

ВОДНЫЕ РЕСУРСЫ И РЕЖИМ ВОДНЫХ ОБЪЕКТОВ

УДК 556.048

ДВУХМЕРНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН, ИМЕЮЩИХ ТРЕХПАРАМЕТРИЧЕСКИЕ ГАММА-РАСПРЕДЕЛЕНИЯ С.Н. КРИЦКОГО И М.Ф. МЕНКЕЛЯ (СИММЕТРИЧНЫЙ СЛУЧАЙ)¹

© 2020 г. М. В. Болгов^а, *, И. О. Сарманов^а

^аИнститут водных проблем РАН,
Москва, 119333 Россия

*e-mail: bolgovmv@mail.ru

Поступила в редакцию 15.02.2019 г.

После доработки 15.02.2019 г.

Принята к публикации 11.04.2019 г.

Рассмотрен метод построения двухмерного закона распределения в симметричном случае для трехпараметрического распределения С.Н. Крицкого и М.Ф. Менкеля, а также обсуждаются некоторые результаты применения рекомендуемой модели в прикладных гидрологических исследованиях. Для получения линейной корреляции трехпараметрических величин предложена система ортогональных функций с весовой функцией в виде трехпараметрического гамма-распределения. На основе метода ортогонализации получено выражение для двумерной симметричной плотности, которая удовлетворяет уравнению Маркова и характеризуется линейным уравнением регрессии. Результаты стохастического моделирования колебаний уровня воды бессточного оз. Чаны, выполненные на основе рекомендуемой модели, показали, что использование линейной корреляции приводит к величине асимметрии распределения уровня, которое хорошо согласуется с выборочными оценками этого параметра по данным наблюдений.

Ключевые слова: распределение Крицкого–Менкеля, двухмерный закон, линейная регрессия, ортогональные полиномы, стохастическое моделирование, многолетние колебания.

DOI: 10.31857/S0321059620040033

ВВЕДЕНИЕ

Специфика вероятностных гидрологических задач заключается в том, что величины речного стока — положительные случайные величины, распределения вероятностей которых асимметричны. В силу этого нормальное распределение вероятностей и связанное с ним двухмерное нормальное распределение (нормальная корреляция) во многих случаях неприменимы в гидрологических исследованиях. Возникает необходимость разработки специальных приемов для анализа асимметричных распределений. Наиболее широко распространенное в гидрологических исследованиях распределение — трехпараметрическое гамма-распределение.

Метод построения двумерной плотности распределения (двухмерного закона распределения), определяющей стационарный Марковский процесс, — в симметричном случае для трехпараметрических распределений С.Н. Крицкого и

М.Ф. Менкеля, результат его применения в прикладных гидрологических исследованиях — применен к описанию многолетних колебаний уровня воды бессточного озера Чаны.

В стохастической гидрологии проблема построения двухмерного закона распределения с произвольными маргинальным распределениями рассматривалась различными авторами. Наиболее распространен подход, основанный на нормализации, а именно — на преобразовании классического нормального закона распределения в совместную плотность равномерно-распределенных случайных величин, а затем в совместное распределение случайных величин с заданным асимметричным распределением. Другой, более простой метод — замена переменной в совместном распределении равномерно-распределенных случайных величин, построенном на основе известного разложения по ортогональным полиномам Лежандра. Предлагались и другие подходы, подробно исследованные Д.Я. Ратковичем, А.Ш. Резниковским, М.В. Болговым и др. Все методы при корректном построении дают двумерную плотность распределения, позволяющую

¹ Работа выполнена в рамках государственного задания ИВП РАН (тема № 0147-2019-0003, государственная регистрация № АААА-А18-118022090105-5).

решать водохозяйственные задачи. Проблема заключается в том, что модель обладает различными свойствами в зависимости от метода построения получаемая вероятностная. Так, например, нелинейная замена переменной в двумерной плотности, обладающей линейным уравнением регрессии, приводит к совместному распределению случайных величин уже с нелинейным уравнением регрессии.

