

УДК 519.772

С.В. Доценко, проф., д-р физ.-мат. наук

Севастопольский национальный технический университет

Студгородок, г. Севастополь, Украина, 99053

E-mail: root@sevgtu.sebastopol.ua

ЭНТРОПИЯ И ВЗАИМНАЯ ИНФОРМАЦИЯ СИСТЕМ НЕГАУССОВСКИХ СЛУЧАЙНЫХ ВЕЛИЧИН

Найдены общие выражения для определения энтропии и взаимной информации систем непрерывных случайных величин, теоретическая модель которых разработана в статьях [1] и [3]. Приведены конкретные примеры.

Общепринятыми методами исследования случайных процессов, протекающих в информационных системах, являются корреляционный и спектральный анализ. Однако эти методы дают полное описание процессов только в том случае, когда последние имеют гауссовское (нормальное) распределение плотностей вероятностей. Если процессы негауссовские или гауссовские процессы в системе подвергаются нелинейным преобразованиям, то при указанных методах анализа часть информации о процессе необратимо утрачивается. Определение количества информации, содержащейся в негауссовских процессах и его потерь, имеющих место при различных преобразованиях, возможно на основе теории информации К. Шеннона. Тем не менее, в литературных источниках фактически отсутствуют работы, связанные с этой проблемой. Данная статья в определенной степени восполняет указанный пробел.

В теоретических исследованиях и практических приложениях нередко встречаются многомерные непрерывные центрированные случайные величины вида $\vec{\xi}_N^T = (\xi_1, \dots, \xi_N)$, N составляющих которых статистически связаны между собой. В частности, это могут быть множества значений случайного процесса, взятых в N точках отсчета. Допустим, что N -мерная плотность распределения вероятностей такой величины имеет вид [1]

$$p_N(\vec{x}_N) = \frac{1}{C_N} f_N\left(\sqrt{\vec{x}_N^T A_N \vec{x}_N}\right), \quad (1)$$

где $f_N(t)$ – неотрицательная, интегрируемая в бесконечных пределах базовая функция; A_N – квадратная симметрическая положительно определенная матрица размера $N \times N$ и нормирующий коэффициент

$$C_N = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N\left(\sqrt{\vec{x}_N^T A_N \vec{x}_N}\right) d\vec{x}_N \quad (2)$$

зависит от вида функции $f_N(t)$, числа составляющих N и матрицы A_N . Согласно [1], он выражается формулой

$$C_N = \frac{2\pi^{N/2} v_{N-1}(N)}{\Gamma(N/2) \sqrt{\det A_N}}, \quad (3)$$

в которой

$$v_n(N) = \int_0^\infty t^n f_N(t) dt \quad (4)$$

– начальный момент n -го порядка функции $f_N(t)$ в пределах полуоси $0 \leq t \leq \infty$. Определитель матрицы A_N равен

$$\det A_N = (q_N^N \sigma_1^2 \sigma_2^2 \dots \sigma_N^2 \det R_N)^{-1}, \quad (5)$$

где

$$R_N = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1N} \\ \rho_{21} & 1 & \dots & \rho_{2N} \\ \dots & \dots & \dots & \dots \\ \rho_{N1} & \rho_{N2} & \dots & 1 \end{bmatrix}$$

– квадратная симметрическая матрица коэффициентов корреляции между всеми парами составляющих величины $\vec{\xi}_N$ и через $\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2$ обозначены дисперсии этих составляющих. $q_N = N v_{N-1}(N) / v_{N+1}(N)$ – коэффициент, определяемый моментами (4).

От центрированных случайных величин в формуле (1) просто перейти к нецентрированным [1]. Однако в силу независимости энтропии от среднего значения случайных величин для наших целей в этом нет необходимости.

Для нахождения взаимной информации двух систем случайных величин следует найти энтропию их совокупности и энтропии каждой из этих систем.

Вначале вычислим совместную дифференциальную энтропию случайного вектора $\vec{\xi}_N$, т.е. совокупности всех его случайных составляющих. Ее величина в натуральных единицах определяется как [2]

$$H_N(\vec{\xi}_N) = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_N(\vec{x}_N) \ln p_N(\vec{x}_N) d\vec{x}_N. \quad (6)$$

Для перехода к значениям энтропии и взаимной информации в битах следует полученные результаты разделить на $\ln 2$, т.е. умножить на коэффициент 1,44270.

