**18. Энтропия, информация**

Рассмотрим некоторое случайное событие , вероятность которого равна . Каким образом выразить понятие ”количество информации, связанное с событием ”? Естественно определить это количество информации (обозначим его ) как некоторую функцию от вероятности, Как следует выбрать функцию , чтобы обладала теми свойствами, которые мы ожидаем от понятия ”информация ”?

Сформулируем эти свойства в виде двух аксиом:

1. Информация является функцией вероятности , удовлетворяющей условию
2. Если событие является композицией двух независимых событий, то вероятность должна быть равна сумме:

Решением этого уравнения является . Действительно, в Параграфе 4 мы видели, что решение уравнения

с условием есть . Положим теперь и обозначим , тогда, согласно ( ),

откуда и окончательно получаем уравнение ( ) для функции *G*. Функцией, обратной к является .

Константа c служит для выбора основания логарифма, или, что - то же самое, для выбора единицы измерения информации. Стандартно в качестве основания используют число 2, тогда для равновероятного случайного выбора (количество информации равно 1 (одному биту). Но в этом параграфе мы будем писать натуральный логарифм для упрощения некоторых формул.

Часто наряду с информацией рассматривают энтропию, обозначим ее *H* – меру неопределенности, имея в виду, что полученная в результате статистического эксперимента информация равна уменьшению неопределенности, то есть, энтропии, . Определим энтропию, связанную с событием :

.

Рассмотрим теперь некий статистический эксперимент с *N* возможными исходами, имеющими соответствующие вероятности ; точно также можно рассуждать о дискретной случайной величине с распределением , но сейчас мы интересуемся только вероятностями событий, а не значениями . В результате проведения эксперимента может наступить дно из возможных событий (для случайной величины , ). С каждым событием связана соответствующая энтропия , так что мы имеем случайную величину с распределением ; в качестве энтропии, связанной с этим статистическим экспериментом (или, что тоже самое, со случайной величиной ), естественно принять математическое ожидание .

**Определение.** Энтропия, соответствующая распределению вероятностей , равна

**Теорема 18.1.** Энтропия обладает следующими свойствами:

1. , причем тогда и только тогда, когда соответствующее распределение вырожденное.
2. Максимум энтропии достигается на равномерном распределении.
3. Если два статистических эксперимента независимы, то их совместная энтропия равна сумме энтропий (аддитивность энтропии).

**Доказательство.** Рассмотрим функцию

она отрицательна и равна нулю лишь при действительно, вычислив производную,

видим, что - единственный экстремум и это максимум, поскольку вторая производная

Таким образом, для любого ,

Рассмотрим невырожденное распределение: пусть , тогда из полученного неравенства выводим для энтропии

Первое утверждение Теоремы доказано.

Для равномерного распределения , следовательно

Пусть - энтропия произвольного распределения, тогда с помощью того же равенства получаем

Совместную энтропию двух случайных экспериментов обозначим

где - совместное распределение вероятностей двух случайных величин, имеющих распределения по отдельности

Для независимых экспериментов , поэтому энтропия равна сумме,

**Определение.** Для непрерывной случайной величины с плотностью распределения энтропия равна

**Замечание.** Для непрерывных распределений энтропия не обязательно имеет знак положительный, но в качестве меры неопределенности в связи с понятием информации она также широко применяется.

**Упражнение 18.1.** Сформулировать и доказать свойство аддитивности энтропии для непрерывных распределений.

**Упражнение 18.2.** Как связаны , если , то есть, как меняется энтропия при линейном преобразовании ?

Еще одно уникальное свойство гауссовского распределения связано с понятием энтропии: оказывается, именно гауссовское распределение имеет максимальную энтропию, то есть наибольшую меру неопределенности.

**Теорема 18.2.** Среди всех распределений с заданными математическим ожиданием и дисперсией наибольшую энтропию имеет нормальное распределение: обозначим энтропию распределения ,

и рассмотрим множество распределений, удовлетворяющих условиям

Максимальное среди всех таких распределений значений энтропии достигается для нормального распределения .

**Доказательство.** Обозначим гауссовское распределение

его энтропия равна

Если – две плотности распределения, то справедливо неравенство

Действительно, с помощью неравенства ( ) получаем,

что и требовалось.

Пусть теперь – любое распределение, удовлетворяющее условиям Теоремы, а . Тогда для энтропии распределения получаем неравенство,