

統計學（一）

第五章 連續型機率分佈 (Continuous Probability Distributions)

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

☆ 本講義未經同意請勿自行翻印 ☆

本課程內容參考書目

• 教科書

- P. Newbold, W. L. Carlson and B. Thorne(2007). *Statistics for Business and the Economics*, 7th Edition, Pearson.

• 參考書目

- Berenson, M. L., Levine, D. M., and Krehbiel, T. C. (2009). *Basic business statistics: Concepts and applications*, 11th Edition Prentice Hall.
- Larson, H. J. (1982). *Introduction to probability theory and statistical inference*, 3rd Edition, New York: Wiley.
- Miller, I., Freund, J. E., and Johnson, R. A. (2000). *Miller and Freund's Probability and statistics for engineers*, 6th Edition, Prentice Hall.
- Montgomery, D. C., and Runger, G. C. (2011). *Applied statistics and probability for engineers*, 5th Edition, Wiley.
- Watson, C. J. (1997). *Statistics for management and economics*, 5th Edition. Prentice Hall.
- 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」，儒林圖書公司。
- 唐麗英、王春和（2007），「Excel 統計分析」，第二版，儒林圖書公司。
- 唐麗英、王春和（2005），「STATISTICA與基礎統計分析」，儒林圖書公司。

- **Continuous Random Variable**
 - A **continuous R.V.** is a R.V. that can take on a **continuum** of values rather than a countably infinite number.

Continuous Random Variables

- 例 1:

- 1) The time required to read the book “How to lie with Statistics”—*continuous R.V.*
- 2) The number of women in a jury of 12.
- 3) The speed of a passing car. —*continuous R.V.*
- 4) The number of heads observed when flip a coin two times

- Cumulative Probability Function, c.d.f.

(累加機率函數)

- The Cumulative Probability Function, $F(x_0)$, for a random variable X is given as:

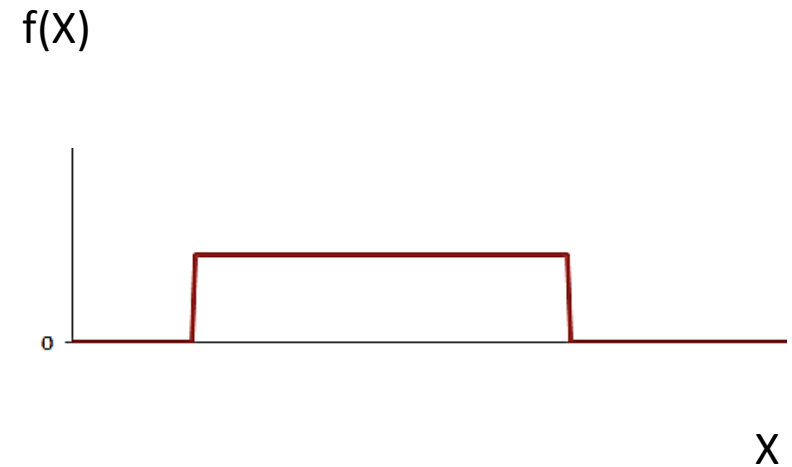
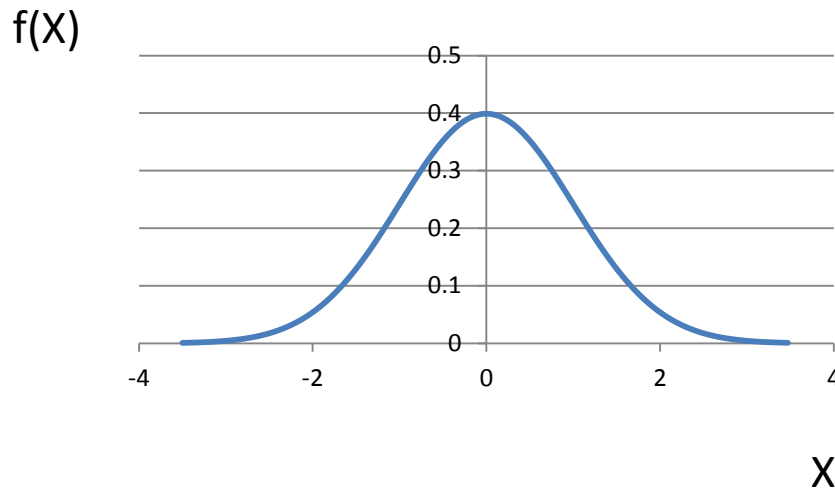
$$F_X(t) = P(X \leq t) \quad \text{for} \quad -\infty \leq t \leq \infty$$

- **Remark:** If X is a continuous R.V., then $F_X(t) = \int_{-\infty}^t f(x)dx$.

(i.e. $F_X(t)$ 是一累加機率函數)

- Cumulative Probability Function, c.d.f.

— 常用圖形表示



Probability Distributions for Continuous R.V.

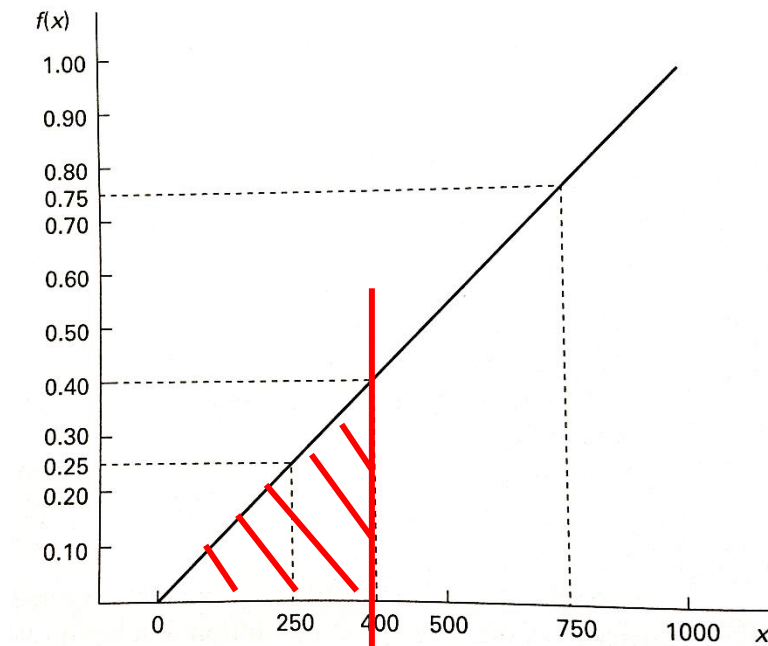
- 例 2:

Consider a gasoline station that has a 1,000-gallon storage tank that is filled each morning at the start of the business day. The random variable X indicates the gasoline sales in gallons for a particular day. We are concerned with the probability of various levels of daily gasoline sales, where the probability of a specific number of gallons sold is **the same** over the range **from 0 to 1,000**. What is the probability that sell **less than** 400 gallons?

Probability Distributions for Continuous R.V.

- [Ans]**

$$F(X) = \begin{cases} 0 & \text{if } x < 0 \\ 0.001x & \text{if } 0 \leq x \leq 1000 \\ 1 & \text{if } x > 1000 \end{cases}$$



$$P(X < 400) = 0.001 * 400 = 0.4$$

Continuous Random Variables

- **The Probability Density Function (p.d.f.) of a Continuous R.V.(連續型隨機變數之機率密度函數)**

- Let X be a continuous random variable with distribution function, $F(x) = P(X \leq x)$.

