

# 【统计简单学】

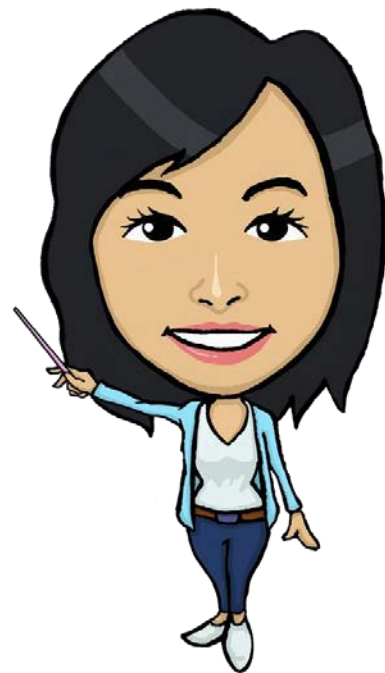
## 第五单元 估计

授课教师：唐丽英 教授

新竹交通大学  
工业工程与管理学系

# 第五单元 内容大纲

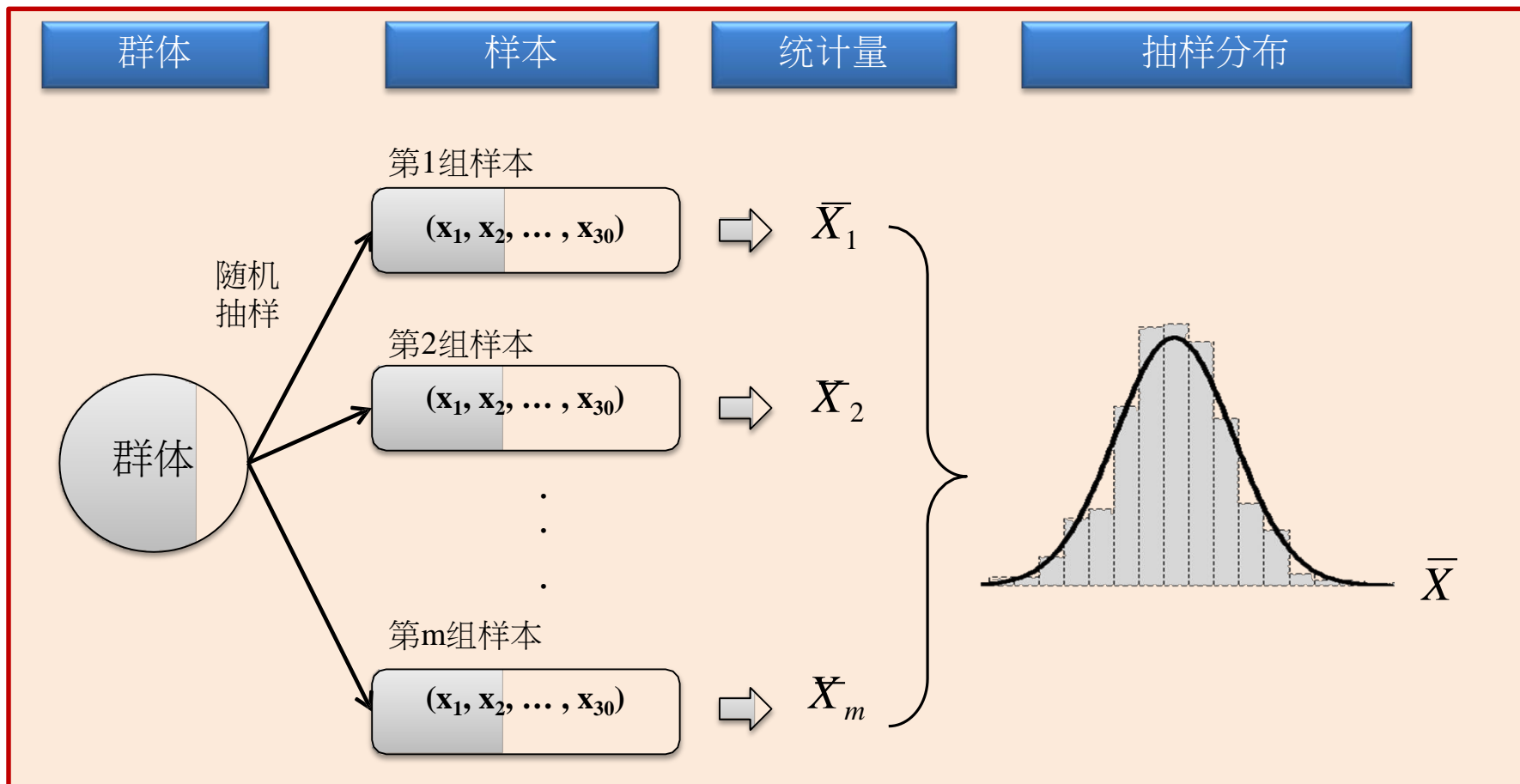
- 第一部份：抽样分布与中央极限定理
- 第二部份：估计方法简介
- 第三部份：区间估计
  1. 群体平均数  $\mu$  之  $(1-\alpha)$  100% 信赖区间
  2. 群体比率值  $P$  之  $(1-\alpha)$  100% 信赖区间
  3. 群体变异数  $\sigma^2$  之  $(1-\alpha)$  100% 信赖区间
- 第四部份：样本大小之决定



