

文章编号:0559-9350(2018)06-0687-07

## 基于 Copula 函数的水文随机变量和概率分布计算

宋松柏<sup>1</sup>, 王小军<sup>2</sup>

(1. 西北农林科技大学 水利与建筑工程学院, 旱区农业水土工程教育部重点实验室, 陕西 杨凌 712100;

2. 南京水利科学研究院 水文水资源与水利工程科学国家重点实验室, 江苏 南京 210029)

**摘要:** 水文随机变量和分布计算是推求设计洪水地区组成、梯级水库下游设计洪水等的重要内容, 对于水库下游水利水电工程规划设计与管理、城市防洪风险评估等至关重要。传统的水文随机变量和分布是根据二维变量函数分布推导而来, 边际分布必须为同一类型分布, 其应用受到限制。本文根据二维随机变量和概率分布定义, 运用条件 Copula 函数和积分变换原理, 严格地推导了二维相依随机变量和概率分布计算公式, 以及 Gamma 分布、P-III 分布两类常用边际分布下变量和的分布概率计算公式, 该计算公式仅为条件 Copula 函数的一维积分, 避免了概率组合离散求和法数据转换的信息失真, 克服了传统多变量分布要求边际分布为同一类型分布。以清江流域水布垭水库至隔河岩水库 3h 洪量组成为例, 给出了水文随机变量和分布的计算方法。文中模型与计算方法以期为我国设计洪水地区组成和梯级水库下游设计洪水计算提供理论支撑。

**关键词:** 相依性; 随机变量和; 概率分布; Copula 函数; 清江流域

**中图分类号:** TV122\*.3

**文献标识码:** A

**doi:** 10.13243/j.cnki.slxb.20171063

## 1 研究背景

二维相依水文随机变量和概率分布计算是流域设计洪水地区组成、梯级水库下游设计洪水等的重要内容和核心计算技术。因此, 如何提高水文变量和分布的计算精度, 受到许多学者的高度关注。《水利水电工程设计洪水计算手册》推荐使用地区组合法、频率组合法和随机模拟法进行设计洪水的地区组成计算<sup>[1]</sup>。其中, 频率组合法推荐使用 1990 年代王锐琛等提出的概率组合离散求和法。张元禧<sup>[2]</sup>是我国最早开展二维水文随机变量和概率分布计算的学者之一, 推导了具有形状参数为正整数的 Gamma 水文变量和、差分布解析计算公式。黄农<sup>[3]</sup>在此基础上, 扩充了张元禧的研究结果, 提出了两独立 Gamma 分布变量之和的数值计算方法。21 世纪以来, 出现了基于 JC 法、Copula 函数以及改进的离散求和法进行梯级水库设计洪水地区组成计算<sup>[4-11]</sup>。闫宝伟等<sup>[6]</sup>系统地总结了现有计算方法的不足: 离散求和法将二维积分转换为两变量有限个“状态”频率的组合求和, 其数据转化过程中难免出现信息失真; 随机模拟法通过建立空间多站随机数学模型进行上下游断面及区间洪水过程进行随机模拟, 模拟序列统计特征的保持性尚难掌握; 地区组合法虽然计算方法简便, 但将该法将区间洪水按某一组成相对固定, 人为因素的不确定性较大。李天元等<sup>[7]</sup>应用 Copula 函数, 获得条件概率显式表达式, 最后对条件概率曲线进行离散来求解二维相依水文随机变量和的分布概率, 这种方法提高了概率组合离散求和法的计算精度, 无需对变量做独立性处理, 但是, 它仍然属于概率组合离散求和法。

收稿日期: 2017-11-02; 网络出版日期: 2018-06-27

网络出版地址: <http://kns.cnki.net/kcms/detail/11.1882.TV.20180627.1118.007.html>

基金项目: 国家自然科学基金项目(51479171, 51579059, 51179160); 国家重点研发计划项目(2017YFC0403506)

