time**, where

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

Example 3: The number of possible 5-card hands (in 5-card poker) drawn from a deck of 52 playing cards is

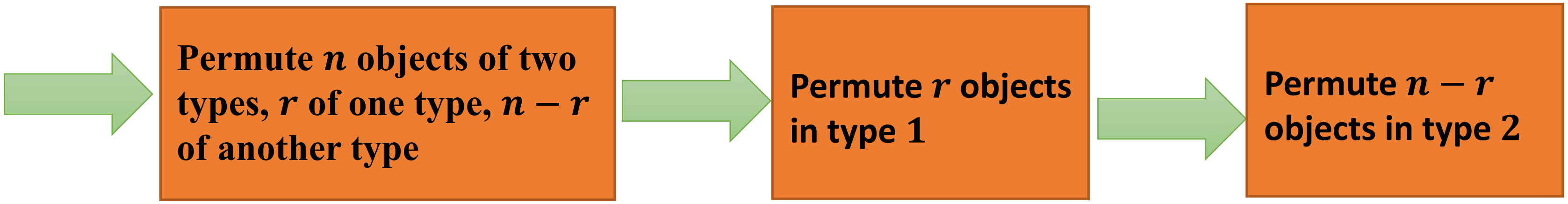
$$ {}_{52}C_5 = \binom{52}{5} = \frac{52!}{5!\,47!} = 2{,}598{,}960. $$

■

Remark: The numbers $\binom{n}{r}$ are frequently called **binomial coefficients**, since they arise in the expansion of a binomial. We illustrate this property by giving a justification of the binomial expansion

$$ (a+b)^n = \sum_{r=0}^{n} \binom{n}{r} b^r a^{n-r}. \qquad (1.2\text{-}1) $$

- Consider permutation of n objects of two types.
- Type 1 has r objects, and type 2 has $n-r$ objects.

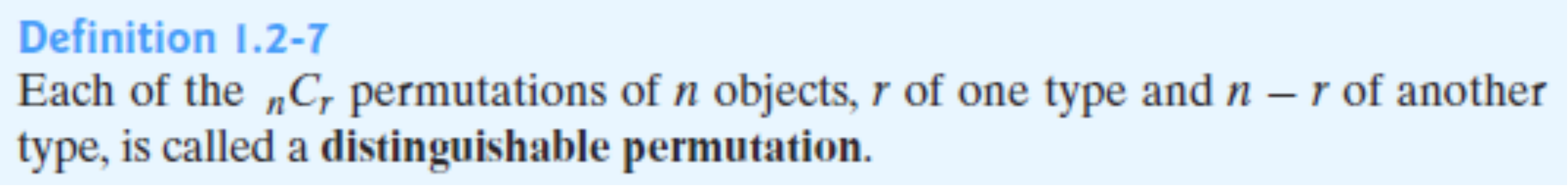
1. Recall permutation of n different objects is $n!$
2. 

Permute n objects of two types, r of one type, $n-r$ of another type

Permute r objects in type 1

Permute $n-r$ objects in type 2

$$n! = X \cdot (n-r)!r! \implies X = \frac{n!}{(n-r)!r!}$$



Definition 1.2-7

Each of the ${}_nC_r$ permutations of n objects, r of one type and $n-r$ of another type, is called a **distinguishable permutation.**

- Consider permutation of n objects of s types ($s \geq 2$).
- Type 1 has n_1 objects, type 2 has n_2 objects, ..., type s has n_s objects.
- $n_1, \ldots, n_s \in \mathbb{N}, n_1 + \cdots + n_s = 1.$
- The number of permutations of the n objects is

$$\binom{n}{n_1, \ldots, n_s} \triangleq \frac{n!}{n_1! n_2! \cdots n_s!}. \longrightarrow \boxed{\textbf{Multi-nomial coefficients}}$$

Remark: The name of multi-nominal coefficients arise in the multi-nominal theorem.

$$(a_1 + \cdots + a_s)^n = \sum_{\substack{n_1=0,\ldots,n_s=0 \\ n_1+\cdots+n_s=n}}^{n} \binom{n}{n_1, \ldots, n_s} a_1^{n_1} a_2^{n_2} \cdots a_s^{n_s}$$

1.3 Conditional Probability

1.3.1 A Motivation Example [Tulip Bulb Combination]

Assumption: Given 20 bulbs and all bulbs are “equally likely”

