

七 平均数分析(Means)与T 检验

平均数分析是最简单,但也是使用频率最高的一个统计分析方法。我们对随机样本进行统计描述以后,还要由此来推论总体情况。平均数分析就是用样本均值来推论总体均值的方法。具体地说来,平均数分析包括下面四部分内容。一是单个样本的T 检验。这是用样本的均值对总体均值的假设进行检验的方法。二是独立样本的T 检验。这是用两个样本的均值差的大小来检验对应的两个总体的均值是否相等的方法。三是配对样本的T 检验。这是通过对配对样本的两次测量结果的差异的大小来检验两个总体的差异是否显著的分析方法。四是一般的均值分析方法。是通过分析多个样本的均值差异的大小,来确定分类变量与分析变量是否独立的方法。下面分别介绍。

学习要求:

- 1、熟练掌握单个样本的T 检验
- 2、熟练掌握独立样本和配对样本的T 检验
- 3、掌握平均数分析(Means)的方法

7.1 单个样本的T 检验

7.1.1 单个样本T 检验的过程

单个样本的T 检验的目的是通过计算出的样本均值来估计总体均值是否为某个确定的值。其基本思想是:计算出样本均值以后,先根据经验或以往的调查结果,对总体的均值提出一个假设,即 $\mu = \mu_0$ 。 μ_0 就是待检的总体均值。然后分析计算出的样本均值来自均值为 μ_0 的总体的概率有多大。如果概率很小,就认为总体的均值不是 μ_0 。

在大样本的情况下($n \geq 30$),根据样本均值来检验总体均值时,在总体方差未知的情况下采用T 检验。在小样本的情况下($n < 30$),要使用T 检验的方法,要求总体服从正态分布。单个样本T 检验的步骤如下:

1. 打开T 检验对话框,执行下述操作: Analyze→Compare Means→One Sample T Test 打开对话框,如图8-1 所示。

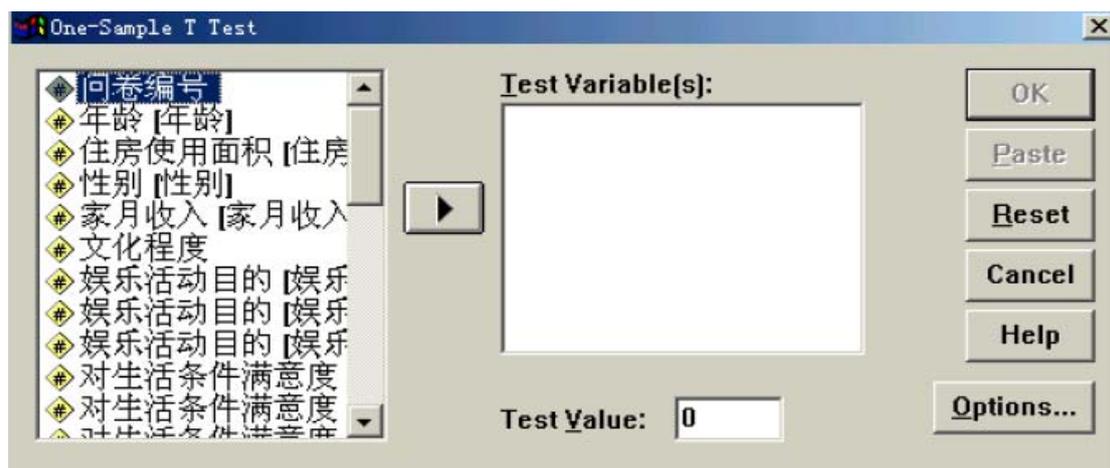


图8-1 单个样本T 检验对话框

2. 选择分析变量

图9-1 中的Test Variable(S)窗口是分析变量窗口。可以从左边的源变量窗口中选择将要分析的一个或多个变量进入到该窗口中。

3. 确定待检参数

Test Value 窗口中输入的数据是待检的总体均值，也就是假设检中原假设的值。

4. 确定置信度和缺失值的处理方法

单击Options 按钮，打开选项对话框，如图8-2 所示。

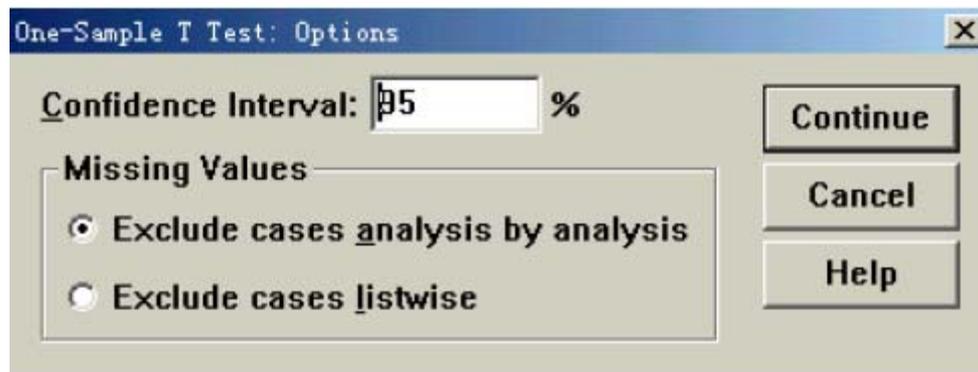


图8-2 单个样本T 检验选项对话框

(1) Confidence Interval 窗口用于设置检验的置信度。系统默认值为95%。即在原假设成立的条件下，样本均值出现的概率如果小于5%，则不能接受原假设。

