

权函数水文频率分析方法的一种应用

梁忠民¹, 宁方贵², 王钦钊³

(1. 河海大学水文水资源及环境学院, 江苏 南京 210098; 2. 松辽水利委员会水文局, 吉林 长春 130021;
3. 江西省水文局, 江西 南昌 330002)

摘要:以(单)权函数法的理论结果为基础, 不需引入第二权函数, 通过联解两个超越方程, 在权函数法的框架下同时解决皮尔逊Ⅲ型(P-Ⅲ)分布的变差系数(C_v)和偏态系数(C_s)的估计问题. 计算结果表明, 修改后的方法统计性能较优, 而且计算较简单.

关键词:(单、双)权函数法; 矩法; 概率权重矩法; 线性矩法; 变差系数; 偏态系数

中图分类号: P333.9 **文献标识码:** A **文章编号:** 1000-1980(2001)04-0095-04

给定总体分布下参数的估计问题是水文频率分析计算的主要内容. 矩法是一个被广泛采用的计算方法, 但研究及应用表明^[1], 矩法有较大的估计偏差, 而且估计偏差与矩的阶数成正比. 由于水文频率分析中一般均需处理三阶样本矩计算, 所以传统矩法往往不适用. 围绕对矩法的改进, 国内外有过很多研究, 其中近期的代表性成果主要有 3 项. 一是 Greenwood^[2]1979 年提出的概率权重矩法(PWM), 二是 Hosking^[3]1990 年提出的线性矩法(L-M). 实用中的 PWM 法及 L-M 法均是样本次序统计量的某种线性组合, 因此与传统矩法相比, 在估计中该两种方法不包含样本值的高次方运算, 仅是一阶样本矩计算. 所以估计结果受样本中个别点据误差的影响小, 估计的偏差小且更稳健. 第三项代表性的研究是马秀峰^[4]1984 年提出的权函数法. 在样本矩的计算中通过引入一个正态权函数, 使 C_s 的估计由矩法中的三阶样本矩降低为二阶矩运算或一阶矩运算(取决于估计中采用的具体公式), 提高了 C_s 的估计精度. 刘光文^[5]在此基础上提出了双权函数法, 通过引入第二权函数达到同时估计 C_s 和 C_v 的目的.

文献[4]的(单)权函数法中隐含着只用一个权函数即可同时估计 C_v 和 C_s 的思路, 整体看, 估计属二阶矩运算, 与双权函数法一致. 本文对此进行初步分析, 并考查了估计的稳健性.

1 权函数法理论及应用

1.1 权函数法基本公式

对 P-Ⅲ型总体, 令权函数 $\phi(x)$ 为正态函数 $N(\mu, \sigma^2)$, 则 C_s 可按式(1)估计^[4]:

$$C_s = -4\sigma \frac{E}{G} \quad (1)$$

式中 σ 为总体均方差, E 和 G 由下式确定:

$$E = E[(x - \mu)\phi(x)] = \frac{1}{2\sqrt{2\pi}\Gamma(\gamma)} \left(\sqrt{\frac{2\gamma}{e}} \Gamma\left(\frac{\gamma}{2}\right) \left[\sqrt{2}\Gamma\left(\frac{\gamma+1}{2}\right) - \sqrt{\gamma}\Gamma\left(\frac{\gamma}{2}\right) \right] \right) \quad (2)$$

$$G = E[(x - \mu)^2\phi(x)] = -2\sigma\sqrt{\gamma}E \quad (3)$$

式中 μ 为总体均值, $\gamma = 4/C_s^2$. 若 E 由样本估计, 则由式(2)即可估计 γ , 从而得到 C_s 的估计, 这是(单)权函数法估计 C_s 的基本做法, 属一阶矩运算. 若由式(1)估计 C_s , 尚需估计 σ 和 G , 属二阶矩运算.

1.2 同时估计 C_v 和 C_s

做法是: 由样本估计 E 和 G ; 由式(2)估计 γ (或 C_s); 由式(1)反求 σ , 则 $C_v = \sigma/\bar{x}$, $\bar{x}$ 为样本均值. 这样, 不需引入第二权函数即可在(单)权函数法框架下同时估计 C_v 和 C_s .