Table 1.3-1 Tulip combinations

	Early (E)	Late (L)	Totals
Red (R)	5	8	13
Yellow (Y)	3	4	7
Totals	8	12	20

Experiment 1: *Select one bulb randomly*

The probability that the selected bulb will be a red tulip (R) *is*

$$P(R)=\frac{N(R)}{N(S)}=\frac{13}{20}$$

*Experiment*2: *Select one bulb randomly from the bulbs that will boom early.*

The probability that the selected bulb will be a red tulip(R) *given that the selected bulb is known to bloom early*(E) *is*

$$P(R|E)=\frac{N(R\cap E)}{N(E)}=\frac{5}{8}=\frac{N(R\cap E)/N(S)}{N(E)/N(S)}=\frac{P(R\cap E)}{P(E)}.$$

$$Experiment2: \quad P(\underset{\overset{\uparrow}{A}}{R}|\underset{\overset{\nwarrow}{B}}{E}) = \frac{N(R \cap E)}{N(E)} = \frac{5}{8} = \frac{N(R \cap E)/N(S)}{N(E)/N(S)} = \frac{P(R \cap E)}{P(E)}$$

For experiment 2, we are interested only in those outcomes which are elements of a subset B of the sample space S.

1. The essential sample space is B (reduced from S to B)
2. Study the problem of how to define a new probability function associated with this sample space B,

Under the assumption that all out comes are "equally likely", the above example give us the idea:

$$P(A|B) = \frac{N(A \cap B)}{N(B)} = \frac{N(A \cap B)/N(S)}{N(B)/N(S)} = \frac{P(A \cap B)}{P(B)}$$

leading to the next definition:

Definition 1.3-1 [conditional probability]

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
- **Need not** to be "equally likely" !

Example 2: $P(A)=0.4,\ \ P(B)=0.5,\ \ P(A\bigcap B)=0.3,$

$$\Rightarrow P(A|B)=\frac{P(A\bigcap B)}{P(B)}=\frac{0.3}{0.5}=0.6$$

$$P(B|A)=\frac{P(B\bigcap A)}{P(A)}=\frac{0.3}{0.4}=0.75.$$

Can conditional probability be larger than 1 or negative ?

* *Conditional probability satisfy the axims for a probability function:*

1. $P(A|B)\geq 0$
2. $P(B|B)=1$
3. *If* $A_1, A_2, \ldots A_k$ *are mutually exclusive events, then*

$$P(A_1\cup A_2\cup\cdots\cup A_k|B)=P(A_1|B)+\cdots+P(A_k|B)$$

for each integer k, and

$$P(A_1\cup A_2\cup\cdots|B)=P(A_1|B)+P(A_2|B)+\cdots$$

for an infnitely, but countable number of events.

4. *The propability properties also holds for conditional probabilities. For example.*

$$P(A'|B)=1-P(A|B)$$

is true.

Example 3: 25 balloons on a board, of which 10 balloons are yellow, 8 are red, and 7 are green. A player randomly hits one of them.

A={The first balloon hit is yellow}

B={The second hit is yellow}

$P(A\cap B)$

Question:

What is the probability that the first two balloons are all yellow?

$$Solution: \; P(A)=\frac{10}{25}, P(B|A)=\frac{9}{24} \quad \Rightarrow \quad P(A\cap B)=P(A)P(B|A)=\frac{10}{25}\times\frac{9}{24}.$$

Definition 1.3-2 [multiplication rule]

The probability that two events, A and B, both occur is given by the **multiplication rule**,

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

provided P(A) > 0 or by

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

provided P(B) > 0.

Example 4: A bowl contains 10 chips in total, 7 blue and 3 red. Two chips drawn successively at random and without replacement.

