統計學（一）

第七章 信賴區間估計: 單一群體
(Confidence Intervals Estimation: One Population)

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**
  - P. Newbold, W. L. Carlson and B. Thorne(2007). *Statistics for Business and the Economics*, 7^{th} Edition, Pearson.
- **參考書目**
  - Berenson, M. L., Levine, D. M., and Krehbiel, T. C. (2009). *Basic business statistics: Concepts and applications*, 11^{th} EditionPrentice Hall.
  - Larson, H. J. (1982). *Introduction to probability theory and statistical inference*, 3^{rd} Edition, New York: Wiley.
  - Miller, I., Freund, J. E., and Johnson, R. A. (2000). *Miller and Freund's Probability and statistics for engineers*, 6^{th} Edition, Prentice Hall.
  - Montgomery, D. C., and Runger, G. C. (2011). *Applied statistics and probability for engineers*, 5^{th} 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。

點估計
(Point Estimators and their Properties)

Types of Estimation

- **Recall : The Objective of Statistics**
  - To make inference about a **population** based on information contained in a **sample**.

- **Two Types of Estimation:**
  1) Point Estimation (點估計)
  2) Interval Estimation (區間估計)

Point Estimation

- **What is a Point Estimation (點估計)**

  A point estimator of a population parameter is a **rule** (or **formula**) that tells you how to calculate a **single number** based on sample data. The resulting number is called a **point estimate** of the parameter.

  Example： What are the point estimators for the following population parameters ?

  - μ：$\overline{X}$
  - σ：s
  - P：$\hat{p}$

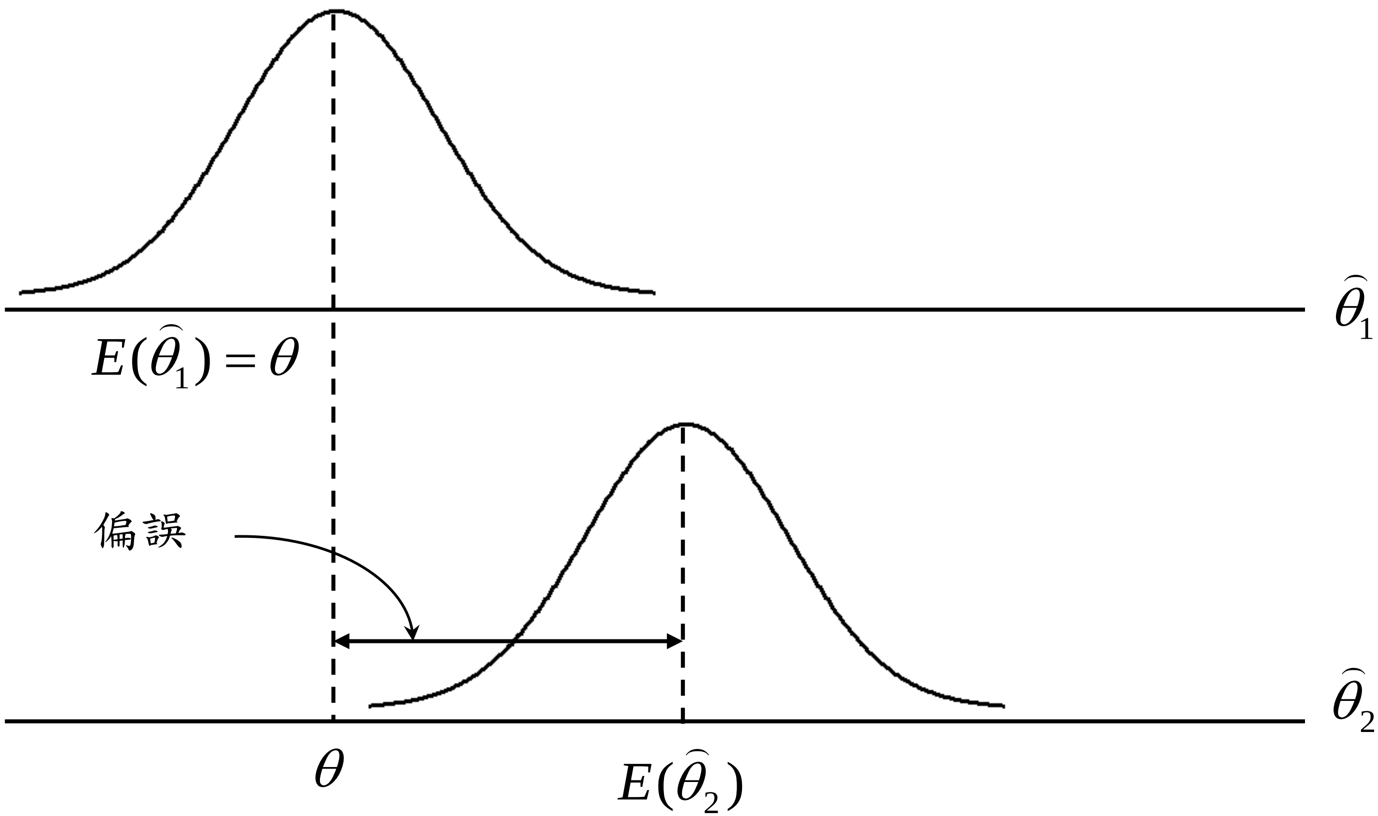
Point Estimation

- **How to Evaluate (評價) the Goodness of a Point Estimator**
  - Unbiasedness (不偏性)
  - Efficiency (有效性)
  - Consistency (一致性)
  - Sufficiency (充份性)
- **What is an Unbiased Estimator (不偏估計量)**
  - **An estimator** of a population **parameter** is said to be ***unbiased*(不偏)** if the **mean** of its sampling distribution is equal to the **parameter**. Otherwise the estimator is said to be **biased (偏頗)** .
  - That is, an estimator is unbiased if

