**相关分析**

**1、相关资料链接**

①spss的相关系数案例链接

<http://wenku.baidu.com/view/244173a95022aaea988f0f63.html?re=view>

②spss简单相关案例

<http://www.docin.com/p-208142836.html>

③Pearson相关系数

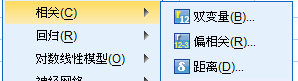
<http://wenku.baidu.com/link?url=2-0EmpeYgFABB-XCs4fykOs3HxTLqp167tplCMKQpRbNpOuVud0USwxluqk49-EAAFCVS2jAHdoub5PzmVGgY3X_rxeFFrXOTYJoe33sCNu>

④t检验表

<http://wenku.baidu.com/link?url=bEsZ1XoF9mE3J56YynsPf_d0r8wg9cyHrFf91FAZvMIf4TI9JwD7duhgijwzW__HBkl9M21PFiAViA4NckwRKgxGL9sqSH0l2MnGIfA6rCq>

1. **基本菜单**

点击“相关”，有三个子变量：双变量（简单相关）、偏相关、距离分析



①双变量：



②偏相关：

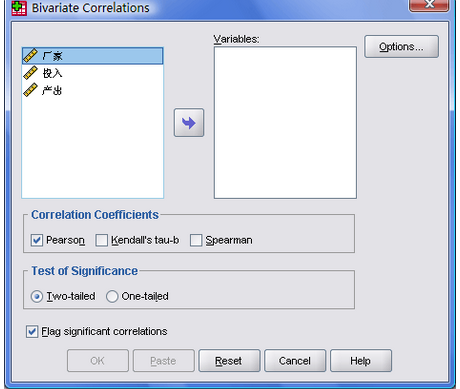


③距离分析：

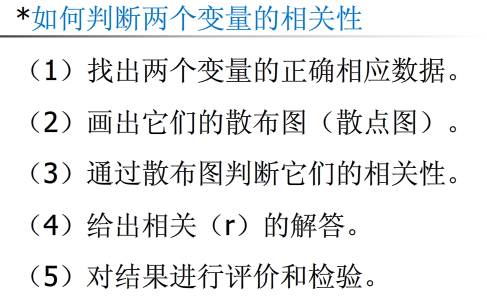


点击“双变量”，我们需要计算出三个相关系数，分别为：Pearson系数、Spearman系数和Kendall系数。具体复选哪个系数，用户可以在复选框中打钩。

检验有：单侧检验或者双侧检验。具体的弹出框如下：

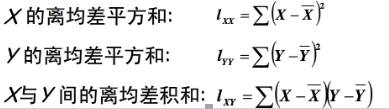


1. **计算“双变量”下的Pearson系数及其案例分析**



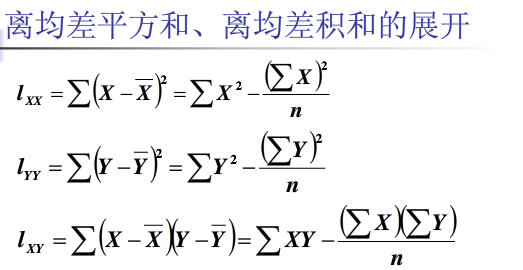
结果：显示相关系数值、t检验值、t临界值和检验结论

**① Pearson系数公式:**





展开：



②所以我们要求：    

③计算出r后，进行t检验

检验值t：

④算出t检验值之后，查表看大小，一般取显著性水平0.05、0.01(单侧），如果是双侧的话，取0.025、0.005。



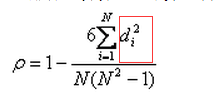
⑤得出结论：如果检验值比查表得到的值大的话，则显著，否则，不显著。

1. 计算“双变量”下的spearman相关系数
2. 相关资源

<http://www.cnblogs.com/zhangchaoyang/articles/2631907.html>

<http://blog.csdn.net/wsywl/article/details/5859751>

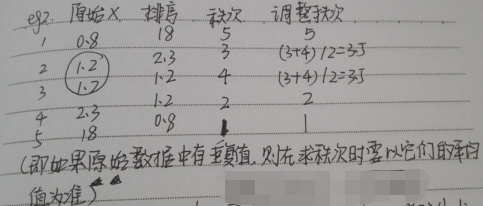
1. spearman公式



【注】我们要求的就是，N为样本的对数（有多少对数据）

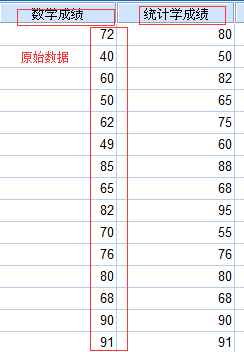
(3)结合案例进行说明计算



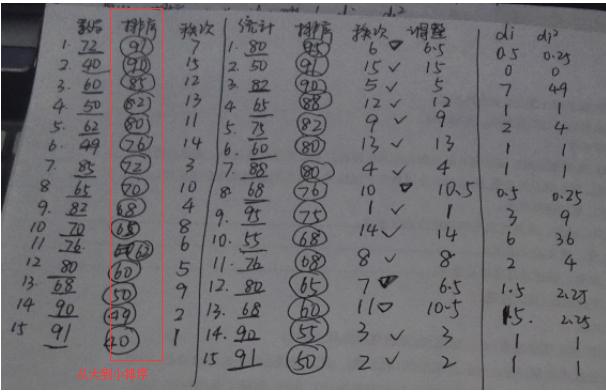


1. 上述案例与计算步骤说明

①变量Xi为原始数据，即我们文件上传后要分析的某列变量，如：



②对原始数据进行排序，注意要么都从大到小，要么都从小到大排序，一般我们用从大到小排序。



③看原始数据值在排序中的第几个。如原始数据“数学”72，在排序中72在第7个位置，所以72的秩次是7，“数学”40，在排序中40在第15个位置，所以40的秩次是15……

