Probability Distributions

Chapter 6

What is a Probability Distribution（機率分配）?

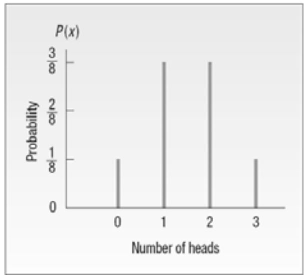
PROBABILITY DISTRIBUTION A listing of all the outcomes of an experiment and the probability associated with each outcome.

Experiment: Toss a coin three times. Observe the number of heads. The possible results are: zero heads, one head, two heads, and three heads.
What is the probability distribution for the number of heads?

Possible Result	Coin Toss First	Coin Toss Second	Coin Toss Third	Number of Heads
1	T	T	T	0
2	T	T	H	1
3	T	H	T	1
4	T	H	H	2
5	H	T	T	1
6	H	T	H	2
7	H	H	T	2
8	H	H	H	3

Probability Distribution of Number of Heads Observed in 3 Tosses of a Coin

Number of Heads, x	Probability of Outcome, $P(x)$
0	$\frac{1}{8} = .125$
1	$\frac{3}{8} = .375$
2	$\frac{3}{8} = .375$
3	$\frac{1}{8} = .125$
Total	$\frac{8}{8} = 1.000$



Probability Distribution of Number of Heads Observed in 3 Tosses of a Coin

- The probability distribution for the number of heads is

$$P(X=a)=\begin{cases} 1/8 & a=0 \\ 3/8 & a=1 \\ 3/8 & a=2 \\ 1/8 & a=3 \end{cases}$$

Characteristics of a Probability Distribution

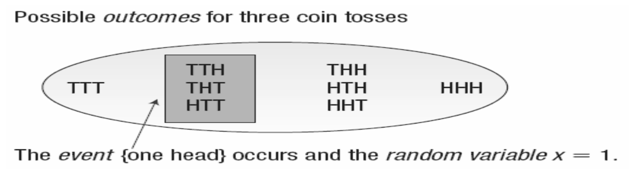
CHARACTERISTICS OF A PROBABILITY DISTRIBUTION

1. The probability of a particular outcome is between 0 and 1 inclusive.
2. The outcomes are mutually exclusive events.
3. The list is exhaustive. So the sum of the probabilities of the various events is equal to 1.

5

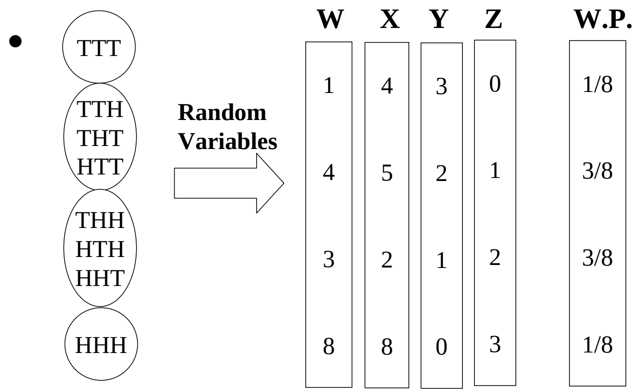
Random Variables（隨機變數）

- Experiment – a coin is tossed 3 times.
- Outcomes – eight outcomes
- Event – one head occurs in the three tosses.
- Random variable - a quantity resulting from an experiment that, by chance, can assume different values, associated with a probability value.