# 第一部份：抽样分布与中央极限定理

# 抽样分布 (Sampling Distribution)

- 样本统计量之机率分布称为「抽样分布」。



# 中央极限定理

- $\bar{X}$  的抽样分布会因群体分配的不同而异。
- 但随着抽样数  $n$  的增加， $\bar{X}$  会有逐渐向中央聚集且呈现常态分布的趋势。
- 中央极限定理(Central Limit Theorem, C.L.T.)：

自平均数为  $\mu$  及变异数为  $\sigma^2$  之群体中，以放回抽样之方式重复抽取样本大小为  $n$  之样本，当  $n$  够大时（ $n \geq 30$ ），则

(1)  $\bar{X}$  的抽样分布近似常态分布。其平均数为  $\mu$ ，标准偏差为  $\frac{\sigma}{\sqrt{n}}$ 。当  $n$  越大， $\bar{X}$  就越接近常态分布。

$$\text{即：}\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

(2)  $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$  近似标准常态分布  $N(0,1)$ 。

(3) 若群体呈常态分布，则不论  $n$  大小， $\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$ 。

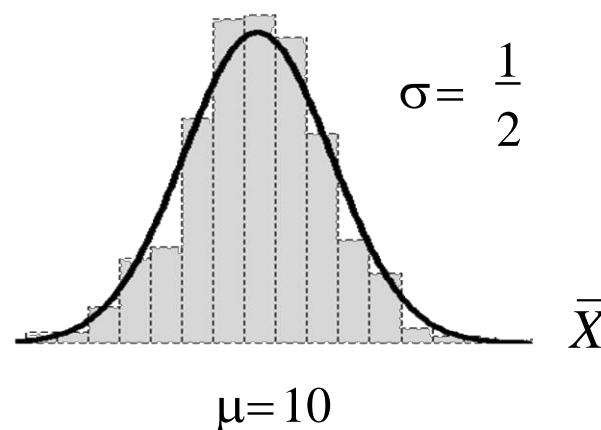
# 中央极限定理

- **例1**：假设某随机变数 $\mathbf{X}$ 之群体平均数 $\mu=10$ ，变异数 $\sigma^2=9$ 。令 $\mathbf{x}_1, \dots, \mathbf{x}_{36}$ 为抽自此群体的**36**个随机样本，试求此样本之平均数的抽样分布为何？

【解】

根据中央极限定理，当  $n \geq 30$  时， $\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$ ，

