

DOI :10.3880/j.issn.1006-7647.2008.05.008

P-Ⅲ型分布参数的线性矩估计法及期望概率研究

白 丽,夏乐天,魏玉华

(河海大学理学院,江苏 南京 210098)

摘要 在总体为 P-Ⅲ型分布的假定下,利用矩法、概率权重矩法、绝对值准则适线法及线性矩法 4 种方法估计未知参数,计算对比了 P-Ⅲ型分布的期望概率。统计试验结果表明:适线法与线性矩法计算得出的期望概率结果较为接近,总体上都略优于概率权重矩法,而由矩法得到的期望概率远远高于设计频率,因而从期望概率这一角度论证线性矩法相对来说是合理有效的。

关键词 参数估计 线性矩法 统计试验 期望概率

中图分类号 :TB114 **文献标识码** :A **文章编号** :1006-7647(2008)05-0029-03

L-moment method for the Pearson Type-Ⅲ distribution and its expected probability in parameter estimation//BAI Li, XIA Le-tian, WEI Yu-hua (College of Sciences, Hohai University, Nanjing 210098, China)

Abstract :Based on the Monte-Carlo method, the main methods of parameter estimation for the Pearson Type-Ⅲ (P-Ⅲ) distribution are analyzed, including the traditional moment (MOM), curve-fitting (FIT), the probability weighted moment (PWM) and the L-moment (LM) methods. The expected probabilities of the P-Ⅲ distribution calculated by each method are compared. Comparison results show that the probabilities predicted by the FIT and LM methods are close to the design frequency, while the expected probability predicted by the MOM method is much higher than the design frequency. With regard to the expected probability, the LM method is more reasonable and effective.

Key words :parameter estimation ; L-moment method ; statistical test ; expected probability

在水利工程规划设计及水资源量的分析计算中,都需要进行水文频率计算工作,以确定达到工程要求的设计值 x_p ,其中参数估计是水文频率计算中的重要内容。在总体为 P-Ⅲ型分布的假定下,目前参数估计方法有矩法(MOM),极大似然法(MXL),适线法(FIT),概率权重矩法(PWM),权函数法(WF1, WF2)等。Hosking 于 1990 年提出了线性矩法(LM),近 10 余年来国内外有关学者就这一方法的适用性进行了一系列的研究工作^[1-3],他们所做的主要是在假定总体分布已知的条件下,就简单样本和非简单样本的情形,通过线性矩法得出的参数估计值来讨论总体的统计特性,并将线性矩法与传统的参数估计方法进行比较,此外还讨论了线性矩法运用于地区综合、线型分析鉴别等方面的内容。

设计值 x_p 是根据样本估计得到的,若实际来水量 X 的分布函数为 $F_0(x)$,文献[4]提出,一旦样本给定,则无论 $F_0(x)$ 具有何种形式,也不管采用何种方法估算设计值 x_p ,均有

$$P\{X \geq x_p\} = E\{F_0(x_p)\} \tag{1}$$

所以利用容量为 n 的样本,估算得到频率为 p 的设计值的期望概率为 $E\{F_0(x_p)\}$ ^[5],文献[4]又称之为期望破坏概率。

文献[1]通过统计试验方法,以设计值的不偏性和有效性为标准说明了线性矩法是合理有效的;文献[6]探讨了在适线法估计参数的条件下 P-Ⅲ型分布期望概率与设计频率的关系;文献[4]探讨了矩法估计参数时 P-Ⅲ型分布期望概率与设计频率的关系。本文分别利用矩法、适线法、概率权重矩法和线性矩法计算 P-Ⅲ型分布的期望概率,并对这 4 种参数估计方法得到的结果进行对比分析,此外还就设计频率与期望概率的关系论证了运用线性矩法估计总体分布的未知参数的可行性。

1 线性矩法的基本原理

线性矩起源于 Greenwood 等^[7]1979 年定义的“概率权重矩”,是概率权重矩的线性组合,因而称其

为线性矩。

1.1 线性矩的基本概念

令 $\beta_0, \beta_1, \beta_2$ 为分布函数的概率权重矩, $\lambda_1, \lambda_2, \lambda_3$ 为分布函数的线性矩, 则 Hosking 等定义的线性矩与概率权重矩的前 4 阶关系为

$$\begin{cases} \lambda_1 = \beta_0 \\ \lambda_2 = 2\beta_1 - \beta_0 \\ \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \\ \lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \end{cases} \quad (2)$$

