第 31 卷第 1 期 河 海 大 学 学 报 (自 然 科 学 版) Vol.31 No.1
2003 年 1 月 Journal of Hohai University(Natural Sciences) Jan.2003

可考虑历史洪水对数正态分布线性矩法的研究

陈元芳[1],沙志贵[2],顾圣华[3],许圣斌[1]

(1.河海大学水资源环境学院,江苏,南京 210098;
2.长江水利委员会水文局,湖北 武汉 430010;3.上海市水文总站,上海 200232)

摘要:简要介绍对数正态分布线性矩与分布参数的关系,提出两种可考虑历史洪水的样本线性矩估计公式.统计试验结果表明,估计公式均具有较高精度.从两种估算公式中推荐一种效果更好的供实际频率计算使用,从而解决了具有历史洪水时对数正态分布的样本线性矩的计算问题.此外,对线性矩法与矩法、适线法作了对比分析,结果表明,线性矩法比矩法、适线法有较大的优越性,其不偏性能最好且具有良好的有效性.

关键词:线性矩法;适线法;矩法;对数正态分布;历史洪水;不偏性;有效性

中图分类号:P333.9 **文献标识码**:A **文章编号**:1000-1980(2003)01-0080-04

在我国水文频率计算中,最常见的总体分布是 P-Ⅲ型及对数正态分布.关于 P-Ⅲ型分布的参数估计,我国学者曾进行过一系列研究,取得不少进展.其中值得一提的是文献[1]在 Hosking[2] 研究基础上对 P-Ⅲ型分布线性矩法进行了研究,介绍了该分布在简单样本时线性矩与分布参数的计算方法,并提出了一种有效合理的可考虑历史洪水的线性矩估计公式等,该研究填补了国内在此方面研究的空白.但对于对数正态分布,所做的研究相对较少,尤其是对对数正态分布的线性矩法的研究,在国内至今仍无人问津.考虑到此分布在实际水文频率计算中仍有较多应用,且线性矩法具有优良的统计性能,因此,本文着重就对数正态分布下,线性矩与分布参数相互关系作了介绍与分析,并就如何应用历史洪水信息进行样本线性矩的计算进行研究.

1 线性矩与统计分布参数的关系

线性矩 $\lambda_1,\lambda_2,\lambda_3,\tau,\tau_3$ 的定义、特性及线性矩与概率权重矩 b_0,b_1,b_2 的函数关系详见文献[1].

设 $\eta=\ln(\xi-c)$ 为符合正态分布 $N(a_\eta,\sigma_\eta^2)$ 的随机变量,则称 ξ 符合三参数对数正态分布.对于水文变量而言,c 应大于或等于零.ξ 的分布密度函数为

$$f(x)=\frac{1}{(x-c)\sigma_\eta\sqrt{2\pi}}e^{-\frac{(\ln(x-c)-a_\eta)^2}{2\sigma_\eta^2}}\qquad(c<x<\infty)\tag{1}$$

由于一般对数正态分布密度函数中的参数与线性矩关系复杂,故 Hosking 等[2] 给出了近似公式,其误差可控制在 10^{-6}以下,具有足够高的精度.下面是笔者结合我国常见的以上对数正态分布的表达形式,经整理得出以上两者关系式.

$$\lambda_1=c+e^{a_\eta+\sigma_\eta^2/2}\tag{2}$$

$$\lambda_2=-e^{a_\eta+\sigma_\eta^2/2}[1-2\Phi(\sigma_\eta/\sqrt{2})]\tag{3}$$

$$\tau_3=\sigma_\eta\frac{A_0+A_1\sigma_\eta^2+A_2\sigma_\eta^4+A_3\sigma_\eta^6}{1+B_1\sigma_\eta^2+B_2\sigma_\eta^4+B_3\sigma_\eta^6}\tag{4}$$

公式(4)中常系数 $A_0,A_1,A_2,A_3,B_1,B_2,B_3$ 取值分别是 0.488 602 51,$4.449\,307\,6\times10^{-3}$,$8.802\,709\,3\times10^{-4}$,$1.150\,708\,4\times10^{-6}$,$6.466\,292\,4\times10^{-2}$,$3.309\,040\,6\times10^{-3}$,$7.429\,068\,0\times10^{-5}$,其中 $\Phi(x)$ 表示标准正态分布的分布函数.反之,当已知 $\lambda_1,\lambda_2,\tau_3$ 后,则参数 σ_η,a_η,c 很容易用以下(5)(6)(7)式计算得到.当 $|\tau_3|\leqslant0.94$,

收稿日期:2001-11-19

作者简介:陈元芳(1963—),男,福建莆田人,教授,博士,主要从事工程水文与随机水文方面的研究工作.