6

Random Variables（隨機變數）



7

Example: Random Variables

Consider a random experiment of tossing three coins. X is a random variable, which assigns the number of heads to a number, associated with a probability value .

$$P(X=x)=\begin{cases}1/8 & x=0\\ 3/8 & x=1\\ 3/8 & x=2\\ 1/8 & x=3\end{cases}$$

Number of Heads, x	Probability of Outcome, $P(x)$
0	$\frac{1}{8} = .125$
1	$\frac{3}{8} = .375$
2	$\frac{3}{8} = .375$
3	$\frac{1}{8} = .125$
Total	$\frac{8}{8} = 1.000$

8

Example: Random Variables

Consider the same random experiment of tossing three coins. Let Y be another random variable, which assigns the number of heads to a number, associated with a probability value .

$$P(Y=y)=\begin{cases}1/8 & y=3\\ 3/8 & y=2\\ 3/8 & y=1\\ 1/8 & y=0\end{cases}$$

9

Example : Random Variable

- Let X denote the number of spots of tossing two fair dice. The probability distribution for the number of spots as tossing two fair dice is

$$P(X=x)=\begin{cases}1/36 & x=2\\ 2/36 & x=3\\ 3/36 & x=4\\ 4/36 & x=5\\ 5/36 & x=6\\ 6/36 & x=7\\ 5/36 & x=8\\ 4/36 & x=9\\ 3/36 & x=10\\ 2/36 & x=11\\ 1/36 & x=12\end{cases}$$

10

Types of Random Variables

- Discrete Random Variable （離散隨機變數） can assume only certain clearly separated values. It is usually the result of counting something.

- Continuous Random Variable （連續隨機變數） can assume an infinite number of values within a given range. It is usually the result of some type of measurement

11

Discrete Random Variables - Examples

- The number of students in a class.
- The number of children in a family.
- The number of cars entering a carwash in a hour.
- Number of home mortgages approved by Coastal Federal Bank last week.

12

Example：離散隨機變數

- 已知某公司總機平均每10 分鐘有 4 通電話，則 5 分鐘內有 X 通電話的機率分配為：

$$P(X=x)=\frac{2^x e^{-2}}{x!}, \quad x=0,1,2,3,.....$$

Continuous Random Variables - Examples

- The distance students travel to class.
- The time it takes an executive to drive to work.
- The length of an afternoon nap.
- The length of time of a particular phone call.

Example：連續隨機變數

- 假設某股票的平均報酬率 5%，報酬率標準差 30 %。假設今日股價 $10，則一年後的股價 S 的機率分配為：

$$P(S=s)=\frac{1}{0.3S\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln S-\ln 10-0.05}{0.3}\right)^2}, \quad 0<S<\infty$$

期望值（Expectation）

假設X是個隨機變數，其機率分配函數爲 f(x)。對任意給定的一個函數 g(X)，g(X) 的期望值定義如下：

$$E[g(X)]=\begin{cases}\sum_x g(x)f(x), & X\text{是離散隨機變數}\\ \int_x g(x)f(x)dx, & X\text{是連續隨機變數}\end{cases}$$

Example: Random Variables

Consider a random experiment of tossing three coins. X is a random variable, which assigns the number of heads to a number, associated with a probability value .

$$P(X=x)=\begin{cases}1/8 & x=0\\ 3/8 & x=1\\ 3/8 & x=2\\ 1/8 & x=3\end{cases}$$

$$\mathrm{E[X]}=\sum_{x=0}^{3} xP(X=x)$$

$$=0\cdot\frac{1}{8}+1\cdot\frac{3}{8}+2\cdot\frac{3}{8}+3\cdot\frac{1}{8}=\frac{3}{2}$$

17

期望值的性質

- 假設 X 是一隨機變數，其機率分配為 f(x)，a, b 與 c 是常數，則：
  - E[a] = a
  - E[aX] = a E[X]
  - E[aX + b] = aE[X]+b
  - E[aX^2+bX+c]= aE[X^2]+bE[X]+c

18

Features of a Discrete Distribution

The main features of a discrete probability distribution are:

- The sum of the probabilities of the various outcomes is 1.
- The probability of a particular outcome is between 0 and 1.
- The outcomes are mutually exclusive.

19

離散機率分配

- 一般來說，離散機率分配可以寫成以下一般式：

$$P(X=n_j)=p_j, \quad j=1,2,3,\ldots.$$

且滿足 $\sum_{j=1}^{\infty} P(X=n_j) = \sum_{j=1}^{\infty} p_j = 1$

20

The Mean of a Probability Distribution

MEAN

- The mean is a typical value used to represent the central location of a probability distribution.