作者简介: 宋松柏(1965-), 男, 陕西永寿人, 博士, 教授, 主要从事水文水资源研究。E-mail: ssb6533@nwsuaf.edu.cn

通讯作者: 王小军(1983-), 男, 陕西西安人, 博士, 教授级高级工程师, 主要从事水文水资源研究。E-mail: xjwang@nhri.cn

Nadarajah 推导边际分布必须为同一类型分布<sup>[12-19]</sup>。本文试图根据二维随机变量和分布函数的定义,从二维随机变量和概率分布的定义出发,运用 Copula 函数和数学积分变换原理,将二维随机变量和概率分布计算转换为条件 Copula 的一维定积分,力求避免离散求和法数据转化过程中的信息失真。在此基础上,推导了 Gamma 分布、P-Ⅲ分布两类常用边际分布变量和的分布概率计算公式。以清江流域水布垭水库至隔河岩水库 3 h 洪水组成为例,说明文中模型的应用。文中模型以期为我国设计洪水地区组成和梯级水库下游设计洪水计算等提供理论支撑。

## 2 二维随机变量和分布的 Copula 函数表达式

设二维随机变量 $(X, Y)$ 的联合密度为 $f(x, y)$ , $X$ 与 $Y$ 相互独立,则 $Z = X + Y$ 的分布概率为<sup>[12]</sup>:

$$F_Z(z) = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{z-x} f(x, y) dy \right] dx = \int_{-\infty}^{\infty} f_X(x) \cdot \left[ \int_{-\infty}^{z-x} \frac{f(x, y)}{f_X(x)} dy \right] dx \quad (1)$$

式中: $f_X(x)$ 为随机变量 $X$ 的密度函数;随机变量 $X$ 、 $Y$ 的积分上、下限可分别根据它们分布函数变量的取值范围来进行确定。

设二维 Copula 函数为 $C(u, v)$ , $u = F_X(x)$ , $v = F_Y(y)$ ,其中, $F_X(x)$ 、 $F_Y(y)$ 分别为随机变量 $X$ 、 $Y$ 的分布函数。根据 Copula 函数性质,联合密度 $f(x, y)$ 与 Copula 函数 $C(u, v)$ 有下列关系<sup>[21-22]</sup>:

$$f(x, y) = \frac{\partial^2 C(u, v)}{\partial u \partial v} \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} = \frac{\partial^2 C(u, v)}{\partial u \partial v} f_X(x) f_Y(y) \quad (2)$$

将式(2)代入式(1),有

$$F_Z(z) = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{z-x} \frac{\partial^2 C(u, v)}{\partial u \partial v} f_Y(y) dy \right] f_X(x) dx \quad (3)$$

令 $u = F_X(x)$ ,则 $x = F_X^{-1}(u)$ , $F_X^{-1}(u)$ 为随机变量 $X$ 分布的逆函数。当 $x \rightarrow -\infty$ 时, $u = 0$ ,当 $x \rightarrow \infty$ 时, $u = 1$ , $du = dF_X(x) = f_X(x)dx$ ;  $v = F_Y(y)$ ,当 $y \rightarrow -\infty$ 时, $v = 0$ ,当 $v = z - x$ 时, $v = F_Y(z - x) = F_Y[z - F_X^{-1}(u)]$ , $dv = dF_Y(y) = f_Y(y)dy$ ,则

$$\begin{aligned} F_Z(z) &= \int_0^1 \left[ \int_0^{F_Y[z - F_X^{-1}(u)]} \frac{\partial^2 C(u, v)}{\partial u \partial v} dv \right] du = \int_0^1 \frac{\partial}{\partial u} \left[ \int_0^{F_Y[z - F_X^{-1}(u)]} \frac{\partial C(u, v)}{\partial v} dv \right] du \\ &= \int_0^1 \frac{\partial C \left\{ u, F_Y[z - F_X^{-1}(u)] \right\}}{\partial u} du \end{aligned} \quad (4)$$

式(4)即为二维随机变量和概率分布的 Copula 函数计算表达式,显然,积分函数仅为变量 $u$ 的函数。二维 Copula 函数为 $C(u, v)$ 可根据二维随机变量 $(X, Y)$ 的相依特性,按照通常基于 Copula 函数的多变量分布计算方法选择函数类型和确定相应的参数。积分变量 $u$ 的上、下限可根据随机变量 $X$ 的取值范围,通过 $u = F_X(x)$ 来进行确定。