$|\sigma_\eta| \leqslant 3$时,则

$$\sigma_\eta = \tau_3 \frac{E_0 + E_1\tau_3^2 + E_2\tau_3^4 + E_3\tau_3^6}{1 + F_1\tau_3^2 + F_2\tau_3^4 + F_3\tau_3^6} \tag{5}$$

$$a_\eta = \ln\left[\frac{-\lambda_2 e^{-\sigma_\eta^2/2}}{1 - 2\Phi(\sigma_\eta/\sqrt{2})}\right] \tag{6}$$

$$c = \lambda_1 - e^{a_\eta + \sigma_\eta^2/2} \tag{7}$$

其中常系数 $E_0,E_1,E_2,E_3,F_1,F_2,F_3$ 的取值分别是 2.046 653 4,-3.654 437 1,1.839 673 3,-0.203 602 44,-2.018 217 3,1.242 040 1,-0.217 418 01. 在 $|\tau_3| > 0.94$ 时未给出算式,但通常情况下,水文变量 $|\tau_3| \leqslant 0.94$,故仍能满足实际计算要求. 以上关系适用于正偏的三参数对数正数分布. 为了能顺利地进行线性矩及分布参数计算并保证有足够的精度,编程时需将数值变量、数组定义为双精度.

2　样本线性矩的计算

2.1　简单样本时样本线性矩的计算公式

设样本 X 为 $x_{1:n} \leqslant x_{2:n} \leqslant \cdots \leqslant x_{n:n}$,则 $\lambda_1,\lambda_2,\lambda_3$ 对应的样本矩 l_1,l_2,l_3 计算公式如下:

$$l_1 = b_0 \qquad l_2 = 2b_1 - b_0 \qquad l_3 = 6b_2 - 6b_1 + b_0 \tag{8}$$

其中

$$b_0 = \frac{1}{n}\sum_{j=1}^{n} x_{j:n} \qquad b_1 = \frac{1}{n}\sum_{j=2}^{n}\frac{(j-1)}{(n-1)}x_{j:n} \qquad b_2 = \frac{1}{n}\sum_{j=3}^{n}\frac{(j-1)(j-2)}{(n-1)(n-2)}x_{j:n} \tag{9}$$

这样只要给定样本 $x_{1:n},x_{2:n},\cdots,x_{n:n}$,则可计算出 $\lambda_1,\lambda_2,\lambda_3$ 及 $\tau = \lambda_2/\lambda_1,\tau_3 = \lambda_3/\lambda_2$.

2.2　具有历史洪水情况下样本线性矩计算公式的提出

实际工作中,常遇到具有历史洪水的水文序列,但国内外至今未提出三参数对数正态分布具有历史洪水信息的线性矩公式,这就大大限制了该公式的在实际中的应用. 为此,在文献[1]研究基础上,本文尝试采用两种可考虑历史洪水的样本线性矩公式. 其一是利用本文作者在文献[1]中建议的具有历史洪水时的线性矩估计公式;其二是利用文献[3]所规定的具有历史洪水时概率权重矩估计公式,再经过转换成线性矩.

设水文样本 X 最大重现期为 N,历史洪水个数为 a,实测期历史洪水个数为 l,实测期样本长度为 n,则系列中水文数据总个数为 $n_0 = n - l + a$,令由大到小排列系列为 $\{x_m, m = 1,2,\cdots,n_0\}$,则经本文作者整理由文献[3]规定的概率权重矩公式如下*:

$$b_0 = \frac{1}{N}\left[\sum_{m=1}^{n_0} x_m + \frac{a}{n_0 - a}\sum_{m=a+1}^{n_0-a} x_m\right] \tag{10}$$

$$b_1 = \frac{1}{N}\left[\sum_{m=1}^{a}\frac{(N-m)}{N-1}x_m + \frac{N+1-a}{N+1}\,\frac{N-a}{n_0-a}\sum_{m=a+1}^{n_0}\frac{n_0-m}{n_0-a-1}x_m\right] \tag{11}$$

$$b_2 = \frac{1}{N}\left[\sum_{m=1}^{a}\frac{(N-m)(N-m-1)}{(N-1)(N-2)}x_m + \left(\frac{N-a+1}{N+1}\right)^2\frac{N-a}{n_0-a}\sum_{m=a+1}^{n_0}\frac{(n_0-m)(n_0-m-1)}{(n_0-a-1)(n_0-a-2)}x_m\right] \tag{12}$$

给定样本 X,由式(10)(11)(12)可估计出 b_0,b_1,b_2,再由式(8)则可估计出线性矩 $\lambda_1,\lambda_2,\tau_3$.