The probability density function for X is

$$f(x) = \frac{dF(x)}{dx} = F'(x)$$

The range for a continuous R.V. X is $R_x = \{ x | f(x) \geq 0 \}$.

- **The properties of the probability density function, $f(x)$**

1) $f(x) \geq 0$

2) $\int_{-\infty}^{\infty} f(x)dx = 1$

3) If X is a continuous R.V. with a density function $f(x)$, then for any $a < b$ the probability that X falls in the interval (a, b) is the area under the density function between a and b :

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

4) $F(x_0) = \int_{x_m}^{x_0} f(x)dx$, where x_m is the minimum value of the random variable X .

Continuous Random Variables

– Remark:

If X is a **continuous** R.V., then the probability that X takes on any particular value is 0:

$$\mathbf{P(X=t) = 0}$$

If X is a continuous R.V., then

$$\mathbf{P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)}$$

– Note : this is not true for a **discrete** R.V. !

Expectations for Continuous Random Variables

- **The Expected Value of a Continuous Random Variable**
 - If X is a continuous R.V. with the density function $f(x)$, the **Expected** Value of X , is

$$E(X) = \mu_X = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

– **Remark :**

$E(X)$ is a **weighted average** of all possible value of X with each value weighted by it associated probability.

- **The Variance and St.D. of continuous R.V. X**

- $\text{Var}(X) = \sigma_X^2 = E[(X - \mu_X)^2] = E(X^2) - \mu_X^2$

- $\text{St. D.}(X) = \sigma_X = \sqrt{\sigma_X^2}$

Expectations for Continuous Random Variables

- 例 1:

A homeowner estimates that within the range of likely temperatures his January heating bill, Y , in dollars, will be

$$Y = 290 - 5T$$

Where T is the average temperature can be represented by a random variable with a mean of 24 and a standard deviation of 4, find the mean and standard deviation of this homeowner's January heating bill.

Expectations for Continuous Random Variables

- 例 2:

Suppose $E(X)=5$, $\text{Var}(X)=10$, Find

(a) $E(3X-5)$

(b) $\text{Var}(3X-5)$

Expectations for Continuous Random Variables

- 例 3:

Let X be a continuous R.V. with the density function

$$f(x) = 2x, \quad 0 < x < 1$$

Find a) $E(X)$

b) $E(X^2)$, $\text{Var}(X)$ and $\text{St.D.}(X)$

- **Some Standard Probability Distributions**

The following are three useful continuous probability distributions:

- Uniform Probability Distribution (齊一分佈)
- Normal Probability Distribution (常態分佈)
- Exponential Probability Distribution (指數分佈)

齊一分佈 (Uniform Probability Distribution)

- **Uniform Probability Distribution**

- The continuous random variable X is called a **uniform** random variable if and only if X is uniformly distributed over the interval (α, β) , i.e., the density for X is

$$f(x) = \frac{1}{\beta - \alpha}, \alpha < X < \beta$$
$$= 0, \text{ otherwise}$$

- **The Mean and Variance for the Uniform R.V.**

- $E(X) = \frac{\alpha + \beta}{2}$

- $\text{Var}(X) = \frac{1}{12} (\beta - \alpha)^2$

Uniform Probability Distribution (齊一分佈)

- 例 1:

If X is **uniformly** distributed over $(0, 10)$, calculate the probability that

a) $X < 3$

b) $X > 6$

c) $3 < X < 8$.

Uniform Probability Distribution (齊一分佈)

- 例 2:

Buses arrive at a specified stop at 15-minute intervals starting at 7, 7:15, 7:30, 7:45, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits

- a) **less than** 5 minutes for a bus.
- b) **more than** 10 minutes for a bus.

常態分佈

(Normal Probability Distribution)

- **Normal Probability Distribution**

- Many continuous random variables observed in nature possess a **bell-shaped**(鐘型) probability distribution. It is known as a **Normal** probability distribution (or **Gaussian distribution** after **Carl Friedrich Gauss**, who proposed it as a model for measurement errors. 量測誤差). The Normal distribution has been used as a model for such diverse phenomena as a person's **height**, **IQ score**, and the **velocity** of a gas molecule.

- **The Normal Probability Law**

- X is called a Normal random variable with the parameters μ and σ^2 if its density is given by

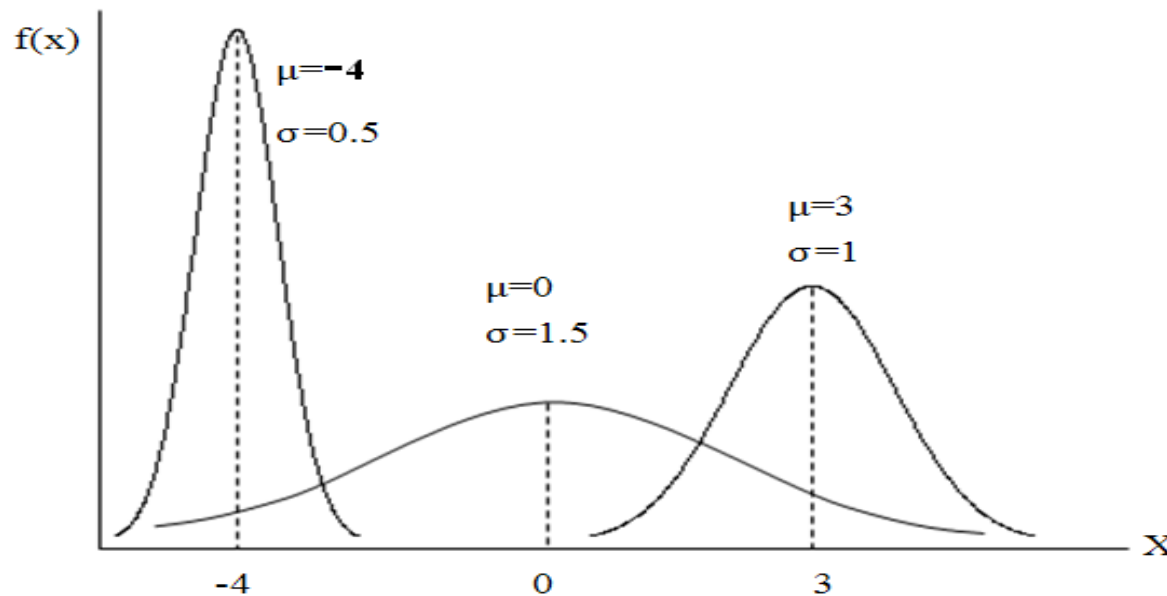
$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < X < \infty$$

- Where π = Mathematical constant approximated by 3.1416
 e = Mathematical constant approximated by 2.718
 μ = Population mean or the true mean
 σ^2 = Population variance
- It is denoted by $N(\mu, \sigma)$

- **Properties of a Normal Distribution(or Normal curve), $N(\mu, \sigma)$**
 - 1) Symmetric about μ .
 - 2) X varies from $-\infty$ to ∞ .
 - 3) Bell shape.
 - 4) The area under the curve is equal to 1.
 - 5) The mean, median and mode are identical.

Normal Probability Distribution(常態分佈)

- **Note:** Everytime we specify a particular combination of μ and σ , a different Normal distribution (or Normal curve) may be generated.



- **Note:** μ determines the **location**; σ determines the **shape**.

