Performance Evaluation

- One in-class exams: 10%
  - 4/13
- One midterm exams: 20%
  - 5/4
- One final exam: 40%
  - 6/8
- Homework assignments: 25%
- In-class performance: 5%

Chapter 2 Discrete Distributions

- Random Variables of the Discrete Type
  - Uniform Distribution
  - Hypergeometric Distribution
- Mathematical Expectation
- Moment Generating Function
- Bernoulli Trials and the Binomial Distribution
- Geometric and Negative Binomial Distribution
- The Poisson Distribution

2.1 Random Variables of the Discrete Type

- Def 2.1-1: Random Variable X (abbreviated by r.v.)
  - a **function** that maps the possible outcomes of an experiment to real numbers.（把實驗結果轉成實數）
  - i.e., assign to each element s in S one and only one real number
  - Notice that: *a function assigns one and only one number in the range (output) to each number in the domain (input)*
  - $X: S \rightarrow R$, where S is the set of all outcomes of an experiment, and R is the set of real numbers.

- The space of X is the set of real numbers S_X , where

$$S_X = \{x: X(s) = x , s \in S\}$$

From now on, S_X is replaced by S in this class

An Example of Random Variable

- If we toss a coin one time, then there are two possible outcomes
  - namely "head up" and "tail up".
- We can define a random variable X that maps "head up" to 1 and "tail up" to 0.
- We also can define a random variable Y that maps "head up" to 0 and "tail up" to 1.
- The spaces of both random variables X and Y are $S = \{0,1\}$.

Further Illustration of Random Variables

- A random variable corresponds to a quantitative interpretation of the outcomes of an experiment.
- For example, a company offers its employees a drawing in its yearend party.
- A computer will randomly select an employee for the first prize of $100,000 based on the employees' ID number, which ranges from 1 to 100.

- In addition, the computer will randomly select two more employees for the second and third prizes of \$50,000 and \$10,000, respectively.
- Assume that each employee can receive only one award and the drawing starts with the third prize and ends with the first prize.
- Number of possible outcomes
  - totally $100 \times 99 \times 98 = 970200$

- To Edward, whose employee ID number is 10, the random variable of his interest is as follows:

  $X(<10, *, *>) = 10{,}000$

  $X(<*, 10, *>) = 50{,}000$ space?

  $X(<*, *, 10>) = 100{,}000$

  $X(\text{all other outcomes}) = 0$

- To Grace, whose employee ID number is 30, the random variable of her interest is as follows:

  $Y(<30, *, *>) = 10{,}000$

  $Y(<*, 30, *>) = 50{,}000$ space?

  $Y(<*, *, 30>) = 100{,}000$

  $Y(\text{all other outcomes}) = 0$

- X and Y map some outcomes to different real numbers.
- However, the spaces of X and Y
  - are identical
  - both are {0, 10000, 50000, 100000}.
- The probability functions of X and Y are also equal.

  $Prob(X=10{,}000) = Prob(Y=10{,}000) = 0.01$

  $Prob(X=50{,}000) = Prob(Y=50{,}000) = 0.01$

  $Prob(X=100{,}000) = Prob(Y=100{,}000) = 0.01$

  $Prob(X=0) = Prob(Y=0) = 0.97$

- The expected values of X and Y
  - are equal to

$$E[X] = E[Y]$$

$$= 10{,}000 * 0.01 + 50{,}000 * 0.01$$

$$+ 100{,}000 * 0.01$$

$$= 1600.$$

Discrete Random Variables

- Given a random variable X,
  - let S denote the space of X.
- If S is a finite or countable infinite set, then X is said to be a discrete random variable.

Countable Infinite

- A set is said to be countable infinite if
  - it contains infinite number of elements
  - there exists a one-to-one mapping between each element of the set and the positive integers.

Examples of Countable / Uncountable Infinite

- The set of integer numbers is ?
  - countable.
- The set of fractional numbers is ?
  - countable.
- The set of real numbers is ?
  - uncountable.

Mapping of Integer Numbers to Positive Integer Numbers

- 0 1 –1 2 –2 3 –3 ……
- 0 1 2 3 4 5 6 ……

Mapping of Fractional Numbers to Positive Integer Numbers ?

Mapping of Fractional Numbers to Positive Integer Numbers

	1	-1	2	-2	3	-3	4	-4
1	1/1	-1/1	2/1	-2/1				
-1	1/-1	-1/-1	2/-1					
2	1/2	-1/2						
-2	1/-2							
3								
-3								
4								
-4								

Probability Mass Function

- The probability mass function (p.m.f.) of a discrete random variable X is defined to be

$$P_X(k) \equiv Prob(X = k) = \sum_{q \in Q_k} Prob(q),$$

where Q_k contains all outcomes that are mapped to k by random variable X.

課本的定義？

- In the previous example of drawing,

$$P_X(10{,}000) = Prob(X = 10{,}000)$$

$$= \sum_{\substack{<10,i,j> \\ i \neq 10,\, j \neq 10 \\ i \neq j}} Prob(<10, i, j>)$$

$$= \sum_{\substack{<10,i,j> \\ i \neq 10,\, j \neq 10 \\ i \neq j}} \frac{1}{100 \times 99 \times 98} = 0.01.$$

$$P_X(k) \equiv Prob(X = k) = \sum_{q \in Q_k} Prob(q),$$

where Q_k contains all outcomes that are mapped to k by random variable X.