即： $\bar{X} \sim N(10, \frac{3}{\sqrt{36}}) = N(10, \frac{1}{2})$



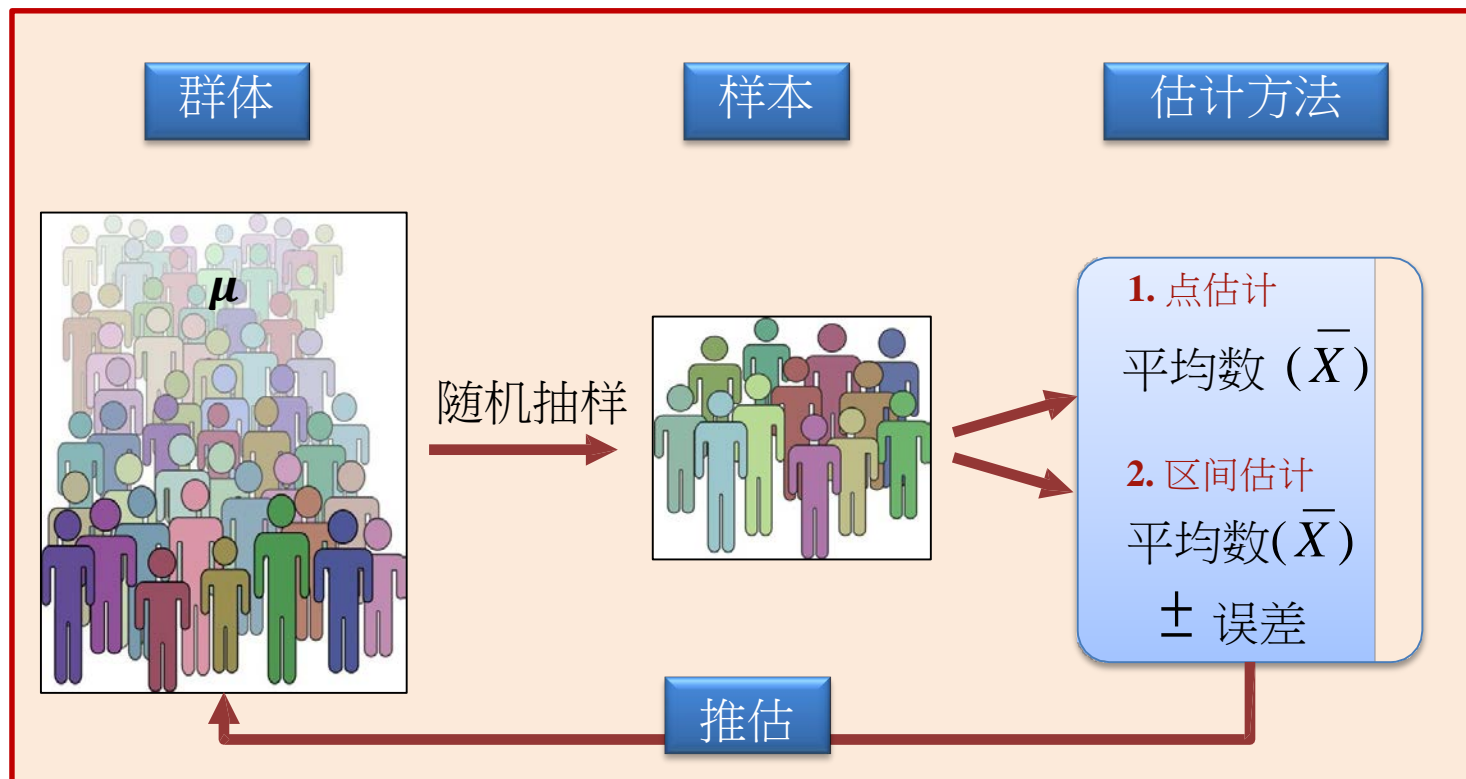
## 第二部份：估计方法简介

# 两种形态之估计

- 估计(**Estimation**)是指用样本资讯来推测群体参数的方法。

常用的估计方法有两种：

- 1) 点估计 ( **Point Estimation** )
- 2) 区间估计 ( **Interval Estimation** )





# 点估计

- 点估计是根据样本资料求算出统计量以推测群体参数值。
  - 例如：利用样本资料之平均数来估计群体平均数，此即为点估计。
- 由于点估计方法只提供未知参数一个可能的数值，故称之为点估计。
- 点估计方法无法提供估计之误差及可信度。
- 常用的点估计式：
$$\mu \Rightarrow \bar{X}$$
$$\sigma \Rightarrow S$$
$$p \Rightarrow \hat{p}$$

# 点估计

- 例2：搜集某制程在一特定温度下之良率(%)资料五笔，分别为：

1	2	3	4	5
90	88	90	92	93

请利用点估计方法来推估此制程在特定温度下之真实良率为何？

[解]：

真实良率之估计值：
$$\hat{\mu} = \bar{X} = \frac{90+88+90+92+93}{5} = 90.6(\%)$$

## 第三部份：区间估计

# 区间估计

- 区间估计是指根据样本资料来估计群体参数值可能之范围，此范围称为信赖区间（**Confidence Interval**）。
- 信赖区间包含群体参数的机率称为信赖水准，一般以  **$(1-\alpha)$  100%** 表之。
- 由于点估计方法仅提供未知参数的一个可能数值，并未能考虑到估计误差及可信度，因此，根据未知参数的点估计式及其抽样分布，可进一步推估未知参数值之可能范围。
- 在估计参数值时，若要考虑估计误差及可信度，则必须使用区间估计的方法。

# 群体参数之区间估计

1. 群体平均数  $\mu$  之  $(1-\alpha)$  100% 信赖区间
  - 当  $\sigma$  已知时
  - 当  $\sigma$  未知时
2. 群体比率值  $P$  之  $(1-\alpha)$  100% 信赖区间
3. 群体变异数  $\sigma^2$  之  $(1-\alpha)$  100% 信赖区间

# 群体平均数之区间估计

群体平均数  $\mu$  之  $(1-\alpha)$  100% 信赖区间

— 当  $\sigma$  已知时

$$\bar{X} \pm z_{\frac{1-\alpha}{2}} \left( \frac{\sigma}{\sqrt{n}} \right)$$

- 1) 公式中之  $z$  值称为临界值(**critical value**)。
- 2)  $(1-\alpha)$  称为信赖系数(**confidence coefficient**)。
- 3)  $(1-\alpha)$  100% 称为信赖水准。

# 群体平均数之区间估计

## 群体平均数 $\mu$ 之 $(1-\alpha)$ 100% 信赖区间

— 当  $\sigma$  未知且为大样本时（用样本标准偏差  $S$  取代  $\sigma$

)