## 3 边际分布 Gamma、P-Ⅲ分布下变量和分布 Copula 函数表达式

Gamma、P-Ⅲ分布是水文常用的分布函数,以下将根据式(4)推导边际分布服从 Gamma、P-Ⅲ分布的变量和分布 Copula 函数表达式。

假定随机变量 $X$ 和 $Y$ 服从 Gamma 分布,其密度函数分别为

$$f_X(x) = \frac{1}{b_1 \Gamma(a_1)} \left( \frac{x}{b_1} \right)^{a_1 - 1} e^{-\frac{x}{b_1}}, \quad x > 0 \quad (5)$$

$$f_Y(y) = \frac{1}{b_2 \Gamma(a_2)} \left( \frac{y}{b_2} \right)^{a_2 - 1} e^{-\frac{y}{b_2}}, \quad y > 0 \quad (6)$$

若随机变量  $X$  和  $Y$  服从 P-III 分布，则其密度函数分别为：

$$f_X(x) = \frac{1}{b_1 \Gamma(a_1)} \left( \frac{x-c_1}{b_1} \right)^{a_1-1} e^{-\frac{x}{b_1}}, \quad x > c_1 \quad (7)$$

$$f_Y(y) = \frac{1}{b_2 \Gamma(a_2)} \left( \frac{y-c_2}{b_2} \right)^{a_2-1} e^{-\frac{y}{b_2}}, \quad y > c_2 \quad (8)$$

式中： $a_1$ 、 $a_2$  分别为形状参数； $b_1$ 、 $b_2$  分别为尺度参数； $c_1$ 、 $c_2$  分别为位置参数； $\Gamma(a_1) = \int_0^\infty t^{a_1-1} e^{-t} dt$ ， $\Gamma(a_2) = \int_0^\infty t^{a_2-1} e^{-t} dt$ 。

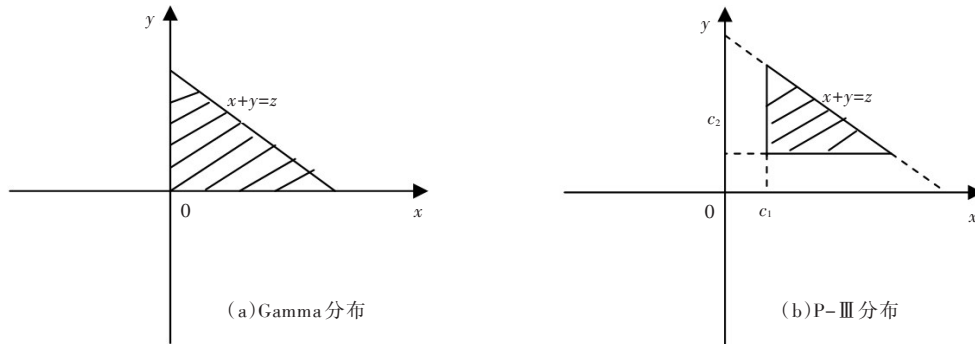


图1 积分区域

给定 Gamma、P-III 边际分布下，变量和分布  $Z = X + Y$  概率计算的积分区域见图 1 所示。Gamma 分布变量  $X$  和  $Y$  分别取值  $x > 0$ ， $y > 0$ ，P-III 分布变量  $x > c_1$ ， $y > c_2$ ， $z = x + y$  为有限值，且已知，根据图 1 所示的积分区域和二维定积分原理， $x$  积分则转换为变量  $u$  积分，其上限值为  $z$  的  $x$  边际分布值，不再取值 1。

$$F_Z(z) = \int_0^{F_X(z)} \frac{\partial C \left\{ u, F_Y \left[ z - F_X^{-1}(u) \right] \right\}}{\partial u} du \quad (9)$$

式中，对于边际分布服从 Gamma 分布， $u = F_X(x) = \frac{1}{b_1 \Gamma(a_1)} \int_0^x \left( \frac{t}{b_1} \right)^{a_1-1} e^{-\frac{t}{b_1}} dt$ ， $v = F_Y(y) = \frac{1}{b_2 \Gamma(a_2)}$

$\int_0^y \left( \frac{t}{b_1} \right)^{a_2-1} e^{-\frac{t}{b_2}} dt$ ；对于边际分布服从 P-III 分布， $u = F_X(x) = \frac{1}{b_1 \Gamma(a_1)} \int_{c_1}^x \left( \frac{t-c_1}{b_1} \right)^{a_1-1} e^{-\frac{t-c_1}{b_1}} dt$ ， $v = F_Y(y)$

$$= \frac{1}{b_2 \Gamma(a_2)} \int_{c_2}^y \left( \frac{t-c_2}{b_2} \right)^{a_2-1} e^{-\frac{t-c_2}{b_2}} dt。$$