Согласно представлению (1) имеем

$$\ln p_N(\vec{x}_N) = \ln f_N(\sqrt{\vec{x}_N^T A_N \vec{x}_N}) - \ln C_N.$$

Подставляя эту сумму и (1) в формулу (6), находим

$$H_N(\vec{\zeta}_N) = \frac{\tilde{C}_N}{C_N} + \ln C_N, \quad (7)$$

где обозначено

$$\tilde{C}_N = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \tilde{f}_N \left(\sqrt{\vec{x}_N^T A_N \vec{x}_N} \right) d\vec{x}_N. \quad (8)$$

Подынтегральная функция

$$\tilde{f}_N(t) = -f_N(t) \ln f_N(t), \quad (9)$$

входящая в последнее выражение, определяется только базовой функцией $f_N(t)$.

Преобразуем N -мерный интеграл (8) в одномерный. Он отличается от интеграла (2) тем, что в него вместо $f_N \left(\sqrt{\vec{x}_N^T A_N \vec{x}_N} \right)$ входит функция $\tilde{f}_N \left(\sqrt{\vec{x}_N^T A_N \vec{x}_N} \right)$. Поэтому по аналогии с формулой (3) получаем

$$\tilde{C}_N = \frac{2\pi^{N/2} \tilde{v}_{N-1}(N)}{\Gamma(N/2) \sqrt{\det A_N}}. \quad (10)$$

Согласно соотношению (9), входящая сюда величина

$$\tilde{v}_n(N) = \int_0^{\infty} t^n \tilde{f}_N(t) dt = - \int_0^{\infty} t^n f_N(t) \ln f_N(t) dt \quad (11)$$

есть начальный момент n -го порядка функции $\tilde{f}_N(t)$. С учетом изложенного первое слагаемое формулы (7) запишем как

$$\frac{\tilde{C}_N}{C_N} = \frac{\tilde{v}_{N-1}(N)}{v_{N-1}(N)}. \quad (12)$$

Учитывая выражения (3), (5) и (12), энтропию (7) можно представить в виде суммы:

$$H_N(\vec{\zeta}_N) = H_N^{(1)}(\vec{\zeta}_N) + H_N^{(2)}(\vec{\zeta}_N) + H_N^{(3)}(\vec{\zeta}_N). \quad (13)$$

В ней первое слагаемое

$$H_N^{(1)}(\vec{\zeta}_N) = \frac{1}{2} \ln \det R_N$$

определяется только совокупностью коэффициентов корреляции между компонентами случайного вектора $\vec{\zeta}_N$, второе слагаемое

$$H_N^{(2)}(\vec{\zeta}_N) = \frac{N}{2} \ln(\pi q_N) + \ln \frac{2v_{N-1}(N)}{\Gamma(N/2)} + \frac{\tilde{v}_{N-1}(N)}{v_{N-1}(N)}$$

зависит только от формы базовой функции $f_N(t)$ рассматриваемого распределения и числа составляющих N вектора $\vec{\zeta}_N$ и третье слагаемое

$$H_N^{(3)}(\vec{\xi}_N) = \frac{1}{2} \ln(\sigma_1^2 \sigma_2^2 \dots \sigma_N^2)$$

зависит только от совокупности дисперсий всех составляющих этого вектора.

Теперь найдем энтропию части совокупности рассматриваемых случайных величин. Согласно [3], частичная (маргинальная) $(N-m)$ -мерная плотность вероятностей, соответствующая (1), может быть записана в виде:

$$p_{N-m}(\vec{x}_{N-m}) = \frac{1}{C_{N-m}} f_{N-m} \left(\sqrt{\vec{x}_{N-m}^T A_{N-m} \vec{x}_{N-m}} \right).$$

Базовая функция этого $(N-m)$ -мерного распределения связана с базовой функцией $f_N(t)$ исходного N -мерного распределения соотношением

$$f_{N-m}(t) = \int_0^\infty x^{m-1} f_N \left(\sqrt{t^2 + x^2} \right) dx. \quad (14)$$

