皮尔逊-III型分布曲线的变步长数值积分

刘仕平，王文川

(华北水利水电学院水利学院，河南 郑州 450011)

摘 要：通过对皮尔逊-III型曲线数值积分的研究，提出了一种新的积分方法——事先确定误差和变步长积分法。其主要思想是先将皮尔逊-III型分布曲线的广义积分转换为伽玛函数和常义积分，利用伽玛函数的递推公式和逼近公式计算出伽玛函数值，然后根据预定容许的相对误差和伽玛函数值确定绝对误差，再利用绝对误差确定基本步长，最后建立步长变动函数，使数值积分的步长按照抛物线规律自动增加，同时，充分考虑参数的适应性，以解决小参数收敛慢和大参数数据溢出问题。测试试验结果表明：事先确定误差免去了数值积分的试算过程，变步长积分能显著节省计算机的运行时间，且具有很宽的参数适应范围，在水利工程设计中具有较大的使用价值。

关键词：皮尔逊-III型曲线；数值积分；伽玛函数；误差估计；步长变动函数

中图分类号：TV12；TP391 **文献标识码：** A **文章编号：**1000-0852(2013)01-0018-03

在水利工程水文计算中，常采用皮尔逊-III型曲线作为径流、洪水及降雨量的理论频率曲线。适线法是确定皮尔逊-III型曲线参数的主要方法，适线过程中需要大量的计算。可以利用离均系数表插值计算[1]，也可以采用数值积分。数值积分一般采用龙贝格积分[2]，并利用事后估计误差法来确定精度[3]。此外，皮尔逊-III曲线的积分可以转换为不完全伽玛函数的计算，而不完全伽玛函数本身就有对应的级数表达式，因此也可将皮尔逊-III曲线的积分转化为级数的计算[4-5]。利用离均系数表计算不仅有计算误差，还受表格有效数字的限制，而且需要庞大的存储空间；数值积分法的特点正好相反，不需要庞大的存储空间，但积分计算比较复杂，而且在微机上运行程序有可能出现运行时间过长，还可能出现数据溢出问题，导致程序无法运行。本文通过对适线法的数值积分的研究，将皮尔逊-III型曲线积分转换为伽玛函数和不完全伽玛函数的计算问题，提出了事先估计误差的思路，给出了积分计算的基本步长和步长变动函数，免去了积分试算的过程，解决了适线计算数据溢出问题，大大节省程序的运行时间。

1 皮尔逊-III型曲线积分公式的变换

文献[6]给出的皮尔逊-III型分布计算公式如下

$$P=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\int_{x_p}^{\infty}(x-a_0)^{\alpha-1}e^{-\beta(x-a_0)}dx \tag{1}$$

式中：$\alpha=4/C_S^2$；$\beta=2/(\bar{x}C_VC_S)$；$a_0=\bar{x}(1-2C_V/C_S)$；$\bar{x}$ 为均值；C_V 为变差系数；C_S 为偏态系数。

由于式(1)积分无法找到原函数，需要用数值积分来计算，公式中分母为标准的伽玛函数，伽玛函数的求值也需要用数值积分来计算。为了统一格式，便于计算，进行积分换元变换。

令 $t=\beta(x-a_0)$，$dx=dt/\beta$，当 $x=x_p$ 时，$t=\beta(x_p-a_0)=u$，则式(1)变为

$$P=\frac{\int_u^{\infty}t^{\alpha-1}e^{-t}dt}{\Gamma(\alpha)}=\frac{\int_u^{\infty}t^{\alpha-1}e^{-t}dt}{\int_0^{\infty}t^{\alpha-1}e^{-t}dt}=\frac{I_2}{I} \tag{2}$$

收稿日期：2011-12-20
基金项目：河南省教育厅自然科学研究计划项目(2010B570002)
作者简介：刘仕平 (1955-)，男，河北唐山人，教授，博士，主要从事水利工程机械、水文计算等方面的研究工作。E-mail: liushiping@ncwu.edu.cn；
王文川(1976-)，男，河南鹿邑人，副教授，博士，主要从事水文水资源系统分析、优化建模等方面的研究。E-mail: wangwen1621@163.com