Our goal is to compute the probability that $\underbrace{\text{the first draw is red}}_{A}$ **and** $\underbrace{\text{the second draw is blue}}_{B}$:

$$Solution:\ P(A)=\frac{3}{10},P(B|A)=\frac{7}{9}\ \Rightarrow\ P(A\cap B)=P(A)P(B|A)=\frac{3}{10}\times\frac{7}{9}=\frac{7}{30}.$$

Quiz: Roll a pair of 4-sided dice and observe the sum of the dice

A={ A sum of 3 is rolled }

B={ A sum of 3 or a sum of 5 is rolled }

C={ A sum of 3 is rolled before a sum of 5 is rolled }

Question: Compute P(A)、P(B) and P(C).

$$Consider\ \ P(A)\ \ and\ \ P(B):$$

$$the\ \ sample\ \ space\ \ S=\{(1,1),(1,2),\ldots,(4,4)\}$$

$$P(A)=\frac{\mathbb{N}(A)}{\mathbb{N}(S)}=\frac{2}{16},\qquad P(B)=\frac{\mathbb{N}(B)}{\mathbb{N}(S)}=\frac{6}{16}.$$

Quiz (c.n.t.)

Consider P(C):

- Method 1 [by definition]:

① Figure out the random experiment
② Figure out the sample space and the event.

For ①, repeat the experiment of rolling a pair of 4-sided dice and record the sum of dice. For each repetition, we keep rolling the dice till we see either a sum of 3 or a sum of 5. Then we stop because we have an answer to the problem whether a sum of 3 is rolled before a sum of 5 is rolled.

For instance,
repetition 1: 2,4,6,3.
repetition 2: 8,6,7,4,5.
repetition 3: 6,5

The sums other than 3 and 5 don't matter and we can remove them.
Repetition 1: a sum of 3 first.
Repetition 2: a sum of 5 first.
Repetition 3: a sum of 5 first.

The problem reduces to roll the pair of dice once and compute the probability that the sum is 3.

Quiz (c.n.t.)

For ②, the reduced sample space

$$S_r=\left\{\begin{matrix}(1,2), & (2,1)\\ (2,3), & (3,2)\\ (1,4), & (4,1)\end{matrix}\right\}\begin{matrix}\longrightarrow \text{Gives a sum of 3}\\ \left.\begin{matrix} \\ \\ \end{matrix}\right\}\longrightarrow \text{Gives a sum of 5}\end{matrix}$$

$$\Rightarrow P(C)=P(\{\ roll\ \ the\ \ pair\ \ of\ \ dice\ \ once\ \ and\ \ the\ \ sum\ \ is\ \ 3\})$$

$$=\frac{N(\{\ roll\ \ the\ \ pair\ \ of\ \ dice\ \ once\ \ and\ \ the\ \ sum\ \ is\ \ 3\})}{N(S_r)}$$

$$=\frac{2}{6}$$

Quiz (c.n.t.)

Method 2 [by conditional probability]

$$P(C) = P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/16}{6/16} = \frac{2}{6}$$

Note that the event “A|B” is the same as event “C”
This is because

① Event C is concerned with the cases where the sum is either 3 or 5

②”B” happened means that the sum is either 3 or 5. If B happened, then A|B is nothing but the event { roll the pair of dice once and the sum is 3}.

Section 1.4 independent events

- Intuition and motivation examples.

Intuition: For certain pair of events, the occurrence of one of them **may** or **may not** change the probability of the occurrence of the other. In the latter case, they are said to be **independent events**.

Example 1:

Experiment: flip a coin twice and observe the sequence of heads and tails.

Sample Space: S={ HH, HT, TH, TT}

Events: A={ heads on the first flip }={ HH, HT },
B={ tails on the second flip }={ HT, TT },
C={ tails on both flip }={ TT }.

Then $P(A)=\frac{2}{4},\ P(B)=\frac{2}{4},\ P(C)=\frac{1}{4}.$

Given that C has occurred, then $P(B|C)=1\ \ because\ \ C\subset B$

Given that A has occurred, then $P(B|A)=\frac{P(B\cap A)}{P(A)}=\frac{1/4}{3/4}=\frac{1}{2}$

So the occurrence of A has not changed the probability of B.

$P(B|A)=P(B),\quad Similarly,\quad \boxed{P(A|B)=P(A).}$ → Verify by yourself

Example 1(c.n.t):

- Intuitionally, this means that the probability of B doesn't depend on the knowledge about event A.

→ A and B are independent events

- That is, events A and B are independent if the occurrence of one of them does not affect the probability of the occurrence of the other. In math,

$$P(B\mid A)=\overset{P(A)>0}{P(B)} \qquad P(A\mid B)=\overset{P(B)>0}{P(A)}$$

- This example motivates the following definition of independent events.

