

## **ТЕОРИЯ ИНФОРМАЦИИ. МЕРА КОЛИЧЕСТВА ИНФОРМАЦИИ**

**Лобач Г.С., Саттаров И.Д.,  
научные руководители канд. ф.-м. наук Мансурова Т.П.,  
кандидат пед. наук Бутакова С.М.  
Сибирский Федеральный университет**

**Теория информации** – комплексная, в основном математическая теория, включающая в себя описание и оценки методов извлечения, передачи, хранения и классификации информации.

Рассматривает носители информации как элементы абстрактного множества, а взаимодействия между носителями как способ расположения элементов в этом множестве. Такой подход дает возможность формально описать код информации, то есть определить абстрактный код и исследовать его математическими методами. Для этих исследований применяет методы теории вероятностей, математической статистики, линейной алгебры, теории игр и других математических теорий.

Основы этой теории заложил американский учёный Э. Хартли в 1928 г., который определили меру количества информации для некоторых задач связи. Позднее теория была существенно развита американским ученым К. Шенноном, российскими учёными А.Н. Колмогоровым, В.М. Глушковым и др.

**Теория вероятностей** – раздел математики, изучающий закономерности случайных явлений: случайные события, случайные величины, их свойства и операции над ними. Современный вид теория вероятностей получила благодаря аксиоматизации, предложенной Андреем Николаевичем Колмогоровым. В результате теория вероятностей приобрела строгий математический вид и окончательно стала восприниматься как один из разделов математики.

**Математическая статистика** – наука, разрабатывающая математические методы систематизации и использования статистических данных для научных и практических выводов.

**Линейная алгебра** – важная в приложениях часть алгебры, изучающая векторы, векторные, или линейные пространства, линейные отображения и системы линейных уравнений. Линейная алгебра широко используется в абстрактной алгебре и функциональном анализе и находит многочисленные приложения в естественных науках. Линейные уравнения, как уравнения прямых и плоскостей, стали естественным предметом изучения после изобретения Декартом и Ферма метода координат (1636 год) Системы линейных уравнений в матрично-векторном виде впервые появились в работах Лагерра (1867 год).

В теории информации выделяются три основных направления: структурное, статистическое и семантическое.

**Структурное направление** – рассматривает дискретное строение массивов информации и их измерение простым подсчётом информационных элементов. Структурные меры учитывают только дискретное строение информации. Элементами информационного комплекса являются кванты – неделимые части информации.

**Статистическое направление** – оперирует понятием энтропии как меры неопределенности, то есть здесь учитывается вероятность появления тех или иных сообщений. При статистическом вероятном подходе получение конкретного количества информации рассматривается как результат определенного выбора среди возможных сообщений. Получатель информации может заранее знать или угадать ее часть.

**Семантическое направление** – учитывает целесообразность, ценность или существенность информации. Целесообразность, полезность информации для решения какой – то задачи можно оценить по эффекту, который оказывает полученная информация на решение задачи. Если вероятность достижения цели увеличивается, то информацию следует считать полезной.

Задача

Условие задачи:

Найти вероятность образования сообщения «МАТЕМАТИКА» состоящего из 10 элементов и количество информации, которую оно содержит. При условии, что из 10 элементов 3 элемента – буква «А», 2 элемента – буква «М», 2 элемента буква «Т», 1 элемент – буква «И», 1 элемент – буква «Е» и 1 элемент буква «К»

Введем события А – в сообщении присутствует буква «А», В – в сообщении присутствует буква «М», С – в сообщении присутствует буква «Т», D – в сообщении присутствует буква «Е», E – в сообщении присутствует буква «И», F – в сообщении присутствует буква «К»

Вычисляем вероятности:

M – кол – во определенных букв, N – кол – во всех букв в сообщении

$$P(A) = M/N = 3/10 = 0,3 \quad P(B) = 2/10 = 0,2$$

$$P(C) = 2/10 = 0,2 \quad P(D) = 1/10 = 0,1$$

$$P(E) = 1/10 = 0,1 \quad P(F) = 1/10 = 0,1$$

Вероятность того что в сообщении попадет n элементов X равна

$$P(A)^n = P(A)^3 = 0,3^3 = 0,0027 \quad P(B)^n = P(B)^2 = 0,2^2 = 0,04$$

$$P(C)^n = P(C)^2 = 0,2^2 = 0,04 \quad P(D)^n = P(D)^1 = 0,1^1 = 0,1$$

$$P(E)^n = P(E)^1 = 0,1^1 = 0,1 \quad P(F)^n = P(F)^1 = 0,1^1 = 0,1$$

Вероятность образования сообщения из данного набора элементов равна

$$P = P_1^{n_1} P_2^{n_2} \dots P_i^{n_i} \dots P_m^{n_m}$$

$$P = P(A)^3 \cdot P(B)^2 \cdot P(C)^2 \cdot P(D)^1 \cdot P(E)^1 \cdot P(F)^1 = 0,0027 \cdot 0,04 \cdot 0,04 \cdot 0,1 \cdot 0,1 \cdot 0,1 = 0,000000432$$

Найдем количество информации, содержащиеся в сообщении «МАТЕМАТИКА»

$$I = n \log_2 m$$

n – кол – во элементов

$m$  – кол – во букв в алфавите.

(повторяющиеся буквы считаются за один элемент)

$$n = 10 \quad m = 6 \quad (\text{м, а, е, т, и, к})$$

$$I = 10 \cdot \log_2 6 = \log_2 6^{10} = 25.5$$