为了便于定义无因次的线性矩, Hosking 还提出了线性矩系数:

$$\tau_r = \frac{\lambda_r}{\lambda_2} \quad (r = 3, 4) \quad (3)$$

$$\tau_2 = \frac{\lambda_2}{\lambda_1} \quad (4)$$

式中: λ_1 为线性均值; λ_2 为线性均方差; τ_2 为线性变差系数; τ_3 为线性偏态系数; τ_4 为线性峰度系数。可见, 线性矩是概率权重矩的线性组合。

1.2 线性矩与统计分布参数的关系

设 P-III 型分布密度函数如下:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (x - a_0)^{\alpha-1} e^{-\beta(x-a_0)} \quad (5)$$

$(\alpha, \beta > 0; x > a_0)$

其中 α, β, a_0 为参数, 由于它们与线性矩之间的关系复杂, 故 Hosking 等^[3]给出了近似算法, 不过其误差可控制在 10^{-6} 以下, 因此具有足够高的精度。 α, β, a_0 与 $\lambda_1, \lambda_2, \tau_3$ 的关系如下:

$$\lambda_1 = a_0 + \frac{\alpha}{\beta} \quad \lambda_2 = \frac{\sqrt{\pi}\Gamma(\alpha + 0.5)}{\Gamma(\alpha)\beta} \quad (6)$$
$$\tau_3 = \begin{cases} \alpha^{-0.5} \frac{A_0 + A_1\alpha^{-1} + A_2\alpha^{-2} + A_3\alpha^{-3}}{1 + B_1\alpha^{-1} + B_2\alpha^{-2}} & \alpha \geq 1 \\ \frac{1 + E_1\alpha + E_2\alpha^2 + E_3\alpha^3}{1 + F_1\alpha + F_2\alpha^2 + F_3\alpha^3} & \alpha < 1 \end{cases} \quad (7)$$

式(7)中常系数 $A_0, A_1, A_2, A_3, B_1, B_2, E_1, E_2, E_3, F_1, F_2, F_3$ 的取值分别为 0.325 735 01, 0.618 691 50, 0.078 327 243, - 0.002 912 053 9, 0.466 971 02, 0.242 554 06, 2.380 757 6, 1.593 179 2, 0.161 837 1, 5.153 329 9, 7.142 526 0, 1.974 505 6。

1.3 样本线性矩的计算

设 $x_{(i)} (i = 1, 2, \dots, n)$ 为样本的次序统计量, 那么线性矩 $\lambda_1, \lambda_2, \lambda_3$ 对应的样本矩 l_1, l_2, l_3 的计算公式详见文献[1]。于是, 对于给定的次序统计量 $x_{(1)}, x_{(2)}, \dots, x_{(n)}$, 就可以估算出 $\lambda_1, \lambda_2, \lambda_3, \tau_3 (\tau_3 = l_3/l_2)$ 。确定了 $\lambda_1, \lambda_2, \tau_3$ 后, 则 EX, C_v, C_s 的计算

公式可在给出 α 近似值的前提下得到。

当 $|\tau_3| < \frac{1}{3}$ 时, 令 $z = 3\pi\tau_3^2$, 有

$$\alpha \approx \frac{1 + 0.2906z}{z + 0.1882z^2 + 0.0442z^3} \quad (8)$$

当 $\frac{1}{3} < |\tau_3| < 1$, 令 $z = 1 - |\tau_3|$, 有

$$\alpha \approx \frac{0.36067z - 0.59567z^2 + 0.25361z^3}{1 - 2.78861z + 2.56096z^2 - 0.77045z^3} \quad (9)$$

$$\begin{cases} EX = l_1 \\ C_v = \frac{l_2\pi^{\frac{1}{2}}\alpha^{\frac{1}{2}}\Gamma(\alpha)}{\Gamma(\alpha + \frac{1}{2})l_1} \\ C_s = 2\alpha^{-\frac{1}{2}}\text{sign}(\tau_3) \end{cases} \quad (10)$$