非连续样本下 E 和 G 的估计如下：

$$E = \frac{1}{N} \left[\sum_{i=1}^a (x_i - \bar{x}) \phi(x_i) + \frac{N-\alpha}{n-l} \sum_{i=a+1}^{n+a-l} (x_i - \bar{x}) \phi(x_i) \right]$$

(4)

$$G = \frac{1}{N} \left[\sum_{i=1}^a (x_i - \bar{x})^2 \phi(x_i) + \frac{N-\alpha}{n-l} \sum_{i=a+1}^{n+a-l} (x_i - \bar{x})^2 \phi(x_i) \right]$$

(5)

$$\bar{x} = \frac{1}{N} \left[\sum_{i=1}^a x_i + \frac{N-\alpha}{n-l} \sum_{i=a+1}^{n+a-l} x_i \right]$$

(6)

式中： N ——样本的调查重现期； n ——实测期长度； α ——作为特大值处理的样本个数，其中有 l 个来自实测样本。

权函数 $\phi(x)$ 中的参数 μ 由式 (6) 估计。由于在计算中 $\phi(x)$ 仅起样本加权作用，该函数中参数 σ 的估计并不作为最终的估计结果，所以可采用任何其它估计方法（如矩法）的估计结果。

2 方法统计性能检验

2.1 方法的无偏性检验

采用文献 [5] 中理想样本的计算方案进行对比研究，部分结果列于表 1，其中设计频率取 $p = 1\%$ 及 $p = 0.1\%$ 两种情况。表中数值是以估计值与总体值离差占总体值的百分比表示。

表 1 假想样本 P-Ⅲ 型分布参数估计方法的对比检验

Table 1 Comparisons of parameter estimation methods for P-Ⅲ distribution of "ideal sample"							
样本序号	样本容量	项 目	假想理论值	矩 法	概率权重矩法	本文权函数法	数值积分双权函数法
1	19	Ex	1	-2.1	-2.1	-2.1	-0.7
		Cv	1/2	-11.5	-7.0	-5.7	-3.7
		Cs	1	-39.9	-14.9	-16.3	-3.1
		$X_{1\%}$	2.51	-13.5	-8.0	-7.4	-3.2
		$X_{0.1\%}$	3.27	-17.5	-9.8	-9.2	-3.8
2	29	Ex	1	-1.5	-1.5	-1.4	-0.4
		Cv	1/2	-8.8	-5.4	-4.0	-2.4
		Cs	1	-34.5	-12.0	-10.4	-1.5
		$X_{1\%}$	2.51	-10.9	-6.2	-5.1	-2.1
		$X_{0.1\%}$	3.27	-14.4	-7.7	-6.3	-2.4
3	19	Ex	1	-3.0	-3.0	-3.0	-0.9
		Cv	1/2	-13.3	-7.9	-7.8	-4.5
		Cs	3/2	-39.7	-13.3	-14.9	-1.5
		$X_{1\%}$	2.67	-16.9	-9.8	-9.9	-3.9
		$X_{0.1\%}$	3.62	-22.1	-12.0	-12.3	-4.5
4	29	Ex	1	-2.2	-2.2	-2.2	-0.6
		Cv	1/2	-10.4	-6.2	-5.6	-3.0
		Cs	3/2	-34.6	-10.8	-9.1	-0.8
		$X_{1\%}$	2.67	-13.9	-7.6	-6.9	-2.6
		$X_{0.1\%}$	3.62	-18.4	-9.1	-8.4	-3.0
5	19	Ex	1	-4.5	-4.5	-4.5	-1.4
		Cv	3/4	-12.0	-6.5	-6.4	-3.2
		Cs	3/2	-39.7	-13.3	-14.9	-1.4
		$X_{1\%}$	3.50	-19.3	-11.2	-11.4	-3.9
		$X_{0.1\%}$	4.93	-24.3	-13.2	-13.5	-4.4
6	29	Ex	1	-3.3	-3.3	-3.3	-0.9
		Cv	3/4	-9.4	-5.1	-4.6	-1.5
		Cs	3/2	-34.6	-10.8	-9.1	-0.8
		$X_{1\%}$	3.50	-15.8	-8.7	-7.9	-2.1
		$X_{0.1\%}$	4.93	-20.3	-10.4	-9.3	-2.3
7	19	Ex	1	-7.3	-7.3	-7.3	-2.3
		Cv	3/4	-21.6	-12.0	-3.4	-3.9
		Cs	3	-40.9	-7.3	+4.5	+5.0
		$X_{1\%}$	4.04	-30.0	-17.0	-8.7	-4.2
		$X_{0.1\%}$	6.36	-37.4	-18.7	-8.2	-3.6
8	29	Ex	1	-5.5	-5.5	-5.5	-1.6
		Cv	3/4	-17.5	-9.6	+8.0	-2.9
		Cs	3	-36.7	-6.3	+16.7	+2.4
		$X_{1\%}$	4.04	-25.1	-13.5	+4.0	-3.2
		$X_{0.1\%}$	6.36	-32.1	-15.4	+8.3	-3.0