Его нормирующий коэффициент

$$C_{N-m} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{N-m} \left(\sqrt{\vec{x}_{N-m}^T A_{N-m} \vec{x}_{N-m}} \right) d\vec{x}_{N-m}$$

выражается формулой

$$C_{N-m} = \frac{2\pi^{(N-m)/2} v_{N-m-1}(N-m)}{\Gamma\left(\frac{N-m}{2}\right) \sqrt{\det A_{N-m}}},$$

которая может быть получена из (3) заменой N на $(N-m)$. Здесь начальные моменты даются выражением, аналогичным (4),

$$v_n(N-m) = \int_0^\infty t^n f_{N-m}(t) dt. \quad (15)$$

Учитывая соотношение (14) и переходя к полярным координатам, из (15) получаем

$$v_n(N-m) = \frac{1}{2} B\left(\frac{m}{2}, \frac{n+1}{2}\right) v_{n+m}(N), \quad (16)$$

где $B(x,y)$ – бета-функция (эйлеров интеграл 1-го рода).

Совместная дифференциальная энтропия случайного вектора $\vec{\xi}_{N-m}$, имеющего $(N-m)$ составляющих, может быть найдена по формуле, аналогичной (6)

$$H_{N-m}(\vec{\xi}_{N-m}) = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_{N-m}(\vec{x}_{N-m}) \ln p_{N-m}(\vec{x}_{N-m}) d\vec{x}_{N-m} \quad (17)$$

ИЛИ

$$H_{N-m}(\vec{\xi}_{N-m}) = \frac{\tilde{C}_{N-m}}{C_{N-m}} + \ln C_{N-m}, \quad (18)$$

где аналогично (8)

$$\tilde{C}_{N-m} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \tilde{f}_{N-m} \left(\sqrt{\vec{x}_{N-m}^T A_{N-m} \vec{x}_{N-m}} \right) d\vec{x}_{N-m} \quad (19)$$

и по аналогии с (9) обозначено

$$\tilde{f}_{N-m}(t) = -f_{N-m}(t) \ln f_{N-m}(t). \quad (20)$$

Входящая в (19) матрица A_{N-m} определяется корреляционными связями только тех составляющих вектора $\vec{\xi}_N$, которые содержатся в $\vec{\xi}_{N-m}$.

Подобно формуле (10) имеем

$$\tilde{C}_{N-m} = \frac{2\pi^{\frac{N-m}{2}} \tilde{v}_{N-m-1}(N-m)}{\Gamma\left(\frac{N-m}{2}\right) \sqrt{\det A_{N-m}}},$$

где начальные моменты определяются как

$$\tilde{v}_n(N-m) = \int_0^{\infty} t^n \tilde{f}_{N-m}(t) dt = - \int_0^{\infty} t^n f_{N-m}(t) \ln f_{N-m}(t) dt. \quad (21)$$

Поэтому первое слагаемое выражения (18) можно записать в виде:

$$\frac{\tilde{C}_{N-m}}{C_{N-m}} = \frac{\tilde{v}_{N-m-1}(N-m)}{v_{N-m-1}(N-m)}.$$

Учитывая равенство (16) и соотношение [3]

$$\det A_{N-m} = \left(q_N^{N-m} \sigma_1^2 \sigma_2^2 \dots \sigma_{N-m}^2 \det R_{N-m} \right)^{-1},$$

где коэффициент q_N тот же, что и в выражении (5), и равенство (1), получаем второе слагаемое формулы (18):

$$\ln C_{N-m} = \ln \left(\frac{\Gamma(m/2) v_{N-1}(N)}{\Gamma(N/2)} \right) + \frac{N-m}{2} \ln(\pi q_N) + \ln(\sigma_1 \sigma_2 \dots \sigma_{N-m}) + \frac{1}{2} \ln \det R_{N-m}.$$

Отсюда следует, что энтропия (17) так же, как и (13), может быть представлена суммой трех слагаемых

$$H_{N-m}(\vec{\xi}_{N-m}) = H_{N-m}^{(1)}(\vec{\xi}_{N-m}) + H_{N-m}^{(2)}(\vec{\xi}_{N-m}) + H_{N-m}^{(3)}(\vec{\xi}_{N-m}), \quad (22)$$