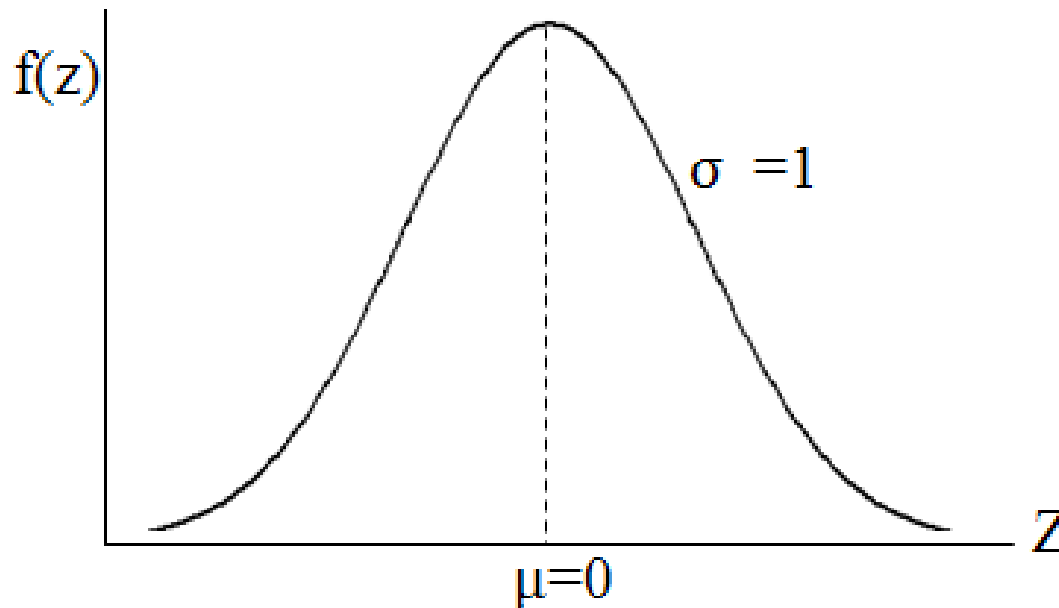
- **How μ and σ effect the normal curve?**

- 1) The smaller the variances (or standard deviations), the sharper the peak is.
- 2) If Normal distribution have the same variance but different means, then the normal curves have the same shape, but different locations .
- 3) μ — Location parameter (位置參數), σ —dispersion parameter (離散或變異參數)

Normal Probability Distribution(常態分佈)

- **The Standard Normal Distribution**

- The Normal distribution that has the mean 0 and variance 1 is called a *standard Normal distribution*. The standard Normal distribution is denoted by $N(0, 1)$



Normal Probability Distribution(常態分佈)

- Z is called the standard normal random variable if its density function is**

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

- The distribution function (or c.d.f.) of a Standard Normal R.V. Z is:**

$$F(z) = P(Z \leq z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

- **How to find the probability for a Standard Normal R.V. Z:**
 - Table 1 in Appendix (page 837) gives the results of the numerical integration of a Standard Normal R.V. Z.

Normal Probability Distribution(常態分佈)

- 例 1:

Suppose $Z \sim N(0, 1)$, find the following probabilities:

a) $P(-0.77 < Z < 1.44)$

b) $P(Z > -1.85)$

c) $P(Z < 1.85)$

d) $P(-0.77 < Z < 0)$

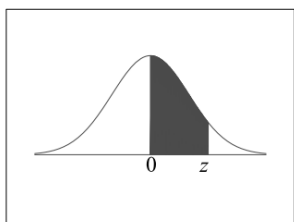
e) $P(Z > -0.36)$

f) $P(-3 < Z < 1.28)$

Note: $F(-C) = 1 - F(C)$

[illegible]

Standard Normal Distribution Table

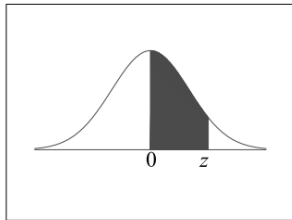


利用表求

$$P(-0.77 < Z < 1.44) =$$

[illegible]

Standard Normal Distribution Table

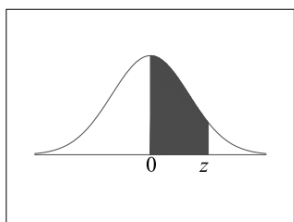


利用表求

$$P(Z > -1.85) =$$

[illegible]

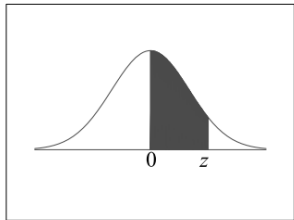
Standard Normal Distribution Table



利用表求

$$P(Z < 1.85) =$$
[illegible]

Standard Normal Distribution Table

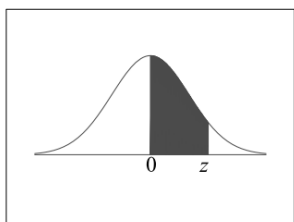


利用表求

$$P(-0.77 < Z < 0) =$$

[illegible]

Standard Normal Distribution Table



利用表求
 $P(Z > -0.36) =$

[illegible]

Normal Probability Distribution(常態分佈)

- 例 2:

Use Table 1 to find C for the following probabilities:

a) $P(Z < C) = 0.95$

b) $P(Z > C) = 0.7019$

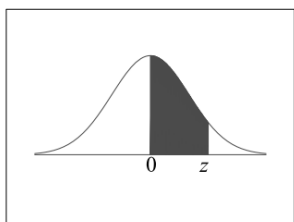
c) $P(Z > C) = 0.1379$

d) $P(Z < C) = 0.0110$

[illegible]

[illegible]

Standard Normal Distribution Table

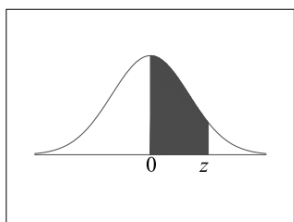


利用表求 C

$$P(Z > C) = 0.1379$$

[illegible]

Standard Normal Distribution Table



利用表求 C

$$P(Z < C) = 0.0110$$

[illegible]

- **How to Standardize the Normal Distribution**

如何標準化一般常態隨機變數？

- We may convert the general Normal distribution $N(\mu, \sigma)$ to a standard Normal distribution $N(0, 1)$ by using the transformation formula :

$$Z = \frac{X - \mu}{\sigma}$$

where $Z \sim N(0, 1)$ and $X \sim N(\mu, \sigma)$

Normal Probability Distribution(常態分佈)

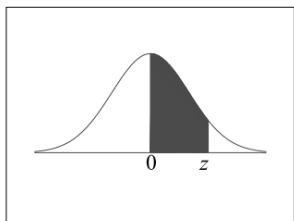
- 例 3:

Suppose $X \sim N(10, 2)$. Find the probability that

a) X lies between 11 and 13.6 ?

b) X is greater than 12 ?

