

ХАРАКТЕРИСТИКИ СТОХАСТИЧЕСКОЙ СВЯЗИ ПАРЫ СЛУЧАЙНЫХ ВЕЛИЧИН

Приведен сравнительный анализ двух характеристик стохастической связи пары случайных величин: их коэффициента корреляции и величины взаимной информации. Показано, что взаимная информация отлична от нуля даже в тех случаях, когда коэффициент корреляции функционально связанных случайных величин обращается в нуль и уже не может служить мерой их связи. Поэтому взаимную информацию можно рассматривать как универсальную характеристику связи, сохраняющую смысл при любых нелинейных преобразованиях исходных величин.

Характеристиками стохастической связи интересуются для получения информации о характере (форме) и тесноте связи между случайными величинами и влиянии на нее различных факторов [1]. Наиболее часто применяемой характеристикой связи случайных величин является коэффициент корреляции. Однако возможны ситуации, когда корреляционный анализ не эффективен или приводит к ошибочным результатам.

Две независимые случайные величины всегда некоррелированы. Однако из некоррелированности случайных величин не вытекает их независимость. Более того, некоррелированными могут быть функционально зависимость случайные величины. Объясняется это тем, что корреляционный момент и коэффициент корреляции характеризуют не любую, а только линейную стохастическую связь. Для установления полной стохастической связи необходимо введение иной количественной меры, учитывающей как линейные, так и все нелинейные эффекты. Такой мерой может служить взаимная информация между случайными величинами. Покажем на конкретном примере, как применять эту меру и к каким результатам это приводит.

Пусть имеются две распределенные нормально (по гауссовскому закону) центрированные случайные величины ξ и η , имеющие единичные дисперсии. Их двумерная плотность вероятностей

$$p_{2\xi\eta}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right), \quad (|\rho| \leq 1, -\infty < x, y < \infty) \quad (1)$$

зависит от коэффициента корреляции ρ . Стохастическая связь между такими случайными величинами линейна и определяется только величиной ρ . Пример двумерного распределения (1) для $\rho=0,75$ приведен на рисунке 1,а.

Из формулы (1) имеем одномерные распределения величин ξ и η

$$p_{1\xi}(x) = p_{1\eta}(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2), \quad (-\infty < x < \infty), \quad (2)$$

которые также нормальны.

Средняя взаимная информация величин ξ и η выражается двойным интегралом [2]

$$I(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{2\xi\eta}(x, y) I_{0\xi\eta}(x, y) dx dy, \quad (3)$$

где частная взаимная информация дается формулой

$$I_{0\xi\eta}(x, y) = \ln \frac{p_{2\xi\eta}(x, y)}{p_{1\xi}(x) p_{1\eta}(y)}. \quad (4)$$

Здесь выбрана натуральная система единиц измерения взаимной информации как более удобная для аналитических расчетов. Для перехода к двоичной системе необходимо окончательный результат разделить на $\ln 2$.

Комбинация формул (1)-(4) дает

$$I(\xi, \eta) = -\ln(\sqrt{1 - \rho^2}). \quad (5)$$

Эта зависимость изображена на рисунке 2 (кривая 1). Из формулы (5) следует, что количество взаимной информации двух гауссовских случайных величин определяется только величиной коэффициента корреляции между ними и не зависит от знака последнего.

Произведем квадратичное преобразование второй из величин распределения (1), т.е. введем случайную величину $\zeta = \eta^2$. Найдем плотность вероятностей последней. В общем случае одномерная плотность распределения квадрата случайной величины при известной одномерной плотности первой степени этой величины может быть получена из выражения [3]

$$p_{1\zeta}(y) = \begin{cases} \frac{p_{1\eta}(\sqrt{y}) + p_{1\eta}(-\sqrt{y})}{2\sqrt{y}}, & y > 0, \\ 0, & y < 0. \end{cases} \quad (6)$$

Аналогично двумерная плотность распределения пары случайных величин ξ и $\zeta = \eta^2$ определяется формулой

$$p_{2\xi\zeta}(x, y) = \begin{cases} \frac{p_{2\xi\eta}(x, \sqrt{y}) + p_{2\xi\eta}(x, -\sqrt{y})}{2\sqrt{y}}, & y > 0, \\ 0, & y < 0. \end{cases} \quad (7)$$