    E(*Sample estimator*) = *Population parameter*

  - **Example :** $\overline{X}$ **and** S **and** $\widehat{p}$ are **unbiased** estimators

Unbiasedness (不偏性)



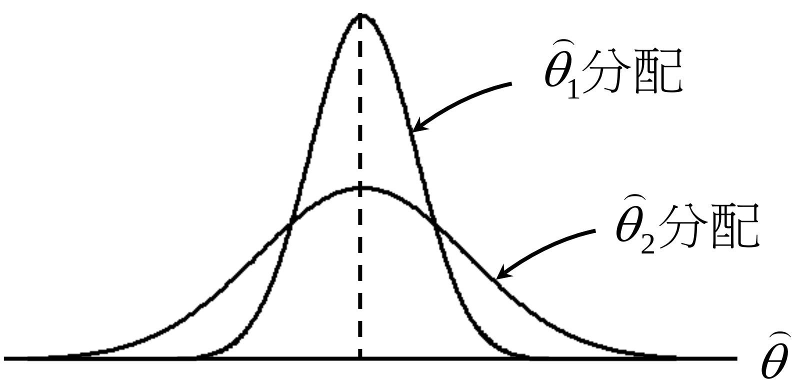
不偏估計量 $\hat{\theta}_1$ 與具有偏誤估計量 $\hat{\theta}_2$ 之比較

Efficiency (有效性)

- **What is an Efficient Estimator (有效估計量)?**
  - A point estimator $\hat{\theta}_1$ **is said to be a more efficient unbiased estimate of** θ **than** $\hat{\theta}_2$ if

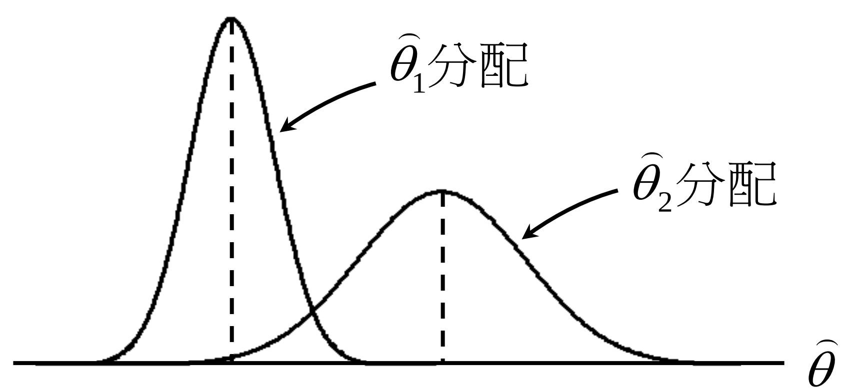
    1) $\hat{\theta}_1$ and $\hat{\theta}_2$ are both **unbiased** estimates of θ .

    2) **the variance** of the sampling distribution of $\hat{\theta}_1$ is less than that of $\hat{\theta}_2$.



$E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$

$\hat{\theta}_1$ 為 θ 之不偏，有效估計量



$E(\hat{\theta}_1)$ $E(\hat{\theta}_2)$

$\hat{\theta}_1$ 為 θ 之有效估計量

Maximum Error of Estimate for μ

- **What is the Maximum Error of Estimate for μ**
  - When we use $\overline{X}$ to estimate μ, the **maximum error of estimate (最大可能估計誤差)** can be expressed as follows:

$$E = Max\left|\overline{X} - \mu\right| = Z_{\frac{1-\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) \text{ or } Z_{\frac{1-\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right) \quad \text{for } \sigma \text{ is known}$$

$$= t_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right) \quad \text{for } \sigma \text{ is unknown}$$

Maximum Error of Estimate for μ

- **例 1：**

  An industrial engineer intends to use the mean of a random sample of size **150** to estimate the average mechanical aptitude of assembly line workers in a large industry. If, on the basis of experience, the engineer can assume that **σ= 6.2** for such data, what can he assert with probability **0.99** about the **maximum** size of the estimation **error**?

  [Ans]

$$E = Max\left|\overline{X} - \mu\right| = Z_{\frac{1-\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) = Z_{0.005}\left(\frac{6.2}{\sqrt{150}}\right) = 2.58 \times 0.506 = 1.306$$

Maximum Error of Estimate for μ

- **例 2：**

  In six determinations of the melting point of tin, a chemist obtained a mean of 232.26 degrees Celsius with a standard deviation of 0.14 degree. If he uses this mean as the actual melting point of tin, what can the chemist assert with 98% confidence about the maximum error?

