技术交流

数值积分单权函数法估算年最大日雨量[①]

郭化文　　魏牲生

（山东省泰安教育学院， 271000）　（中国气象科学研究院，北京 100081）

陈建昌

（山东省泰安市气象局，271000）

提　要

简要介绍用数值积分单权函数法推求皮尔逊Ⅲ型分布参数，计算我国部分城市的年最大日雨量不同重现期值。数值积分单权函数法基本上能通过检验，取得精度合格的结果。

关键词： 数值积分单权函数法　年最大日雨量　重现期值

引言

在农业生产和水利工程方面，例如农田灌溉与排水、水库、下水道的设计与施工都需要掌握百年一遇或二百年一遇的日最大降雨量，以便做到既节省投资又安全可靠。目前国内曾用皮尔逊Ⅲ型分布，用矩法估计参数的值 $\bar{x}$、变差系数 C_v 和偏态系数 C_s，并用它计算北京、上海等地百年一遇、二百年一遇的年最大日雨量，但计算值偏小。

为了寻求一种较好的参数估计方法，多年来许多学者进行了大量的研究和探索，马秀峰[1]从分析矩法的求矩差出发，提出一种新的估计方法——权函数法，提高了三阶中心矩或 C_s 的估计精度。本文对权函数作了改进，应用数值积分单权函数法将开型数值积分法与权函数法相结合计算 C_s，使估计精度进一步提高[2,3]。

1　权函数法

估计概率密度参数的传统方法是矩法。设 n 代表观测资料系列个数，用矩法计算上述参数的公式为：

$$\bar{x}=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x\approx\frac{1}{n}\sum_{i=1}^{n}x_i \tag{1}$$

$$\sigma^2=\int_{-\infty}^{\infty}(x-\bar{x})^2f(x)\mathrm{d}x\approx\frac{1}{(n-1)}\sum_{i=1}^{n}(x_i-\bar{x})^2 \tag{2}$$

$$C_v=\frac{\sigma}{\bar{x}} \tag{3}$$

$$C_s=\frac{1}{\sigma^3}\int_{-\infty}^{\infty}(x-\bar{x})^3f(x)\mathrm{d}x\approx\frac{1}{(n-3)\sigma^3}\sum_{i=1}^{n}(x_i-\bar{x} \tag{4}$$

马秀峰从分析矩法的求矩差出发，提出一种新的估计方法——权函数法。据他分析，

① 本文系泰安市科委1995-1998课题资助项目。

求矩差的主要部分是端矩差。即由于样本极差小于变量全距,使得用样本矩代替总体矩时,不能计入样本两端以外部分的矩而产生的误差。因此如能增加均值附近数据在求矩中的作用,降低远离均值处数据的影响,必须提高估计精度。

经推导可得：

$$C_s=\frac{2[\sigma^2\int_{x_0}^{\infty}\varphi'(x)f(x)\mathrm{d}x-\int_{x_0}^{\infty}(x-\bar{x})\varphi(x)f(x)\mathrm{d}x]}{\sigma\int_{x_0}^{\infty}(\bar{x}-x)\varphi'(x)f(x)\mathrm{d}x} \tag{5}$$

经分析,权函数 $\varphi(x)$ 可取正态密度函数,即

$$\varphi(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{(x-\bar{x})^2}{2\sigma^2}} \tag{6}$$

由于 $\varphi'(x)=-\frac{(x-\bar{x})}{\sigma^2}\varphi(x)$

代入式(5)可得 $C_s=-4\sigma\frac{E(x)}{H(x)}$ (7)

式中

$$\left.\begin{aligned}E(x)&=\int_{x_0}^{\infty}f(x)(x-\bar{x})\varphi(x)\mathrm{d}x\\H(x)&=\int_{x_0}^{\infty}f(x)(x-\bar{x})^2\varphi(x)\mathrm{d}x\end{aligned}\right\} \tag{8}$$