Standard Normal Distribution Table

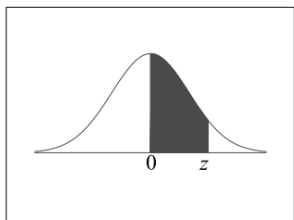


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998

$X \sim N(10, 2)$ ，利用右表
 求 $P(11 \leq X \leq 13.6) = ?$

$$\begin{aligned}
 &P(11 \leq X \leq 13.6) \\
 &= P\left(\frac{11-10}{2} \leq Z \leq \frac{13.6-10}{2}\right) \\
 &= P(0.5 \leq Z \leq 1.8)
 \end{aligned}$$

Standard Normal Distribution Table



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998

$X \sim N(10, 2)$ ，利用右表
求 $P(X > 12) = ?$

$$\begin{aligned}
 &P(X > 12) \\
 &= P\left(Z > \frac{12-10}{\sqrt{2}}\right) \\
 &= P(Z > 1) \\
 &= 1 - P(Z < 1)
 \end{aligned}$$

- **Application**

- **Recall:** If X is a continuous random variable, the probability that X takes some specific value, say 10, is 0. This is because there is **no area** under the probability density function over the point $X=10$.

Therefore,

$$P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = P(a < X < b),$$

since $P(X=a)=P(X=b)=0$.

Normal Probability Distribution(常態分佈)

- 例 4:

Studies show that gasoline usage for compact cars (小客車) sold in the United States is **normally** distributed, with a **mean** usage of 25 miles per gallon(mpg) and a **standard deviation** of 4.5 mpg, what percentage of compacts obtain 30 or more mile per gallon?

Normal Probability Distribution(常態分佈)

- 例 5:

Suppose that the Mid-term examination Scores for a Statistics class are **normally** distributed with **mean** 70 and **standard deviation** 10.

- a) What is the probability that a student has a score above 80?
- b) Find a score **S** such that 10% of students have scores **above S**.

Normal Probability Distribution(常態分佈)

- 例 6:

A client has an investment portfolio which is **normally** distributed, with the **mean** value \$1,000,000 and the **standard deviation** \$30,000. He has asked you to determine the probability that the value of his portfolio is between \$970,000 and \$1,060,000.

Normal Probability Distribution(常態分佈)

- [Ans]

$$Z_{970,000} = \frac{970,000 - 1,000,000}{30,000} = -1$$

$$Z_{1,060,000} = \frac{1,060,000 - 1,000,000}{30,000} = 2$$

$$\begin{aligned} P(970,000 \leq X \leq 1,060,000) &= P(-1 \leq Z \leq 2) \\ &= 1 - P(Z \leq -1) - P(Z \geq 2) \\ &= 1 - 0.1587 - 0.0228 \\ &= 0.8185 \end{aligned}$$

The probability for the indicated range is, thus, 0.8185.

Normal Probability Distribution(常態分佈)

- 例 7:

A company produces lightbulbs whose life follows a **normal** distribution, with a **mean** of 1,200 hours and a **standard deviation** of 250 hours. If we choose lightbulbs at random, what is the probability that its lifetime will **between 900** and **1,300** hours?

Normal Probability Distribution(常態分佈)

- [Ans]

$$\begin{aligned}P(900 < X < 1,300) &= P\left(\frac{900 - 1,200}{250} \leq Z \leq \frac{1,300 - 1,200}{250}\right) \\&= P(-1.2 < Z < 0.4) \\&= 0.6554 - (1 - 0.8849) \\&= 0.5403\end{aligned}$$

Hence, the probability is approximately 0.54 that a lightbulb will last between 900 and 1,300 hours.

Normal Probability Distribution(常態分佈)

- 例 8:

Whole Life Organic Inc. produces high quality organic frozen turkeys for distribution in organic food markets in the upper Midwest. The company has developed a range feeding program with organic grain supplements to produce their products. The **mean** weight of one of its frozen turkeys is 15 pounds with a **variance** of 16. Historical experience indicated that weights can be approximated by the **normal** probability distribution. Market research indicated that sales for frozen turkeys over 18 pounds are limited. What percentage of the company's turkey units will be **over 18** pounds?

Normal Probability Distribution(常態分佈)

- [Ans]

$$\begin{aligned}P(X > 18) &= P\left(Z > \frac{18 - 15}{4}\right) \\&= P(Z > 0.75) \\&= 1 - P(Z < 0.75) \\&= 1 - 0.7734 \\&= 0.2266\end{aligned}$$

Thus, Whole life can expect that 22.7% of its turkeys will weigh more than 18 pounds.

檢查數據是否呈常態分佈

檢查數據是否呈常態分佈

- 如何利用統計圖檢查數據是否呈常態分配？

1. 利用直方圖：

只要出現鐘形分佈圖形，即判定數據呈常態分佈

2. 利用常態機率圖：

只要圖形呈直線，即判定數據呈常態分佈

檢查數據是否呈常態分佈

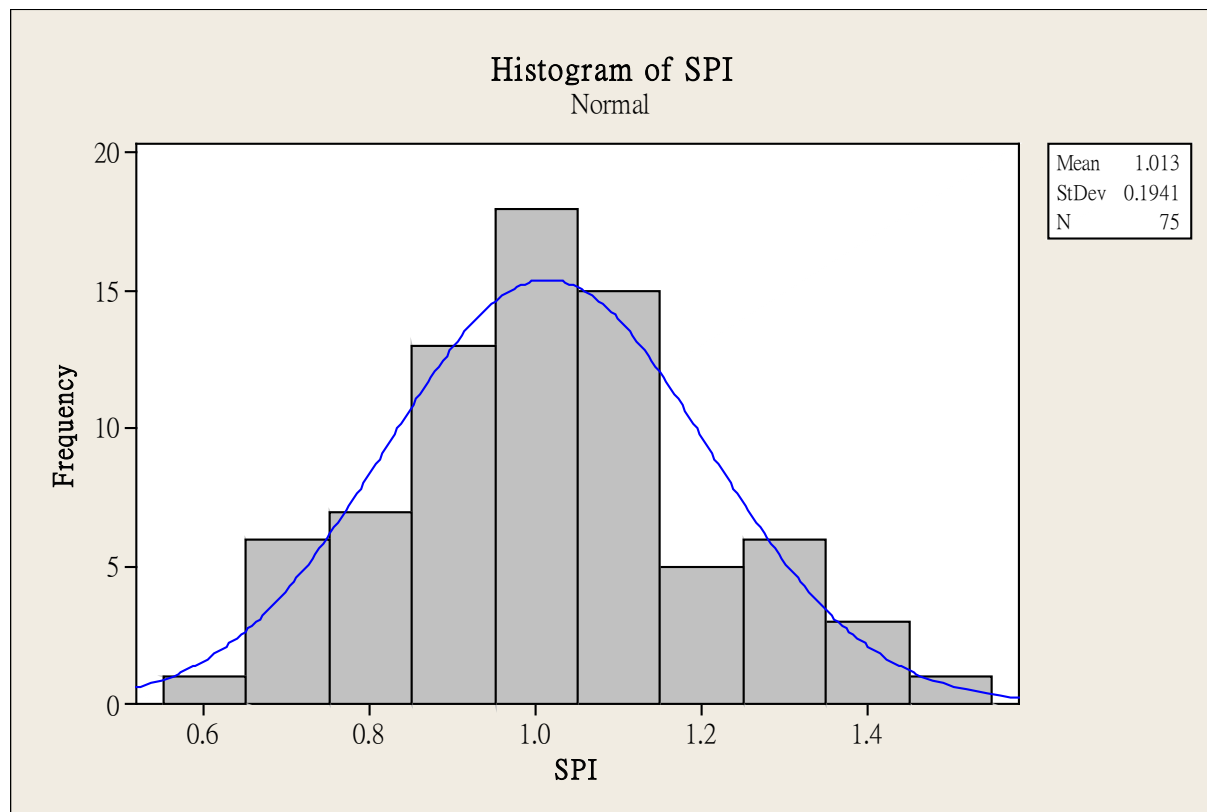
• 例 9:

下列數據為某模具上的孔徑尺寸值，請檢查數據是否呈常態分佈？

0.88	0.87	1.09	1.10	1.20
0.95	0.69	1.15	1.12	0.77
0.72	0.89	1.00	0.94	0.79
1.39	0.96	0.93	1.15	1.10
0.81	1.15	1.32	1.34	1.28
0.88	1.26	1.24	0.98	1.13
0.94	1.18	1.07	0.74	1.06
1.12	0.85	1.03	1.28	0.83
0.69	0.87	0.89	1.16	0.76
0.95	0.76	1.09	0.99	0.67
0.98	0.95	1.04	1.40	1.10
1.29	0.64	0.95	0.95	1.42
1.54	1.01	0.72	1.06	0.88
0.87	0.95	1.21	0.96	1.04
1.09	0.96	1.02	0.99	0.97

檢查數據是否呈常態分佈

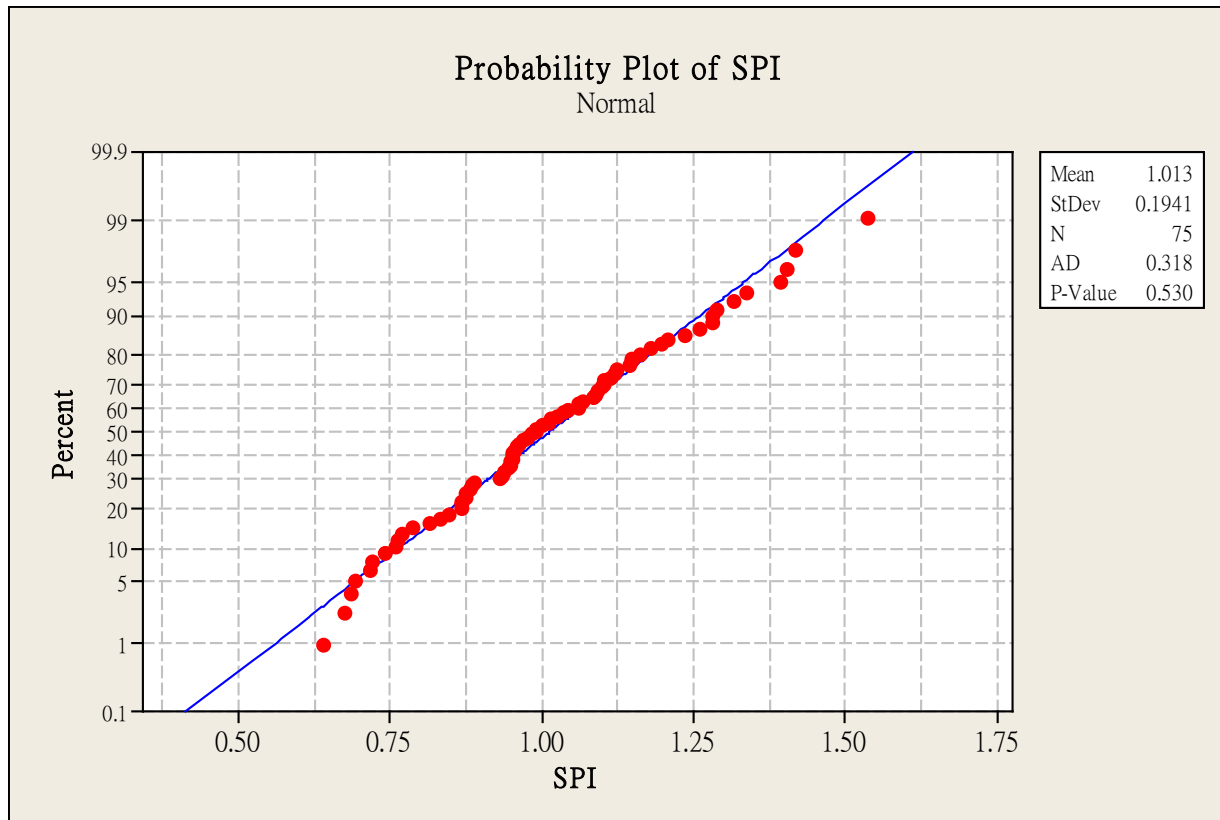
- 繪製直方圖



直方圖呈_____形分佈曲線，因此可判定數據_____常態分佈。

檢查數據是否呈常態分佈

- 利用常態機率圖

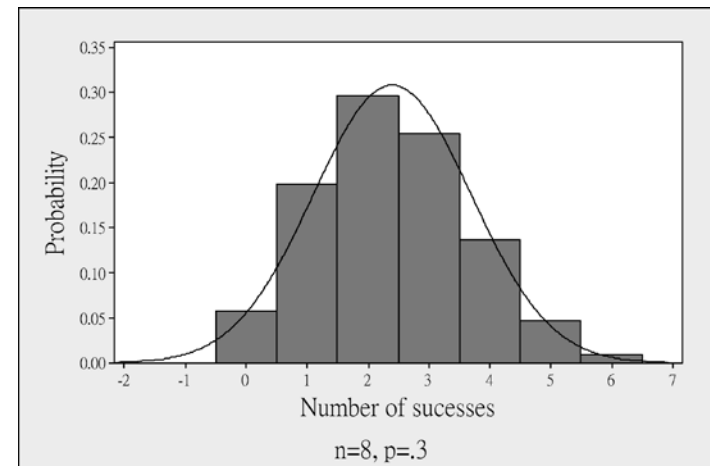
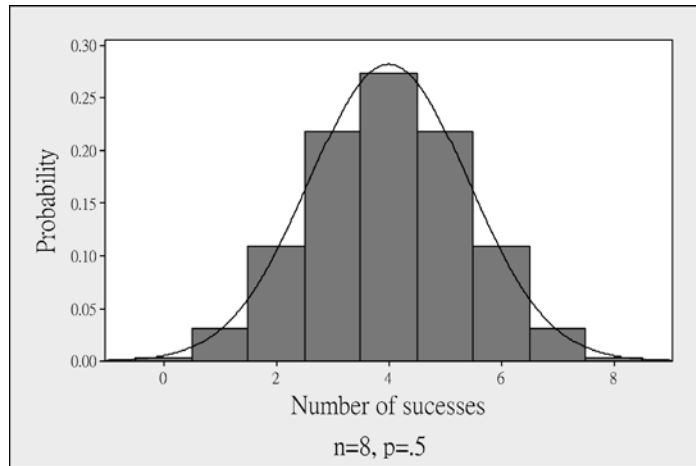
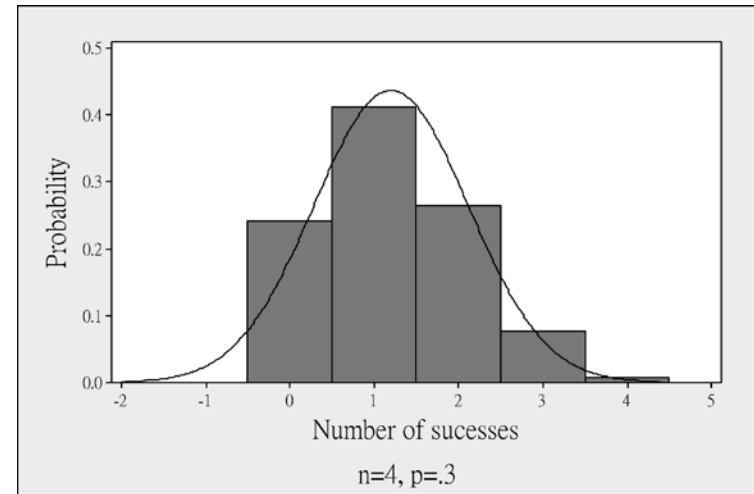
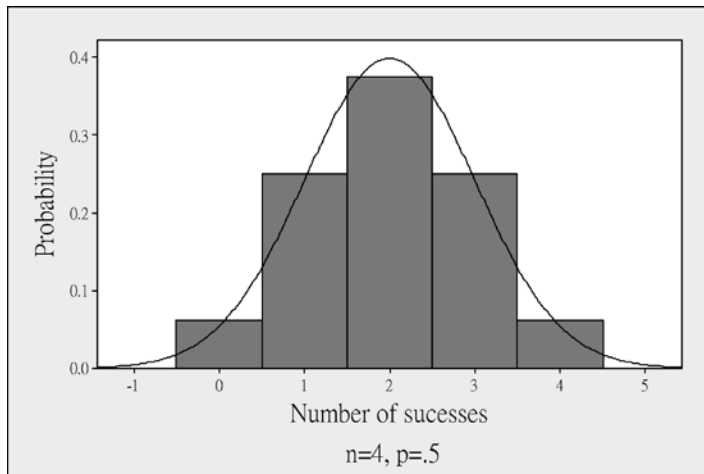