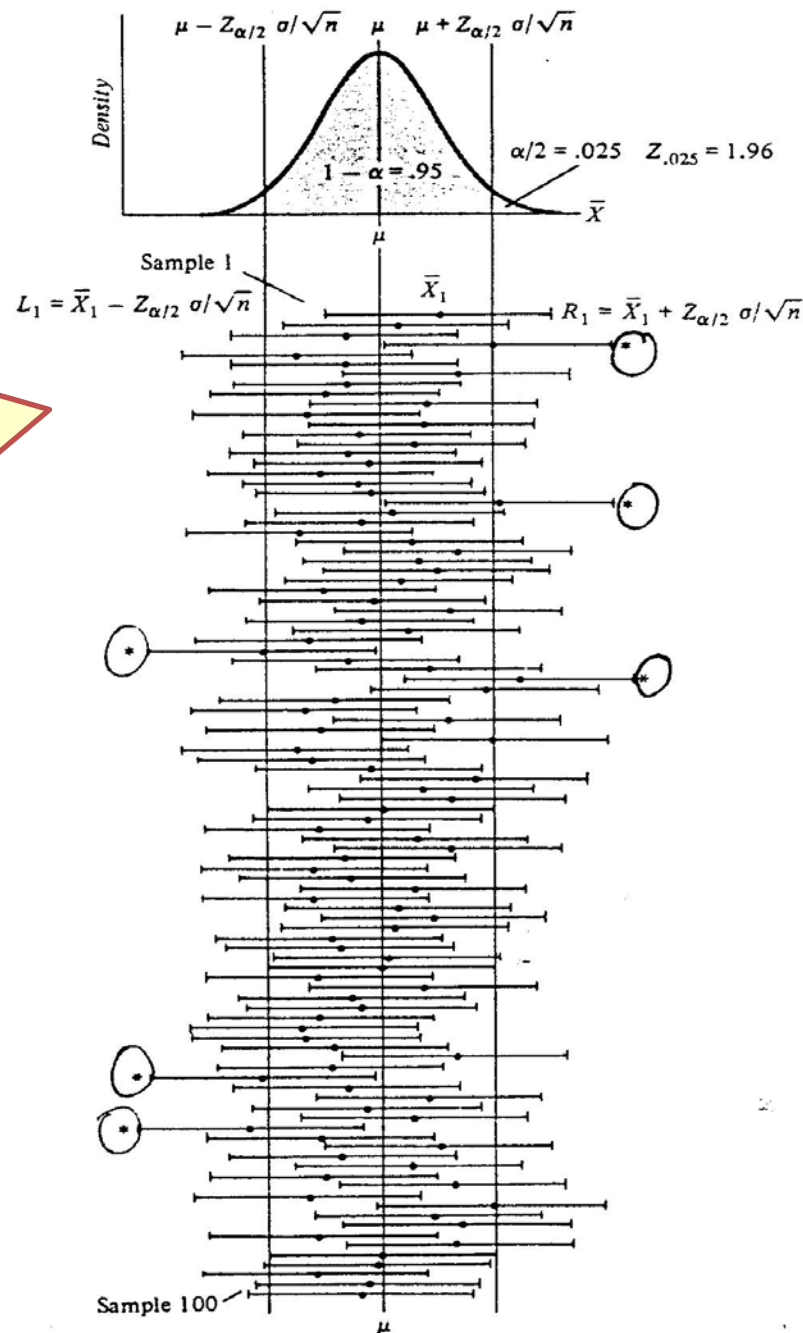
$$\bar{X} \pm z_{\frac{1-\alpha}{2}} \left( \frac{S}{\sqrt{n}} \right)$$

- 1) 公式中之  $z$  值称为临界值(**critical value**)。
- 2)  $(1-\alpha)$  称为信赖系数(**confidence coefficient**)。
- 3)  $(1-\alpha)$  100% 称为信赖水准。

# 群体平均数之区间估计

当 $\sigma$ 已知时， $\mu$ 之95%信赖区间示意图

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



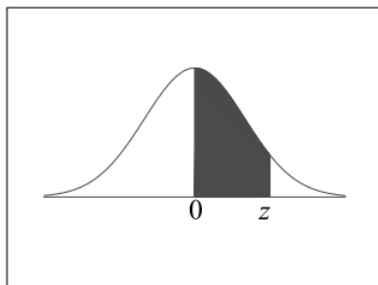


# 群体平均数之区间估计

- 常用之信赖系数( $1-\alpha$ )与相对应的  $Z$  值

$1 - \alpha$	$Z_{\frac{1-\alpha}{2}}$
0.90	1.645
0.95	1.96
0.98	2.33
0.99	2.575

Standard Normal Distribution Table



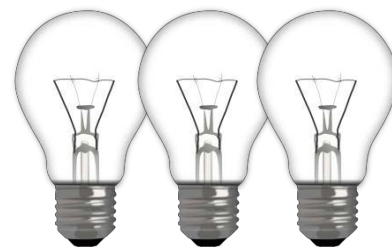
$z$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.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

# 群体平均数之区间估计

- 例3：假设某制造灯泡的厂商欲估计其所生产之A型灯泡的平均寿命，故从生产线上随机抽取50个灯泡进行寿命试验，获得以下样本寿命数据。试计算该A型灯泡之平均寿命的95%信赖区间。