权函数法仍然属于单参估计,不能全面地解决皮尔逊Ⅲ型分布参数估计问题。即是说它未能解决 $\bar{x}$,特别是 C_v 的估计精度问题。事实上,影响估值 $\hat{x}_p$ 的参数首先是 $\bar{x}$ 和 C_v,其次才是 C_s。

2 数值积分单权函数法

2.1 密度函数的建立

该方法的特点是将开型数值积分与权函数相结合,改进累加求矩及修正端矩误差,提高求矩的计算精度。

任何近似积分公式都是加权平均公式,即对各指定纵坐标(被积函数)乘上适当的积分权数求和,乘以全距,再用总权数除,得到数值积分结果。最常用的数值积分公式是等距求积公式,正好适用于经验分布点据。若统一取各权数为1,即得累加求矩,即是不加权平均。另一例是人们熟悉的辛卜森公式,适用于 n 为偶数。这里的 n 代表公式所涉及的积分全距中的组距数。当 $n=2$,公式为：

$$\int_{x_0}^{x_2}f(x)\mathrm{d}x=I_0^2\approx\frac{2h}{6}(y_0+4y_1+y_2)-\frac{h^5}{90}f^{(4)}(\xi)$$

$$(x_0\leqslant\xi\leqslant x_2) \tag{9}$$

式中,纵坐标 $y=f(x)$;1,4,1 即是各积分权数;6 为总权数;h 为组距;$2h$ 为全距。右方末项代表余项,表明精度高达4阶,比起累加求积的精度不过零到一阶之间的情况,显然优越得多。

对于更大的偶数纵坐标,可以重复使用上式求和。

欲修正端矩误差,可以采用开型求积公式[4],通常指在积分一端(或两端)向外延伸一个组距的近似积分公式。例如,当 $n=4$,公式为：

$$\int_{x_0}^{x_4}f(x)\mathrm{d}x=I_0^4\approx\frac{4h}{3}(2y_1-y_2+2y_3)+\frac{14}{15}h^5f^{(4)}(\xi)$$

$$(x_0\leqslant\xi\leqslant x_4) \tag{10}$$

这时,全距$=4h$,但积分值是由 y_1,y_2,y_3 求出,不含端纵坐标 y_0,y_4。公式精度达到4阶,但远逊于辛卜森公式。

如不带端积延伸,即为闭型求积公式。累加求积、梯形公式和辛卜森公式均属此类。

将适当的开型及闭型求积公式组合起来求矩,就可达到前述提高求积精度连带修正端矩误差的要求。

但是等距求积公式种类很多,为了便于应用,同时保证精度,采用刘光文建议[5]下列积分权数组合：

① 当 n 为奇数,积分权数组合如下：

8,−4,8,1,4,2,4,2,…,2,4,2,4,1,8,−4,8

总权数=3（$n+1$）。$n=13,15,17,\cdots$

② 当n为偶数，采用64，−32，64，8，32，16，…，32，17，27，27，17，32，…，16，32，8，64，−32，64

总权数=24（$n+1$）。$n=14,16,18,\cdots$

取北京、上海等14个城市1960～1995年的年最大日雨量资料，$n=36$，权数恰好呈对称排列，总权数等于888。

照此，当施用求矩时，$(n+1)h=1$（总频率）积分值就可以简化为：

$$\int_{x_0}^{x_n} f(x)\mathrm{d}x \approx \frac{\sum_{i=1}^{n}\omega_i f(x_i)}{\sum_{i=1}^{n}\omega_i} \qquad (11)$$

式中，ω_i为各积分权数；$\sum\omega_i$为总权数，注意这时公式已经简化，不再显含组距h在内。

将上述数值积分步骤与权函数法结合起来，可望有效地提高频率计算的精度达到容许范围以内。具体地说，就是通过数值积分求$\bar{x}$和C_v，然后利用数值积分计算权函数矩$E(x)$及$H(x)$来求C_s。

