СИСТЕМНО-ИНФОРМАЦИОННЫЙ ПОДХОД В ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ

канд. техн. наук, проф. Голик Е.С.

Санкт-Петербургский государственный горный университет
Факультет приборостроения, информационных и электронных систем
Кафедра системного анализа и управления инновациями

Аннотация

В статье определяется один из возможных приемов идентификации закона распределения случайной величины в рамках процедуры обоснования предпочтений ЛПР в условиях неопределенности на основе информационного подхода.

Ключевы слова

Постулат Джейнса, энтропия Шеннона, распределение, мера информации, выборка.

Abstract

This article is one of the possible methods of identification of the distribution of a random variable as part of the justification for the decision maker's preferences under uncertainty based approach.

Keywords

Postulate of Janes, Shannon's entropy, distribution, a measure of information retrieval.

Теория Неймана-Моргенштерна, определяя принципы и способы выбора наилучшей альтернативы в выборе варианта решения в условиях риска, не даёт представления о степени общности получаемых результатов. При этом очевидно, что некоторые альтернативы могут отличаться друг от друга таким образом, что лица, принимающее решение (ЛПР), независимо от вида функции полезности, будет предпочиать один вариант другому, следуя некоторым общим правилам обоснования предпочтений. Решение следует искать в классе задач оценки в условиях неопределённости, то есть обозначить сопосталение распределений заданных альтернатив. Понятно, что обобщение распределений вероятностей производится на основе функций распределения.

Процедура введения частичной упорядочности во множестве функций распределения случайной величины есть стохастическое доминирование.

И здесь уместно воспользоваться постулатом Э.Т. Джейнса, который, опираясь на теорию информаций Клодта Шеннона, предложил при рассмотрении и решени задач в условиях неопределённости истолковывать теоретические результаты лишь как наиболее вероятностные и беспристрастные выводы, которые можно сделать на основе всей имеющейся информации.

Утверждать что-либо другое – значит объявлять, что мы знаем на самом деле.

Подобно любой задаче вероятностного характера достоверность результатов, получаемых в форме статистического вывода будет тем выше, а неопределённость тем меньше, чем больше объём выборки.

В качестве меры неопределёности распределения вероятностей случайной величины используют энтропию Шеннона.

Как термин – энтропия (от греч. entropia – поворот, превращение) впервые была введена в XIX в. в теплотехнике и характеризовала ту часть энергии тепловых машин, которая рассеивалась в пространстве, не совершая полезной работы.

В современных исследованиях энтропия нашла широкое применение, в том числе и в теории информации, как мера неопределённости какого-либо опыта, который может иметь разные исходы.

Предполжим, что в результате испытаний получены n чисел $x_1, \rightleftharpoons, x_n$ и проверка статистических гипотез с помощью χ-статистик о принадлежности функции распределения $F(x)$ генеральной совокупности некоторому классу распределений (для определённости будем считать равномерному, нормальному и экпоненциальному) не выявила противоречий. Требуется определить условия, при которых каждый из рассматриваемых законов доминируемт над альтернативным. С этой целью определим значение дифференциальной энтропии и введём H_{ε}-упорядочение для равномерного, нормального и экспоненциального распределений.

Теоретическое значение дифференциальной энтропии равно

$$H_{\varepsilon}^{(1)} = \int_{a}^{b} \frac{1}{b-a} \ln\left(b-a\right) = \ln\left(b-a\right)$$

для равномерного на интервале $\left(a, b\right)$ закона распределения;

$$H_{\varepsilon}^{(2)} = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \ln \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \ln\left(\sqrt{2\pi}\sigma\right)$$

для нормального распределения с параметрами $\left(m, \sigma^2\right)$;

$$H_{\varepsilon}^{(3)} = -\int_{0}^{\infty} \lambda e^{-\lambda x} \ln\left(\lambda e^{-\lambda x}\right) dx = 1 - \ln\lambda$$

для экспоненциального распределения с параметром λ.

Используя значения параметров, опредлённые по выборке $x_1, \ldots, x_n$, можно определить эмпирические («выборочные») значения энтропии:

$$H_{\varepsilon}^{(1)}\left(n\right) = \ln\left(x_n^{(n)} - x_1^{(n)}\right) = \ln t_n;$$

$$H_{\varepsilon}^{(2)}\left(n\right) = \ln\sqrt{2\pi e} + \frac{1}{2}\ln S_x^2;$$

$$H_{\varepsilon}^{(3)}\left(n\right) = 1 + \ln \overline{x}_n,$$

где $\overline{x}_n$, S_x^2 – оценка математического ожидания и выборочная дисперсия соответственно;

$t_n = x_n^{(n)} - x_n^{(1)}$ – размах выборки, определяемый по крайним членам $x_n^{(n)}$ и $x_n^{(1)}$ вариационного ряда.

Сопротивление выборочных значений энтропий $H_\varepsilon^{(1)}(n)$, $H_\varepsilon^{(2)}(n)$ и $H_\varepsilon^{(3)}(n)$ позволяет найти условия стохастического доминирования по энтропии рассматриваемых распределений.

С математической точки зрения применение принципа максимума неопределённости в общем случае приводит к решению вариационныз задач при ограничениях, обусловленных формой задания вероятносных харакеристик и областью занчений случайной величины. Решение такой экстремальной задачи однозначно опредеяет выбор распределения. Для дискретных случайных величин выбор распределения сводится к решению задачи на условных экстремум.

Литература

1. Мартыщенко, Л.А. Стохастическое доминирование /Л.А. Мартыщенко. – М.: МОСССР, ВАА, 1985. – 42 с.

2. System Information Approach. http://maikov.shat.ru/russian/system/index.htm

Рецензент проф. Первухин Д.А