2 各种参数估计及设计频率与期望概率的关系

为了确定设计频率 p 和期望概率的关系, 假设总体 EX, C_v, C_s 的值为 EX_0, C_{v0}, C_{s0} , 其中 $EX_0 = 1$ 。用 Monte-Carlo 方法随机抽取 k 组容量为 n 的样本, 对每组样本分别采用矩法、适线法、概率权重矩法及线性矩法估计参数。当用线性矩法时, 由式(10)可以得到 3 个参数的估计值, 即 $\bar{x}_k, C_{vk}, C_{sk}$ 。根据给定的设计频率 p 及以上各组估计的参数, 可以计算出 k 个设计值 $x_p^{(k)}$, 然后根据已知的总体 EX_0, C_{v0}, C_{s0} 及已求出的 $x_p^{(k)}$, 求得 $p_k = P\{X \geq x_p^{(k)}\}$, 最后再计算其平均值 $\bar{p} = (p_1 + p_2 + \dots + p_k)/k$ 。若 k 很大, 则计算的平均值可作为设计频率 p 的期望概率近似值^[4]。表 1 为设计频率 $p = 1\%, 0.1\%$ 以及 $k = 1000$ 时用矩法、适线法、概率权重矩法、线性矩法计算得到的各种 P-III 型分布总体的期望概率。

从表 1 计算结果可以看出: ①用矩法计算得到的 P-III 型分布总体的期望概率远远大于设计频率 p , 这在文献[4]和文献[8]中也有着重说明, 而用适线法计算得出的期望概率与线性矩法的计算结果较为接近, 总体上都略优于概率权重矩法, 但是用适线法计算的 C_v 和 C_s 的不偏性较 PWM 和 LM 差, 即从参数及设计值的不偏性和有效性方面来看, 线性矩法比适线法效果显著^[1]。②随着样本容量 n 增加, 对于同组参数组合的期望概率减小, 更接近设计频率值, 而对于相同样本容量, 随着 C_s 的增大则期望概率值增大, 因此本文从比较期望概率值这一角度论证了线性矩法用来估计总体参数是合理有效的。

3 结 语

水文工作者通过各种参数估计方法, 从推求设

表 1 各种参数估计方法基于统计试验得出的 P-Ⅲ型分布期望概率及相对误差

总体参数			参数估计方法	$\bar{X}$	EC_v	EC_s	$p=1\%$		$p=0.1\%$	
C_{v0}	C_{s0}	n					$E\{F_0(x_p)\}/\%$	相对误差/%	$E\{F_0(x_p)\}/\%$	相对误差/%
0.25	1.0	30	MOM	1.0343	0.2378	0.8195	1.6829	-68.29	0.4856	-385.6
0.25	1.0	30	FIT	1.0343	0.2600	0.0100	1.3369	-33.69	0.3272	-272.2
0.25	1.0	30	PWM	1.0343	0.2455	0.9898	1.4300	-43.00	0.3812	-281.2
0.25	1.0	30	LM	1.0343	0.2411	0.9867	1.4513	-45.13	0.3100	-210.0
0.5	1.0	30	MOM	0.9950	0.4773	0.7957	1.7820	-78.20	0.5212	-421.2
0.5	1.0	30	FIT	0.9950	0.5200	0.9900	1.3656	-36.56	0.3628	-262.8
0.5	1.0	30	PWM	0.9950	0.4912	1.0176	1.4620	-46.20	0.3840	-284.0
0.5	1.0	30	LM	0.9950	0.4927	0.9987	1.4402	-44.02	0.3530	-253.0
0.5	1.5	30	MOM	1.0505	0.4590	1.1069	1.7253	-72.53	0.5109	-410.9
0.5	1.5	30	FIT	1.0505	0.5200	1.4900	1.4823	-48.23	0.2808	-180.8
0.5	1.5	30	PWM	1.0505	0.4908	1.4856	1.6302	-63.02	0.3400	-240.0
0.5	1.5	30	LM	1.0505	0.4951	1.5241	1.4900	-49.00	0.2642	-164.2
0.5	2.0	30	MOM	1.0017	0.4655	1.4971	2.1499	-114.99	0.6667	-566.7
0.5	2.0	30	FIT	1.0017	0.5259	2.0212	1.4856	-48.56	0.3592	-259.2
0.5	2.0	30	PWM	1.0017	0.4964	1.9804	1.5872	-58.72	0.3752	-275.2
0.5	2.0	30	LM	1.0017	0.4983	1.9969	1.5049	-50.49	0.3840	-284.0
1.0	2.5	30	MOM	0.9956	0.9232	1.8089	2.2223	-122.23	0.7310	-631.0
1.0	2.5	30	FIT	0.9956	1.0200	2.4800	1.5512	-55.12	0.3026	-202.6
1.0	2.5	30	PWM	0.9956	0.9865	2.4815	1.6160	-61.60	0.3370	-237.0
1.0	2.5	30	LM	0.9956	0.9832	2.4811	1.5400	-54.00	0.3320	-232.0
0.25	1.0	50	MOM	1.0025	0.2438	0.8640	1.4873	-48.73	0.3072	-207.2
0.25	1.0	50	FIT	1.0025	0.2600	1.0100	1.2162	-21.62	0.2296	-129.6
0.25	1.0	50	PWM	1.0025	0.2475	0.9823	1.2940	-29.40	0.2571	-157.1
0.25	1.0	50	LM	1.0025	0.2483	1.0020	1.2018	-20.18	0.2100	-110.0
0.5	1.0	50	MOM	1.0026	0.4846	0.8717	1.5673	-56.73	0.3240	-224.0
0.5	1.0	50	FIT	1.0026	0.5200	1.0100	1.3025	-30.25	0.2423	-142.3
0.5	1.0	50	PWM	1.0026	0.4935	1.0165	1.3730	-37.30	0.2183	-118.3
0.5	1.0	50	LM	1.0026	0.4950	0.9912	1.2900	-29.00	0.2530	-153.0
0.5	1.5	50	MOM	1.0452	0.4746	1.2237	1.6320	-63.20	0.4172	-317.2
0.5	1.5	50	FIT	1.0452	0.5100	1.4700	1.3924	-37.24	0.1944	-94.4
0.5	1.5	50	PWM	1.0452	0.4859	1.4656	1.4761	47.61	0.2298	-129.8
0.5	1.5	50	LM	1.0452	0.4899	1.4720	1.3333	-33.33	0.2280	-128.0
0.5	2.0	50	MOM	0.9973	0.4799	1.6379	1.7160	-71.60	0.4243	-324.3
0.5	2.0	50	FIT	0.9973	0.5200	2.0200	1.3333	-33.33	0.2349	-134.9
0.5	2.0	50	PWM	0.9973	0.4974	1.9788	1.4100	-41.00	0.2671	-167.1
0.5	2.0	50	LM	0.9973	0.4951	1.9823	1.3400	-34.00	0.2370	-137.0
1.0	2.5	50	MOM	1.0074	0.9395	1.9528	1.8442	-84.42	0.4539	-353.9
1.0	2.5	50	FIT	1.0074	1.0300	2.4900	1.3029	-30.29	0.2126	-112.6
1.0	2.5	50	PWM	1.0074	0.9942	2.4884	1.3855	-38.55	0.2371	-137.1
1.0	2.5	50	LM	1.0074	1.0026	2.4868	1.3256	-32.56	0.2380	-138.0

