

统计与统计学

统计：总括起来计算，是对某一现象或事物的有关资料进行搜集、整理、计算、分析的工作过程。

统计有三层含义：1 统计资料 (statistic)，即反映各种现象的数据资料；2 统计工作 (statistic)，即具体搜集、整理、分析统计资料的工作过程；属于统计实践的过程 3 统计学 (statistics)，即研究统计原理与方法的科学。为这种实践提供理论与方法的学问

统计学：依据其作用分类 1 应用统计学：与研究对象的特征密切结合的各科专门统计，如人口、医药、经济、工业、体育、生物等，以本学科现象的数量形式为基础，对它们的规律性进行数量上的分析研究。教育统计学亦属于应用统计学的范畴 2 数理统计学：为各门应用统计学提供数理方法基础的一门学科，其内容主要是运用概率的知识解释统计数据数量关系的模式。

教育统计

教育统计学 (educational statistics)：专门研究如何搜集、整理、分析在心理和教育方面对实验或调查所获得的数字资料，如何根据这些资料所传递的信息，进行数学推论，找出客观规律的一门学科。简言之，教育统计学是运用统计学的一般原理和方法研究教育科学领域数量关系的一门科学。

教育统计工作：各级教育行政部门的教育统计报表和相应制度，它作为教育行政部门管理系统的一部分，着重于统计数据的收集和管理的工作。

统计学的内容

多元统计分析：进行教育或心理的研究和实验，常会受到多种因素的影响。而寻找主要的影响因素，把相近或相关的因素合并或归类则是多元统计分析的主要任务。多元分析增强了心理与教育研究和实验的可靠性与准确性，使研究更贴近客观实际。

多元分析的理论深奥，计算复杂，因此在很长一段时间其应用受到限制。直到计算机技术发展普及的今天，有了专门成熟的统计软件包，才使得多元统计分析法获得广泛的应用。

多元统计分析的主要内容有主成分分析、因素分析、聚类分析、多元方差分析、多元回归分析等等。

模糊统计学：教育、心理现象的不确定性，不仅表现为随机性，而且更多地表现为模糊性。因此，对心理与教育现象进行模糊性的描述也是必然的，不可缺少的。

预备知识：变量与数据

变量 (variable)：又称随机变量 (random variable) 数学中把不断变化的，可取不同值的量称为变量；统计学则因观测数据具有变异性的特点（即数据总是在一定的时空范围内不断变化而具有变异性）称其为变量（即，具有变异性的数据）。如：学生身高。

数据 (data) 与变量：在统计学中把在取值之前不能预料到将取什么值的量称为变量（随机变量）；如果某个数值一旦被取定，则称这个数值为随机变量的一个观察值，即数据。

变量的分类

1 名称变量 (nominal variable)，指一事物与其他事物在属性、类别上不同；如：性别、颜色。名称变量的统计方法主要有次数计算、百分比、 χ^2 检验、相关等。

2 顺序变量 (o r d i n a l) : 事物的某一属性的多少或大小按顺序排列起来的变量；如：教师按能力大小或成绩高低排列等级。顺序变量的常用统计方法有中位数、百分位数、等级相关、秩次检验等。

3 等距变量 (i n t e r v a l) : 数据具有相等的单位，但没有绝对零点；如：能力测验或知识测验。等距变量常用的统计方法有平均数、标准差、t检验、F检验、积差相关等。

4 比率变量 (r a t i o n a l) : 一种既有相等的单位，又有绝对零点的变量，又称等比变量；如人的身高与体重、距离、时间、教育投资、学校固定资产金额等。比率变量常用的统计方法除了上述方法外，还有几何平均数、相对差异量数等。

预备知识：统计术语初步

总体与个体

总体 (p o p u l a t i o n) : 客观存在的，并在同一性质的基础上结合起来的许多个别单位的整体，即具有某一特性的一类事物的全体，又叫母体或全域。统计总体具有同质性、大量性和差异性的特征，只有同时具备这三个特征才能形成真正意义上的统计总体。

个体 (i n d i v i d u a l) : 构成总体的基本单位或单元，又称元素或个案。

样本与样本容量

样本 (s a m p l e) : 从总体中抽出的一部分个体。

样本容量：样本中所包含的个体数目，或称样本含量，用符号 N 或 n 表示。大样本 (l a r g e s a m p l e) 的容量一般在 30 以上或 40 以上，小样本 (s m a l l s a m p l e) 的容量在 30 以下。

参数与统计量

参数 (p a r a m e t e r) : 表示总体特征的量数，一般用希腊字母表示，如总体平均数 μ (m u)，总体标准差 (s i g m a)，总体相关系数 (r h o) 等。字符表示，如，样本平均数 \bar{X} 或 M ，样本标准差 S ，样本相关系数 r 等。

标志 (m a r k) : 统计总体中各个个体共同具有的属性和特征，它是说明个体属性和特征的名称，如个体的性别、文化程度、年龄、职业、民族、个体的记忆力、思维力、想象力。

品质标志：表明个体属性特征，不能用数量说明，只能用文字说明，如性别。

数量标志：表明个体数量特征，用数值表示，如年龄。

指标 (i n d e x) : 表明总体数量特征的概念和具体数值，又称统计指标，它是把各个个体的特征汇总起来的综合结果，必须用数量表示。作为一个统计指标必须具备五个要素：指标名称、计量单位、时间、地点、数量。如某大学 2005 年本科生招生人数 6000 其中指标名称是本科生招生人数，计量单位是 1 人，时间是 2005 年，地点是某大学，数量是 6000。

绝对数与相对数

绝对数 (a b s o l u t e) : 用来表明在一定时间、地点条件下某种教育、心理现象的总体规模和发展水平的统计指标，又称总量指标。绝对数量能具体反映各种事物的实在数量，在统计工作中具有非常重要的意义。

相对数 (r e l a t i v e) : 教育与心理现象中两个单位相同的互相关系的指标数值的比率，它比绝对数更清晰地反映现象之间的对比关系，能使个别现象从绝对数的差别中抽象出来，把不能直接比较的现象变成可比较的现象。

SP 简介

是“社会科学统计软件包”(Statistical Package for the Social Sciences)的简称，是一种集成化的计算机数据处理应用软件。SP是世界上公认的三大数据分析软件之一(SAS、SPSS、TAB)。伴随SP服务领域的扩大和深度的增加，SP公司已于2014年正式将英文全称更改为Statistical Product and Service Solutions，意为“统计产品与服务解决方案”，标志着SP的战略方向正在做出重大调整。

发展历程

1978年 斯坦福大学三位学生创建SPSS
1978年 诞生第一个用于大型机的统计软件；
1979年 在芝加哥成立SP总部；
1984年 推出用于个人电脑的SPSS / PC+
1992年 推出WINDOWS版本，同时全球自SPSS 6.1起，SP全称为“Statistical Product and Service Solutions”
2009年 SP公司宣布重新包装旗下的SP产品线，定位为预测统计分析软件(Predictive Analytics)，包括三部分：统计分析、数据挖掘、数据收集、企业应用服务；
2014年 随着SP公司被IBM并购，各子产品家族名称前面不再以PA为前缀，修改为统一加上IBM字样。SAS




SP统计功能概述

基础统计：包括描述性统计、探索性统计、列联表分析、线性组合测量、t检验、单因素方差分析、多维反应模型分析、线性回归分析、相关分析、非参数检验等。

专业统计：包括判别分析、因子分析、聚类分析、距离分析、可靠性分析等。

高级统计分析：包括Logit回归分析、多变量方差分析、重复测量方差分析、多协变量方差分析、非线性回归、Probit回归分析、Cox回归分析、曲线估计等。

SP统计功能概述

Analyze	Direct Marketing	Graphs
Reports		►
Descriptive Statistics		►
Tables		►
Compare Means		►
General Linear Model		►
Generalized Linear Models		►
Mixed Models		►
Correlate		►
Regression		►
Loglinear		►
Neural Networks		►
Classify		►
Dimension Reduction		►
Scale		►
Nonparametric Tests		►
Forecasting		►
Survival		►
Multiple Response		►
 Missing Value Analysis...		
Multiple Imputation		►
Complex Samples		►
 Simulation...		
Quality Control		►
 ROC Curve...		
IBM SPSS Amos...		

