

## 如果两组数据样本量比较大怎么办——数据量较大的excel文件，如何让它操作时快起来？-股识吧

一、您好！请问用spss做聚类分析时，数据量比较大，树状图上的有些分支没有标示可以怎么解决？谢谢

数据量太大的时候 不适合用系统聚类法 一个就是你遇到的问题 无法判断，另一个就是该方法聚类的结果不是一种情况，你可以根据自己的专业进行适当的选择归类，比较有难度如果你的变量类型都是连续性的数据 可以直接采用k-mean法 指定聚类的类别进行如果你的变量类型既有分类 又有连续性的，则采用 两阶段法 聚类，这种方法会自动推荐出聚类的个数，而且应用非常广泛，结果直观长期兼职 市场研究、数据分析、论文数据分析等qq 94168195

## 二、如果服务器传过来的数据量过大，怎么处理

说白了就是服务器的承受能力。

第一，确认服务器硬件是否足够支持当前的流量。

普通的P4服务器一般最多能支持每天10万独立IP，如果访问量比这个还要大，那么必须首先配置一台更高性能的专用服务器才能解决问题，否则怎么优化都不可能彻底解决性能问题。

第二，优化数据库访问。

服务器的负载过大，一个重要的原因是CPU负荷过大，降低服务器CPU的负荷，才能够有效打破瓶颈。

而使用静态页面可以使得CPU的负荷最小化。

前台实现完全的静态化当然最好，可以完全不用访问数据库，不过对于频繁更新的网站，静态化往往不能满足某些功能。

缓存技术就是另一个解决方案，就是将动态数据存储到缓存文件中，动态网页直接调用这些文件，而不必再访问数据库，WordPress和Z-

Blog都大量使用这种缓存技术。

我自己也写过一个Z-Blog的计数器插件，也是基于这样的原理。

如果确实无法避免对数据库的访问，那么可以尝试优化数据库的查询SQL.避免使用 Select \*from这样的语句，每次查询只返回自己需要的结果，避免短时间内的大量SQL查询。

第三，禁止外部的盗链。

外部网站的图片或者文件盗链往往会带来大量的负载压力，因此应该严格限制外部

对于自身的图片或者文件盗链，好在目前可以简单地通过refer来控制盗链，Apache自己就可以通过配置来禁止盗链，IIS也有一些第三方的ISAPI可以实现同样的功能。

当然，伪造refer也可以通过代码来实现盗链，不过目前蓄意伪造refer盗链的还不多，可以先不去考虑，或者使用非技术手段来解决，比如在图片上增加水印。

第四，控制大文件的下载。

大文件的下载会占用很大的流量，并且对于非SCSI硬盘来说，大量文件下载会消耗CPU，使得网站响应能力下降。

因此，尽量不要提供超过2M的大文件下载，如果需要提供，建议将大文件放在另外一台服务器上。

目前有不少免费的Web2.0网站提供图片分享和文件分享功能，因此可以尽量将图片和文件上传到这些分享网站。

### 三、样本量不同的两组数据如何比较好坏

用独立样本t检验应该可以。

一个变量为组别，一个变量为原始的单个回复量。

### 四、请问（数学问题）：在知道两组样本数据的情况下，怎样较快速的比较两组数据的标准差的大小？

- 1) 如果都服从正态分布，用各自的最大值减去最小值除以6，谁的得数大(小)谁的标准差就大(小)，该得数可作为标准差的近似值；
  - 2) 把数据画出曲线，找出平均线，观察数据相对平均线谁的波动大(小)谁的标准差就大(小)；
  - 3) 再快一点，算出各自的极差，谁的大(小)，谁的标准差就大(小)；
  - 4) 同种变量，用上述方法可比较标准差的大小；
- 对于不同类的变量，最好用变异系数进行比较 更为合理。

### 五、统计问题：两样本比较的秩和检验，样本过大无法用查表法

## ，采用了u检验法，为什么仍属非参数统计？急求...

上面的p值时用z统计量算出来的渐进p值，下面的是精确概率，也就是样本容量很小时，给出的精确概率，这两个p值均是双尾概率，单边检验时，取一半即可。

## 六、样本间差异比较大时，如何处理数据，才能使作的统计图更好看

可以做柱状图，很鲜明的表现出来，不同的数据表格加不同的颜色，总结或者数据分析写的详细，改进方法多写等

## 七、数据量较大的excel文件，如何让它操作时快起来？

减少公式的使用。

有些数据算出来后就 选择性粘贴 值，这样excel就不用老是重算了。

再有关闭自动重算，需要刷新数据时按一下F9

## 八、两组样本量相差大，可以用秩和检验比较吗

可以使用SPSS软件（或其他统计软件）的Crosstabs进行分析，采用其列出的fisher精确检验的结果就可以了，结果是肯定可靠的。

但是，由于另一组只有9例，这有可能会降低卡方检验的检验效能。

## 参考文档

[下载：如果两组数据样本量比较大怎么办.pdf](#)

[《股票z值是什么》](#)

[《炒股中的佩服是什么意思》](#)

[《科士达是什么龙头》](#)

[《日线周线月线是什么颜色》](#)

[下载：如果两组数据样本量比较大怎么办.doc](#)

[更多关于《如果两组数据样本量比较大怎么办》的文档...](#)

声明：

本文来自网络，不代表

【股识吧】立场，转载请注明出处：

<https://www.gupiaozhishiba.com/store/37445584.html>