常態機率圖呈_____線，因此可判定數據_____常態分佈。

- **Normal Approximation for Binomial Distribution**

- As the sample size (or number of trials) n **increases** and the values of P are not close to 0 or 1, then the Binomial distribution becomes very close to a Normal distribution with mean $\mu = np$ and $\sigma = \sqrt{npq}$.

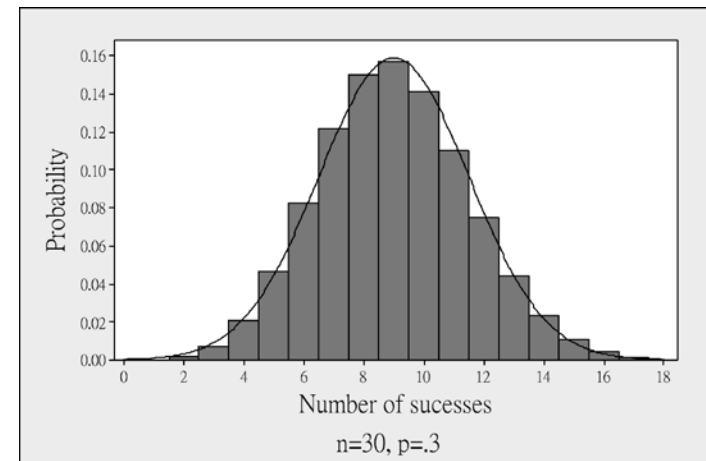
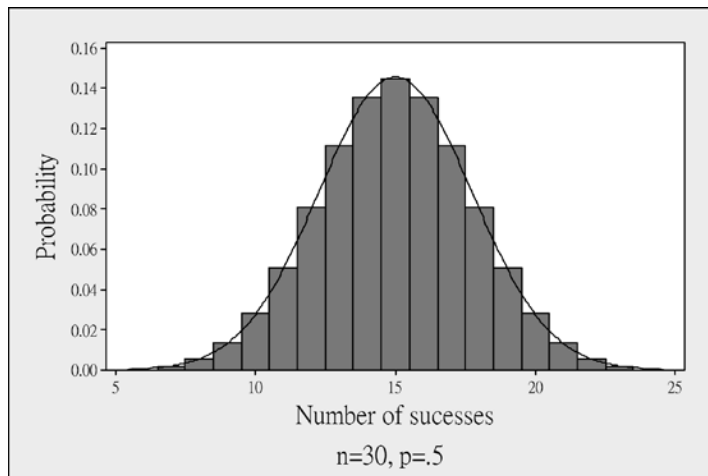
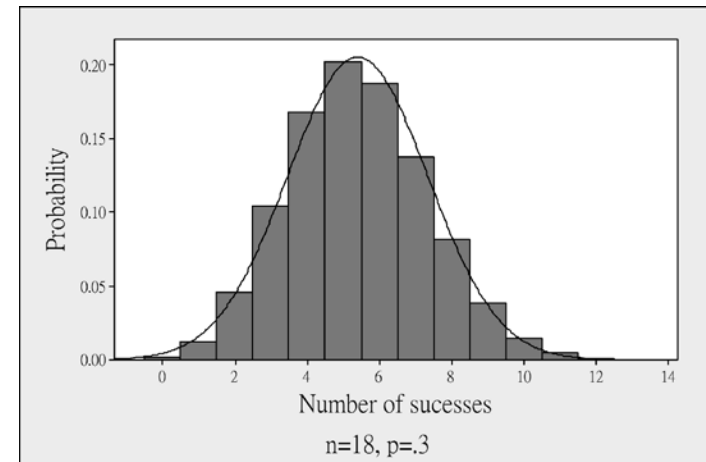
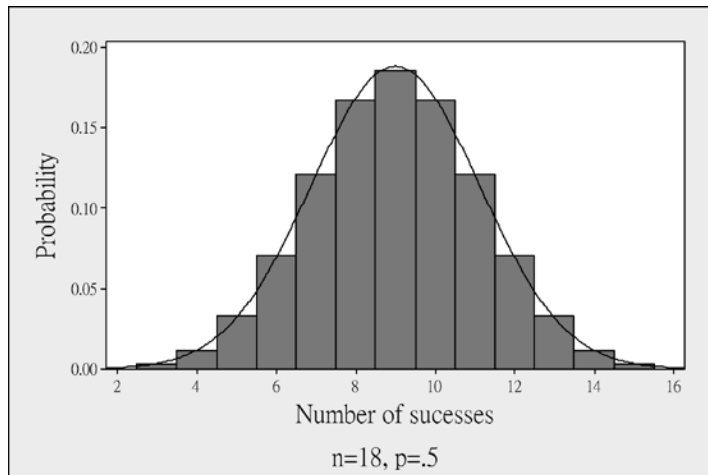
Normal Approximation for Binomial Distribution



Constant $p = .5$

Constant $p = .3$

Normal Approximation for Binomial Distribution



Constant $p = .5$

Constant $p = .3$

- **How to Use the Normal Probability to Approximate the Binomial Probability?**
 - Let X be the number of success in a Binomial Distribution with p not close to 0 or 1. If n is large, $X \sim N(\mu, \sigma)$, then we can use Normal to approximate the Binomial probability after the continuity correction to the random variable X .

- **Continuity Correction (連續性校正)**

- Subtract $1/2$ from the lower value on the X-scale, and add $1/2$ to the upper value. That is,

$P(a \leq X \leq b)$ pass to

$$P\left(a - \frac{1}{2} \leq X \leq b + \frac{1}{2}\right)$$

Normal Approximation for Binomial Distribution

- 例 1:

Apply the continuity correction to the following binomial R.V. X

a) $P(8 \leq X \leq 10) =$

b) $P(X \geq 21) =$

c) $P(X = 30) =$

d) $P(X \leq 15) =$

Normal Approximation for Binomial Distribution

- 例 2:

Flip a coin 100 times. What is the probability of getting

a) at least 40 “**heads**”?

b) exactly 50 “**heads**”?