- The mean of a probability distribution is also referred to as its **expected value.**
- Expected value of X is denoted by μ = E[X]

MEAN OF A PROBABILITY DISTRIBUTION $\mu = \Sigma[xP(x)]$ **[6–1]**

21

The Variance, and Standard Deviation of a Probability Distribution

Variance and Standard Deviation

- Measures the amount of spread in a distribution
- The **variance** is found by:
  1. Subtract the mean from each value, and square this difference.
  2. Multiply each squared difference by its probability.
  3. Sum the resulting products. to arrive at.
- The **standard deviation** is found by taking the positive square root of the variance.
- $\sigma^2 = E[(X-\mu)^2]$, the expected value of $(X-\mu)^2$.

VARIANCE OF A PROBABILITY DISTRIBUTION $\sigma^2 = \Sigma[(x-\mu)^2 P(x)]$ **[6–2]**

22

Computation Form for Variance

$$\sigma^2 = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2]$$
$$= E[X^2] - 2\mu E[X] + \mu^2$$
$$= E[X^2] - 2\mu\mu + \mu^2$$
$$= E[X^2] - \mu^2$$
$$= E[X^2] - \left(E[X]\right)^2$$

23

Mean, Variance, and Standard Deviation of a Probability Distribution - Example

John Ragsdale sells new cars for Pelican Ford. John usually sells the largest number of cars on Saturday. He has developed the following probability distribution for the number of cars he expects to sell on a particular Saturday.

Number of Cars Sold, *x*	Probability, *P*(*x*)
0	.10
1	.20
2	.30
3	.30
4	.10
Total	1.00

24

Mean of a Probability Distribution - Example

$\mu = \Sigma[xP(x)]$
$= 0(.10) + 1(.20) + 2(.30) + 3(.30) + 4(.10)$
$= 2.1$

Number of Cars Sold, x	Probability, $P(x)$	$x \cdot P(x)$
0	.10	0.00
1	.20	0.20
2	.30	0.60
3	.30	0.90
4	.10	0.40
Total	1.00	$\mu = 2.10$

25

Variance and Standard Deviation of a Probability Distribution - Example

Number of Cars Sold, x	Probability, $P(x)$	$(x-\mu)$	$(x-\mu)^2$	$(x-\mu)^2P(x)$
0	.10	0 − 2.1	4.41	0.441
1	.20	1 − 2.1	1.21	0.242
2	.30	2 − 2.1	0.01	0.003
3	.30	3 − 2.1	0.81	0.243
4	.10	4 − 2.1	3.61	0.361
				$\sigma^2 = 1.290$

26

Bernoulli Distribution （百努里分配）

- 簡單隨機實驗只有兩種結果，一種為成功，機率為 π；另一種失敗，機率為 1- π 。
  - 例如擲一公平銅板，正面機率 0.5，反面機率 0.5 。
  - 生產線上抽取一產品樣本，良品機率 π ，不良品機率1- π 。
- 以隨機變數 X 表示簡單隨機實驗的結果：
  - X = 1 表示成功，機率 π ； X = 0 表示失敗，機率 1- π 。
  - 隨機變數 X 的機率分配：

$$f(x) = \pi^x (1-\pi)^{1-x},\ \ X = 0,1,\ \ 0 \le \pi \le 1$$

27

Binomial Probability Distribution （二項分配）

- 丟擲一公平硬幣 10 次，假設以 X 表示出現正面的次數，則隨機變數 X 即是服從二項分配 (Binomial Distribution)。X 的機率分配為：

$$P(X = x) = C_x^{10}(0.5)^x(0.5)^{10-x},\quad x = 0,1,2,...,10$$

- 若丟擲 n 次，則 X 的機率分配為：

$$P(X = x) = C_x^{n}(0.5)^x(0.5)^{n-x},\quad x = 0,1,2,...,n$$

- 丟擲一公平硬幣 n 次，即是作 n 次 Bernoulli 實驗。因此假設每次實驗結果為隨機變數 X_i，則

$$X = \sum_{i=1}^{n} X_i$$

28

Binomial Probability Distribution （二項分配）

Characteristics of a Binomial Probability Distribution

- The random variable is the result of counts.
- Each trial is *independent and identical* of any other trial.
- There are only two possible outcomes on a particular trial of an experiment.
- The outcomes are mutually exclusive.

