統計學（一）

第四章 離散型隨機變數
(Discrete Random Variables)

授課教師：唐麗英教授

國立交通大學 工業工程與管理學系
聯絡電話：(03)5731896
e-mail：litong@cc.nctu.edu.tw

2013

本課程內容參考書目

- **教科書**
  - Mendenhall, W., & Sincich, T. (2007). *Statistics for engineering and the sciences*, 5th Edition. Prentice Hall.
- **參考書目**
  - Berenson, M. L., Levine, D. M., & Krehbiel, T. C. (2009). *Basic business statistics: Concepts and applications*, 11th Edition. Upper Saddle River, N.J: Pearson Prentice Hall.
  - Larson, H. J. (1982). *Introduction to probability theory and statistical inference*, 3rd Edition. New York: Wiley.
  - Miller, I., Freund, J. E., & Johnson, R. A. (2000). *Miller and Freund's Probability and statistics for engineers*, 6th Edition. Upper Saddle River, NJ: Prentice Hall.
  - Montgomery, D. C., & Runger, G. C. (2011). *Applied statistics and probability for engineers*, 5th Edition. Hoboken, NJ: Wiley.
  - Watson, C. J. (1997). *Statistics for management and economics*, 5th Edition. Englewood Cliffs, N.J: Prentice Hall.
  - 林惠玲、陳正倉（2009），「統計學：方法與應用」，第四版，雙葉書廊有限公司。
  - 唐麗英、王春和（2013），「從範例學MINITAB統計分析與應用」，博碩文化公司。
  - 唐麗英、王春和（2008），「SPSS 統計分析 14.0 中文版」，儒林圖書公司。
  - 唐麗英、王春和（2007），「Excel 2007統計分析」，第二版，儒林圖書公司。
  - 唐麗英、王春和（2005），「STATISTICA6.0與基礎統計分析」，儒林圖書公司。
  - 陳順宇（2004），「統計學」，第四版，華泰書局。
  - 彭昭英、唐麗英（2010），「SAS123」，第七版，儒林圖書公司。

離散型機率分佈

- **常用的離散型機率分佈**
  - 白努力分佈(Bernoulli Probability Distribution)
  - 二項分佈(Binomial Probability Distribution)
  - 超幾何分佈(Hypergeometric Probability Distribution)
  - 波瓦松分佈(Poisson Probability Distribution)
  - 負二項分佈(Negative Binomial Probability Distribution)
  - 幾何分佈(Geometric Probability Distribution)

二項分佈
(Binomial Probability Distribution)

二項分佈 (Binomial Probability Distribution)

- **何謂二項實驗？**

  一個實驗必須滿足以下四個條件，才能稱為二項實驗：

  1) 此一實驗獨立、重複的試行n次。
  2) 每一試行均產生兩個結果：成功(Success)或失敗(Failure)。
  3) 每一試行成功的機率均為 p，失敗的機率為(1-p)或 q。
  4) 我們對試行n次中，成功 X 次之機率有興趣。

二項分佈 (Binomial Probability Distribution)

- **二項機率分佈**
  - 在n次獨立的二項實驗試行中，出現x次成功的機率為

$$\boxed{p(x)=C(n,x)p^{x}q^{n-x} \qquad x=0,1,2,\ldots,n}$$

其中

- n 表全部的試行數
- x 表在n次試行中成功的次數；
- C(n, x) 表n次試行中取 x 次成功次數的組合數；
- p 表每一試行成功的機率；
- q=1-p 表每一試行失敗的機率。

二項分佈 (Binomial Probability Distribution)

- **二項隨機變數的平均數與變異數**
  - 平均數 $E(x) = \mu_x = np$
  - 變異數 $Var(x) = \sigma_x^2 = npq$

- **白努力(Bernoulli) 隨機變數**

1) 當X 服從(n=1，p) 之二項分佈，則 X 稱為 白努力(Bernoulli) 隨機變數。
2) 一個服從二項分佈（n, p）之隨機變數Y是n個白努力隨機變數之和

二項分佈 (Binomial Probability Distribution)

- **例1：某製成品中約有10%為不良品，今任取10個成品檢查，求其中含有**
  a) 兩個不良品數之機率？
  b) 少於兩個不良品數之機率？

二項分佈 (Binomial Probability Distribution)

- 例2：設某考試有20題五選一之單選題，若能答對至少12題，就算通過。假設你每一題均隨機猜答案，請問你通過該考試的機會有多大？

超幾何分佈
(Hypergeometric Probability Distribution)

超幾何分佈 (Hypergeometric Probability Distribution)

- 超幾何隨機變數
  - The experiment consists of randomly drawing n elements **without replacement**（取後不放回）from a set of ***N*** elements, a of which are S's (for Success) and ***(N-a)*** of which are F's (for Failure).
  - The hypergeometric random variable **X is the number of S's in the draw of *n* elements**.