Use normal to approximate the binomial.

Normal Approximation for Binomial Distribution

- 例 3:

Mary David makes the initial telephone contact with customers who have responded to an advertisement on her company's Web page in an effort to assess whether a follow-up visit to their homes is likely to be worthwhile. Her experience suggests that **40%** of the initial contacts lead to follow-up visits. IF she has 100 Web page contacts, what is the probability **between 45 and 50** home visits will result?

Normal Approximation for Binomial Distribution

- [Ans]

$$\begin{aligned}P(45 \leq X \leq 50) &\cong P\left(\frac{45 - (100)(0.4)}{\sqrt{100 * 0.4 * 0.6}} \leq Z \leq \frac{50 - (100)(0.4)}{\sqrt{100 * 0.4 * 0.6}}\right) \\&= P(1.02 \leq Z \leq 2.04) \\&= F(2.04) - F(1.02) \\&= 0.9793 - 0.8461 \\&= 0.1332\end{aligned}$$

Continuity Correction

$$\begin{aligned}P(45 \leq X \leq 50) &\cong P\left(\frac{44.5 - (100)(0.4)}{\sqrt{100 * 0.4 * 0.6}} \leq Z \leq \frac{50.5 - (100)(0.4)}{\sqrt{100 * 0.4 * 0.6}}\right) \\&= P(0.92 \leq Z \leq 2.14) \\&= F(2.14) - F(0.92) \\&= 0.9838 - 0.8212 \\&= 0.1626\end{aligned}$$

指數分佈 (Exponential Probability Distribution)

- **The Exponential Distribution**

- X is called an exponential random variable if and only if

$$f(x) = \frac{1}{\beta} e^{-\left(\frac{1}{\beta}\right)x}, \quad \text{for } X > 0 \text{ and } \beta > 0$$
$$= 0, \quad \text{otherwise}$$

Exponential Probability Distribution(指數分佈)

- **The Mean and Variance for the Exponential R.V.**
 - $E(X) = \beta$
 - $\text{Var}(X) = \beta^2$

Exponential Probability Distribution(指數分佈)

- 例 1:

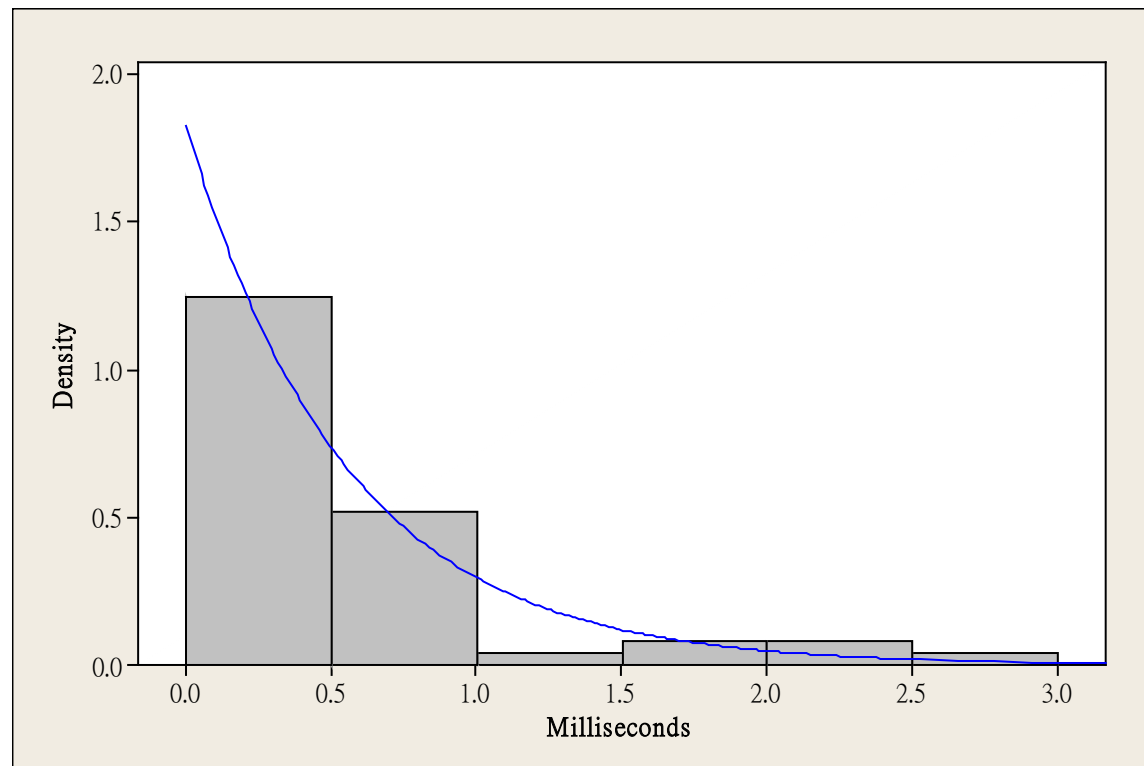
A nuclear engineer observing a reaction measures the time intervals between the emissions of beta particles.

0.894	0.991	0.061	0.186	0.311	0.817	2.267	0.091	0.139	0.083
0.235	0.424	0.216	0.579	0.429	0.612	0.143	0.055	0.752	0.188
0.071	0.159	0.082	1.653	2.010	0.158	0.527	1.033	2.863	0.365
0.459	0.431	0.092	0.830	1.718	0.099	0.162	0.076	0.107	0.278
0.100	0.919	0.900	0.093	0.041	0.712	0.994	0.149	0.866	0.054

Exponential Probability Distribution(指數分佈)

- 例 1:

These decay times (in milliseconds) are presented as a histogram in the following figure:



Exponential Probability Distribution(指數分佈)

- **Remark:** It can be shown that in connection with Poisson processes the **waiting time** between successive arrivals has an exponential distribution.
- More specifically, it can be shown that if in a **Poisson** process the mean arrival rate (average number of arrivals per unit time) is $\lambda = 1/\beta$, the time until the first arrival, or the waiting time between successive arrivals, has an **exponential** distribution with $1/\beta$.

Exponential Probability Distribution(指數分佈)

- 例 2:

If on the average **three trucks arrived per hour** to be unloaded at a warehouse, what are the probabilities that the time between the arrival of successive trucks will be

a) Less than 5 minutes;

b) At least 45 minutes;

What is the expected waiting time between successive arrivals?

Exponential Probability Distribution(指數分佈)

- 例 3:

Service times for customers at a library information desk can be modeled by an exponential distribution with a mean service time of **5 minutes**. What is the probability that a customer service time will take **longer than** 10 minutes?

- [Ans]

$$\begin{aligned} P(T > 10) &= 1 - P(T < 10) \\ &= 1 - (1 - e^{-(0.2)(10)}) \\ &= e^{-2} = 0.1353 \end{aligned}$$

Thus, the probability that a service time exceeds 10 minutes is 0.1353.

Exponential Probability Distribution(指數分佈)

- 例 4:

An industrial plant in Britain with 2,000 employees has a mean number of losttime accidents per week equal to $\lambda = 0.4$, and the number of accidents follows a Poisson distribution. What is the probability that the time between accidents is **less than** 2 weeks?