Подставив в (7) соответственно плотности (2) и (1), получим выражение для одномерной плотности вероятностей величины ζ

$$p_{1\zeta}(y) = \frac{1}{\sqrt{2\pi y}} \exp(-y/2), \quad (0 < y < \infty) \quad (8)$$

и двумерную – величин ξ и ζ :

$$p_{2\xi\zeta}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sqrt{y}} \exp\left(-\frac{x^2+y}{2(1-\rho^2)}\right) \operatorname{ch}\left(\frac{\rho x\sqrt{y}}{1-\rho^2}\right), \quad (-\infty < x < \infty, 0 < y < \infty). \quad (9)$$

Распределения (8) и (9), в отличие от (2) и (1), не являются гауссовскими. Поэтому роль коэффициента ρ в распределении (9) иная, чем в распределении (1). Трансформация плотности распределения (1) в (9) для $\rho=0,75$ изображена на рисунке 1. Заметим, что вид функции (1) зависит от знака коэффициента корреляции ρ , а функции (9) – не зависит.

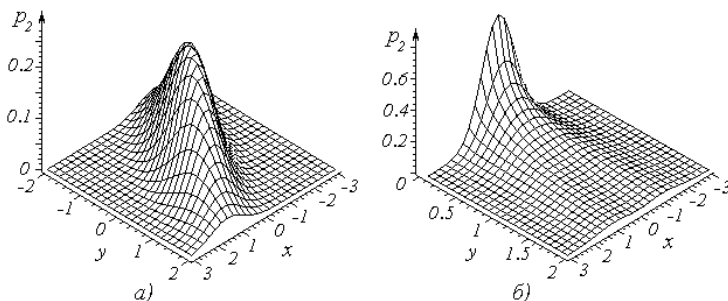


Рисунок 1 – Двумерные распределения: а) гауссовское, б) трансформированное

Математическое ожидание величины ζ есть $M[\zeta]=1$. Корреляционный момент случайных величин ξ и ζ равен нулю в силу четной зависимости $p_{2\xi\eta}(x,y)$ от переменной x .

Итак, величины ξ и $\zeta=\eta^2$ некоррелированы, хотя они функционально связаны и коэффициент корреляции ξ и η равен отличному от нуля значению ρ . Следовательно, корреляционный момент не отражает стохастической связи величин ξ и η^2 .

Найдем среднюю взаимную информацию ξ и ζ . Аналогично формуле (3) имеем

$$I(\xi, \zeta) = \int_{x=-\infty}^{\infty} \int_{y=0}^{\infty} p_{2\xi\zeta}(x,y) I_{0\xi\zeta}(x,y) dx dy,$$

где частная взаимная информация дается формулой

$$I_{0\xi\zeta}(x,y) = \ln \frac{p_{2\xi\zeta}(x,y)}{p_{1\xi}(x)p_{1\zeta}(y)}.$$

Подставив в последнюю формулу плотности вероятностей (2), (6) и (9), получим

$$I_{0\xi\zeta}(x,y) = -\ln(\sqrt{1-\rho^2}) - \frac{\rho^2}{2(1-\rho^2)}(x^2+y) + \ln \operatorname{ch}\left(\frac{\rho x\sqrt{y}}{1-\rho^2}\right).$$

С учетом этого выражения имеем

$$I(\xi, \zeta) = -\ln(\sqrt{1-\rho^2}) - \frac{\rho^2}{1-\rho^2} + \frac{1}{2\pi\sqrt{1-\rho^2}} T(\rho), \quad (10)$$

где обозначен двойной интеграл

$$T(\rho) = \int_{x=-\infty}^{\infty} \int_{y=0}^{\infty} \frac{1}{\sqrt{y}} \exp\left(-\frac{x^2+y}{2(1-\rho^2)}\right) \operatorname{ch}\left(\frac{\rho x\sqrt{y}}{1-\rho^2}\right) \ln \operatorname{ch}\left(\frac{\rho x\sqrt{y}}{1-\rho^2}\right) dx dy.$$

Значения этого двойного интеграла не удастся найти непосредственно с необходимой точностью известными численными методами. Представим его в таком виде, когда это возможно.