Данная работа основана на идеях, многие годы разрабатывавшихся в области теории вероятностей И.О. Сармановым. Основные сложности в реализации предлагаемого подхода были связаны с вычислительными трудностями при использовании распределения Крицкого и Менкеля. Развитие вычислительной техники позволило преодолеть некоторые из них.

ДВУХМЕРНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН С ТРЕХПАРАМЕТРИЧЕСКИМИ РАСПРЕДЕЛЕНИЯМИ ВЕРОЯТНОСТЕЙ

Пусть случайные величины ξ и η распределены по трехпараметрическому закону Крицкого и Менкеля и имеют соответствующие плотности распределения p_1 и p_2 :

$$p_1(x) = \left[\frac{\Gamma(\gamma_1 + b_1)}{\Gamma(\gamma_1)} \right]^{\frac{\gamma_1}{b_1}} \frac{1}{\Gamma(\gamma_1) |b_1| x_0} \left(\frac{x}{x_0} \right)^{\frac{\gamma_1}{b_1} - 1} \times \exp \left\{ - \left[\frac{x}{x_0} \frac{\Gamma(\gamma_1 + b_1)}{\Gamma(\gamma_1)} \right]^{\frac{1}{b_1}} \right\}, \quad x \geq 0, \quad (1)$$

$$p_2(y) = \left[\frac{\Gamma(\gamma_2 + b_2)}{\Gamma(\gamma_2)} \right]^{\frac{\gamma_2}{b_2}} \frac{1}{\Gamma(\gamma_2) |b_2| y_0} \left(\frac{y}{y_0} \right)^{\frac{\gamma_2}{b_2} - 1} \times \exp \left\{ - \left[\frac{y}{y_0} \frac{\Gamma(\gamma_2 + b_2)}{\Gamma(\gamma_2)} \right]^{\frac{1}{b_2}} \right\}, \quad y \geq 0, \quad (2)$$

x_0, y_0 — средние значения; $\gamma_1 > 0, b_1 > 0, \gamma_2 > 0, b_2 > 0$ — параметры распределения. Начальные моменты α_k и β_k соответственно выражены так:

$$\alpha_k = \int_0^\infty x^k p_1(x) dx = \frac{\Gamma(\gamma_1 + b_1 k)}{\Gamma(\gamma_1)} \left[\frac{\Gamma(\gamma_1) x_0}{\Gamma(\gamma_1 + b_1)} \right]^k, \quad (3)$$

$$\beta_k = \int_0^\infty y^k p_2(y) dy = \frac{\Gamma(\gamma_2 + b_2 k)}{\Gamma(\gamma_2)} \left[\frac{\Gamma(\gamma_2) y_0}{\Gamma(\gamma_2 + b_2)} \right]^k. \quad (4)$$

Далее будем рассматривать симметричный случай: $\gamma = \gamma_1 = \gamma_2, b = b_1 = b_2, x_0 = y_0$.

Будем говорить, что между неотрицательными случайными величинами ξ и η существует линей-

ная корреляция, если эти величины имеют совместную плотность распределения

$$p(x, y) = p_1(x) p_2(y) \left[1 + \sum_{k=1}^{\infty} R^k P_k(x) P_k(y) \right] \quad (5)$$

при единственном ограничении — интегрируемости квадрата ядра:

$$\iint_0^\infty \left[\frac{p(x, y)}{\sqrt{p_1(x) p_2(y)}} \right]^2 dx dy < \infty,$$

$p(x, y)$ — неотрицательно определенная функция, $P_k(x)$ и $P_k(y)$ — ортонормированные полиномы, R — коэффициент корреляции между случайными величинами ξ и η . Положительность ядра следует из теоремы [2]: всякое ядро, все собственные значения которого положительны, определено положительно.