表中数值的绝对值越小,表明估计方法的无偏性越好.在计算的 20 种总体方案中,权函数法对 $X_p(p = 1\%)$ 和 $X_p(p = 0.1\%)$ 估计的相对偏差绝对值的平均值分别约为 9% 和 11%,而概率权重矩法估计的分别为 11% 和 13%(表 1 中仅列出其中的 8 种方案结果).所以,本文方法的无偏性整体上与文献 [5] 中所列的概率权重矩法的结果相当.两者皆稍逊于数值积分双权函数法,差别主要源自对总体均值的估计偏低,但优于文献 [5] 中所比较的其它方法,如矩法、福氏改正法、P-Ⅲ型因子法等.

2.2 方法的稳健性检验

水文频率分析中实际应用的样本一般均是统计上所谓的“污染样本”.可以认为实际样本的总体分布 F_t 与事先假设的总体分布 F_h 有一定的偏离,表示为

$$F_t = (1 - \epsilon)F_h + \epsilon H$$

(7)

式中 $F_h(x) \neq H(x)$; $\epsilon(0 \leq \epsilon \leq 0.5)$ 称为“污染比例”,表示有 100 $\epsilon\%$ 比例的样本来自另一总体 $H(x)$.

导致水文分析中 F_t 与 F_h 有偏离的因素很多,本文仅考虑了其中的一种可能情况,即未能准确确定样本中个别点据的量值而使计算样本受到“污染”.计算中,将样本的首 K 位样本值放大 50%,而其重现期不变,并与其它样本一起参与计算.考查样本放大后的估计结果与未放大时估计结果的差别,若差别小,表明方法受样本中个别点据误差的影响小,即方法具有相对的稳健性;反之,稳健性相对差.本次仍取样本容量 $n = 19, 29$ 两种情况,并约定 $n = 19$ 时 $k = 1$, $n = 29$ 时 $k = 2$,则“污染”比例 $\epsilon \approx 5\%$.本文的权函数法和矩法的部分计算结果列于表 2,表中数值是以样本放大后的估值与原估值之差占原估值的百分比来表示的.

表 2 假想样本参数估计方法稳健性对比检验

Table 2 Comparison of stability of parameter estimation methods for P-Ⅲ distribution of“ideal sample”

