

水文频率分析中参数估计最大熵法的梅林变换推导

张明¹, 张阳²

(1. 云南省水利厅, 云南 昆明 650021; 2. 云南省水利水电勘测设计研究院, 云南 昆明 650021)

摘要: 针对水文频率分析中最大熵法估计概率分布参数求解问题,探讨了用最大信息熵原理推导概率分布参数估计最大熵法的数学特性,分析了该方法与梅林变换之间的关系。从梅林变换出发,对各种水文频率分布参数估计的最大熵法进行了推导,说明了推导的具体步骤。用梅林变换推导最大熵法的过程更为简单、易懂。

关键词: 最大熵法; 水文频率分布; 梅林变换; 参数估计

中图分类号: P333.9

文献标识码: A

文章编号: 1672-643X(2015)04-0104-04

Mellin transform derivation of maximum entropy for parameter estimation in hydrologic frequency analysis

ZHANG Ming¹, ZHANG Yang²

(1. Water Conservancy Bureau of Yunnan Province, Kunming 650021, China; 2. Institute of Investigation, Design and Research of Water Resources & Hydropower Engineering of Yunnan Province, Kunming 650021, China)

Abstract: Principle of maximum entropy (POME) is a method for estimating the parameters of probability distribution. The paper discussed the method that used the maximum entropy principle to infer the mathematical properties of estimation parameters of probability distribution of POME. The relationships between POME and Mellin transform were analyzed. From Mellin transform, it derived POME of estimation parameters of 13 kinds of hydrological frequency distribution curve and illustrated the specific steps of derivation. The process of inferring POME by Mellin transform is simpler and easier to be understood.

Key words: principle of maximum entropy (POME); hydrologic frequency distribution; Mellin transform; parameter estimation

Singh V P^[1] 1998 年系统总结了运用最大信息熵原理推导水文频率分布参数估计方法,即最大熵法,并指出该算法在诸多水文频率分布应用中与极大似然法相当,计算还简单快捷。张明等^[2] 2014 年在上述推导过程中用梅林变换代替 Lagrange 乘法重新推导了指数 Γ 分布等 3 种分布参数估计的最大熵法,这一举措不仅提供了推导最大熵法的新途径,更为重要的是使推导过程更加直接、易懂。但是文献^[2] 极少涉及最大熵法与梅林变换之间的数学关系,因此阐明相关的理论是很有必要的。本文从最大信息熵原理入手分析了推导水文频率分布参数估计最大熵法的原理和过程,对算法中出现的数学特性做出归纳,从梅林变换的角度看待最大熵法并进行了推演。

1 最大熵法的原理与推导

Shannon 定义的信息熵为:

$$H(x) = - \int_R f(x) \ln f(x) dx \quad (1)$$

式中: $H(x)$ 为熵; x 为随机变量; $f(x)$ 为概率密度函数; R 为积分区间。

Jaynes E T^[3] 提出了最大信息熵原理,该原理认为:选择在一些约束条件下使熵达到最大值的概率分布是最佳概率分布。从某种意义上讲,概率分布与约束条件是相互等价的。在某些问题中,一般总是假设随机变量服从某一概率分布,在已知概率分布的条件下求取约束条件以及由约束条件研究熵的性质,则是反问题。

对于概率密度函数 $f(x)$ 代入式(1)分解有:

$$\int_R f(x) dx = 1 \quad (2)$$

$$\int_R G_i(x) f(x) dx = E[G_i(x)] \quad i = 1, 2, \dots, N \quad (3)$$

式中: $G_i(x)$ 为某一函数; $E[\cdot]$ 为均值; i 为自然

数; N 为最大自然数。式(2) 是正则化条件, 式(3) 是约束条件。

现用 Lagrange 乘子法求解。引入 Lagrange 不定乘子 $\lambda_0, \lambda_1, \dots, \lambda_N$ 使目标泛函 L , 即:

$$L = H(x) - \int_R (\lambda_0 - 1)f(x) dx - \int_R \sum_{i=1}^N \lambda_i G_i(x)f(x) dx \quad (4)$$

