УДК 556.048

СОВРЕМЕННОЕ СОСТОЯНИЕ ТЕОРИИ КОРРЕЛЯЦИИ ДЛЯ ГИДРОЛОГИЧЕСКИХ РАСЧЕТОВ И СТОХАСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

© 2012 г. М.В. Болгов

Институт водных проблем Российской академии наук, Москва

Ключевые слова: двумерные распределения вероятностей, гидрологические характеристики, марковские процессы.

Обсуждаются методы построения двумерных законов распределения для широкого круга гидрологических задач. Показаны эффекты, возникающие при использовании корреляций различного типа. Рекомендуются методы решения вероятностных задач в гидрологии, экологии и др.

Введение

Многомерные распределения вероятностей возникают в различных гидрологических задачах. При изучении и пространственной изменчивости гидрологических характеристик в общем случае возникает задача построения случайных полей, т. е. совместного распределения случайных величин в зависимости от пространственных координат. При анализе временных случайных процессов необходимо совместное распределение значений стока (осадков, уровней и пр.) в различные моменты времени. Во многих инженерно-гидрологических задачах требуются совместные распределения различных характеристик. Так, например, при задании расчетного ливня для расчета гидрографа по детерминированным моделям стока требуется построить совместный закон распределения продолжительности выпадения осадков, максимальной интенсивности и слоя осадков за дождь. В задачах пропуска паводка при регулировании стока требуется совместное распределение вероятностей сбросов через гидроузел и боковой приточности в нижнем бьефе, а также совместное распределение величин притока к различным водохранилищам каскада (или к узлам водохозяйственной системы).

Наиболее важные задачи стохастической гидрологии связаны с теорией корреляции, т. е. построением совместного распределения двух и более случайных величин. Специфика вероятностных гидрологических задач заключается в том, что основной объект их – величины речного стока – существенно положительны, а распределения их вероятностей – асимметричны. Возникает потребность в разработке моделей, специально приспособленных к анализу асимметричных распределений, а именно задача построения двумерного распределения вероятностей случайных величин, имеющих негауссовские одномерные распределения, такие как трехпараметрическое гамма-распределение Крицкого–Менкеля, распределения Пирсона, обобщенное распределение экстремумов и пр.

Основным методом построения двумерного закона распределения в данном случае является решение уравнения Маркова в виде билинейного разложения по системе ортогональных функций. При этом в качестве весовой функции при построении ортогональных полиномов используются маргинальные (одномерные) распределения вероятностей марковского процесса. Метод позволяет задавать широкий класс линейных (по отношению к уравнению регрессии) марковских процессов с различными одномерными законами распределения.

Значительно расширяет область применения марковской теории преобразование двумерной плотности путем замены переменной в линейных моделях. Распространение получил прием, заключающийся в переходе от корреляции исходных величин к корреляции их функций распределения, а затем к корреляции величин, имеющих нормальные распределения вероятностей. Другая идея заключается в замене переменной в двумерной плотности равномерно-распределенных случайных величин. Известно, что функция распределения есть равномерно-распределенная случайная величина в интервале [0, 1] потому, имея корреляцию равномерно-распределенных случайных величин, достаточно одной замены переменной для перехода к корреляции с произвольным одномерным законом распределения. Такая модель была применена Д.Я. Ратковичем при исследовании многолетних колебаний речного стока [1]. Кроме упомянутых решений, в статистической и гидрологической литературе встречаются и другие подходы к построению двумерной плотности в негауссовском случае.

В последующем были предприняты попытки распространения предложенных схем на несимметричный случай, также появились марковские модели с маргинальными трехпараметрическими распределениями Крицкого–Менкеля.

Построенные таким образом двумерные модели легли в основу решения разнообразных прикладных задач [2, 3]. Рассмотрим основные подходы к построению двумерных распределений, используемые в гидрологических приложениях.

Построение двумерных законов распределения, удовлетворяющих уравнению Маркова

В основе математического описания многолетних колебаний речного стока и многих других гидрометеорологических процессов лежит гипотеза стационарности рассматриваемого случайного процесса. Для стационарного процесса $X(t)=\{x_t\}$ плотность распределения пары случайных величин $x_{t_1}=x$ и $x_{t_2}=y$ зависит только от разности $\tau=t_2-t_1$, т. е. $p(t_1, x; t_2, y)=p(\tau, x, y)$, а $p(t, x)=p(x)$ не зависит от t (другими словами, не зависит от начала отсчета). При всех τ функция $p(\tau, x, y)$ удовлетворяет условию симметрии $p(\tau, x, y)=p(\tau, y, x)$.

