# Copulas函数及其在水文中的应用

宋松柏　蔡焕杰　金菊良　康　艳　编著

科学出版社

# Copulas 函数及其在水文中的应用

宋松柏　蔡焕杰　金菊良　康　艳　编著

国家自然科学基金(50879070、51179160、50579065 和 51079037)
高等学校博士学科点专项科研基金(20110204110017)
“西北农林科技大学创新团队支持计划”
水利部重大基建前期项目“全国干旱区划及旱灾风险评估研究”资助

科 学 出 版 社
北 京

## 内 容 简 介

本书吸收了国内外关于 Copulas 函数理论和应用的前沿研究进展，系统地总结了作者近年来和相关科研课题的研究结果. 全书注重循序渐进，理论联系实际，系统研究了 Copulas 函数应用的关键技术问题. 主要包括单变量概率分布、多变量概率分布、Copulas 函数及其特性、对称 Archimedean Copulas 函数及其应用、非对称 Archimedean Copulas 函数及其应用、Plackett Copulas 函数及其应用、meta-elliptic Copulas 函数及其应用、Pair-Copulas 函数及其应用和一些实用计算算法，书中并附有大量的计算实例.

本书可供水文学及水资源、农业水土工程、水利水电工程、环境科学、气象科学、经济管理和统计等专业的高年级本科生、研究生以及相关领域教学、科研与工程技术人员使用.

图书在版编目(CIP)数据

Copulas 函数及其在水文中的应用/宋松柏 等编著. —北京：科学出版社，2012

ISBN 978-7-03-034095-5

Ⅰ. ①C… Ⅱ. ①宋… Ⅲ. ①时间序列分析-应用-水文学-研究 Ⅳ.①P33

中国版本图书馆 CIP 数据核字(2012) 第 075201 号

责任编辑：汪 操 彭胜潮／责任校对：包志虹

责任印制：钱玉芬／封面设计：耕者设计工作室

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

中国科学院印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2012 年 5 月第 一 版 开本：B5(720 × 1000)

2012 年 5 月第一次印刷 印张：24 1/4

字数：472 000

**定价：88.00 元**

(如有印装质量问题，我社负责调换)

# 前　　言

水文频率分析是水利工程规划设计的主要依据, 已有一百多年的研究和实践历史. 单变量频率分布在洪水、干旱和暴雨等复杂水文事件频率分析中具有一定的局限性. 其原因是这类复杂水文事件具有多个特征属性, 且特征属性间存在一定的相依性, 单变量频率分析不能为水利工程规划设计提供足够的信息. 因此, 多变量水文频率分析研究开始受到重视. 现有传统多变量联合分布是建立在一定的假定基础上：① 变量边际分布必须为同一类型的分布; ② 联合分布一般采用正态分布和经过数学转换成为正态分布; ③ 一些非正态联合分布模型求解复杂, 构造高维分布困难, 其应用受到限制. 自 1959 年 Sklar 形成了现代 Copulas 理论以来, Copulas 函数被引入许多领域的多变量联合分布分析, 也被开始应用于洪水、暴雨等极值水文事件的联合概率分布分析, 成为水文计算领域的一个研究热点和新型研究课题. 研究表明, Copulas 函数能有效地描述水文事件的内在规律和特征属性之间的相互关系, 能够克服传统多变量联合分布构建的缺点.

基于 Copulas 函数的多变量水文频率计算是一个崭新的研究领域, 涉及高等数学、概率论与数理统计、水文学、数值计算、优化计算等学科的交叉和渗透, 面临一系列亟待解决的科学问题. Copulas 在水文中的应用综述已有不少文献进行了归纳总结, 本书不再叙述. 作者结合在美国 Texas A & M 大学期间同国际著名学者 Vijay P. Sigh 教授的一些合作研究, 在国家自然科学基金 (50879070、51179160、50579065 和 51079037)、高等学校博士学科点专项科研基金 (20110204110017)、“西北农林科技大学创新团队支持计划”、水利部重大基建前期项目 “全国干旱区划及旱灾风险评估研究”、水利部公益性行业科研专项经费项目 (201001043) “抗旱能力评价及干旱风险管理研究” 的支持下, 吸收了大量国外同行的理论和方法, 系统地探索了对称 Archimedean Copulas、非对称 Archimedean Copulas、Plackett Copula、meta-elliptic Copulas 和 Pair-Copulas 等在我国西北地区典型流域的研究, 推导了一些 Copulas 函数的应用计算公式和计算机实现方法, 主要包括以下几个方面：

(1) 单变量概率分布.

单变量水文序列频率分析的主要依据是随机单变量的概率分布. 这类分布的取值范围一般为 $(-\infty,\infty)$ 或 $(a,\infty)$, 而实际水文变量的实测值不可能为无穷大值, 因此, 现有随机单变量的概率分布实际上是对水文变量分布的逼近, 频率拟合在 25%～75%的频率段一般能够取得较好的拟合效果, 而频率曲线的上尾或下尾部 (频率曲线的两端外延部分) 拟合出现较大的偏差. 我国设计洪水频率曲线的适线方法

有目估法和优化适线法, 适线选用离 (残) 差平方和最小准则 (OLS)、离 (残) 差绝对值和最小准则 (ABS)、相对离 (残) 差平方和最小准则 (WLS), 其参数计算归结为非线性优化问题. 不管是采用高斯–牛顿法进行迭代计算, 还是采用其他智能优化求解, 最终结果大都是以计算者按照上述评定准则, 目估或评价整体曲线或着重考虑曲线的上尾、下尾段的适线效果. 因而, 同一概率分布的具体实际应用, 不同经验水平的计算者有不同的参数计算结果, 评定结果主观性较强. 按照概率与数理统计原理, 水文频率分布的选用和估算参数均需要通过分布函数的拟合度检验后, 才能用于水文设计值和重现期的计算. 除 $\chi^2$ 检验外, 一些教科书列举了 Kolmogorov-Smirnov (K-S) 检验, 也给出了一些显著性水平 $\alpha$ 下的临界值, 但是, 所列 K-S 检验法的临界值是适用小样本检验, 其临界值是根据蒙特卡罗法按照选定分布进行大量抽样, 通过统计量随机试验获得, 不一定适用所有分布的临界值选取. 所以, 单变量频率分布选定与参数估算直接影响计算结果的合理性, 仅靠计算者的目估适线是不够的, 必须通过统计意义上的拟合度检验. 为了便于读者使用, 本书还列举了 Cramer-von Mises (C-M) $W_n^2$、Aderson-Darling (A-D)$A_n^2$、修正 Watson $U_n^2$ 和 Liao-Shimokawa $L_n$ 的检验方法.

(2) 多变量间的相依性度量.

Copulas 函数能够描述多变量间的相依性. 根据 Sklar 理论, 设 $X_1,\cdots,X_d$ 为具有边际分布 $F_{X_1},\cdots,F_{X_d}$ 的随机变量, 联合分布为 $F(x_1,x_2,\cdots,x_d)$, $u_1=F_{X_1}(x_1),\cdots,u_d=F_{X_d}(x_d)$. 若 $X_1,\cdots,X_d$ 相互独立或它们之间的相依性可以忽略不计, 则有联合分布 $F(x_1,x_2,\cdots,x_d)=C(u_1,\cdots,u_d)=\prod\limits_{i=1}^{d}F_{X_i}(x_i)=\prod\limits_{i=1}^{d}u_i$; 若 $X_1,\cdots,X_d$ 间的相依性不能忽略, 则联合分布为 $F(x_1,x_2,\cdots,x_d)=C[F_{X_1}(x_1),F_{X_2}(x_2),\cdots,F_{X_d}(x_d)]=C(u_1,u_2,\cdots,u_d)$. 所以, 对于相互独立的多个变量, 其联合分布等于它们各自边际分布的乘积, 而无需利用 Copulas 函数构建它们的联合分布. 本书详细地介绍了 Pearson 古典相关系数 $r_n$、Spearman 秩相关系数 $\rho_n$、Kendall $\tau$ 系数、Chi 图和 K 图五种变量间的相依性度量方法及其有关推导过程, 统一了一些计算公式, 便于读者理解多变量相依性度量的基本原理.