【注】如果原始数据有相同的情况，则将相同秩次相加，求平均值。

如在“统计”中80有2个，第一个80的秩次是6，第二个秩次是7，则调整后的秩次是（6+7）/2=6.5，其他不变。

④将比较的两个变量的秩次对应相减，得到di

如7-6.5=0.25； 15-15=0；……

⑤求,累加，得到

例如上述数学和统计的=112

⑥利用spearman的公式，计算最后的spearman系数





偏相关

1. 资源

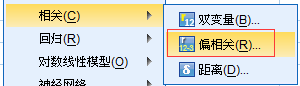
<http://wenku.baidu.com/link?url=kfJ6nXVo5AjEQn0OXV6f057IUsbEpwf45ZNGPoJjknCaqWc34c5z119vmjYOjE9KvxAgMkYJF337PFQh0rHjB6ZehKKoan6W9HATHfVcssm>

1. 偏相关应用

多个变量之间的相关关系是错综复杂的。任何两个变量之间都存在简单相关关系，这种关系又受其他变量的干扰，使得简单相关关系不能确切反映两个变量之间的纯相关关系。在实际工作中，**需要固定一些因素**，留下一个变量看一看这个变量对另一个变量的作用。这时就需要利用偏相关系数。

1. 偏相关的一些定义与公式

**①选择偏相关**



**②弹框**



【说明】

变量只能选择两个，分析A与B之间的关系。

控制：可以是一个、两个或者多个。（我们最多算到2个吧，3个以上的公式我没有找到）

这里的含义是：控制语文成绩，计算英语成绩与统计学的相关关系。（可能英语成绩和统计学成绩受到语文成绩的影响，所以要固定语文成绩。）

【阶数】

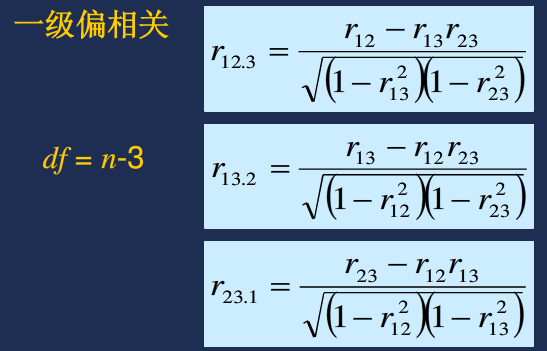
当“控制”变量数目为1个的时候，是一阶，当“控制”变量数目为2的时候，为二阶，依次类推。

【自由度】

此处的自由度为n-q-2（q为阶数）。当q=1时候，df=n-3; 当q=2时候，df=n-4

**③公式**

控制变量数目为1的时候，即1阶时候：



【说明】的意思是，控制变量3，计算变量1与变量2之间的相关关系。

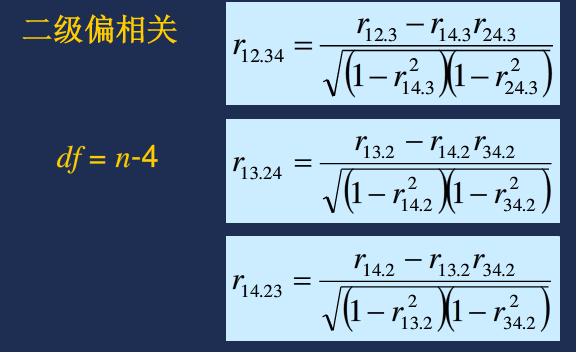
【注意】

**这里的r其实就是前面的pearson相关系数（简单相关）计算而来的，即**

我们可以复用pearson相关系数的方法。计算各个变量之间的pearson相关系数值，然后代入上述公式，既可以得到偏相关一阶系数。



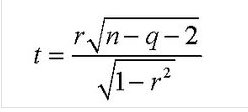
控制变量数目为2的时候，即1阶时候：



【注意】

求二阶偏相关前，也是要求一阶偏相关系数的。所以，一阶偏相关系数我们要求。

**④计算t检验值**



**⑤查t检验表，得出显著性结论。**

对于相关分析的小结

相关分析，我们算两个大类：简单相关和偏相关。另外一个大类距离分析，我们不算。

其中简单相关中，我们计算连个相关系数：pearson系数和spearman系数。还有一个kendell系数我们就不算了~

偏相关我们算一阶和二阶偏相关系数，多阶的先不算。