Как известно, случайный процесс $X(t)$ обладает марковским свойством, если условное распределение величины $X(t_k)$ при известных значениях величин $X(t_{k-1})$, $X(t_{k-2}),\ldots X(t_{k-m})$, где $t_{k-m}<\ldots<t_{k-2}<t_{k-1}<t_k$ зависит только от значения случайной величины $X(t_{k-1})$ в предшествующий момент времени.

Двумерная плотность стационарного марковского процесса (равно как и условная плотность – плотность вероятности перехода из состояния $x_{t_1}=x$ в состояние $x_{t_2}=y$) не является произвольной функцией, а должна удовлетворять так называемому уравнению Маркова [4, 5]:

$$p(\tau_1+\tau_2, x, y)=\int_{-\infty}^{\infty}\frac{p(\tau_1, x, z)\,p(\tau_2, z, y)}{p(z)}dz, \qquad (1)$$

где $\tau_1=t-t_1$, $\tau_2=t-t_2$.

Одномерные плотности в силу стационарности процесса выражаются одинаковыми функциями

$$p(x)=\int_{-\infty}^{\infty}p(\tau,x,y)dy,\quad p(y)=\int_{-\infty}^{\infty}p(\tau,x,y)dx. \qquad (2)$$

Двумерная плотность распределения $p(\tau, x, y)$ по определению является неотрицательной функцией $p(\tau, x, y)\geq 0$ и удовлетворяет условию $\iint\limits_{\Omega} p(\tau,x,y)\ dxdy=1$. Это позволяет говорить, что функция $p(\tau, x, y)$ определяет корреляционную зависимость между двумя случайными переменными x и y, заданную в прямоугольной области $\Omega=\left[a\leq x\leq b; a\leq y\leq b\right]$.

Интегральное уравнение

$$\varphi(x)=\lambda\int_{a}^{b}\varphi(y)\frac{p(\tau,x,y)}{p(x)}dy, \qquad (3)$$

где $p(\tau, x, y)=p(\tau, y, x)>0$ – плотность распределения двух случайных переменных, называется корреляционным уравнением с симметризуемым ядром [6]. Соответствующее уравнение с симметрическим ядром имеет вид [6]:

$$\omega(x)=\lambda\int_{a}^{b}\omega(y)K(x,y)dy, \qquad (4)$$

где

$$\omega(x)=\varphi(x)\sqrt{p(x)}, \qquad (5)$$

$$K(x,y)=\frac{p(\tau,x,y)}{\sqrt{p(x)p(y)}}. \qquad (6)$$

Если наложить на ядро $K(x, y)$ этого уравнения единственное ограничение интегрируемости квадрата ядра

$$\int_{a}^{b}\int_{a}^{b}\frac{p^2(\tau,x,y)}{p(x)p(y)}dxdy=K^2<\infty, \qquad (7)$$

то на корреляционное уравнение (4) можно распространить все приемы аппарата интегральных уравнений. В этом случае ядро представляется в виде билинейного разложения ряда по ортонормированным собственным функциям.

Для того чтобы симметричная плотность $p(\tau, x, y)$ стационарного случайного процесса удовлетворяла уравнению Маркова (1) с заданными маргинальными распределениями при условии (7), необходимо и достаточно, чтобы она была суммой вида

$$p(\tau,x,y)=p(x)p(y)\left[1+\sum_{k=1}^{\infty}e^{-\lambda_k|\tau|}\varphi_k(x)\varphi_k(y)\right]. \qquad (8)$$

Это утверждение составляет содержание теоремы, доказанной О.В. Сармановым [5] и является фундаментальным результатом, служащим основой для построения двумерных плотностей стационарного марковского процесса.