Reports: 生成数据报告
 Descriptives Statistics: 描述性统计
 Tables: 生成统计表格
 Compare Means: 均值比较分析
 General Linear Model: 一般线性模型
 Generalized Linear Model: 广义线性模型
 Mixed Models: 混合模型
 Correlate: 相关分析
 Regression: 回归分析
 Loglinear: 对数线性模型
 Neural Networks: 人工神经网络
 Classify: 聚类与判别分析
 Dimension Reduction: 降维分析
 Scale: 测度分析
 Nonparametric Tests: 非参数检验
 Forecasting: 预测分析
 Survival: 生存分析
 Multiple Response: 多维反应模型分析
 Missing Value Analysis: 缺省值分析
 Multiple imputation: 多重插补分析
 Complex Samples: 复杂样本分析
 Simulation: 模拟分析
 Quality Control: 质量控制分析
 ROC Curve: ROC曲线
 IBM SPSS Amos: Amos分析

数据的搜集、整理与表达

统计工作过程的三大相互依存的基本环节

数据的搜集不仅是统计工作的首要环节，而且还是全部统计工作的基础。数据的来源：

经常性资料：日常工作中的记录和统计报表等。工作记录，如学校中关于教师资历、教学情况、教学进度的记录，关于学生学期、学年成绩和操行评定的记录，关于学生健康状况、流失情况的记录，关于学校经费的使用、学校建设的记录，教师的工作记录等；统计报表，如教师职称变动的年报表、学生流失率的季报表、学生入学率的年报表等。

一时性资料：在一定时限内所收集的有关问题的资料。主要来源于三个方面，即教育与心理调查、教育与心理测量、教育与心理实验。

数据的整理，可使数据变得简约直观，且显现出数据的分布特征。数据整理包括检查数据和整理方式。

检查数据：

- 1 准确性：观测数据的真实可靠；包括检查逻辑性和检查计算。
- 2 完整性：数据是否齐全、有无遗漏和重复。
- 3 及时性：数据是否按规定时期报送。

整理方式：

- 1 顺序分布法：将数据（X）按大小顺序先排列起来，然后用

次数 (f r e q u e n c y) 表示相同数据的出现数目，形成一个简单次数表。2 等级分布法：首先按顺序排列数据，然后标上相应的等级 (R a n k) , 可分无重复等级的整理方式和有重复等级的整理方式。3 次数分布法：针对数据量较大的情况。次数分布法是从一群数据中求其全距 (r a n g e) 决定组数、组距 (i n t e r v a l) 组限 (b o u n d a r i e s) , 然后进行归类汇总形成一个有规律的表式的数据整理方式，又称分组次数分布。

数据的表达，观测数据经过整理之后需要用一定的方式表达出来，包括列表法（即统计表）和图示法（即统计图）。

列表法（即统计表）：统计表 (s t a t i s t i c) 是以表格的形式表达统计资料数量关系的方式或工具。统计表可以简化资料，反映出事物的全貌及蕴含的特性，省去冗长的文字叙述，同时也便于分析、对比和计算。

统计表的基本内容：线条、标题、标目、数字。

统计表的种类

简单表 (s i m p l e) — “ 定语 ” 部分未经分组，只罗列单位名称、地点、时序等的表

分组表 (g r o u p i n g) — “ 定语 ” 按照一个分类标志进行分组的表格

复合表 (c o m p l e x) — “ 定语 ” 按照两个或两个以上的分组标志进行多层分组所形成的表格

图示法（即统计图）：统计图 (s t a t i s t i c) 是以几何图形和形象图形表示统计资料数量关系的工具。统计图在表现事物的数量关系时比统计表更为形象、直观，给人的印象更加深刻，而且易于记忆。

从图形形状上分有 直条图 (t h e r m o m e t e r) 饼图 (p a e t d i a g r a m) 直方图 (h i s t o g r a m) 折线图 (l i n e g r a m) 散点图 (s c a t t e r) d i a g r a m

从图形维度上分有二维图（即平面图）和三维图（即立体图）

从图形来源上分有几何图、象形图和统计地图

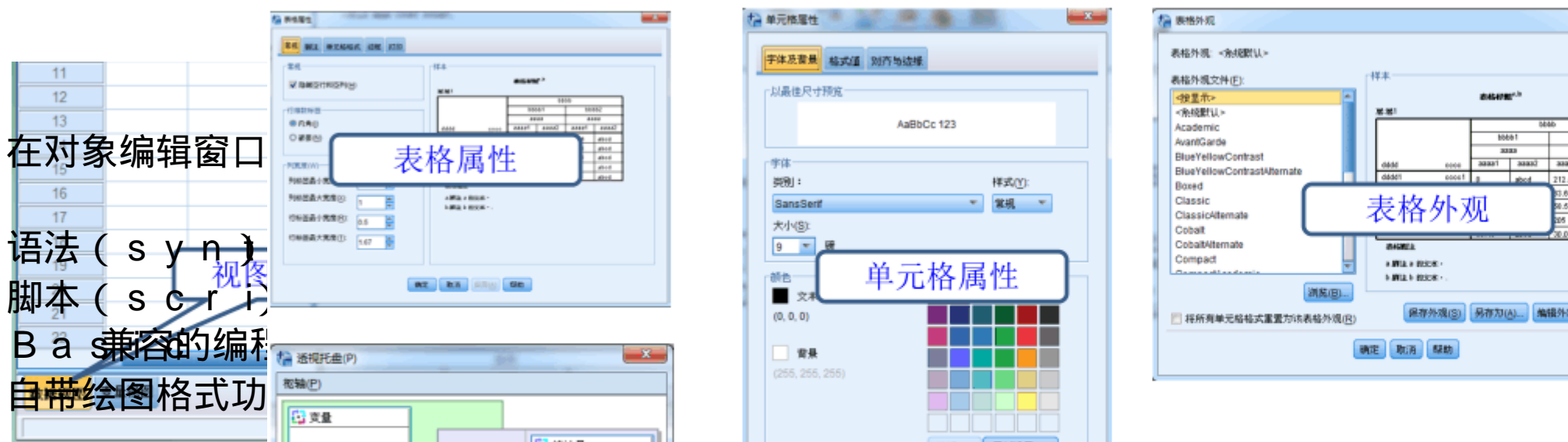
S P 的启动：启动 S P S 可双击桌面上 I B M P S S t a t i 图标；也可以在“开始”菜单中依次选择 A l l P r o g r a m s S P S S o f t w a r e S t a t i s t i c s

S P 的常用界面与窗口 (1)

数据编辑窗口功能：对 S P 的数据文件进行录入、修改、管理等基本操作的窗口



对象编辑窗口：
可打开该输出结



在对象编辑窗口

语法 (syntax)
脚本 (script)
兼容的编程
自带绘图格式功

数据输入：SPSS
结果

- 1 SPSS界面：打开SPSS为一空白数据编辑窗口“SPSS Data Editor”
- 2 定义变量：单击“Variable View”窗口，该视图的每行各表示对一个变量的定义，每列表示定义该变量时所用的某种属性，如名称（Name）

变量类型 (Type) 等。定义变量后，单击“Data”(数据视图)。

3 输入数据：输入变量的值，确认一行一列单元格为当前单元格。

4 保存数据：单击“File”，选择“Save”，若该数据从未被保存过，故弹出“Save Data”对话框。

数据整理

1 单元值的修改，直接在每个数据单元中进行修改。

2 增加和删除一行数据(一个个案)。单击“Data”，选择“insert & delete”命令，也可选中一行，再击右键，选插入(删除)。

3 数据的排序。单击“Data”，选择“sort”命令，弹出“sort”case对话框，在左框中选取一个需要按其数值大小排序的变量(也可多个)添加到右框图中，并选择按升序还是降序，按“OK”即可。

4 数据的行列互换。将数据管理器中原先按行(列)方向排列的数据转换成列(行)方向排列，即将个案和变量进行转换。这时单击“Data”，选择“transpose”弹出其对话框，再在变量名列表中选择1个或多个需转换的变量，使之添加到“variable”框中，单击“OK”即可。

5 选择个案子集。在数据统计中可从所有资料中选择部分进行统计分析，单击“Data”，选择“select”命令。

集中量数

集中趋势 (central tendency) 在实验测量或调查中获得的大量观测数据，具有一种向数据中央某一点靠拢的趋势，它是数据分布的特征之一。集中量数 (central tendency) 用于描述观测数据集中趋势的量数是一组数据的代表值，用以说明一组数据分布的典型情况或一般水平，它比个别数据更能反映客观现象或事物的实际情况还可以用于组与组之间的差异比较

描述集中趋势的数量指标 算术平均数、中数、众数、几何平均数、调和平均数

均数、中数、众数

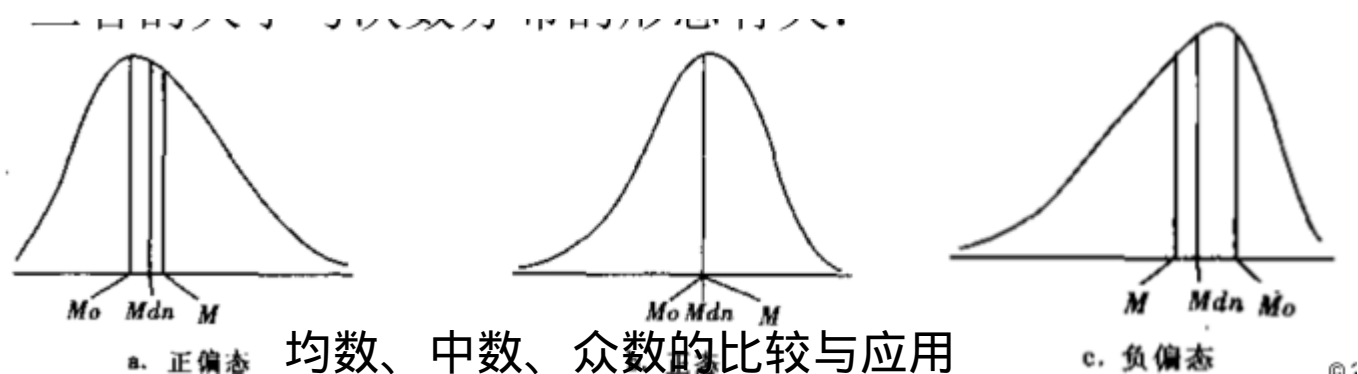
算术平均数 (arithmetic mean) 是所有观测值(或变量)的总和除以总个数所得的商，简称平均数、均数或均值。其符号系统既有表示样本平均数的数学符号 \bar{X} 和英文符号 M ，又有表示总体平均数的希腊字符 μ 。