29

Binomial Probability Formula

BINOMIAL PROBABILITY FORMULA $P(x) = {}_nC_x \pi^x (1 - \pi)^{n-x}$ [6–3]

where:

C denotes a combination.
n is the number of trials.
x is the random variable defined as the number of successes.
π is the probability of a success on each trial.

30

Binomial Probability - Example

There are five flights daily from Pittsburgh via US Airways into the Bradford, Pennsylvania, Regional Airport. Suppose the probability that any flight arrives late is .20.

What is the probability that none of the flights are late today?

$$P(x=0) = {}_nC_x \pi^x (1-\pi)^{n-x}$$
$$= {}_5C_0 (.20)^0 (1-.20)^{5-0}$$
$$= (1)(1)(.3277)$$
$$= 0.3277$$

31

Binomial Dist. – Mean and Variance

MEAN OF A BINOMIAL DISTRIBUTION $\mu = n\pi$ [6–4]

VARIANCE OF A BINOMIAL DISTRIBUTION $\sigma^2 = n\pi(1 - \pi)$ [6–5]

32

Binomial Dist. – Mean and Variance: Example

For the example regarding the number of late flights, recall that π =.20 and n = 5.

What is the average number of late flights?

What is the variance of the number of late flights?

$$\mu = n\pi$$
$$= (5)(0.20) = 1.0$$

$$\sigma^2 = n\pi(1-\pi)$$
$$= (5)(0.20)(1-0.20)$$
$$= (5)(0.20)(0.80)$$
$$= 0.80$$

33

Binomial Dist. – Mean and Variance: Another Solution

Number of Late Flights, x	$P(x)$	$xP(x)$	$x-\mu$	$(x-\mu)^2$	$(x-\mu)^2P(x)$
0	0.3277	0.0000	−1	1	0.3277
1	0.4096	0.4096	0	0	0
2	0.2048	0.4096	1	1	0.2048
3	0.0512	0.1536	2	4	0.2048
4	0.0064	0.0256	3	9	0.0576
5	0.0003	0.0015	4	16	0.0048
		μ = 1.0000			σ^2 = 0.7997

34

Binomial Distribution - Table

Five percent of the worm gears produced by an automatic, high-speed Carter-Bell milling machine are defective. What is the probability that out of six gears selected at random none will be defective? Exactly one? Exactly two? Exactly three? Exactly four? Exactly five? Exactly six out of six?

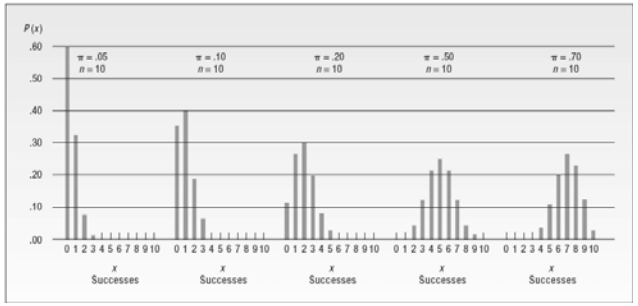
TABLE 6–2 Binomial Probabilities for n = 6 and Selecte Values of π

	n = 6 Probability										
$x\backslash\pi$	.05	.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
0	.735	.531	.262	.118	.047	.016	.004	.001	.000	.000	.000
1	.232	.354	.393	.303	.187	.094	.037	.010	.002	.000	.000
2	.031	.098	.246	.324	.311	.234	.138	.060	.015	.001	.000
3	.002	.015	.082	.185	.276	.313	.276	.185	.082	.015	.002
4	.000	.001	.015	.060	.138	.234	.311	.324	.246	.098	.031
5	.000	.000	.002	.010	.037	.094	.187	.303	.393	.354	.232
6	.000	.000	.000	.001	.004	.016	.047	.118	.262	.531	.735

