

统计简单学

授课教师：唐丽英 教授

新竹交通大学
工业工程与管理学系

课程简介

- 统计方法无所不在

- 不论产品开发、制程管制、市场营销、财务金融或人文社会方面的研究，均需藉助统计方法分析资料。
- 一般人在解释统计数字或统计分析结果时，常因观念不清或方法不熟悉而得到错误结论。

- 统计方法一点也不难

- 许多人认为统计方法就是一堆数学公式和繁复的计算，因此对于学习统计课程产生恐惧。
- 本课程将以深入浅出的方式及多元的实务例题，搭配统计软件的报表解读，让学习者不需要记忆繁复的公式，也能轻松地学会统计基本观念及方法，并运用于各种领域或工作场合。

- 本课程之设计

- 乃是针对初学者，因此学习本课程者不需具备任何统计背景，仅须具备基本高中数学能力即可。

课程单元介绍

叙述统计

第一单元	统计基本概念介绍	包括：统计学的专有名词、范围与抽样方法等
第二单元	常用统计量	如：学生身高的平均值、标准偏差
第三单元	常用统计图表	点图、盒型图、直方图、条图、单圆图及柏拉图
第四单元	常态机率分布	

推论统计

第五单元	估计	分为点估计及区间估计
第六单元	假说检定	推论一或两组样本之母体及特定值间是否有显著差异
第七单元	变异数分析	推论三组或以上之母体及特定值是否有显著差异
第八单元	回归分析	

【统计简单学】

第一单元

统计基本概念介绍

授课教师：唐丽英 教授

新竹交通大学
工业工程与管理学系

第一单元 内容大纲

- 第一部份：统计学简介
 1. 统计学中常用的术语
 2. 统计学的范围
 3. 统计学的工作流程
- 第二部份：随机变数的种类
 1. 质变数－类别变数
 2. 量变数－数值变数
- 第三部份：数据之取得方式
 1. 简单随机抽样
 2. 系统抽样
 3. 分层随机抽样
 4. 部落抽样

第一部份：统计学简介

统计学中常用的术语

(1) 统计学(Statistics)

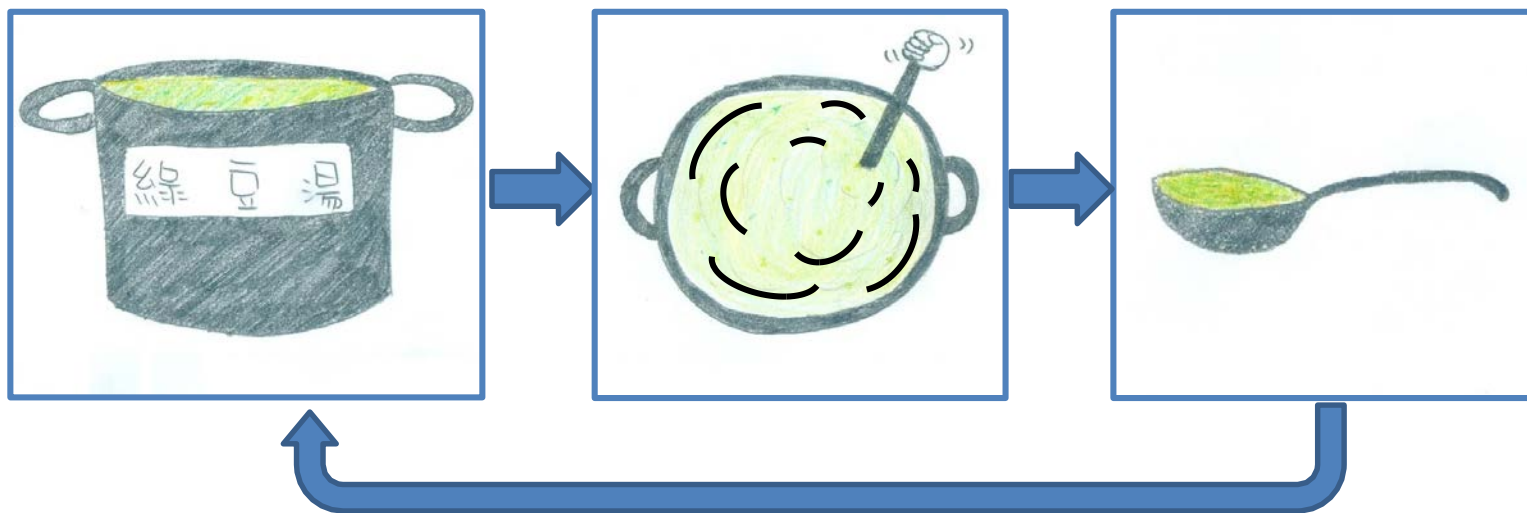
统计学为搜集、整理、展示、分析、解释资料，并由样本推论群体，使在不确定情况下作成决策的科学方法。