令 L 对 $f(x)$ 的变分为零, 即:

$$\partial L = - \int_R [\ln f(x) + \lambda_0 + \sum_{i=1}^N \lambda_i G_i(x)] \partial f(x) dx = 0 \quad (5)$$

最大熵分布概率密度函数的解析式:

$$f(x) = \exp[-\lambda_0 - \sum_{i=1}^N \lambda_i G_i(x)] \quad (6)$$

将式(6) 代入式(2) 得:

$$\lambda_0 = \ln \left\{ \int_R \exp[-\sum_{i=1}^N \lambda_i G_i(x)] dx \right\} \quad (7)$$

由式(7) 得:

$$\lambda_0 = g(\lambda_1, \lambda_2, \dots, \lambda_N) \quad (8)$$

式中: $g(\lambda_1, \lambda_2, \dots, \lambda_N)$ 为某一函数。将式(8) 代入式(6) 有:

$$f(x) = \exp[-g(\lambda_1, \lambda_2, \dots, \lambda_N) - \sum_{i=1}^N \lambda_i G_i(x)] \quad (9)$$

比较最大熵分布与概率密度函数, 可建立 Lagrange 乘子与概率密度函数参数之间的关系。还必须建立概率密度函数参数与约束条件之间的关系。将式(7) 对 λ_i 求偏导数, 并加入式(6)、(2)、(3), 有:

$$\lambda_0 / \lambda_i = -E[G_i(x)], \quad i = 1, 2, \dots, N \quad (10)$$

再将式(8) 对 λ_i 求偏导数, 为:

$$\lambda_0 / \lambda_i = q_i(\lambda_1, \lambda_2, \dots, \lambda_N) \quad (11)$$

式中: $q_i(\lambda_1, \lambda_2, \dots, \lambda_N)$ 为某一函数。式(10) 等于式(11) 有:

$$-E[G_i(x)] = q_i(\lambda_1, \lambda_2, \dots, \lambda_N) \quad (12)$$

式(12) 共有 N 个关系式, 可联立求解 $\lambda_1, \lambda_2, \dots, \lambda_N$ 。

若 $G_i(x)$ 还包括一些待定参数 $\omega_k (k = 1, 2, \dots, r)$ 时, 仅依靠式(12) 不可能同时解出, 还需导出补充关系式。将式(7) 对 λ_i 求二阶偏导数, 同时注意到式(6)、(2)、(3), 有:

$$\partial^2 \lambda_0 / \partial \lambda_i^2 = \text{Var}[G_i(x)] \quad (13)$$

式中: $\text{Var}[*]$ 为方差。式(8) 对 λ_i 求二阶偏导数, 得:

$$\partial^2 \lambda_0 / \partial \lambda_i^2 = z_i(\lambda_1, \lambda_2, \dots, \lambda_N), \quad i = 1, 2, \dots, N \quad (14)$$

式中: $z_i(\lambda_1, \lambda_2, \dots, \lambda_N)$ 为某一函数。式(13) 等于式(14) 有:

$$\text{Var}[G_i(x)] = z_i(\lambda_1, \lambda_2, \dots, \lambda_N) \quad (15)$$

式(15) 共有 N 个关系式。当 $r \leq N$ 时, 则在式(15) 中选择 r 个关系式, 联同式(12) 的 N 个关系式求解 $r + N$ 个未知数。当 $r > N$ 时, 此种情况很少, 还需导出另一补充关系式。

将式(7) 对 λ_i 和 $\lambda_j (i \neq j)$ 求偏导数, 并加入式(6)、(2)、(3), 有:

$$\partial^2 \lambda_0 / \partial \lambda_i \partial \lambda_j = \text{Cov}[G_i(x), G_j(x)] \quad (16)$$

式中: $\text{Cov}[*]$ 为协方差; j 为自然数。再对式(11) 求 $\lambda_j (j \neq i)$ 偏导数, 得

$$\partial^2 \lambda_0 / \partial \lambda_i \partial \lambda_j = u_{ij}(\lambda_1, \lambda_2, \dots, \lambda_N) \quad (17)$$