3　线性矩法的统计试验研究

对矩法,绝对值准则适线法与两种不同的可考虑历史洪水线性矩法作了统计试验比较研究.

3.1　参数优劣评判标准

参数估计方法比较时,以参数、设计值不偏性及有效性为标准,对于设计值采用相对偏差均值 B_{X_p} 和相对根均方误 S_{X_p} 表示,计算公式见文献[4],P 取 0.1% 和 1% 两种. 统计试验组数 $Ns = 1000$,B_{X_p} 为正值表示偏大,负值表示偏小,$|B_{X_p}|$ 愈大表示偏差愈大,S_{X_p} 愈小愈有效. 对两种不同的可考虑历史洪水线性矩法还从线

* 文献[3]第 19 页的公式(A15)中概率权重矩 M_2 计算公式有误,把公式(12)中$(N-m-1)$错误印刷成$(n-m-1)$,本文(12)式是作了更正后的公式.

性矩的估计效果作比较.

3.2 试验方案设计

3.2.1 连序样本

实测序列长度 n 取 30 和 50,总体参数 $EX_0=1.0$,$C_{V0}=0.3$,0.5,1.0,C_{S0}/C_{Vo}取 3.5,4,5,共 18 种方案.

3.2.2 非连序样本

最大重现期 $N=100$,n 取 30 和 50 两种,历史洪水个数 1,3,总体参数 $EX_0=1.0$,$C_{Vo}=0.3$,0.5,1.0,C_{S0}/C_{Vo}取 3.5,4.0,5.0,共 36 种方案.此外还对最大重现期为 150、历史洪水个数为 3 的方案作了计算.

4 试验结果分析

大量统计试验计算结果表明(部分结果列于表 1,表 2):

a. 两种可考虑历史洪水的线性矩估计公式,在不同总体参数、样本组成情况下,均具有较高精度.简单样本时两种公式是一致的,但当有历史洪水时,不管从样本线性矩还是从分布参数、设计值的不偏性及有效性来看,由文献[1]建议的线性矩公式比由文献[3]中规定采用的公式均略好.

b. 虽然矩法具有较好的有效性,但不偏性差,而线性矩法不论是参数,还是设计值都具有良好的不偏性和有效性,因此,线性矩法比矩法在统计性能上有很大提高.在各种情况下,不论从参数、设计值的不偏性还是有效性来讲,线性矩法均比适线法有较大程度的提高.总之,线性矩法是一种性能优良的对数正态分布参数估计方法.

c. 本文所建议的具有历史洪水时线性矩估计公式有效合理,这是因为在有历史洪水时,线性矩法能保持在简单样本时具有良好的不偏性及有效性.

d. 与 P-Ⅲ分布时的线性矩法比较,对数正态分布线性矩法其参数及设计值不偏性略差.从文献[1]的结果可知,P-Ⅲ分布线性矩法所估计的参数 C_V 和 C_S 基本上是无偏的,但本文结果却发现所估计的参数 C_V 和 C_S 仍有一定程度的偏差.

表 1 参数、设计值不偏性及有效性计算成果(部分方案)

Table 1 Calculated results of unbias and effectiveness of parameters and their designed values

总体参数与样本组成					估计方法	EC_V	EC_S	SC_V	SC_S	B_{Xp1}	B_{Xp2}	S_{Xp1}	S_{Xp2}
C_{Vo}	C_{So}	N	n	a									
0.5	2.0	30	30	0	MVS	0.48	1.62	0.10	0.66	-5.05	-8.34	20.47	27.67
0.5	2.0	30	30	0	CFM	0.53	2.22	0.12	1.50	3.95	8.44	25.25	41.78
0.5	2.0	30	30	0	LMGH	0.51	2.15	0.11	1.35	0.58	3.77	24.17	39.44
0.5	2.0	30	30	0	LM	0.51	2.15	0.11	1.35	0.58	3.77	24.17	39.44
0.5	2.0	100	30	1	MVS	0.49	1.85	0.07	0.77	-3.19	-4.81	15.16	21.59
0.5	2.0	100	30	1	CFM	0.54	2.45	0.11	2.72	5.54	11.1	21.29	39.61
0.5	2.0	100	30	1	LMGH	0.51	2.16	0.09	1.05	1.36	4.00	18.75	29.88
0.5	2.0	100	30	1	LM	0.51	2.13	0.08	1.04	1.03	3.37	18.67	29.61
0.3	1.5	100	30	3	MVS	0.29	1.43	0.04	0.59	-2.15	-2.86	9.65	14.96
0.3	1.5	100	30	3	CFM	0.32	1.84	0.05	1.14	4.35	9.16	12.99	24.98
0.3	1.5	100	30	3	LMGH	0.31	1.67	0.04	0.76	1.20	3.59	11.30	19.45
0.3	1.5	100	30	3	LM	0.30	1.59	0.04	0.73	0.20	1.65	11.12	18.75
1.0	5.0	100	50	1	MVS	0.94	3.29	0.21	2.06	-7.21	-15.85	26.61	34.63
1.0	5.0	100	50	1	CFM	1.09	6.65	0.31	5.34	5.44	12.98	31.3	53.05
1.0	5.0	100	50	1	LMGH	1.03	5.89	0.26	3.83	1.18	5.50	28.78	45.68
1.0	5.0	100	50	1	LM	1.03	5.82	0.25	3.77	0.91	4.94	28.72	45.39

