

4.4 选择个案

选择个案的功能是选择出一些符合一定条件的个案，分析时将只对这些个案进行分析。

4.4.1 选择个案的方法

1. 打开选择个案对话框

执行下述操作：Data→Select Cases 打开对话框，如图4-12所示。

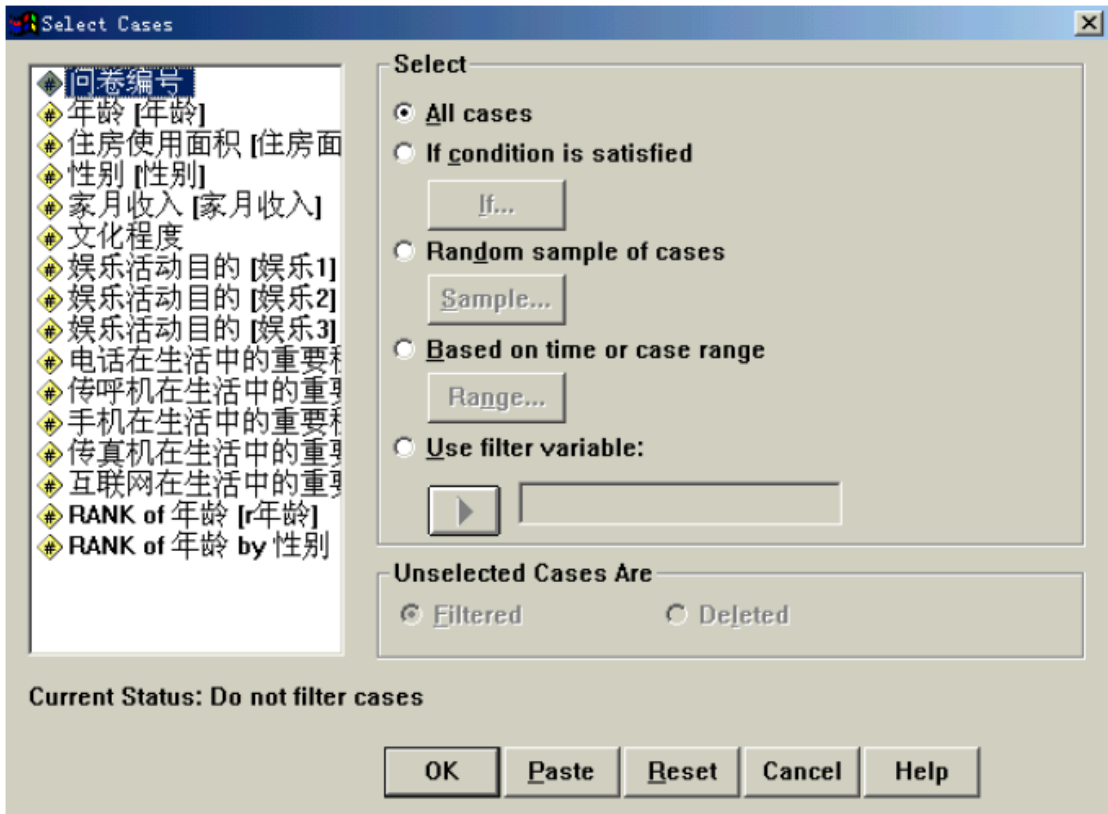


图4-12 选择个案对话框

2. 确定选择个案的方法

Select 是确定选择个案的方法的选项栏，其中包含五个选项。

- (1) all cases 是所有的个案都选择。这是系统默认状态。执行后回到原来状态。
- (2) If condition is satisfied 是满足一定条件的选项。选择此项后，激活If 按钮，单击If按钮，进入选择条件对话框，如图4-13 所示。