(假设标准差已知  $\sigma=100$ )

$$\bar{x} = 350 \text{ 小时}, \quad n = 50$$



【解】

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

该 A 型灯泡平均寿命的 95% 信赖区间 = (322.3 , 377.7) 小时

# 群体平均数之区间估计

群体平均数  $\mu$  之  $(1-\alpha)$  100% 信赖区间

— 当  $\sigma$  未知时

$$\bar{X} \pm t_{\alpha/2, n-1} \left( \frac{S}{\sqrt{n}} \right)$$

- 1) 公式中之  $t$  值称为临界值(**critical value**)。
- 2)  $(1-\alpha)$  称为信赖系数(**confidence coefficient**)。
- 3)  $(1-\alpha)$  100% 称为信赖水准。
- 4)  $(n-1)$  称为自由度。

# t分布简介 (Student's t Distribution)

- 由一常态群体中，以放回方式重覆抽取样本大小为 $n$ 之随机样本，则

$\frac{\bar{X} - \mu}{S / \sqrt{n}}$  会服从 **t** 分布，

其自由度 (Degrees of freedom, d.f.) 为  $n-1$ 。

即：
$$\frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t_{(n-1)}$$

t 分布的发明人

William Sealy Gosset



Student in 1908

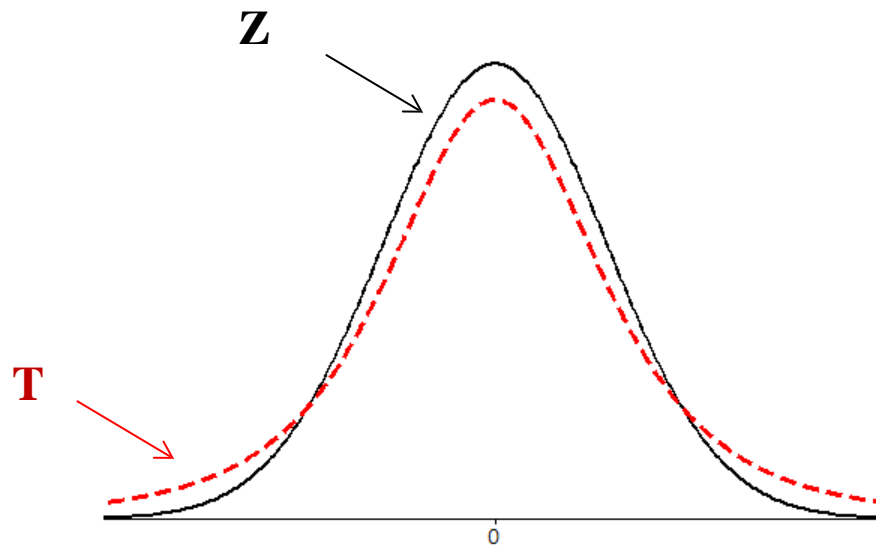
**Born** June 13, 1876  
Canterbury, Kent, England  
**Died** October 16, 1937 (aged 61)  
Beaconsfield, Buckinghamshire, England

source:

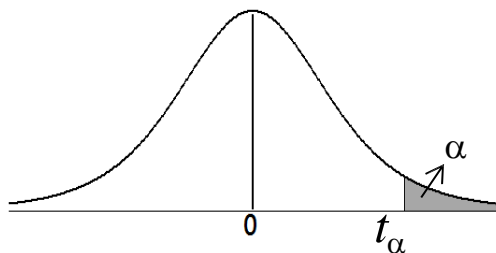
[http://en.wikipedia.org/wiki/William\\_Sealy\\_Gosset](http://en.wikipedia.org/wiki/William_Sealy_Gosset)

# t分布有什么特性？

- $t$  分布与  $Z$  分布非常类似，此二分布之比较如下：
  - 两者皆呈对称之钟形分布。
  - 两者之平均数皆为 0。
  - $t$  分布之变异较  $Z$  分布为大。
  - 当自由度  $n-1$  愈大（即  $n$  愈大）， $t$  分布就愈接近  $Z$  分布。



# t 分布之机率分布表



$\alpha$	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
v							
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
$\infty$	1.282	1.645	1.960	2.326	2.576	3.090	3.291

## 利用 $t$ 分布之机率分布表查 $t$ 值

- 例4：当 $n=6$ 及 $\alpha=0.05$ 时，请利用  $t$  分布表找出 $t_{\alpha}$ 及 $t_{\alpha/2}$ 之值。

$v$	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869

【解】  $t_{0.05,5} = 2.015$        $t_{0.25,5} = 2.571$



## 利用 $t$ 分布之机率分布表查 $t$ 值

- 例5：当  $n=20$  及  $\alpha=0.01$  时，请利用  $t$  分布表找出  $t_{\alpha}$  及  $t_{\alpha/2}$  之值。