(2) Missing Values 是设置缺失值的处理方法的选项栏。

1) Exclude cases analysis by analysis 是只剔除分析变量为缺失值的个案，这是系统默认状态。

2) Exclude cases listwise 是剔除任何含有缺失值的个案。

上述选项作完以后，单击Continue 按钮返回单个样本T 检验对话框。

5. 单击OK 提交运行。系统在输出文件窗口中输出T 检验的结果。

7.1.2 单个样本T 检验的实例

实验：对“休闲调查”中的“住房面积”变量进行单个样本的T 检验。原假设为：城市居民的户均住房面积为38 平方米。

7.2 独立样本的T 检验

7.2.1 独立样本的T 检验的过程

独立样本的T 检验的目的是，通过比较两个样本均值差的大小来确定两个总体的均值是否相等。如果两个样本相互独立且均为大样本，可以采用T 检验的方法。如果两个样本相互独立但都是小样本，或有一个样本是小样本，则要求总体服从正态分布。检验的基本思想是，按照一定的分组原则将所有的个案分为两部分，可将这两部分视为两个独立的样本，对两个样本分别进行统计描述。然后对两个样本进行方差齐性检验，或称为等方差检验。确认两个总体具有等方差性以后，用t 检验的方法对两个样本的均值差进行检验。如果均值差过大，则说明这两个样本来源于均值不同的两个总体。对两个样本均值差的检验基本过程为：

1. 打开独立样本的T 检验对话框

执行下述操作：

Analyze→Compare Means→Independent- Samples T Test 打开对话框，如图8-3 所示。

2. 选择分析变量

图8-3 中的Test Variable(S)窗口是分析变量窗口。可以从左边的源变量窗口中选择将要分析的一个或多个变量进入到该窗口中。

3. 确定分组变量

Grouping Variable 是分组变量的窗口。选择一个变量作为分组依据进入GroupingVariable 窗口，同时激活Define Groups 按钮。由于独立样本的T 检验是对两个总体的均值差进行检验，所以作为分组依据的变量只能取两个值。



图8-3 独立样本的T 检验对话框

4. 确定分组变量的取值

单击Define Groups 按钮打开对话框，如图8-4 所示。以确定分组变量的取值。

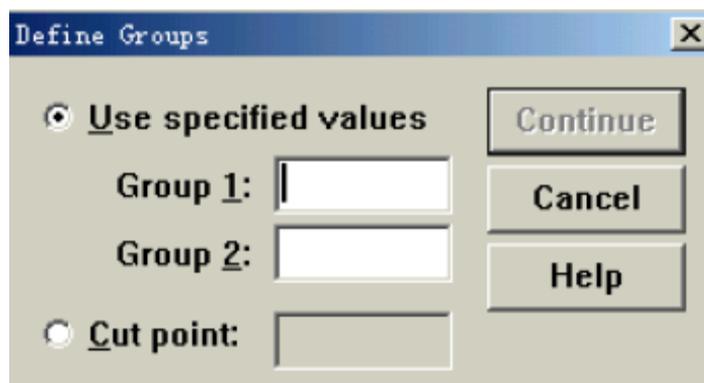


图8-4 确定分组对话框

(1) 如果分组变量是二分定类变量

选择Use specified values 选项。在Group1 和Group2 窗口中分别输入变量的两个取值。系统将以这两个取值为依据，将所有的个案分为两部分，并对分析变量分别进行统计。

(2) 分组变量是一般的离散型的数值变量

分组变量是一般的离散型的数值变量。而且希望对其中某两个类别的均值差异进行分析的话，也可选择Use specified values 选项。在Group1 和Group2 窗口中分别输入分组变量的值。系统将以这两个值为依据，对分析变量分别进行统计。

(3) 分组变量为连续型变量，或虽为离散型变量，但变量的取值较多在这种情况下，只希望以分组变量为依据将个案分为两部分时，应选择Cut points (切分点) 选项。同时激活Cut point 窗口。并在窗口中输入分割值。系统将所有观测值分为大与该数值和小于该数值的两组，并分别计算分析变量的均值然后进行比较。