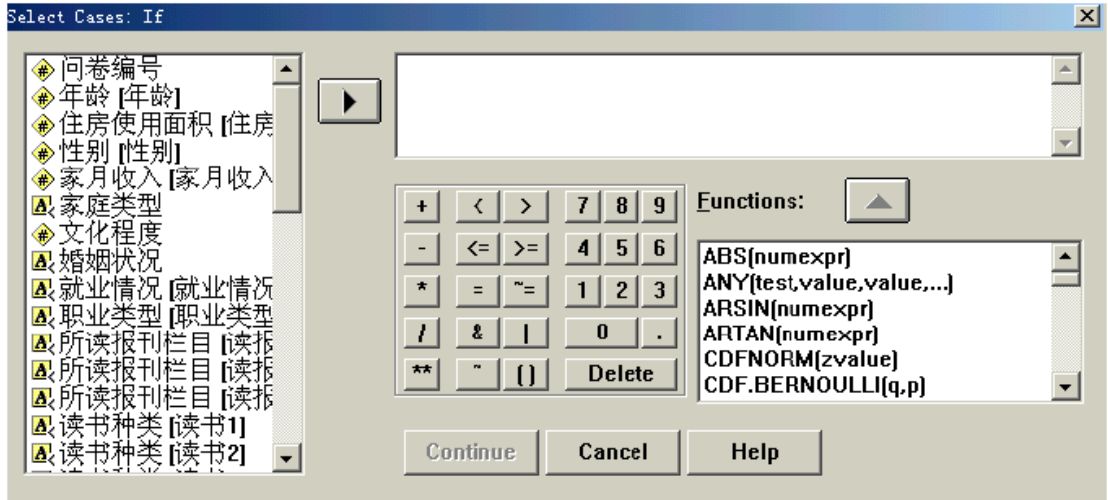


图4-13 选择条件对话框

在右上方的窗口中输入条件表达式。条件表达式中的变量名可以从左侧的原变量窗口中选择，其他符号或数字可以从图4-22 中的模版上选择。单击“Continue”，回到条件对话框。

(3) Random sample of cases 是随机选择个案的选项。即在所有的个案中按随机抽样的方式选择个案。选择此选项后，激活Sample 按钮。单击Sample 按钮，进入Random sample（随机样本）对话框，如图4-14 所示。

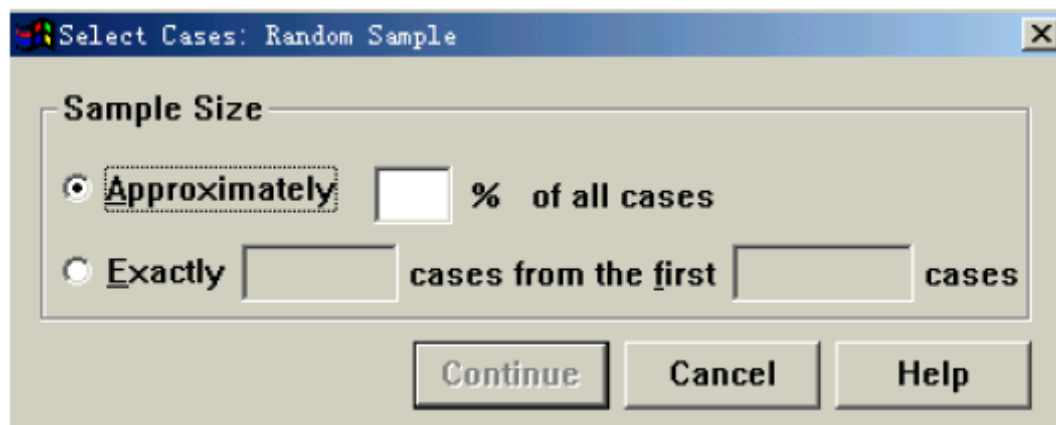


图4-14 随机选择个案对话框

1) Approximately 是按照给定的百分比从所有个案中随机选择个案的选项。将百分比数输入到Approximately 后面的窗口中。

2) Exactly 是在前若干个个案中随机选择一定数量的个案的选项。选择此选项后，激活后面的两个窗口。在前一个窗口中填入要选择的个案数。在后一个窗口中填入选定范围的个案数。

单击“Continue”，回到条件对话框。

(4)Based on time or case range 是在设定的范围内随机选择个案的选项。选择此选项后，激活Range（范围）按钮。单击Range 按钮，进入Range 对话框，如图4-15 所示。