$v$	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
...			...				
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850

【解】  $t_{0.01,19} = 2.539$        $t_{0.005,19} = 2.861$



## 利用 $t$ 分布之机率分布表查 $t$ 值

- 例6：当  $n=62$  及  $\alpha=0.10$  时，请利用  $t$  分布表找出  $t_\alpha$  及  $t_{\alpha/2}$  之值。

$v$	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
...				...			
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
$\infty$	1.282	1.645	1.960	2.326	2.576	3.090	3.291

【解】  $t_{0.10,\infty} = 1.282$        $t_{0.05,\infty} = 1.645$

# 群体平均数之区间估计

- 例7：某印刷电路板厂商之品管人员，欲对镀铜生产线进行品检，他随机抽取**30**片印刷电路板并量得铜厚资料（如下表），**求此生产线平均铜厚之95%的信赖区间**。

单位： $\mu\text{m}$

123	119	122	121	120
120	121	121	124	118
122	120	125	123	117
121	120	124	122	119
119	118	122	123	121
120	117	119	122	121

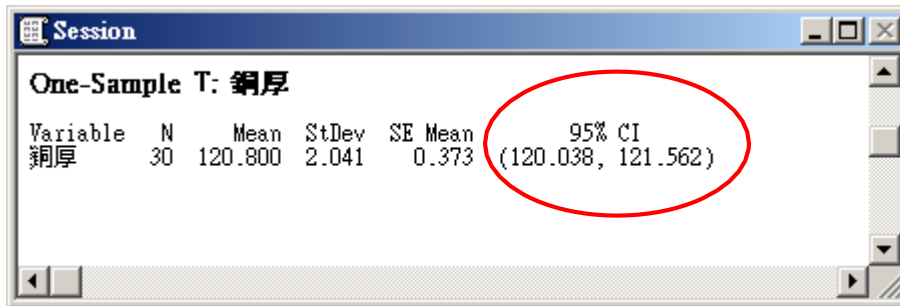
# 群体平均数之区间估计

【解】：由上述例题可求得  $\bar{X} = 120.8$      $t_{0.025, 29} = 2.045$      $S = 2.041$      $n = 30$

则可计算铜厚平均数的信赖区间为：

$$\bar{X} \pm t_{\alpha/2, n-1} \times \frac{S}{\sqrt{n}} = 120.8 \pm 2.045 \times \frac{2.041}{\sqrt{30}} = (120.038, 121.562)$$

【Minitab 报表】



Session

One-Sample T: 铜厚

Variable	N	Mean	StDev	SE Mean	95% CI
铜厚	30	120.800	2.041	0.373	(120.038, 121.562)

# 群体比率值之区间估计

- 群体比率值  $P$  之  $(1-\alpha)100\%$  信赖区间

$$\hat{P} \pm Z_{\frac{1-\alpha}{2}} \times \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

- 1) 公式中之  $z$  值称为临界值(**critical value**)。
- 2)  $(1-\alpha)$  称为信赖系数(**confidence coefficient**)。
- 3)  $(1-\alpha) 100\%$  称为信赖水平。

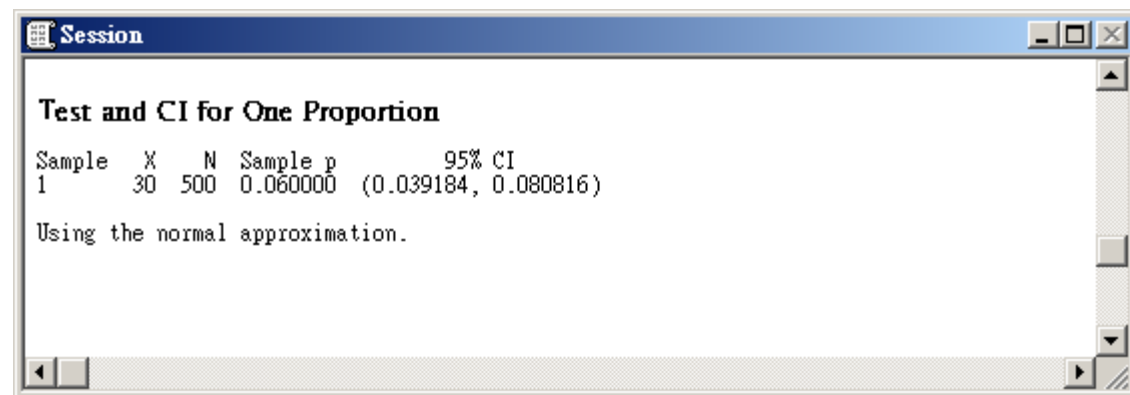
# 群体比率值之区间估计

- 例8：台北市政府希望估计该市目前之劳工失业率。假设由台北市劳工中随机抽取**500**人进行调查，发现失业者有**30**人。试求台北市失业率的**95%**信赖区间。

【解】  $\hat{P} = \frac{30}{500} = 0.06 \quad n = 500$

$$0.06 \pm (1.96) \times \sqrt{\frac{(0.06)(0.94)}{500}} = 0.06 \pm 0.021 = (0.039, 0.081)$$

【Minitab 报表】



The image shows a screenshot of the Minitab Session window. The title bar is labeled 'Session'. The main text area displays the results of a 'Test and CI for One Proportion'. It includes a table with columns for Sample, X, N, Sample p, and 95% CI. The data row shows Sample 1, X=30, N=500, Sample p=0.060000, and 95% CI=(0.039184, 0.080816). Below the table, it states 'Using the normal approximation.'