[Ans]

$$E = Max\left|\overline{X} - \mu\right| = t_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right) = t_{0.01}\left(\frac{0.14}{\sqrt{6}}\right) = 3.365 \times 0.057 = 0.1923$$

估計群體之平均數
(Estimation of a Population Mean)

Interval Estimation

- **Interval Estimation**

  An interval estimator of a **population parameter** is a **rule** that tells you how to calculate **two numbers** based on sample data.

- **Confidence Coefficient 信賴係數**

  The **probability** that a **confidence interval** will enclose the estimated parameter is called the **confidence coefficient**.

Interval Estimation of μ

- **Interval Estimation of μ :**
  - **Case 1: σ is known** (1-α)100% Confidence Interval for the Population Mean, **μ**

$$\bar{y} \pm z_{\frac{1-\alpha}{2}} \sigma_{\bar{y}} = \bar{y} \pm z_{\frac{1-\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}}\right) \approx \bar{y} \pm z_{\frac{1-\alpha}{2}} \left(\frac{s}{\sqrt{n}}\right)$$

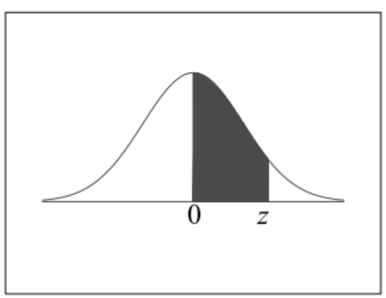
- **Assumptions: None.**
  - The CLT guarantees that $\bar{y}$ is approximately normally distributed **regardless of** the distribution of the sampled population.
- **Note**
  - The **value of z** selected for constructing such a confidence interval is called the **critical value (臨界值)**of the distribution.
  - (1-α) is called the **confidence coefficient. 信賴係數**
  - (1-α)100% is called the **confidence level. 信賴水準**

Interval Estimation of μ

- **Some Confidence Coefficients for Estimating μ**

$1-\alpha$	$Z_{\frac{1-\alpha}{2}}$
.90	1.645
.95	1.96
.98	2.33
.99	2.58

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

Interval Estimation of μ

- **例 1 :**

The quality control manager at a light bulb factory needs to estimate the **average** life of a large shipment of light bulbs. The process standard deviation is known to be 100 hours. A random **sample** of 50 light bulbs indicated a sample average life of 350 hours. Find a 95% confidence interval estimate of the true average life of light bulbs in this shipment.
Interpret your result.

[Ans] $\bar{y} \pm z_{\frac{1-\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) = 350 \pm z_{\frac{1-0.05}{2}}\left(\frac{100}{\sqrt{50}}\right) = 350 \pm 1.96 \times 14.14$

A 95% confidence interval for average life of light bulbs = (322.3 , 377.7)

ANS: We can feel __95%__ confident that the **true average life** of light bulbs in this shipment lies between __322.3__ hours and __377.7__ hours.

Interval Estimation of μ

- **例 2: 請參考課本319頁 例7.4**

A process produces bags of refined sugar. The weights of the contents of these bags are normally distributed with standard deviation 1.2 ounces. The contents of a random sample of 25 bags had a mean weight of 19.8 ounces. Find the upper and lower confidence limits of a 99% confidence interval for the true mean weight for all bags of sugar produced by the process.

[Ans] For a 99% confidence interval the reliability factor is $Z_{0.005} = 2.58$

and with a sample mean of 19.8, n=25, and a standard deviation of 1.2, the confidence limits are as follows:

$$UCL = \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 19.8 + 2.58\frac{1.2}{\sqrt{25}} = 20.42$$

$$LCL = \bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 19.8 - 2.58\frac{1.2}{\sqrt{25}} = 19.18$$

Interval Estimation of μ

- **Interval Estimation of μ :**
  - **Case 2: σ is unknow**n (1-α)100% Confidence Interval for the Population Mean, **μ**

$$\bar{y} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

- ***Assumptions*:**
  - The **population** from which the sample is selected has an approximate **normal** distribution.
- **Note**
  - $t_{\alpha/2}$ is obtained from the Student's *t* distribution with **(*n*-1)** degrees of freedom.

Interval Estimation of μ

- **例 3 :**

A sample of 25 employees from a company was selected and the annual salary for each was recorded. The **mean** and **standard deviation** of this sample were found to be **7750** and **900**, respectively.

Construct the **95 %C.I.** for the **population average salary**. **<u>Interpret your result.</u>**

[Ans] $$\bar{y} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) = 7750 \pm t_{0.05/2}\left(\frac{900}{\sqrt{25}}\right) = 7750 \pm 2.064 \times 180$$

A 95% confidence interval for the average salary = (7378.48 , 8121.52)

<u>ANS:</u> We are <u> 95% </u> confident that the true average salary lies between <u> 7378.48 </u> and <u> 8121.52 </u> .