(3) Copulas 函数的参数估算.

对于给定的多变量水文序列和选定的 Copulas 函数, 除对称 Archimedean Copulas 可采用 Spearman 秩相关系数 $\rho_n$ 和 Kendall $\tau$ 系数与 Copulas 函数参数关系式计算外, Copulas 函数参数估算的通用方法是极大似然法, 包括精确极大似然法 (一阶段法)、边际函数推断法 (两阶段法) 和半参数法. 极大似然法的基本思路是在 Copulas 函数参数取值范围内构建最大似然函数值, 这类计算可归结为似然函数最大优化问题. 根据极值原理, 似然函数分别对 Copulas 函数各参数求偏导数, 并令这些偏导数等于零组成非线性方程 (或非线性方程组), 应用数值方法求解非线性

方程 (或非线性方程组) 即可得到 Copulas 函数各参数. 但是, 对于一参数 Copulas 函数, 非线性方程容易求解, 而两参数以上的 Copulas, 则非线性方程组求解困难. 因此, 可采用非线性规划和其他智能优化求解似然函数最大优化问题的方法来获得 Copulas 函数参数.

(4) Copulas 随机数模拟.

Copulas 模拟是选定多元联合分布拟合度检验的重要步骤. 对于对称 Archimedean Copula, 可采用 Marshal-Olkin 法进行模拟, 除此之外, Copulas 的通用模拟方法可采用 CPI Rosenblatt 转换法 (条件逆函数法). CPI Rosenblatt 转换法需要通过条件 $C(u_i|u_1, u_2, \cdots, u_{i-1})$ 来模拟 $u_i$. 为了便于使用, 本书根据概率原理, 详细推导了多变量条件分布、条件 Copulas 和 CPI Rosenblatt 转换法以及常用几类条件 Copulas 的计算公式.

(5) 最佳 Copula 函数的选择.

对称多元 Archimedean Copulas 也称为可交换 Archimedean Copulas (exchangeable Archimedean Copulas), 其特点是具有一个生成函数 $\varphi$ 和一个参数 $\theta$ 描述变量间的正相依性, 计算简单. 由于对称 Archimedean Copulas 要求变量为对称相依, 实际中, 水文变量间可能存在正、负相依性或独立, 它们也可能是不对称的样本. 因此, 对称 Archimedean Copulas 不一定是三变量以上水文序列联合概率分布拟合的最佳函数, 其他非对称类 Copulas 在水文中的应用研究是非常必要的.

Copulas 模型选择虽然有均方根误差法、AIC 法和 BIC 法. 但这三种方法仅是依据实测数据经验概率与理论联合概率从整体上评价 Copulas 模型的拟合效果. 实际中, 经常出现当一个随机变量 $X$ 取较大 (或较小) 值时, 它对另一个随机变量 $Y$ 的取值是否有影响? 这类问题归结为 Copulas 函数的尾部相关性问题. 因此, Copulas 模型选择除应用均方根误差法、AIC 法和 BIC 法外, 还应考虑选取的 Copulas 函数是否能够反映极值事件的相关性, 对于多变量洪水频率分析具有重要的研究意义.

与单变量概率分布拟合类似, 最终选择的 Copulas 也要通过拟合度检验. Copulas 拟合度检验有经验 Copulas 检验法、Kendall 转换检验法和 CPI Rosenblatt 转换检验法. CPI Rosenblatt 转换检验法是一种通用的 Copulas 拟合度检验法, 其关键技术是 Copulas 的模拟和检验统计量在不同显著性水平 $\alpha$ 下临界值的推求. 对于这类不同显著性水平 $\alpha$ 下临界值, 没有现成的表格和计算公式查用, 需要模拟大量的 Copulas 随机数, 构建某一统计量, 通过提取 $(1-\alpha)$%的分位数确定临界值.

全书由单变量概率分布、多变量概率分布、Copulas 函数理论和基于 Copulas 的多变量水文频率计算等部分组成, 共分为 10 章. 第 1 章主要介绍几种常见单变量分布、偏态非对称分布和它们的参数估算方法. 第 2 章主要介绍多变量联合概率分布、经验频率、条件概率和重现期的计算, 在此基础上, 叙述几种常用多变量联合

分布. 第 3 章详细总结了 Copulas 的定义、特性和构造方法. 第 4~10 章详细叙述对称 Archimedean Copulas、非对称 Archimedean Copulas、Plackett Copulas、meta-elliptic Copulas 和 Pair-Copulas 函数的理论方法及其应用.

本书编写分工如下：全书由宋松柏、蔡焕杰和金菊良统稿. 第 1 章由宋松柏、李宏伟、袁超、李扬编写; 第 2 章由宋松柏、蔡焕杰、成静清编写; 第 3 章由宋松柏、金菊良、谢萍萍编写; 第 4 章由宋松柏、张雨、李雪月编写; 第 5 章由宋松柏、于艺、曾智、王剑峰编写; 第 6 章由宋松柏、金菊良、肖可以编写; 第 7 章由宋松柏、蔡焕杰、赵丽娜编写; 第 8 章由康艳、宋松柏、刘丹丹编写, 第 9 章由宋松柏、侯芸芸、刘斌、陈子全编写; 第 10 章由宋松柏、张雨、马明卫、于艺编写. 在本书编写过程中, 研究生王红兰、郭成、原秀红、肖玲和王俊珍等先后帮助搜集资料、整理和校对文稿, 在此向他们表示衷心的感谢!

感谢 Vijay P. Sigh 教授的悉心指导和鼓励, 也感谢西北农林科技大学旱区农业水土工程教育部重点实验室、农业部旱区农业节水工程重点开放实验室和合肥工业大学土木与水利工程学院水资源与环境系统工程研究所的大力支持. 书中参考了大量国内外学者的研究成果和文献, 大部分在文中和参考文献中列出, 在此一并致谢. 作者还感谢科学出版社编辑的文字润色和辛勤劳动, 借此机会对他们的辛勤工作和大力支持表示深深的谢意.

由于 Copulas 函数是近年来发展起来的新型分支学科, 其相关理论方法还在进一步研究和发展中. 作者深信：在不久的将来, Copulas 函数许多研究成果将会不断地得到完善和充实.

由于作者水平有限, 书中计算公式较多, 虽经多次核对和修改, 难免存在一定的不足之处, 敬请有关专家学者和读者批评指正, 以利本书今后进一步修改和完善.

作　者

2011 年 12 月 28 日

# 目　　录

# 第 1 章　单变量概率分布

现行水文频率分析主要采用单变量概率分布. 本章在回顾几种常见单变量分布的基础上, 介绍几种偏态非对称分布和它们的参数估算方法. 本章约定水文序列为 $x_1, x_2, \cdots, x_n$, 其中, $n$ 为样本长度; $T$ 为重现期, 相应的不超越概率为 $F = 1 - 1/T$, $\hat{x}_T$ 为对应的设计估计值; $\beta_r$ 为 $r$ 阶概率权重矩; $b_r$ 为 $\beta_r$ 的样本式计算矩; $l_r$ 为 $r$ 阶线性矩 $\lambda_r$ 的样本计算矩; $t = l_2/l_1, t_r = l_r/l_2, r \geqslant 3$; $m_1'$ 为样本均值; $m_2$ 为样本方差; $m_r'$ 为样本的 $r$ 阶原点矩; $m_r$ 为样本的 $r$ 阶中心矩; $C_{\mathrm{v}}$ 为分布离势系数 (变差系数); $C_{\mathrm{s}}$ 为分布偏态系数; $F(x)$ 和 $f(x)$ 分别为随机变量的分布函数和密度函数.