Sample	X	N	Sample p	95% CI
1	30	500	0.060000	(0.039184, 0.080816)

Using the normal approximation.

# 群体变异数之区间估计

- 群体变异数  $\sigma^2$  之  $(1-\alpha)100\%$  信赖区间

$$\frac{(n-1)s^2}{\chi_{\alpha/2}^2} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi_{(1-\alpha/2)}^2}$$

- 1) 公式中之卡方  $\chi^2$  值称为临界值(**critical value**)。
- 2)  $(1-\alpha)$  称为信赖系数(**confidence coefficient**)。
- 3)  $(1-\alpha) 100\%$  称为信赖水平。

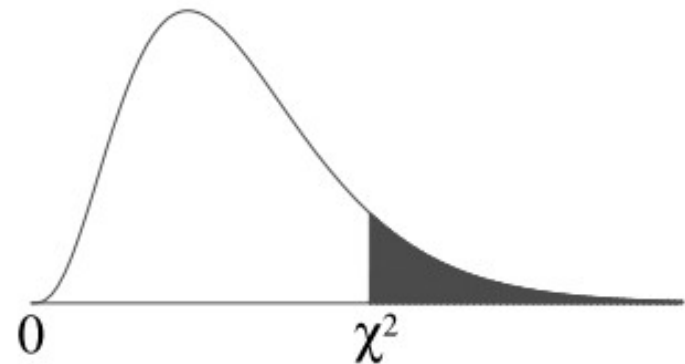
# 卡方分布 ( $\chi^2$ Distribution 或 Chi-Square Distribution)

- 由一常态群体中，以放回取样方式，随机抽取样本大小为 **n** 之一组样本，则

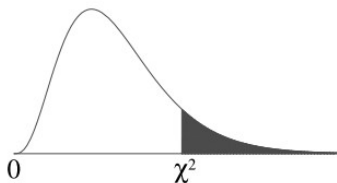
$\frac{(n-1)S^2}{\sigma^2}$  会服从 卡方分布( $\chi^2$  Distribution)，

其自由度 (Degrees of freedom, d.f.)  $\nu = n-1$ 。

$$\text{即：} \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$



# 卡方分布之机率分配表

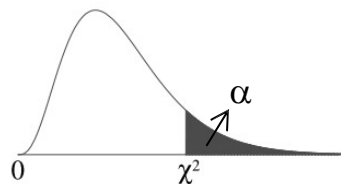


$df$	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	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



# 群体变异数之区间估计

- 例9：若 $n=20$ ，利用下表求  $\chi_{0.05}^2 = ?$



$df$	$\chi_{.995}^2$	$\chi_{.990}^2$	$\chi_{.975}^2$	$\chi_{.950}^2$	$\chi_{.900}^2$	$\chi_{.100}^2$	$\chi_{.050}^2$	$\chi_{.025}^2$	$\chi_{.010}^2$	$\chi_{.005}^2$
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

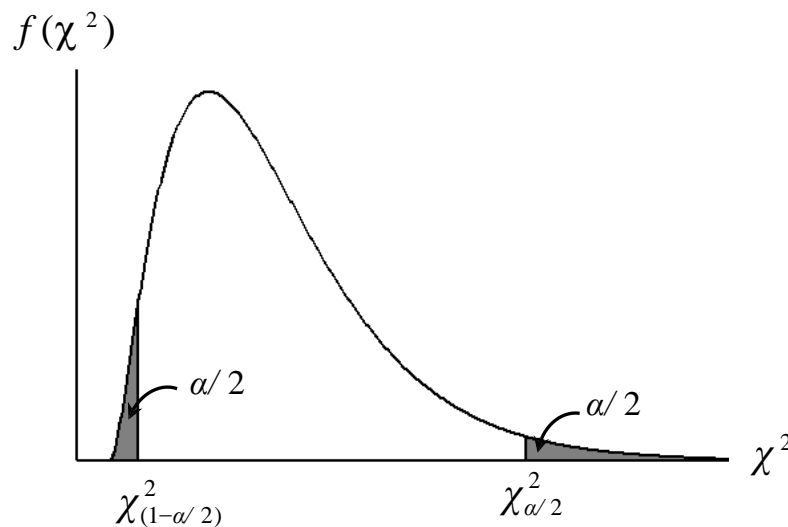
【解】  $\chi_{19,0.05}^2 = 30.144$

# 群体变异数之区间估计

- 群体变异数  $\sigma^2$  之  $(1-\alpha)100\%$  信赖区间

$$\frac{(n-1)s^2}{\chi_{\alpha/2}^2} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi_{(1-\alpha/2)}^2}$$