Гамма-корреляция и ее свойства

Для анализа многолетних колебаний стока важнейшим результатов изложенной выше теории является гамма-корреляция, в основу которой в качестве безусловного закона распределения положено двухпараметрическое гамма-распределение. Блохиновым Е.Г. и Сармановым О.В. [7] были получены основные формулы для двухпараметрического двумерного распределения вероятностей:

$$p(x)=\frac{\gamma^{\gamma}}{\Gamma(\gamma)}x^{\gamma-1}e^{-\gamma x}. \qquad (9)$$

Параметры гамма-распределения в (9): среднее x_0=1, параметр γ связан с коэффициентом изменчивости C_v соотношением $\gamma = {}^{1}\!/_{C_v^2}$. Двумерная плотность в этом случае будет иметь вид

$$p(\tau,x,y)=p(x)p(y)\left[1+\sum_{k=1}^{\infty}R^k\tilde{L}_k^{\gamma-1}(\gamma x)\tilde{L}_k^{\gamma-1}(\gamma y)\right], \qquad (10)$$

где $L_k(\ )$ – k-й полином Лагерра, а соответствующие параметры условного распределения запишутся в виде [7]:

условное среднее:

$$y_0(x)=1+R(x-1), \qquad (11)$$

условная дисперсия:

$$\sigma_{y/x}^2=\sigma^2\left[(1-R)^2+2R(1-R)x\right], \qquad (12)$$

условный коэффициент изменчивости:

$$C_{v_{y/x}} = \frac{C_v\left[(1-R)^2 + 2R(1-R)x\right]^{1/2}}{1+R(x-1)}, \tag{13}$$

Здесь σ^2, C_v – соответствующие безусловные параметры, $0 \le R < 1$ есть первая (при $\tau=1$) ордината корреляционной функции марковского процесса $R(\tau) = e^{-\lambda|\tau|}$.

Из выражения для условного математического ожидания (11) следует, что гамма-корреляция линейна, а из (10), что при $R=0$ величины x и y не только некоррелированы, но и независимы.

Корреляция равномерно-распределенных случайных величин

Двумерная плотность для случайных величин, имеющих равномерное распределение на отрезке [–1, 1] с плотностью $p(x)=1/2$, получается путем разложения в ряд по полиномам Лежандра $P_k(x)$, $k=1, 2,\ldots$, ортонормированным с этим весом на отрезке [–1, 1]:

$$P_k(x) = \frac{1}{k!2^k}\frac{d^k}{dx^k}\left[(x^2-1)^k\right], \tag{14}$$

и определяется следующим выражением [13]

$$p(\tau, x, y) = \frac{1}{4}\left[1 + \sum_{k=1}^{\infty} e^{-k\lambda\tau}\sqrt{2k+1}\,P_k(x)\,\sqrt{2k+1}\,P_k(y)\right]. \tag{15}$$

Важным для теории многолетних колебаний стока случаем является корреляция случайных величин, имеющих равномерное распределение на отрезке [0, 1] с плотностью $p(x)=1$ при ограничении в разложении в ряд первыми тремя полиномами Лежандра.

Для построения двумерного закона распределения равномерно-распределенных случайных величин в квадрате $[0 \le u, v \le 1]$ необходимо в (15) сделать линейную замену переменных $x=2u-1$, $y=2v-1$. Плотность $f(u,v)$ запишется в виде [8]:

$$f(u,v) = 1 + 3\lambda(2u-1)(2v-1) + \frac{5}{4}\lambda^2\left[3(2u-1)^2-1\right]\left[3(2v-1)^2-1\right]. \tag{16}$$

Отметим некоторые характеристики плотности (16). Условная функция распределения записывается в виде:

$$F(u/v) = u + 3\lambda(2v-1)(u^2-u) + \frac{5}{4}\lambda^2\left[3(2v-1)^2-1\right](2u^3-3u^2+u). \tag{17}$$

Дисперсия условного распределения

$$\sigma_u^2 = \frac{1-\lambda^2}{12}. \quad (18)$$

Условная дисперсия не зависит от *v*, и, следовательно, корреляция (16) гомоскедастична.

Линейная корреляция случайных величин с распределениями Крицкого–Менкеля

Ввиду важности для приложений рассмотрим вариант линейной корреляции случайных величин с распределением Крицкого–Менкеля.