中位数 (median) 又称中点数，简称中数，用符号 M_d 或 M 表示。中位数是位于按一定顺序排列的一组数的中央位置的数值，它是把一组数据按次数划分为两半，即在中位数上下的数据分布各占一半。

众数 (mode) 一群数据中出现次数最多的那个数值，又称范数，用符号 M_o 表示。

均数、中数、众数的关系

三者的大小与次数分布的形态有关：



均数、中数、众数的比较与应用

作为优良集中量数应当具备六个条件：

- 1 反应灵敏，即一群数据中任何一个数值的变动都会影响集中量跟着变动
- 2 严密确定，即该集中量指标应由全部观测值计算得来，而且同一组数用该指标不同计算方法计算出的结果相同，如定义式或计算式或加权式等
- 3 意义简明，易于理解，即其方法不应带有过多的数学抽象性质
- 4 容易计算，即在其他条件相同的情况下，则以计算简便作为选择的依据
- 5 适合代数法则的处理，即可以用数学性质处理或简化数据
- 6 受抽样变动的的影响较小，即从同一总体中抽样的多个样本，个样本的计算结果相同或相近

	均数	中数
优点	1. 符合①②③④⑤⑥ 2. 已知 ΣX 和 N ，即可求 \bar{X} 3. 便于进行加权处理 4. 统计推断的结果更可靠、稳定	1. 符合②③④⑥ 2. 少受极端数值的影响
应用	1. 加权平均数 2. 离差、相关计算、进行统计推断 3. 考试评估	1. 有极端数值时 2. 测量单位的性质有怀疑 3. 上下端距离不确定 4. 采用百分体制时
不足	1. 易受极端值的影响 2. 组距不确定时无法计算	1. 易受抽样偏差影响 2. 不适合代数法处理

S P S S均数、中数、众数

1 录入数据；2选“ A n a l”展开下拉菜单。 击“ D e s c r i b t t a i t v i e s , t i c s
再单击“ F r e q e”添加变量到“ V a r a b ; B e单击“) S t a t i s t , i c s
选择“ C e n t T r e i d e”中的“ y M e , M e d i m d S e u”再单击“ C o n t i , n u e
返回 “ F r e q u e”单击“ O K”即可

差异量

离中趋势 (d i v ... r g t e m d e 统计分布的 ；一种特征，又称离散程度。差
g e n c e) 描述一组数 离中趋势的量数。 描述离中趋
、标准差 、 平均差 、 全
析，不仅需要集中量数的 标，而且还需要差异量数的指
应用 来才能对一系列数据做出较 全面的分析
均差平方的算术平均数 表示一系列数据平均差距的平方，
表示，总体方差用符号 S^2 表示。
用符号 V (V a r i a) 表示，或称均方差、均方，用符号 $M S$
不足 u a r e) 表示 e v i a t i o n

S^2
 σ^2

标准差(Standard deviation)方差的算术平均跟，表示一系列数据的平均差距。
样本标准差用符号 s 或 S 表示，总体标准差则用符号 σ 表示。

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{N}}$$

全距(Range)是一列数据中的最大值与最小数的差距，

$$Rg = X_{\max} - X_{\min}$$

又称极差，用符号 R 表示。
平均差(Average Deviation)以离差绝对值的和除以总和次数所得的商，平均差用符号 AD 表示。

$$AD = \frac{\sum |X - \bar{X}|}{N}$$

百分位差(Percentile deviation)表示某两个百分位数之间差异程度的指标，常用的百分位差如 $P_{93} - P_7$ (分析数据中 80% 的数据)， $P_{90} - P_{10}$ 分析数据中 80% 的数据。

四分位差(Quartile deviation)百分位差的特例，用于分析 $P_{75}(Q_3)$ 与 $P_{25}(Q_1)$ 之间的差距。

$$Q = \frac{Q_3 - Q_1}{2}$$

各种差异量指标的比较

类别	优点	缺点	应用
标准差与方差	1. 反应灵敏 2. 严密确定 3. 适合代数法处理 4. 受抽样变动影响小	1. 原理难理解 2. 计算较复杂 3. 受极端值影响较平均差大	1. 反应变量值的离散程度 2. 进行统计估计和假设检验 3. 用于求正常范围 4. 计算变异系数，标准分数和标准误等
全距	1. 意义简明 2. 计算简单	1. 易受极端值的影响 2. 不反映全部数据的差异情况	使用价值极小
平均差	1. 意义简明 2. 计算简单 3. 严密确定	1. 受极端值的影响 2. 不适合代数法的处理	使用较少
百分位差	1. 意义简明 2. 不受极端数的影响 3. 组距不确定时刻用	1. 不反映全部数据差异情况 2. 不适合代数法的处理 3. 受抽样变动影响较大	1. 不需精细的结果 2. 比较团体各阶段之间的差异

标准差(Standard deviation)方差的算术平均跟，表示一系列数据的平均差距。
样本标准差用符号s或S表示，总体标准差则用符号σ表示。

平均 全距

全距(Range)一系列数据中的最大差即一系列数据中最大数与最小数的差距，

又称极差，用符号R表示。

平均差(Average Deviation)以离差绝对值的和除以总和次数所得的商，平均差用符号A表示。

$$P_R =$$

百分位差(Percentile deviation)表示某两个百分位数之间差异程度的指标，常用的百分位差如P93-P7(分析数据中86的数据)，P90-P10分析数据中90的数据。

四分位差(Quartile Deviation)百分位差的特例，用于分析P75(Q3)与P25(Q1)半。

$$Z = \frac{X - \bar{X}}{S}$$

各种差异量指标的比较

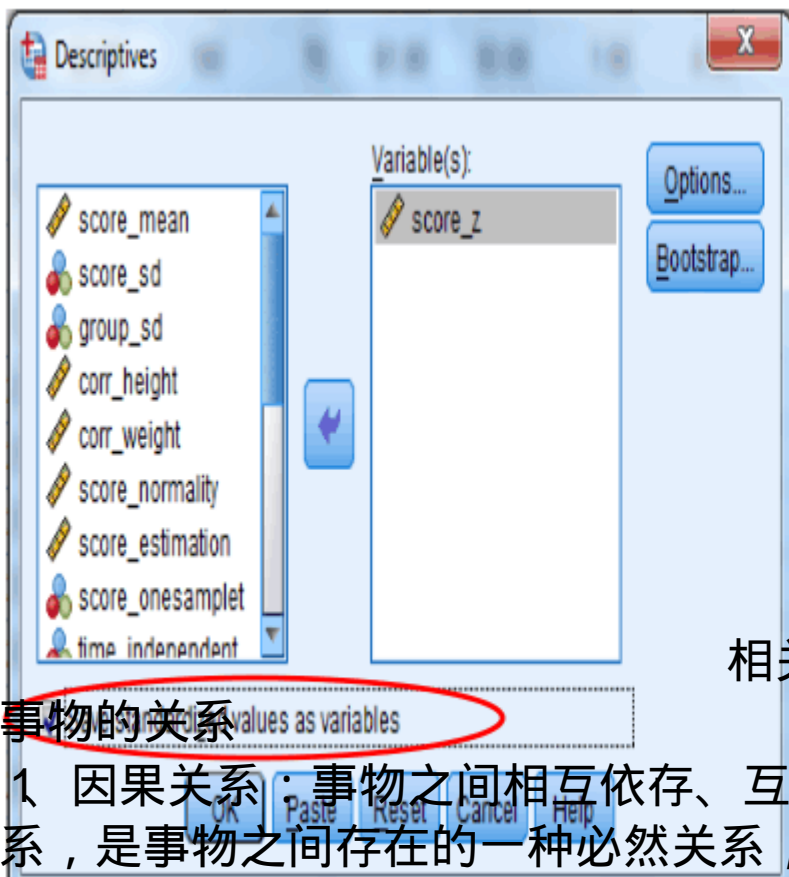
集中量数与差异量数的关系

1 在多组数据比较时，当集中量相同差异量不同，或是差异量相同集中量不同时，均不能冒然下结论说各组分布相同。

$$P_R = 100 - \frac{100 - P_1 - P_9}{N}$$

$$Z = \frac{X - \bar{X}}{S}$$

中的“ Descrip出现某对话框



score_mean	score_sd	group_sd	score_z	Zscore_z	corr_height	corr_weight
43.00	8.00	1	80.00	-.30125	170	72
23.00	9.00	1	90.00	.85740	165	69
35.00	10.00	1	78.00	-.53298	150	66
37.00	13.00	1	64.00	-2.15510	180	70
42.00	13.00	1	88.00	.62567	185	68
48.00	14.00	1	92.00	1.08914	.	.
54.00	14.00	1	83.00	.04635	.	.
28.00	15.00	1	75.00	-.88058	.	.
44.00	3.00	2	90.00	.85740	.	.
36.00	5.00	2	86.00	.39394	.	.
38.00	5.00	2
40.00	7.00	2
.	9.00	2
48.00	21.00	2
33.00	33.00	2

相关量数

事物的关系

- 1 因果关系：事物之间相互依存、互为因果的关系，是事物之间存在的一种必然关系，即一种引起与被引起的关系，因在前果在后的顺序不能颠倒
- 2 函数关系：事物之间的一种共变关系，其特点是函数与反函数可以互换位置
- 3 伴随关系：两个事物或变量相伴随而变化的关系，它既不同于因果关系和函数关系，又不排斥因果关系和函数关系，是事物之间的一种更为复杂的关系，相关关系即属这种关系。