## 1.1　几种常见的单变量分布和参数估算方法

### 1.1.1　正态分布类

#### 1.1.1.1　正态分布

1) 概率密度

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty \tag{1.1}$$

式中, $\mu$ 为随机变量的均值, 也称位置参数; $\sigma$ 为标准差, 也称形状参数. 正态分布以 $x = \mu$ 为对称轴左右完全对称.

2) 参数估计方法

(1) 矩法

$$\hat{\mu} = m_1' \tag{1.2}$$

$$\hat{\sigma}^2 = m_2 \tag{1.3}$$

(2) 极大似然法

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = m_1' \tag{1.4}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \hat{\mu})^2 = m_2 \tag{1.5}$$

由式 (1.2)~(1.5) 可以看出, 对于正态分布, 采用矩法和极大似然法的计算参数一致.

(3) 概率权重矩法

$$\lambda_1 = \beta_0 = \mu \tag{1.6}$$

$$\lambda_2 = 2\beta_1 - \beta_0 = \frac{\sigma}{\sqrt{\pi}} \tag{1.7}$$

$$\tau_3 = \frac{\lambda_3}{\lambda_2} = 0 \tag{1.8}$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} = \frac{30}{\pi}\arctan\sqrt{2} - 9 = 0.1226 \tag{1.9}$$

概率权重矩法的参数估计结果由样本矩 $l_1$ 和 $l_2$ 表示为

$$\hat{\mu} = l_1 \tag{1.10}$$

$$\hat{\sigma} = \sqrt{\pi} l_2 \tag{1.11}$$

3) 设计值的推求

不超越概率 $F$ 对应的标准正态变量 $u$ 是易于计算的. Abramowitz 和 Stegun 于 1965 年 (Rao and Hamed, 2000) 给出了 $u$ 的计算公式.

$$u = W - \frac{C_0 + C_1 W + C_2 W^2}{1 + d_1 W + d_2 W^2 + d_3 W^3} + \varepsilon(P) \tag{1.12}$$

式中

$$C_0 = 2.515517, \quad C_1 = 0.802853, \quad C_2 = 0.010328$$

$$d_1 = 1.432788, \quad d_2 = 0.189269, \quad d_3 = 0.001308$$

当 $P < 0.5$ 时, $W$ 由下式计算：

$$W = \sqrt{-2\ln(P)} \tag{1.13}$$

式中, $P = 1 - F$, 误差 $\varepsilon(P)$ 小于 $4.5 \times 10^{-4}$. 当 $P > 0.5$ 时, $P$ 由 $1 - P$ 替换, 并且 $u$ 取相反的符号. 设计值 $\hat{x}_T$ 可由式 (1.14) 计算.

$$\hat{x}_T = \hat{\mu} + u\hat{\sigma} \tag{1.14}$$

1.1.1.2　两参数对数正态分布 (ln(2))

1) 概率密度

$$f(x) = \frac{1}{x\sigma_y\sqrt{2\pi}} \exp\left\{\frac{-\left[\ln x - \mu_y\right]^2}{2\sigma_y^2}\right\}, \quad x > 0 \tag{1.15}$$

式中, $\mu_y$ 和 $\sigma_y$ 分别是 $x$ 序列取自然对数 ($y = \ln x$) 后形成序列的均值和标准差.

2) 参数估计方法

(1) 矩法

$$\hat{\sigma}_y^2 = \ln\left(\frac{m_2}{(m_1')^2} + 1\right) \tag{1.16}$$

$$\hat{\mu}_y = \ln m_1' - \frac{\hat{\sigma}_y^2}{2} \tag{1.17}$$

(2) 极大似然法

$$\hat{\mu}_y = \frac{1}{n}\sum_{i=1}^{n} \ln x_i = m_{1y}' \tag{1.18}$$

$$\hat{\sigma}_y^2 = \frac{1}{n}\sum_{i=1}^{n} (\ln x_i - \hat{\mu}_y)^2 = m_{2y} \tag{1.19}$$

式中, $m_{1y}'$ 和 $m_{2y}$ 分别为 $x$ 样本取自然对数后形成序列的均值和方差.

(3) 概率权重矩法. 设 $\mathrm{erf}(x)$ 为误差函数, $F(x)$ 为正态分布函数.

$$\lambda_1 = \mathrm{e}^{\mu_y + \sigma_y^2/2} \tag{1.20}$$

$$\lambda_2 = \mathrm{e}^{\mu_y + \sigma_y^2/2}\mathrm{erf}(\sigma_y/2) \tag{1.21}$$

其中

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{e}^{-\mu^2}\mathrm{d}u = 2F(x\sqrt{2}) - 1 \tag{1.22}$$

将样本线性矩 $l_1$ 和 $l_2$ 代入式 (1.20) 和 (1.21), 有

$$\hat{\sigma}_y = 2\mathrm{erf}^{-1}\left(\frac{l_2}{l_1}\right) \tag{1.23}$$

$$\hat{\mu}_y = \ln l_1 - \frac{\hat{\sigma}_y^2}{2} \tag{1.24}$$

3) 设计值的推求

标准正态变量 $u$ 可由式 (1.12) 计算, 则设计值 $\hat{x}_T$ 可由式 (1.25) 和 (1.26) 推求.

$$\ln \hat{x}_T = \hat{\mu}_y + u\hat{\sigma}_y \tag{1.25}$$

$$\hat{x}_T = \mathrm{e}^{\hat{\mu}_y + u\hat{\sigma}_y} \tag{1.26}$$

### 1.1.2 Γ 分布类

#### 1.1.2.1 指数分布

1) 概率密度

Γ 分布类包括皮尔逊III型分布、一参数 Γ 分布和两参数 Γ 分布、对数皮尔逊

III型分布 (三参数 Γ 分布) 和广义 Γ 分布. 指数分布是一类特殊的 Γ 分布, 皮尔逊III型分布的概率密度函数为

$$f(x)=\frac{1}{\alpha^{\beta}\Gamma(\beta)}(x-\varepsilon)^{\beta-1}\mathrm{e}^{-\frac{x-\varepsilon}{\alpha}},\quad \varepsilon<x<\infty \tag{1.27}$$

式中, $\alpha>0$ 为尺度参数; $\beta>0$ 为形状参数; $\varepsilon$ 为位置参数. $\Gamma(x)$ 由式 (1.28)~(1.33) 定义.

$$\Gamma(y+1)=\int_0^{\infty}t^y\mathrm{e}^{-t}\mathrm{d}t,\quad y+1>0 \tag{1.28}$$

$\Gamma(x)$ 具有以下性质：

$$\Gamma(1)=\Gamma(2)=1 \tag{1.29}$$

$$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi} \tag{1.30}$$

$$\Gamma(y+1)=y\Gamma(y),\quad y>0 \tag{1.31}$$

$$\Gamma(y)=\Gamma(y+1)/y,\quad y<1 \tag{1.32}$$

$$\Gamma(k)=(k-1)!,\quad k\ \text{为正整数} \tag{1.33}$$

式 (1.27) 取 $\beta=1$ 时即为指数分布

$$f(x)=\frac{1}{\alpha}\mathrm{e}^{-\frac{x-\varepsilon}{\alpha}},\quad \varepsilon<x<\infty \tag{1.34}$$