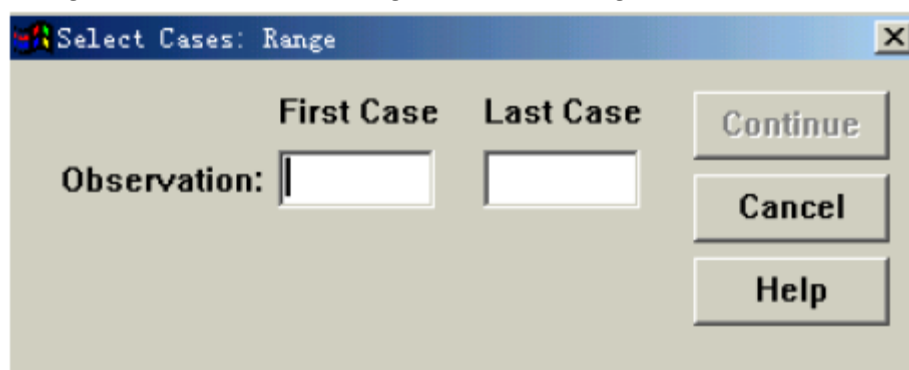


图4-15按范围选择个案对话框

Observation 是样本观测值。将所设定的样本观测值的最小值输入到First Case 下面的窗口中。将所设定的样本观测值最大值输入到Last Case 下面的窗口中。这就确定了选择个案的范围。

单击“Continue”，回到条件对话框。

(5)Use filter variable 是使用过滤变量的选项。

选择Use filter variable 后，激活下面的窗口。Use filter variable 是要求选择一个变量来过滤个案。这个变量称为过滤变量。过滤变量的值为0 时，个案将不被选中。可以

从左侧的源变量窗口选择一个过滤变量进入Use filter variable 下面的窗口中。

3. 确定未被选中的个案的处理方法

Unselected Cases Are 是确定未被选中的个案的处理方法的选项栏。该栏中包括两个选项。

(1) Filtered 是生成过滤变量的选项。选择此项后，系统在执行完选择个案的命令后将生成一个新的变量“Filter”，其变量值说明了哪些是被选择的个案和哪些是未被选择的个案。未被选择的个案在数据窗口最左侧的个案序号上打上了斜线。以后的分析将只对选择的个案进行。这也是系统默认选项。

(2) Deleted 是删除未选个案的选项。选择此项后，系统在执行完选择个案的命令后将删除未被选择的个案，数据窗口中只保留被选中的个案。如果没有特殊需求，建议初学者不要选择此项，以避免数据丢失。

4. 上述选项作完以后单击“Ok ”提交运行。

4.4.2 选择个案的实验

实验：在“休闲调查1”的数据文件中，按不同的条件选择个案。

1. 选择年龄在50 岁以上的个案
2. 从前100 个个案中随机选择30 个个案
3. 选择第3 到第6 个个案
4. 选择“文化程度”为过滤变量。

4.5 数据文件的转置

对社会调查的结果进行分析时，有时既需要对变量进行分析，也需要对个案进行分析。如由10 名评委对20 名选手进行打分。既需要分析每名选手的得分情况以排出高低。也许要研究每个评委的打分，以确定是否公正。SPSS 的数据分析主要都是针对变量进行的。当对个案进行分析时，需要把个案转换成为变量。这个功能由Transpose(转置)命令来完成。

