

# 统计数据及其分类



南京财经大学  
NANJING UNIVERSITY OF FINANCE & ECONOMICS

# 分类数据

## categorical data

分类数据是对事物进行分类的结果，数据的主要特征是采用文字、数字的代码和其他符号对事物进行简单的分类和分组。

分类数据以定类尺度（nominal scale）来衡量。



# 顺序数据

## rank data

即等级数据，是对事物进行分类的结果，这些分类在语义上表现出明显的等级或顺序关系，例如，学生成绩可以分为优秀、良好、中等、及格和不及格。

顺序数据以定序尺度（ordinal scale）来衡量。



# 数值型数据

metric data

数值型数据是使用自然或度量衡单位对事物进行测量的结果，其结果表现为具体数值。

数值型数据有两种衡量尺度：

一是定距尺度 (interval scale)

二是定比尺度 (ratio scale)

区别：“零点”



# 时间序列数据

time series data

时间序列数据是在不同时间上收集到的数据，它所描述的是现象随时间而变化的情况。



# 截面数据

## cross-sectional data

截面数据是在相同或近似相同的时间点上收集的数据，它所描述的是现象在某一时刻或某一时间段的变化情况。



# 面板数据

# panel data

面板数据是对若干个单位在不同时间进行重复跟踪调查所形成的数据。



# 绝对数

## absolute number

绝对数是统计数据的基本表现形式，是其他指标形式形成的基础。现象的总体规模和水平一般都以绝对数形式表现，一个地区的总人口、国内生产总值、货物周转量等都是绝对数。





# 相对数

## relative number

相对数是由两个相互联系的绝对数对比得到，反映事物的相对数量。常用的相对数包括：结构相对数、动态相对数、比较相对数、计划完成相对数。



## 平均数 average

平均数反映现象总体的一般水平。如职工的平均工资。



# 变量



变量 (*variable*) 总是可以用数值形式表现的, 可变的数量标志和所有的统计指标都是变量。

分类: 自变量与因变量

离散型变量、连续型变量

确定性变量、随机变量



# 离散型变量

discrete variable

离散型变量一般用“点数”的方法取得，数值往往是整数，如职工人数、企业个数、设备台数等。



# 连续型变量

## continuous variable

连续型变量用“测量”取得，测量的仪器越精密，变量值就会越精确，如人的身高、体重；产品的长度、重量等。



# 自变量与因变量

**Independent variable**

**Dependent variable**

自变量是自身变化会引起其他变量变化的量；因变量是受其他变量影响而变化的量。

如一个国家的消费水平受这个国家的生产水平的影响，因此，人均国内生产总值是自变量，人均消费金额是因变量。



# 确定性变量

## continuous variable

受确定性因素影响的变量称为确定性变量。这种影响变量值变化的因素是明显的，是可以解释的，是人为的或者受人控制的，其影响变量值的大小、方向都可以确定。



# 随机变量

## continuous variable

受随机性因素影响的变量称为随机变量，所谓随机因素，是指各种不确定的、偶然性的因素，这种因素对变量影响的大小和方向都是不确定的，通常是微小的。