Пусть случайная величина ξ имеет трехпараметрическую плотность распределения вероятностей Крицкого–Менкеля [9]:

$$p_1(x) = \left[\frac{\Gamma(\gamma_1 + b_1)}{\Gamma(\gamma_1)}\right]^{\frac{\gamma_1}{b_1}} \frac{1}{\Gamma(\gamma_1)|b_1|x_0} \left(\frac{x}{x_0}\right)^{\frac{\gamma_1}{b_1}-1} exp\left\{-\left[\frac{x}{x_0}\frac{\Gamma(\gamma_1 + b_1)}{\Gamma(\gamma_1)}\right]^{\frac{1}{b_1}}\right\}, \quad x \ge 0, \quad (19)$$

а случайная величина η имеет плотность распределения $p_2(y)$ вида (19) с параметрами с индексом 2. Здесь x_0, y_0 – средние значения, $\gamma_1 > 0$, $b_1 > 0$, $\gamma_2 > 0$, $b_2 > 0$ – параметры распределения. Начальный момент *k*-го порядка α_k равен:

$$\alpha_k = \int_0^\infty x^k p_1(x)dx = \frac{\Gamma(\gamma_1 + b_1 \cdot k)}{\Gamma(\gamma_1)}\left[\frac{\Gamma(\gamma_1)x_0}{\Gamma(\gamma_1 + b_1)}\right]^k, \quad (20)$$

Рассмотрим симметричный случай $\gamma = \gamma_1 = \gamma_2$, $b = b_1 = b_2$, $x_0 = y_0$, следуя работе [10]. Как известно, между неотрицательными случайными величинами ξ и η существует линейная корреляция, если эти величины имеют совместную плотность распределения

$$p(x,y) = p_1(x)p_2(y) \cdot \left[1 + \sum_{k=1}^{\infty} R^k P_k(x)P_k(y)\right], \quad (21)$$

где $p(x, y)$ – неотрицательно определенная функция;

$P_k(x)$ и $P_k(y)$ – ортонормированные полиномы;

R – коэффициент корреляции между случайными величинами ξ и η.

Для вычисления ортонормированных полиномов используется их представление через моменты весовой функции. Согласно [11] $P_n(x)$ и $P_n(y)$ определяются формулами

$$P_n(x)=K_n\begin{vmatrix}\alpha_0 & \alpha_1 & \cdots & \alpha_n\\ \alpha_1 & \alpha_2 & \cdots & \alpha_{n+1}\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{n-1} & \alpha_n & \cdots & \alpha_{2n-1}\\ 1 & x & \cdots & x^n\end{vmatrix},\quad K_n=\left(D_{n-1}D_n\right)^{-\frac{1}{2}}, \tag{22}$$

где D_n является определителем Грама:

$$D_n=\begin{vmatrix}\alpha_0 & \alpha_1 & \alpha_2 & \cdots & \alpha_n\\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{n+1}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \alpha_n & \alpha_{n+1} & \alpha_{n+2} & \cdots & \alpha_{2n}\end{vmatrix},\quad D_{n-1}=\begin{vmatrix}\alpha_0 & \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1}\\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \alpha_{n-1} & \alpha_n & \alpha_{n+1} & \cdots & \alpha_{2(n-1)}\end{vmatrix} \tag{23}$$

Построенные таким образом полиномы обладают свойством полноты и замкнутости. Выпишем выражения для первых двух ортогональных полиномов. Полагая $\alpha_0=1$, $\alpha_1=1$, получим $D_0=1$, $D_1=m_2-m_1^2=\sigma^2$ и первый полином можно представить в следующем виде

$$P_1(x)=\frac{1}{\sigma}(x-1). \tag{24}$$

Второй полином $P_2(x)$:

$$P_2(x)=\left(D_1D_2\right)^{-\frac{1}{2}}\left[x^2\left(\alpha_0\alpha_2-\alpha_1{}^2\right)-x\left(\alpha_0\alpha_3-\alpha_1\alpha_2\right)+\left(\alpha_1\alpha_3-\alpha_2{}^2\right)\right], \tag{25}$$

Соответствующие вычисления показывают [3], что корреляция случайных величин ξ и η, задаваемая формулой (21) линейная, т. е. условное среднее является линейной функцией от *x*. Линейность корреляции (21) является существенным моментом для гидрологических приложений. Так в [12] показано, что при переходе от нелинейной корреляции к линейной для величин речного стока (притока), распределение вероятностей уровней воды в озере, принимающем этот приток, будет характеризоваться существенно большей асимметрией.