35

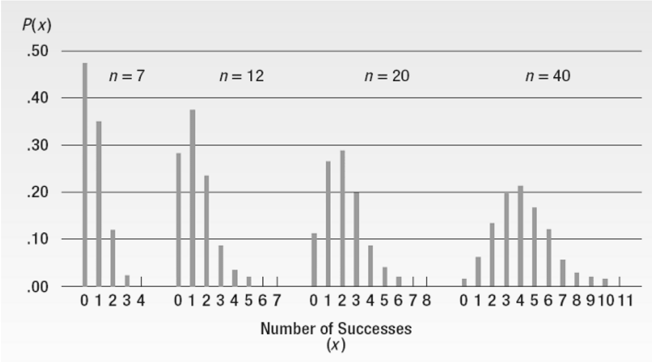
Binomial – Shapes for Varying π (n constant)

CHART 6–2 Graphing the Binomial Probability Distribution for a π of .05, .10, .20, .50, and .70 and an n of 10



36

Binomial – Shapes for Varying *n* (π constant)



37

Cumulative Distribution Function (累積分配函數)

- 假設 X 是一隨機變數，其機率分配函數通常以函數形式表示，以小寫的 f(x) 表示，

$$f(X = x) = P(X = x)$$

- 以大寫的 F(x) 表示累積機率函數，F(x) 表示 X 發生小於 x 的機率，定義如下：

$$F(x) = P(X \leqq x)$$

38

Cumulative Binomial Probability Distributions

A study in June 2003 by the Illinois Department of Transportation concluded that 76.2 percent of front seat occupants used seat belts. A sample of 12 vehicles is selected. What is the probability the front seat occupants in at least 7 of the 12 vehicles are wearing seat belts?

$$P(X \geq 7 | n = 12 \text{ and } \pi = .762)$$
$$= P(X=7) + P(X=8) + P(X=9) + P(X=10) + P(X=11) + P(X=12)$$
$$= .0902 + .1805 + .2569 + .2467 + .1436 + .0383$$
$$= .9563$$

$$P(x = 7 | n = 12 \text{ and } \pi = .762) = {}_{12}C_7(.762)^7(1 - .762)^{12-7} = 792(.149171)(.000764) = .0902$$

39

Finite Population

A finite population is a population consisting of a fixed number of known individuals, objects, or measurements. Examples include:

- The number of students in this class.
- The number of cars in the parking lot.
- The number of homes built in Taipei.

40

Characteristics of Hypergeometric Distribution (超幾何分配)

- 假設 50 個燈泡只有 20 個品質合格。試問任意抽取 20 個檢驗，有 X 個合格的機率為何？
- 隨機變數 X 表示抽到的 20 個中，品質合格的個數。X 的機率分配，即是超幾合分配。

41

Characteristics of Hypergeometric Distribution (超幾何分配)

- There are only 2 possible outcomes.
- The sample is selected from a finite population without replacement.
- The random variable is the number of successes in a fixed number of trials.
- The probability of a success is not the same on each trial.
- The trials are not independent.
- The size of the sample n is greater than 5% of the size of the population N *(i.e. $n/N \geq .05$).*

42

Hypergeometric Distribution

HYPERGEOMETRIC DISTRIBUTION $$P(x) = \frac{({}_SC_x)({}_{N-S}C_{n-x})}{{}_NC_n} \quad [6\text{–}6]$$

where:

N is the size of the population.
S is the number of successes in the population.
x is the number of successes in the sample. It may be 0, 1, 2, 3, ...
n is the size of the sample or the number of trials.
C is the symbol for a combination.