(2) 群体(population)

由具有共同特性之个体所组成的整体。

(3) 样本(sample)

群体之一部分。



统计学中常用的术语

(4) 参数(parameter)

由群体资料所计算出之群体表征值。例： μ 、 σ 、 P

读音： μ **mu** 读音： σ **sigma**

(5) 统计量(statistic)

由样本资料所算出之样本表征值。例： \bar{X} 、 S 、 \hat{P}

读音： \bar{X} **X bar** 读音： \hat{P} **P hat**

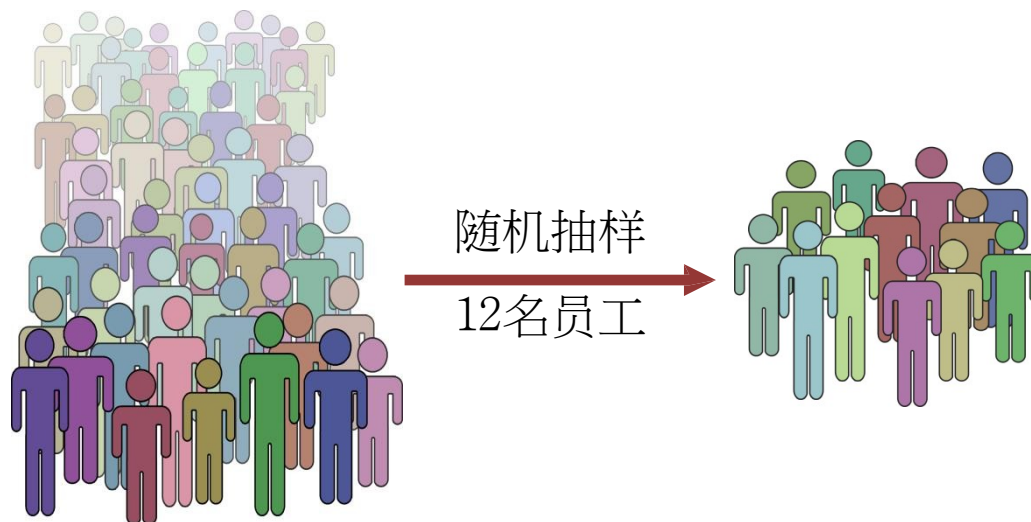
(6) 统计学的目的

由样本来推论群体之参数。

例题

例

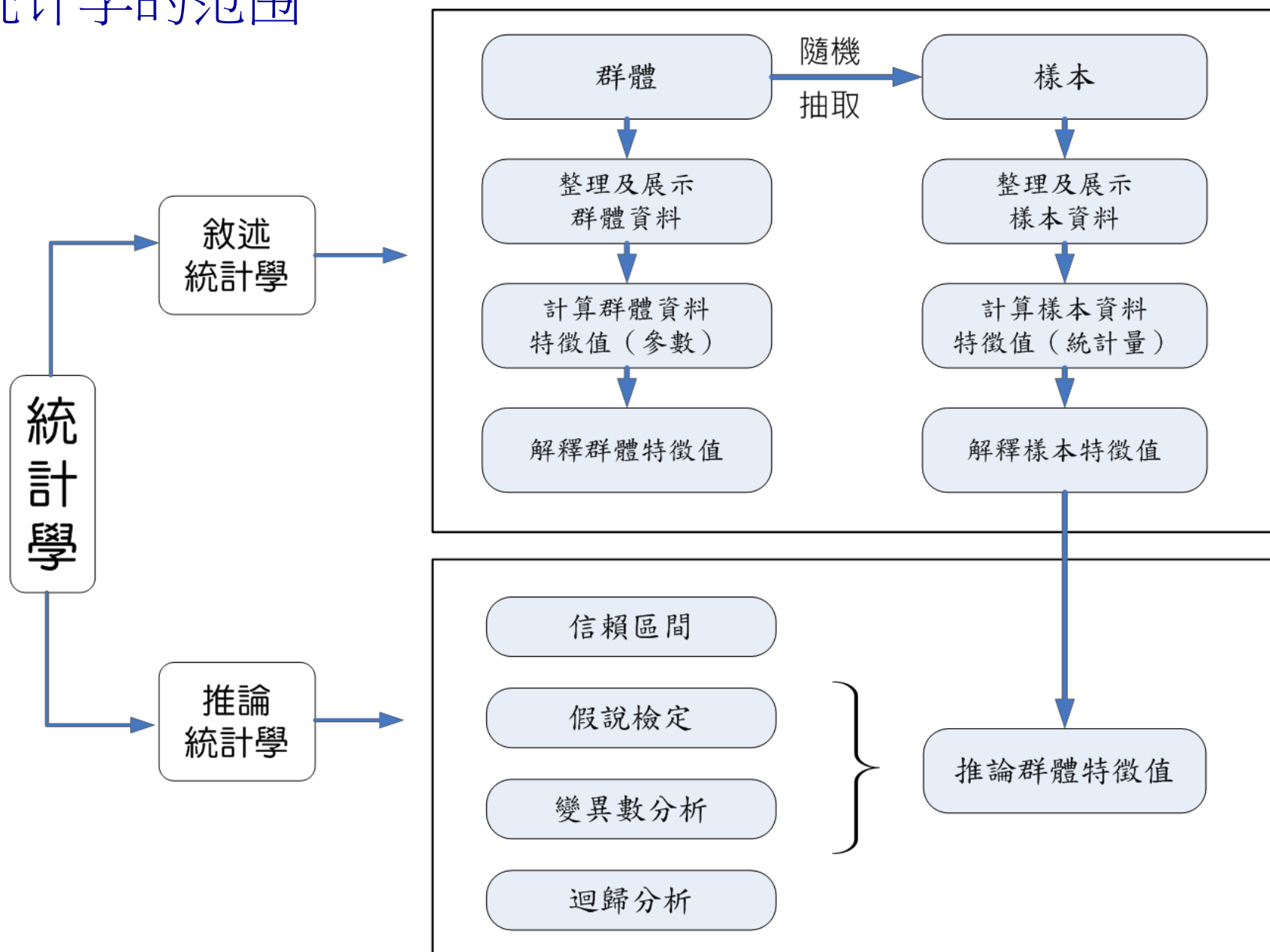
某人力资源部门，欲由12个随机抽出之员工来估计该公司全体员工一年内之请假天数。请指出本例之群体、样本、参数及统计量。



