

DOI :10.3876/j.issn.1000-1980.2012.01.005

论混合型总体的频率分析按成因组合的必要性

秦 毅 ,李子文 ,颜 恒 ,万 俊

(西安理工大学西北水资源与环境生态教育部重点实验室 陕西 西安 710048)

摘要 :针对混合型总体的频率分析是否需要按成因分开后再组合计算这一问题 ,依据概率论与数理统计理论 ,首先明确基本概念 ,之后对频率组合方法的应用范围进行理论分析与讨论 ,并以实例对混合型总体频率分析分成因与不分成因这 2 种方法的计算结果进行了比较、分析验证。结果表明 :对于混合型总体 ,分与不分成因的频率分析结果是一致的 ;在没有人干扰和超乎寻常的突发事件影响时 ,只要保持自然规律性 ,混合型总体也是一致性总体 ,频率分析没有必要采用分类频率组合的方法。

关键词 :混合型总体 ;频率分析 ;频率组合法

中图分类号 :P333.6 文献标志码 :A 文章编号 :1000-1980(2012)01-0023-05

Necessity of frequency combination according to different causes for analysis of a mixed population

QIN Yi , LI Zi-wen , YAN Heng , WAN Jun

(Key Laboratory of Northwest Water Resources and Environment Ecology of Ministry of Education , Xi 'an University of Technology , Xi 'an 710048 , China)

Abstract : In order to examine whether it is necessary to conduct frequency analysis through combination according to different causes of storm floods , the probability theory and the mathematical statistics theory were used to analyze the application range of the frequency combination method . A case study was conducted on a mixed population to compare , analyze , and verify the calculated results of the two methods , i . e . , the method of frequency combination analysis according to different causes and the method of frequency analysis without considering the causes . The results show that the frequency analysis results of the two methods are consistent , and the mixed population is a population of consistency when there is no influence of human disturbance or unusual emergency . Thus , it is not necessary to conduct classified frequency combination analysis .

Key words : mixed population ; frequency analysis ; frequency combination method

水文频率分析时要求样本具有一致性 ,即各样品均出自同一总体或说具有同一成因 ,这是常规分析方法^[1-4] (例如适线法)只适应于简单样本决定的。但是 ,实践中碰到多成因样本的情况屡见不鲜 ,如河流交汇处的流量 ,受顶托河段的洪峰水位 ,由暴雨、融雪、冰川湖溃决 3 种因素引发的洪水等。为满足简单样本适应的数理统计理论的要求 ,针对多成因总体的频率分析 ,SL 44—2006《水利水电工程设计洪水计算规范》^[5]规定 :“许多地区的洪水常由不同成因(如融雪、暴雨)、不同类型(如台风、锋面)暴雨形成。一般认为它们是不同的分布的 ,不宜把它们混在一起作为一个洪水系列进行频率计算 ,也不能把由于垮坝所形成的洪水加入系列作频率计算。”另外 ,国内外一些研究指出 ,洪水季节 ,形成洪水的地区不同都可能影响“同分布”假定^[6]。由此可知 ,水文频率分析时不宜将成因、统计特性明显不同的洪水当作一个洪水系列。

一些学者提出了应用洪水频率组合原理来解决这类问题的方法。例如 :谢迪全等^[7]应用频率组合原理分

析了湘潭站设计洪水位的问题 姚杰等^[8]将该原理应用于水库实际防洪能力分析 杨建青等^[9]也采用同样的原理得出了黄浦江防潮工程设计更合理的上海市区最高潮(水)位 刘燕等^[10]甚至将该方法应用到了干旱频率分析及预测中.所有方法具有共同特点——分析计算过程复杂.

那么,存在这种异分布洪水,甚至发生任何随机组合事件的情况下,是否需要进行频率组合分析计算才能得出更合理的结果,是一个值得探讨的问题.为此,本文对此问题进行了研究.

1 基本概念

首先从数理统计和概率论的角度明确基本概念并进行探讨.从数理统计术语出发,总体即为研究对象的全体.从概率论术语出发,总体也是随机变量 X 所有可能取值的全体,或者说是随机试验的所有可能结果的集合.所以一个随机试验也可以称之为一个总体.

1.1 随机试验

随机试验是对某种随机现象进行的观测、调查或试验.随机试验具有以下特征 (a)试验可以在相同条件下重复进行 (b)试验的可能结果有多种,且试验之前明确知道它有可能的结果 (c)每次试验肯定能得出一种结果,但这种结果是什么预先不能确定.