43

Example

- 假設 50 個燈泡只有 20 個品質合格。試問任意抽取 20 個檢驗，有 X 個合格的機率為何？

$$f(X) = \frac{C_x^{20} C_{20-x}^{30}}{C_{20}^{50}}, \quad x = 0,1,2,.....,20$$

- 抽到個合格的機率為：

$$f(2) = \frac{C_2^{20} C_{20-2}^{30}}{C_{20}^{50}} = \frac{C_2^{20} C_{18}^{30}}{C_{20}^{50}}$$

44

Hypergeometric Distribution - Example

PlayTime Toys, Inc., employs 50 people in the Assembly Department. Forty of the employees belong to a union and ten do not. Five employees are selected at random to form a committee to meet with management regarding shift starting times. What is the probability that four of the five selected for the committee belong to a union?

45

Hypergeometric Distribution - Example

N is 50, the number of employees.
S is 40, the number of union employees.
x is 4, the number of union employees selected.
n is 5, the number of employees selected.

$$P(4) = \frac{({}_{40}C_4)({}_{50-40}C_{5-4})}{{}_{50}C_5} = \frac{\left(\frac{40!}{4!36!}\right)\left(\frac{10!}{1!9!}\right)}{\frac{50!}{5!45!}} = \frac{(91{,}390)(10)}{2{,}118{,}760} = .431$$

46

Poisson Probability Distribution （卜瓦松/柏松 機率分配）

- 假設大台北地區，下午5點～7點，平均發生 μ 件交通事故。試問某日下午5點～7點之間，大台北地區發生X件車禍的機率為何？
- 一般以X服從 Poisson Probability Distribution 來描述這類事件的機率分配。

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}, \qquad x = 0,1,2,....,$$

其中 μ 表示所『探討時間單位』內，平均發生次數。

- 若改成下午 5 ~ 6 點間，發生 Y 次事故的機率，則 Y 的機率分配為：

$$f(y) = \frac{(\mu/2)^y e^{-\mu/2}}{y!}, \qquad y = 0,1,2,....$$

47

Poisson Probability Distribution （卜瓦松/柏松 機率分配）

The **Poisson probability distribution** describes the number of times some event occurs during a specified interval. The interval may be time, distance, area, or volume.

- Assumptions of the Poisson Distribution
  - (1) The probability is proportional to the length of the interval.
  - (2) The intervals are independent.

48

Poisson Probability Distribution

The Poisson distribution can be described mathematically using the formula:

POISSON DISTRIBUTION $$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$ [6–7]

where:

μ (mu) is the mean number of occurrences (successes) in a particular interval.
e is the constant 2.71828 (base of the Napierian logarithmic system).
x is the number of occurrences (successes).
P(*x*) is the probability for a specified value of *x*.

49

Poisson Probability Distribution - Example

Assume baggage is rarely lost by Northwest Airlines. Suppose a random sample of 1,000 flights shows a total of 300 bags were lost. Thus, the arithmetic mean number of lost bags per flight is 0.3 (300/1,000). If the number of lost bags per flight follows a Poisson distribution with *u* = 0.3, find the probability of not losing any bags for a flight.

$$P(0) = \frac{\mu^x e^{-u}}{x!} = \frac{0.3^0 e^{-.3}}{0!} = .7408$$

50

Example

- 交通尖峰時間下午五點到七點，全台北市平均事故 180 次。假設發生機率服從 Poisson Dist. 。請問，在任一各三分鐘的區間內，發生三次事故的機率為何？
- u = 180次/120分 = 1.5次/分 =4.5次/3分
-

Poisson Distribution with u = 4.5

$$P(X=3)=\frac{u^x e^{-u}}{x!} = \frac{4.5^3 e^{-4.5}}{3!} = 0.1687$$

51

Exercises

- 1,5,9,11,13,15,19,25,27,31,33,41,43,45,47, 49,51

52

End of Chapter 6

53