Exponential Probability Distribution(指數分佈)

- [Ans]

$$\begin{aligned}P(T < 2) &= F(2) \\&= (1 - e^{-(0.4)(2)}) \\&= 1 - e^{-0.8} \\&= 1 - 0.4493 = 0.5507\end{aligned}$$

Thus, the probability of less than 2 weeks between accidents is about 55%.

Jointly Distributed Random Variables

- **Jointly Cumulative Distribution Functions**

1. Let X_1, \dots, X_k be continuous R.V., then **their jointly cumulative distribution** is:

$$F(x_1, x_2, \dots, x_k) = P(X_1 < x_1 \text{ and } X_2 < x_2 \text{ and } \dots X_k < x_k)$$

2. $F(x_1), F(x_2), \dots, F(x_k)$ are called **marginal distribution functions**. 邊際累加函數

3. X_1, \dots, X_k are **independent** *if only if*
 $F(x_1, x_2, \dots, x_k) = F(x_1) F(x_2) \dots F(x_k).$

Independence and conditional Distributions

– Recall: Measures of Relationship between Variables

兩變數間之關聯性指標

– Two measures of association between two random variables 衡量兩變數間關聯性之指標有二：

1. Covariance (共變異數)

$$\begin{aligned}\mathbf{Cov(X, Y)} &= E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X \mu_Y \\ &= \sum_x \sum_y xy P(x, y) - \mu_x \mu_y\end{aligned}$$

2. Correlation (相關係數)

$$\rho = \frac{\mathbf{Cov(X, Y)}}{\sigma_X \sigma_Y}, \quad \text{provide that } \sigma_X < \infty \text{ and } \sigma_Y < \infty$$

Jointly Continuous Random Variables

- **Sum of random variables : K個隨機變數之和的平均數與變異數公式**

Let X_1, \dots, X_k be R.V. with means $\mu_1, \mu_2, \dots, \mu_k$, and variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$, then

$$E(X_1 + X_2 + \dots + X_k) = E(X_1) + E(X_2) + \dots + E(X_k) = \mu_1 + \mu_2 + \dots + \mu_k$$

$$\begin{aligned} \text{Var}(X_1 + X_2 + \dots + X_k) &= \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_k) + 2 \sum \sum \text{COV}(X_i, X_j) \\ &= \sigma_1^2 + \sigma_2^2 + \dots + \sigma_k^2 + 2 \sum \sum \text{COV}(X_i, X_j) \end{aligned}$$

Jointly Continuous Random Variables

兩個隨機變數之和的平均數與變異數公式

$$E(X_1 + X_2) = \mu_1 + \mu_2$$

$$\text{Var}(X_1 + X_2) = \sigma_1^2 + \sigma_2^2 + 2\text{Cov}(x_1, x_2)$$

兩個隨機變數之差的平均數與變異數公式：

$$E(X_1 - X_2) = \mu_1 - \mu_2$$

$$\text{Var}(X_1 - X_2) = \sigma_1^2 + \sigma_2^2 - 2\text{Cov}(x_1, x_2)$$

Note: If X_1 and X_2 are **uncorrelated** or **independent**, $\text{Cov}(x_1, x_2) = 0$

Jointly Continuous Random Variables

- 兩個隨機變數X與Y線性組合的平均數與變異數公式：

Let $W = aX + bY$, where a and b are constants, then

$$\mu_w = a\mu_x + b\mu_y$$

$$\begin{aligned}\sigma_w^2 &= a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab \operatorname{Cov}(X, Y) \\ &= a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab \rho \sigma_x \sigma_y\end{aligned}$$

Let $W = aX - bY$, where a and b are constants, then

$$\mu_w = a\mu_x - b\mu_y$$

$$\begin{aligned}\sigma_w^2 &= a^2\sigma_x^2 + b^2\sigma_y^2 - 2ab \operatorname{Cov}(X, Y) \\ &= a^2\sigma_x^2 + b^2\sigma_y^2 - 2ab \rho \sigma_x \sigma_y\end{aligned}$$

Jointly Continuous Random Variables

- 例 1:

A contractor is uncertain of the precise total costs for either materials or labor for a project. In addition, the total line of credit for financing the project is \$260,000, and the contractor wants to know the probability that total costs exceed \$260,000. It is believed that material costs can be represented by a normally distributed random variable with mean \$100,000 and standard deviation \$10,000. Labor cost are \$1,500 a day, and the number of days needed to complete the project can be represented by a normally distributed random variable with mean 80 and standard deviation 12. Assuming that material and labor costs are independent, what are the mean and standard deviation of the total project cost? In addition, what is the probability that the total project cost is greater than \$260,000?

Jointly Continuous Random Variables

- [Ans]

$$\mu_1 = 100,000, \sigma_1 = 10,000$$

$$\mu_2 = 1,500 * 80 = 120,000$$

$$\sigma_2 = 1,500 * 12 = 18,000$$

$$W = X_1 + X_2,$$

$$\mu_w = 100,000 + 120,000 = 220,000,$$

$$\sigma_w = \sqrt{10,000^2 + 18,000^2} = 20,591$$

$$Z = \frac{260,000 - 220,000}{20,591} = 1.94$$

We find that the probability that the total cost exceeds \$260,000 is 0.0262. Since this probability is small, the contractor has some confidence that the project can be completed within the available line of credit.

Jointly Continuous Random Variables

- 例 2:

Henry Chang has asked for your assistance in establishing a portfolio containing two stocks. Henry has \$1,000, which can be allocated in any proportion to two alternative stocks. The returns per dollar from these investments will be designated as random variable X and Y . Both of these random variables are independent and have the same mean and variance. Henry wishes to know the risk for various allocation options. When is the risk will be **minimized**?

Jointly Continuous Random Variables

- [Ans]

令 first investment = α , second = $1,000 - \alpha$

$$R = \alpha X + (1,000 - \alpha)Y$$

$$E(R) = \alpha * \mu + (1,000 - \alpha) * \mu = 1,000\mu$$

$$\begin{aligned} \text{Var}(R) &= \alpha^2 \sigma^2 + (1,000 - \alpha)^2 \sigma^2 \\ &= (2\alpha^2 - 2,000\alpha + 1,000,000) \sigma^2 \end{aligned}$$

→ $\alpha = \$500$, Var 最小

Jointly Continuous Random Variables

- 例 3:

Judy Chang, the account manager for Northern Securities, has a portfolio that includes 20 shares of Allied Information Systems and 30 shares of Bangalore Analytics. Both firms provide Web access devices that compete in the customer market. The price of Allied stock is normally distributed with mean $\mu_X = 25$ and *variance* $\sigma_X^2 = 81$. The price of Bangalore stock is also normally distributed with the mean $\mu_Y = 40$ and *variance* $\sigma_Y^2 = 121$. The stock prices have a negative correlation, $\rho_{XY} = -0.4$. Judy has asked you to determine the probability that the portfolio value exceeds 2,000.

Jointly Continuous Random Variables

- [Ans]

$$W = 20X + 30Y$$

$$\mu_W = 20\mu_X + 30\mu_Y = 1,700$$

$$\sigma_W^2 = 20^2\sigma_X^2 + 30^2\sigma_Y^2 + 2 * 20 * 30 * \text{Corr}(X, Y)\sigma_X\sigma_Y = 93,780$$
$$\sigma_W = 306.24$$

$$Z_W = \frac{2,000 - 1,700}{306.24} = 0.980$$

$$P(W > 2,000) = 0.1635$$

Thus, the probability that the portfolio value exceeds 2,000 is 0.1635.

本單元結束