需要强调指出的是在随机试验特征的表述中,并未强调随机试验所观测结果的属性.换句话说,这个观测结果可以是简单随机现象,也可以是复杂现象.例如:对一批炮弹射程(其中第 i 发炮弹的射程为 d_i)的观测,就是简单现象的试验,也称单一试验,而同时投骰子和抛硬币来观测骰子的点数与硬币出现正反面的情况,就是复杂现象的试验,也称复合试验.简单试验(单一试验)的结果单一,比如 d_i ;复杂试验(复合试验)的结果复杂,是复合结果,如骰子点数为 3,同时硬币出现反面.

总之,试验是简单的还是复杂的,取决于研究对象的简单和复杂与否.

1.2 混合总体

水文频率分析时将一条河流生命期内的年最大流量作为研究对象,则对该河年最大流量的观测就相当于做随机试验,各次试验结果也即 $Q_{\max,i}$ 是随机变量.显然,如果该河流洪水完全由同一性质的暴雨产生,这个随机试验就是简单(单一)试验,如果洪水由不同性质的暴雨如台风雨、梅雨产生,或由融冰雪、融冰溃决形成,则这个随机试验就是复杂试验.

根据总体与试验的关系,复杂试验所对应的总体为复杂(或说复合)总体,在水文统计学中也称之为混合总体.例如 (a)掷骰子和抛硬币试验结果就是一个混合总体,试验结果为 (1,反面)(1,正面)……(6,反面),(6,正面)共 12 个 (b)叶尔羌河卡群站年最大洪峰流量^[7]也是一个混合总体,试验结果为 $Q_i = Q_{\text{冰雪},i} + Q_{\text{暴雨},i} + Q_{\text{冰溃},i}$ ($i = 1, 2, \dots, n$) (河流生命期内) (c)干支流交汇后的洪峰流量也是混合总体,试验结果为 $Q_i = Q_{A\text{支流},i} + Q_{B\text{支流},i} + Q_{C\text{干流},i}$.

可见,复杂试验的结果也是复杂的.

2 频率组合方法的应用范围

2.1 视研究目标而定

这个问题不能简单定论,其结论应视研究目标而定.如图 1 所示,存在 2 种情况 (a)已知干流 C 的流量 Q_C 和支流 B 的流量 Q_B ,求支流 A 的设计流量 Q_A (b)已知 Q_A, Q_B ,下游设计目标区域 C 没有观测站,求干流 C 的流量 Q_C .这时,进行求解当然需要频率组合,即 $Q_A = Q_C - Q_B$,由 Q_B 的总体分布情况和 Q_C 的总体分布情况求 Q_A 的总体分布情况,或者 $Q_C = Q_A + Q_B$,由 Q_A 的总体分布情况和 Q_B 的总体分布情况求 Q_C 的总体分布情况.

若基本情况如图 2 所示,显然不论哪条支流发生洪水, Z 点必发生洪水,此时若在 Z 处设站观

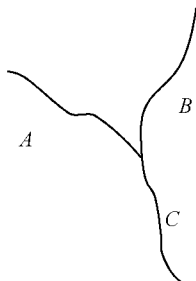


图 1 需要进行频率组合的情况
Fig. 1 Case in need of frequency combination

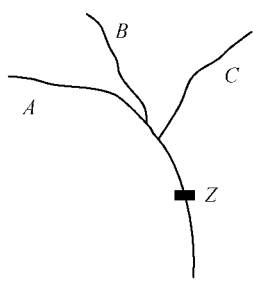


图 2 无需进行频率组合的情况
Fig. 2 Case in no need of frequency combination

测 就如同对以上各支流和干流洪水形成的混合总体做试验 ,即 $Q_Z = Q_A + Q_B + Q_C$. 而 Q_Z 的总体分布特性也必然是上游各支流和干流各随机变量总体分布特性的组合 ,所以 ,若以 Q_Z 为研究对象 ,则直接对 Q_Z 进行频率分析即可 ,而无需按成因分开后再进行组合.

举例说明如下.

例 随机试验为掷骰子和抛硬币 ,观测出现的花色 Z ,求其频率分布律(总体分布特性).

解 试验的所有可能结果为 $C_6^1 C_2^1 = 6 \times 2 = 12$ (种). 现采用 2 种途径求解.

途径 1 :直接对 Z 进行分析 ,则其分布律 P 如表 1 所示. 若令 $Y = 0$ 表示硬币出现反面(-), $Y = 1$ 表示硬币出现正面(+),则 Z 的分布律如表 2 所示.