4.5.1 数据文件转置的方法

1. 打开数据文件转置对话框

执行下述操作： Data→Transpose 打开对话框，如图4-16 所示。

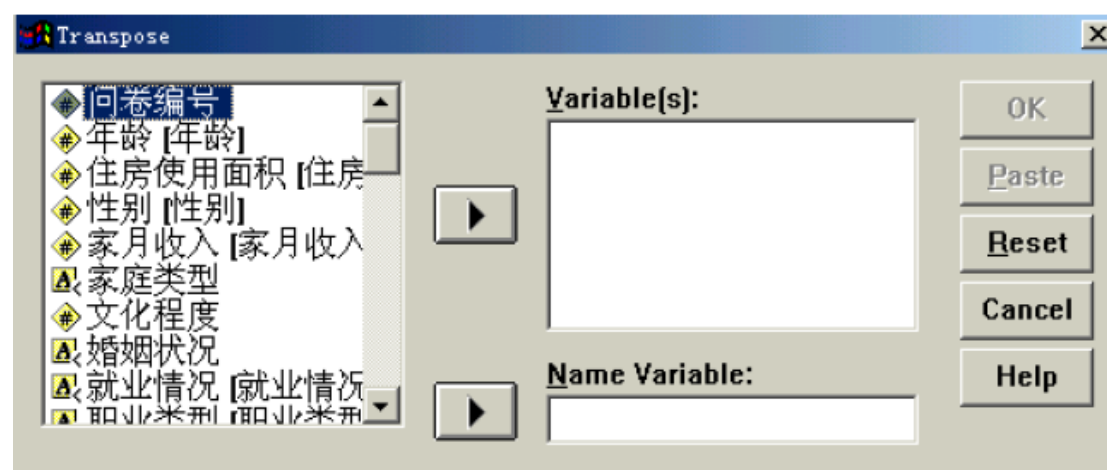


图4-16 数据文件转置对话框

2. 选择转置置变量

从左侧的源变量窗口中选择欲转置的变量，使之进入Variable(s)窗口。执行转置命令后，这些被选择的变量将变为个案。未选择的变量将消失。

3. 选择命名变量

如果原数据文件中包含一个序号变量，或者某个变量在每个个案上的取值都是唯一的。可以将这个变量作为“名称变量”。并使之进入到Name Variable 窗口。转置后，“名称变量”的值变为新变量的变量名。如果“名称变量”是数值型变量，则新变量的名以字母V开头，后面的是“名称变量”的变量值。

如果不选择“名称变量”，系统会自动给转置后的变量命名为var001、var002、var003……等等。

4. 上述选项作完以后，单击“Ok ”提交运行。可以在数据窗口中看到转置后的数据文件。

4.5.2 数据文件转置的实验

实验：将“休闲调查 1”中的“满意度 1”、“满意度 2”、“满意度 3”、“满意度 4”、“满意度 5”、“满意度 6”、“满意度 7”进行转置。

4.6 变量内容的重新编码

在统计分析时，有时需要把定距以上变量化为定序变量或定类变量。比如，把人们的年龄与它们的身体状况进行交叉分析时，研究者所关心的并不是某一年龄的人们的身体状态如何，而关心的是某一年龄段的人们的身体状态如何。这就需要对年龄变量进行重新编码，将其转化为定序变量。转换方法有两种：一是用重新编码的变量取代原来的变量。二是用重新编码的变量生成一个新变量。单击主菜单中的Transform 的下拉菜单中的Recode，在左侧引出二级菜单，如图4-22所示。其中Into Same Variables 是用重新编码的变量取代原来的变量，Into Different Variables 是用重新编码的变量生成一个新变量。