按照积分的可加性式(2)可以写成

$$P=\frac{\int_u^{\infty} t^{\alpha-1}e^{-t}dt}{\int_0^u t^{\alpha-1}e^{-t}dt+\int_u^{\infty} t^{\alpha-1}e^{-t}dt}=\frac{I_2}{I_1+I_2} \tag{3}$$

式中

$$\begin{cases} I=\int_0^{\infty} t^{\alpha-1}e^{-t}dt \\ I_1=\int_0^u t^{\alpha-1}e^{-t}dt \\ I_2=\int_u^{\infty} t^{\alpha-1}e^{-t}dt \end{cases} \tag{4}$$

应当指出,对于 $1<\alpha<2$ 情况利用式(4)积分收敛比较慢,利用分步积分法对上式进行变换可得

$$\begin{cases} I=\int_0^{\infty} t^{\alpha-1}e^{-t}dt=\frac{1}{\alpha}\int_0^{\infty} t^{\alpha}e^{-t}dt \\ I_1=\int_0^u t^{\alpha-1}e^{-t}dt=\frac{u^{\alpha}}{\alpha e^u}+\frac{1}{\alpha}\int_0^u t^{\alpha}e^{-t}dt \\ I_2=\int_u^{\infty} t^{\alpha-1}e^{-t}dt=-\frac{u^{\alpha}}{\alpha e^u}+\frac{1}{\alpha}\int_0^{\infty} t^{\alpha}e^{-t}dt \end{cases} \tag{5}$$

对于 $0<\alpha<1$ 时的情况,式(5)仍然收敛比较慢,再进行分步积分处理可得

$$\begin{cases} I=\frac{1}{\alpha(\alpha+1)}\int_0^{\infty} t^{\alpha+1}e^{-t}dt \\ I_1=\frac{u^{\alpha}}{\alpha e^u}+\frac{u^{\alpha+1}}{\alpha(\alpha+1)e^u}+\frac{\int_0^u t^{\alpha+1}e^{-t}dt}{\alpha(\alpha+1)} \\ I_2=-\frac{u^{\alpha}}{\alpha e^u}-\frac{u^{\alpha+1}}{\alpha(\alpha+1)e^u}+\frac{\int_0^{\infty} t^{\alpha+1}e^{-t}dt}{\alpha(\alpha+1)} \end{cases} \tag{6}$$

可以看出被积函数包括了幂函数和指数函数的运算,为了防止数据过大溢出,式(4)需要进一步变换为

$$\begin{cases} I=\int_0^{\infty} e^{(\alpha-1)\ln t-t}dt \\ I_1=\int_0^u e^{(\alpha-1)\ln t-t}dt \\ I_2=\int_u^{\infty} e^{(\alpha-1)\ln t-t}dt \end{cases} \tag{7}$$

上述各个式子中各包括有三个积分,I_1 是常义积分,而 I 和 I_2 是广义积分,且三个积分是相互关联的,只要知道其中的两个即可。在实际中,I 用于积分值的估计,I_1 用于实际的积分计算。

当 $0<\alpha\leq 1$ 时,用式(6)计算;

当 $1<\alpha\leq 2$ 时,用式(5)计算;

当 $\alpha>2$ 时,用式(7)计算。

2　容许误差和基本步长的确定

2.1　截断误差公式

由于被积函数的表达式是已知的,在区间$[a,b]$数值积分复化梯形公式的绝对误差为[3]

$$R_T=\left|\sum_{k=0}^{n-1}\frac{h^3}{12}f''(\eta_k)\right|=\left|\frac{(b-a)}{12}h^2f''(\eta)\right| \approx \frac{h^2}{12}\left|f'(b)-f'(a)\right| \tag{8}$$

式中:h 为积分步长。

对于区间$[0,u]$数值积分复化梯形公式的绝对误差为

$$R_T=\left|\frac{(u-0)}{12}h^2f''(\eta)\right|\approx\frac{h^2}{12}\left|f'(u)-f'(0)\right| \tag{9}$$