表 1 途径 1 计算的 Z 的概率分布律

Table 1 Probability distribution of Z calculated by first method

Z	(1 , +)	(1 , -)	(2 , +)	(2 , -)	(3 , +)	(3 , -)	(4 , +)	(4 , -)	(5 , +)	(5 , -)	(6 , +)	(6 , -)
P	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

表 2 表 1 的另一种表达方式

Table 2 Another expression of Table 1

Z	2	1	3	2	4	3	5	4	6	5	7	6
P	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

对表 2 整理后得 Z 的最终分布律 ,见表 3.

表 3 Z 的最终分布律

Table 3 Final probability distribution of Z

途径 2 :从另一个角度求 Z 的分布律. 设 X 为骰子出现的点数 , Y 为硬币出现正、反面的情况 ,则 X 和 Y 的分布律如表 4 和表 5 所示. 显然 , $Z = X + Y$,则 Z 的取值为 1 2 3 4 5 6 7.

Z	1	2	3	4	5	6	7
P	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$

表 4 X 的分布律

Table 4 Probability distribution of X

X	1	2	3	4	5	6
P_X	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

表 5 Y 的分布律

Table 5 Probability distribution of Y

Y	(反)	(正)
P_Y	$\frac{1}{2}$	$\frac{1}{2}$

Z 的分布律计算过程为

$$\begin{aligned} P(Z = 1) &= P(X = 1, Y = 0) = P_X(X = 1)P_Y(Y = 0) = \frac{1}{6} \times \frac{1}{2} = \frac{1}{12} \\ P(Z = 2) &= P(X = 2, Y = 0) + P(X = 1, Y = 1) = \frac{1}{6} \times \frac{1}{2} + \frac{1}{6} \times \frac{1}{2} = \frac{2}{12} \\ P(Z = 3) &= P(X = 3, Y = 0) + P(X = 2, Y = 1) = \frac{1}{6} \times \frac{1}{2} + \frac{1}{6} \times \frac{1}{2} = \frac{2}{12} \\ P(Z = 4) &= P(X = 4, Y = 0) + P(X = 3, Y = 1) = \frac{2}{12} \\ P(Z = 5) &= P(X = 5, Y = 0) + P(X = 4, Y = 1) = \frac{2}{12} \\ P(Z = 6) &= P(X = 6, Y = 0) + P(X = 5, Y = 1) = \frac{2}{12} \\ P(Z = 7) &= P(X = 6, Y = 1) = \frac{1}{6} \times \frac{1}{2} = \frac{1}{12} \end{aligned}$$

由上述计算可得出 Z 的分布律 ,如表 6 所示.

表 6 由 X 和 Y 组合得出的 Z 分布律

Table 6 Probability distribution of Z with combination of X and Y

Z	1	2	3	4	5	6	7
P_Z	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$

其结果与直接对 Z 进行分析完全相同 .这就证明当研究对象为 Z 时 ,并不需要

采用途径2,即不需要先分成因 X 和 Y 再进行组合.

2.2 按成因分开组合的实际可操作性

按成因分开组合这种方法无疑是正确的,但这种方法实际操作很困难,原因有2个:
从理论上讲,水文站观测到的结果往往是复杂总体的复杂结果 Z ,是许多变量的函数,如

$$Z=f(A,B,C,\dots)$$

在仅知道 Z 的条件下,要正确区分出 A,B,C 等是不可能的,因为1个方程只能解1个未知数,1个方程求多个未知数只能是不定解.即便有辅助资料作参考,也会受很多不确定性因素的影响而使分解出的 A,B,C 带有很大的任意性,同时受人为因素影响也大,从而失去可靠性.地表水与地下水的划分就是一例.用这种人为影响较大的分离结果再进行组合,其结果当然很难说是可靠的,也很难进行合理性分析.除非各个成因都能进行观测,否则按成因先分开再组合不具有实际可操作性.

从资料系列实际情况看,现有资料系列太短,如果把不同成因的洪水强行分开,则资料容量势必大幅度减小,其误差比用混合系列进行频率分析可能更大^[11].

3 实例分析

以新疆某流域混合型设计洪水计算为例,来说明是否需要按成因组合进行洪水频率组合分析.
该流域有控制水文站A及长期洪水观测资料.对该流域洪水的多年研究表明,A站以上流域内存在多种类型的洪水,除存在通常的融雪型、暴雨型、融雪与暴雨混合型洪水之外,还存在冰川阻塞溃坝型的突发洪水(受气温影响,这种洪水每年都会发生)及融雪与溃坝的混合型突发洪水.