#### 4 二维随机变量和概率分布模型求解

由式(9)可以看出， $z$  已知，二维随机变量和概率分布  $F_Z(z)$  仅为变量  $u$  的一维定积分。当  $\frac{\partial C \left\{ u, F_Y \left[ z - F_X^{-1}(u) \right] \right\}}{\partial u}$  函数为简单初等函数时，可获得  $F_Z(z)$  的解析积分值。实际中，Copula 函数表达式一般较为复杂， $\frac{\partial C \left\{ u, F_Y \left[ z - F_X^{-1}(u) \right] \right\}}{\partial u}$  形式复杂，只能采用数值积分获得  $F_Z(z)$  的数值积分值。

本文采用连分式数值积分法计算  $F_Z(z)$ ，其步骤如下<sup>[23-24]</sup>：

为了叙述方便，式(9)进一步可写为  $\int_a^b f(u) du$ ，其中， $a = 0$ ， $b = F_X(z)$ ， $f(u) = \frac{\partial C \left\{ u, F_Y \left[ z - F_X^{-1}(u) \right] \right\}}{\partial u}$ 。

利用变步长梯形法计算出步长为<sup>[23-24]</sup> $h_i = \frac{b-a}{2^{i-1}}$ ,  $i=1, 2, \dots$ 的系列积分近似值 $S_i$ ,  $i=1, 2, \dots$ 。积分近似值 $S$ 为步长 $h$ 的函数 $S(h)$ 。 $S(h)$ 可用连分式表示:

$$S(h) = b_1 + \frac{h-h_1}{b_2 + \frac{h-h_2}{b_3 + \dots + \frac{h-h_i}{b_{i+1}} + \dots}} \quad (10)$$

式中,  $b_1, b_2, \dots, b_i, \dots$ , 由式(11)计算的一系列积分近似点 $(h_i, S_i)$ 来确定,  $i=1, 2, \dots$ 。

$$\begin{cases} u = S_i \\ u = b_1 + \frac{h_i - h_{j-1}}{u - b_{j-1}}, \quad j=2, 3, \dots, i \\ b_i = u \end{cases} \quad (11)$$

当 $h=0$ 时,  $S(0)$ 即为积分值。即

$$S = S(0) = b_1 - \frac{h_1}{b_2 - \frac{h_2}{b_3 - \dots - \frac{h_i}{b_{i+1}} + \dots}} \quad (12)$$

实际计算中, 连分式计算一般取 $i=7-10$ 就能满足精度要求。

## 5 应用实例

本文选用文献[7]采用的清江流域梯级水库3h洪量序列分布参数和Copula函数参数, 说明二维相依水文随机变量和概率分布计算模型的应用。 $X$ 代表水布垭水库3h洪量,  $Y$ 代表水布垭水库至隔河岩水库区间(水-隔区间)3h洪量, 水布垭水库水量加上水-隔区间水量等于下游隔河岩水库水量。3种序列的统计参数和分布参数见表1<sup>[7]</sup>。

表1 序列参数计算结果

| 序列    | 统计参数               |                     |      |        | Gamma分布 |        | P-III分布 |        |        |
|-------|--------------------|---------------------|------|--------|---------|--------|---------|--------|--------|
|       | 均值/亿m <sup>3</sup> | 均方差/亿m <sup>3</sup> | 变差系数 | 偏态系数   | $a$     | $b$    | $a$     | $b$    | $c$    |
| 水布垭   | 7.94               | 3.4936              | 0.44 | 1.3200 | 5.1653  | 1.5372 | 2.2957  | 2.3058 | 2.6467 |
| 水-隔区间 | 3.08               | 1.9096              | 0.62 | 1.8600 | 2.6015  | 1.1840 | 1.1562  | 1.7759 | 1.0267 |
| 隔河岩   | 11.10              | 5.4390              | 0.49 | 1.4700 | 4.1649  | 2.6651 | 1.8511  | 3.9977 | 3.7000 |

选用二维Gumbel-Hougaard Copula函数, 其分布函数为:

$$C(u, v) = e^{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}}, \quad \theta \geq 1 \quad (13)$$