Вычисление требуемых ортонормированных полиномов выполняется методом ортогонализации Грама—Шмидта с использованием их явного выражения через моменты весовой функции [5]. В данном случае весовой функцией является трехпараметрическая плотность распределения, описываемая уравнением (1) для полиномов $P_n(x)$ и уравнением (2) для полиномов $P_n(y)$. Таким образом, ортонормированные полиномы $P_n(x)$ определяются следующими формулами:

$$P_n(x) = K_n \begin{vmatrix} \alpha_0 & \alpha_1 & \Lambda & \alpha_n \\ \alpha_1 & \alpha_2 & \Lambda & \alpha_{n+1} \\ \Lambda & \Lambda & \Lambda & \Lambda \\ \alpha_{n-1} & \alpha_n & \Lambda & \alpha_{2n-1} \\ 1 & x & \Lambda & x^n \end{vmatrix}, \quad (6)$$

$$K_n = (D_{n-1} D_n)^{-\frac{1}{2}}, \quad n = 1, 2, \dots,$$

где D_n — определитель Грама:

$$D_n = \begin{vmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \Lambda & \alpha_n \\ \alpha_1 & \alpha_2 & \alpha_3 & \Lambda & \alpha_{n+1} \\ \Lambda & \Lambda & \Lambda & \Lambda & \Lambda \\ \Lambda & \Lambda & \Lambda & \Lambda & \Lambda \\ \alpha_n & \alpha_{n+1} & \alpha_{n+2} & \Lambda & \alpha_{2n} \end{vmatrix},$$

$$D_{n-1} = \begin{vmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \Lambda & \alpha_{n-1} \\ \alpha_1 & \alpha_2 & \alpha_3 & \Lambda & \alpha_n \\ \Lambda & \Lambda & \Lambda & \Lambda & \Lambda \\ \Lambda & \Lambda & \Lambda & \Lambda & \Lambda \\ \alpha_{n-1} & \alpha_n & \alpha_{n+1} & \Lambda & \alpha_{2(n-1)} \end{vmatrix},$$

или в более удобном для вычислений виде:

$$P_n(x) = K_n (A_{n+1, n+1} x^n + A_{n+1, n} x^{n-1} + \Lambda + A_{n+1, 1}), \quad (7)$$

$A_{i,j}$ — соответствующие алгебраические дополнения элементов последней строки определителя в (6). В силу симметричности рассматриваемого случая полиномы $P_n(y)$ определяются аналогично. Построенные таким образом полиномы обладают свойством полноты и замкнутости.

Покажем, что корреляция случайных величин ξ и η , задаваемая формулой (5), линейная, для чего вычислим первый момент условного распределения (условное математическое ожидание) величины η при фиксированном значении $\xi = x$:

$$\bar{y}(x) = M(\eta|x) = \int_0^{\infty} yp(x, y)/p_1(x)dy. \quad (8)$$

С учетом выражения для совместной плотности распределения (5) и в силу свойства ортогональности многочленов $P_k(y)$ формула (8) будет иметь следующий вид:

$$\bar{y}(x) = y_0 + R \frac{(x - \alpha_1)}{\sigma_x} \sigma_y, \quad (9)$$

где σ_x^2 и σ_y^2 — дисперсии случайных величин ξ и η соответственно.

Таким образом, функция регрессии $\bar{y}(x)$ — линейная функция x .

Вычислим моменты условного распределения случайной величины η более высокого порядка. Условный момент k -го порядка m_{0k} выражается так:

$$m_{0k} = \int_0^{\infty} y^k p(x, y)/p_1(x)dy. \quad (10)$$

Любую непрерывную функцию можно разложить в ряд на основе системы ортогональных полиномов. Разложим y^k по ортонормированной системе полиномов $P_k(y)$:

$$y^k = C_{1,k}P_1(y) + C_{2,k}P_2(y) + \Lambda + C_{k,k}P_k(y). \quad (11)$$

Для определения коэффициентов разложения $C_{i,k}$ умножим обе части (11) на $p_2(y)P_i(y)$ и проинтегрируем полученное выражение на интервале $[0, \infty]$. Тогда с учетом построенной ортогональной системы полиномов $P_k(y)$ (7) формула для определения коэффициентов $C_{i,k}$ будет иметь следующий вид:

$$C_{i,k} = K_i \sum_{j=1}^{i+1} A_{i+1,j} \beta_{k+j-1}. \quad (12)$$

С учетом соотношений (11)–(12) выражение (10) для моментов условного распределения k -го порядка преобразуется к виду:

$$m_{0k}(x) = \beta_k + \sum_{i=1}^k C_{i,k} P_i(x) R^i, \quad (13)$$

коэффициенты $C_{i,k}$ находятся по формуле (12).