(4) 如果选择的分组变量是字符串变量，单击Define Groups 按钮后，打开的是如图8-5 所示对话框。

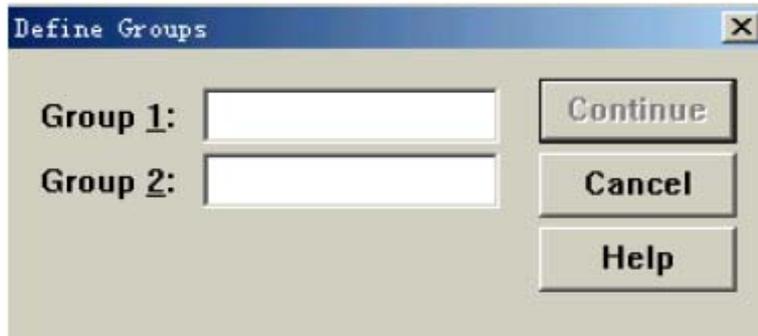


图8-5 确定分组对话框

可直接在两个窗口中输入变量值即可。

上述选择做完以后，单击Continue 按钮返回主对话框。

5. 确定置信度和缺失值的处理方法

单击Options 按钮，将打开一个与单个样本T 检验选项对话框完全相同的对话框。由来设置置信度和缺失值的处理方法。如果采用系统默认选项，该对话框可以不做任何设置。单击Continue 返回主对话框。

6. 单击OK 提交运行。系统在输出文件窗口中输出独立样本的T 检验的结果。

7.2.2 独立样本的T 检验的分析实验

实验1：用“休闲调查” 中的数据分析不同性别的被调查者家庭的住房面积是否有差异。

实验2：用“休闲调查” 中的数据分析初中学历与高中学历的被调查者的住房面积是否有差异。

实验3：用“休闲调查” 中的数据分析40岁以下与40岁以上被调查者的住房面积是否有差异。

7.3 配对样本的T 检验

7.3.1 配对样本T 检验过程

配对样本是指对同一样本的某个变量进行前后两次测试所获得的两组数据。或是对两个完全相同的样本在不同条件下进行测试所获得的两组数据。可以把这两组数据看作是分别来自两个总体的两个样本。配对样本的检验是先求出每对观测值的差，再对差值求平均值。通过检验配对变量均值之间的差异的大小，来确定两个总体的均值的差异是否显著。检验过程如下：

1. 打开配对样本的T 检验对话框

执行下述操作：Analyze→Compare Means→Paired-Samples T Test 打开对话框，如图8-6 所示。

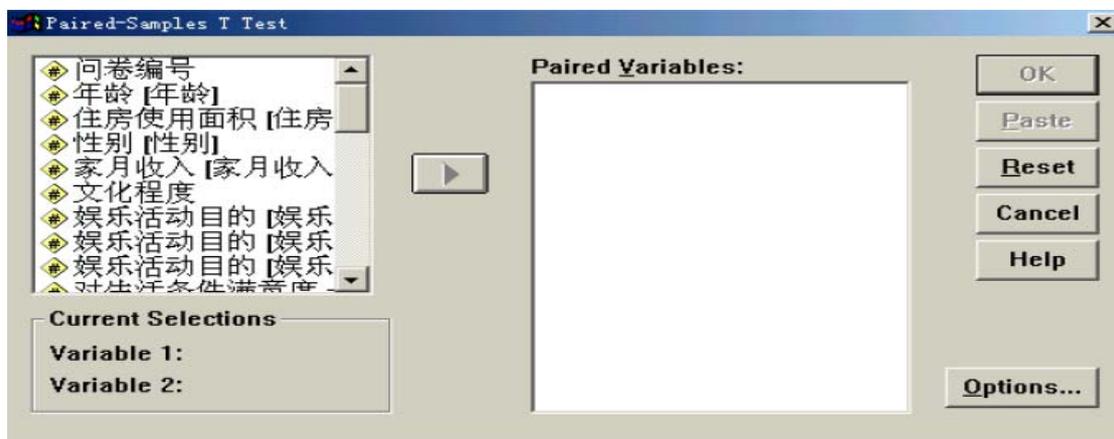


图8-6 配对样本的T 检验对话框

图中的Paired Variables 窗口是用于显示已经配对的变量。

2. 选择分析变量

从左边的源变量框中选择两个变量。使之进入Paired Variable 窗口中。两个变量只占一行，可选择多对配对变量。

3. 确定置信度和缺失值的处理方法

单击Options 按钮，将打开一个与单个样本T 检验选项对话框完全相同的对话框。由自己来设置置信度和缺失值的处理方法。如果采用系统默认选项，该对话框可以不做任何设置。单击Continues 按钮返回主对话框。

4. 单击OK 提交运行。系统在输出文件窗口中输出配对样本的T 检验的结果。

7.3.2 配对样本T 检验的分析实验

实验：数据文件“贫困调查”中的“满意度1”是贫困人口获得最低生活保障金以前的生活满意度测量结果，“满意度2”是获得最低生活保障金以后的生活满意度测量结果。检验获得最低生活保障金前后生活满意度是否有变化。