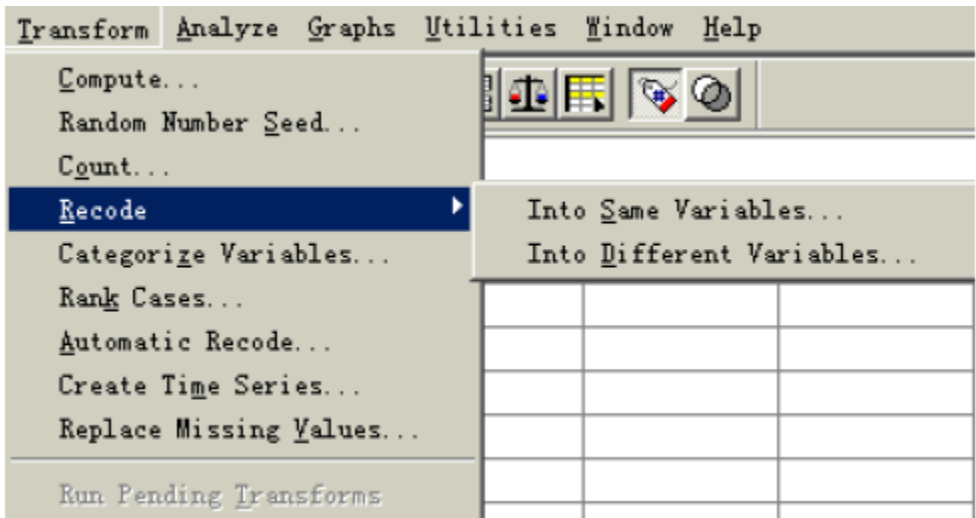


图4-22 重新编码的指令菜单

4.6.1 用重新编码的变量取代原来的变量

1. 用重新编码的变量取代原来的变量的方法

(1) 执行下述操作。

Transform→Recode→Into Same Variables 打开重新编码对话框，如图4-23 所示。

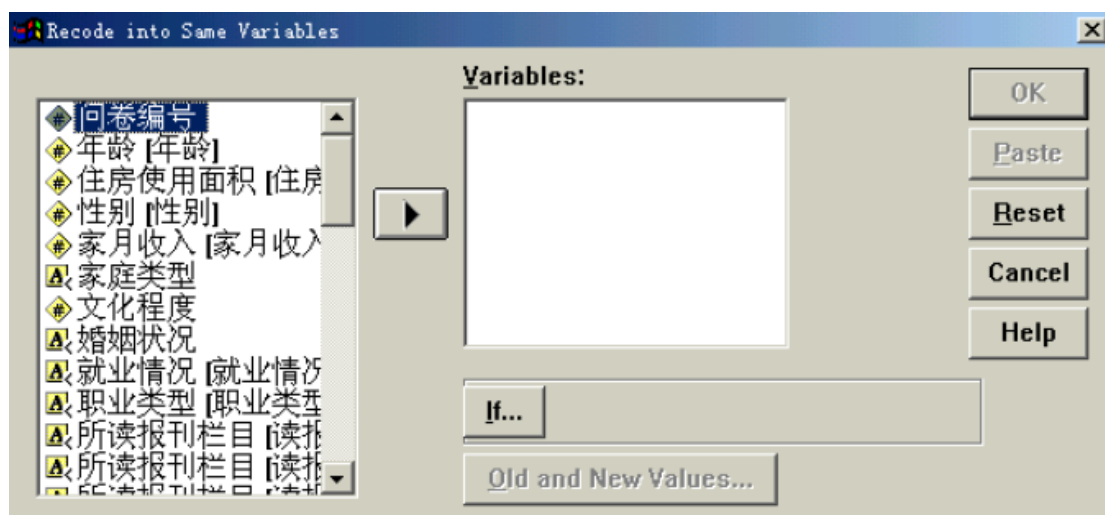


图4-23 重新编码对话框

(2) 确定从新编码的变量

从左侧的源变量窗口中选择将要重新编码的变量进入到Variables 窗口中，同时激活Old and New Values（新旧变量值的转换）按钮。

(3) 新旧变量值的转换方法

单击Old and New Values 按钮，进入新旧变量值转换对话框，如图4-24 所示。

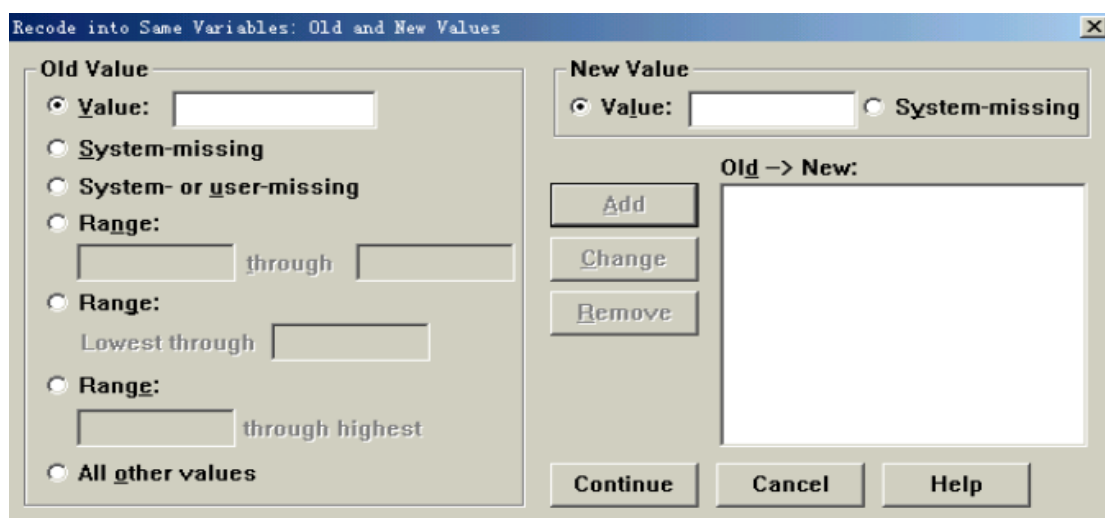


图4-24 新旧变量值转换对话框

该对话框包含三组内容。左侧的部分的Old Value 是旧变量值选项栏。右上部分的New Value 是新变量值选项栏。右下部分的Old→New 是新旧变量的转换关系。下面根据变量值的测量层次的不同，分别介绍新旧变量值的转换方法。