Модели корреляции с нелинейным уравнением регрессии

Формы нелинейной корреляции гидрологических величин

Рассмотренные выше примеры двумерных распределений характеризуются линейными уравнениями регрессии. Системы случайных величин могут иметь нелинейные уравнения регрессии, что рассмотрено ниже для ряда случаев. И.О.

Сармановым [13] предложен переход от корреляции равномерно-распределенных величин к корреляции с заданными маргинальными распределениями.

Пусть $f(u, v)$ – двумерная плотность для системы равномерно-распределенных в [0, 1] случайных зависимых величин ξ_1, η_1, определяемая, например, формулой (15). Рассмотрим случайные величины ξ, η, задаваемые равенствами

$$\xi_1 = \int_{\alpha}^{\xi} p(x)dx = F(\xi), \eta_1 = \int_{\alpha}^{\eta} p(y)dy = F(y), \qquad (26)$$

где $p(x)>0$ – заданная плотность распределения в отрезке $[a, b]$. Задача заключается в вычислении плотности распределения ξ и η.

Поскольку $F(\xi)$ и $F(\eta)$ – строго монотонные функции, то существуют обратные функции $F^{-1}(\xi)$ и $F^{-1}(\eta)$ и, следовательно, необходимо изучить совместный закон распределения величин $\xi=F^{-1}(\xi_1)$, $\eta=F^{-1}(\eta_1)$.

Плотность совместного распределения ξ и η, обозначаемая ниже $f_1(x, y)$, в соответствии с правилом замены переменного в двукратном интеграле, записывается в виде:

$$f_1(x, y) = p(x)p(y)f[F(x), F(y)]. \qquad (27)$$

Представим плотность $f(u, v)$ билинейным разложением по полиномам Лежандра, нормированным с весом единица. В таком случае плотность $f_1(x, y)$ определяется в виде:

$$f_1(x, y) = p(x)p(y)\left\{1 + \sum_{k-1}^{\infty} \lambda^k P_k[F(x)] P_k[F(y)]\right\}. \qquad (28)$$

где λ – коэффициент корреляции между $F(\xi)$ и $F(\eta)$;

$P_k(\ )$ – k-ый полином Лежандра.

Поскольку известно выражение для суммы билинейного ряда по полиномам Лежандра, то плотность совместного распределения записывается в виде [5]:

$$f_1(x, y) = \frac{p(x)p(y)}{\pi} \times$$
$$\times \int_0^{\pi} \frac{\left(1-\lambda^2\right)d\omega}{\left\{1 - 2\lambda\left[(2F(x)-1)(2F(y)-1) + 4\sqrt{\left(F(x)-F^2(x)\right)\left(F(y)-F^2(y)\right)}\cos\omega\right] + \lambda^2\right\}^{\frac{3}{2}}}. \qquad (29)$$

Совместное распределение случайных величин, имеющих маргинальные распределения вида Крицкого–Менкеля (19) определяется соотношением, получаемым путем подстановки (19) в (28) или (29).

Преобразование линейной нормальной корреляции в нелинейную корреляцию гамма-распределенных величин

В целом ряде работ, посвященных приложениям гауссовских статистических моделей в гидрометеорологии, преобразование случайных величин с произвольным законом распределения в гауссовские одновременно с предположением о линейном характере регрессии последних рекомендуется как основной метод, позволяющий обойти проблему негауссовости безусловных распределений. Однако последствия такой процедуры, степень искажения свойств модели исходных величин, адекватность модели данным натурных наблюдений, практически не рассматривались.

Исследуемая модель (двумерное распределение) строится следующим образом: от величин, связанных классическим двумерным нормальным законом распределения, мы переходим к значениям их обеспеченностей (или функций распределения этих величин), имеющим равномерное распределение, а от обеспеченностей – к величинам с произвольным законом распределения.