相关量数：分析或研究两个或两个以上变量之间相互

相关的种类

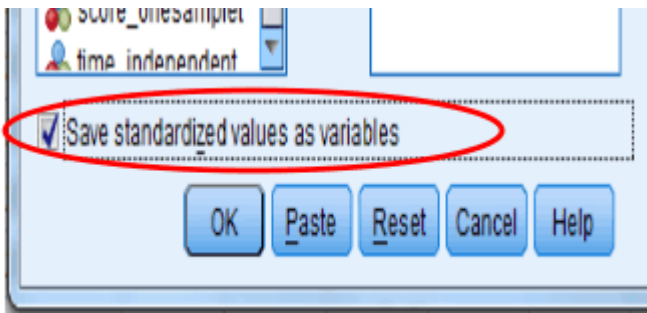
事物或现象的相关种类可以从方向、形态及变量个数

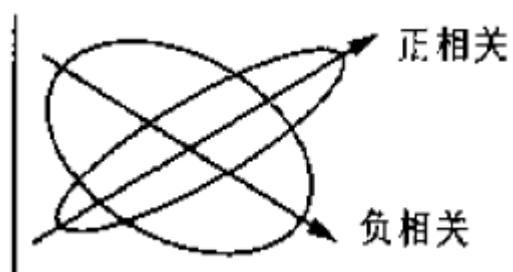
- 1 正相关、负相关和零相关

正相关(positive) 一列变量由大而小或由小而大变化时，另一列变量亦由大而小或由小而大的变化，即两列变量是同方向变化的，属“同增共减”的关系

负相关(negative) 一列变量由大而小或由小而大的变化，另一列变量却由小而大或由大而小的变化，即两列变量的变化方向是相反的，属“此增彼减”的关系

零相关(zero correlation), 又称无相关, 指一列变量由大而小或由小而大变化时，另一列变量或大或小的变化，即两列变量的变化看不出一定的趋势，甚至毫无关系





2 直线相关、曲线相关

直线相关 (linear correlation) 两列变量中的一列变量在增加时，另一列变量随之而增加；或一列变量在增加，另一列变量却相应地减少，形成一种直线关系

曲线相关 (curve correlation) 两列相伴随变化的变量，未能形成直线关系



3 简单相关、复杂相关

简单相关 (simple correlation) 只有两个变量的相关

复杂相关 (complex correlation) 有三个或三个以上变量的相关

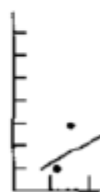
4 完全相关、强相关、弱相关、无相关

完全相关 (complete correlation) 两列变量的关系是一一对应、完全确定的关系；在坐标轴上描绘两列变量时会形成一条直线

强相关 (strong correlation) 又称高度相关，即当一列变量变化时，与之相应的另一列变量增大（或减少）的可能性很大；在坐标图上则表现为散点较为集中地呈现在某条直线的周围

弱相关 (weak correlation) 又称低度相关，即当一列变量变化时，与之相对应的另一列变量增大（或减少）的可能性较小，亦即两列变量之间虽然有一定的联系，但联系的紧密程度较低；在坐标图上表现出散点比较分散地分布在某条直线的周围

无相关 (no correlation) 一列变量变动时，相对应的另一列变量可能无变动，也可能无规律



相关分析的方法

相关分析 研究两个或两个以上变量之间是否存在相互关系，如果存在关系，其相关的性质（即方向）和程度如何，这个研究过程在统计学上称相关分析。相关分析的方法主要是图示法和计算法

1 图示法：主要利用散点图来描述变量之间的关系。散点图是将成对变量的变动值描绘在坐标图上形成的一种图形

从散点图上，可以了解相关的方向（正相关、负相关、零相关）相关的形态（直线相关、曲线相关）相关的大致程度（强相关、弱相关）

2 计算法：通过计算变量之间的相关系数来表述其相关情形

相关系数(c o r r e l a t i o n i)是表示相关方向和大小的一种数值，用符号 r 表示。相关系数的取值范围为 - 1 r + 1其中符号表示相关的方向，绝对值表示相关的程度

一些统计学家对相关程度作了规定，认为0~± 0 . 表低度相关，± 0 . 4 0 ~ ± 0 . 表中度相关，± 0 . 7 0± 1~. 表高度相关

积差相关P r o d u c t M o m e n t直线相关中最基本的方法，又叫均方相关、积矩相关，其公式由英国统计学家皮尔逊(P e a r)提出，故又称皮尔逊相关，用符号 r 表示积差相关是利用离差乘积的关系来说明事物的关系，是将原始记分转换为离差乘积（即积差），再转换成标准积差后所得的标准积差的平均数

$$r_{XY} = \frac{\sum Z_X Z_Y}{N} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2} \sqrt{\sum (Y - \bar{Y})^2}}$$

使用条件 两列变量必须是成对的，而且样本容量不宜少于30 两列变量必须是比率变量或等距变量； 两列变量的总体分布均为正态分布或近似正态分布

等级相关R a n k C o r r 根据等级资料来研究变量之间相互关系的方法
等级资料的主要来源 研究中所收集的数据本身就是等级评定的资料； 研究所收集的数据原本为等距或比率变量的资料，因不满足积差相关的使用条件而需要将其转化为等级性资料进行分析的情形

使用条件 可以用于多列等级或顺序变量； 可以用于成对变量少于3的情形；
可以用于两列变量总体分布为非正态时

计算方法

1 斯皮尔曼等级相关(S p e a r m a n a n k c o),根据两列变量的成对等级差数计算的相关系数，又叫“等级差数法”，用符号 r 或 r_s 表示，它是由英国统计学家和心理学家斯皮尔曼创立的；S p e a 等级相关因等级或数据中有无重复而分为无重复的等级相关和有重复的等级相关

2 肯德尔和谐系数(K e n d e l C o e f f i c i e n t)又称肯德尔一致 c o r d a n c 性系数或肯德尔相关系数，它是由统计学家肯德尔(K e n d)提出的肯德尔交错系数、相容性系数和一致性系数等三种相关系数的总称。

肯德尔交错系数、相容性系数主要用于描述两列变量的一致性程度，作用与斯皮尔曼等级相关系数相似

肯德尔一致性系数是用于描述多列等级变量相关程度或一致性程度的相关方法，其中又有肯德尔V系数和肯德尔U系数

肯德尔V系数用于一般等级评定的资料（亦分无重复与有重复）

K个评分人评价N个被评价人或N件作品，以分析和评价K个评分人的评价是否一致

同一个人先后K次评价N个被评人或N件作品，以分析其前后评价是否一致

质与量的相关

研究列等距或比率变量与一列名称变量之间相互关系的统计方法，包括

点二列相关、二列相关、多系列相关等
 点二列相关：研究一列等距或比率变量与一列“二分”名称变量之间相关的统计方法，用符号 r_{pb} 表示

$$r_{pb} = \frac{X_p - \bar{X}_q}{S_X} \sqrt{pq}$$

二列相关：研究一列正态的比率或等距变量和一列人为“二分”名称变量之间相互关系的统计方法，用符号 r 或 r_b 表示

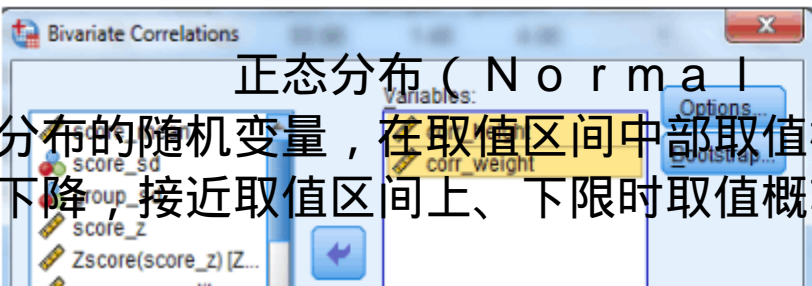
多系列相关
 品质相关 (Correlation of) 研究两列品质变量间相互关系的统计方法，是衡量两个分类（名称）变量关联程度的指标。品质相关因变量分类数目的不同又分为 相关和列联相关
 相关 (phi correlation) 专门研究两列“二分”名称变量之间相关的统计方法，是衡量两个分类变量均分为两类时其关联程度的指标

$$r_b = \frac{X_p - \bar{X}_q}{S_X} \cdot \frac{pq}{Y} = \frac{\bar{X}_p - \bar{X}_q}{S_X} \cdot \frac{p}{Y}$$

SPSS相关分析

1 录入或读取数据；2 单击“Analyze”选择“Correlate”中的“Bivariate”进入“Bivariate Correlations”；3 将变量调入“Variables”；4 在“Correlation Coefficients”中选择所需的相关系数（默认为Pearson相关）；5 在“Test of Significance”中选择“Two-tailed”，单击“OK”即可。

正态分布 (Normal Distribution) 从概率的角度来说，服从正态分布的随机变量，在取值区间中部取值概率最高，从中间到两侧取值概率逐渐下降，接近取值区间上、下限时取值概率越来越小，且两侧取值概率是