式中: $u_{ij}(\lambda_1, \lambda_2, \dots, \lambda_N)$ 为某一函数。式(16) 等于式(17) 有:

$$\text{Cov}[G_i(x), G_j(x)] = u_{ij}(\lambda_1, \lambda_2, \dots, \lambda_N) \quad (18)$$

在式(18) 中选择 $r - N$ 个关系式, 与式(12)、(15) 的 $2N$ 个关系式一起, 联立求解 $r + N$ 个未知数。

从上述推导过程可以看出, 通过 Lagrange 乘子分别构建了与参数和约束条件之间的关系, 消去 Lagrange 乘子后, 建立了参数与统计特征之间的关系。

需要注意的是, 最大熵法不适用所有的概率密度函数, 但是对常用的水文频率分布均适用。

2 基于梅林变换的最大熵法

在最大熵法的推导中, 重要结果为式(12)、(15)、(18), 实质上归结为建立了在信息熵约束条件下概率密度函数参数与统计特征的关系, 这种关系是概率密度函数自身性质决定的, 与在信息熵约束条件下直接用概率统计方法建立的关系是等价的。根据式(1) 的性质, 概率密度函数 $f(x)$ 取对数后, 使 $G_i(x)$ 表现出一一般函数或者对数函数特点, 约束条件也表现出传统矩或者对数矩的性质。求解此问题, 能用一种数学工具解决是最好的, 梅林变换具有如此特性, 为此可以用梅林变换推导最大熵法。

梅林变换定义为:

$$\Phi_x(s) = M[f_x; s] = \int_0^{\infty} x^{s-1} f(x) dx \quad (19)$$

式中: $\Phi_x(s)$ 或者 $M[\cdot]$ 为梅林变换; s 为变量。

式(19)中当 $s = n + 1$ 时,可以得出传统矩(第一类矩)为:

$$m_n = M[f_x; s] |_{s=n+1} \quad (20)$$

式中: n 为自然数; m_n 为第 n 阶原点矩。

根据梅林变换的性质可得对数矩(第二类矩)为^[4]:

$$\begin{aligned} \tilde{m}_n &= M[(\ln x)^n f_x; s] |_{s=1} \\ &= \int_0^{\infty} x^{s-1} (\ln x)^n f(x) dx |_{s=1} = \frac{d^n}{ds^n} \Phi_x(s) |_{s=1} \end{aligned} \quad (21)$$

式中: $\tilde{m}_n$ 为第 n 阶对数矩。令

$$\Psi_x(s) = \ln \Phi_x(s) \quad (22)$$

可导出对数累积量,有:

$$\tilde{k}_n = \frac{d^n}{ds^n} \Psi_x(s) |_{s=1} \quad (23)$$

式中: $\tilde{k}_n$ 为第 n 阶累积量。对数矩和对数累积量有下列关系:

$$\begin{aligned} \tilde{k}_1 &= \tilde{m}_1 \\ \tilde{k}_2 &= \tilde{m}_2 - \tilde{m}_1^2 \\ \tilde{k}_3 &= \tilde{m}_3 - 3\tilde{m}_1\tilde{m}_2 + 2\tilde{m}_1^3 \\ \tilde{k}_4 &= \tilde{m}_4 - 4\tilde{m}_1\tilde{m}_3 + 6\tilde{m}_1^2\tilde{m}_2 - 3\tilde{m}_1^4 \end{aligned} \quad (24)$$

应用梅林变换推导概率密度函数参数估计最大熵法的步骤如下:

(1) 由已知的概率密度函数分布,根据最大信息熵原理反求约束条件,即式(3);

(2) 确定概率密度函数分布参数的个数,依照参数个数选择参数估计统计特征方程式,保证统计特征方程式个数与参数个数相等,取舍规则:首先选择信息熵约束条件下的均值特征方程式,即式(12),均值特征方程式不够参数个数时再选择方差特征方程式,即式(15),再不够时还需要选择协方差特征方程式,即式(18);

(3) 分析判断其约束条件的边界是否符合梅林变换的要求,对不符合的,做数学处理达到要求;