Произведя замену переменной $y=z^2$, учитывая четный характер подынтегральной функции в зависимости от x и переходя к полярным координатам, получим

$$T(\rho) = 2(1-\rho^2) T_2(\rho), \quad (11)$$

где обозначено

$$T_2(\rho) = \int_0^{\pi/2} T_1(\theta) d\theta \quad (12)$$

и, в свою очередь,

$$T_1(\theta) = 2 \int_0^{\infty} e^{-x} \operatorname{ch}(\rho x \sin \theta) \ln \operatorname{ch}(\rho x \sin \theta) dx. \quad (13)$$

Подстановка формулы (11) в (10) дает величину средней взаимной информации случайных величин ξ и ζ в виде:

$$I(\xi, \zeta) = -\ln(\sqrt{1-\rho^2}) - \frac{\rho^2}{1-\rho^2} + \frac{\sqrt{1-\rho^2}}{\pi} T_2(\rho),$$

где $T_2(\rho)$ определяется формулой (12).

Найдем вначале аналитическое выражение функции $T_1(\theta)$, перейдя тем самым от вычисления двойного интеграла к одинарному. Для этого учтем известное представление гиперболического косинуса через экспоненциальную функцию $\operatorname{ch} x = (e^x + e^{-x})/2$. Кроме того, под знаком логарифма в формуле (13) запишем $\operatorname{ch}(x)$ в виде $\operatorname{ch} x = e^x(1 + e^{-2x})/2$. Получим

$$T_1(\theta) = -\ln 2 [V_1(\theta) + V_2(\theta)] + \rho \sin \theta [V_3(\theta) + V_4(\theta)] + V_5(\theta) + V_6(\theta), \quad (14)$$

где обозначены интегралы

$$\begin{aligned}
V_{1,2}(\theta) &= \int_0^{\infty} \exp[-x(1 \pm |\rho| |\sin \theta|)] d\theta, \quad V_{3,4}(\theta) = \int_0^{\infty} x \exp[-x(1 \pm |\rho| |\sin \theta|)] d\theta, \\
V_{5,6}(\theta) &= \int_0^{\infty} \exp[-x(1 \pm |\rho| |\sin \theta|)] \ln[1 + \exp(-2|\rho| |x \sin \theta|)] d\theta.
\end{aligned} \tag{15}$$

Два первых интеграла вычисляются аналитически:

$$V_{1,2}(\theta) = (1 \pm |\rho| |\sin \theta|)^{-1}, \quad V_{3,4}(\theta) = (1 \pm |\rho| |\sin \theta|)^{-2}.$$

Далее, преобразуем выражение $V_{5,6}(\theta)$. Так как в нашем случае $x \geq 0$ и $\sin \theta \geq 0$, а, соответственно, и $\exp(-2|\rho| |x \sin \theta|) \leq 1$, то возможно представление входящего под знак интеграла логарифма в виде ряда:

$$\ln(1 + \exp(-2|\rho| |x \sin \theta|)) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \exp(-2|\rho| |kx \sin \theta|).$$

Подставив это разложение в последнюю формулу (15) и выполнив почленное интегрирование, получим

$$V_{5,6}(\theta) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \frac{1}{1 + (2k \pm 1)|\rho| |\sin \theta|}.$$

Нетрудно убедиться, что имеет место представление

$$V_5(\theta) + V_6(\theta) = \frac{1}{1 + |\rho| |\sin \theta|} + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)} \frac{1}{1 + (2k+1)|\rho| |\sin \theta|},$$

представляющее собой знакочередующийся быстро сходящийся ряд. Подставляя найденные выражения в формулу (14), а затем в (12), видим, что выражение для искомой взаимной информации может быть записано в виде суммы

$$\begin{aligned}
I(\xi, \zeta) &= -\ln(\sqrt{1 - \rho^2}) - \frac{\rho^2}{1 - \rho^2} + \frac{\sqrt{1 - \rho^2}}{\pi} \times \\
&\times \left[-\ln(2)(U_1(\rho) + U_2(\rho)) + U_1(\rho) + |\rho|(U_3(\rho) + U_4(\rho)) + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)} S_k(\rho) \right],
\end{aligned} \tag{16}$$

