Секция «Теория вероятностей и математическая статистика»

Неравенства между различными f-дивергенциями

Научный руководитель – Булинский Александр Вадимович

Хлюстов Дмитрий Кириллович
Студент (специалист)
Московский государственный университет имени М.В.Ломоносова,
Механико-математический факультет, Кафедра теории вероятностей, Москва, Россия
E-mail: hlustov.d@gmail.com

Понятие дифференциальной энтропии, введенное Шенноном в 1948 году [5], в настоящее время находит широкое применение в теории вероятностей, теории информации, моделировании сложных систем и исследованиях в области искусственного интеллекта. Для случайной величины X, принимающей значения в пространстве $\mathbb{R}^n$ и имеющей плотность $p(x)$ по мере μ, ее дифференциальная энтропия по определению равна $H(X) = -\int_{\mathbb{R}^n} p(x)\log p(x)d\mu$. Часто данную величину интерпретируют как меру неопределенности случайной величины X.
Обобщением понятия энтропии служат f-дивергенции (см., например, [4]), чаще всего рассматриваемые как меры сходства двух распределений. Для мер P, Q на пространстве $\mathbb{R}^n$ и выпуклой функции $f : \mathbb{R} \to \mathbb{R}$ полагают

$$D_f(P||Q) = \begin{cases} \int_{\mathbb{R}^n} f(\frac{dP}{dQ})dQ, \text{ если } P \ll Q \\ +\infty, \text{ иначе.} \end{cases}$$

В зависимости от выбора f величина D_f может обладать различными свойствами, делающими ее удобнее для тех или иных приложений. В частности, подстановка $f(t) = t\log t, t > 0$, позволяет получить дивергенцию Кульбака-Лейблера [2], широко применяемую в статистике и теории информации. Особый интерес представляют теоремы, обеспечивающие возможность сравнивать значения f-дивергенций между собой. Это позволяет использовать важные свойства каждой дивергенции [3].
В данной работе рассматриваются новые неравенства между f-дивергенциями и развивается общий подход к их получению, основанный на понятии совместной области [1]. Для доказательств привлекается аппарат выпуклого анализа (например, теоремы Каратеодори и Крейна-Мильмана). Полученные результаты допускают наглядную геометрическую интерпретацию и позволяют сводить анализ изучаемых распределений к распределениям, сосредоточенным на конечных множествах.

Источники и литература

1) Harremoës P., Vajda I. On pairs of f-divergences and their joint range //IEEE Transactions on Information Theory. 2011, Vol. 57, №, 6. p. 3230-3235.
2) Kullback S., Leibler R. A. On information and sufficiency //The annals of mathematical statistics. 1951, Vol. 22, № 1. p. 79-86.
3) Nishiyama T., Sason I. On relations between the relative entropy and χ^2-divergence, generalizations and applications //Entropy. 2020, Vol. 22, № 5. p. 563.
4) Rényi A. On measures of entropy and information //Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. 1961, Vol. 4. p.547-562.
5) Shannon C. E. A mathematical theory of communication //The Bell system technical journal. 1948, Vol. 27, № 3. p. 379-423.