Definition 1.4-1 [independent events]

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

Example 2: A red die and a white die are rolled

- S={ (1,1), (1,2), ……, (6,6) } (Number of all outcomes is 36)
- A={ 4 on the red die }. B={ sum of dice is odd }.

Assuming the two dies are fair. Are events A and B independent ?

$$Solution: P(A) = \frac{6}{36}, \quad P(B) = \frac{18}{36} \quad P(A \cap B) = \frac{3}{36}$$

$$P(A \cap B) = \frac{3}{36} = P(A)P(B) = \frac{6}{36} \times \frac{18}{36} \quad \Rightarrow \quad A \text{ and } B \text{ are independent.}$$

Theorem 1.4-1

If A and B are independent events, then the following pairs of events are also independent:

(a) A and B;

(b) A' and B;

(c) A' and B'.

The proofs are in the later page.

Proofs of theorem 1.4-1

$(a) Proof:$

$$
\begin{aligned}
P(A\cap B') &= P(A)P(B'|A) && (multiplication\ \ rule)\\
&= P(A)\left[1-P(B|A)\right] && (axims\ \ for\ \ conditional\ \ probability)\\
&= P(A)[1-P(B)] && (definiton\ \ of\ \ independent\ \ events)\\
&= P(A)P(B'). && (properties\ \ of\ \ probability\ \ function)
\end{aligned}
$$

The proofs of part(b) and (c) will be written simply:

$Proof:$

$(b)P(A'\cap B)=P(B)P(A'|B)=P(B)\left[1-P(A|B)\right]=P(B)[1-P(A)]=P(B)P(A')$

$(c)P(A'\cap B')=P\left[(A\cup B)'\right]=1-(A\cup B)=1-P(A)-P(B)+P(A\cap B)$

$=1-P(A)-P(B)+P(A)P(B)=[1-P(A)][1-P(B)]=P(A')P(B').$

$Q.E.D.$

Then let’s extend the definition of independent events to more than two events.

Definition 1.4-2 [mutually independent]

Events A, B, and C are **mutually independent** if and only if the following two conditions hold:

(a) A, B, and C are **pairwise** independent; that is,

$$\left\{\begin{array}{l} P(A\cap B)=P(A)P(B) \\ P(A\cap C)=P(A)P(C) \\ P(B\cap C)=P(B)P(C) \end{array}\right.$$

(b) $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Remark:

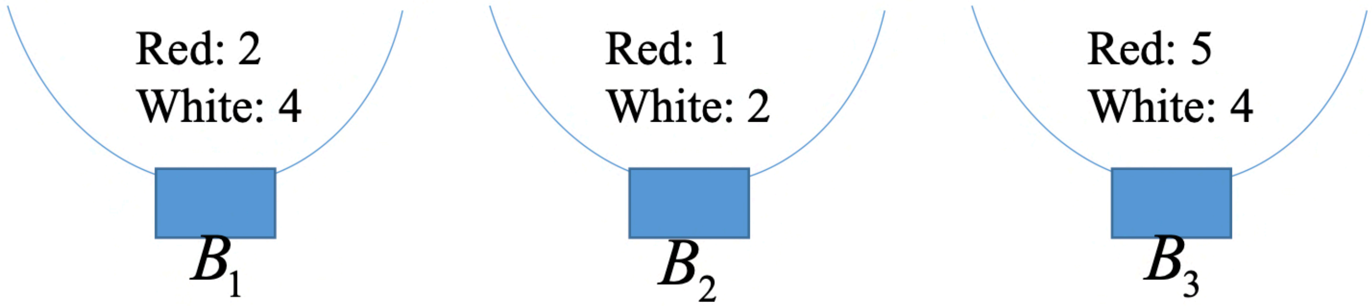
1. This definition can be extended to the mutual independence of four or more events. In such a extension, each pair, triple, quartet, and so on, must satisfy this type of multiplication rule. In the following we will discuss its definition in mathematical formal.
2. If there is no possibility of misunderstanding, independent is often used without the modifier mutually when several events are considered.