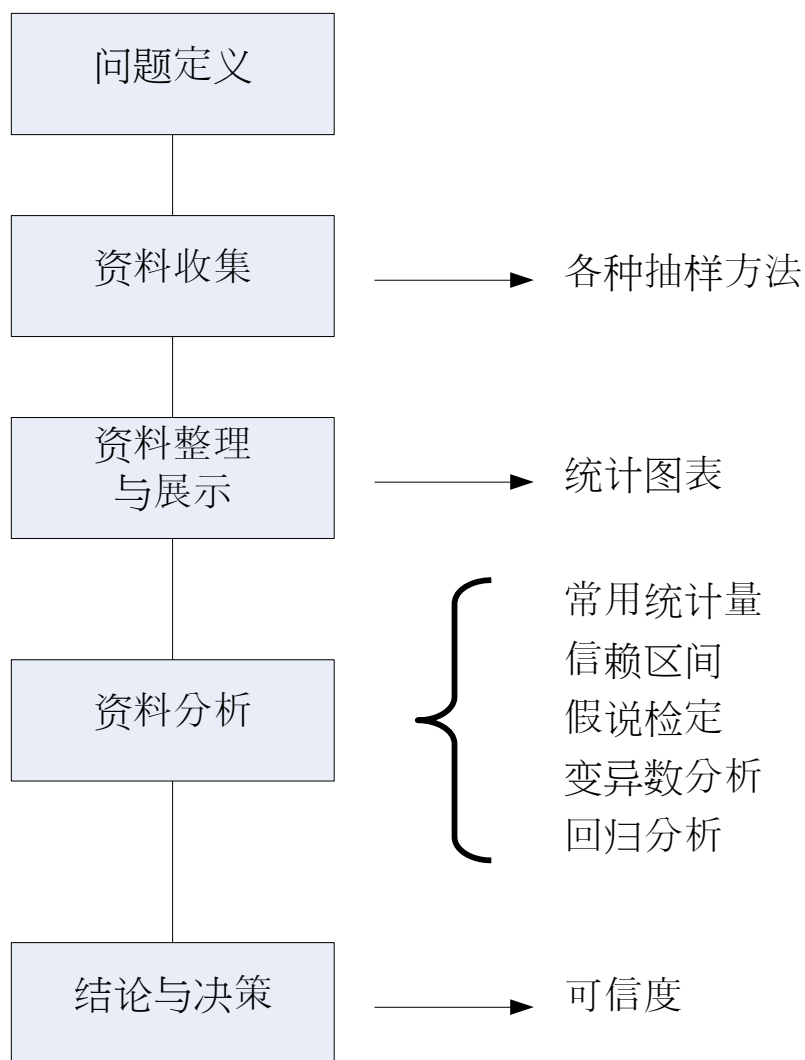
统计学的范围

- 统计学可分为**叙述统计(Descriptive Statistics)**及**推论统计(Inferential Statistics)**两部份。
 - **叙述统计**：包含如何搜集数据、展示数据、及找出可描述数据特征之值的方法。
 - **推论统计**：包含如何由样本资讯来推论群体，并估计该推论之可信度大小的方法。

统计学的范围



统计学的工作流程



第二部份：随机变数的种类

随机变数的种类

随机变数(Random Variable, R.V)

- 研究者对所欲研究之群体所感兴趣的一项或多项特质，即为随机变数。

随机变数之种类：

1. 质变数，定性变数或类别变数(Qualitative R.V ; Categorical R.V)

- 随机变量的各结果不以数量表示，而依其特性之类别表之。
- 例如：①性别 ②职别 ③造成业务延迟结案之可能原因。

随机变数的种类

2. 量变数，数值变数(Quantitative R.V ; Numerical R.V)

➤ 随机变量的各结果可以数量表之。

1) 离散型(Discrete)：经由计数的方式取得数据。

- 例如：①不良品个数 ②一公司之员工人数 ③一本书上的错字个数。

2) 连续型(Continuous)：经由量测的方式取得数据。

- 例如：①物品的重量 ②员工生产力 ③个人的血压值。

※离散型量变量之可能值：**0**或整数。
连续型量变量之可能值：任何实数。

质变数与量变数的例子

例 某医院护理部门，将病人的信息纪录如下表，请指出各属性字段为何种随机变量？

病人编号	性别	血型	身高	体重	就诊次数
S001	男	A	175.2	69.3	2
S002	女	B	158.6	42.5	1
S003	女	O	162.4	51.2	4
S004	男	AB	170.3	65.4	3
S005	男	B	180.5	90.6	2
...