1) 如果要转换的旧变量值是离散数据，且新旧编码虽不同但一一对应。则在Old Value选项栏中选择Value，这也是系统默认状态，并将旧值输入到Value 窗口中。然后在New Value选项栏中选择Value，并将新值输入到Value 窗口中，同时Old→New 下面窗口的Add 按钮被激活。单击Add 确认重新编码结果。新旧编码的对应值将出现在Old→New 下面的窗口中。对重新编码的每一个值都重复上述操作，直至完成所有值的转换。

2) 如果要转换的旧变量是连续性的数据，或虽然是离散型数据，但变量的取值范围很大。而且新变量的一个值对应的是旧变量的一个区间。这实际上是把旧变量按一定的组距划分成组，每一组对应着新变量的一个值。转换方法如下。

第一步：确定旧变量的最低组与新变量的对应值。即确定旧变量从最小值到某一数据之间的区间。在Old Value 选项栏中选择中间的Range,同时激活下面的窗口。将最低组的上限输入到Lowest through 后面的窗口中。然后在New Value 选项栏中选择Value,并将新值输入到Value 窗口中,同时Old→New 下面窗口的Add 按钮被激活。单击Add 确认重新编码结果。新旧编码的对应值将出现在Old→New 下面的窗口中。

第二步：将旧变量的各个组与新变量的各个对应值一一输入。在Old Value 选项栏中选择最上面的Range,同时激活下面的两个窗口。将表示区间下限的数值输入到前面的窗口中,将表示区间上限的数值输入到后面的窗口中。即确定两个数值之间的区间。然后在New Value选项栏中选择Value,并将新值输入到Value 窗口中,同时Old→New 下面窗口的Add 按钮被激活。单击Add 确认重新编码结果。

第三步：确定旧变量的最高组与新变量的对应值。即确定旧变量从某一数值到最大值之间的区间。在Old Value 选项栏中选择最下面的Range,同时激活下面的窗口。将最高组的下限输入到Range 下面的窗口中。然后在New Value 选项栏中选择Value,并将新值输入到Value 窗口中,同时Old→New 下面窗口的Add 按钮被激活。单击Add 按钮确认重新编码结果。

(4) 对缺失值的处理方法

如果对系统缺失值或系统与缺失值进行重新定义时,可在Old Value 选项栏中选择System-missing (系统缺失值)或选择System-or User-missing (系统与用户缺失值),并将新定义的缺失值内容输入到New Value 框中的Value 窗口中,单击Add 确认。如果对缺失值的内容不做转换,则在Old Value 框中做完选择后,在New Value 框中直接选择System-missing,并单击Add 确认。

2. 用重新编码的变量取代原来的变量的实验

实验：将“休闲调查1”中的“性别”变量的编码由1=男、2=女转变为11=男、22=女。

4.6.2. 用重新编码的变量生成一个新变量。

1. 用重新编码的变量生成一个新变量的方法。

(1) 执行下述操作。

Transform→Recode→Into Different Variables 打开Recode Into Different Variables (重新编码生成不同变量)对话框,如图4-25 所示。