(4) 用梅林变换推导已选取的统计特征方程式,建立参数与统计特征的关系,组成参数估计的统计特征方程组,即最大熵法。

3 典型水文频率分布线型参数估计最大熵法的梅林变换推导

为了实证用梅林变换推导最大熵法的有效性和

合理性,本文选取了 5 大类 13 种水文频率分布线型^[5]进行参数估计最大熵法的推导,第一类为正态分布类,包括两参数对数正态分布、正态分布;第二类为极值分布类,包括极值 I 型分布、对数极值 I 型分布、广义极值分布、威布尔分布;第三类为 Γ 分布类,包括指数分布、两参数 Γ 分布、P-III 型分布、对数 P-III 型分布;第四类逻辑斯特分布,包括二参数对数逻辑斯特分布;第五类为帕雷托分布类,包括两参数帕雷托分布、两参数广义帕雷托分布。有关线型、梅林变换方程、参数估计方程组等列于表 1。

下面以表 1 中所列的 P-III 型分布参数估计最大熵法推导为例说明。

第一步:将 P-III 型分布代入式(1),有:

$$\begin{aligned} H(x) &= [b \ln a + \ln \Gamma(b)] \int_c^{\infty} f(x) dx + \frac{1}{a} \int_c^{\infty} (x-c) \cdot \\ &\quad f(x) dx - (b-1) \int_c^{\infty} \ln(x-c) f(x) dx \end{aligned} \quad (25)$$

得约束条件为:

$$E[x-c] = \int_c^{\infty} (x-c) f(x) dx \quad (26)$$

$$E[\ln(x-c)] = \int_c^{\infty} \ln(x-c) f(x) dx \quad (27)$$

第二步:P-III 型分布有 3 个参数,分别为 a, b, c ,按照规则分别选取约束条件的均值特征方程式 $E[x-c]$ 、 $E[\ln(x-c)]$ 和补充关系式中的方差特征方程式 $\text{Var}[x-c]$ 等 3 个方程式。

第三步:式(26)、(27)可以看出,边界条件不符合式(19)的要求,做变换,令

$$y = x - c \quad (28)$$

$$\text{有: } E[y] = \int_0^{\infty} y f(y) dy \quad (29)$$

$$E[\ln y] = \int_0^{\infty} (\ln y) f(y) dy \quad (30)$$

$$\text{Var}[y] = \int_0^{\infty} y^2 f(y) dy - \left[\int_0^{\infty} y f(y) dy \right]^2 \quad (31)$$

第四步:将 $f(y)$ 代入式(19),用梅林变换推导得:

$$\Phi_y(s) = M[f_y; s] = \frac{a^{s-1}}{\Gamma(b)} \Gamma(s+b-1) \quad (32)$$

式(32)中 s 分别取 2、3,推导有:

$$E[y] = m_1 = ab \quad (33)$$

$$\text{Var}[y] = m_2 - m_1^2 = a^2 b \quad (34)$$

将式(32)代入式(21),取 $n = 1$,有:

$$E[\ln y] = \tilde{m}_1 = \Psi(b) + \ln a \quad (35)$$

将式(28) 代入式(33)、(34)、(35) 逆变换得： $E[\ln(x-c)] = \Psi(b) + \ln a$ (38)

$E[x] = ab + c$ (36)

$\text{Var}[x] = a^2b$ (37)