2) 参数估计方法

(1) 矩法

$$\hat{\alpha}^2=m_2 \tag{1.35}$$

$$\hat{\varepsilon}=m_1'-\hat{\alpha} \tag{1.36}$$

(2) 极大似然法

$$\hat{\alpha}=\frac{n(m_1'-x_1)}{n-1},\quad x_1=\min(x_i) \tag{1.37}$$

$$\hat{\varepsilon}=\frac{nx_1-m_1'}{n-1} \tag{1.38}$$

(3) 概率权重矩法

$$\hat{\alpha}=2l_2 \tag{1.39}$$

$$\hat{\varepsilon}=l_1-2l_2 \tag{1.40}$$

3) 设计值的推求

$$\hat{x}_T=\hat{\varepsilon}+\hat{\alpha}\ln(T) \tag{1.41}$$

#### 1.1.2.2 两参数 Γ 分布 ($G(2)$)

1) 概率密度

两参数 Γ 分布式属一类特殊的皮尔逊III型分布, 即式 (1.27) 中的参数 $\varepsilon = 0$. 两参数 Γ 分布的概率密度函数为

$$f(x) = \frac{1}{\alpha^{\beta}\Gamma(\beta)} x^{\beta-1}\mathrm{e}^{-x/\alpha}, \quad x > 0, \alpha > 0, \beta > 0 \tag{1.42}$$

2) 参数估计方法

(1) 矩法

$$\hat{\alpha} = \frac{m_2}{m_1'} \tag{1.43}$$

$$\hat{\beta} = \frac{(m_1')^2}{m_2} \tag{1.44}$$

(2) 极大似然法

$$U = \ln A - \ln G \tag{1.45}$$

式中, $A$ 为样本算术平均值; $G$ 为样本几何平均值.

$$A = \frac{1}{n}\sum_{i=1}^{n} x_i = m_1' \tag{1.46}$$

$$G = (x_1 x_2 \cdots x_n)^{\frac{1}{n}} \tag{1.47}$$

$\beta$ 由式 (1.48) 和 (1.49) 计算.

对于 $0 \leqslant U \leqslant 0.5772$,

$$\hat{\beta} = \frac{1}{U}(0.5000876 + 0.1648852U - 0.054427U^2) \tag{1.48}$$

对于 $0.5772 \leqslant U \leqslant 17.0$,

$$\hat{\beta} = \frac{8.898919 + 9.059950U + 0.9775373U^2}{U(17.7928 + 11.968477U + U^2)} \tag{1.49}$$

$$\hat{\alpha} = \frac{A}{\hat{\beta}} \tag{1.50}$$

(3) 概率权重矩法

$$t = \frac{l_2}{l_1} \tag{1.51}$$

如果 $0 < t < 1/2$, 那么 $z = \pi t^2$, $\hat{\beta}$ 可由下式得到：

$$\hat{\beta} = \frac{1 - 0.3080z}{z - 0.05812z^2 + 0.01765z^3} \tag{1.52}$$

如果 $1/2 \leqslant t < 1$, 那么 $z = 1 - t$, $\hat{\beta}$ 可由下式得到：

$$\hat{\beta} = \frac{0.7213z - 0.5947z^2}{1 - 2.1817z + 1.2113z^2} \tag{1.53}$$

$$\hat{\alpha} = \frac{l_1}{\hat{\beta}} \tag{1.54}$$

3) 设计值的推求

Fisher 和 Cornish 于 1960 年 (Rao and Hamed, 2000) 给出了当 $|C_{\mathrm{s}}| < 2$ 时两参数 $\Gamma$ 分布 $K_T$ 的计算公式, 见式 (1.55).

$$\begin{aligned} K_T =& u + \frac{C_{\mathrm{s}}}{2}\left(\frac{u^2-1}{3}\right) + \frac{C_{\mathrm{s}}^2}{2^4}\left(\frac{u^3-7u}{9}\right)\left(\frac{C_{\mathrm{s}}^3}{2^5}\right)\left(\frac{6u^4+14u^2-32}{405}\right) \\ & + \frac{C_{\mathrm{s}}^4}{2^7}\frac{(9u^5+256u^3-433u)}{4860} + \frac{C_{\mathrm{s}}^5}{2^8}\frac{(12u^6-143u^4-923u^2+1472)}{25515} \\ & - \frac{C_{\mathrm{s}}^6}{2^{10}}\frac{(3753u^7+4353u^5-289517u^3-289717u)}{9185400} \end{aligned} \tag{1.55}$$

式中, $u$ 为 $F = 1 - \dfrac{1}{T}$ 非超越概率标准正态分布对应的分位数.

设计值 $\hat{x}_T$ 可由下式推求：

$$\hat{x}_T = \hat{\alpha}\hat{\beta} + K_T\sqrt{\hat{\alpha}^2\hat{\beta}} \tag{1.56}$$

1.1.2.3　皮尔逊Ⅲ型分布 (P-Ⅲ)

1) 概率密度

$$f(x) = \frac{1}{\alpha\Gamma(\beta)}\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1}\mathrm{e}^{-\frac{x-\gamma}{\alpha}}, \quad \gamma < x < \infty \tag{1.57}$$

式中, $\gamma$ 为位置参数; $\alpha > 0$ 为尺度参数; $\beta > 0$ 为形状参数.

2) 参数估计方法

(1) 矩法

$$\hat{\beta} = (2/C_{\mathrm{s}})^2 \tag{1.58}$$

$$\hat{\alpha} = \sqrt{m_2/\hat{\beta}} \tag{1.59}$$

$$\hat{\gamma} = m_1' - \sqrt{m_2\hat{\beta}} \tag{1.60}$$

(2) 极大似然法. 联立式 (1.61)~(1.63) 方程求解, 可得皮尔逊Ⅲ型分布的极大似然估计参数

$$\frac{n\beta}{\alpha} - \frac{1}{\alpha^2}\sum_{i=1}^{n}(x_i - \gamma) = 0 \tag{1.61}$$

$$-n\Psi(\beta)+\sum_{i=1}^{n}\ln\left(\frac{x_i-\gamma}{\alpha}\right)=0 \tag{1.62}$$

$$\frac{n}{\alpha}-(\beta-1)\sum_{i=1}^{n}\left(\frac{1}{x_i-\gamma}\right)=0 \tag{1.63}$$

式中,

$$\Psi(\beta)=\frac{\partial\ln\Gamma(\beta)}{\partial\beta}=\Gamma'(\beta)/\Gamma(\beta) \tag{1.64}$$

(3) 概率权重矩法.

对于 $t_3\geqslant 1/3$, 令 $t_\mathrm{m}=1-t_3$, 得

$$\hat{\beta}=\frac{0.36067t_\mathrm{m}-0.5967t_\mathrm{m}^2+0.25361t_\mathrm{m}^3}{1-2.78861t_\mathrm{m}+2.56096t_\mathrm{m}^2-0.77045t_\mathrm{m}^3} \tag{1.65}$$

对于 $t_3<1/3$, 令 $t_\mathrm{m}=3\pi t_3^2$ 得

$$\hat{\beta}=\frac{1+0.2906t_\mathrm{m}}{t_\mathrm{m}+0.1882t_\mathrm{m}^2+0.0442t_\mathrm{m}^3} \tag{1.66}$$

求出 $\hat{\beta}$ 后, 由式 (1.67), (1.68) 分别计算 $\hat{\alpha}$ 和 $\hat{\gamma}$.

$$\hat{\alpha}=\sqrt{\pi}l_2\frac{\Gamma(\hat{\beta})}{\Gamma(\hat{\beta}+1/2)} \tag{1.67}$$

$$\hat{\gamma}=l_1-\hat{\alpha}\hat{\beta} \tag{1.68}$$

注意, 如果 $t_3$ 为负值, 则式 (1.67) 计算的 $\hat{\alpha}$ 也为负值.

