

关于适线法中经验频率计算公式的对比研究

黄振平, 萨迪伊, 王春霞, 马军建

(河海大学水资源环境学院, 江苏 南京 210098)

摘要: 阐述适线法的理论基础并简介 10 个经验频率计算公式. 采用统计试验方法, 选取常见的 8 个经验频率公式, 对 P-Ⅲ型分布适线计算的设计值优劣进行对比分析. 结果表明: 用数学期望公式计算的结果偏大; 华氏公式、海森公式、Маматов公式的计算结果相对较好. 研究还表明, 一般用绝对值准则适线求得的结果要优于平方和准则适线的结果.

关键词: 经验频率公式; 适线法; P-Ⅲ型分布; 设计值

中图分类号: P333.9

文献标识码: A

文章编号: 1006-7647(2002)05-0005-03

适线法推求设计洪水, 是目前我国广泛使用的一种方法. 其成果的好坏决定于 3 个方面: 一是总体分布线型的选取是否适当; 二是经验频率的计算是否合理; 三是用什么准则判断理论频率曲线与经验点据拟合的好坏. 由于选择统一线型的理论基础不充分, 目前广泛采用 P-Ⅲ型曲线, 是一种带有规定性的假定. 经验点据绘点位置是适线的基础, 比适线方法和准则重要得多. 到目前为止, 已有十几个经验频率公式, 有的有一定的理论基础, 有的则纯属经验性的^[1]. 研究表明, 用不同的经验频率公式计算的频率, 在样本的中间部分相差较小, 但在两端相差比较大. 而在水文频率计算中两端(尤其是上端)拟合的优劣恰恰是最重要的, 因为它直接影响所推求的设计值的大小. 因此, 有必要对一些常用的经验频率公式进行对比研究, 为我们适线时选择合理的经验频率公式提供参考.

1 适线法的理论基础

设 $F(x, \theta)$ 为总体分布函数, θ 为参数, $(x_1, x_2, \dots, x_n)$ 为 X 的一组样本, x_i 与其所对应的超值概率 p_i 的关系可表示为

$$x_i = g(p_i, \theta) \quad i = 1, 2, \dots, n \quad (1)$$

$$\text{或 } p_i = P(X > x_i) = P(x_i, \theta) = \int_{x_i}^{\infty} dF(x, \theta)$$

$$i = 1, 2, \dots, n \quad (2)$$

由选择的经验频率公式计算出 $\hat{p}_i$ 作为 p_i 的估计值, 则公式(1)和(2)可改写为

$$x_i = g(\hat{p}_i, \theta) + \epsilon_i \quad i = 1, 2, \dots, n \quad (3)$$

$$P(x_i, \theta) = \hat{p}_i + \delta_i \quad i = 1, 2, \dots, n \quad (4)$$

为计算参数 θ 的估计值 $\hat{\theta}$, 我们自然会想到取使估计误差 ϵ 或 δ 达到最小的 θ 值. 即:

$$\sum_{i=1}^n (x_i - g(\hat{p}_i, \hat{\theta}))^2 = \min_{\theta} \sum_{i=1}^n (x_i - g(\hat{p}_i, \theta))^2 \quad (5)$$

此即纵标最小二乘准则. 或

$$\sum_{i=1}^n |x_i - g(\hat{p}_i, \hat{\theta})| = \min_{\theta} \sum_{i=1}^n |x_i - g(\hat{p}_i, \theta)| \quad (6)$$

此即纵标绝对值准则. 类似, 可得横标最小二乘准则和横标绝对值准则^[2].

2 经验频率公式简介

由于经验点据的绘点位置, 即经验频率公式的选取, 直接影响到适线结果, 因此, 长期以来, 许多水文工作者都致力于改进经验频率计算公式, 并从不同角度出发, 推荐了许多公式, 主要有以下几种.

2.1 以 p_i 的特征值计算 $\hat{p}_i$

a. 数学期望公式.

$$P_i = \frac{i}{n+1} \quad (7)$$

Weibull 通过解析途径导出 $\hat{p}_i$ 的计算式, 通常称之为数学期望公式, 因为此公式简单且与分布无关, 因此被广泛采用^[3].

b. 中值公式^[4].

$$P_i = \frac{i-0.3}{n+0.4} \quad (8)$$

该公式是 $\hat{p}_i$ 中值近似公式.

c. Алексеев 公式^[4].

$$P_i = \frac{i-0.25}{n+0.5} \quad (9)$$

Алексеев 公式只有 n 较大时才适用。

2.2 以 x_i 的特征值计算 $\hat{p}_i$

a. Blom 公式.