注 $\bar{X}$ 表示总体试验样本均值; EC_v, EC_s 分别表示 C_v 和 C_s 的期望值。

计值的不偏性和有效性的角度论证了线性矩法是一种具有良好统计性能的、合理有效的参数估计方法。笔者在总体为 P-Ⅲ型分布的假定下,利用统计试验手段,通过运用矩法、适线法、概率权重矩法及线性矩法来计算期望概率,并研究了其与设计频率的误差,得出结论:就这 4 种估计方法而言,线性矩法和适线法计算得出的期望概率精度较好,两者结果大致相当。

参考文献：

[1] 陈元芳,沙志贵,陈剑池,等. 具有历史洪水时 P-Ⅲ型分布线性矩法的研究[J]. 河海大学学报:自然科学版, 2001,29(4):76-80.
[2] PEARSON C P. New Zealand regional flood frequency analysis using L-moment[J]. Journal of Hydrology,1991,30:53-64.

[3] HOSKING J R M,WALLIS J R. Regional frequency analysis: an approach based on L-moment [M]. London: Cambridge University Press,1997:1-280.
[4] 黄振平,林小丽,侯云青,等. P-Ⅲ型分布参数的矩估计与设计洪水的计算频率[J]. 河海大学学报:自然科学版, 2005,33(1):49-51.
[5] 美国水文小组委员会. 确定洪水频率指南[J]. 徐映波,许大明,译. 水文计算,1984(2):16-18.
[6] 刘治中,王俊德,丛树铮. P-Ⅲ型分布的期望概率计算[J]. 河海大学学报:自然科学版,1989,17(4):59-64.
[7] GREENWOOD J A, LANDWEHR J M, MATALAS N C,et al. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form [J]. Water Resources Research, 1979,15(5):1049-1054.
[8] 陈元芳,侯玉. P-Ⅲ分布参数估计方法的研究[J]. 河海大学学报:自然科学版,1992,20(3):24-31.

(收稿日期:2007-10-17 编辑:高建群)