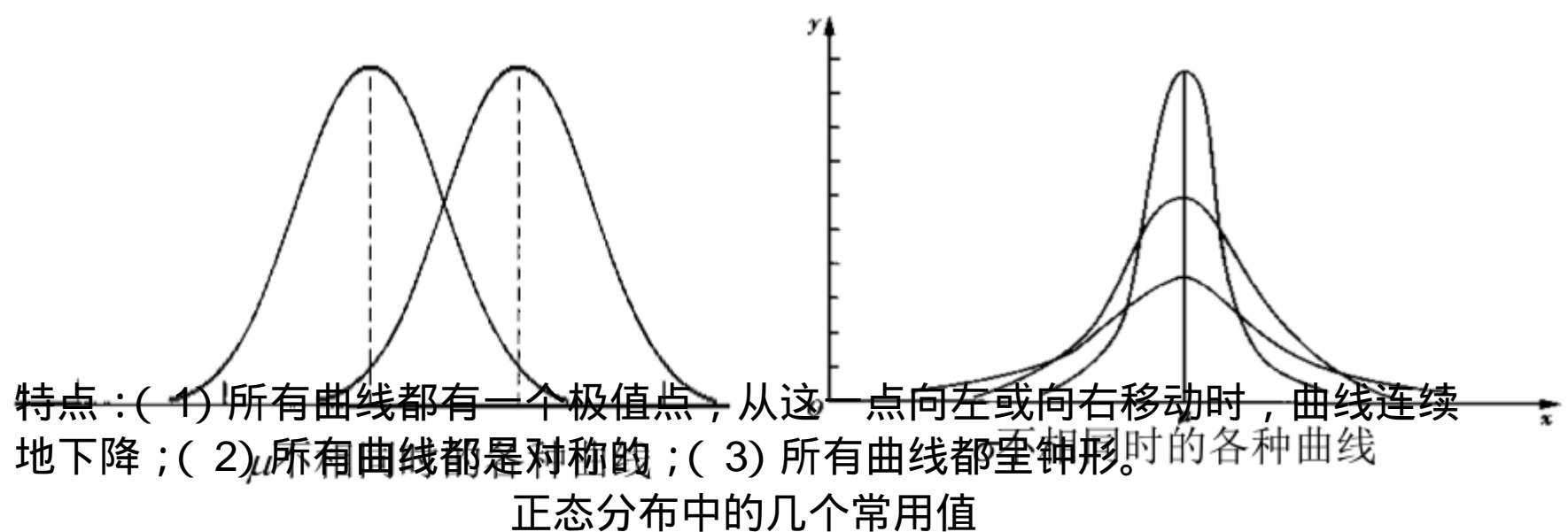


	corr_height	corr_weight
corr_height	1	.449

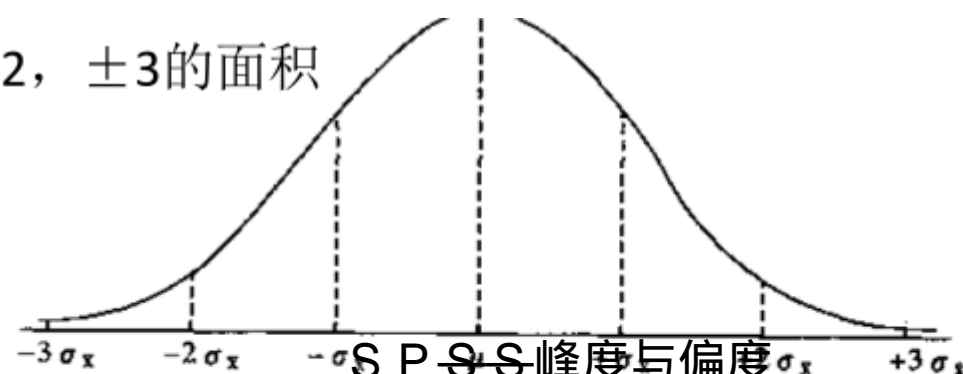
对称的。

通俗地说，就是中间量数次数分布多，两端量数次数分布少，呈对称型的概率分布，又叫高斯分布。

在正态分布中，平均数 μ 和标准差 σ 决定着分布曲线的位置和形状。其中，平均数决定着曲线在 x 轴上的位置，标准差 σ 决定着曲线的形状。



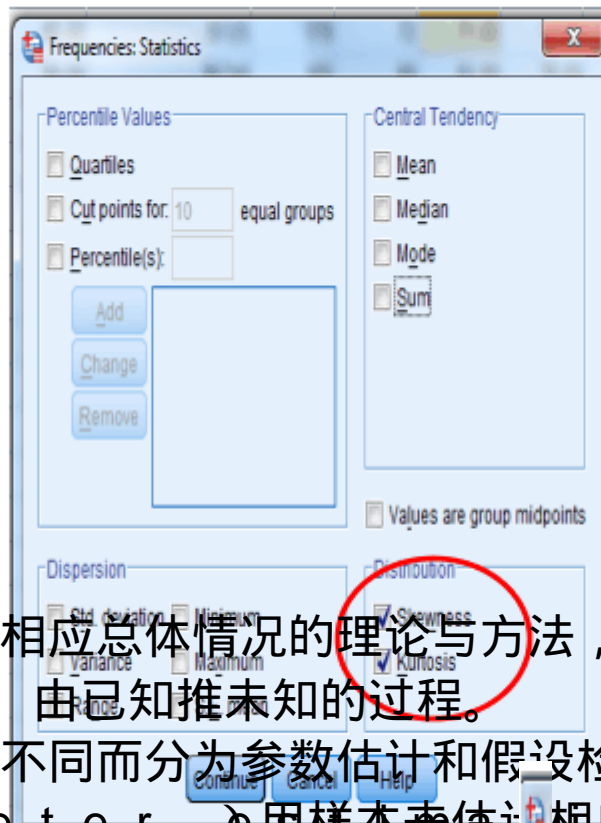
$Z = \pm 1, \pm 2, \pm 3$ 的面积



✓ 离开均数的距离越远，其离差越大

1 录入数据(或读取数据)；2 单击“Analyze”选择 Descriptive Statistics 中的“Frequencies”，出现其对话框，添加“Variables”；3 单击“Display”中的“Skewness”和“Kurtosis”，再单击“OK”即可。

在 $\pm 1.96\sigma$ 和 $\pm 2.58\sigma$ 范围之外的面积只有 0.05 和 0.01



Statistics

score_normality		
N	Valid	50
	Missing	0
Skewness		-.444
Std. Error of Skewness		.337
Kurtosis		-.456
Std. Error of Kurtosis		.662

推断统计：由样本去推测相应总体情况的理论与方法，也就是由部分推全本，由已知推未知的过程。

推断统计根据推测的性质不同而分为参数估计和假设检验两方面。

参数估计 (parameter) 是用样本去估计相应总体的状况。具体方法：点估计和区间估计

假设检验 (hypothesis) 主要用途是对出现差异的两个或多个现象或事物进行真实性情况的检验，又称统计检验 (statistical test)。根据是否需要依赖于对总体分布形态和总体参数检验的假设，分为参数检验和非参数检验 (non-parameter test)。参数检验法主要有 Z 检验、t 检验、F 检验和 χ^2 检验等；非参数检验主要有 2 检验、符号检验、符号等级检验、秩和检验、中位数检验等。

参数估计

根据样本统计量去估计相应总体的参数。如，可以根据样本平均数 (\bar{X}) 去估计总体的均数 (μ)，根据样本方差 (S^2) 去估计总体方差 (σ^2)，根据样本的相关系数 (r) 去估计总体的相关系数 (ρ) 等。

点估计 (point estimate) 在参数估计中直接用样本的统计量 (数轴上的一个点) 作为总体参数的估计值。

区间估计 (interval estimate) 在参数估计中直接以样本的统计量 (数轴上的一个点) 作为总体参数的估计值。它要求按照一定的数值范围，也就是表示未知参数可能落在此范围之间，而落在这一范围之外，表示进行区间估计时存在误差。

置信区间 (confidence interval) 在特定的可靠性 (即置信系数) 要求下，估计总体参数落在的区间范围，亦即进行估计的全距。

以样本均数 (\bar{X}) 为例，在估计总体均数 (μ) 时，假定 \bar{X} 的抽样分布服从

从正态分布，就可以用正态分布的理论，在一定的置信系数下推断出 μ 的置信区间，即

$D=0.95$ （即 $\alpha=0.05$ ）时，置信区间为 $\bar{X} - 1.96\sigma_{\bar{X}} < \mu < \bar{X} + 1.96\sigma_{\bar{X}}$

$D=0.99$ （即 $\alpha=0.01$ ）时，置信区间为 $\bar{X} - 2.58\sigma_{\bar{X}} < \mu < \bar{X} + 2.58\sigma_{\bar{X}}$