Interval Estimation of μ

- **例 4: 請參考課本325頁 例7.5**

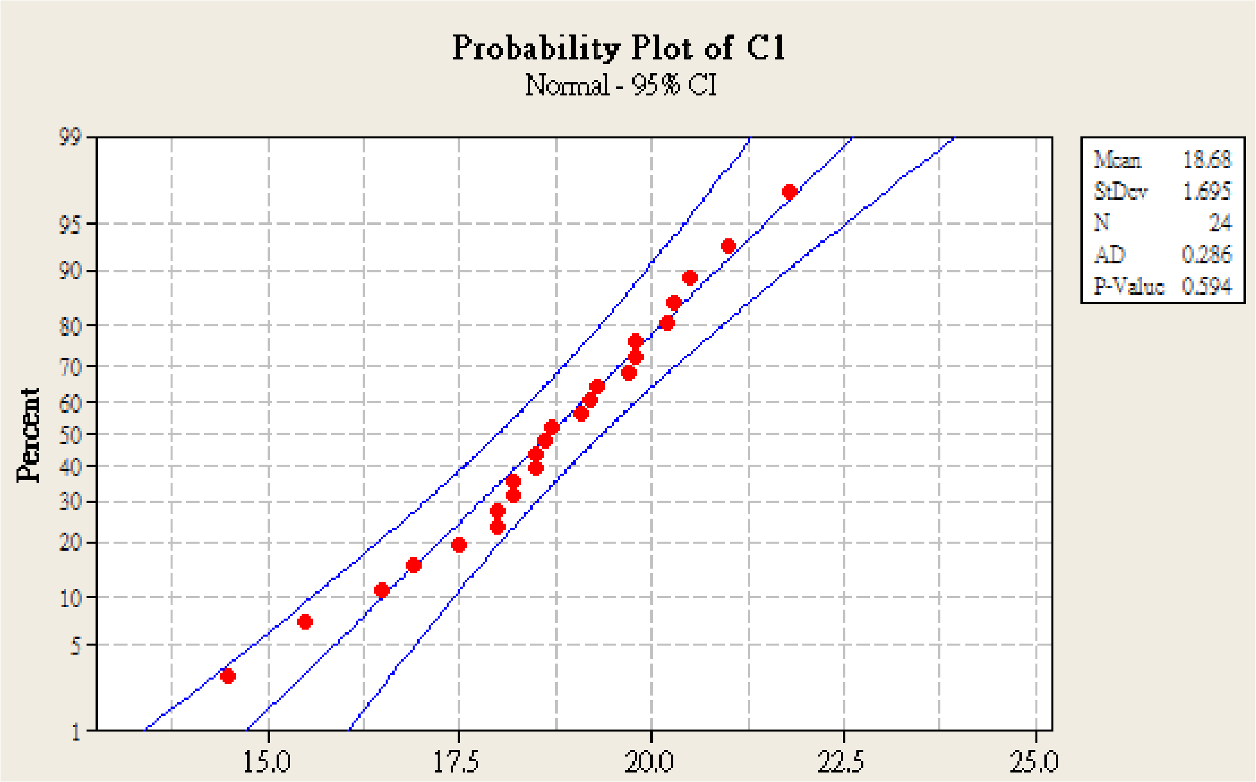
  Gasoline prices rose drastically during the early years of this century. Suppose that a recent study was conducted using truck drivers with equivalent years of experience to test run 24 trucks of a particular model over the same highway. Estimate the population mean fuel consumption for this truck model with 90% confidence if the fuel consumption, in miles per gallon, for these 24 trucks was as follows:

15.5	21.0	18.5	19.3	19.7	16.9	20.2	14.5
16.5	19.2	18.7	18.2	18.0	17.5	18.5	20.5
18.6	19.1	19.8	18.0	19.8	18.2	20.3	21.8

Interval Estimation of μ

- **[Ans](1/2)**

  The normal probability plot figure does not provide evidence of non-normality.



Interval Estimation of μ

- **[Ans](2/2)**

  Calculating the mean and standard deviation, we find the following:

  $\bar{x} = 18.68 \quad s = 1.69526 \qquad t_{n-1,\alpha/2} = t_{23,0.05} = 1.714$

  The confidence interval is as follows:

  $$\bar{x} \pm t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} = 18.68 \pm 1.714 \times (0.3460)$$
  $$= 18.68 \pm 0.5930$$