Несмотря на то, что данная идея была предложена достаточно давно, а сам метод широко использовался при проведении имитационных экспериментов, соответствующая стохастическая модель обсуждалась лишь в работах Морана [14], Клемеша и Борювки [15]. Согласно этим работам, двумерный закон распределения случайных величин, имеющих одномерные гамма-распределения, может быть получен следующим образом. Двумерное гауссово распределение случайных величин w и z имеет вид

$$g(w,z,\rho)=\frac{1}{2\pi\left(1-\rho^2\right)^{1/2}}\exp\left[\frac{w^2-2\rho wz+z^2}{2\left(1-\rho^2\right)}\right], \qquad (30)$$

Случайные переменные u и v определим как

$$u=\left(2\pi\ \right)^{-1/2}\int\limits_{\infty}^{w}\exp\left(-\frac{1}{2}t^2di\right)=\Phi(w)$$
$$v=\Phi(z). \qquad (31)$$

Тогда u и v имеют совместный закон распределения с равномерными маргинальными плотностями. Пользуясь правилами замены переменной в двукратном интеграле, запишем формулу для двумерной плотности u и v

$$h(u,v;\rho)=g(w,z,\rho)\frac{dw}{du}\frac{dz}{dv}. \tag{32}$$

Обозначая $w=\Phi^{-1}(u)$ и $z=\Phi^{-1}(v)$ и вычисляя производные, имеем

$$h(u,v;\rho)=\frac{1}{\left(1-\rho^2\right)^{1/2}}\exp\left\{-\frac{\left[\rho^2\left(\Phi^{-1}(u)\right)^2-2\rho\Phi^{-1}(u)\Phi^{-1}(v)+\rho^2\left(\Phi^{-1}(v)\right)^2\right]}{2\left(1-\rho^2\right)}\right\} \tag{33}$$

Если мы теперь имеем гамма-распределенные величины X и Y, то u и v можно определить иначе:

$$\begin{aligned} u &= \frac{1}{m_1^{k_1}\Gamma(k_1)}\int_0^x t^{k_1-1}\exp\left(-\frac{t}{m_1}\right)dz = F(x,m_1,k_1), \\ v &= F(y,m_2,k_2). \end{aligned} \tag{34}$$

Делая далее еще одну замену переменой, получим совместную плотность гамма-распределенных величин:

$$\begin{aligned} f(x,y,\rho,m_1,m_2,k_1,k_2) &= \frac{1}{\left(1-\rho^2\right)^{1/2}}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\rho^2\left[\Phi^{-1}\left(F(x,m_1,k_1)\right)\right]^2 - \right.\right. \\ &\left.\left. -2\rho\Phi^{-1}\left[F(x,m_1,k_1)\right]\Phi^{-1}\left[F(y,m_2,k_2)\right]+\rho^2\left[\Phi^{-1}\left(F(y,m_2,k_2)\right)\right]^2\ \right]\right\}\times \\ &\times\frac{1}{m_1^{k_1}m_2^{k_2}\Gamma(k_1)\Gamma(k_2)}x^{k_1-1}y^{k_2-1}\exp\left[-\left(\frac{x}{m_1}+\frac{y}{m_2}\right)\right] \end{aligned} \tag{35}$$

Аналогично формуле (35) могут быть записаны двумерные плотности для случайных величин с другими произвольными одномерными законами распределения стока, например, для распределения Крицкого–Менкеля.

Как уже отмечалось, данная модель задает нелинейный по отношению к уравнениям регрессии тип стохастической модели, что приводит к уменьшению коэффициента корреляции между гамма-величинами по сравнению с параметром ρ, характеризующим тесноту связи нормальных величин. Соотношения, связывающие коэффициенты корреляции при таких преобразованиях, изучались А.В. Рождественским [16] методом Монте-Карло при произвольном соотношении C_s/C_v.

Копула функции

В последние годы в англоязычной литературе обсуждается метод построения многомерных распределений, основанный на так называемых копула-функциях. В частности, в работах С. Гримальди [17–20] рассматриваются основы теории и возможности их приложений в гидрологии, например, обсуждается совместный закон распределения параметров, характеризующих форму гидрографа стока. Приведем, придерживаясь работы [18], следующую теорему**.**

<u>Теорема.</u> Пусть F является двумерной функцией распределения случайных величин с маргинальными функциями F_{X_1} и F_{X_2}. Тогда, существует копула C такая, что для всех x_1, x_2,$\in R$

$$F(x_1, x_2) = C\lfloor F_{X_1}(x_1),\ F_{X_2}(x_2)\rfloor. \tag{36}$$

Данная теорема позволяет рассматривать функцию распределения как комбинацию двух маргинальных функций и соответствующей копулы.