式中:  $u = F_X(x)$ 、 $v = F_Y(y)$ 分别为边际分布,  $\theta$ 为Gumbel-Hougaard Copula函数参数, 本文中 $\theta = 1.89$ 。根据式(13), 有Copula函数的偏导数:

$$\frac{\partial C(u, v)}{\partial u} = e^{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}} \frac{\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}-1} (-\ln u)^{\theta-1}}{u} \quad (14)$$

将式(14)代入式(9), 随机变量 $X$ 和 $Y$ 分别选用Gamma、P-III分布, 应用上述连分式数值计算方法, 两种边际分布下,  $X+Y \leq Z$ 的概率计算结果分别见表2第(2)和第(5)列, 其对比见图2所示。为了验证文中模型计算结果的正确性, 表2第(3)和第(6)列分别列出 $X+Y$ 序列分别按Gamma、P-III分布计算的理论概率值, 第(4)和第(7)列给出了相应的误差计算值。其误差是由于 $X$ 、 $Y$ 分布参数不同, 式(4)没有显式解析计算公式, 只能由文中变步长梯形法数值计算, 因而具有一定的偏差, 另外

也受边缘分布拟合精度的影响。不难看出，两种边缘分布下，文中模型与变量和序列的和概率分布计算结果基本一致，表明文中模型和计算方法是正确的。

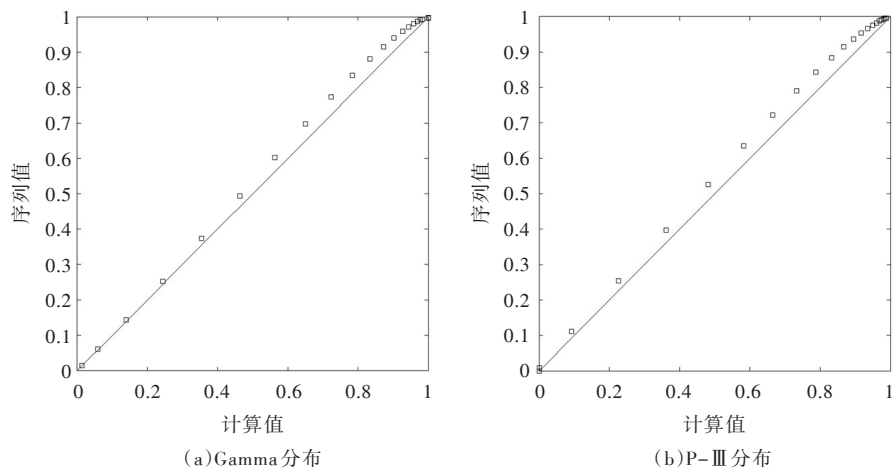


图2  $X+Y \leq Z$  分布概率计算结果对比

表2  $X+Y \leq Z$  分布概率计算结果

| z     | Gamma分布 |        |         | P-III分布 |        |         |
|-------|---------|--------|---------|---------|--------|---------|
|       | 计算值     | 序列值    | 误差      | 计算值     | 序列值    | 误差      |
| (1)   | (2)     | (3)    | (4)     | (5)     | (6)    | (7)     |
| 2.65  | 0.0128  | 0.0143 | -0.0015 | 0.0000  | 0.0000 | 0.0000  |
| 4.15  | 0.0588  | 0.0602 | -0.0014 | 0.0000  | 0.0092 | -0.0092 |
| 5.65  | 0.1401  | 0.1427 | -0.0027 | 0.0924  | 0.1107 | -0.0183 |
| 7.15  | 0.2434  | 0.2522 | -0.0088 | 0.2266  | 0.2537 | -0.0270 |
| 8.65  | 0.3545  | 0.3737 | -0.0192 | 0.3613  | 0.3977 | -0.0364 |
| 10.15 | 0.4630  | 0.4939 | -0.0309 | 0.4809  | 0.5265 | -0.0456 |
| 11.65 | 0.5626  | 0.6035 | -0.0409 | 0.5820  | 0.6346 | -0.0526 |
| 13.15 | 0.6498  | 0.6975 | -0.0477 | 0.6655  | 0.7220 | -0.0565 |
| 14.65 | 0.7236  | 0.7744 | -0.0508 | 0.7334  | 0.7908 | -0.0574 |
| 16.15 | 0.7846  | 0.8351 | -0.0506 | 0.7883  | 0.8439 | -0.0557 |
| 17.65 | 0.8337  | 0.8815 | -0.0478 | 0.8322  | 0.8844 | -0.0522 |
| 19.15 | 0.8728  | 0.9162 | -0.0434 | 0.8673  | 0.9149 | -0.0477 |
| 20.65 | 0.9034  | 0.9415 | -0.0381 | 0.8952  | 0.9377 | -0.0425 |
| 22.15 | 0.9270  | 0.9596 | -0.0326 | 0.9173  | 0.9546 | -0.0373 |
| 23.65 | 0.9452  | 0.9724 | -0.0272 | 0.9348  | 0.9670 | -0.0322 |
| 25.15 | 0.9591  | 0.9814 | -0.0223 | 0.9487  | 0.9761 | -0.0274 |
| 26.65 | 0.9696  | 0.9875 | -0.0179 | 0.9596  | 0.9828 | -0.0231 |
| 28.15 | 0.9775  | 0.9917 | -0.0142 | 0.9683  | 0.9876 | -0.0193 |
| 29.65 | 0.9832  | 0.9945 | -0.0113 | 0.9751  | 0.9911 | -0.0160 |
| 31.15 | 1.0000  | 0.9964 | 0.0036  | 0.9806  | 0.9936 | -0.0130 |
| 32.65 | 1.0000  | 0.9977 | 0.0023  | 0.9844  | 0.9954 | -0.0110 |
| 34.15 | 1.0000  | 0.9985 | 0.0015  | 0.9877  | 0.9967 | -0.0090 |