Таким образом, на основе системы ортогональных функций с весовой функцией в виде трехпараметрического гамма-распределения получена симметричная плотность, которая удовлетворяет уравнению Маркова и характеризуется линейным уравнением регрессии (9).

ВЕРОЯТНОСТНАЯ МОДЕЛЬ МНОГОЛЕТНИХ КОЛЕБАНИЙ УРОВНЯ ОЗЕРА ЧАНЫ

В качестве приложения полученных результатов рассмотрим эффекты, возможные при стохастическом моделировании колебаний уровня воды бессточного оз. Чаны.

Озеро Чаны находится в пределах Обь-Иртышского междуречья и обладает своеобразным водным режимом с внутривековыми циклическими (пульсационными) колебаниями с периодами различной продолжительности. Управление уровнем режима озера требует знания закономерностей формирования его водного баланса в вероятностной форме. Колебания уровня воды h бессточного водоема могут быть описаны с помощью уравнения водного баланса:

$$\frac{dh}{dt} = v(t)/S(h) - e(t), \quad (14)$$

где $v(t)$ — приток воды в единицу времени, $e(t)$ — слой видимого испарения (испарение минус осадки), $S(h)$ — площадь озера. В практических расчетах используется конечно-разностный аналог уравнения (14) и модельные расчеты уровня воды $h(t)$ осуществляются с помощью искусственных рядов притока и испарения заданной продолжительности.

Моделирование искусственных последовательностей компонентов водного баланса озера основывается на схеме простой цепи Маркова, для чего необходимо располагать условным распределением, которое является функцией перехода, определяющей процесс моделирования искусственных реализаций исследуемых процессов:

$$F(y/x) = \int_0^y p_2(y) \left[1 + \sum_{k=1}^{\infty} R^k P_k(x) P_k(y) \right] dy. \quad (15)$$

Для моделирования случайного значения y при заданном x (предыдущее значение процесса) из таблицы равномерно распределенных в интервале $[0, 1]$ случайных чисел берется значение условной функции распределения $F(y/x)$, и y находится как верхний предел интеграла в правой части уравнения (15). Полученное значение y

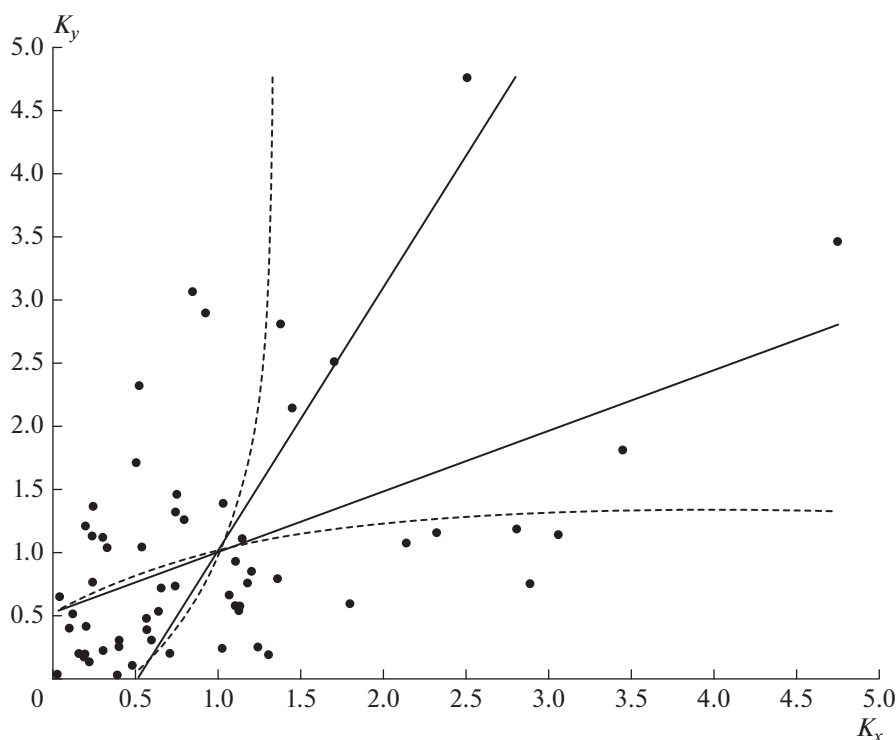
Параметры распределения среднегодовых уровней воды оз. Чаны при различных схемах моделирования притока