质变数
(类别资料)



质变数
(类别资料)



质变数
(类别资料)



量变数
(连续型)



量变数
(连续型)



量变数
(离散型)

第三部份：数据之取得方式

普查与抽样

- 普查
 - 对群体中的每一个体取得资料，也就是100%的全检。
- 抽样（**Sampling**）
 - 利用一种程序或方法，由群体中抽出样本。
 - 常用的抽样方法有下列四种：
 - 简单随机抽样
 - 系统抽样
 - 分层随机抽样
 - 部落抽样

简单随机抽样

- 是指群体中每一个体被抽中之机会均等。
- 对群体内的每个个体编号，再以随机数表、计算机仿真随机数或制作纸签的方法决定欲抽出之样本。
- 优点：取样方法简便。
- 缺点：有时会因抽到的样本过于集中某部分之群体，造成样本之代表性不足。

系统抽样

- 只做第一次随机抽样后，就依固定间隔数抽出一样本。
- 例如：群体有30个样本，预计抽出6个样本，所以每隔5个即抽出一样本。
- 优点：抽出第一个种子号码，以后仅需每间隔数个样本抽样即可。
- 缺点：样本在编号排序时必须与研究者所关心的变项无关，否则会造成样本之代表性不足。

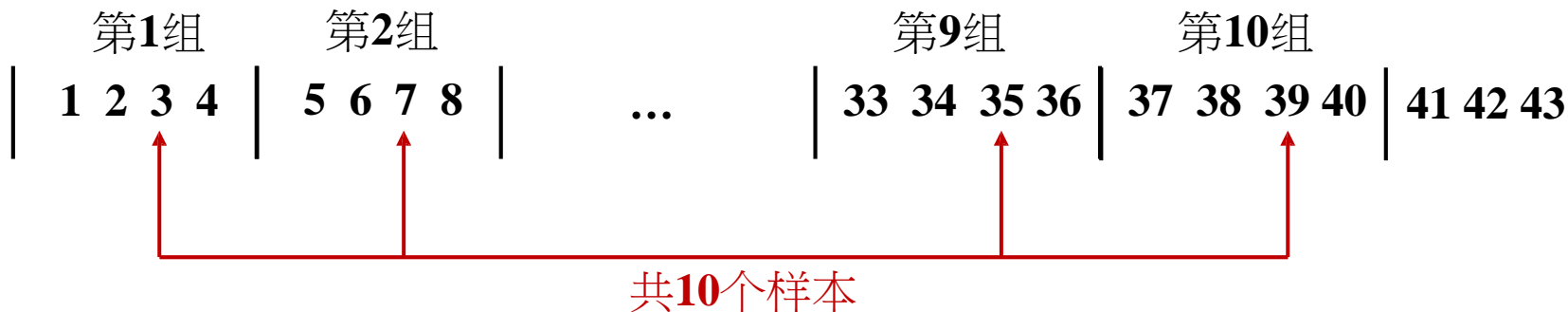
例：假设群体含 **43** 个项目，预计抽取**10**个样本，请以系统抽样方法进行抽样。