$$P_i = \frac{i - 0.375}{n + 0.25} \quad (10)$$

Blom 公式是由正态分布导出的近似公式。

b. Gringorten 公式.

$$P_i = \frac{i - 0.44}{n + 0.12} \quad (11)$$

Gringorten 公式是对极值曲线由 $P(EX_i)$ 出发得到的近似公式,仅适用于 $i = 1$ 的情况。

c. 华氏公式^[4].

$$P_i = \frac{i - 0.45}{n + 0.1} \quad (12)$$

这是河海大学华家鹏教授通过数值积分途径求得的适用于 $C_s = 1 \sim 4$ 时的近似无偏计算公式。

2.3 其他类型的公式

a. 简单公式.

$$P_i = \frac{i}{n} \quad (13)$$

该公式是根据概率的古典定义得出的,它只有在 n 较大时才能适用。

b. 海森公式.

$$P_i = \frac{i - 0.5}{n} \quad (14)$$

海森公式是简单公式的修正式,各项经验频率取简单公式相邻两项经验频率的中间值。Cunnane^[2]认为,对偏态系数为正的分布,如 P-III 型分布,海森公式是较合适的。

c. Tukey 公式.

$$P_i = \frac{3i - 1}{3n + 1} \quad (15)$$

d. Маматов 公式.

$$P_i = \frac{i - 0.5}{n + 1} \quad (16)$$

3 研究方法

上面所讨论的经验频率公式,有的有一定理论

基础,有的是纯属经验性的。目前尚不能从理论上确定哪个公式更适合 P-III 型分布。我们从上述公式中选择了常见的 8 个公式(即简单公式、数学期望公式、华氏公式、海森公式、中值公式、Blom 公式、Tukey 公式和 Маматов 公式),通过统计试验方法来评价它们的优劣。

假定一组 P-III 型分布参数,本次研究设 $EX = 1000$, $C_v = 1.0$, $C_s = 2C_v$ 。计算对应设计频率 P 的理论值 x_p^0 ,又称为真值。从总体中抽取样本容量为 N 的 m 组样本,对用不同经验频率公式计算的经验频率,采用适线法,求得 m 个对应设计频率 P 的设计值 $x_{p,1}, x_{p,2}, \dots, x_{p,m}$,根据如下评判指标来衡量各经验频率公式的优劣:

$$\text{设计值均值: } \bar{x}_p = \frac{1}{m} \sum_{i=1}^m x_{p,i};$$

$$\text{设计值均值的相对误差: } \epsilon_x = \frac{\bar{x}_p - x_p^0}{x_p^0} \times 100\%;$$

$$\text{设计值的相对误差: } \delta_i = \frac{|x_{p,i} - x_p^0|}{x_p^0} \times 100\%;$$

$$\text{设计值相对误差的均值: } \bar{\delta} = \frac{1}{m} \sum_{i=1}^m \delta_i;$$

$$\text{设计值相对误差的方差: } \epsilon_\delta = \frac{1}{m} \sum_{i=1}^m (\delta_i - \bar{\delta})^2.$$

根据我国资料的实际情况,本文样本容量 N 取 50。设计频率 P 分别取 1% 和 0.1%。样本组数 m 取为 500。每组样本系列的参数估计方法是:均值采用矩法估计, C_v 和 C_s 采用绝对值准则和平方和准则适线求得。

4 结果分析

计算结果见表 1 和表 2。

从表 1 和表 2 可以看出:用简单公式或数学期望公式计算经验频率,用平方和准则适线,求得的设计值误差较大,而华氏公式、海森公式、Маматов 公式计算结果比较理想;一般来讲,绝对值准则适线计算结果要优于平方和准则适线的结果;重现期越大,

表 1 设计值均值及设计值均值相对误差

经验频率公式	平方和准则				绝对值准则			
	$P = 1\%$		$P = 0.1\%$		$P = 1\%$		$P = 0.1\%$	
	均值	相对误差/%	均值	相对误差/%	均值	相对误差/%	均值	相对误差/%
简单公式	5160	12.05	8022	16.12	4650	0.98	7006	1.41
数学期望公式	5153	11.89	8006	15.90	4735	2.82	7154	3.56
华氏公式	4665	1.30	7063	2.24	4589	-0.34	6891	-0.25
海森公式	4608	0.07	6954	0.66	4569	-0.78	6853	-0.79
中值公式	4833	4.96	7389	6.97	4646	0.90	6995	1.27
Blom 公式	4749	3.13	7226	4.60	4619	0.30	6944	0.53
Tukey 公式	4804	4.32	7337	6.21	4579	-0.57	6877	-0.45
Маматов 公式	4603	-0.05	6940	0.47	4654	1.06	7007	1.43