置信系数（confidence）被估计的总体参数落在置信区间内的概率或以 1 表示，又叫置信水平、置信度、可靠性系数和置信概率。

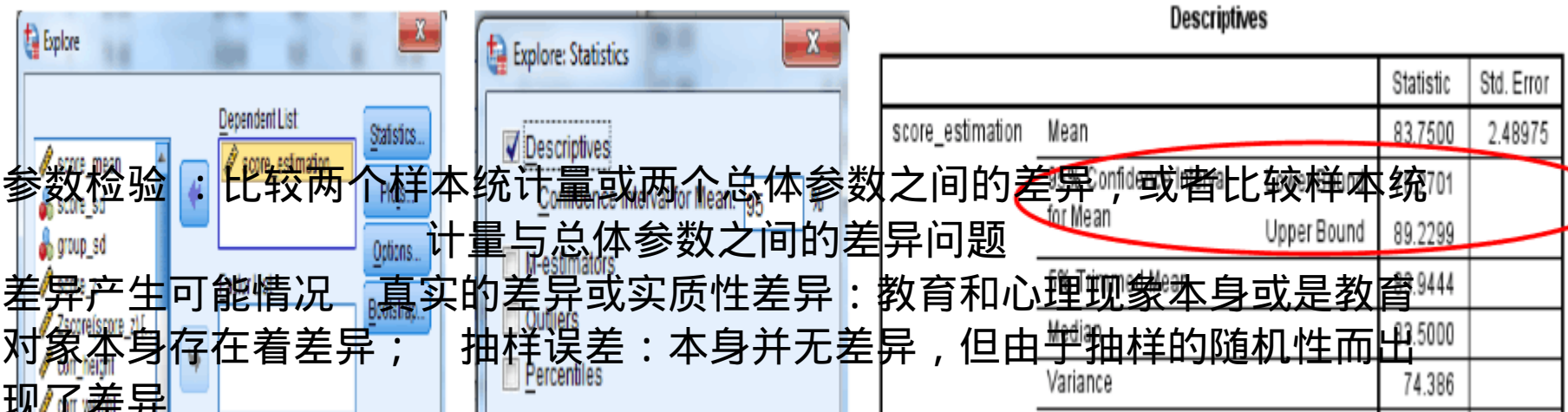
置信系数是用来说明置信区间可靠程度的概率，也是进行正确估计的概率。一个置信系数同时反映了在做出一个估计时犯错误的小概率（ α ），即可靠性为 $1-\alpha$ ，意味着犯错误的概率为 α 。可靠性为 95%，意味着犯错误的概率为 5%；可靠性为 99%，意味着犯错误的概率为 1%。

置信限（confidence limit）

被估计的总体参数所落区间的上、下界限，即

SPSS参数估计

- 1 录入（或读取）数据；
- 2 选择“Analyze”中的“Descriptives”，单击“Options”出现其对话框，添加变量到“Display of statistics”（因变量框）；
- 3 单击“Statistics”，再单击“Descriptives”（置信度默认为 95%，可接受，也可修改）。单击“Continue”按钮，再单击“OK”即可。



The image shows the SPSS 'Descriptives' dialog box and its output table. The dialog box has 'score_estimation' in the 'Dependent List' and 'Descriptives' checked under 'Statistics'. The output table shows statistics for 'score_estimation'.

	Statistic	Std. Error
Mean	83.7500	2.48975
95% Confidence Interval for Mean	Lower Bound: 81.701, Upper Bound: 85.799	
Fit Trimmed Mean	83.9444	
Median	83.5000	
Variance	74.386	

参数检验：比较两个样本统计量或两个总体参数之间的差异，或者比较样本统计量与总体参数之间的差异问题

差异产生可能情况：真实的差异或实质性差异：教育和心理现象本身或是教育对象本身存在着差异；
抽样误差：本身并无差异，但由于抽样的随机性而出现了差异

差异显著性检验：如果经过统计检验（t 检验或卡方检验）发现差异超过了统计上所规定的某一误差的限度时，则表示差异已经不属于抽样误差了，亦即事物之间已出现了真实的差异，统计上将此称为差异显著。这种对事物差异所进行的检

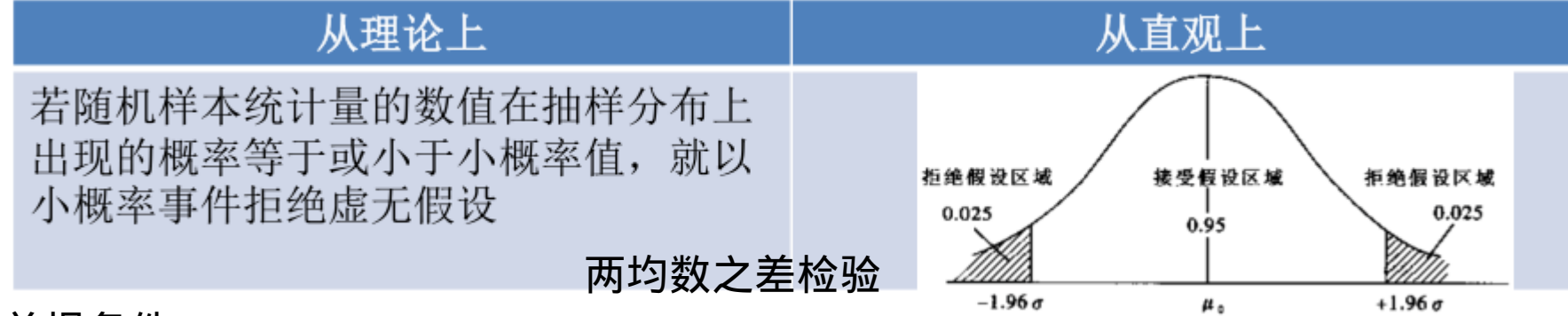
验就是差异显著性检验。
差异显著的界限：通常把概率不超过0.05或0.01作为抽样误差的限度。在进行差异检验时需要先对事物是否存在差异作出假设，然后再作统计检验，因此称为假设检验，亦称差异的显著性检验。

假设与假设检验

假设：研究者在研究过程中根据不多的事实材料，运用已有的科学原理，充分发挥思维的想象力和创造力对所研究事物的本质和规律提出来的一种初步设想

研究假设 (a l t e r n a t i v e) : 实验人员希望证实的假设：假设两个样本统计（或两个总体参数）之间，或者是样本统计量与总体参数之间存在真实的差异，即存在显著的差异，这是一种有差假设，记为H1

虚无假设 (n u l l h y p o t h e s i s) : 研究人员为了证实研究假设是真的而利用概率论的反证法所进行的假设，即从研究假设的反面进行假设，用符号H0表示
显著性水平 (l e v e l o f s i g) 拒绝虚无假设的小概率值



前提条件

- 统计量来自随机样本
- 总体呈正态分布（以中心极限定理为理论依据；当总体为正态分布时，样本均数的抽样分布也呈正态分布；若总体为非正态分布，则样本容量必须很大时，其样本均数的分布才接近正态分布）
- 总体的方差齐性（两个总体之间的方差相等或一致；由于总体方差往往是未知的，因此只能借助样本方差去估计总体方差，所以需要对样本方差一致性进行检验）

考虑的问题

- 总体情况：总体分布（是否正态）、总体方差（是否已知）、总体方差一致性（是否齐性）
- 样本类型

独立样本 (i n d e p e n d e n t s a m p l e) : 从两个无关的总体中随机抽取的两个或多个样本，或者说是独立抽取的，彼此间的数据不存在对应关系的样本

相关样本 (c o r r e l a t e d s a m p l e) : 具有一定程度相关的总体中抽取的两个或多个样本，亦即彼此的观测值之间存在一一对应关系的样本。常见形式有“同组比较”和“配对比较”

两均数之差的显著性检验

✓ 单总体均数之差的显著性检验

检验一个样本均数与相应总体均数之差（即 $\bar{X} - \mu$ ）是否显著的统计方法

○ 总体正态， σ^2 已知，Z检验： $Z = \frac{\bar{X} - \mu_0}{SE_{\bar{X}}}$ ，其中 $SE_{\bar{X}} = \frac{\sigma_0}{\sqrt{n}}$

○ 总体正态， σ^2 未知，t检验： $t = \frac{\bar{X} - \mu_0}{SE_{\bar{X}}}$ ，其中 $SE_{\bar{X}} = \frac{S}{\sqrt{n-1}}$

○ 总体非正态（ $n > 30$ ）， σ^2 已知，Z'检验： $Z' = \frac{\bar{X} - \mu_0}{SE_{\bar{X}}}$ ，其中 $SE_{\bar{X}} = \frac{\sigma_0}{\sqrt{n}}$

○ 总体非正态（ $n > 30$ ）， σ^2 未知，Z'检验： $Z' = \frac{\bar{X} - \mu_0}{SE_{\bar{X}}}$ ，其中 $SE_{\bar{X}} = \frac{S}{\sqrt{n}}$

✓ 双总体均数之差的显著性检验

双总体均数之差（ $\mu_1 - \mu_2$ ）的检验目的在于通过对两个样本均数差异（ $\bar{X}_1 - \bar{X}_2$ ）的检验来说明并推测它们所代表的两个总体均数之差（ $\mu_1 - \mu_2$ ）的情况

SPSS两均数之差的检验
两个样本均数之差检验方法也有Z检验、t检验和Z'检验三种方法，单击ans
“One-Sample t-test”
1 录入（或读取）数据；2 选择“Analyze”中的“Compare”，单击ans
“One-Sample t-test”
3 在左框变量选入“Test Variable”并在“Test Value”中
输入总体均数4 单击“Options”在“Confidence”输入
置信度（默认为95%）单击“OK”即可

○ Z'检验，总体为非正态分布（样本容量大于30），总体方差已知或未知

✓ 双总体均数之差的显著性检验

双总体均数之差（ $\mu_1 - \mu_2$ ）的检验目的在于通过对两个样本均数差异（ $\bar{X}_1 - \bar{X}_2$ ）的检验来说明并推测它们所代表的两个总体均数之差（ $\mu_1 - \mu_2$ ）的情况