которые имеют вид:

$$H_{N-m}^{(1)}(\vec{\xi}_{N-m}) = \frac{1}{2} \ln \det R_{N-m},$$

$$H_{N-m}^{(2)}(\vec{\zeta}_{N-m}) = \frac{N-m}{2} \ln(\pi q_N) + \ln \left(\frac{\Gamma(m/2) v_{N-1}(N)}{\Gamma(N/2)} \right) \frac{2v_{N-m-1}(N-m)}{\Gamma\left(\frac{N-m}{2}\right)} + \frac{\tilde{v}_{N-m-1}(N-m)}{v_{N-m-1}(N-m)}$$

и

$$H_{N-m}^{(3)}(\vec{\zeta}_{N-m}) = \frac{1}{2} \ln(\sigma_1^2 \sigma_2^2 \dots \sigma_{N-m}^2).$$

Применим найденные выражения для определения взаимной информации между системами случайных величин.

Допустим, что имеются две многомерных случайных величины: L -мерная $\vec{\xi}_L^T = (\xi_1, \dots, \xi_L)$ и M -мерная $\vec{\eta}_M^T = (\eta_1, \dots, \eta_M)$, компоненты которых в общем случае статистически связаны как между собой, так и друг с другом. Совокупность этих случайных величин можно рассматривать как N -мерную случайную величину $\vec{\zeta}_N^T = (\zeta_1, \zeta_2, \dots, \zeta_N) = (\vec{\xi}_L^T, \vec{\eta}_M^T)$, где $N=L+M$. Ее первые L составляющих совпадают с компонентами вектора $\vec{\zeta}_N^T$: $\zeta_1 = \xi_1, \zeta_2 = \xi_2, \dots, \zeta_L = \xi_L$, а остальные – с компонентами вектора $\vec{\eta}_M^T$: $\zeta_{L+1} = \eta_1, \zeta_{L+2} = \eta_2, \dots, \zeta_{L+M} = \eta_M$. Будем считать, что совокупная величина $\vec{\zeta}_N$ описывается N -мерной плотностью распределения вероятностей (1), матрица A_N которой выражается через корреляционную матрицу R_N [1]. В свою очередь, последняя может быть представлена как положительно определенная блочная матрица

$$R_N = \begin{bmatrix} R_L & R_{LM} \\ R_{ML} & R_M \end{bmatrix}.$$

Входящая в нее квадратная положительно определенная матрица R_L размера $L \times L$ имеет своими элементами коэффициенты корреляции между компонентами ξ_i и ξ_j вектора $\vec{\xi}_L$, квадратная положительно определенная матрица R_M размера $M \times M$ – коэффициенты корреляции между компонентами η_i и η_j вектора $\vec{\eta}_M$, прямоугольная матрица R_{LM} размера $L \times M$ – коэффициенты корреляции между компонентами ξ_i вектора $\vec{\xi}_L$ и компонентами η_j вектора $\vec{\eta}_M$, а матрица R_{ML} транспонирована последней, т.е. $R_{ML} = R_{LM}^T$.

Найдем среднюю взаимную информацию между L -мерной случайной величиной $\vec{\xi}_L$ и M -мерной случайной величиной $\vec{\eta}_M$, в совокупности образующими указанную выше $L+M$ -мерную случайную величину $\vec{\zeta}_N$. Эта информация равна [2]

$$I(\vec{\xi}_L, \vec{\eta}_M) = H_L(\vec{\xi}_L) + H_M(\vec{\eta}_M) - H_N(\vec{\zeta}_N)$$

или

$$I(\vec{\xi}_L, \vec{\eta}_M) = H_{N-M}(\vec{\xi}_L) + H_{N-L}(\vec{\eta}_M) - H_N(\vec{\xi}_N).$$

Учитывая представления (13) и (22), запишем взаимную информацию в виде суммы

$$I(\vec{\xi}_L, \vec{\eta}_M) = I^{(1)}(\vec{\xi}_L, \vec{\eta}_M) + I^{(2)}(\vec{\xi}_L, \vec{\eta}_M) + I^{(3)}(\vec{\xi}_L, \vec{\eta}_M). \quad (23)$$

