



市场调查与分析

实验课——假设检验和方差检验

谢光明

成都信息工程大学物流学院



核心的关注点——统计推断

❖ 通过样本推断总体



❖ 假设检验的基本原理

假设检验的方法

零假设与备择假设

显著性与显著性水平

假设检验的两类错误



1.1 假设检验的方法

- ❖ 首先对某个未知参数提出一个假设，然后利用样本信息，判断这一假设是否成立，做出接受或拒绝这个假设的决定。
- ❖ 未知参数通常是指均值或方差。



❖ 1.2 零假设与备择假设

零假设 H_0 :

通常是假设未知参数（均值或方差）**等于**某一个值（检验值），或者是假设两个变量的均值或方差**相等**，这样的假设被称为零假设。

备择假设 H_1 :

与零假设对立的假设被称为备择假设。常是假设未知参数（均值或方差）**不等于**某一个值（检验值），或者是假设两个变量的均值或方差**不相等**。



❖ 例如，年龄在25-35岁和35-45岁人群的在网上购物花费有无显著差异？

H_0 : 两组人群的网上购物花费的均值无显著差异；

H_1 : 两组人群的网上购物花费的均值有显著差异。

假设检验就是要判断零假设是否成立，做出接受或拒绝这个零假设的决定。



1.3 显著性与显著性水平

例:科技文教人员与公司职员平均假日支出是否存在差异?

H_0 无显著差异;

H_1 有显著差异。

如果科技文教人员的平均假日支出是3600元,而公司职员的平均假日支出也是3600元,很明显,两者没有差别,与零假设一致。所以,接受零假设,即认为科技文教人员与公司职员的平均假日支出无显著差异。



——思考——



那么，差距多小才能认为“无显著差异”？差距多大才能认为“有显著差异”？

如果科技文教人员的平均假日支出是3625元，而公司职员平均假日支出是3580元，是应该接受零假设还是拒绝零假设呢？

显著性 (Significance)



那么，差距多小才能认为“无显著差异”？差距多大才能认为“有显著差异”？

显著性 (Significance)：

即P值，在SPSS的输出结果中，P值以显著性Sig.表示。

显著性水平：就是拒绝零假设的标准。

SPSS默认的显著性水平为0.05。

当Sig. < 0.05时，拒绝零假设，即认为“有显著差异”

当Sig. > 0.05时，接受零假设，即认为“无显著差异”



1.4 假设检验的两类错误

在假设检验中，是拒绝还是接受零假设是依赖抽样得到的样本而做出的判断，因为随机抽样数据会受到抽样误差的影响，所以假设检验的结果也具有随机性，那么就有可能做出与真实情况不符的决策。

接受零假设只是表明在给定的显著性水平下，样本数据不足以拒绝它，并不意味着它肯定是正确的；同理，拒绝零假设也只是表明在给定的显著性水平下，样本数据不足以接受它，并不意味着它肯定是错误的。



所以，假设检验就有可能发生两类错误。

第一类错误，也称为**弃真错误**，是指零假设实际上是真实的，而检验结果却拒绝了它。

第二类错误，也称为**取伪错误**，是指零假设实际上是不真实的，而检验结果却接受了它。

显著性水平就是弃真错误发生的概率



例：检验某种新药的疗效。

H_0 ：该药未提高疗效； H_1 ：该药提高了疗效。

第一类错误：本来无效，但结论为有效，此时若推广此药，对患者不利。
(弃真)

第二类错误：本来有效，但结论为无效，此时若不推广此药，会带来经济上的损失。
(取伪)



假设检验

- ❖ 1、均值比较
- ❖ 2、单样本T检验
- ❖ 3、独立样本T检验
- ❖ 4、配对样本T检验



□ 均值过程

- ❖ 均值过程的基本功能是分组计算指定变量的描述统计量。包括**均值、标准差、总和、观测量数、方差**等一系列单变量描述统计量。还可以给出方差分析表和线性检验结果。

例：要计算学生的平均身高，“性别”变量把学生按性别分为男、女生两组，均值过程可以分别计算男、女生平均身高。

- ❖ **注意：均值比较中，用于形成分组的变量，应该是其值数量少且能明确表明其特征的变量。**



- ❖ 1、从菜单栏里选择“分析”-“比较均值”-“均值”命令，打开“均值”对话框。
- ❖ 2、在“均值”对话框左侧变量列表中选中结果变量，将其移入到“因变量”列表框，并将要素变量移入“自变量列表”列表框。单击“选项”按钮，打开“均值：选项”对话框。
- ❖ 3、选择所需要分析的统计量，一般选择“均值、标准差、最大值、最小值”等统计量，见下图。
- ❖ 4、单击“继续”，回到“均值”对话框，单击“确定”，查看分析结果。



- ❖ 结果分析:
- ❖ 观察数据描述值，对每一个变量进行了分组，直观进行比较。
- ❖ 使用均值过程求若干组的描述统计量，目的在于比较。因此必须分组求均值。这是与描述性统计过程不同之处。



课堂练习5



请使用“**出口量**”文件，做**均值比较**分析，结果导出为WORD文档，并在文档中对比较结果进行分析阐述。

15分钟



spss操作：假设检验——T检验



进行均值比较常用T检验，按照不同的比较方式spss提供3个t检验过程，包括：

- 单个样本t检验;
- 独立样本t检验;
- 配对样本t检验



单个样本的T检验目的是：

推断总体均值与检验值之间是否存在显著差异

即：检验单个变量的均值是否与给定的常数（一般为理论值、标准值或经过大量观察所得的稳定值等）之间存在差异。**样本均数与总体均数之间的差异显著性检验属于单个样本T检验。**



操作步骤:

- 1、建立检验假设，确定原假设 H_0 ；
- 2、根据样本值计算检验统计量的值；
- 3、根据显著性水平 α ，确定拒绝域；
- 4、做出统计推断；



打开“温度”.这是特定地区城市的平均温度，假设已知普通城市平均温度是20度，根据该数据文件推断特定地区城市温度的平均数是否与普通城市有所不同。

操作：

- 1、选择分析——比较均值——单样本t检验
- 2、将变量“年平均温度”选入“检验变量”中
- 3、在“检验值”框中输入检验值——平均值20
- 4、单击确定，输出结果。



单个样本统计量

| | N | 均值 | 标准差 | 均值的标准误 |
|-------|----|--------|--------|--------|
| 年平均温度 | 33 | 14.752 | 4.8959 | .8523 |

单个样本检验

| | 检验值 = 20 | | | | | |
|-------|----------|----|----------|---------|--------------|--------|
| | t | df | Sig.(双侧) | 均值差值 | 差分的 95% 置信区间 | |
| | | | | | 下限 | 上限 |
| 年平均温度 | -6.158 | 32 | .000 | -5.2485 | -6.984 | -3.512 |

分析结果：样本均值和检验值的差为-5.2485，显著性水平小于0.05，所以拒绝原假设。——两者有显著差异。

- 请自行复习、练习10分钟。



- ❖ 已知某水样中含碳酸钙 CaCO_3 的真值为20.7mg/L，现用某方法重复测定该水样11次 CaCO_3 的含量（mg/L）为：20.99，20.41，20.10，20.00，20.91，22.60，20.99，20.41，20.00，23.00，22.00。问该方法测得的均值是否偏高？
- ❖ 15分钟，包括录入数据，结果导出为WORD文档，并进行结果分析。