结果表明,基于梅林变换的最大熵法推导既简单又易懂。

表 1 不同分布线型概率密度函数、梅林变换方程及参数估计方程组

线型		概率密度函数 $f(x)$	梅林变换方程 $M[f_y;s]$	参数估计方程组	数学变换 y
正态分布类	两参数对数正态分布	$\frac{1}{xb\sqrt{2\pi}}\exp[-\frac{(\ln x-a)^2}{2b^2}]$	$\exp\left\{\frac{[a+(s-1)b^2]^2-a^2}{2b^2}\right\}$	$E[\ln y] = a,$ $E[(\ln y)^2] = b^2+a^2$	x
	正态分布	$\frac{1}{b\sqrt{2\pi}}\exp[-\frac{(x-a)^2}{2b^2}]$	$\exp\left\{\frac{[a+(s-1)b^2]^2-a^2}{2b^2}\right\}$	$E[\ln y] = a,$ $E[(\ln y)^2] = b^2+a^2$	e^x
极值分布类	极值Ⅰ型分布	$a\exp\{-a(x-b)-\exp[-a(x-b)]\}$	$\Gamma(s)$	$E[y] = 1,$ $E[\ln y] = -0.5772$	$\exp[-a(x-b)]$
	对数极值Ⅰ型分布	$\frac{a}{x}\exp\{-a(\ln x-b)-\exp[-a(\ln x-b)]\}$	$\Gamma(s)$	$E[y] = 1,$ $E[\ln y] = -0.5772$	$\exp[-a(\ln x-b)]$
	广义极值分布	$\frac{1}{a}[1-\frac{b}{a}(x-c)]^{\frac{1-b}{b}}\cdot\exp\{-[1-\frac{b}{a}(x-c)]^{\frac{1}{b}}\}$	$\Gamma(s)$	$E[y] = 1,$ $E[\ln y] = -0.5772,$ $\text{Var}[y] = 1$	$[1-\frac{b}{a}(x-c)]^{\frac{1}{b}}$
	威布尔分布	$\frac{a}{b}(\frac{x}{b})^{a-1}\exp[-(\frac{x}{b})^a]$	$\Gamma(s)$	$E[y] = 1,$ $E[\ln y] = -0.5772$	$(\frac{x}{b})^a$
Γ 分布类	指数分布	$a\exp[-a(x-c)]$	$\Gamma(s)$	$E[y] = 1,$ $E[\ln y] = -0.5772$	$a(x-c)$
	两参数 Γ 分布	$\frac{1}{a\Gamma(b)}(\frac{x}{a})^{b-1}\exp(-\frac{x}{a})$	$\frac{a^{s-1}\Gamma(s+b-1)}{\Gamma(b)}$	$E[y] = ab,$ $E[\ln y] = \Psi(b)+a$	x
	$P-Ⅲ$ 分布	$\frac{1}{a\Gamma(b)}(\frac{x-c}{a})^{b-1}\exp[-(\frac{x-c}{a})]$	$\frac{a^{s-1}\Gamma(s+b-1)}{\Gamma(b)}$	$E[y] = ab,$ $E[\ln y] = \Psi(b)+a$ $\text{Var}[y] = a^2b$	$x-c$
	对数 $P-Ⅲ$ 分布	$\frac{1}{ax\Gamma(b)}(\frac{\ln x-c}{a})^{b-1}\cdot\exp[-(\frac{\ln x-c}{a})]$	$\frac{a^{s-1}\Gamma(s+b-1)}{\Gamma(b)}$	$E[y] = ab,$ $E[\ln y] = \Psi(b)+a,$ $\text{Var}[y] = a^2b$	$\ln x-c$
逻辑斯特分布类	两参数对数 Logistic 分布	$\frac{b(\frac{x}{a})^{b-1}}{a[1+(\frac{x}{a})^b]^2}$	$\begin{cases} \frac{\Gamma(s)\Gamma(2-s)}{\Gamma(2)}, \\ -\frac{1}{s-2}, s-2<0 \end{cases}$	$\begin{cases} E[\ln y] = 0, \\ E[\ln y] = 1 \end{cases}$	$\begin{cases} \frac{x}{a}, \\ 1+(\frac{x}{a})^b \end{cases}$
帕雷托分布类	两参数 Pareto 分布	ba^bx^{-b-1}	$-\frac{ba^{s-1}}{s-b-1}, s-b-1<0$	$E[\ln y] = \frac{1}{b}+\ln a,$ $\text{Var}[\ln y] = \frac{1}{b^2}$	x
	两参数广义 Pareto 分布	$\frac{1}{b}(1-\frac{ax}{b})^{\frac{1}{a}-1}, a<0$	$\frac{1}{1+(s-1)a}, \frac{1}{a}+s-1<0$	$E[\ln y] = -a,$ $\text{Var}[\ln y] = a^2$	$1-\frac{ax}{b}$

5、17 及 26 a; 冬季气候对应 5、9 及 29 a 的周期。

[参考文献]:

- [1] 丁一汇,任国玉,石广玉,等. 气候变化国家评估报告(1):中国气候变化的历史和未来趋势[J]. 气候变化研究进展,2006,2(1):3-8.
- [2] 任朝霞,杨达源. 西北干旱区近 50 年气候变化特征与趋势[J]. 地球科学与环境学报,2007,29(1):99-102.
- [3] 姚俊强,杨青,陈亚宁,等. 西北干旱区气候变化及其对生态环境影响[J]. 生态学杂志,2013,32(5):1283-1291.
- [4] 黄小燕,张明军,贾文雄,等. 中国西北地区地表干湿变化及影响因素[J]. 水科学进展,2011,22(2):151-159.
- [5] 马柱国,符淙斌. 中国北方干旱区地表湿润状况的趋势分析[J]. 气象学报,2000,59(6):737-746.
- [6] 施雅风,沈永平,李栋梁,等. 中国西北气候由暖干向暖湿转型的特征和趋势探讨[J]. 第四纪研究,2003,23(2):152-164.
- [7] 张永,陈发虎,勾晓华,等. 中国西北地区季节间干湿变化的时空分布:基于 PDSI 数据[J]. 地理学报,2007,62(11):1142-1152.
- [8] 郑红星,刘静. 东北地区近 40 年干燥指数变化趋势及其气候敏感性[J]. 地理研究,2011,30(10):1765-1774.
- [9] 王鹏祥. 西北地区干湿气候转型研究[D]. 南京:南京信

息工程大学,2006.

- [10] 雷江群,刘登峰,黄强. 渭河流域气候变化及干湿状况时空分布分析[J]. 西北农林科技大学学报(自然科学版),2015,43(3):175-181.
- [11] 张晓晓,张钰,徐浩杰. 1950~2005 年大通河流域径流变化特征及影响因素[J]. 水文,2013,33(6):90-96.
- [12] Allen R G, Pereira L S, Raes D, et al. Crop evapotranspiration: Guidelines for computing crop water requirements [R]. Rome: FAO, 1998.
- [13] 童成立,张文菊,汤阳,等. 逐日太阳辐射的模拟计算[J]. 中国农业气象,2005,26(3):165-169.
- [14] 李修仓. 中国典型流域实际蒸散发的时空变异研究[D]. 南京:南京信息工程大学,2013.
- [15] 马宁,王乃昂,王鹏龙,等. 黑河流域参考蒸散量的时空变化特征及影响因素的定量分析[J]. 自然资源学报,2012,27(6):975-989.
- [16] 刘超,秦毅,邓娜. 黄河上游主要干支流近期降水、径流统计特征变化分析[J]. 水土保持学报,2004,18(1):96-99.
- [17] 王文圣,丁晶,金菊良. 随机水文学[M]. 北京:中国水利水电出版社,2008.
- [18] 邵晓梅,许月卿,严昌荣. 黄河流域降水序列变化的小波分析[J]. 北京大学学报(自然科学版),2006,42(4):503-509.

(上接第 107 页)

4 结 语

(1) 最大熵法是以最大信息熵原理为基础,在已知概率密度函数前提下,反求约束条件,通过 Lagrange 乘子分别构建与参数和约束条件之间的关系,消去 Lagrange 乘子后,建立参数与统计特征之间的关系。

(2) 梅林变换可以作为推导最大熵法的一种新方法,阐述了推导步骤和注意问题。

(3) 对 5 大类 13 种水文频率分布线型进行了参数估计最大熵法的推导,结果证明,该推导过程既简单又易懂。

参考文献:

- [1] Singh V P. Entropy-based parameter estimation in hydrology [M]. Boston: Kluwer Academic Publishers, 1998.
- [2] 张明,柏绍光,张阳,等. 用 Mellin 变换推导几种分布参数估计的最大熵法[J]. 水电能源科学,2014,32(6):16-18.
- [3] Jaynes E T. Information theory and statistical mechanics [J]. Physical Review, 1957,106(4):620-630.
- [4] 时公涛,高贵,周晓光,等. 基于 Mellin 变换的 G0 分布参数估计方法[J]. 自然科学进展,2009,19(6):677-690.
- [5] 肖可以,宋松柏. 最大熵原理在水文频率参数估计中的应用[J]. 西北农林科技大学学报(自然科学版),2010,38(2):197-205.