Вычислим входящие в нее слагаемые:

$$I^{(1)}(\vec{\xi}_L, \vec{\eta}_M) = H_{N-M}^{(1)}(\vec{\xi}_L) + H_{N-L}^{(1)}(\vec{\eta}_M) - H_N^{(1)}(\vec{\xi}_N) = \frac{1}{2} \ln \frac{\det R_L(\vec{\xi}_L) \det R_M(\vec{\eta}_M)}{\det R_N(\vec{\xi}_N)}.$$

Далее,

$$\begin{aligned} I^{(2)}(\vec{\xi}_L, \vec{\eta}_M) &= H_{N-M}^{(2)}(\vec{\xi}_L) + H_{N-L}^{(2)}(\vec{\eta}_M) - H_N^{(2)}(\vec{\xi}_N) = \\ &= \ln(Q_N(L, M)) + \frac{\tilde{v}_{L-1}(N-M)}{v_{L-1}(N-M)} + \frac{\tilde{v}_{M-1}(N-L)}{v_{M-1}(N-L)} - \frac{\tilde{v}_{N-1}(N)}{v_{N-1}(N)}, \end{aligned}$$

где обозначено

$$Q_N(L, M) = \frac{1}{2} B(L/2, M/2) v_{N-1}(N). \quad (24)$$

Наконец,

$$I^{(3)}(\vec{\xi}_L, \vec{\eta}_M) = H_{N-M}^{(3)}(\vec{\xi}_L) + H_{N-L}^{(3)}(\vec{\eta}_M) - H_N^{(3)}(\vec{\xi}_N) = 0.$$

Итак, средняя взаимная информация определяется выражением, не зависящим от математических ожиданий и дисперсий составляющих случайных величин

$$I(\vec{\xi}_L, \vec{\eta}_M) = I^{(1)}(\vec{\xi}_L, \vec{\eta}_M) + I^{(2)}(\vec{\xi}_L, \vec{\eta}_M). \quad (25)$$

В частности, можно показать, что для гауссовского распределения формула (23) сводится к одному слагаемому

$$I_{\text{гаусс}}(\vec{\xi}_L, \vec{\eta}_M) = \frac{1}{2} \ln \frac{\det R_L(\vec{\xi}_L) \det R_M(\vec{\eta}_M)}{\det R_N(\vec{\xi}_N)}, \quad (26)$$

которое зависит только от корреляционных матриц. Этот результат впервые найден в [4].

В общем случае выражение (25) можно записать как

$$I(\vec{\xi}_L, \vec{\eta}_M) = I_{\text{гаусс}}(\vec{\xi}_L, \vec{\eta}_M) + \Delta I(\vec{\xi}_L, \vec{\eta}_M),$$

где

$$\Delta I(\vec{\xi}_L, \vec{\eta}_M) = \ln(Q_N(L, M)) + \frac{\tilde{v}_{L-1}(N-M)}{v_{L-1}(N-M)} + \frac{\tilde{v}_{M-1}(N-L)}{v_{M-1}(N-L)} - \frac{\tilde{v}_{N-1}(N)}{v_{N-1}(N)} \quad (27)$$

— дополнительное к гауссовскому слагаемое, не зависящее от корреляционных матриц и обусловленное негауссовским характером распределения.

Известно [4], что средняя взаимная информация для любых распреде-

лений неотрицательна: $I(\tilde{\xi}_L, \bar{\eta}_M) \geq 0$. Если система случайных величин $\tilde{\xi}_L$ не коррелирует с системой $\bar{\eta}_M$, то величина (26) обращается в нуль. Поэтому для распределений любого вида должно выполняться условие $\Delta I(\tilde{\xi}_L, \bar{\eta}_M) \geq 0$. Отсюда следует, что системы случайных величин, распределенных по нормальному закону, несут минимальное количество взаимной информации среди всех возможных распределений, описываемых формулой (1), и это количество определяется только корреляционными связями между компонентами данных систем. Случайные величины, имеющие любые иные распределения вида (1), более информативны, и степень этой дополнительной информативности зависит от формы их распределения. Гауссовские величины содержат в себе только линейные взаимные связи. Все остальные величины в дополнение к линейным имеют еще и нелинейные связи. Последние и дают приращение взаимной информации. Чем значительнее форма распределения отличается от гауссовской, тем большую величину имеет это приращение.

