从算术平均数谈最小一乘法与最小二乘法的区别

王文平

（武汉船舶职业技术学院，公共课部 430050）

摘 要 本文从算术平均值出发，讨论了最小一乘法与最小二乘法的区别，并用这二种方法对数据进行了回归分析。

关键词 算术平均数；中位数；最小一乘法；最小二乘法；线性回归

中图分类号 G633.66 **文献标志码** A **文章编号** 1671－8100(2009)06－0040－02

1 算术平均数与中位数

"这个球队队员的平均年龄……"，"某家庭每月的平均收入……"，无论是看电视、读报纸或者和周围的人进行交谈，差不多每天都要遇到"平均"这个词，正是因为对"平均"这个词太熟悉，所以实际上没有去深入思考它的真正内涵。

设有 n 个数 $a_1, a_2, \cdots a_n$ 要找一个数 x 反映这组数的总的情况，使得 x 和这 n 个数的偏差

$$x-a_1, x-a_2, \cdots, x-a_n \tag{1}$$

在总体上说来尽可能地小。

采用的常见方法，就是适当地选取值，使得这些偏差的平方和 $D=\sum_{i=1}^{n}(x-a_i)^2$ 达到它的最小值。令 $\frac{dD}{dx}=2\sum_{i=1}^{n}(x-a_i)=0$，得 $\bar{x}=\frac{a_1+a_2+\cdots+a}{n}a_n$，此 $\bar{x}$ 即为 n 个数 $a_1, a_2, \cdots a_n$ 的算术平均数。

事情到此似乎结束了，其实不然，有必要做进一步的分析。

为什么不把(1)式的n个偏差本身相加，而把它们的平方和相加，这是因为这些偏差有些为正值，有些为负值，直接相加，就会正负相消，不能反映总体的情况。

有人又问能否把(1)式的 n 个偏差的绝对值相加，即找一个数使得 $D=\sum_{i=1}^{n}|x-a_i|$ 最小？所求的这个数就是中位数 med：①当 n 为偶数时，中位数 med 是 $[a_{(\frac{n}{2})}, a_{(\frac{n}{2}+1)}]$ 中的任何一个数；②当 n 为奇数时，中位数 med 是 $a_{(\frac{n+1}{2})}$ 这个数。这里 $a_{(1)}\leqslant a_{(2)}\leqslant\cdots\leqslant a_{(n)}$ 是 $a_1, a_2, \cdots, a_n$ 按由小到大的排列（统计学上称里 $a_{(1)}, a_{(2)}\cdots, a_{(n)}$ 是 $a_1, a_2\cdots, a_n$ 的顺序统计量）。

由上面的分析可知，算术平均数 $\bar{x}$ 是到 n 个数 $a_1, a_2\cdots, a_n$ 的距离的平方和最小，而中位数 med 是到 n 个数 $a_1, a_2\cdots, a_n$ 的距离的和最小，这才是算术平均数和中位数的实质（注意：算术平均数 $\bar{x}$ 满足 $\sum_{i=1}^{n}(\bar{x}-a_i)=0$ 的性质）。

2 最小二乘法与最小一乘法

谈到算术平均数和中位数，就要从计算方法角度做进一步的分析，以期对其实质有更深一步的理解.

①对于一维情况

求 x 使 $D=\sum_{i=1}^{n}(\bar{x}-a_i)^2$ 达到最小的方法称为最小二乘法，而使 $D=\sum_{i=1}^{n}|x-a_i|$ 达到最小的方法称为最小一乘法.

由最小二乘法得到的就是算术平均数，由最小一乘法得到的就是中位数.

②对于二维情况

在中学课本中，已知两点 (x_1, y_1)，(x_2, y_2) 可确定一条直线 $y=a+bx$，这只需将两点坐标代入直线方程，解出 a, b 即可。将两点推广到n个点，如何确定线性回归直线呢？

最常用的是利用最小二乘法准则，即求出 a，

收稿日期：2009－08－16

作者简介：王文平，男，副教授，主要从事高等教学与研究。

b,使得$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$各点沿 y 轴到直线 $y=a+bx$ 的偏差平方和最小,即

$$\sum_{i=1}^{n}[y_i-(a+bx_i)]^2=min \tag{2}$$

对(2)式中的a,b分别求偏导并令其为0,若其系数行列式不为0,则可求得唯一的a,b的显式表达,它们即为所求的回归系数。

那么能否给出另一种准则,求出a,b,使得$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$各点沿y轴到直线y=a+bx的偏差的绝对值和最小,即

$$\sum_{i=1}^{n}|y_i-(a+bx_i)|=min \tag{3}$$

答案是可行的,这即为最小一乘法准则。不过对(3)式的求解比较困难,也没有一个显式表达,具体计算见文[3]。

也许是计算上的困难,一般只学习最小二乘法,很少涉足最小一乘法。但从历史上看最小一乘法比最小二乘法的起源要早40多年。

1755年起,波斯科维奇投身于子午线长的问题研究,在1760年提出来最小一乘法准则。但由于其计算上的困难及其它原因,对这个准则的研究长期处于停止状态,直到本世纪50年代以后,局面才改变,越来越受到统计学家的重视。

1805年,法国数学家勒让德在研究天文学和测地学处理数据时最先发明最小二乘法,后来高斯等数学家对最小二乘法进行了大量的理论研究和应用,在统计学中发挥着重要的作用,是19世纪统计学的"中心主题"。最小二乘法与最小一乘法均是进行回归分析的方法,它们各有千秋。

①异常值对回归线的影响:在回归分析的最小二乘法估计中,远离中心的突出点,对参数的计算结果有较大的影响。

这样,就好理解下面的问题:在歌手电视大奖赛上,10个评委亮分之后,为什么要去掉最高分和最低分?而在最小一乘法估计中,异常值的影响一般较小。

②解的多重性:回归分析中用最小一乘法,有时会出现多个解,并且缺乏显式表达,而对于最小二乘法,解一般是唯一的,且有显式表达。

不过,有人做过这样的实验:拿大量的 $x-y$ 散点图,让一些人各自用目测的方法配直线,结果表明,大多数人目测的结果更接近于最小一乘法而不是最小二乘法。

3　启示

学生学了很多有关"平均"的概念(算术平均数、几何平均数、调和平均数和中位数等),也进行了很多关于它们之间关系不等式的推导证明,但这些概念真正的内涵似乎不太清楚,对一些概念的理解只是形式的、肤浅的"平均"概念,其实有着悠久的历史,蕴涵着很深的道理。教师不仅要教给学生知识,而且要教给他们知识的来龙去脉,这样学生不仅知其然,而且知其所以然。

参　考　文　献

1　史济怀. 平均[M]. 北京:科学出版社,2002.
2　陈希孺. 数理统计学简史[M]. 湖南:湖南教育出版社,2002.
3　李仲来. 最小一乘法介绍[J]. 数学通报,1992(2):42~45.

Discussing from the Arithmetic Average of the Difference between the Least One Multiplication and the Least Square Method

WANG Wen-ping

(Wuhan Institute of Shipbuilding Technology, Wuhan 430050, China)

Abstract: From the arithmetic mean, this paper discusses the differences between the least one multiplication and the least square method, and uses the two methods to make a regression analysis of data.

Key words: arithmetic mean ; median ; least one multiplication ; least square method ; linear regression

(责任编辑:谭银元)