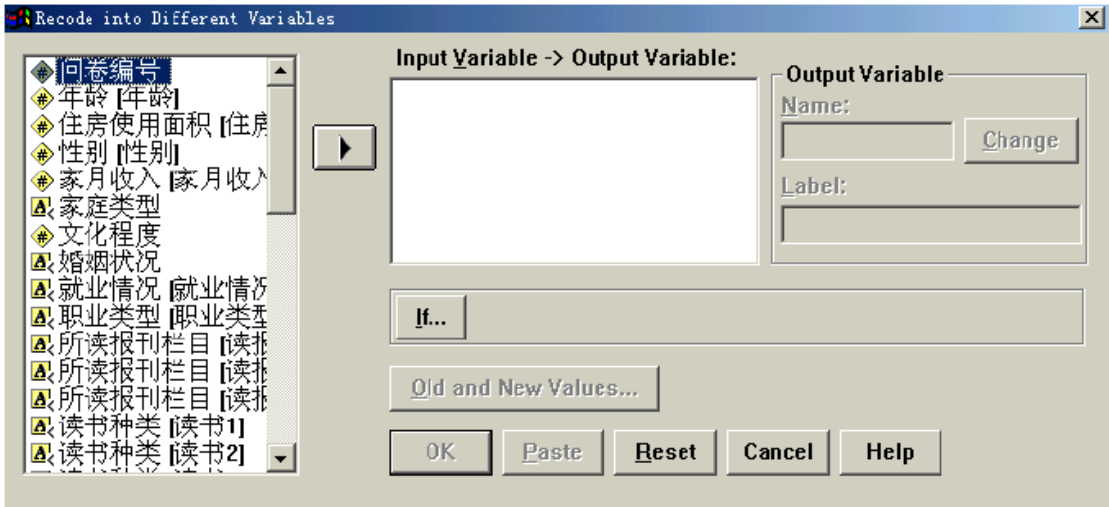


图4-25 重新编码生成新变量对话框

(2) 确定重新编码的变量

从左侧的源变量窗口中选择将要重新编码的变量进入到Input Variable→Output Variable

下面的窗口中，此时Name 和Label 窗口被激活。

(3) 确定新变量名和标签

在Output Variable 栏中的Name 和 Label 的窗口中分别输入新的变量名和变量名的标签，此时Change 按钮被激活。单击Change 按钮予以确认。

(4) 新旧变量值的转换方法

单击Old and New Values 按钮，进入新旧变量值的转换对话框，如图4-26 所示。

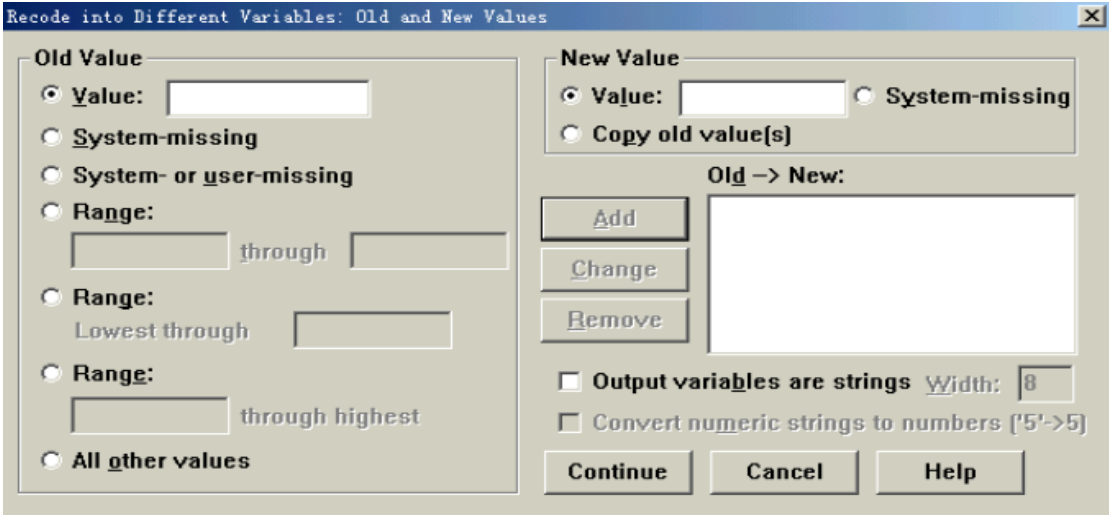


图4-26 新旧变量值转换对话框

该对话框与用重新编码的变量取代原来的变量的新旧变量值转换对话框基本相同。新旧变量值的转换方法也与用重新编码的变量取代原来的变量的新旧变量值的转换方法基本相同。下面仅对不同的部分进行介绍。

- 1) 如果原变量的分布中有较大的离群值或特异值，无法包含到任何一个分组中时，可以在Old Value 选项栏中选择All other values 选项。并在新变量中为其定义一个值。
- 2) 如果新变量的值与旧变量相同，在Old Value 选项栏中的Value 窗口中输入旧变量的值，然后在New Value 选项栏中选择Copy old value(s)选项
- 3) 如果新生成的变量是字符串变量，必须选择Output Variables are strings 选项，并在Width 窗口中输入变量值的宽度。然后按照新旧变量值的转换方法进行转换和确认。

(5) 对缺失值的处理方法

与用重新编码的变量取代原来的变量的方法完全相同

2. 用重新编码的变量生成一个新变量的实验。

实验：将“休闲调查1”中的“年龄”变为每10 年为一个年龄段的定序变量，并存储为新变量。