Рассмотрим два примера вычисления взаимной информации. В основе первого из них лежит N -мерное распределение Стюдента с k степенями свободы. Для него N -мерная базовая функция имеет вид [1]

$$f_N(t) = (1 + t^2)^{-(N+k)/2}, \quad (-\infty < t < \infty, \quad k > 2) \quad (28)$$

с начальными моментами (4) n -го порядка

$$v_n(N) = S_N(k, n+1), \quad (n < N + k - 1), \quad (29)$$

где обозначен коэффициент

$$S_N(l, r) = \frac{1}{2} B\left(\frac{N+l-r}{2}, \frac{r}{2}\right). \quad (30)$$

Особо отметим, что область определения функции $f_N(t)$ здесь бесконечна.

Подставляя (28) в (14), находим $N-m$ -мерные базовые функции этого распределения

$$f_{N-m}(t) = S_N(k, m) (1 + t^2)^{-\frac{N+k-m}{2}}, \quad (m < N + k) \quad (31)$$

и их начальные моменты

$$v_n(N-m) = S_N(k, m) S_N(k-m, n+1), \quad (n < N + k - m - 1). \quad (32)$$

Подставив функцию (28) в (11), произведя замену переменной $x = t^2$ и учитывая значение интеграла

$$\int_0^\infty x^{\mu-1} (1+x)^{-\nu} \ln(1+x) dx = B(\mu, \nu-\mu) [\psi(\nu) - \psi(\nu-\mu)], \quad (0 < \mu < \nu),$$

где $\psi(x)$ – пси-функция Эйлера, находим моменты функции (9):

$$\tilde{v}_n(N) = \frac{N+k}{2} S_N(k, n+1) \left[\Psi \left(\frac{N+k}{2} \right) - \Psi \left(\frac{N+k-n-1}{2} \right) \right]. \quad (33)$$

Далее вычислим моменты (21) функции (20). Из (31) имеем

$$\ln f_{N-m}(t) = \ln S_N(k, m) - \frac{N+k-m}{2} \ln(1+t^2)$$

и

$$\tilde{f}_{N-m}(t) = \frac{N+k-m}{2} S_N(k, m) (1+t^2)^{-\frac{N+k-m}{2}} \ln(1+t^2) - S_N(k, m) \ln S_N(k, m) (1+t^2)^{-\frac{N+k-m}{2}}.$$

Учитывая приведенные ранее формулы, при $m < N+k$ получаем

$$\begin{aligned} \tilde{v}_n(N-m) &= \\ &= S_N(k, m) S_N(k-m, n+1) \left\{ \frac{N+k-m}{2} \left[\Psi \left(\frac{N+k-m}{2} \right) - \Psi \left(\frac{N+k-m-n-1}{2} \right) \right] - \ln S_N(k, m) \right\}. \end{aligned} \quad (34)$$

В силу изложенного и соотношений (29) и (30) определяем коэффициент (24):

$$Q_N(L, M) = \frac{1}{4} B \left(\frac{L}{2}, \frac{M}{2} \right) B \left(\frac{k}{2}, \frac{N}{2} \right).$$

Далее согласно (29) и (33) имеем $v_{N-1}(N) = S_N(k, N)$ и

$$\tilde{v}_{N-1}(N) = \frac{N+k}{2} S_N(k, N) \left[\Psi \left(\frac{N+k}{2} \right) - \Psi \left(\frac{k}{2} \right) \right].$$

Поэтому

$$\frac{\tilde{v}_{N-1}(N)}{v_{N-1}(N)} = \frac{N+k}{2} \left[\Psi \left(\frac{N+k}{2} \right) - \Psi \left(\frac{k}{2} \right) \right].$$

В свою очередь, из формул (32) и (34) находим:

$$\left. \begin{aligned} \frac{\tilde{v}_{L-1}(N-M)}{v_{L-1}(N-M)} &= \frac{L+k}{2} \left[\Psi \left(\frac{L+k}{2} \right) - \Psi \left(\frac{k}{2} \right) \right] - \ln S_N(k, M), \\ \frac{\tilde{v}_{M-1}(N-L)}{v_{M-1}(N-L)} &= \frac{M+k}{2} \left[\Psi \left(\frac{M+k}{2} \right) - \Psi \left(\frac{k}{2} \right) \right] - \ln S_N(k, L). \end{aligned} \right\}$$