2.2　绝对误差的估计

由于参数 α 不同, 积分结果相差数量级比较大,用同一绝对误差控制是不合理的。只有用相对误差才能给人以定量和合理的感觉。给定相对误差后,再考虑积分的量值,从而确定绝对误差。

积分 $I=\Gamma(\alpha)$数值大小可以用递推公式来计算

$$\Gamma(\alpha)=\begin{cases} (\alpha-1)(\alpha-2)\cdots\Gamma(s) & (\alpha>2,\quad 1\leq s\leq 2) \\ \Gamma(s) & (1\leq\alpha\leq 2, 1\leq s\leq 2) \\ \Gamma(\alpha+1)/\alpha=\Gamma(s)/\alpha & (\alpha<1, 1\leq s\leq 2) \end{cases} \tag{10}$$

伽玛函数 $\Gamma(s)$在区间[1,2]内变化最为平缓,可用圆锥曲线来近似[7],伽玛函数 $\Gamma(s)$的近似表达式为

$$\Gamma(s)=\frac{-(Bs+E)+\sqrt{(Bs+E)^2-4C(As^2+Ds-1)}}{2C}\quad(1\leq s\leq 2) \tag{11}$$

式中:$A=-0.2266069$, $B=0.1494856$, $C=-0.3082491$, $D=0.5303577$, $E=0.8550017$。

这样,利用式(11)就可以确定 $\Gamma(s)$,再利用式(10)就可以确定积分 $I=\Gamma(\alpha)$。如果取定了相对误差(如 $\varepsilon=0.1\%$),则绝对误差为

$$R_T=\varepsilon I=\varepsilon\Gamma(\alpha) \tag{12}$$

2.3　基本步长的确定

积分 I_1 的积分区间为$[0,u]$,被积函数为 $f(t)=t^{\alpha-1}e^{-t}$,被积函数的一阶导数为 $f'(t)=(\alpha-1-t)t^{\alpha-2}e^{-t}$,由式(9)可得积分的基本步长为

$$h_0=\sqrt{\frac{12R_T}{f'(u)-f'(0)}}=\sqrt{\frac{12R_T}{|(\alpha-1-u)u^{\alpha-2}e^{-u}|}} \quad (13)$$

有了基本步长，利用梯形积分公式可以对 I_1 进行积分，从而可以求得概率 $P=I-I_1$。

3 变步长积分

由于被积函数 $f(t)=t^{\alpha-1}e^{-t}$ 变化幅度比较大，要加快积分运算还有潜力可以挖掘，可通过变步长积分实现。

3.1 步长函数的确定

由误差公式(9)可得，理论上的步长变动函数为

$$h_e=\sqrt{\frac{12R_T}{uf''(t)}}=\sqrt{\frac{12R_T}{u(\alpha-1-t)^2t^{\alpha-3}e^{-t}}}$$

$$=\sqrt{\frac{12R_T}{u}}\frac{e^{t/2}}{(\alpha-1-t)t^{(\alpha-3)/2}} \quad (14)$$

从式(14)可以看出，理论步长变动函数比较复杂，不太适合数值积分的计算，需要进行简化。

将上式展开为泰勒级数并取到二次项可得实际步长变动函数

$$h_p=C_0+C_1(t-\alpha)+C_2(t-\alpha)^2 \quad (15)$$

式中

$$C_0=-\sqrt{12R_T/u}\ e^{\alpha/2}\alpha^{-(\alpha-3)/2}$$

$$C_1=\sqrt{12R_T/u}\ (\alpha-3/2)e^{\alpha/2}\alpha^{-(\alpha-1)/2}$$

$$C_2=-\sqrt{12R_T/u}\ e^{\alpha/2}[(3/4)\alpha^{-(\alpha-3)/2}+(5/4)\alpha^{-(\alpha-1)/2}-(5/8)\alpha^{-(\alpha+1)/2}]$$