总体参数		项 目	矩 法	本文权函数法
$C_v = 0.5$ $C_s = 1.0$	$N = 19$	$X_{1\%}$	48.5	5.2
		$X_{0.1\%}$	60.4	7.9
$C_v = 0.5$ $C_s = 1.0$	$N = 29$	$X_{1\%}$	52.9	10.4
		$X_{0.1\%}$	65.1	15.2
$C_v = 0.5$ $C_s = 1.5$	$N = 19$	$X_{1\%}$	48.0	1.0
		$X_{0.1\%}$	58.5	1.2
$C_v = 0.5$ $C_s = 1.5$	$N = 29$	$X_{1\%}$	52.2	6.1
		$X_{0.1\%}$	62.6	8.3
$C_v = 0.75$ $C_s = 1.5$	$N = 19$	$X_{1\%}$	48.0	0.1
		$X_{0.1\%}$	54.6	-0.3
$C_v = 0.75$ $C_s = 1.5$	$N = 29$	$X_{1\%}$	52.2	4.8
		$X_{0.1\%}$	58.6	5.9
$C_v = 0.75$ $C_s = 2.25$	$N = 19$	$X_{1\%}$	47.5	-9.5
		$X_{0.1\%}$	52.2	-13.6
$C_v = 0.75$ $C_s = 2.25$	$N = 29$	$X_{1\%}$	51.4	-5.4
		$X_{0.1\%}$	55.4	-8.2

两种样本情况下估计结果相对偏差绝对值的大小,可作为评价方法稳健性的指标.相对偏差绝对值越小,方法稳健性越好,反之亦然.对所有 20 个方案,权函数的稳健性均优于矩法(表中仅列出其中的 8 种).以对 $X_p(p = 1\%)$ 的估计为例,相对偏差绝对值的平均,权函数法约为 10%,矩法为 50%,前者优势较明显.

3 结 语

- a. 权函数法是在样本矩的计算中引入特定函数,使参数估计由高阶矩计算降为低阶矩计算.这种在估计中的“降阶”作用使权函数法具有较优良的统计性能且方法的稳健性直观、明显.
- b. 本文不引入第二权函数,直接在(单)权函数法的框架下同时估计 P-Ⅲ型总体的 C_v 和 C_s .计算结果表明,该法整体的估计水平与概率权重矩法相当,同时也具有较好的稳健性.
- c. 文献 [5] 中构造的数值积分双权函数法具有较优的无偏性,但需引入第二个权函数且采用数值积分算法,计算过程相对复杂.本文仍只使用原文献 [4] 中的一个权函数,且估计中统计量 E 与 C_s 的关系可预先制表查用,计算过程相对较简单.

d. 由公式 (2) 表示的 $E \sim C_s$ 函数是单峰凸函数, E 值约在 $C_s = 4.0$ 时达到极大值; $C_s = 4.0$ 后, E 值又减小. 所以本文方法仅适用于 $C_s < 4.0$ 的情况.

致谢: 本文工作曾得到马秀峰教授级高级工程师的指点, 特此致谢.

参考文献:

- [1] Wallis J R, Matalas N C, Matas N C, *et al.* Just a moment! [J]. Water Resources Research, 1974, 10: 211 ~ 219.
- [2] Greenwood J A, Landwehr J M, *et al.* Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form [J]. Water Resources Research, 1979, 15: 1049 ~ 1054.
- [3] Hosking J R M. L-moments: Analysis and estimation of distribution using linear combinations of order statistics [J]. Journal of the Royal Statistical Society, Series B, 1990, 52: 105 ~ 124.
- [4] 马秀峰. 计算水文频率参数的权函数法 [J]. 水文, 1984, 21: 1 ~ 8.
- [5] 刘光文. 皮尔逊 III 型分布参数估计 [J]. 水文, 1990, 25: 1 ~ 15.

Application of Weighted Function Method to Hydrological Frequency Analysis

LIANG Zhong-min¹, NING Fang-gui², WANG Qin-zhao³

(1. College of Water Resources and Environment, Hohai Univ., Nanjing 210098, China;

2. Hydrology Bureau, Shongliao Water Resources Committee, Changchun 130021, China;

3. Hydrology Bureau of Jiangxi Province, Nanchang 330002, China)

Abstract: A procedure for the estimation of coefficient of variation (C_v) and coefficient of skewness (C_s) of a Pearson type of three probability distribution is presented based on the weighted function method (WF). The estimated coefficients are the joint solutions to two transcendental equations. In comparison with the method of moments (MOM) and the method of probability weighted moments (PWM), the proposed method is of good statistical properties, and simple in calculation.

Key words: (single, double) weighted function method; method of moments; probability weighted moments; L-moments; coefficient of variation; coefficient of skewness