  The confidence interval then is $18.1 < \mu < 19.3$

Interpretation of the confidence interval

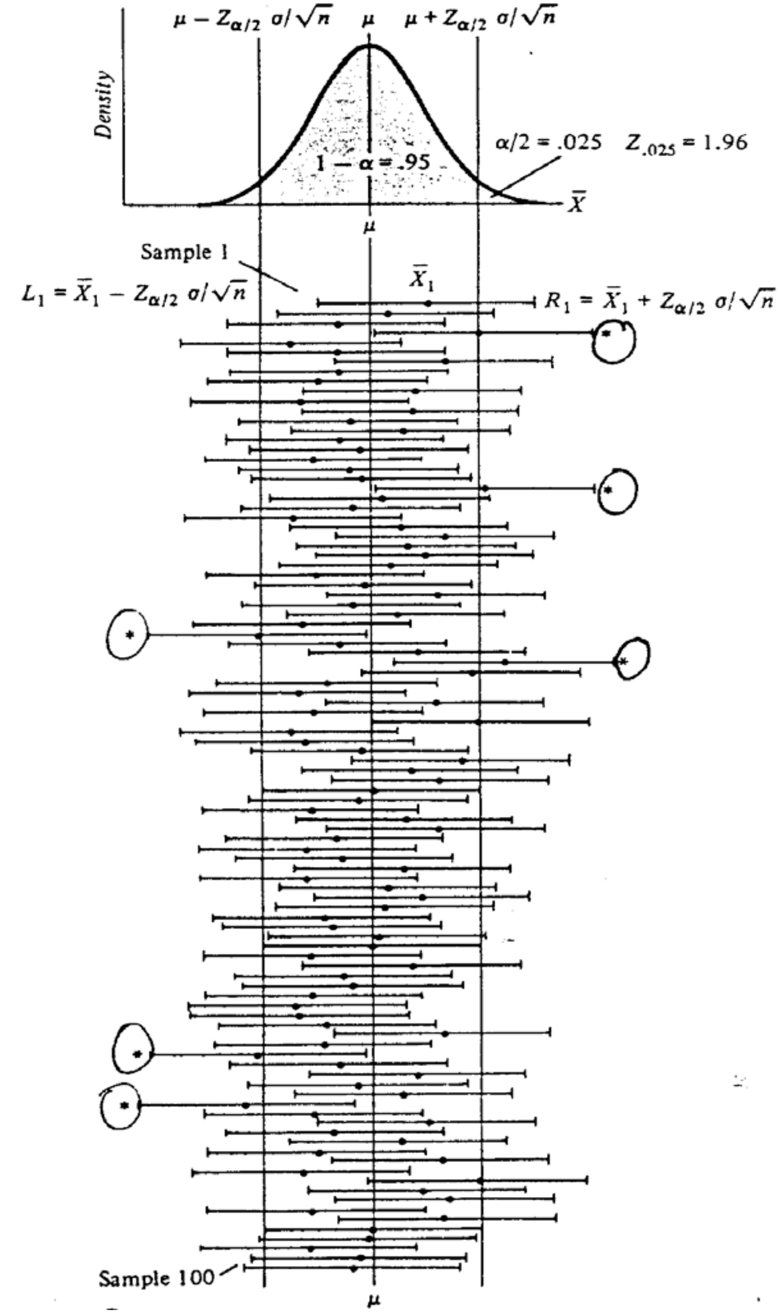
- **Interpretation of the confidence interval for population mean**
  - The **confidence** we have in the limits _322.3_ hours and _377.7_ hours in 例 **1** derives from our confidence in the statistical procedure that gave rise to them. The procedure gives random variables L and R that have a **95 %** chance of enclosing the **true** but **unknown** mean μ; whether their specific values _____ and enclose μ we have **no way of knowing**.
  - The reason that **"we can feel 95% confident"** is as follows: If we were to take 100 different samples from the sample population and calculate the confidence limits for each sample, then we would expect that about **95** of these **100** intervals would contain the true value of μ, and **5** would not contain the true value of μ. Since we usually only have one sample and hence one confidence interval, we do not know whether our interval is one of the 95 or one of the 5. In this sense, then, **we are 95% confident.**

Interpretation of the confidence interval

The meaning of being **95% confidence** is shown in the results of a sampling exercise. Statistic students took 100 different random samples of size n=10 each from a normal population with a mean μ and a standard deviation σ= 1.66. One hundred sample means, X's, and corresponding confidence intervals were then computed, and the results are presented in the following Figure (see next page).

The mean μ was contained in **94** of the 100 intervals, as shown in the figure. This result conforms to our expectation that about 95 of our 100 intervals should encompass the mean μ.

∗These six intervals do not contain the mean μ. Of the 100 confidence intervals. 94 contain the mean μ



Source: Watson, C., Billingsley, P., Croft, J., Huntsberger, O. (1986) *Statistics for Management and Economics* 4th Ed. Allyn and Bacon, Newton, MA, p 328.

估計群體之比率值
(Estimation of a Population Proportion)

Estimation of a Population Proportion

- **Estimation of a Population Proportion**
  - **Large-Samples (1-α)100% Confidence Interval for the Population Proportion, *P***

$$\hat{p} \pm z_{\alpha/2}\,\sigma_{\hat{p}} \approx \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

- ***Assumptions*:**
  - The sample size **n** is sufficiently large so that the approximation is valid. As a rule of thumb, the condition of a "**sufficiently large**" sample size will be satisfied if $\boldsymbol{n\hat{p}\hat{q} \geq 5}$.

Estimation of a Population Proportion

- **例 1:**

  In a random sample of 400 industrial accidents, it was found that 231 were due at least partially to unsafe working conditions. Construct a 99% confidence interval for the corresponding true proportion. **Explain your result.**

[Ans]

$$\hat{p} = \frac{231}{400} = 0.5775 \quad \hat{q} = 1 - \hat{p} = 0.4225$$

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.5775 \pm z_{0.01/2}\sqrt{\frac{0.5775 \times 0.4225}{400}} = 0.5775 \pm 0.0637$$

A 99% confidence interval for the true proportion = (0.5138 , 0.6412)

– **ANS:** We are __99%__ confident that the true proportion of accidents due at least partially to unsafe working conditions lies between __0.5138__ and __0.6412__.

Estimation of a Population Proportion

- **例 2: 請參考課本329頁 例7.6**

Management wants an estimate of the proportion of the corporation's employees who favor a modified bonus plan. From a random sample of 344 employees it was found that 261 were in favor of this particular plan. Find a 90% confidence interval estimate of the true population proportion that favors this modified bonus plan.