超幾何分佈 (Hypergeometric Probability Distribution)

- **超幾何機率分佈**
  - 超幾何隨機變數X之機率分佈如下：

$$p(x)=\frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} \qquad x=0,1,2,\ldots,a$$

  - 其中
    - N = 群體總數(total number of elements)
    - a = 群體中成功的元素個數(Number of S's in the N elements)
    - n = 從群體中抽取n個元素(Number of elements drawn)
    - x = 抽取n個元素中成功的個數(Number of S's drawn in the n elements)

超幾何分佈 (Hypergeometric Probability Distribution)

- **超幾何隨機變數X的平均數與變異數**
  - 平均數：$E(x)=\mu_x=\dfrac{na}{N}$
  - 變異數：$Var(x)=\sigma_x^2=\dfrac{na(N-a)(N-n)}{N^2(N-1)}$

超幾何分佈 (Hypergeometric Probability Distribution)

- **例1**

Suppose an employer randomly selects <u>three</u> new employees from a <u>total of ten</u> applications, <u>six men</u> and <u>four women</u>. Let X be the number of women who are hired.

1) Give the probability function of X.
2) Find the probability that no women are hired.
3) Find the probability that women outnumber men in the selection.
4) Compute E(X) and Var(X)

- 例2

A shipment of 100 tape recorders contains 25 that are defective. If 10 of them are randomly chosen for inspection, find the probability that 2 of the 10 will be defective by using

1) the formula for the hypergeometric distribution;

2) the formula for the binomial distribution as an approximation.

超幾何分佈 (Hypergeometric Probability Distribution)

- **以二項分佈近似超幾何分佈**
  - In general, it can be shown that the hypergeometric distribution approaches binomial distribution with **$p = a/N$ when $N \to \infty$**
  - A good rule of thumb is to use the binomial distribution as an approximation to the hypergeometric distribution if $\mathbf{n} \le \frac{\boldsymbol{N}}{\mathbf{10}}$.

波瓦松分佈
(Poisson Probability Distribution)

波瓦松分佈 (Poisson Probability Distribution)

- 波瓦松分佈是用來形容在某一特定時間或面積內稀有事件發生之機率。

- 波瓦松隨機變數的一些例子：
  1) 幾週內保險公司收到的要保信數
  2) 幾分鐘內經過剪票口的旅客數
  3) 一段短時間內經轉接的電話次數
  4) 一段時間內地震發生次數

波瓦松分佈 (Poisson Probability Distribution)

- 波瓦松機率分佈
  - 假設事件是隨機且彼此獨立的發生，單位時間的平均次數為 μ，而 X 表示一段時間事件發生的次數，則波瓦松機率密度函數如下：

$$\boxed{P(x)=\frac{(\lambda t)^{x}e^{-\lambda t}}{x!}=\frac{\mu^{x}e^{-\mu}}{x!} \qquad for \quad x=0,1,2,\ldots}$$

其中

- μ = 波瓦松分佈事件在某一特定時間(或面積)內發生的平均數
- λ = 單位時間(或面積)內發生的平均數
- t = 特定之時間(或面積)
- e = 2.718

波瓦松分佈 (Poisson Probability Distribution)

- **波瓦松何隨機變數的平均數與變異數**
  - 平均數 $E(x)=\mu=\lambda t$
  - 變異數 $Var(x)=\sigma^2=\lambda t$

波瓦松分佈 (Poisson Probability Distribution)

利用機率表尋找波瓦松機率 (Cumulative Poisson Probabilities)

$$P(X \le k) = P(X=0) + P(X=1) + \cdots + P(X=k) = \sum_{x=0}^{k} \frac{\mu^x e^{-\mu}}{x!}$$

TABLE 4 in Appendix B

k	μ 0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
0	0.6065	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.9098	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9856	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9982	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9998	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	1.0000	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	1.0000	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8	1.0000	1.0000	1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9	1.0000	1.0000	1.0000	1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992	0.9980
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9993
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

波瓦松分佈 (Poisson Probability Distribution)

- **以波瓦松分佈近似二項分佈**
  - The **Poisson** Probability distribution provides good approximations to the Binomial Probability **when n is large and p is small (preferable with $np \leq 7$).**

波瓦松分佈 (Poisson Probability Distribution)

- **例1：ABC公司生產的平板玻璃窗內的氣泡數目為一波瓦松分佈，平均每平方呎有0.004個氣泡，試求**
  a) 所生產的玻璃窗內無氣泡的機率
  b) 所生產的玻璃窗內氣泡數不超過1個的機率

波瓦松分佈 (Poisson Probability Distribution)

- **例2**

The number of customers arriving at a teller's window at Bay Bank is Poisson distributed with a mean rate of .75 person per minute.

a) What is the probability that two customers will arrive in the next 6 minutes?

b) Use table to find the answer in (a)

波瓦松分佈 (Poisson Probability Distribution)

- **例3**

Suppose that a large food-processing and canning plant has 20 automatic canning machines in operation at all times. If the probability that an individual canning machine breaks down during a given day is .05, find the probability that during a given day 2 canning machines fail.

a) Use the **binomial** distribution to compute the exact probability. (Use table in the Appendix.)

b) Compute the **Poisson approximation**.

c) Compare the answer obtained in a) and b).