Подстановка полученных выражений в (27) дает негауссовскую составляющую искомой взаимной информации:

$$\begin{aligned} \Delta I(\tilde{\xi}_L, \tilde{\eta}_M) &= \ln \left[\frac{\Gamma \left(\frac{N+k}{2} \right) \Gamma \left(\frac{k}{2} \right)}{\Gamma \left(\frac{L+k}{2} \right) \Gamma \left(\frac{M+k}{2} \right)} \right] + \frac{L+k}{2} \Psi \left(\frac{L+k}{2} \right) + \frac{M+k}{2} \Psi \left(\frac{M+k}{2} \right) - \\ &\quad - \frac{N+k}{2} \Psi \left(\frac{N+k}{2} \right) - \frac{k}{2} \Psi \left(\frac{k}{2} \right). \end{aligned}$$

Значения этой величины для $M=1$ и различных L и k представлены в таблице 1.

Таблица 1 – Значения негауссовской составляющей взаимной информации многомерного распределения Стьюдента

k L	2	3	4	5	6	7	...	∞
1	0,0827	0,0424	0,0256	0,0171	0,0122	0,0092	...	0
2	0,1251	0,0681	0,0428	0,0294	0,0214	0,0163	...	0
3	0,1507	0,0852	0,0550	0,0385	0,0285	0,0220	...	0
4	0,1679	0,0974	0,0642	0,0457	0,0342	0,0266	...	0
5	0,1801	0,1066	0,0713	0,0513	0,0389	0,0305	...	0
...	...	...	...	...	...	...	...	0
∞	0,2527	0,1701	0,1276	0,1020	0,0849	0,0726	...	0

Из этой таблицы видно, что, как и следовало ожидать, негауссовская составляющая взаимной информации неотрицательна при $N=2$, имеет максимальное значение, если $k=3$, и убывает с ростом k , стремясь к нулю. Последнее обстоятельство связано с тем, что при $k \rightarrow \infty$ распределение Стьюдента стремится к гауссовскому (нормальному). Величина ΔI здесь невелика, что также определяется близостью распределения Стьюдента к нормальному.

Рассмотрим второй пример – многомерное цилиндрическое распределение [1]. Его порождает финитная прямоугольная базовая функция

$$f_N(t) = \begin{cases} 1 & \text{при } |t| \leq 1, \\ 0 & \text{при } |t| > 1, \end{cases} \quad (35)$$

начальные моменты которой

$$v_n(N) = (n+1)^{-1} \quad (36)$$

и N -мерная плотность распределения случайных величин

$$p_N(\vec{x}_N) = \begin{cases} \frac{\Gamma(N/2+1)}{\sqrt{\pi}^N (N+2)^N \det R_N \sigma_1 \dots \sigma_N} & \text{при } |l_N| \leq \sqrt{N+2}, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Функция (14) здесь имеет вид:

$$f_{N-m}(t) = \begin{cases} \frac{1}{m} (1-t^2)^{\frac{m}{2}} & \text{при } |t| \leq 1, \\ 0 & \text{при } |t| > 1. \end{cases}$$

Ее моменты (15) согласно (16) и (36) равны

$$v_n(N-m) = \frac{1}{2(m+n+1)} B\left(\frac{m}{2}, \frac{n+1}{2}\right). \quad (37)$$

Подстановка (35) в (9) дает $\tilde{f}_N(t) = 0$, а следовательно и

$$\tilde{v}_n(N) = 0. \quad (38)$$

В свою очередь, комбинируя (36) и (19), получаем

$$\tilde{f}_{N-m}(t) = \frac{\ln m}{m} (1-t^2)^{\frac{m}{2}} - \frac{1}{2} (1-t^2)^{\frac{m}{2}} \ln(1-t^2).$$

Подставляя последнее выражение в формулу (21) и учитывая значения табличных интегралов, находим моменты

$$\tilde{v}_n(N-m) = \frac{1}{2(m+n+1)} B\left(\frac{m}{2}, \frac{n+1}{2}\right) \left\{ \ln m + \frac{m}{2} \left[\psi\left(\frac{m+n+1}{2}\right) - \psi\left(\frac{m}{2}\right) \right] - \frac{n+1}{m+n+1} \right\},$$

где $\psi(x)$ – пси-функция.