Events $A_1, A_2, \ldots, A_k$ *are independent if and only if the following condition hold*:

$$P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_j})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_j}),\qquad j=2,\ldots,k.$$

Section 1.5 Bayes's theorem

Motivation example:

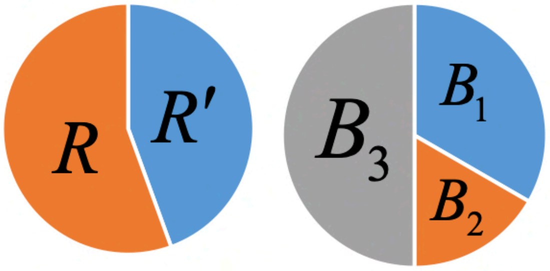


Experiment: Select a bowl first, and then draw a chip from the selected bowl.

Goal: compute the probability of event R={ draw a red chip }.

Assumption: $P(B_1)=\frac{1}{3},\ P(B_2)=\frac{1}{6},\ P(B_3)=\frac{1}{2}.$

Question1: Compute the probability of event R



$$\mathit{Solution}: P(R)=P(S\cap R)$$

$$=P[(B_1\cup B_2\cup B_3)\cap R]=P[(B_1\cap R)\cup(B_2\cap R)\cup(B_3\cap R)]$$

$$=P(B_1\cap R)+P(B_2\cap R)+P(B_3\cap R)$$

$$=P(B_1)P(R|B_1)+P(B_2)P(R|B_2)+P(B_3)P(R|B_3)$$

$$=\frac{1}{3}\times\frac{2}{6}+\frac{1}{6}\times\frac{1}{3}+\frac{1}{2}\times\frac{5}{9}=\frac{4}{9}.$$

Question2: Suppose now that the outcome of the experiment is a red chip, but we do not know from which bowl the chip was drawn. We are interested in the **conditional probability** that the chip was drawn from the bowl, namely, $P(B_1|R),\ P(B_2|R),\ P(B_3|R).$

From the definition of conditional probability, we consider

$$P(B_1|R)=\frac{P(B_1\cap R)}{P(R)}=\frac{P(B_1)P(R|B_1)}{P(B_1)P(R|B_1)+P(B_2)P(R|B_2)+P(B_3)P(R|B_3)}$$

$$=\frac{\frac{1}{3}\times\frac{2}{6}}{\frac{4}{9}}=\frac{1}{4}\qquad Similarly,\ P(B_2|R)=\frac{1}{8},\ P(B_3|R)=\frac{5}{8}$$

Recall that:

$$P(B_1)=\frac{1}{3},\ P(B_2)=\frac{1}{6},\ P(B_3)=\frac{1}{2}\qquad\rightarrow\ \textit{Prior Probability}$$

$$P(B_1|R)=\frac{1}{4},\ P(B_2|R)=\frac{1}{8},\ P(B_3|R)=\frac{5}{8}\qquad\rightarrow\ \textit{posterior probability}$$

We observe: ①$P(B_i)$ *different from* $P(B_i|R)$

②*The changer coincide with our intuition.*

Generalization: Assume that

1. S is a sample space, and $B_1, B_2, \ldots, B_m$ are mutually exclusive and exhaustive. That is:

$$S = B_1 \cup B_2 \cup \cdots \cup B_m \quad and \quad B_i \cap B_j = \phi, i \neq j$$

2. The prior probabilities of B_i is positive. That is:

$$P(B_i) > 0, \quad i = 1, 2, \ldots, m.$$

Then we have

a) For any event A,

$$
\begin{aligned}
P(A) = P(A \cap S) &= P\left[A \cap (B_1 \cup B_2 \cup \cdots \cup B_m)\right] \\
&= P\left[(A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_m)\right] \\
&= P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_m) \\
&= \sum_{i=1}^{m} P(A \cap B_i)
\end{aligned}
$$

$$\Rightarrow \boxed{P(A) = \sum_{i=1}^{m} P(B_i) P(A \mid B_i)} \longrightarrow \text{Total probability}$$

b) If P(A)＞0， then

$$P(B_k \mid A) = \frac{P(B_k \cap A)}{P(A)}, \quad k = 1, \ldots, m.$$

$$Hence \; \boxed{P(B_k \mid A) = \frac{P(B_k) P(A \mid B_k)}{\sum_{i=1}^{m} P(B_i) P(A \mid B_i)}.} \longrightarrow \text{Bayes's Theorem}$$

- $P(B_k) \quad \rightarrow$ *Prior Probability*
- $P(B_k \mid A) \quad \rightarrow$ *posterior probability*
- $P(A \mid B_i) \quad \rightarrow$ *likelihood of B_k, A is called a data.*