示意图：



# 群体变异数之区间估计

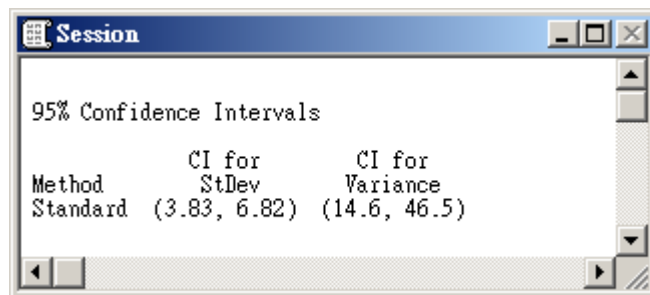
- **例10**：某饮料制造厂之检验人员，欲对填充饮料之生产线进行品检，他随机抽样**25**罐饮料，并量得其平均容量为**18.4cc**，标准差为**4.9cc**。试求此生产线饮料容量变异数之 **95%** 信赖区间。

【解】：已知  $S = 4.9$ ， $n = 25$ ，

饮料容量变异数之 95% 信赖区间为：

$$\frac{(n-1)s^2}{\chi_{\alpha/2}^2} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi_{(1-\alpha/2)}^2} \Rightarrow \frac{(25-1)(4.9)^2}{\chi_{0.05/2}^2} \leq \sigma^2 \leq \frac{(25-1)(4.9)^2}{\chi_{(1-0.05/2)}^2}$$
$$\Rightarrow \frac{576.24}{39.3641} \leq \sigma^2 \leq \frac{576.24}{12.4011} \Rightarrow 14.64 \leq \sigma^2 \leq 46.47 \Rightarrow 3.83 \leq \sigma \leq 6.82$$

【Minitab 报表】



95% Confidence Intervals		
Method	CI for StDev	CI for Variance
Standard	(3.83, 6.82)	(14.6, 46.5)

## 第四部份：样本大小之决定

# 如何决定样本 **n** 之大小？

估计  $\mu$  时，若估计误差之范围在  $\pm e$  之内，信赖系数为  $(1 - \alpha)$ ，则 **n** 可由下式得之：

$$n = \frac{(z_{(1-\alpha)/2} \cdot \sigma)^2}{e^2}$$

※当  $\sigma$  未知时，以 **S** 代之。

# 如何决定样本 $n$ 之大小？

- **例11**：某大学教师欲以信赖区间方式来估计学生每天上网的平均时间（小时），他希望估计误差最多为**1.5**小时，信赖水准为**95%**，则需要多少的样本数？假设标准差  $\sigma=3$ 。

【解】

$$n = \frac{(z \times \sigma)^2}{e^2} = \frac{(1.96 \times 3)^2}{1.5^2} \cong 16 ,$$

即需要抽样至少16个学生。

# 如何决定样本 **n** 之大小？

估计 **P** 时，若估计误差之范围在  $\pm e$  之内，信赖系数为  $(1 - \alpha)$ ，则 **n** 可由下式得之：

$$n = \frac{(z_{(1-\alpha)/2})^2 \cdot p \cdot q}{e^2}$$

※若p与q无法事先估出，则以p=q=0.5 代入n之公式。

## 如何决定样本 $n$ 之大小？

- **例12**：某公司欲了解员工对某一新制度的喜好程度，特进行问卷调查并希望以调查结果来估计员工对此新制度喜好比例之**95%**信赖区间，并希望估计误差在 $\pm$  **0.08**之内，试问需要多少样本？

【解】

$$n = \frac{Z^2 \times P \times q}{e^2} = \frac{1.96^2 \times 0.5 \times 0.5}{(0.08)^2} = 150.06 \cong 151 ,$$

即需要抽样至少151个员工。



本单元结束

## 第五单元 简单回顾

# 简单回顾

- 抽样分布
- 中央极限定理
- 点估计
- 区间估计：
  - 群体平均数  $\mu$  之  $(1-\alpha)$  100% 信赖区间
    - 当  $\sigma$  已知时： $X \pm Z_{\frac{1-\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$
    - 当  $\sigma$  未知且为大样本时： $X \pm Z_{\frac{1-\alpha}{2}}(\frac{s}{\sqrt{n}})$
    - 当  $\sigma$  未知时： $X \pm t_{\alpha/2, n-1}(\frac{s}{\sqrt{n}})$
  - 群体比率值  $P$  之  $(1-\alpha)$  100% 信赖区间
  - 群体变异数  $\sigma^2$  之  $(1-\alpha)$  100% 信赖区间

# 简单回顾

- 样本大小之决定：
  - 估计  $\mu$
  - 估计  $\mathbf{P}$