負二項分佈
(Negative Binomial Probability Distribution)

- **何謂負二項實驗？**

一個實驗必須滿足下列個條件，才能稱為負二項實驗。

1) 某一實驗獨立、重複的試行 y 次
2) 每一試行均產生兩結果：成功(Success)或失敗(Failure)
3) 每一試行成功的機率均為 p，失敗的機率為(1-p)或 q
4) 我們對出現第 r 次成功所經歷的試行次數 y 有興趣

- **負二項機率分佈**
  - 負二項隨機變數Y之機率分佈如下：

$$p(y)=\binom{y-1}{r-1}p^{r}q^{y-r} \quad (y=r,r+1,r+2,\cdots)$$

其中

- p = Probability of success on a single Bernoulli trail
- q = 1- p
- y = Number of trials until the ***r***th success is observed

- **負二項隨機變數的平均數與變異數**
  - 平均數 $E(x) = \mu = \dfrac{r}{p}$
  - 變異數 $Var(x) = \sigma^2 = \dfrac{rq}{p^2}$

幾何分佈
(Geometric Probability Distribution)

幾何分佈 (Geometric Probability Distribution)

- 幾何分佈是負二項分佈的一個特例
  - The geometric distribution is **a special case of the negative binomial distribution**. It deals with the number of trials required for a single success. Thus, the geometric distribution is negative binomial distribution where **the number of successes *(r)* is equal to 1**.

幾何分佈 (Geometric Probability Distribution)

- 幾何機率分佈
  - For the special case ***r = 1***, the probability distribution of ***Y*** is Known as a geometric probability distribution. (Y表示出現第1次成功所經歷的試驗次數)

$$\boxed{p(y) = pq^{y-1} \quad (y = 1, 2, \cdots)}$$

其中

- p = Probability of success on a single Bernoulli trail
- q = 1- p
- y = Number of trials until the ***first*** success is observed

幾何分佈 (Geometric Probability Distribution)

- 幾何隨機變數的平均數與變異數
  - 平均數 $E(x)=\mu=\dfrac{1}{p}$
  - 變異數 $Var(x)=\sigma^2=\dfrac{q}{p^2}$

負二項與幾何分佈

- **例1 : To attach the housing on a motor, a production line assembler must use an electrical hand tool to set and tighten four bolts. Suppose that the probability of setting and tightening a bolt in any 1-second time interval is p =.8. If the assembler fails in the first second, the probability of success during the second 1-second interval is .8 and so on.**

a) Find the probability distribution of Y, the length of time until a complete housing is attached.

b) Find p(6).

c) Find the mean and variance of Y.

負二項與幾何分佈

- **解：**

a) Since the housing contains r=4 bolts, we will use the formula for the negative binomial probability distribution. Substituting p = .8 and r=4 into the formula for p(y), we obtain

$$p(y)=\binom{y-1}{r-1}p^r q^{y-r}=\binom{y-1}{3}(.8)^4(.2)^{y-4}$$

b) To find the probability that the complete assembly operation will require Y=6 seconds, we substitute y=6 into the formula obtained in part a and find

$$p(y)=\binom{5}{3}(.8)^4(.2)^2=(10)(.4096)(.04)=.16384$$

負二項與幾何分佈

c) For this negative binomial distribution,

$$\mu = \frac{r}{p} = \frac{4}{.8} = 5 \text{ seconds}$$

and

$$\sigma^2 = \frac{rq}{p^2} = \frac{4(.2)}{(.8)^2} = 1.25$$

負二項與幾何分佈

- **例2 : A manufacturer uses electrical fuses in an electronic system. The fuses are purchased in large lots and tested sequentially until the <u>first</u> defective fuse is observed. Assume that the lot contains <u>10%</u> defective fuses.**

  a) What is the probability that the first defective fuse will be one of the first five fuses tested?

  b) Find the mean, variance and standard deviation for Y, the number of fuses tested until the first defective fuse is observed.

負二項與幾何分佈

- **解：**

a) The number Y of fuses tested until the first defective fuse is observed is a geometric random variable with

p = .1 (probability that a single fuse is defective)

q = 1 – p =.9

and

$$p(y) = pq^{y-1} = (.1)(.9)^{y-1} \quad (y = 1, 2, \cdots)$$

The probability that the first defective fuse is one of the first five fuses tested is

$$p(Y \le 5) = p(1) + p(2) + \cdots + p(5)$$

$$= (.1)(.9)^0 + (.1)(.9)^1 + \cdots + (.1)(.9)^4 = .41$$

負二項與幾何分佈

b) The mean, variance and standard deviation of this geometric random variable are:

$$\mu = \frac{1}{p} = \frac{1}{.1} = 10$$

$$\sigma^2 = \frac{q}{p^2} = \frac{.9}{(.1)^2} = 90$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{90} = 9.49$$

本單元結束