# Лабораторная работа №1

**Тема:** Структурные и статистические меры информации.

**Цель работы**: Исследование структурных и статистических методов измерения информации.

**Порядок выполнения работы**

1. Изучить теоретический материал.
2. Выполнить задания.
3. Ответить на контрольные вопросы.

**Задание 1**

Решите следующие задачи на определение количества информации.

* Сообщение, записанное буквами из 16 символьного алфавита, содержит 7 символов. Найти количество информации в сообщении.
* Алфавит содержит 8 букв. Найти количество информации, которое несут 5 букв этого алфавита.
* В коробке лежат 10 красных шариков и 16 синих шариков. Вынули один шарик. Найти количество информации, что это синий шарик.

**Задание 2**

Источник сообщений выдает символы из алфавита A={a1, a2, a3, a4} с вероятностями:

P1=0.2; P2=0.3; P3=0.1; P4=0.4.

Найти количество информации. Все вычисления выполнить в программе MS Excel.

**Задание 3**

Выберите вариант задания и решите в MS Excel  задачу. В коробке лежат X красных, Y желтых, Z синих шаров. Какое количество информации несет сообщение о цвете вынутого шарика?

**Варианты заданий**

|  |  |  |  |
| --- | --- | --- | --- |
| № варианта | X | Y | Z |
| 1 | 9727 | 16154 | 8407 |
| 2 | 15463 | 30981 | 4960 |
| 3 | 8462 | 28728 | 7733 |
| 4 | 8431 | 22394 | 6334 |
| 5 | 7556 | 13181 | 8184 |
| 6 | 11873 | 31755 | 8086 |
| 7 | 14832 | 20337 | 5737 |
| 8 | 13899 | 19917 | 4307 |
| 9 | 8908 | 26594 | 4248 |
| 10 | 9272 | 28438 | 8365 |
| 11 | 14887 | 19417 | 4478 |
| 12 | 15798 | 14733 | 7166 |

**Задание** **4**

Выберите вариант задания и для заданного текста (количество символов в тексте *М*) найдите:

1) Мощность алфавита (К) источника (исходя из заданного текста);

2) Количество появлений каждой i-ой буквы в тексте (*ni*);

3) Определите для каждой буквы алфавита вероятность ее появления по следующей формуле:



4) Для каждой буквы алфавита найти значение энтропии (количества информации), содержащееся в i-ой букве по формуле:



5) Вычислите энтропию источника русских букв по формуле:



6) Определите максимальное значение энтропии источника при условии, что появление всех символов равновероятно.

7) Постройте график зависимости энтропии буквы *hi* от вероятности ее появления *pi*.

8) Все вычисления и график выполнить в программе MS Excel. Примерный вид расчета представлен в таблице 2.

Таблица 2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | А | B | C | D | E | F |
| 1 |  |  |  |  |  |  |
| 2 |  | Буквы  алфавита | Количество символов | Pi | Энтропия буквы | Энтропия источника |
| 3 |  | А | 4 | 0,08 |  |  |
| 4 |  | Б | 2 | 0,04 |  |  |
| … |  | … | … | … | … | … |
| 33 |  | Я | 3 | 0,06 |  |  |
| 34 |  | пробел | 9 | 0,12 |  |  |
| 35 |  |  | Значение М | 1 |  | Значение H |

В ячейке С35 формула = СУММ(С3:С34)

В ячейках D3-D34 формула = Ci/C35 (i – меняется от 3 до 34)

В ячейке D35 формула = СУММ(D3:D34)

В ячейках E3-E34 формула =-log(Di;2) (i – меняется от 3 до 34)

В ячейках F3-F34 формула =Ei\*Di (i – меняется от 3 до 34)

В ячейке F34 формула = СУММ(F3:F34)

**Варианты заданий**

1. В Карагандинской области работают предприятия по добыче угля, машиностроения, металлообработки и пищевой промышленности.
2. Для обнаружения и исправления ошибок, возникающих при передаче сообщений, применяют корректирующие коды.
3. На сегодняшний день Караганда является одним из крупнейших промышленных, научных и культурных центров Казахстана.
4. Семантической информацией называют информацию, содержащуюся в высказываниях и передаваемую через значения единиц речи.
5. Измерительной информацией называют информацию о значениях физических величин, а также данные о погрешности измерений.
6. Ключевой информацией называют такие новые сведения, которые не могут быть подсказаны ни контекстом, ни ситуацией.
7. Документальная информация закрепляется посредством какой-либо знаковой системы на бумаге или ином материальном носителе.
8. Генетическая информация определяет морфологическое строение, рост, развитие, обмен веществ, а также иные признаки организма.
9. Деятельность по сбору и обработке существующей и созданию новой информации называется информационной деятельностью.
10. Экономическая информация – это документированные сведения, отражающие состояние экономических процессов.
11. Если же под воздействием помех разрешенная кодовая комбинация перешла в другую разрешенную комбинацию, то ошибка не будет обнаружена.
12. В статистической теории информации в качестве меры количества информации используют энтропию.

**Содержание отчета**

1. Титульный лист.
2. Результаты выполненных работ.
3. Ответы на контрольные вопросы.

# Контрольные вопросы

# Какая мера используется для количественной оценки информационной емкости технических систем?

# Какую формулу предложил Р. Хартли для измерения количества информации?

# Какую формулу предложил Клод Шеннон для вычисления количества информации и энтропии?

# В каком случае формула Шеннона совпадает с формулой Хартли?

# Что такое энтропия?