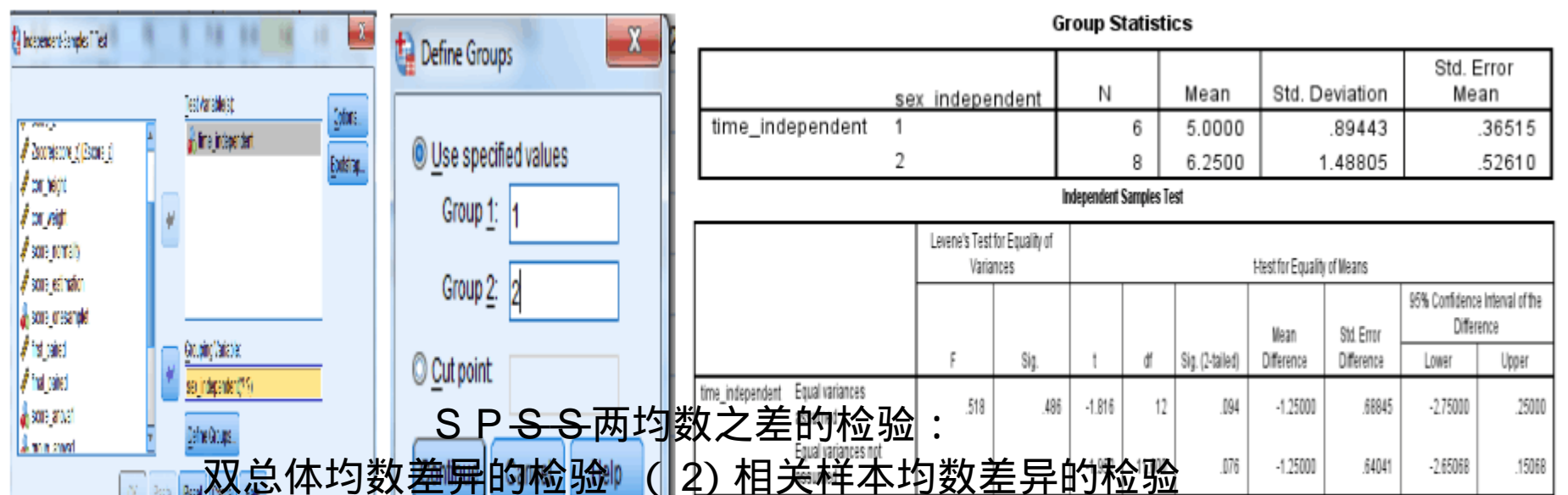
两个样本均数之差检验方法也有Z检验、t检验和Z'检验三种方法

○ Z检验：两总体分布均为正态分布，且两总体方差均一致

○ t检验：两总体分布均为正态分布，但两总体方差均未知

○ Z'检验：总体为非正态分布（样本容量大于30），总体方差已知或未知

SPSS两均数之差的检验
双总体均数差异的检验（1）独立样本均数差异的检验
1 录入（或读取）数据；2 选择“Analyze”中的“Compare”，单击ans
“Independent-Samples t-test”
3 在左框变量分别选入“Test Variable”和“Group Variable”单击“Define Groups”，在“Group 1”和“Group 2”键入相应组别。单击“OK”返回再单击“OK”即可



S P S S两均数之差的检验：

双总体均数差异的检验 (1) 独立样本均数差异的检验 (2) 相关样本均数差异的检验

1 录入 (或读取) 数据；2 选择 “Analyze” 中的 “Compare”，单击 “Paired-Samples t-test”，选择变量入 “Paired Variables”，单击 “OK”

方差分析 (ANOVA) 对多个平均数进行比较的一种统计方法，又称变异数分析

2 检验和 t 检验的局限性

比较的组合次数增多 (A-B-C 组，需 3 次；1 组：需 4 次)

降低可靠程度 (对数据做的 t 检验或 F 检验越多，越容易犯 I 型错误；做 n 次检验，每次都为 0.05 的显著性水平，那么不犯 I 型错误的概率为 0.95^n ，此时犯 I 型错误的概率则为 $1 - 0.95^n$)

缺少综合或整体信息 (两个以上的平均数检验中若仍采用 t 检验或 F 检验，都只提供了两个组的信息，而忽略了其余的综合信息，而这些被忽视的信息可能对检验结果产生更大的影响力)

使用条件

总体分布的正态性

各个实验组的方差齐性

变异具有可加性 (方差分析是将事物的总变异分解为各个不同变异来源，分解后的各部分变异相互独立，相加后又构成总变异) 总变异 = 组间变异 (实验条件 + 随机误差) + 组内变异 (个体差异 + 实验误差) (组内变异属随机误差)

变异率与 F 分布：根据方差分析的原理，需比较组间变异和组内变异，若用一个统计量来揭示组间变异，则称这一统计量为变异率。方差分析的 F 值是以统计学家 R . A . 的名字命名的，用符号 F 表示

变异率 (F) = 组间变异 / 组内变异 = (实验处理变异 + 误差变异) / (误差变异)

当实际计算的 F 值超过了 F 分布表中 0 . 05 的临界值时，说明各样本均数属于相同总体的机率小于 5 % 则要拒绝虚无假设，接受研究假设：认为实验因素确实产生了效应，各样本均数之间存在着显著差异

方差齐性检验：不同总体的变异水平——个体差异是否一致

方差齐性检验 (t e s t o f h o m o g e n e) i t y o f v a r i a n c e

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2$$

单因素方差分析与多因素方差分析：实验所考察的自变量只有一个（多个）的实验设计的方差分析

多重比较：F 检验反映的总体差异情况，多重比较可找出使 F 值显著的根源，即有关各对平均数之间差异的状况

S P S S 单因素方差分析

- 1 录入 (或读取) 数据 ; 2 选择 “ a n a l y s e o f v a r i a n c e ” 单击 “ O n e - W a y A N O V A ” 选择变量入 “ D e p e n d e n t ” 和 “ F a c t o r ”
- 3 单击 “ O p t i o n s ” 进入方差分析，选择 “ D e s c r i p t i v e ” 和 “ H o m o g e n e i t y o f v a r i a n c e ” 单击 “ C o n t i n u e ” 返回 “ O n e - W a y A N O V A ” 单击 “ P o s t H o c ” 进入多重比较，选择相应比较法，单击 “ C o n t i n u e ” 返回 “ O n e - W a y ” , A N O V A 再单击 “ O K ” 即可

score_anova								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1 A	6	5.0000	1.41421	.57735	3.5159	6.4841	3.00	7.00
2 B	6	10.0000	.89443	.36515	9.0614	10.9386	9.00	11.00
3 C	6	15.0000	1.67332	.68313	13.2440	16.7560	13.00	17.00
Total	18	10.0000	4.39251	1.03532	7.8157	12.1843	3.00	17.00

Test of Homogeneity of Variances			
score_anova			
Levene Statistic	df1	df2	Sig.
1.154	2	15	.342

ANOVA					
score_anova					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	300.000	2	150.000	80.357	.000
Within Groups	28.000	15	1.867		
Total	328.000	17			

Post Hoc Tests

Multiple Comparisons						
Dependent Variable: score_anova						
Bonferroni						
(I) group_anova	(J) group_anova	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1 A	2 B	-5.00000 [*]	.78881	.000	-7.1249	-2.8751
	3 C	-10.00000 [*]	.78881	.000	-12.1249	-7.8751
2 B	1 A	5.00000 [*]	.78881	.000	2.8751	7.1249
	3 C	-5.00000 [*]	.78881	.000	-7.1249	-2.8751
3 C	1 A	10.00000 [*]	.78881	.000	7.8751	12.1249

变异率与 F 分布：根据方差分析的原理，需比较组间变异和组内变异，若用一个统计量来揭示组间变异，则称这一统计量为变异率。方差分析的 F 值是以统计学家 R. A. 的名字命名的，用符号 F 表示

变异率 (F) = 组间变异 / 组内变异 = (实验处理变异 + 误差变异) / (误差变异)

当实际计算的 F 值超过了 F 分布表中 0.05 的水平临界值时，说明各样本均数属于相同总体的机率小于 5% 则要拒绝虚无假设，接受研究假设：认为实验因素确实产生了效应，各样本均数之间存在着显著差异

方差齐性检验：不同总体的变异水平——个体差异是否一致

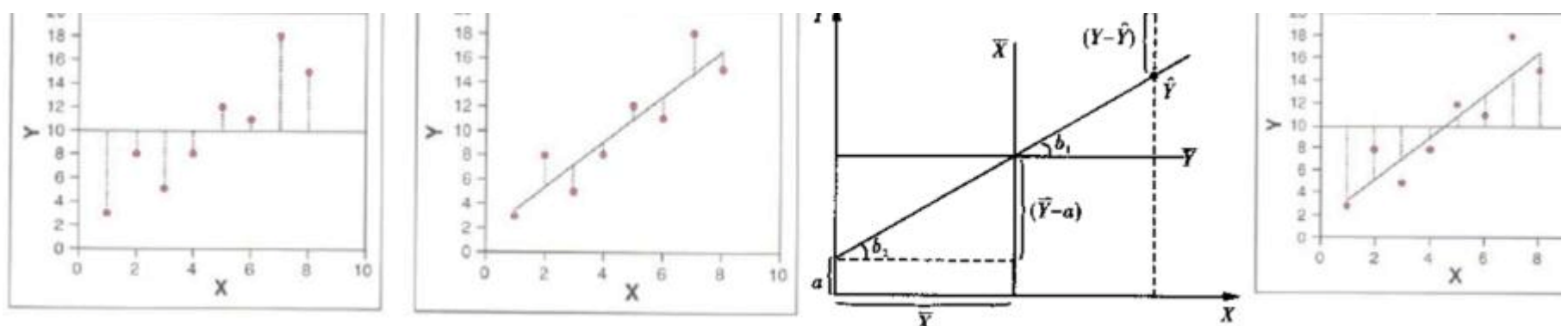
方差齐性检验 (test of homogeneity of variance)