[Ans]

If P denotes the true population proportion and $\hat{p}$ the sample proportion, then confidence intervals of the population are obtained as

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

It follows that n=344 $\hat{p} = {}^{261}/_{344} = 0.759$

$$0.759 - 1.645\sqrt{\frac{(0.759)(0.241)}{344}} < P < 0.759 + 1.645\sqrt{\frac{(0.759)(0.241)}{344}}$$

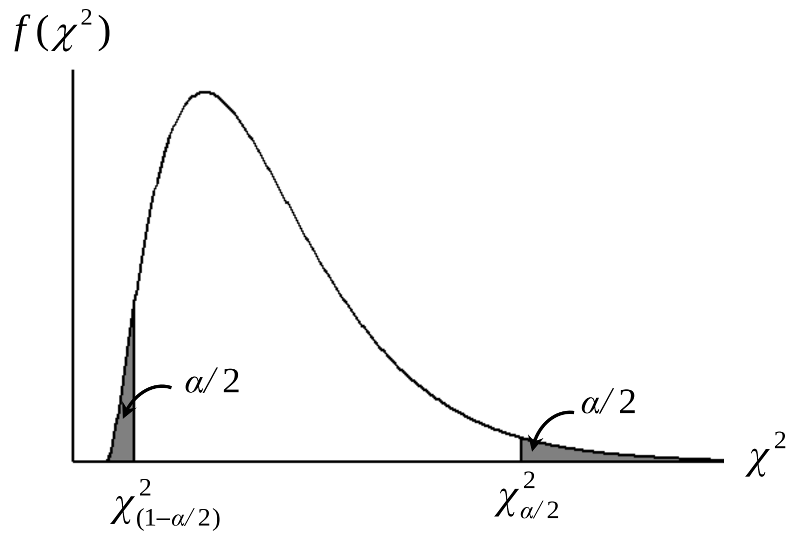
or $0.721 < P < 0.797$

估計群體之變異數
(Estimation of a Population Variance)

Confidence Interval for σ^2

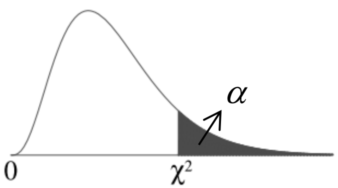
- **A (1-α)100% Confidence Interval for a Population Variance, σ^2**

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}}$$



- **Assumptions:**
  – The Population from which the sample is selected has an approximate **normal** distribution.
- **Note:**
  – $\boldsymbol{\chi^2_{\alpha/2}}$ and $\boldsymbol{\chi^2_{(1-\alpha/2)}}$ can be obtained from Table 8.

Chi-Square Distribution Table



df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928

Confidence Interval for σ²

- **例 1:**

While performing a strenuous (激烈的) task, the pulse rate (脈博跳動率) of **25** workers increased on the average by **18.4** beats per minute with a standard deviation of **4.9** beats per minute. Find a 95% confidence interval for the corresponding **population standard deviation.**

[Ans]

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}} \Rightarrow \frac{(25-1)(4.9)^2}{\chi^2_{0.05/2}} \le \sigma^2 \le \frac{(25-1)(4.9)^2}{\chi^2_{(1-0.05/2)}}$$

$$\Rightarrow \frac{576.24}{39.3641} \le \sigma^2 \le \frac{576.24}{12.4011} \Rightarrow 14.64 \le \sigma^2 \le 46.47 \Rightarrow 3.83 \le \sigma \le 6.82$$

A 95% confidence interval for the population standard deviation = (3.83 , 6.82)

Confidence Interval for σ²

- **例 2: 請參考課本334頁 例7.7**

  The manager of Northern Steel Inc. wants to assess the temperature variation in the firm's new electric furnace. A random sample of 25 temperatures over a one-week period is obtained, and the sample variance is found to be $s^2 = 100$. Find a 95% confidence interval for the population variance temperature.

[Ans] $$\frac{(n-1)s^2}{\chi^2_{n-1,\alpha/2}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{n-1,(1-\alpha/2)}} \Rightarrow \frac{(25-1)(100)^2}{\chi^2_{0.05/2}} \le \sigma^2 \le \frac{(25-1)(100)^2}{\chi^2_{(1-0.05/2)}}$$

$$\Rightarrow \frac{240}{39.3641} \le \sigma^2 \le \frac{240}{12.4011} \Rightarrow 60.97 \le \sigma^2 \le 193.53$$

A 95% confidence interval for the population standard deviation = (7.81 , 13.91)

Confidence Interval for a Finite Population

Estimation of the Population Mean, Finite population

- **Recall:**

Let $x_1, x_2, \ldots, x_n$ denote the values observed from a simple random sample of size n, taken from a population of N members with mean μ.

1. The sample mean is an unbiased estimator of the population mean, μ. The point estimate is

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

2. An unbiased estimation procedure for the variance of the sample mean yields the point estimate

$$\hat{\sigma}_{\bar{x}}^2 = \frac{s^2}{n}\left(\frac{N-n}{N-1}\right)$$

3. A 100(1-α)% confidence interval for the population mean is given by

$$\bar{x} - t_{n-1,\alpha/2}\hat{\sigma}_{\bar{x}} < \mu < \bar{x} + t_{n-1,\alpha/2}\hat{\sigma}_{\bar{x}}$$

Estimation of the Population Mean, Finite population

- **例 1: 請參考課本337頁 例7.8**

In a particular city 1,118 mortgages were financed last year. A random sample of 60 of these had a mean amount \$87,300 and a standard deviation \$19,200. Estimate the mean amount of all mortgages financed in this city last year, and find a 95% confidence interval.