слагаемыми которой являются интегралы

$$U_{1,2}(\rho) = \int_0^{\pi/2} \frac{d\theta}{1 \pm |\rho| |\sin \theta|}, \quad U_{3,4}(\rho) = \int_0^{\pi/2} \frac{\sin \theta d\theta}{(1 \pm |\rho| |\sin \theta|)^2}, \quad S_k(\rho) = \int_0^{\pi/2} \frac{d\theta}{1 + (2k+1)|\rho| |\sin \theta|}.$$

Эти интегралы вычисляются аналитически [4]:

$$U_{1,2}(\rho) = \frac{\pi/2 \mp \arcsin|\rho|}{\sqrt{1 - \rho^2}}, \quad U_1(\rho) + U_2(\rho) = \frac{\pi}{\sqrt{1 - \rho^2}}. \tag{17}$$

Далее, нетрудно убедиться, что

$$U_{3,4}(\rho) = \mp \frac{d}{d\rho} U_{1,2}(\rho).$$

Комбинирование этого выражения с первой формулой (17) дает

$$\begin{aligned} U_{3,4}(\rho) &= \frac{1}{1-\rho^2} + \frac{|\rho| \left(|\arcsin|\rho| \mp \pi/2 \right)}{(1-\rho^2)^{3/2}}, \\ U_3(\rho) + U_4(\rho) &= \frac{2}{1-\rho^2} + \frac{2|\rho| |\arcsin|\rho|}{(1-\rho^2)^{3/2}}. \end{aligned} \quad (18)$$

Интегралы $S_k(\rho)$ также вычисляются аналитически, однако получаемые при этом выражения весьма громоздки. В то же время в силу простоты подынтегральных функций их значения с любой заданной степенью точности могут быть найдены простыми численными методами.

Подставляя формулы (17) и (18) в (16), окончательно находим

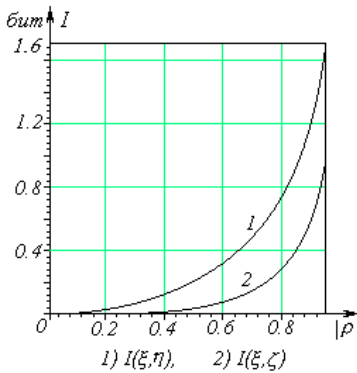


Рисунок 2 – Взаимная информация

$$\begin{aligned} I(\xi, \zeta) &= -\ln(\sqrt{1-\rho^2}) + \frac{1}{2} - \ln(2) - \frac{\rho^2}{1-\rho^2} \left(1 - \frac{2}{\pi} \arcsin|\rho| \right) + \\ &+ \frac{2}{\pi} \frac{|\rho|}{\sqrt{1-\rho^2}} - \frac{\arcsin|\rho|}{\pi} + \frac{\sqrt{1-\rho^2}}{\pi} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)} S_k(\rho). \end{aligned}$$

Результаты вычислений взаимной информации по этой формуле с учетом перевода из натуральных единиц в двоичные представлены на рисунке 2 (кривая 2). В отличие от коэффициента корреляции она не равна нулю.

Из сравнения кривых 1 и 2 следует, что взаимная информация ξ и $\zeta = \eta^2$ имеет существенно меньшую величину, чем взаимная информация ξ и η . Такое уменьшение вызвано тем, что при возведении величины η в квадрат необратимо теряются сведения о ее знаке. Эти потери информации более существенны при малых величинах $|\rho|$.

Полученные результаты показывают, что взаимная информация пар случайных величин количественно характеризует их стохастическую связь. Это имеет место во всех случаях, в том числе и тогда, когда коэффициент

корреляции этих величин может оказаться равным нулю. Следовательно, величина взаимной информации является более универсальной характеристикой связи, чем коэффициент корреляции.

Теперь допустим, что в квадрат возведена не только случайная величина η , но и ξ , и обозначим $\lambda = \xi^2$. Найдем указанные выше характеристики связи пары случайных величин λ и ζ , т.е. квадратов исходных величин ξ и η , описываемых двумерным распределением (1).