3) 设计值的推求

皮尔逊III型分布的设计值 $\hat{x}_T$ 可由式 (1.69) 推求, 式中 $K_T$ 可由式 (1.55) 计算.

$$\hat{x}_T=\hat{\alpha}\hat{\beta}+\hat{\gamma}+K_T\sqrt{\hat{\alpha}^2\hat{\beta}} \tag{1.69}$$

#### 1.1.2.4 对数皮尔逊III型分布

1) 概率密度

如果一个序列取自然对数后服从皮尔逊III型分布, 那么原数据序列服从对数皮尔逊III型分布. 对数皮尔逊III型分布的概率密度函数为

$$f(x)=\frac{1}{\alpha x\Gamma(\beta)}\left[\frac{\ln(x)-\gamma}{\alpha}\right]^{\beta-1}\mathrm{e}^{-\frac{\ln(x)-\gamma}{\alpha}},\quad \gamma<x<\infty \tag{1.70}$$

2) 参数估计方法

(1) 矩法. 令

$$A = \frac{1}{\alpha} - 3; \quad B = \frac{\ln m_3' - 3\ln m_1'}{\ln m_2' - 2\ln m_1'}; \quad C = \frac{1}{B-3} \tag{1.71}$$

当 $3.5 < B < 6.0$ 时,

$$A = -0.23019 + 1.65262C + 0.20911C^2 - 0.04557C^3 \tag{1.72}$$

当 $3.0 < B \leqslant 3.5$ 时,

$$A = -0.47157 + 1.99955C \tag{1.73}$$

$\alpha, \beta$ 和 $\gamma$ 分别由式 (1.74)~(1.76) 计算：

$$\hat{\alpha} = \frac{1}{A+3} \tag{1.74}$$

$$\hat{\beta} = \frac{\ln(1+C_v^2)}{[\ln(1-\alpha)^2 - \ln(1-2\alpha)]} = \frac{\ln m_2' - 2\ln m_1'}{\ln(1-\alpha)^2 - \ln(1-2\alpha)} \tag{1.75}$$

$$\hat{\gamma} = \ln m_1' + \beta\ln(1-\alpha) \tag{1.76}$$

(2) 极大似然法. 联立式 (1.77)~(1.79) 求解, 可得对数皮尔逊III型分布的极大似然估计参数：

$$\sum_{i=1}^{n}(\ln x_i - \gamma) = n\alpha\beta \tag{1.77}$$

$$n\Psi(\beta) = \sum_{i=1}^{n}\ln\frac{\ln x_i - \gamma}{\alpha} \tag{1.78}$$

$$n = \alpha(\beta-1)\sum_{i=1}^{n}\frac{1}{(\ln x_i - \gamma)} \tag{1.79}$$

式中, $\Psi(x)$ 为普西函数, 同式 (1.64).

3) 设计值的推求

设计值 $\hat{x}_T$ 可由式 (1.80) 推求, $K_T$ 可采用式 (1.55) 计算.

$$x_T = \mathrm{e}^{z_T} \tag{1.80}$$

式中,

$$z_T = \ln x_T = \mu_z + K_T\sigma_z \tag{1.81}$$

$$\mu_z = \hat{\gamma} + \hat{\alpha}\hat{\beta} \tag{1.82}$$

$$\sigma_z = \hat{\alpha}\sqrt{\hat{\beta}}, \quad C_s = 2/\sqrt{\hat{\beta}} \tag{1.83}$$

### 1.1.3 极值分布类

1.1.3.1 广义极值分布 (GEV)

1) 概率密度

$$f(x)=\frac{1}{a}\left[1-k\left(\frac{x-u}{a}\right)\right]^{1/k-1}\mathrm{e}^{-\left[1-k\left(\frac{x-u}{a}\right)\right]^{1/k}} \tag{1.84}$$

$$F(x)=\exp\left\{-\left[1-k\left(\frac{x-u}{a}\right)\right]^{1/k}\right\},\quad 1-k\left(\frac{x-u}{a}\right)>0 \tag{1.85}$$

式中, $u$ 为位置参数; $a$ 为尺度参数; $k$ 为形状参数.

2) 参数估计方法

(1) 矩法. 首先, 采用牛顿迭代法求解式 (1.86) 得 $\hat{k}$.

$$C_{\mathrm{s}}=\frac{\mu_3}{\mu_2^{3/2}}=\frac{k}{|k|}\frac{-\Gamma(1+3k)+3\Gamma(1+k)\Gamma(1+2k)-2\Gamma^3(1+k)}{[\Gamma(1+2k)-\Gamma^2(1+k)]^{3/2}} \tag{1.86}$$

然后, 计算 $\hat{a}$ 和 $\hat{u}$.

$$\hat{a}=\left[m_2\hat{k}^2/\left\{\Gamma(1+2\hat{k})-\Gamma^2(1+\hat{k})\right\}\right]^{1/2} \tag{1.87}$$

$$\hat{u}=m_1'-\frac{\hat{a}}{\hat{k}}[1-\Gamma(1+\hat{k})] \tag{1.88}$$

(2) 极大似然法. 联立方程 (1.89)~(1.91) 求解, 可得广义极值分布的极大似然估计参数.

$$-\frac{\partial\ln L}{\partial u}=\frac{Q}{a}=0 \tag{1.89}$$

$$-\frac{\partial\ln L}{\partial a}=\frac{1}{a}\frac{P+Q}{k}=0 \tag{1.90}$$

$$-\frac{\partial\ln L}{\partial k}=\frac{1}{k}\left(R-\frac{P+Q}{k}\right)=0 \tag{1.91}$$

在方程 (1.89)~(1.91) 中,

$$P=n-\sum_{i=1}^{n}\mathrm{e}^{-y_i},\quad y_i=\frac{1}{k}\ln\left[1-k\left(\frac{x_i-u}{a}\right)\right] \tag{1.92}$$

$$Q=\sum_{i=1}^{n}\mathrm{e}^{-y_i+ky_i}-(1-k)\sum_{i=1}^{n}\mathrm{e}^{ky_i} \tag{1.93}$$

$$R=n-\sum_{i=1}^{n}y_i+\sum_{i=1}^{n}y_i\mathrm{e}^{-y_i} \tag{1.94}$$

(3) 概率权重矩法. 求解式 (1.95) 可得 $\hat{k}$:

$$\frac{3+t_3}{2}=\frac{1-3^{-k}}{1-2^{-k}} \tag{1.95}$$

解得 $\hat{k}$ 的近似数值解为

$$\hat{k}=7.8590C+2.9554C^2 \tag{1.96}$$

式中, $C=\dfrac{2b_1-b_0}{3b_2-b_0}-\dfrac{\ln 2}{\ln 3}=\dfrac{2}{3+t_3}-\dfrac{\ln 2}{\ln 3}$.

$\hat{k}$ 求出后, 则可由式 (1.97) 和 (1.98) 分别计算 $\hat{a}$ 和 $\hat{u}$.

$$\hat{a}=\frac{(2b_1-b_0)\hat{k}}{\Gamma(1+\hat{k})(1-2^{-\hat{k}})}=\frac{l_2\hat{k}}{\Gamma(1+\hat{k})(1-2^{-\hat{k}})} \tag{1.97}$$

$$\hat{u}=b_0+\frac{\hat{a}}{\hat{k}}[\Gamma(1+\hat{k})-1]=l_1+\frac{\hat{a}}{\hat{k}}[\Gamma(1+\hat{k})-1] \tag{1.98}$$

3) 设计值的推求

$$\hat{x}_T=\hat{u}+\frac{\hat{a}}{\hat{k}}\left[1-\left\{-\ln\left(1-\frac{1}{T}\right)\right\}^{\hat{k}}\right] \tag{1.99}$$

1.1.3.2　极值 I 型分布 (EV1(2))

1) 概率密度

$$f(x)=\frac{1}{a}\exp\left[-\left(\frac{x-\beta}{a}\right)-\mathrm{e}^{-\frac{x-\beta}{a}}\right],\quad -\infty<x<\infty \tag{1.100}$$

分布函数为

$$F(x)=\exp\left[-\mathrm{e}^{-\frac{x-\beta}{a}}\right] \tag{1.101}$$

式中, $\beta$ 为位置参数; $a$ 为尺度参数. 极值 I 型分布是广义极值分布 $k$=0 时的特例.

