

分位数 (Quantiles)

NAG Fortran 算法库第 22 版、NAG MATLAB 工具箱与 NAG C 算法库第 9 版中，新加入了快速自未排序的数值中取得分位数的算法。简单来说，分位数可被视为资料中的一个指标，用来指示在资料中低于与高于特定比例位置的值。许多分位数有许多相似的名称，最广为人知的就是中位数，其指的就是在已排序的数据集合中，中间的那个数值。如果资料量是偶数的，当然，中间的值就不存在；然而，依照定义我们会取中间两个数值的平均值。以 NAG 函数的用语来说，中位数就是 0.5 的分位数。

分位数以百分位值来表示，例如：0.01, 0.02, 0.03 等。前述第一个百分位值乃是表示在已经升序好的资料中第百分之一的资料。大型的金融机构往往需要一个快速的方法来求得资料中的这些分位数。最直觉计算百分位数的方式就是将资料先升序，然后挑出所要百分比位置的数值。但是，对大量的数据集合来说，这样的方法不是很有效率的，因为这需要许多时间先进行资料排序。

为了回应这样的需求，NAG 与牛津大学的 Mike Giles 教授共同合作开发部份快速 (partial Quicksort) 算法，这是基于 NAG 算法库中已经有的快速排序函数。这个新的算法有更高的运行效率，因为它并不需要将全部的资料都排序过就能取得分位数。

下图显示新的算法所提高的性能比较。我们利用此新的函数 G01AMF 与 NAG 排序函数 M01CAF 来取得不同数据集中的中位数。

