ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 22 МАТЕМАТИКИ Выпуск 5
2015

А. Н. Тырсин (Екатеринбург, УрФУ). **Структура дифференциальной энтропии многомерных случайных величин.**

Пусть дана многомерная непрерывная случайная величина $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)$. Исследуем свойства ее дифференциальной энтропии $H(\mathbf{Y})$.

В [1] установлено, что если все компоненты Y_i многомерной случайной величины $\mathbf{Y}$ имеют конечные дисперсии, то

$$H(\mathbf{Y}) = \sum_{i=1}^{m} H(Y_i) + \frac{1}{2}\sum_{k=2}^{m} \ln(1 - R^2_{Y_k/Y_1 Y_2 \ldots Y_{k-1}}), \tag{1}$$

где $H(\mathbf{Y})_V = \sum_{i=1}^{m} H(Y_i)$ — энтропия хаотичности, равная сумме энтропий случайных величин Y_i; $H(\mathbf{Y})_R = \frac{1}{2}\sum_{k=2}^{m} \ln\left(1 - R^2_{Y_k/Y_1 Y_2 \ldots Y_{k-1}}\right)$ — энтропия самоорганизации, характеризующая вклад в энтропию тесноты корреляционных связей между компонентами многомерной случайной величины $\mathbf{Y}$, $R^2_{Y_k/Y_1 Y_2 \ldots Y_{k-1}}$ — индексы детерминации соответствующих регрессионных зависимостей, $k = 2, 3, \ldots, m$.

Анализ формулы (1) свидетельствует о недостаточной детализации энтропии хаотичности.

Теорема. *Дифференциальная энтропия $H(\mathbf{Y})$ многомерной непрерывной случайной величины $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)$, компоненты которой имеют дисперсии $\sigma^2_{Y_i}$ и параметры масштаба, равна*

$$H(\mathbf{Y}) = \sum_{i=1}^{m} \ln \sigma_{Y_i} + \sum_{i=1}^{m} \varkappa_i + \frac{1}{2}\sum_{k=2}^{m} \ln\left(1 - R^2_{Y_k/Y_1 Y_2 \ldots Y_{k-1}}\right), \tag{2}$$

где $\varkappa_i = H\left(\frac{Y_i}{\sigma_{Y_i}}\right)$ — энтропийный показатель типа закона распределения случайной величины Y_i, $i = 1, 2, \ldots, m$.

Доказательство. Пусть непрерывная случайная величина X имеет дисперсию σ^2_X и параметр масштаба. Это означает, что случайные величины X и $\widetilde{X} = X/\sigma_X$ имеют однотипные распределения. При этом у случайной величины $\widetilde{X}$ среднее квадратическое отклонение $\sigma_{\widetilde{X}} = 1$. Поэтому, как показано в [1], разность дифференциальных энтропий

$$H(X) - H(\widetilde{X}) = \ln \frac{\sigma_X}{\sigma_{\widetilde{X}}} = \ln \sigma_X.$$

Введем энтропийный показатель типа закона распределения случайной величины X как $\varkappa = H(\widetilde{X})$. Тогда последнее выражение можно представить в виде

$$H(X) = \ln \sigma_X + \varkappa.$$

Поскольку все случайные величины Y_i имеют дисперсии $\sigma^2_{Y_i}$ и параметры масштаба, то для них также будет справедливо полученное выражение

$$H(Y_i) = \ln \sigma_{Y_i} + \varkappa_i, \tag{3}$$

где $\varkappa_i = H\left(\frac{Y_i}{\sigma_{Y_i}}\right), \ \ i = 1, 2, \ldots, m.$

Подставив (3) в (1), получим формулу (2).

Из (2) видно, что энтропия хаотичности состоит из суммы двух разнородных компонент

$$H(\mathbf{Y})_V = \sum_{i=1}^{m} \ln \sigma_{Y_i} + \sum_{i=1}^{m} \varkappa_i. \tag{4}$$

Первая компонента в (4) $H(\mathbf{Y})_V^{\sigma} = \sum_{i=1}^{m} \ln \sigma_{Y_i}$ определяется значениями средних квадратических отклонений случайных величин Y_i. Она характеризует вклад в энтропию $H(\mathbf{Y})$ степени рассеяния компонент многомерной случайной величины $\mathbf{Y}$.

Вторая компонента в (4) $H(\mathbf{Y})_V^{\varkappa} = \sum_{i=1}^{m} \varkappa_i$ равна сумме энтропийных показателей типов законов распределений случайных величин Y_i. Она характеризует вклад в энтропию $H(\mathbf{Y})$ форм распределений компонент многомерной случайной величины $\mathbf{Y}$. Этот вклад может быть существенным. Например, разность между энтропийными показателями распределений Гаусса и Лапласа равна $\Delta\varkappa = \ln\sqrt{2\pi e} - \ln e\sqrt{2} = \ln\frac{\pi}{\sqrt{e}} = 0,645$, что равносильно увеличению дисперсии случайной величины в $\frac{\pi^2}{e} = 3,631$ раз.

Таким образом, дифференциальная энтропия обладает *триализмом*. Существуют три причины изменения энтропии многомерной случайной величины $\mathbf{Y}$: изменение степени рассеяния ее компонент, изменение форм распределений ее компонент и изменение тесноты корреляционных связей между ее компонентами.

Достоинством формулы (2) является то, что энтропийное моделирование многомерных стохастических систем на ее основе не требует знания или определения закона распределения многомерной случайной величины $\mathbf{Y}$, что практически нереализуемо в реальных задачах. При этом в отличие от методов многомерного статистического анализа, здесь не теряется формальная строгость и соответствие модели (2) реальным экспериментальным данным. Это позволяет использовать формулу (2) для моделирования и исследования реальных многомерных стохастических систем и процессов по экспериментальным данным ограниченного объема.

СПИСОК ЛИТЕРАТУРЫ

1. *Тырсин А. Н.*, *Ворфоломеева О. В.* Исследование динамики многомерных стохастических систем на основе энтропийного моделирования. — Информатика и ее применения, 2013, т. 7, в. 4, с. 3–10.