注:EC_V 和 EC_S 表示 N_S 次统计试验样本 C_V 和 C_S 的平均值,SC_V 和 SC_S 表示样本 C_V 和 C_S 的均方误差,LM 表示由本文作者[1]建议的线性矩法,LMGH 表示由文献[3]规定的线性矩法,CFM 表示绝对值准则适线法,MVS 表示矩法,$P_1=1\%$,$P_2=0.1\%$.

表 2　线性矩不偏性及有效性统计试验结果(部分方案)

Table 2　Statistic experiment results of unbias and effectiveness of L-moments

总体参数及样本组成							$E\tau$	$E\tau_3$	$S\tau$	$S\tau_3$	估计方法
C_{Vo}	C_{So}	N	n	a	τ	τ_3					
0.5	2.0	100	30	1	0.25448	0.26491	0.25403	0.26073	0.030	0.082	LM
0.5	2.0	100	30	1	0.25448	0.26491	0.25585	0.25971	0.031	0.084	LMGH
0.3	1.5	100	30	3	0.15841	0.21434	0.15780	0.21387	0.017	0.072	LM
0.3	1.5	100	30	3	0.15841	0.21434	0.15827	0.21679	0.017	0.075	LMGH
1.0	5.0	100	50	1	0.42000	0.42913	0.41766	0.42331	0.045	0.081	LM
1.0	5.0	100	50	1	0.42000	0.42913	0.41260	0.40942	0.052	0.098	LMGH

注：$E\tau$，$E\tau_3$，$S\tau$，$S\tau_3$ 分别表示线性矩 τ 和 τ_3 的 N_S 次统计试验的平均值和均方误差.

5　结　　语

a. 本文对对数正态分布下的线性矩与分布参数的相互关系、样本线性矩的计算作了较全面的分析与介绍，这在国内文献中尚属首次，希望能起到抛砖引玉之作用.

b. 与矩法、绝对值准则适线法所作的统计试验相比较，结果表明，线性矩法比矩法、适线法有较大优越性，可以说是一种性能优良的好方法，建议在实际工作中加以采用.

c. 本文作者所建议的可考虑历史洪水的线性矩公式，经论证是有效合理的.

d. 如何利用线性矩进行水文频率计算的区域综合有待进一步研究.

参考文献：

[1] 陈元芳，沙志贵，陈剑池，等. 具有历史洪水时 Pearson-Ⅲ分布线性矩法的研究[J]. 河海大学学报(自然科学版)，2001(4)：76—80.

[2] HOSKING J R M，WALLIS J R. Regional frequency analysis—an approach based on L-moment[M]. London：Cambridge Unirersity Press，1997. 1—280.

[3] 中华人民共和国水利部. 水利水电工程设计洪水计算规范[M]. 北京：水利电力出版社，1993. 19—20.

[4] 陈元芳. 统计试验方法及应用[M]. 哈尔滨：黑龙江人民出版社，2000. 129—130.

Study on L-moment estimation method with consideration of historical floods for log-normal distribution

CHEN Yuan-fang[1]，SHA Zhi-gui[2]，GU Sheng-hua[3]，XU Sheng-bin[1]

(1. *College of Water Resources and Environment*，*Hohai Univ.*，*Nanjing* 210098，*China*；
2. *Yangtze River Water Resources Commission*，*Wuhan* 430010，*China*；
3. *Shanghai Hydrological General Station*，*Shanghai* 200232，*China*)

Abstract：The relationship between L-moments and parameters of the log-normal distribution was briefly introduced，then two formulae considering historical floods were proposed for estimation of the L-moments of samples. The Monte-Carlo calculation shows that both of the two formulae are of high accuracy in the estimation of the parameters and their design values，and the better one is recommended to be used in practice. A comparison of the moment method，the curve-fitting method，and the L-moment estimation method was carried out by use of the Monte-Carlo experiment，and the results show that the L-moment estimation method is the best.

Key words：L-moment；curve-fitting method；moment method；log-normal distribution；historical flood；unbias；effectiveness