1) 编号

2) $k = 4$ (取 $43/10$ 之最大整数)，将所有**43**个项目，依顺序每 **4** 个分一组。

3) 随机抽取 **1~4** 中的一个数字，例如 **3**

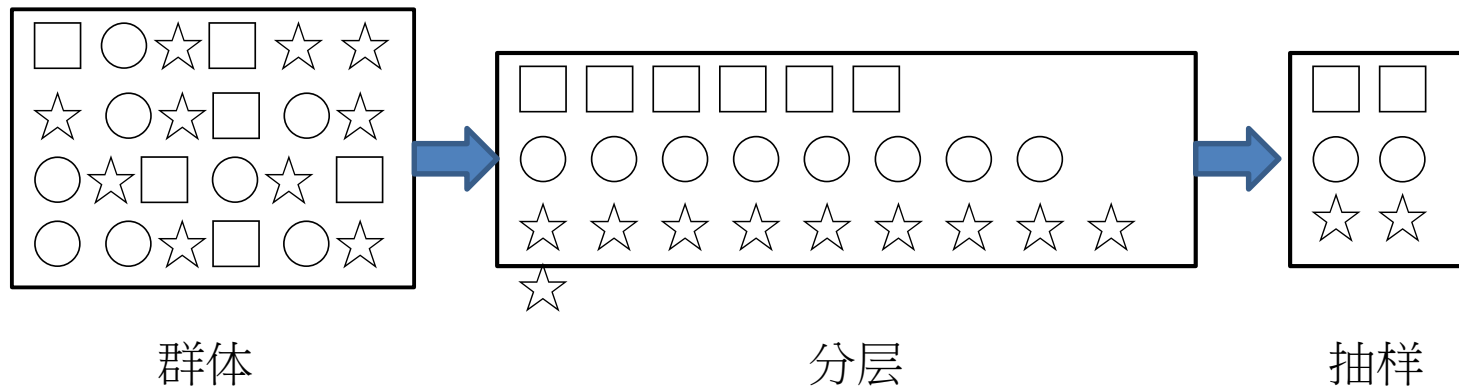
4) 从编号**3**开始，间隔**4**个项目抽取一个样本，即可获得**10**个系统抽样的样本组



分层随机抽样

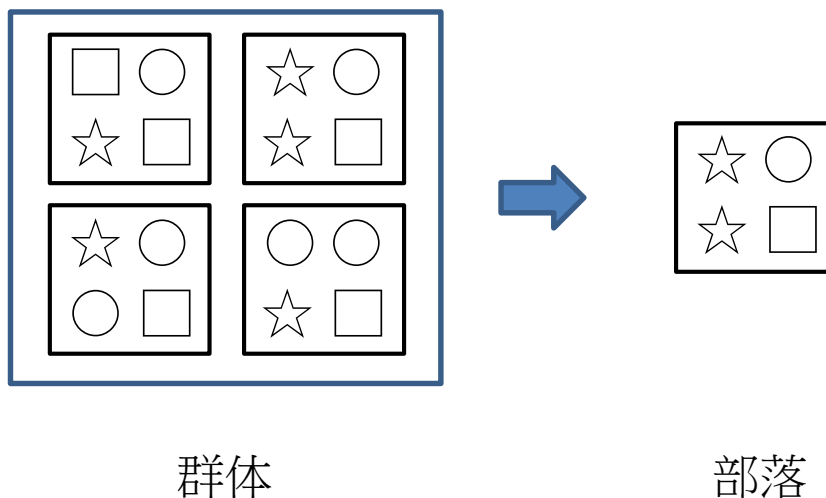
- 群体依某一衡量标准分成数个不重迭的子群（称为层）。
- 将群体分层之后，再从每一层中利用简单随机方式取样，即为分层随机抽样。
- 分层随机抽样之原则是同层内的性质差异要小，而不同层间之差异则要越大越好。

例：



部落抽样

- 常用在群体中之个体分离很远，且很难搜集到样本时。
- 部落抽样是将群体先分成几个部落，再从这些同一个部落中抽出一个或数个部落进行普查。
- 部落抽样是假设每一个部落都是群体的缩影，因此不同部落间个体的性质差异性要小，而部落内个体的性质差异性要大。



本单元结束

第一单元 简单回顾

简单回顾

- 统计学常用术语：
 - 群体
 - 样本
 - 参数
 - 统计量

简单回顾

- 随机变数的种类：
 - 质变数(又称定性变数或类别变数)
 - 量变数(又称数值变数)
 - 离散型
 - 连续型
- 数据取得之方式：
 - 普查
 - 抽样
 - 简单随机抽样
 - 系统抽样
 - 分层随机抽样
 - 部落抽样