有了C_v、C_s和$\bar{x}$值，密度函数即已确定，则各个频率p的估计值$\hat{x}_p$值由下式求得：

$$\hat{x}_p = \bar{x}(\varphi \cdot C_v + 1) \qquad (12)$$

其中φ值可以从皮尔逊Ⅲ型曲线的离均系数值表查得[6,7]。

2.2 单权函数法求算方法

下面具体介绍用数值积分单权函数法计算上海台36年一遇的年最大日雨量。

① 计算数积参数$\bar{x}$、σ、C_v。

$\bar{x}=85.9538$，$\sigma=37.3674$，$C_v=0.4347$

② 计算权函数

$$\varphi(x_i) = \frac{1}{37.3674\sqrt{2\pi}} e^{-\frac{(x_i-\bar{x})^2}{2\times 85.9538^2}}$$

③ 计算单权参数C_s

$$E(x) = \sum_{i=1}^{36}\omega_i(x_i-\bar{x})\varphi(x_i) = -53.2180$$

$$H(x) = \sum_{i=1}^{36}\omega_i(x_i-\bar{x})^2\varphi(x_i) = 2929.8266$$

根据式(7)，$C_s=2.7144$

④ 计算重现期36年的最大日雨量

$T=36$年　频率$p=\frac{1}{T}=2.78\%$　查表内插$\varphi_p=2.7163$

根据式(12)有

$\hat{x}_p = 85.9538\times(2.7163\times 0.4347) = 187.5\text{mm}$

此值和1960～1995年实测最大日雨量$X=204.4$mm大体相拟合。

当$\left|\frac{\text{理论推算值}-\text{实测值}}{\text{实测值}}\right|$不超过8%～10%时[5]，可以认为估值$\hat{x}_p$的误差在容许范围内，理论推算值和实测值拟合很好，用数值积分单权函数法求算T年一遇的年最大日雨量是接近实际的，是可取的。

而用矩法算得上海台的参数$\bar{x}=82.17$，$C_v=0.3745$，$C_s=1.97$；用数值积分单权函数法算得$\bar{x}=85.95$，$C_v=0.4348$，$C_s=2.71$，用两套参数进行适线对比（图略），可以看出：尽管该台实测系列的点据分布很不规则，用数值积分单权函数法算得的频率曲线，仍能经验点据匀称地分布在曲线的两侧，向左上方外延的趋势也比较规顺合理；但是，用矩法求得的频率曲线向左上方外延不能与实测点相吻合，没有合理的外延趋势，这表明数值积分单权函数法对实测系列比较矩法为优。

3 计算结果

我们选用北京、上海等14个台站的年最大日雨量资料利用数值积分单权公式(3)、(7)、(11)估计了参数C_v、C_s，然后用公式(12)计算了36年一遇的年最大日雨量，并与实测值$R_{max}(36)$对照（见表1）。降水资料年代为1960～1995年。

从表1看出用数值积分单权函数法推算年最大日雨量与实测值拟合较好，14个台的

误差都不超过10%；而用矩法估计参数推算年最大日雨量与实测值拟合较差，14个台中有8个台拟合差。

其次我们又计算了各台站50年一遇、100年一遇、200年一遇的年最大日雨量，分别记为R_{max}(S，50)、R_{max}(S，100)、R_{max}(S、200)，其结果见表2。

表1 理论推算值R_{max}(S,36)与实测值R_{max}(36)对照(单位:mm)

序号	台站	R_{max}(S,36)		R_{max}(36)	评价	
		矩法	单权函数法		矩法	单权函数法
1	北京	170.7	196.2	212.2	偏小	√
2	济南	229.3	281.5	298.4	偏小	√
3	南京	180.6	192.5	179.3	√	√
4	上海	162.2	187.5	204.4	偏小	√
5	汉口	258.8	281.4	298.5	偏小	√
6	福州	183.4	189.5	170.9	√	√
7	广州	242.2	253.9	253.6	√	√
8	成都	185.8	195.8	201.3	√	√
9	贵阳	126.2	130.3	130.2	√	√
10	西安	98.5	108.7	110.1	偏小	√
11	兰州	73.3	88.2	96.8	偏小	√
12	银川	65.5	70.8	66.8	√	√
13	沈阳	166.2	198.2	215.5	偏小	√
14	哈尔滨	110.2	126.5	131.2	偏小	√