[Ans] Denote the population mean by μ. It is know n that

$$N = 1{,}118 \quad n = 60 \quad \bar{x} = \$87{,}300 \quad s = \$19{,}200$$

$$\hat{\sigma}_{\bar{x}}^2 = \frac{s^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{(19{,}200)^2}{60}\frac{1{,}058}{1{,}118} = 5{,}819{,}474$$

$$\$87{,}300 - (2.00)(2{,}412) < \mu < \$87{,}300 + (2.00)(2{,}412)$$

or $$\$82{,}476 < \mu < \$92{,}124$$

- **例 2: 請參考課本337頁 例7.9**

Toivo Steendahl Associates, a major auditing firm, has been engaged to audit Big Woods Furniture, an upper Midwest furniture retailer, in order to determine the value of the firm's assets preceding a takeover by National Distribution. As part of this audit we have been asked to conduct receivable to estimate mean value of the accounts receivable. The company presently has 1,420 accounts receivable on the ledger.

[Ans] Denote the population mean by μ. It is know n that

$$N = 1{,}420 \quad n = 100 \quad \bar{x} = 784 \quad s^2 = 2{,}300$$

$$\hat{\sigma}_{\bar{x}}^2 = \frac{s^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{(2{,}300)}{100}\frac{1{,}320}{1{,}419} = 21.395$$

$$784 - 1.96(4.626) < \mu < 784 + 1.96(4.626)$$

or $$775 < \mu < 793$$

Estimation of the Population Total, Finite population

- **Recall:**

  Suppose a simple random sample of size n from a population of size N is selected and that the quantity to be estimated is the population total Nμ.

  1. An unbiased estimation procedure for the population total Nμ yields the point estimate $N\bar{x}$.
  2. An unbiased estimation procedure for the variance of our estimator of the population total yields the point estimate

$$\mathbf{N}\hat{\boldsymbol{\sigma}}_{\bar{\mathbf{x}}} = \frac{\boldsymbol{Ns}}{\sqrt{\boldsymbol{n}}}\sqrt{\frac{(\boldsymbol{N}-\boldsymbol{n})}{(\boldsymbol{N}-\mathbf{1})}}$$

  3. A 100(1-α)% confidence interval for the population total is obtained from

$$\boldsymbol{N}\bar{\mathbf{x}} - \mathbf{t}_{\mathbf{n-1},\alpha/2}\boldsymbol{N}\hat{\boldsymbol{\sigma}}_{\bar{\mathbf{x}}} < \mathbf{N}\boldsymbol{\mu} < \mathbf{N}\bar{\mathbf{x}} + \mathbf{t}_{\mathbf{n-1},\alpha/2}\boldsymbol{N}\hat{\boldsymbol{\sigma}}_{\bar{\mathbf{x}}}$$

- **例 3: 請參考課本339頁 例7.10**

  Supposed that there are 1,395 colleges in the United States. From a simple random sample of 400 of these schools it was found that the sample mean enrollment during the past year in business statistics courses was 320.8 students, and the sample standard deviation was found to be 149.7 students. Estimate the total number of students enrolled in business statistics courses in the previous year, and find a 99% confidence interval.

Estimation of the Population Total, Finite population

- **[Ans]**

If the population mean is μ, an estimate of Nμ includes that following:

$$N = 1{,}395 \qquad n = 400 \qquad \bar{x} = 320.8 \qquad s = 149.7$$

Our point estimate for the total is

$$\mathrm{N}\bar{x} = (1{,}395)(320.8) = 447{,}516$$

$$N\hat{\sigma}_{\bar{x}} = \frac{Ns}{\sqrt{n}}\sqrt{\frac{(N-n)}{(N-1)}} = \frac{(1{,}395)(149.7)}{\sqrt{400}}\sqrt{\frac{995}{1{,}394}} = 8{,}821.6$$

Then $$N\bar{x} - t_{n-1,\alpha/2}N\hat{\sigma}_{\bar{x}} < N\mu < N\bar{x} + t_{n-1,\alpha/2}N\hat{\sigma}_{\bar{x}}$$

$$447{,}516 - (2.58)(8{,}821.6) < N\mu < 447{,}516 + (2.58)(8{,}821.6)$$

or $$424{,}756 < N\mu < 470{,}276$$

Estimation of the Population Proportion, Finite population

- **Recall:**

  Let $\hat{p}$ be the proportion possessing a particular characteristic in a random sample of n observations from a population with a proportion, P, of whose members process that characteristic.

  1. The sample proportion,$\hat{p}$, is an unbiased estimator of the population proportion, P.
  2. An unbiased estimation procedure for the variance of our estimator of the population proportion yields the point estimate

$$\boldsymbol{\sigma_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n} \times \frac{(N-n)}{(N-1)}}$$

  3. Provided the sample size is large, 100(1-α)% confidence interval for the population proportion are given by

$$\boldsymbol{\hat{p} - Z_{\alpha/2}\hat{\sigma}_{\hat{p}} < P < \hat{p} + Z_{\alpha/2}\hat{\sigma}_{\hat{p}}}$$

- **例 4: 請參考課本340頁 例7.11**

  From a simple random sample of 400 of the 1,395 colleges in our population, it was found that business statistics was a two-semester courses in 141 of the sampled colleges. Estimate the proportion of all colleges for which the course is two semesters long, and find a 90% confidence interval.