Варианты расчета уровня воды	Средний уровень, м БС	Коэффициент асимметрии C_s
Моделирование притока по модели 1 с линейным уравнением регрессии	106.14	0.08
Моделирование притока по модели 2 с нелинейным уравнением регрессии [5]	106.16	-0.32
Последовательность наблюдаемых уровней воды	106.2	0.85

подставляется вместо x , и вся процедура повторяется требуемое число раз, а затем на основе уравнения водного баланса озера рассчитываются последовательности уровней воды в озере. Результаты оценки параметров распределения компонент водного баланса оз. Чаны показывают, что распределение величин годового стока соответствует трехпараметрическому гамма-распределению, а распределения испарения и осадков близки к нормальному закону.

Для моделирования искусственных рядов входного случайного процесса речного притока использовались две стохастические модели: модель 1 – на основе полученной авторами статьи симметричной плотности распределения с линейной регрессией (9) между случайными вели-

чинами, имеющими трехпараметрические гамма-распределения (2); модель 2 – с линейной корреляцией между равномерно распределенными значениями случайных величин (между их обеспеченностями), но с нелинейной регрессией между исходными случайными величинами [4]. Обе модели достаточно точно воспроизводят статистические характеристики моделируемых рядов по отношению к безусловным выборочным характеристикам случайного процесса, но характер условных распределений различается существенно. Результаты моделирования колебаний уровня воды в оз. Чаны по уравнению (14) приведены в таблице, данные которой показывают, что использование линейной корреляции (5) приводит к заметно большей величине асимметрии распределения уровня воды в озере, что лучше соот-



Линии регрессии для корреляции смежных значений притока к оз. Чаны, рассчитанные на основе линейной модели 1 (сплошные линии) и нелинейной модели 2 (пунктир); точки – наблюдаемые значения стока.

ветствует оценке этого параметра, полученной по данным наблюдений.

ЗАКЛЮЧЕНИЕ

Рассмотрен вариант двумерного распределения вероятностей, отражающего специфику вероятностных задач, решаемых в прикладной гидрологии, и задающего линейный марковский процесс. Этот результат позволяет рассматривать широкий круг практических задач, оставаясь в рамках марковского приближения, в рамках корреляционной теории случайных процессов.

Дальнейшая задача исследований в данной области — задание процесса Маркова в форме уравнений Фоккера—Планка—Колмогорова для одномерных трехпараметрических распределений Крицкого и Менкеля, что позволит перейти к вычислению разнообразных вероятностных характеристик процесса стока без трудоемких имитационных расчетов.

Статья посвящена памяти Ибрагима Олеговича Сарманова, крупнейшего специалиста, внесшего серьезный вклад в развитие российской стохастической гидрологии, учителя и старшего товарища.

СПИСОК ЛИТЕРАТУРЫ

1. Болгов М.В., Сарманов И.О., Сарманов О.В. Марковские процессы в гидрологии. М.: ИВП РАН, 2009. 211 с.
2. Пугачев В.С. Теория случайных функций и ее применение к задачам автоматического управления. М.: Физматлит, 1962. 883 с.
3. Пульсирующее озеро Чаны. Л.: Наука, 1982. 304 с.
4. Раткович Д.Я. Стохастическая модель колебаний годового стока рек // Вод. ресурсы. 1972. № 1. С. 52–94.
5. Сегё Г. Ортогональные многочлены. М.: ГИФМЛ, 1962. 500 с.