注:标√表示推算与实测值拟合较好。

表2 不同重现期最大日雨量理论推算值R_{max}(S,T)(单位:mm)

序号	台站	R_{max}(S,50)		R_{max}(S,100)		R_{max}(S,200)	
		矩法	单权函数法	矩法	单权函数法	矩法	单权函数法
1	北京	181.2	211.0	204.4	243.8	227.4	276.9
2	济南	246.4	308.9	283.9	369.6	321.5	432.0
3	南京	188.1	202.2	204.1	222.9	219.3	243.2
4	上海	171.8	201.5	193.0	232.9	214.2	264.7
5	汉口	275.7	304.7	312.6	356.5	349.4	409.2
6	福州	189.9	196.9	203.3	212.4	215.8	227.2
7	广州	253.5	267.7	277.0	297.6	301.2	327.2
8	成都	195.1	207.2	215.0	232.1	234.6	257.0
9	贵阳	130.6	135.2	139.7	145.6	148.5	155.9
10	西安	107.8	115.7	115.2	131.0	126.4	146.2
11	兰州	78.4	95.6	89.4	112.0	100.4	128.7
12	银川	68.8	75.1	75.6	84.4	82.3	93.6
13	沈阳	177.2	214.3	201.3	249.8	225.3	285.9
14	哈尔滨	116.0	134.5	128.2	152.3	140.4	170.4

4 结语

4.1 皮尔逊Ⅲ型分布是一种理论分布，采用传统矩法估计参数，用它推算不同重现期的年最大日雨量，所计算重现期值偏小，因此需要研制较为精确的皮尔逊Ⅲ型分布参数估计方法，以求得最好结果。

4.2 单纯权函数法在理论上是正确的，它能降低矩阶，适当加权，从而有效地提高了皮尔逊Ⅲ型分布矩法定参精度。可惜此法主要只考虑到参数 C_s 的改善，未曾涉及 $\bar{x}$ 及 C_v，因而其估计精度仍然不足，难以达到合格地步。

4.3 数值积分单权函数法在现行的定参方法中，从理论上以及对实测降水资料的拟合都能通过检验，取得精度合格的结果。对于通常资料条件，数值积分单权函数法业已足用。若发现 C_s 估值特大(>2.0)应考虑采用数值积分双权函数法[5]。

参考文献

1 马秀峰. 计算水文频率参数的权函数法. 水文，1984年第4期：1～8.

2 刘治中. 数值积分权函数推求P-Ⅲ型分布参数. 水文，1987年第5期：11～15.

3 马开玉等. 气候统计原理与方法. 北京：气象出版社，1993：69～76.

4 Buckingham. . A. Numerical methods. 1957, London.

5 刘光文. 皮尔逊Ⅲ型分布参数估计. 水文，1990年第4期：1～15.

6 西安冶金学院编. 水文学. 北京：中国建筑工业出版社，1979：152.

7 华东水利学院主编. 水文学的概率统计基础. 北京：水利出版社，1981：412～415.

Estimating Annual Maximum Precipitation with Numerical Integral Single Weighted Function Method

Guo Huawen　　Wei Shensheng

(Taian Institute of Education, Shandong Province 2710000)　(Chinese Academy of Meteorological Sciences, Beijing 100081)

Chen Jianchang

(Taian Meteorological Office, 271000)

Abstract

The method of numerical integral single weighted function can be used to estimate the Pearson Type Ⅲ distribution and the estimation of values of different annual maximum diurnal precipitation return preiod in some of the cities in China was calculated. It suggested that the numerical integarl single weighted function method can basically pass the inspection and the result qualifies the precision.

Key Words: numerical integral single weighted function method　reappearance period values　annual maximun diurnal precipition