2) 参数估计方法

(1) 矩法

$$\hat{a}=\frac{\sqrt{6}}{\pi}\sqrt{m_2}=0.7797\sqrt{m_2} \tag{1.102}$$

$$\hat{\beta}=m_1'-0.45005\sqrt{m_2} \tag{1.103}$$

(2) 极大似然法

$$a_{t+1}=a_t-F(a_t)/F'(a_t) \tag{1.104}$$

$$F(a)=\sum_{i=1}^{n}x_i\mathrm{e}^{-x_i/a}-\left(\frac{1}{n}\sum_{i=1}^{n}x_i-a\right)\sum_{i=1}^{n}\mathrm{e}^{-x_i/a}=0 \tag{1.105}$$

$$F'(a)=\frac{\mathrm{d}F(a)}{\mathrm{d}a}=\frac{1}{a^2}\sum_{i=1}^{n}x_i^2\mathrm{e}^{-x_i/a}+\sum_{i=1}^{n}\mathrm{e}^{-x_i/a}+\frac{1}{a}\sum_{i=1}^{n}x_i\mathrm{e}^{-x_i/a} \tag{1.106}$$

由式 (1.104)~(1.106) 采用牛顿迭代法可计算出 $\hat{a}$, 然后, 可由式 (1.107) 计算 $\hat{\beta}$.

$$\hat{\beta}=\hat{a}\ln\left(\frac{n}{\sum\limits_{i=1}^{n}\mathrm{e}^{-x_i/\hat{a}}}\right) \tag{1.107}$$

(3) 概率权重矩法

$$\hat{a}=(2b_1-b_0)/\ln 2=l_2/\ln 2 \tag{1.108}$$

$$\hat{\beta}=b_0-0.5772157\hat{a}=l_1-0.5772157\hat{a} \tag{1.109}$$

3) 设计值的推求

$$\hat{x}_T=\hat{\beta}+\hat{\alpha}\ln\left[-\ln(1-1/T)\right] \tag{1.110}$$

1.1.3.3 Weibull 分布

1)Weibull 分布

概率密度函数为

$$f(x)=\frac{b}{a}\left(\frac{x-m}{a}\right)^{b-1}\mathrm{e}^{-\left(\frac{x-m}{a}\right)^b} \tag{1.111}$$

分布函数为

$$F(x)=1-\mathrm{e}^{-\left(\frac{x-m}{a}\right)^b},\quad x-m>0 \tag{1.112}$$

式中, $m$ 为位置参数; $a$ 为尺度参数; $b$ 为形状参数.

Weibull 分布实际上是 $k=1/b$, $\alpha=a/b$, $u=m-a$ 时的广义极值分布特例.

2) 参数估计方法

(1) 矩法

$$(1/b)_{t+1}=(1/b)_t-F(1/b)/F'(1/b) \tag{1.113}$$

$$g_r=\Gamma(1+r/b) \tag{1.114}$$

$$d_r=\Gamma'(1+r/b)=\Psi(1+r/b)\Gamma(1+r/b) \tag{1.115}$$

$$F(1/b)=-C_s+\frac{g_3-3g_2g_1+2g_1^3}{(g_2-g_1^2)^{3/2}} \tag{1.116}$$

$$F'(1/b)=\frac{3d_3-3g_2d_1-6g_1d_2+6g_1^2d_1}{(g_2-g_1^2)^{3/2}}-\frac{3}{2}\frac{(g_3-3g_2g_1+2g_1^3)(2d_2-2g_1d_1)}{(g_2-g_1^2)^{5/2}} \tag{1.117}$$

由式 (1.113)~(1.117) 采用牛顿迭代法可计算出 $\hat{b}$.

其他参数计算公式如下：

$$\hat{a} = m_2^{1/2} / \left[\Gamma(1+2/\hat{b}) - \Gamma^2(1+1/\hat{b})\right]^{1/2} \tag{1.118}$$

$$\hat{m} = m_1' - \hat{a}\Gamma(1+1/\hat{b}) \tag{1.119}$$

(2) 极大似然法. 联立方程 (1.120)~(1.122) 求解, 可得 Weibull 分布的极大似然估计参数.

$$\frac{\partial \ln L}{\partial b} = \frac{n}{b} + \sum_{i=1}^{n} \ln\left(\frac{x_i - m}{a}\right) - \sum_{i=1}^{n}\left(\frac{x_i - m}{a}\right)^b \cdot \ln\left(\frac{x_i - m}{a}\right) = 0 \tag{1.120}$$

$$\frac{\partial \ln L}{\partial a} = -\frac{nb}{a} + \frac{b}{a}\sum_{i=1}^{n}\left(\frac{x_i - m}{a}\right)^b = 0 \tag{1.121}$$

$$\frac{\partial \ln L}{\partial m} = -(b-1)\sum_{i=1}^{n}(x_i - m)^{-1} + \frac{b}{a}\sum_{i=1}^{n}\left(\frac{x_i - m}{a}\right)^{b-1} = 0 \tag{1.122}$$

(3) 概率权重矩法. 参数 $b$ 采用式 (1.123) 计算.

$$\hat{b} = \frac{1}{7.8590C + 2.9554C^2} \tag{1.123}$$

式中,

$$C = \frac{2}{3 - t_3} - \frac{\ln 2}{\ln 3} \tag{1.124}$$

其他参数采用式 (1.125), (1.126) 计算.

$$\hat{a} = \frac{l_2}{\Gamma(1+1/\hat{b})(1 - 2^{-1/\hat{b}})} \tag{1.125}$$

$$\hat{m} = l_1 - \hat{a}\Gamma(1+1/\hat{b}) \tag{1.126}$$

3) 设计值的推求

$$\hat{x}_T = \hat{m} + \hat{a}\left[\ln(T)\right]^{1/\hat{b}} \tag{1.127}$$

### 1.1.4　Wakeby 分布类

#### 1.1.4.1　五参数 Wakeby 分布 (WAK(5))

1) 分位数表达式

如果变量 $x$ 可以用式 (1.128) 表示, 则 $x$ 服从 Wakeby 分布.

$$x = m + a[1 - (1-F)^b] - c[1 - (1-F)^{-d}] \tag{1.128}$$

式中, $F=F(x)=P(X\leqslant x)$; $m$ 为位置参数; $a$, $b$, $c$ 和 $d$ 为分布参数.

2) 参数估计方法

(1) 矩法与极大似然法. 该分布无法用矩法和极大似然法估计参数.