Для получения двумерного распределения величин λ и ζ произведем преобразование типа (7) двумерной плотности вероятностей (9) по переменной x :

$$p_{2\lambda\zeta}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sqrt{xy}} \exp\left(-\frac{x+y}{2(1-\rho^2)}\right) \operatorname{ch}\left(\frac{\rho\sqrt{xy}}{1-\rho^2}\right), (0 < x, y < \infty). \quad (19)$$

Это распределение для различных значений коэффициента корреляции представлено на рисунке 3. Одномерные распределения обеих величин λ и ζ имеют вид (8). Математические ожидания как λ , так и ζ равны 1, а их дисперсии равны 2. Нетрудно показать, что коэффициент корреляции величин λ и ζ есть $\rho_{\lambda\zeta} = \rho^2$, где ρ – коэффициент корреляции исходных величин ξ и η .

Найдем величину взаимной информации λ и ζ . Заменим в формуле (4) ξ и η соответственно на λ и ζ , и подставим в нее распределения (19) и (8):

$$I_{0\lambda\zeta}(x, y) = -\ln\left(\sqrt{1-\rho^2}\right) - \frac{\rho^2(x+y)}{2(1-\rho^2)} + \ln \operatorname{ch}\left(\frac{\rho\sqrt{xy}}{1-\rho^2}\right). \quad (20)$$

Принимая во внимание (20) и (19) в формуле, аналогичной (3), и производя ряд несложных замен переменных, обнаруживаем, что формула (20) совпадает с (10), т.е.

$$I(\lambda, \zeta) = I(\xi, \eta).$$

Таким образом, количество информации в одной гауссовской случайной величине о квадрате другой гауссовской случайной величины совпадает с количеством взаимной информации квадратов этих величин. Иными словами, возведение в квадрат второй случайной величины уже не изменяет взаимную информацию, если в квадрат возведена первая случайная величина.

Подводя итог изложенному, можно сказать, что как коэффициент корреляции, так и величина взаимной информации характеризуют степень стохастической связи двух гауссовских случайных величин. Нелинейные преобразования этих величин, если они не взаимно однозначны, могут привести к утрате коэффициентом корреляции роли характеристики такой связи. В то же время взаимная информация при этом не равна нулю и может с успехом выполнять такую роль. Это обстоятельство необходимо обязательно учитывать при анализе и разработке нелинейных информационных систем.

Библиографический список

1. Мирский Г.Я. Характеристики стохастической взаимосвязи и их измерения / Г.Я. Мирский. — М.: Энергоиздат, 1982. — 319 с.
2. Колесник В.Д. Курс теории информации / В.Д. Колесник, В.Ш. Полтырев. — М.: Наука, 1982. — 416 с.
3. Левин Б.Р. Теоретические основы статистической радиотехники. Кн. I. / Б.Р. Левин. — М.: Советское радио, 1969. — 750 с.
4. Двайт Г.Б. Таблицы интегралов и другие математические формулы / Г.Б. Двайт. — М.: Наука, 1973. — 228 с.

Поступила в редакцию 27.11.2003 г.

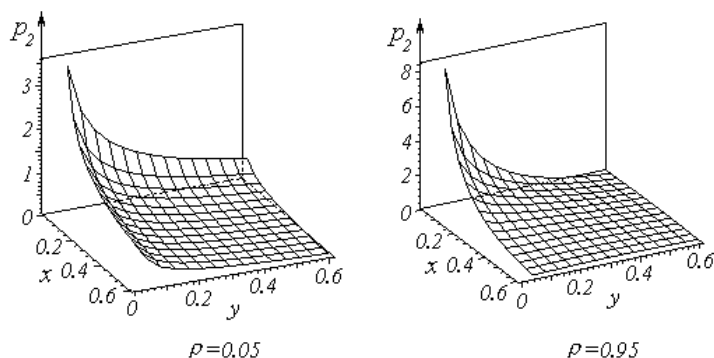


Рисунок 3 – Двумерная плотность распределения вероятностей преобразованных величин