本文分别采用分型与不分型2种方法对A站进行频率分析.

3.1 不分型设计洪水计算

将水文站A的特大历史洪水和实际观测资料中冰川溃决型洪水组成不连续系列直接进行适线法频率分析,结果见表7和图3.

3.2 按成因分型频率组合法设计洪水计算

SL 44—2006《水利水电工程设计洪水计算规范》^[5]规定:
“冰川湖溃决的洪水,不宜直接加入频率计算.”因此,流域的洪水可分成冰川阻塞湖溃决型洪水和融雪与暴雨型洪水2类.通过分析得到的2类洪水的频率曲线如图4所示.以文献^[12]给出的频率组合列表法进行组合频率分析,相应的计算结果见表7.

3.3 不分型洪水计算结果与频率组合计算结果对比

图3显示了2种分析方法所得结果的比较.由图3和表7可以看出,不分型的频率曲线与分型条件下用频率组合法所得的频率曲线相对误差并不大,范围在5.08%~8.79%之间.误差产生的原因在于(a)分型条件下采用频率组合法计算时,所用到的2种成因洪水的频率曲线被离散化取值,相对于连续函数解析结果必然会有误差.(b)无法认证人为分型后得出的各成因统计规律是否合理,故可能使频率组合结果带有误差.

4 结 语

硬币与骰子的总体频率分析和新疆A站频率分析结果表明,对于混合型总体,分与不分成因的频率分析结果是一致的.换句话说,在没有人类干扰和超乎寻常的突发事件影响时,只要保持自然规律性,混合型总体也是一致性总体,因此频率分析没有必要采取分成因再组合的频率分析方法.

表7 水文站A的设计洪水计算结果
Table 7 Calculated results of design flood at hydrological station A

设计频率/%	不分型设计 洪峰流量/ ($\text{m}^3\cdot\text{s}^{-1}$)	分型设计 洪峰流量/ ($\text{m}^3\cdot\text{s}^{-1}$)	相对 误差/%
0.1	13 540	12 350	8.79
1	8 786	8 202	6.65
2	7 391	6 954	5.91
5	5 590	5 304	5.12
10	4 272	4 055	5.08

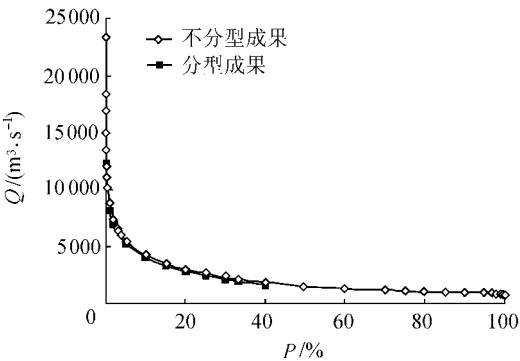


图3 结果比较
Fig. 3 Comparison results of two methods

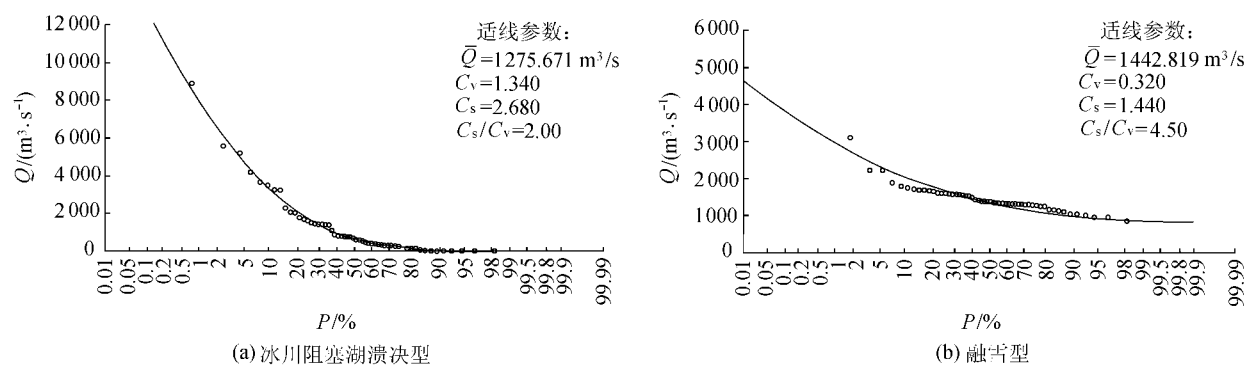


图 4 A 站不同成因引起的年最大洪峰流量频率曲线

Fig. 4 Frequency curves of annual maximum peak discharge formed by different causes at hydrological station A