- In fact, the p.m.f. of a random variable is defined on a set of events of the experiment conducted.
  - $f(x) \equiv P(X=x)$ as shown in the Hogg's textbook
- In the previous drawing example, the set of outcomes that are mapped to 10,000 by X is an event.

- To Edward, whose employee ID number is 10, the random variable of his interest is as follows:

  $X(<10, *, *>) = 10{,}000$

  $X(<*, 10, *>) = 50{,}000$

  $X(<*, *, 10>) = 100{,}000$

  $X(\text{all other outcomes}) = 0$

- To Grace, whose employee ID number is 30, the random variable of her interest is as follows:

  $Y(<30, *, *>) = 10{,}000$

  $Y(<*, 30, *>) = 50{,}000$

  $Y(<*, *, 30>) = 100{,}000$

  $Y(\text{all other outcomes}) = 0$

- Furthermore, in the previous drawing example, random variables X and Y map some outcomes to different real numbers. However, X and Y have the same distribution, i.e. the p.m.f. of X and the p.m.f. of Y are equal. More precisely,

?

$$P_X(k) = P_Y(k)$$
$$\text{for every } k \in \{0, 10000, 50000, 100000\}.$$

- The distribution of the random variables X
  - *The distribution of probability associated with space S of the r.v. X*

Properties of Probability Mass Function

$$P_X(k) \equiv Prob(X = k) = \sum_{q \in Q_k} Prob(q),$$

- The p.m.f. of a random variable X satisfies the following three properties:

$$(1)\ \ P_X(x) > 0 \ \ ,\ \ x \in S\ :\ the\ \ space\ \ of\ \ X.$$

$$(2)\ \ \sum_{x_i \in S} P_X(x_i) = 1\ .$$

$$(3)\ \ Prob(A) = \sum_{x_j \in A} P_X(x_j)\ ,\ \ where\ \ A \subseteq S.$$

課本將以上的Properties視為定義，如下一頁：

For a random variable X of the discrete type, the probability $P(X = x)$ is frequently denoted by $f(x)$, and this function $f(x)$ is called the **probability mass function**. Note that some authors refer to $f(x)$ as the probability function, the frequency function, or the probability density function. In the discrete case, we shall use "probability mass function," and it is hereafter abbreviated pmf.

Let $f(x)$ be the pmf of the random variable X of the discrete type, and let S be the space of X. Since $f(x) = P(X = x)$ for $x \in S$, $f(x)$ must be nonnegative for $x \in S$, and we want all these probabilities to add to 1 because each $P(X = x)$ represents the fraction of times x can be expected to occur. Moreover, to determine the probability associated with the event $A \in S$, we would sum the probabilities of the x values in A. This leads us to the following definition.

Definition 2.1-2

The pmf $f(x)$ of a discrete random variable X is a function that satisfies the following properties:

(a) $f(x) > 0, \qquad x \in S;$

(b) $\displaystyle\sum_{x \in S} f(x) = 1;$

(c) $\displaystyle P(X \in A) = \sum_{x \in A} f(x), \qquad \text{where } A \subset S.$

課本第29-30頁

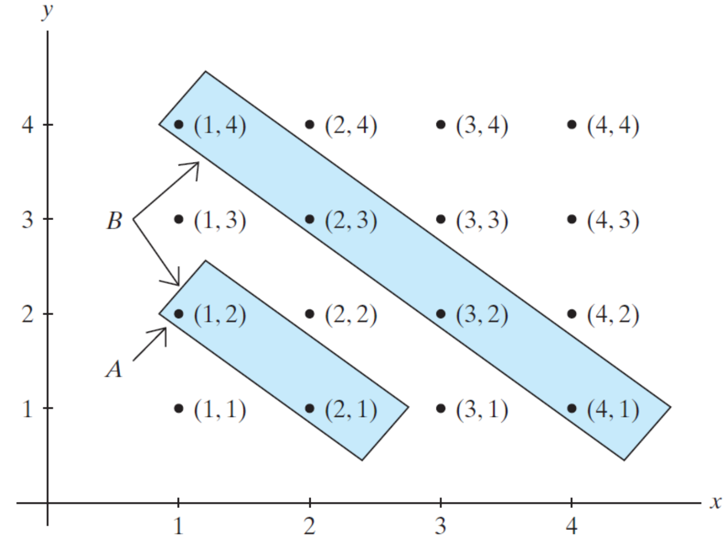

Figure 1.3-2 Dice example

Example 1.3-4

A pair of fair four-sided dice is rolled and the sum is determined. Let A be the event that a sum of 3 is rolled, and let B be the event that a sum of 3 or a sum of 5 is rolled. In a sequence of rolls, the probability that a sum of 3 is rolled before a sum of 5 is rolled can be thought of as the conditional probability of a sum of 3 given that a sum of 3 or 5 has occurred; that is, the conditional probability of A given B is

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

Note that for this example, the only outcomes of interest are those having a sum of 3 or a sum of 5, and of these six equally likely outcomes, two have a sum of 3. (See Figure 1.3-2 and Exercise 1.3-13.)