

## 基本统计概念说明

### 基本统计说明

数据类型是每类分析方法的基石,区分好数据类型,便可找到合适的分析方法. 基本统计名词概念,可有助于理解分析结果指标意义.

#### 1. 数据分两类,定量和定类,区别为数字大小是否具有比较意义

术语	说明	举例
定量数据	数字大小具有比较意义	您对天猫的满意度情况 (非常不满意, 比较不满意, 中立, 比较满意, 非常满意)
定类数据	数字大小代表分类	性别 (男和女), 专业 (文科、理科、工科)

- 定量: 数字有比较意义, 比如数字越大代表满意度越高, 量表为典型定量数据
- 定类: 数字无比较意义, 比如性别, 1 代表男, 2 代表女

用户可结合数据类型及相应需求选择使用对应的分析方法, 如下表所示:

#### 2. P 值: 显著性值或 Sig 值, 描述某事情发生的概率

如果 P 值小于 0.01 即说明某件事情的发生至少有 99% 的把握, 如果 P 值小于 0.05 (并且大于 0.01) 则说明某件事情的发生至少有 95% 的把握.

研究人员想研究不同性别人群的购买意愿是否有明显的差异, 如果对应的 P 值小于 0.05, 则说明呈现出 0.05 水平的显著性差异, 即说明不同性别人群的购买意愿有着明显的差异, 而且对此类差异至少有 95% 的把握. 绝大多数研究均希望 P 值小于 0.05, 即说明有影响, 有关系, 或者有差异等.

- 常见标准: 0.01 和 0.05, 分别代表某事情发生至少有 99% 或 95% 的把握
- 语言表述: 0.01 或 0.05 水平显著
- 符号标示: 0.01 使用 2 个\*号表示, 0.05 使用 1 个\*号表示

#### 3. 量表, 通常指李克特量表, 测量样本对于某构念(通俗讲即某事情)的态度或看法

量表答项类似于“非常同意”、“同意”、“不一定”、“不同意”、“非常不同意”, 也或者“非常满意”、“比较满意”、“中立”、“比较不满意”, “非常不满意”等. 大多数统计方法均只能针对量表, 比如信度分析, 效度分析, 探索性因子分析(Exploratory Factor Analysis, EFA)等.

量表的尺度形式有多种, 常见是五级量表, 即五个答项, 另外还会有七级量表, 九级量表或者四级量表等.

SPSSAU 建议：先描述想研究什么，用一句话描述，话里面拆分成 X 和 Y；然后结合 X 与 Y 的数据类型，选择对应的数据方法。

分析方法	功能	一句话说明	数据类型
频数	百分比	男女比例分别多少	定类
交叉(卡方)	差异关系	不同性别【X】人群是否抽烟【Y】的差异情况	● X(定类) ● Y(定类)
描述	平均值	平均身高	定量
分类汇总	差异关系		● X(定类) 【可选】 ● Y(定量/定类) 【可选】
相关	相关关系	身高【X】和体重【Y】有没有关系	● Y(定量) 【可选】 ● X(定量) 【可选】
回归	影响关系	身高【X】影响体重【Y】吗？	● Y(定量) ● X(定量/定类)
聚类	人群分类	300 个人分成几类？	定量
因子	● 浓缩 ● 权重	● 30 句话概述成 5 个关键词(因子) ● 5 个关键词(因子)分别代表 30 句话的信息比重？	定量
信度	可靠性	数据真实吗？	定量
效度	有效性	数据有效吗？	定量
方差	差异关系	不同收入【X】群体的身高【Y】是否有差异？	● X(定类) ● Y(定量)
T 检验	差异关系	不同性别【X】群体的身高【Y】是否有差异？	● X(定类) ● Y(定量)
多重响应	百分比	多选题的选择比例情况如何	● X(定类) 【可选】 ● 多选题选项
事后检验	差异关系	不同收入【X】群体的身高【Y】详细差异情况？	● X(定类) ● Y(定量)
单样本 T 检验	差异关系	身高是否明显等于 1.8	定量
配对 T 检验	差异关系	注射新药和没有注射的两组老鼠，血压一样吗	● 配对 1(定量) ● 配对 2(定量)
逐步回归	影响关系	帮我自动找出影响身高【Y】的因素【X】	● Y(定量) ● X(定量/定类)
分层回归	影响关系	身高【X, 分层 1】对于体重【Y】的影响，再加入饮食习惯【X, 分层 2】，看看饮食习惯对体重的影响有多严重	● Y(定量) ● 分层 1(定量/定类) ● 分层 2(定量/定类) ● 分层 3(定量/定类) ● 分层 4(定量/定类)
正态性检验	正态检验	数据正态吗	定量
非参数检验	差异关系	身高数据不正态时，我想研究收入【X】与身高【Y】的差异关系	● Y(定类) ● X(定量)
双因素方差	差异关系	性别【X】和地区【X】对于身高【Y】的差异	● Y(定量) ● X(定类,2 个)
二元 Logit	影响关系	哪些因素【X】影响人们是否购买电影票【Y】	● Y(定类) ● X(定量/定类)