

医学实验中经常会有出现离群值的情况，究竟是由于随机**涨落**引起的，还是由于某些确定因素造成的，有时难以判断，如果处理不好将会引起较大的**系统误差**。对离群值的处理有一些统计判断的方法，如**chanwennt**准则规定，如果一个数值偏离观测**平均值**的**概率小于等于1/(2n)**，则该数据应当舍弃（其中n为观察例数，概率可以很据数据的分布进行估计）。

特征

编辑

发现离群值也可以通过观察值的**频数表**或直方图来初步判断，也可通过统计软件作观察值的**箱式图**来判断，如果**观测值**距箱式图底线**Q1**（第**25百分位数**）或顶线**Q3**（第**75百分位数**）过远，如超出箱体高度（**四分位数间距**）的两倍以上，则可视该观测值为离群值。当数据近似**正态分布**时，有一种较为简单的方法，可用**均数**加减**2.5s**来判断，如观测值在此范围以外，可视为离群值。在统计学上也可用**线性回归**的方法来对离群值进行判断。当出现离群值的时候，要慎重处理，要将专业知识和统计学方法结合起来，首先应认真检查原始数据，看能否从专业上加以合理的解释，如数据存在**逻辑错误**而原始记录又确实如此，又无法在找到该观察对象进行核实，则只能将该**观测值**删除。如果数据间无明显的逻辑错误，则可将离群值删除前后各做一次统计分析，若前后结果不矛盾，则该例观测值可予以保留。

离群值处理方法包括：

- 一、剔除离群值，不追加观测值；
- 二、剔除离群值，追加观测值；或剔除离群值，适宜地**插补**替代；
- 三、找到实际原因修正离群值，否则予以保留的；

离群值的取舍
<div><div>1. 定义</div><div>在一组平行测定数据中，有时会出现个别值与其他值相差较远，这种值叫离群值。</div><div>判断一个测定值是否是离群值，不是把数据摆在一块看一看，那个离得远，那个是离群值，而是要经过计算、比较才能确定，我们用的方法就叫Q检验法。</div></div> <div><div>2. 检验方法</div><div><div>（1）求Q：Q=</div><div>即：求出离群值与其最邻近的一个数值的差，再将它与极差相比就得Q值。</div></div><div><div>（2）比较：根据测定次数n和置信度查Q，若Q>Q，则离群值应舍去，反之则保留离群值。</div></div></div>