单因素方差分析与多因素方差分析：实验所考察的自变量只有一个（多个）的实验设计的方差分析

多重比较：F 检验反映的总体差异情况，多重比较可找出使 F 值显著的根源，即有关各对平均数之间差异的状况

SPSS 单因素方差分析

1 录入 (或读取) 数据；2 选择“Analyze”中的“Compare Means”，单击“One-Way ANOVA”，选择变量入“Dependent List”和“Factor”
3 单击“Options”，进入方差分析，选择“Descriptive Statistics”和“Homogeneity of Variance”，单击“Continue”，返回“One-Way ANOVA”，单击“Post Hoc”，进入多重比较，选择相应比较法，单击“Continue”，返回“One-Way ANOVA”，再单击“OK”即可



Descriptive Statistics

	Mean	Std. Deviation	N
score_regression	76.0000	12.43651	10
time_regression	29.0000	12.03698	10

Correlations

		score_regression	time_regression
Pearson Correlation	score_regression	1.000	.721
	time_regression	.721	1.000
Sig. (1-tailed)	score_regression	.	.009
	time_regression	.009	.
N	score_regression	10	10
	time_regression	10	10

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.721 ^a	.519	.459	9.14441

a. Predictors: (Constant), time_regression

b. Dependent Variable: score_regression

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	723.038	1	723.038	8.647	.019 ^b
	Residual	668.962	8	83.620		
	Total	1392.000	9			

a. Dependent Variable: score_regression

b. Predictors: (Constant), time_regression

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	54.406	7.893		6.893	.000	36.205	72.606
	time_regression	.745	.253	.721	2.941	.019	.161	1.329

a. Dependent Variable: score_regression

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	61.8520	89.4034	76.0000	8.96312	10
Std. Predicted Value	-1.578	1.495	.000	1.000	10
Standard Error of Predicted Value	2.903	5.614	3.978	1.001	10
Adjusted Predicted Value	58.7389	89.6191	75.8276	9.41629	10
Residual	-15.06442	11.25537	.00000	8.62144	10
Std. Residual	-1.647	1.231	.000	.943	10
Stud. Residual	-1.854	1.298	.008	1.041	10
Deleted Residual	-19.07925	12.51663	.17239	10.55707	10
Stud. Deleted Residual	-2.296	1.367	-.033	1.138	10
Mahal. Distance	.007	2.492	.900	.903	10
Cook's Distance	.001	.458	.112	.131	10
Centered Leverage Value	.001	.277	.100	.100	10

a. Dependent Variable: score_regression

S P 数据分析处理流程

Residual	668.962	8	83.620		
Total	1392.000	9			

a. Dependent Variable: score_regression
b. Predictors: (Constant), time_regression

SPSS 基本运行方式

Coefficients ^a								
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	54.406	7.893		6.893	.000	36.205	72.606
	time_regression	.745	.253	.721	2.941	.019	.161	1.329

完全窗口菜单方式

所有分析操作过程都是通过菜单和按钮及对话框方式进行；是经常使用的一种运行方式，适用于一般分析和 S P 的初学者程序运行方式

菜单程序混合运行方式：一般适合于熟练的 S P 程序员 手工编写 S P 命令程序； 一次性提交计算机运行； 适用于大规模的分析工作和熟练的 S P 程序员

概念比较

描述统计(Descriptive statistics)推断统计(Inferential statistics)

描述统计：研究如何简缩数据和描述这些数据，其任务是在大量的观测数据中，用恰当的统计方法来简缩数据，计算其有代表性的量数，以描述事物的典型性、波动范围以及相互关系等，揭示事物的内部规律。同时，描述统计学还可以对观测数据进行整理归类，并用相应的表格、图形来表示其主要分布特点或特征。

如计算平均数、中数和众数等以反映观测数据的集中趋势；计算方差、标准差等来反映观测数据的离散趋势；计算积差相关系数、等级相关系数等来反映观测数据之间的相互关系。

推断统计：主要研究如何利用数据去作出决策的方法，在心理与教育的研究和实验中涉及面非常大。因为不可能对全体对象进行研究和实验，只能从中抽取部分具有代表性的样本进行研究和实验，推断统计学正式一种依据部分数据去推测全体特征的科学方法，是对教育与心理研究或实验作出预测和规划的有力工具。

推断统计学的主要内容分为参数统计法（含参数估计和假设检验）和非参数统计法。

Central tendency 集中趋势 (Divergence 差异量数)

集中趋势 (central tendency) 在实验、测量或调查中获得的大量观测数据，具有一种向数据中某一点靠拢的趋势，它是数据分布的特征之一。

集中量数 (central tendency) 用于描述观测数据集中趋势的量数是一组数据的代表值，用以说明一组数据分布的典型情况或一般水平，它比个别数据更能反映客观现象或事物的实际情况还可以用于组与组之间的差异比较

描述集中趋势的数量指标 算术平均数、中数、众数、几何平均数、调和平均数

离中趋势 (divergence) 统计分布的另一种特征，又称离散程度。

差异量数 (divergence) 描述一组数据离中趋势的量数。

描述离中趋势的数量指标 方差、标准差、平均差、全距

对一系列数据分布的分析，不仅需要集中量数的指标，而且还需要差异量数的指标，只有二者结合起来才能对一系列数据做出较为全面的分析

Kurtosis 偏度 (Skewness)

峰度 (Kurtosis) 描述某变量所有取值的分布形态陡缓程度的统计量，而峰度对陡缓程度的量度是与正态分布进行比较的结果。

峰度 = 0 其数据分布的陡缓程度与正态分布相同

峰度 > 0 其数据分布比正态分布更陡峭

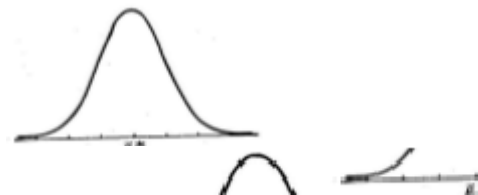
峰度 < 0 其数据分布比正态分布更平坦

偏度 (Skewness) 描述数据分布对称性的统计量，而且也是与正态分布的对称性相比较而得到的。

偏度 = 0 其数据分布的对称性与正态分布相同

偏度 > 0 其分布为正偏，即在峰的右边有大的偏差值，使右边出现一个拖得较远的尾巴

偏度 < 0, 其分布为负偏，即在峰的左边有大的偏差值，使左边出现一个拖得较远的尾巴



错误 (Type I) error 错误 (Type II) error



- ✓ 虚无假设为真时，拒绝虚无假设 ← α 错误
- ✓ 虚无假设为真时，接受虚无假设
- ✓ 虚无假设不真时，接受虚无假设 ← β 错误
- ✓ 虚无假设不真时，拒绝虚无假设

○ α 错误 (α -error)

虚无假设本身是正确，但由于抽样的随机性而使检验值落入了拒绝虚无假设的区域，致使拒绝虚无假设的结论，又称I型错误

若 $\alpha=0.05$ ，平均每百次决策中拒绝正确的虚无假设

若 $\alpha=0.01$ ，平均每百次决策中拒绝正确的虚无假设

○ β 错误 (β -error)

虚无假设本身不正确，但由于抽样的随机性而使检验值落入了接受虚无假设的区域，致使接受虚无假设的结论，又称II型错误 (Type II error)

○ α 错误与 β 错误

减小犯 α 错误的风险，必然会增大犯 β 错误的风险

减小犯 β 错误的风险，又会增大犯 α 错误的风险

- ✓ 1
- ✓ 1
- ✓ 1
- ✓ 1

Z检验 / t检验 vs 方差分析 (ANOVA)

Z检验 (Z Test) 是一般用于大样本 (即样本容量大于 30) 平均值差异性检验的方法。它是用标准正态分布的理论来推断差异发生的概率，从而比较两个平均数的差异是否显著。

t检验，亦称 student 检验 (Student's t test)，主要用于样本含量较小 (例如 $n < 30$) 总体标准差未知的正态分布资料。

方差分析 (ANOVA) 对多个平均数进行比较的一种统计方法，又称变异数分析

Z检验和 t检验的局限性

比较的组合次数增多 (A-B-C三组，需 3次；1组：需 4次)

降低可靠程度 (对数据做的 Z检验或 t检验越多，越容易犯 I型错误；做 n次检验，每次都为 0.05 的显著性水平，那么不犯 I型错误的概率为 0.95^n ，此时犯 I型错误的概率则为 $1 - 0.95^n$)

缺少综合或整体信息 (两个以上的平均数检验中若仍采用 Z检验或 t检验，都

只提供了两个组的信息，而忽略了其余的综合信息，而这些被忽视的信息可能对检验结果产生更大的影响力)

使用条件 总体分布的正态性； 各个实验组的方差齐性； 变异具有可加性（方差分析是将事物的总变异分解为各个不同变异来源，分解后的各部分变异相互独立，相加后又构成总变异） $\text{总变异} = \text{组间变异} (\text{实验条件} + \text{随机误差}) + \text{组内变异} (\text{个体差异} + \text{实验误差})$ (组内变异属随机误差)