参考文献：

[1] 徐晓 林盛吉 张徐杰 等. 气候变化对衢州水文站洪水频率分析的影响[J]. 河海大学学报 : 自然科学版 , 2010 , 38(6) : 625-628.(XU Xiao , LIN Sheng-ji , ZHANG Xu-jie , et al. Impact of climate change on flood frequency analysis at Quzhou hydrological station [J]. Journal of Hohai University : Natural Sciences , 2010 , 38(6) : 625-628.(in Chinese))

[2] 杨涛 陈喜 杨红卫 等. 基于线性矩法的珠江三角洲区域洪水频率分析[J]. 河海大学学报 : 自然科学版 , 2009 , 37(6) : 615-619.(YANG Tao , CHEN Xi , YANG Hong-wei , et al. Regional flood frequency analysis in Pearl River Delta region based on L-moments approach[J]. Journal of Hohai University : Natural Sciences , 2009 , 37(6) : 615-619.(in Chinese))

[3] 张国建 尹红美 李珂凌. 大气降水时空分布的数理统计与分析[J]. 水资源保护 , 2005 , 21(5) : 34-36.(ZHANG Guo-jian , YIN Hong-mei , LI Ke-ling. Analysis of spatial-temporal distribution of atmospheric precipitation using statistical method[J]. Water Resources Protection , 2005 , 21(5) : 34-36.(in Chinese))

[4] 梁忠民 戴昌军. 多项式正态变换水文频率分析方法的稳健性研究[J]. 水利水电科技进展 , 2005 , 25(3) : 15-17.(LIANG Zhong-min , DAI Chang-jun. Robustness of polynomial normal transforms in flood frequency analysis[J]. Advances in Science and Technology of Water Resources , 2005 , 25(3) : 15-17.(in Chinese))

[5] SL 44—2006 水利水电工程设计洪水计算规范[S].

[6] 能源部 水利部. 水利水电工程设计洪水计算手册[K]. 北京 : 中国水利水电出版社 , 1995 : 35.

[7] 谢迪全 余文公. 用频率组合法分析湘潭站设计洪水值[J]. 黑龙江水利科技 , 2006 , 34(3) : 1-2.(XIE Di-quan , YU Wen-gong. The analysis about design flood level of hydrometric station Xiang Tan by the frequency combination method[J]. Heilongjiang Science and Technology of Water Conservancy , 2006 , 34(3) : 1-2.(in Chinese))

[8] 姚杰 江涛. 频率组合法分析水库实际防洪能力的探索[J]. 内江科技 , 2007(8) : 104-117.(YAO Jie , JIANG Tao. The searching about analysing actual flood control capacity of reservoir by the frequency combination method[J]. Science and Technology of Neijiang , 2007(8) : 104-117.(in Chinese))

[9] 杨建青 朱杰. 关于水文随机变量频率组合方法的初探[J]. 水文 , 2008 , 28(1) : 61-63.(YANG Jian-qing , ZHU Jie. A tentative discussion on the method of calculating frequency composition of hydrologic random variables[J]. Journal of China Hydrology , 2008 , 28(1) : 61-63.(in Chinese))

[10] 刘燕 杜学挺 高博平 等. 频率组合法及其在干旱频率分析中的应用 : 以泾河流域干旱频率分析为例[J]. 干旱地区农业研究 , 2006 , 24(2) : 195-198.(LIU Yan , DU Xue-ting , GAO Bo-ping , et al. Frequency combination method and its application in drought frequency analysis : a case of Jinghe River[J]. Agricultural Research in the Arid Areas , 2006 , 24(2) : 195-198.(in Chinese))

[11] 王国安 李文家. 水文设计成果合理性评价[M]. 郑州 : 黄河水利出版社 , 2002 : 18-19.

[12] 秦毅. 河流交汇处设计洪峰流量的分析研究[J]. 西安理工大学学报 , 2001 , 17(1) : 57-61.(QIN Yi. An analytical study of designing peak flow at river confluence reaches[J]. Journal of Xi'an University of Technology , 2001 , 17(1) : 57-61.(in Chinese))