注:百年一遇真值为 4605,千年一遇真值为 6908;均值为设计值均值;相对误差为设计值均值与真值的相对误差。

表2 设计值相对误差均值及方差

经验频率公式	平方和准则				绝对值准则			
	相对误差均值/%		相对误差方差		相对误差均值/%		相对误差方差	
	$P=1\%$	$P=0.1\%$	$P=1\%$	$P=0.1\%$	$P=1\%$	$P=0.1\%$	$P=1\%$	$P=0.1\%$
简单公式	22.42	27.83	0.2202	0.2267	15.67	17.71	0.1241	0.1258
数学期望公式	22.47	28.08	0.2194	0.2265	15.98	18.33	0.1294	0.1315
华氏公式	18.43	22.39	0.1563	0.1612	15.47	17.73	0.1240	0.1260
海森公式	18.15	21.99	0.1509	0.1557	15.41	17.64	0.1230	0.1250
中值公式	19.53	23.94	0.1752	0.1807	15.62	17.91	0.1265	0.1285
Blom公式	18.94	23.11	0.1653	0.1705	15.54	17.81	0.1248	0.1269
Tukey公式	19.16	23.28	0.1713	0.1762	15.33	17.43	0.1233	0.1251
Маматов公式	18.12	22.10	0.1497	0.1549	15.67	18.01	0.1253	0.1275

设计值的计算误差越大;样本容量越大,计算精度越高;百年一遇设计值相对误差的均值为15%左右,一般不超过16%,设计值均值的最大相对误差不超过3%,而千年一遇设计值相对误差的均值为18%左右,一般不超过20%,设计值均值的最大相对误差不超过4%。

5 结 语

虽然以前许多学者也用不同方法定性地研究得出:在适线时,使用数学期望公式,计算出的设计值偏大。但是,出于工程安全性的考虑,我国设计部门仍广泛地使用数学期望公式。本文通过统计试验方法,定量地得出8种经验频率公式的计算误差,结果

表明,用华氏公式、海森公式、Маматов公式计算经验频率,然后用绝对值准则适线求得的设计值有较高的准确性。

参考文献:

- [1] 金光炎.水文频率分析述评[J].水科学进展,1999(3):319~327.
- [2] 潘军宁.关于适线法中经验频率公式的讨论[J].海洋工程,1996(3):65~73.
- [3] 黄振平.水文统计原理[M].南京:河海大学出版社,2002.143~145,160~161.
- [4] 华家鹏.关于不偏的洪水经验频率计算公式的研究[J].水文,1984(4):5~11.

(收稿日期:2002-02-07 编辑:马敏峰)

·简讯·

水利人的网上家园——水信息网

水信息网是由水利部海河水利委员会主办、天津市龙网科技发展有限公司承办的面向水利的大型行业门户网站,自1999年11月19日开通以来,水信息网为读者提供了全面、及时的国内国际水利行业新闻、图片视频信息和科技论文。拥有一大批涵盖行业高层领导、学界名家的忠实用户。水信息网现已成为业内影响较广的网站之一,也是目前我国较大的行业网络媒体之一,水信息网等水利行业四大骨干网站的开通与发展被列入我国20世纪水利大事年表。

水利部部长汪恕诚,副部长敬正书、张基尧、翟浩辉、陈雷,纪检组长刘光和,总工程师高安泽、何文垣;天津市市长李盛霖、副市长孙海麟等领导先后视察网站,并多次对网站的发展作了直接指示。

全心全意为水利人服务是水信息网一贯奉行的宗旨,水信息网希望能为读者提供最完善的服务:水事频道报道国内各类一线媒体的水利新闻,更新速度快;技术频道拥有全世界范围的原创征文,并与多家国内权威水利学术(技术)刊物有合作伙伴关系;国际频道关注国际热点,每天都有最新译稿上网;商务频道商家云集,蕴涵无限商机。水信息网有涵盖水利行业方方面面的16个大型水利数据库,并设有学者之家和记者之家,来自技术界和新闻界的众多名家们在这里筑起自己的温馨家园;水信息网还拥有充分展示网络多媒体优势的水利视频新闻栏目……

水信息网域名: <http://www.hwcc.com.cn> 或 <http://www.waterinfo.net.cn>

网管信箱: news@hwcc.gov.cn

联系电话:022-24102117,24102110

联系人:汪先生,史小姐

(汪浩波供稿)