Среди существующих типов копул, Архимедовы – самый популярный класс, используемый в гидрологических приложениях. Двумерная ассиметричная Архимедова копула может быть определена как:

$$C(u_1, u_2) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2)), \tag{37}$$

где φ – непрерывная функция, называемая генератор, убывающая и выпуклая от I=[0, 1] до [0, φ(0)].

В таблице приведены некоторые двумерные копулы. В этих функциях значения параметра α определяет зависимость между переменными. Для каждой двумерной Архимедовой копулы значения коэффициента корреляции Кендала τ_K, соответствующие значениям α, могут быть получены из определения $\tau_K = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt$.

Допустимые интервалы τ_K представлены в последнем столбце таблицы.

Чтобы определить двумерную функцию распределения, использующую копулу, применяют описанный метод, который состоит из оценки копулы и маргинальных функций. Для оценки параметра α может быть применен полупараметрический подход, называемый каноническим методом максимального правдоподобия.

Таблица. Двумерные Архимедовы копулы Гамбела–Хьюгарда, Клейтона–Парето, Али–Михаила–Хака и Франка [18]

Копула $C(u_1,u_2)$		α∈	$\tau_K = 1+4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt$	$\tau_K \in$
Гамбел-Хьюгард	$\exp\left[-\left((-\ln u_1)^{\alpha}+(-\ln u_2)^{\alpha}\right)^{1/\alpha}\right]$	$[1,\infty)$	$1-\alpha^{-1}$	$(0,1)$
Клейтон-Парето	$\max\left((u_1^{-\alpha}+u_1^{-\alpha}-1)^{-1/\alpha},0\right)$	$[-1,\infty)\setminus\{0\}$	$\frac{\alpha}{\alpha+2}$	$[-1,1)\setminus\{0\}$
Али-Михаил-Хак	$\frac{u_1u_2}{1-\alpha(1-u_1)(1-u_2)}$	$[-1,1)$	$1+4\frac{\frac{-1}{6\alpha}\left((-1+\alpha)^2\ln(1-\alpha)\right)}{6\alpha^2}$	$(-0.181726,\frac{1}{3})$
Франк	$\frac{1}{\alpha}\ln\left(1+\frac{(e^{\alpha u_1}-1)(e^{\alpha u_2}-1)}{e^{\alpha}-1}\right)$	$(-\infty,\infty)\setminus\{0\}$	$1-\frac{4}{\alpha}[D_1(-\alpha)-1]$*	$(-1,1)\setminus\{0\}$

Примечание: $D_1(\alpha)=\frac{1}{\alpha}\int_0^{\alpha}\frac{t}{e^t-1}dt$ и $D_1(-\alpha)=D_1(\alpha)+\frac{\alpha}{2}$.

Плотность двумерного распределения получается путем дифференцирования двумерной функции распределения (36):

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot C_{12}(F_{X_1}(x_1), F_{X_2}(x_2)) \quad (38)$$

где $C_{12}(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C(u_1, u_2)$, $f_{X_1}(x_1)$, $f_{X_2}(x_2)$ – маргинальные плотности распределения и $u_1 = F_{X_1}(x_1)$, $u_1 = F_{X_2}(x_2)$ – маргинальные функции распределения.

Применению копул в гидрологических исследованиях способствуют простота вычислений и возможность обобщения на многомерный случай.

Заключение

Анализ рассматриваемых методов построения двумерных законов распределения вероятностей демонстрирует, что эта задача может быть решена различными методами и единственного решения для заданных одномерных распределений, как показал еще в начале прошлого века французский математик Фреше, не имеет.

Для решения гидрологических задач необходимо обоснование типа корреляции, приемлемого в каждом конкретном случае. Важным является представление о том, что характер корреляционной зависимости и результат решения прикладной задачи, могут существенным образом зависеть от метода построения двумерной плотности. Метод же решения прикладной вероятностной задачи может выбираться, исходя из удобства выполнения соответствующих вычислений. В зависимости от типа прикладной задачи могут использоваться, например, диффузионные уравнения Фоккера–Планка–Колмогорова для оценки достижения процессом заданного уровня и вероятности нахождения его за этими пределами. Имитационные алгоритмы, позволяющие генерировать продолжительные ряды стока, целесообразно разрабатывать на основе явных выражений для двумерных законов распределения.