Учитывая (37), получаем входящие в формулу (27) отношения:

$$\frac{\tilde{v}_n(N-m)}{v_n(N-m)} = \ln m + \frac{m}{2} \left[\psi\left(\frac{m+n+1}{2}\right) - \psi\left(\frac{m}{2}\right) \right] - \frac{n+1}{m+n+1}.$$

В частности, с учетом $L+M=N$ имеем

$$\begin{aligned} \frac{\tilde{v}_{L-1}(N-M)}{v_{L-1}(N-M)} &= \ln M + \frac{M}{2} \left[\psi\left(\frac{N}{2}\right) - \psi\left(\frac{M}{2}\right) \right] - \frac{L}{N}, \\ \frac{\tilde{v}_{M-1}(N-L)}{v_{M-1}(N-L)} &= \ln L + \frac{L}{2} \left[\psi\left(\frac{N}{2}\right) - \psi\left(\frac{L}{2}\right) \right] - \frac{M}{N}. \end{aligned}$$

Подставляя эти величины в формулу (27), а также (36) в формулу (24) и принимая во внимание (38), находим величину негауссовской составляющей информации:

$$\Delta I(\tilde{\xi}_L, \tilde{\eta}_M) = \ln \left[\frac{LM}{2N} B\left(\frac{L}{2}, \frac{M}{2}\right) \right] + \frac{N}{2} \psi\left(\frac{N}{2}\right) - \left[\frac{M}{2} \psi\left(\frac{M}{2}\right) + \frac{L}{2} \psi\left(\frac{L}{2}\right) + 1 \right].$$

Некоторые значения приращения информации, определяемые по этой формуле, представлены в таблице 2.

Таблица 2 – Значения негауссовской составляющей взаимной информации многомерного цилиндрического распределения

M L	1	2	3	4	5	6	7	...	∞
1	0,1447	0,2082	0,2433	0,2655	0,2806	0,2917	0,3001	...	0,3610
2	0,2082	0,3069	0,3641	0,4014	0,4276	0,4470	0,4620	...	0,5772
3	0,2433	0,3641	0,4365	0,4849	0,5195	0,5455	0,5657	...	0,7300
4	0,2655	0,4014	0,4849	0,5416	0,5828	0,6141	0,6387	...	0,8476
5	0,2806	0,4276	0,5195	0,5828	0,6293	0,6649	0,6932	...	0,9431
...	...	...	...	...	...	...	...	...	...
∞	0,3610	0,5772	0,7300	0,8476	0,9431	1,0234	1,0927	...	∞

Абсолютные значения приращения информации, связанные с негауссовским характером рассматриваемых величин, здесь больше, чем для распределения Стюдента, и растут при увеличении числа составляющих этих величин. Причина этого состоит в том, что рассматриваемое распределение

не является родственным гауссовскому (оно, в частности, в отличие от гауссовского существует в конечных пределах) и с ростом числа измерений не стремится к нему.

Формула (27) позволяет вычислить количество взаимной информации, дополнительное к гауссовскому, для любых распределений многомерных непрерывных случайных величин, если их совокупность описывается плотностью распределения вероятностей вида (1). В частности, это могут быть множества значений негауссовских случайных процессов, взятых в любых N точках отсчета.

Библиографический список

1. Доценко С.В. Модель многомерного распределения вероятностей / С.В. Доценко // Вестн. СевГТУ. Сер. Информатика, электроника, связь: Сб. науч. тр. — Севастополь, 2001. — Вып. 31. — С. 25–34.
2. Вентцель Е.С. Теория вероятностей / Е.С. Вентцель. — М.: Высш. шк., 2001. — 575 с.
3. Доценко С.В. Связь между плотностями вероятностей различного числа измерений многомерной случайной величины / С.В. Доценко // Вестн. СевГТУ. Сер. Физика и математика: Сб. науч. тр. — Севастополь, 2003. — Вып. 43. — С. 76–85.
4. Гельфанд И.М. О вычислении количества информации о случайной функции, содержащейся в другой такой функции / И.М. Гельфанд, А.М. Яглом // Успехи математических наук. — 1957. — Т. XII. — Вып. 1 (73). — С. 3–51.

Поступила в редакцию 5.03.2003 г.