3.2 变步长积分计算

首先，取 $t_1=0$，步长为 $h_p=h_0$，并取 $t_2=t_1+h_p$，则第一个小区间的平均高度为 $f(t_1+h_p/2)=(t_1+h_p/2)^{\alpha-1}e^{-(t_1+h_p/2)}$，第一个小区间的面积为 $A_1=h_p\cdot f(t_1+h_p/2)$。

对于第二个以后的小区间，取 $t_1=t_2$，$h_p=h_p(t_1)=C_0+C_1(t_1-\alpha)+C_2(t_1-\alpha)^2$，$t_2=t_1+h_p$，小区间宽度为

$$h_p=\begin{cases}C_0+C_1(t_1-\alpha)+C_2(t_1-\alpha)^2 & (t_2\leq u)\\ u-t_1 & (t_2>u)\end{cases}$$

小区间的平均高度为 $f(t_1+h_p/2)=(t_1+h_p/2)^{\alpha-1}e^{-(t_1+h_p/2)}$，则小区间的面积为

$$A_i=h_p\cdot f(t_1+h_p/2)$$

将所有小区间的面积累加起来就可得到积分值 I_1。

4 运行结果分析

根据上述思路编制了数值积分计算程序，并在微机上作了大量实验，程序一运行立即出结果，反映了该算法的先进性。从 α 很小，到 α 很大都进行了实验，只要 α 不超过 171，积分值不超过 7.25×10^{306}，程序都不会由于数据溢出而停机，这个极限远远超过了水利工程中可能出现的数据，充分体现了程序的适应性。

5 结语

通过对皮尔逊-III型曲线变换和对变换后被积函数性态的研究，提出事先确定误差和步长的思路，给出了变步长的函数，免去了以往积分的试算过程。实践证明该方法具有以下特点：

(1)该方法在积分之前做好一切积分的准备，给出误差和步长，一次就可算出结果；

(2)该方法按照抛物线规律自动增加步长，可以大大节省计算机的运行时间；

(3) 较好地解决了小参数收敛慢的问题和大参数数据溢出问题，α 可达 171，积分值可达 7.25×10^{306}；

(4)该方法具有很强的适应性，能够满足适线过程需要。

参考文献：

[1] 任伯帜，许仕荣，王涛.皮尔逊-III 型分布统计参数的确定[J].中国给水排水,2001, 17(1): 40-42. (REN Bozhi, XU Shirong, WANG Tao. Determination of Pierson III distribution statistical parameters [J]. China Water & Waste Water, 2001, 17(1): 40-42. (in Chinese))

[2] 刘钧哲，马兴冠，傅金祥.皮尔逊-III 型曲线的快速算法[J]. 沈阳建筑工程学院学报，2004，20 (1)：60-62.(LIU Junzhe, MA Xingguan, FU Jinxiang. A quick algorithm of Pierson III distribution curve [J]. Journal of Shenyang Architecture and Civil Engineering University，2004, 20(1): 60-62. (in Chinese))

[3] 李庆扬，王能超，易大义.数值分析[M].武汉：华中理工大学出版社，2006.7.(LI Qingyang, WANG Nengchao, YI Dayi. Numerical Analysis [M].Wuhan: Huazhong University of Science & Technology Press，2006.7. (in Chinese))

[4] 李世才，吴戈堂，林莺.Γ 分布函数算法新解及其应用[J]. 水利学报，1999, (12): 70-76.(LI Shicai, WU Getang, LIN Ying. New interpretation of the algorithm and its application for Γ-distribution function [J]. Journal of Hydraulic Engineering, 1999, (12): 70-76. (in Chinese))

[5] M Aslam Chaudhry, Syed M Zubair. On a class of incomplete Gamma functions with applications [M]. CRC Press, 2002.