(2) 概率权重矩法. 定义

$$\{k\}=(k+1)(k+1+b)(k+1-d)a_k,\quad k=0,1,2,3,4 \tag{1.129}$$

$$N_{4-j}=(4)^j a_3-(3)^{1+j}a_2+3(2)^j a_1-a_0,\quad j=1,2,3 \tag{1.130}$$

$$C_{4-j}=(5)^j a_4-3(4)^j a_3+(3)^{1+j}a_2-(2)^j a_1,\quad j=1,2,3 \tag{1.131}$$

通过式 (1.129)~(1.131) 定义, 可以计算出五参数 Wakeby 分布概率权重估计参数.

$$\hat{m}=[\{3\}-\{2\}-\{1\}+\{0\}]/4 \tag{1.132}$$

$$\hat{b}=\frac{(N_3C_1-N_1C_3)\pm[(N_1C_3-N_3C_1)^2-4(N_1C_2-N_2C_1)(N_2C_3-N_3C_2)]^{1/2}}{2(N_2C_3-N_3C_2)} \tag{1.133}$$

$$\hat{d}=(N_1+\hat{b}N_2)/(N_2+bN_3) \tag{1.134}$$

$$\hat{a}=\frac{(\hat{b}+1)(\hat{b}+2)}{\hat{b}(\hat{b}+\hat{d})}\left[\frac{\{1\}}{2+\hat{b}}-\frac{\{0\}}{1+\hat{b}}-\hat{m}\right] \tag{1.135}$$

$$\hat{c}=-\frac{(1-\hat{d})(2-\hat{d})}{\hat{d}(\hat{b}+\hat{d})}\left[\frac{-\{1\}}{2-\hat{d}}-\frac{\{0\}}{1-\hat{d}}-\hat{m}\right] \tag{1.136}$$

由式 (1.133) 可得出两个 $\hat{b}$, 一般取较大的一个作为估计值.

3) 设计值的推求

$$\hat{x}_T=\hat{m}+\hat{a}[1-T^{-\hat{b}}]-\hat{c}[1-T^{\hat{d}}] \tag{1.137}$$

#### 1.1.4.2 四参数 Wakeby 分布 (WAK(4))

1) 分位数表达式

四参数 Wakeby 分布是当 $m$=0 时的五参数 Wakeby 分布特例. 如果变量 $x$ 可以用式 (1.138) 表示, 那么它服从四参数 Wakeby 分布.

$$x=a[1-(1-F)^b]-c[1-(1-F)^{-d}] \tag{1.138}$$

2) 参数估计方法

(1) 矩法与极大似然法.

该分布无法用矩法和极大似然法估计参数.

(2) 概率权重矩法.

定义

$$N_{4-j} = -(3)^j a_2 + (2)^{1+j} a_1 - a_0, \quad j = 1,2,3 \tag{1.139}$$

$$C_{4-j} = -(4)^j a_3 + 2(3)^j a_2 - (2)^j a_1, \quad j = 1,2,3 \tag{1.140}$$

然后取 $\hat{m} = 0$, 通过式 (1.139), (1.140) 可以计算出四参数 Wakeby 分布概率权重估计参数.

3) 设计值的推求

$$\hat{x}_T = \hat{a}[1 - T^{-\hat{b}}] - \hat{c}[1 - T^{\hat{d}}] \tag{1.141}$$

1.1.4.3　广义 Pareto 分布

1) 概率密度

广义 Pareto 分布是一类特殊的 Wakeby 分布. 如果变量 $x$ 可以用式 (1.142) 表示, 则 $x$ 服从广义 Pareto 分布.

$$x = \varepsilon + \frac{a}{k}\left[1 - (1-F)^k\right] \tag{1.142}$$

分布函数为

$$F(x) = 1 - \left[1 - \frac{k}{a}(x-\varepsilon)\right]^{1/k} \tag{1.143}$$

概率密度函数为

$$f(x) = \frac{1}{a}\left[1 - \frac{k}{a}(x-\varepsilon)\right]^{1/k-1} \tag{1.144}$$

变量 $x$ 的取值范围如下：

$$k \leqslant 0 \text{ 时}, \quad \varepsilon \leqslant x < \infty \tag{1.145}$$

$$k > 0 \text{ 时}, \quad \varepsilon \leqslant x \leqslant \varepsilon + a/k \tag{1.146}$$

2) 参数估计方法

(1) 矩法. 由式 (1.147)~(1.149), 采用牛顿迭代法可以计算出 $\hat{k}$.

$$k_{t+1} = k_t - F(k_t)/F'(k_t) \tag{1.147}$$

$$F(k) = \frac{2(1-k)(1+2k)^{1/2}}{1+3k} - C_{\rm s} \tag{1.148}$$

$$F'(k) = \frac{-2(1+2k)^{1/2}}{1+3k} + \frac{2(1-k)(1+2k)^{-1/2}}{1+3k} - \frac{6(1-k)(1+2k)^{1/2}}{(1+3k)^2} \tag{1.149}$$

其他参数的计算公式如下：

$$\hat{a} = [m_2(1+\hat{k})2(1+2\hat{k})]^{1/2} \tag{1.150}$$

$$\hat{\varepsilon} = m_1' - \frac{\hat{a}}{1+\hat{k}} \tag{1.151}$$

(2) 极大似然法. 采用牛顿迭代法求解联立方程 (1.152)~(1.154), 可以得到广义 Pareto 分布的极大似然估计参数.

$$L = \left(\frac{1}{a}\right)^n \prod_{i=1}^{n} \left[1 - \frac{k}{a}(x_i - \varepsilon)\right]^{1/k-1} \tag{1.152}$$

$$\frac{\partial \ln L}{\partial a} = \frac{1}{a}\left(\frac{1}{k} - 1\right)\sum_{i=1}^{n} y_i^{-1} - \frac{n}{ak} = 0, \quad y_i = 1 - \frac{k}{a}(x_i - \varepsilon) \tag{1.153}$$

$$\frac{\partial \ln L}{\partial k} = \frac{n}{k}\left(\frac{1}{k} - 1\right) - \frac{1}{k^2}\sum_{i=1}^{n} \ln(y_i) - \frac{1}{k}\left(\frac{1}{k} - 1\right)\sum_{i=1}^{n} y_i^{-1} = 0 \tag{1.154}$$

(3) 概率权重矩法. 由式 (1.155)~(1.157) 可以得到广义 Pareto 分布的概率权重估计参数.

$$\hat{k} = \frac{1 - 3t_3}{1 + t_3} \tag{1.155}$$

$$\hat{a} = l_2(1+\hat{k})(2+\hat{k}) \tag{1.156}$$

$$\hat{\varepsilon} = l_1 - l_2(2+\hat{k}) \tag{1.157}$$

3) 设计值的推求

$$\hat{x}_T = \hat{\varepsilon} + \frac{\hat{a}}{\hat{k}}[1 - T^{-\hat{k}}] \tag{1.158}$$

### 1.1.5 Logistic 分布类

1.1.5.1 Logistic 分布

1) 概率密度

分布函数为

$$F(x) = \left[1 + \mathrm{e}^{-\frac{x-m}{a}}\right]^{-1}, \quad -\infty < x < \infty \tag{1.159}$$

概率密度函数为

$$f(x) = \frac{1}{a}\mathrm{e}^{-\frac{x-m}{a}}\left[1 + \mathrm{e}^{-\frac{x-m}{a}}\right]^{-2} \tag{1.160}$$

2) 参数估计方法

(1) 矩法. 矩法估计 Logistic 分布参数的计算公式如下:

$$\hat{a} = \frac{\sqrt{3}}{\pi} m_2^{1/2} \tag{1.161}$$

$$\hat{m} = m_1' \tag{1.162}$$

(2) 极大似然法. 采用牛顿迭代法求解联立方程 (1.163)~(1.165), 可以得到 Logistic 分布的极大似然估计参数.