## 6 结论

二维相依水文随机变量和概率分布计算是梯级水库下游设计洪水计算、设计洪水地区组成的重要计算内容。本文基于 Copula 函数和数学积分变换原理，严格地推导了两 Gamma 分布、P-III 分布相依水文随机变量和分布的概率计算公式。以清江流域水布垭水库至隔河岩水库 3 h 洪量组成为例，说

明文中模型的应用。主要研究结论如下：(1)文中计算公式克服了离散求和法进行数据转换中可能出现的误差。(2)计算公式表达为一维条件 Copula 函数的定积分，利用计算机数值积分，不难获得满意的求解结果。对于 P-Ⅲ分布，通过计算序列变量值减去分布的位置参数，即可转换为 Gamma 分布函数数值计算。(3)文中二维相依水文随机变量和概率分布采用条件 Copula 函数计算，可方便地组合不同分布相依的水文特征变量，可克服传统多变量分布函数要求同一类边际分布函数的限制。(4)研究区两 Gamma 分布、P-Ⅲ分布下，洪量组合和概率分布结果表明，文中随机变量和概率分布计算公式是正确的。文中模型以期为我国流域设计洪水地区组成和梯级水库下游设计洪水计算提供理论支撑。(5)文中仅根据下游设计断面洪量等于上游水库断面洪量和区间洪量之和，没有考虑水库的调节影响，推导了下游设计断面洪量概率分布计算公式。因此，如何考虑水库调洪影响下的下游断面设计洪水有待于进一步深入研究。