СПИСОК ЛИТЕРАТУРЫ

1. *Раткович Д.Я.* Многолетние колебания речного стока. Л.: Гидрометеоиздат, 1976. 256 с.

2. *Болгов М.В.* Марковские процессы в задаче прогнозирования уровня замкнутого водоема // Метеорология и гидрология. 2005. № 11. С. 74–85.
3. *Болгов М.В., Сарманов И.О., Сарманов О.В.* Марковские процессы в гидрологии. М. 2009. 211 с.
4. *Коваленко И.Н., Сарманов О.В.* Краткий курс теории случайных процессов. Киев: Вища школа, 1978. 262 с.
5. *Сарманов О.В.* Исследование стационарных Марковских процессов методом разложения по собственным функциям // Труды МИАН. 1961. Т. 60. С. 238–261.
6. *Сарманов О.В.* О монотонных решениях корреляционных интегральных уравнений // ДАН СССР. 1946. Т. LIII. № 9. С. 781–784.
7. *Блохинов Е.Г., Сарманов О.В.* Гамма-корреляция и ее использование при расчетах многолетнего регулирования стока // Труды ГГИ. 1968. Вып. 143. С. 52–75.
8. *Сарманов И.О.* Построение корреляции между равномерно-распределенными случайными величинами // Труды ГГИ. 1968. Вып. 160. С. 81–89.
9. *Крицкий С.Н., Менкель М.Ф.* Гидрологические основы речной гидротехники. М.–Л.: Изд-во АН СССР, 1950. 391 с.
10. *Сарманов И.О., Болгов М.В.* Несимметричная линейная корреляция между величинами, имеющими трехпараметрическое гамма-распределение Крицкого–Менкеля // Тезисы докладов VI Всеросс. гидролог. съезда, 28 сентября-1 октября 2004 г., Санкт-Петербург. Секция 5. СПб: Гидрометеоиздат, 2004. С. 185–187.
11. *Сегё Г.* Ортогональные многочлены (перевод с англ.). М: ГИФМЛ, 1962. 500 с.
12. *Болгов М.В., Коробкина Е.А.* О моделировании колебаний уровня о. Чаны для управления его гидрологическим режимом // Водное хозяйство России. 2012. № 1. С. 4–22.
13. *Сарманов И.О.* О корреляции между функциями зависимых случайных величин, имеющих гидрологическое приложение // Проблемы регулирования и использования водных ресурсов. М.: Наука, 1973. С. 87–103.
14. *Moran P.A.P.* Statistical inference with bivariate gamma-distributions // Biometrica. 1969. Vol. 56. № 3. P. 627–634.
15. *Klemes V., Boruvka L.* Simulation of gamma-distributed first-order Markov Chain // Water Resources Research. 1974. Vol. 10. № 1. P. 87–91.
16. *Рождественский А.В.* Оценка точности кривых распределения гидрологических характеристик. Л.:Гидрометеоиздат, 1977. 270 с.

17. *Serinaldi F., Grimaldi S., Napolitano F., Ubertini L.* A 3-copula function application for design hyetograph analysis / edited by Savic D.A., Mariño M.A., Savenije H.G., Bertoni J.C. // Sustainable Water Management Solutions for Large Cities. 2005. IAHS-AISH Publication (293). P. 203–211.
18. *Grimaldi S., Serinaldi F.* Design hyetograph analysis with 3-copula function // Hydrological Science Journal. 2006. 51(2). P. 223–238.
19. *Grimaldi S., Serinaldi F.* Asymmetric copula in multivariate flood frequency analysis // Advances in Water Resources. 2006. 29 (8). P. 1155–1167.
20. *Serinaldi F., Grimaldi S.* Fully nested 3-copula: procedure and application on hydrologic data // Journal of Hydrologic Engineering 2007. 12 (4). P. 420–430.

Сведения об авторе:

Болгов Михаил Васильевич, д. т. н., заведующий лабораторией, Институт водных проблем Российской академии наук, 119333, Москва, ул. Губкина, 3; e-mail: bolgovmv@mail.ru