$$L = \frac{1}{a^n} \exp\left[-\frac{1}{a}\sum_{i=1}^{n}(x_i - m)\right]\prod_{i=1}^{n}\left[1 + \mathrm{e}^{-\frac{x_i - m}{a}}\right]^{-2} \tag{1.163}$$

$$\frac{\partial \ln L}{\partial m} = -\frac{n}{a} + \frac{2}{a}\sum_{i=1}^{n} y_i^{-1} = 0, \quad y_i = 1 + \mathrm{e}^{-\frac{x_i - m}{a}} \tag{1.164}$$

$$\frac{\partial \ln L}{\partial a} = -\frac{n}{a} - \frac{1}{a^2}\sum_{i=1}^{n}(x_i - m) + \frac{2}{a^2}\sum_{i=1}^{n}(x_i - m)\, y_i^{-1} = 0 \tag{1.165}$$

(3) 概率权重矩法.

$$\hat{a} = a_0 - 2a_1 = l_2 \tag{1.166}$$

$$\hat{m} = a_0 = l_1 \tag{1.167}$$

3) 设计值的推求

$$\hat{x}_T = \hat{m} + \hat{a}\ln(T - 1) \tag{1.168}$$

1.1.5.2 广义 Logistic 分布

1) 概率密度

分布函数为

$$F(x) = \left[1 + \left\{1 - k\left(\frac{x - \varepsilon}{a}\right)\right\}^{1/k}\right]^{-1} \tag{1.169}$$

变量 $x$ 的取值范围为

$$k < 0 \text{ 时}, \quad \varepsilon + \frac{a}{k} \leqslant x < \infty \tag{1.170}$$

$$k > 0 \text{ 时}, \quad -\infty < x \leqslant \varepsilon + \frac{a}{k} \tag{1.171}$$

概率密度函数为

$$f(x) = \frac{1}{a}\left[1 - k\left(\frac{x - \varepsilon}{a}\right)\right]^{(1/k-1)}\left[1 + \left\{1 - k\left(\frac{x - \varepsilon}{a}\right)\right\}^{1/k}\right]^{-2} \tag{1.172}$$

2) 参数估计方法

(1) 矩法

Ahmad 等于 1988 年 (Rao et al., 2000) 给出了一种 $\hat{k}$ 的计算方法 ($C_{\mathrm{s}} > 0$ 和 $k > -1/3$ 时).

当 $C_{\mathrm{s}} < 2.5$ 时,

$$\hat{k} = -\exp(-2.246 + 0.848z - 0.1271z^2 + 0.4008z^3) \tag{1.173}$$

当 $C_{\mathrm{s}} > 2.5$ 时,

$$\hat{k} = \frac{-C_{\mathrm{s}}^2}{4.007 + 3.411C_{\mathrm{s}} + 2.985C_{\mathrm{s}}^2} \tag{1.174}$$

式中, $z = \ln(C_{\mathrm{s}})$.

其他参数的计算公式为

$$\hat{a} = \left[m_2\hat{k}^2/(g_2 - g_1^2)\right]^{1/2} \tag{1.175}$$

$$\hat{\varepsilon} = m_1' - \frac{\hat{a}}{\hat{k}}\left[1 - g_1\right] \tag{1.176}$$

式中, $g_r = \Gamma\left(1 + rk\right)\Gamma(1 - rk)$.

(2) 极大似然法. 采用牛顿迭代法求解联立方程 (1.77)~(1.181), 可以得到广义 Logistic 分布的极大似然估计参数.

$$L = \frac{1}{a^n}\prod_{i=1}^{n}\left[1 - k\left(\frac{x-\varepsilon}{a}\right)\right]^{1/k-1}\prod_{i=1}^{n}\left[1 + \left\{1 - k\left(\frac{x-\varepsilon}{a}\right)\right\}^{1/k}\right]^{-2} \tag{1.177}$$

$$\ln L = -n\ln(a) + \left(\frac{1}{k} - 1\right)\sum_{i=1}^{n}\ln\left[1 - k\left(\frac{x_i-\varepsilon}{a}\right)\right] - 2\sum_{i=1}^{n}\ln\left[1 + \left\{1 - k\left(\frac{x_i-\varepsilon}{a}\right)\right\}^{1/k}\right] \tag{1.178}$$

$$-\frac{\partial\ln L}{\partial\varepsilon} = \frac{Q}{a} = 0 \tag{1.179}$$

$$-\frac{\partial\ln L}{\partial a} = \frac{P+Q}{ak} = 0 \tag{1.180}$$

$$-\frac{\partial\ln L}{\partial k} = -\frac{1}{k}\left[R - \frac{P+Q}{k}\right] = 0 \tag{1.181}$$

式中,

$$P = -n + 2\sum_{i=1}^{n}(1 + \mathrm{e}^{-y_i})^{-1}, \quad y_i = -\frac{1}{k}\ln\left\{1 - k\left(\frac{x_i-\varepsilon}{a}\right)\right\}^{1/k} \tag{1.182}$$

$$Q = (1+k)\sum_{i=1}^{n}\mathrm{e}^{y_ik} - 2\sum_{i=1}^{n}\mathrm{e}^{y_ik}(1 + \mathrm{e}^{y_i})^{-1} \tag{1.183}$$

$$R = n + \sum_{i=1}^{n}y_i - 2\sum_{i=1}^{n}y_i(1 + \mathrm{e}^{-y_i})^{-1} \tag{1.184}$$

(3) 概率权重矩法.

$$\hat{k} = -t_3 \tag{1.185}$$

$$\hat{a} = l_2 / \left[\Gamma(1+\hat{k})\Gamma(1-\hat{k})\right] \tag{1.186}$$

$$\hat{\varepsilon} = l_1 + (l_2 - \hat{a})/\hat{k} \tag{1.187}$$

3) 设计值的推求

$$\hat{x}_T = \hat{\varepsilon} + \frac{\hat{a}}{\hat{k}}[1-(T-1)^{-k}] \tag{1.188}$$

## 1.2　几种常见的偏态非对称分布

本节根据 A. Azzalini、A.K. Gupta 和 S. Nadarajah 的研究成果, 介绍几种常用的偏态非对称分布 (skew-asymmetric distributions) 模型. 这些模型的主要特点是引入了一个新的参数 $\lambda$ 控制分布模型中的偏态 (skew) 和峰度 (kurtosis) 系数.

**定义**　设 $U, V$ 为绝对连续独立、关于 0 对称的随机变量, 其分布密度和分布函数分别为 $g$ 和 $G$. 对于任意 $\lambda \in \mathbf{R}$, 函数

$$f_X(x) = 2g(x)\,G(\lambda x) \tag{1.189}$$

是一个随机变量 $X$ 的密度. 本节约定 $\lambda > 0$, 对于 $\lambda < 0$, 可以通过 $-X$ 转换为密度 $2g(x)\,G(-\lambda x)$.

式 (1.189) 具有下列特性：

(1) 当 $U = u$, $S_U = +1$ 依概率 $G(\lambda x)$, $S_U = -1$ 依概率 $1 - G(\lambda x)$, 有 $X = S_U U$.

(2) 当 $|U| = |u|$, $S_U = +1$ 依概率 $G(\lambda x)$, $S_U = -1$ 依概率 $1 - G(\lambda|x|)$, 有 $X = S_U|U|$.

式 (1.189) 可以被解释为给定 $\lambda U > V$ 下, $U$ 的条件密度.

为了统一起见, 本节约定以下函数：

(1) 不完全 Beta 函数 $\mathrm{B}_x(a,b)$ 和 $I_x(a,b)$

$$I_x(a,b) = \frac{1}{\mathrm{B}(a,b)}\int_0^x w^{a-1}(1-w)^{b-1}\,\mathrm{d}w, \quad \mathrm{B}_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1}\,\mathrm{d}t \tag{1.190}$$

(2) 第一类完全椭圆函数积分

$$\mathbf{K}(a) = \int_0^1 \frac{1}{\sqrt{1-x^2}\sqrt{1-a^2x^2}}\mathrm{d}x \tag{1.191}$$