### 参 考 文 献：

- [ 1 ] 水利部长江水利委员会水文局，水利部南京水文水资源研究所．水利水电工程设计洪水计算手册[M]．北京：中国水利水电出版社，2001．
- [ 2 ] 张元禧．相互独立量 GAMMA 分布随机变数之和、差的联合分布函数[J]．合肥工业大学学报，1983(3)：53-59．
- [ 3 ] 黄农．一种推算两独立 $\Gamma$ 分布随机变数之和的组合频率数值计算法[J]．合肥工业大学学报，1987，9(5)：110-115．
- [ 4 ] 谢小平，黄灵芝，席秋义，等．基于 JC 法的设计洪水地区组成研究[J]．水力发电学报，2006，25(6)：125-129．
- [ 5 ] 黄灵芝，谢小平，黄强，等．梯级水库设计洪水地区组成研究中的 JC 法[J]．自然灾害学报，2006，15(4)：163-167．
- [ 6 ] 闫宝伟，雷时忠，胡四一，等．基于 Copula 函数的设计洪水地区组成研究[J]．水力发电学报，2010，29(6)：60-65．
- [ 7 ] 李天元，郭生练，刘章君，等．梯级水库下游设计洪水计算方法研究[J]．水利学报，2014，45(6)：641-648．
- [ 8 ] 刘章君，郭生练，李天元，等．基于 Copula 函数的梯级水库设计洪水地区组成研究[J]．水资源研究，2014，3(2)：124-135．
- [ 9 ] 刘章君，郭生练，李天元，等．设计洪水地区组成的区间估计方法研究[J]．水利学报，2015，46(5)：543-549．
- [ 10 ] 刘章君，郭生练，李天元，等．梯级水库设计洪水最可能地区组合法计算通式[J]．水科学进展，2014，25(4)：549-558．
- [ 11 ] 栗飞，郭生练，李天元，等．不连续序列的设计洪水地区组成方法研究[J]．水电能源科学，2011，29(5)：47-49．
- [ 12 ] NADARAJAH S．Products and ratios for a bivariate gamma distribution[J]．Journal of Applied Mathematics and Computing，2005，24：581-595．
- [ 13 ] NADARAJAH S．A bivariate distribution with gamma and beta marginals with application to drought data[J]．Journal of Applied Statistics，2009，36：277-301．
- [ 14 ] NADARAJAH S．A bivariate gamma model for drought[J]．Water Resources Research，2007，43：749-759．
- [ 15 ] NADARAJAH S．A bivariate Pareto model for drought[J]．Stochastic Environmental Research and Risk Assessment，2008，23：811-822．
- [ 16 ] NADARAJAH S，GUPTA A K．Cherian's bivariate gamma distributions as a model for drought data[J]．Agrociencia-Mexico，2006，40：483-490．
- [ 17 ] NADARAJAH S，GUPTA A K．Friday and Patil's bivariate exponential distribution with application to drought data[J]．Water Resources Management，2006，20：749-759．
- [ 18 ] NADARAJAH S，GUPTA A K．Intensity-duration models based on bivariate gamma distributions[J]．Hiroshima

Mathematical Journal, 2006, 36: 387-395 .

- [ 19 ] NADARAJAH S, KOTZ S . Sums, products and ratios for Downton 's bivariate exponential distribution[J] . Stochastic Environmental Research and Risk Assessment, 2006, 20: 164-170 .
- [ 20 ] 黄振平, 陈元芳 . 水文统计学[M] . 北京: 中国水利水电出版社, 2011 .
- [ 21 ] 宋松柏, 蔡焕杰, 金菊良, 等 . Copulas函数理论及其在水文中的应用[M] . 北京: 科学出版社, 2012 .
- [ 22 ] NELSEN R B . An Introduction to Copulas [M] . Springer, New York, 1999 .
- [ 23 ] 徐士良 . FORTRAN常用算法程序集[M] . 2版 . 北京: 清华大学出版社, 1997 .
- [ 24 ] 邹海 . 最优设计中的新计算方法[M] . 北京: 新时代出版社, 1982 .

## Probability distribution calculation of the sum of hydrological random variables based on Copula function approach

SONG Songbai<sup>1</sup>, WANG Xiaojun<sup>2</sup>

(1. College of Water Resources & Architecture Engineering, Key Laboratory for Agricultural Soil and Water Engineering in Arid Area of Ministry of Education, Northwest A & F University, Yangling 712100, China;

2. Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China)

**Abstract:** The distribution probability calculation of the sum of hydrological random variables is a key content for the design flood in regional composition and the design flood in downstream area of cascade reservoirs. It is also important for planning, designing and management of water resources projects in downstream area of cascade reservoirs, and evaluating urban flood risk. But probability distribution of the sum of random is derived by two dimensional random function distribution, its marginal distributions must have the same type probability distribution, and the applications are limited in practical. In this study, based on the probability distribution definition of the sum of two dimensional random variables, a probability distribution calculation model of the sum of two dimensional dependent random variables was proposed by using of Copula function and mathematical transformation. This model was also extended to calculate probability distribution for the sum of two dimensional dependent random variables with Gamma and P-III marginal distribution, respectively. The proposed model, which expressed as an integral of one dimensional conditional Copula, can overcome shortcomings of information distortion in data transformation process of probability combination discrete summation method, and the same type marginal distributions. Taking the 3h flood volumes of downstream area of the Qingjiang cascade reservoirs as an example, the proposed model calculation method and procedures were given. It also indicates that the model in this paper can provide some basic theories for the design flood region composition and design flood of downstream area of cascade reservoirs.

**Keywords:** dependent; sum of random variables; probability calculation; Copula function; Qingjiang watershed

(责任编辑: 王成丽)