[6] 林冬益，孙保沭，林丽蓉.工程水文学[M].南京：河海大学出版社，2003.7.(LIN Dongyi, SUN Baoshu, LIN Lirong, Engineering Hydrology [M].Nanjing: Hohai University Press，2003.7. (in Chinese))

[7] 赵晓慎.伽玛函数的高精度二次曲线逼近[J]. 华北水利水电学院学报, 2006，27(3): 1-4. (ZHAO Xiaoshen.High precision approximation of Gamma function by using cone curve [J]. Journal of North China Institute of Water Conservancy and Hydroelectric Power, 2006, 27 (3): 1-4.(in Chinese))

(下转第 93 页)

Research on Characteristics of Water Resources in Nansihu Lake Basin

SUN Shuyun, ZHANG Guoyu

(*Bureau of Comprehensive Development, MWR, Beijing* 100053, *China*)

Abstract: According to the correlation between the runoff and landform characteristics, the Nansihe Lake basin has three types of runoff forming. This paper did a system investigation for all of them, and analyzed characteristics of the water resources balance. Special attention was paid to mechanisms of the natural runoff forming and impact of geographical and climatic factors on the river runoff. It also discussed the methods for estimation of the stream ecological runoff in the region.

Key words: water resources; natural runoff; geography; climate; ecological runoff

(上接第 20 页)

A Varying-Step Algorithm for Numerical Integration of Pierson III Distribution

LIU Shiping, WANG Wenchuan

(*Hydraulic Engineering Department, North China Institute of Water Conservancy and Hydroelectric Power, Zhengzhou* 450011, *China*)

Abstract: The numerical integration of Pierson Ⅲ distribution was investigated, a new integration algorithm of Pierson Ⅲ distribution was proposed. Firstly, the generalized integral of Pearson Ⅲ distribution curve was transformed into a Gamma function and an ordinary integral. By using the recurrence formula and approximate formula of gamma function, the approximate value of the Gamma function was obtained. Based on the allowable relative error and the Gamma function value, the absolute error was determined. According to the relation between step length and truncation error, the basic step length used for numerical integration was obtained. Finally, step varying function was established, so that the step length of the numerical integration can automatically increase in accordance with the parabolic law. Meanwhile, the adaptability of the parameters had also been considered fully. As a result, the two problems were successfully resolved. One problem is the slow convergence if parameter is very small, and another is the data overflow if parameter is very large. The results of experiments show that the pre-determined error method can avoid the pre-calculation process of numerical integration. The step varying integral can significantly save computer run time. The program possesses a very wide parameter adaptability, which has a great application value in the hydraulic engineering design.

Key words: Pierson Ⅲ distribution curve;numerical integration; Gamma function;error estimation; step varying function

(上接第 89 页)

Changes of Evaporation and Its Influencing Factors in Shiyang River Basin

LIU Ruirui, LU Baohong, XU Dan, ZHANG Jie, ZHAI Mengen, CHANG Na, LI Lihui

(*College of Hydrology and Water Resources*, *Hohai University*, *Nanjing* 210098, *China*)

Abstract: Mann-Kendall trend test, Wavelet analysis and Gray relation analysis were applied to determine the change characteristics and influencing factors of 6 stations over the Shiyang River Basin from 1959 to 2005. The results show that the average annual evaporation of 47 years has a decline trend over the basin. The time of abrupt change of the evaporation happened in the 1960s and late 1980s and around the year 2000. The evaporation of Summer and Spring declines clearly, which is the main contribution of annual evaporation decreasing. The spatial distribution of evaporation and change trend of evaporation both have increasing change from the westsouth to northeast, which is closely related to the terrain conditions and meteorological factors in the basin. The first order main periods of evaporation are 28 years and 20 years, which contain small periods from 10 years to 14 years. There is a close relation between air temperature, sunshine duration, wind speed of the main influencing factors and evaporation. The average temperature and sunshine hours have significant increasing trend, and average wind speed has remarkable decrease. The change of precipitation, relative humidity and vapour pressure are relatively weak. Remarkable decrease of average wind speed compensates the increased evaporation resulted from rise of air temperature and sunshine duration, which eventually reduce the evaporation.

Key words: Shiyang River Basin; characteristics of evaporation variation; wavelet analysis; gray relation analysis