Estimation of the Population Proportion, Finite population

- **[Ans]**

Given $N = 1{,}395 \quad n = 400 \quad \hat{p} = \frac{141}{400} = 0.3525$

$$\sigma_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n} \times \frac{(N-n)}{(N-1)} = \frac{(0.3525)(0.6475)}{400} \times \frac{995}{1{,}394} = 0.0004073$$

Then $$\hat{p} - Z_{\alpha/2}\hat{\sigma}_{\hat{p}} < P < \hat{p} + Z_{\alpha/2}\hat{\sigma}_{\hat{p}}$$

$$0.3525 - (1.645)(0.0202) < P < 0.3525 + (1.645)(0.0202)$$

or $$0.3193 < P < 0.3857$$

選擇樣本大小以估計參數
(Choosing the Sample Size for Estimating μ and P)

Choosing the Sample Size for Estimating μ

1) Choosing the Sample Size for Estimating μ Correct to Within E units with Probability (1-α)

$$\mathbf{n} = \left(\frac{\mathbf{z}_{\alpha/2}}{\boldsymbol{E}}\right)^{2}(\boldsymbol{\sigma}^{2})$$

where n is the numbers of observations sampled from the population, and $\boldsymbol{\sigma}^{2}$ is the population variance.

Note: $\boldsymbol{\sigma}^{2}$ will usually have to be approximated.

Choosing the Sample Size for Estimating μ

- **例 1: 請參考課本364頁 Example8.7**

The lengths of metal rods produced by an industrial process are normally distributed with a standard deviation of 1.8 millimeters(mm). Based on a random sample of nine observations from this population, the 99% confidence interval

$$194.65 < \mu < 197.75$$

was found for the population mean length. Suppose that a production manager believes that the interval is too wide for practical use and instead requires a 99% confidence interval extending no further than 0.5 mm on each side of the sample mean. How large a sample is needed to achieve such an interval?

Choosing the Sample Size for Estimating μ

- **[Ans]**

Since

$$E = \frac{Z_{\alpha/2}(\sigma)}{\sqrt{n}} = 0.5, \qquad \sigma = 1.8 \;\; and \;\; Z_{\alpha/2} = Z_{0.005} = 2.576$$

the required sample size is as follows:

$$n = \frac{Z_{\alpha/2}^2 \sigma^2}{E^2} = \frac{(2.576)^2(1.8)^2}{(0.5)^2} \approx 86$$

Therefore, to satisfy the manager's requirement, a sample of at least 86 observations is needed.

Choosing the Sample Size for Estimating P

1) Choosing the Sample Size for Estimating P Correct to Within E units with Probability (1-α)

$$\mathbf{n} = \left(\frac{\mathbf{Z}\alpha/_{2}}{\mathbf{E}}\right)^{2} (\boldsymbol{P}(\mathbf{1}-\boldsymbol{P}))$$

where **P** is the population proportion; n is the numbers of observations to be sampled from the population.

Note: P is usually set as 0.5

Choosing the Sample Size for Estimating P

- **例 2: 請參考課本366頁 Example8.8**

A confidence interval was calculated for the proportion of graduate admissions personnel who viewed scores on standardized exams as very important in the consideration of a candidate. Based on 142 observations the interval obtained was as follows:

$$0.533 < P < 0.693$$

Suppose, instead, that it must be ensured that a 95% confidence interval for the population proportion extends no further than 0.06 on each side of the sample proportion. How large of a sample must be taken?

Choosing the Sample Size for Estimating P

- **[Ans]**

It is given that

$$E = \frac{Z\alpha_{/2}\sqrt{P(1-P)}}{\sqrt{n}} = 0.06 \ \ and \ \ Z\alpha_{/2} = Z_{0.025} = 1.96$$

Thus, the number of sample observations needed is as follows:

$$n = \frac{Z^2_{\alpha_{/2}}\,(P(1-P))}{E^2} = \frac{(1.96)^2(0.25)}{(0.06)^2} = 266.78 \Rightarrow n = 267$$

To achieve this narrower confidence interval, a minimum of 267 sample observations is required (a significant increase over the original 142 observations).

Choosing the Sample Size for Estimating P

- **例 3: 請參考課本366頁 Example8.9**

  Suppose that an opinion survey following a presidential election reported the views of a sample of U.S. citizens of voting age concerning changing the Electoral College process. The poll was said to have "a 3% margin of error." The implication is that a 95% confidence interval for the population proportion holding a particular opinion is the sample proportion plus or minus at most 3%. How many citizens of voting age need to be sampled to obtain this 3% margin of error?

- **[Ans]**

$$n = \frac{Z^2_{\alpha/2}\,(P(1-P))}{E^2} = \frac{(1.96)^2(0.25)}{(0.03)^2} = 1{,}067.111 \Rightarrow n = 1{,}068$$

Therefore, 1,068 U.S. citizen of voting age need to be sampled to achieve the desired result.

本單元結束