

Расчетное задание 1

Применение формулы Байеса для принятия решений.

В данном задании каждый студент получает исходные данные в виде набора файлов, имеющих названия вида task_1_XXXX.txt или task_1_XXXX.txt.zip, где XXXX – шифр, определяющий тип исходных данных и соответствующую ему последовательность расчетов и действий, представленных ниже.

Все текстовые файлы имеют кодировку Unicode UTF-8.

Возможные следующие варианты исходных данных, представленные в таблице ниже.

№ п/п	Шифр	Задача
1.1	balls	Вытягивание шаров из одной корзины с определением числа шаров каждого цвета.
1.2	cards	Угадывание набора случайно выбранных карт.
1.3	ball_boxes	Вытягивание шаров из случайно выбранной корзины. Угадывание корзины.
1.4	ball_boxes_arrange	Вытягивание шаров из случайно перемешанных вначале корзин. Угадывание порядка расположения корзин.
1.5	words	Угадывание слов по информации о буквах
1.6	info_messages	Передача сообщения в двоичном коде по каналу связи с помехами
1.7	liars_guess_letter	Задача о лгунах, нужно угадать букву
1.8	liars_guess liars	Задача о лгунах, нужно угадать, где какой лгун
1.9	random_int_number	Случайное число из таблицы целых чисел
1.10	shoot_guess prizes	Стрельба по мишеням и угадывание призов
1.11	lamp_check	Проверка лампочек и угадывание числа неисправных

Правила оформления

Формат файла отчета – docx (Word), odt, latex (один на выбор) + PDF-версия (не обязательно).

Отчет должен иметь строгую и лаконичную структуру с нумерацией разделов и подразделов. Типовая структура отчета: титульный лист, задание, теоретические основы, практические результаты, анализ результатов, выводы.

Все рисунки, формулы, графики, таблицы должны иметь подпись и быть пронумерованы. Страницы также должны иметь нумерацию.

В отчете все выполняемые расчеты необходимо подкреплять используемыми при этом формулами и листингами программ.

Не допускается слишком разреженное представление информации, вывод "сырых" данных без предварительной обработки. Большие объемы данных в таблицах приводятся либо с прореживанием, либо приводится только начальный и конечный фрагменты. Следует отдавать предпочтение визуальному представлению.

Не следует приводить в отчете результаты (данные, таблицы, графики), если они фактически повторяют результаты, уже приведенные где-то выше. В этом случае достаточно сослаться на аналогичные результаты и не приводить подробное описание.

Перед выполнением любого эксперимента необходимо заранее мысленно представить ожидаемый результат. После эксперимента следует сравнить ожидаемый результат с реальным и на основании этого выполнить анализ расхождений и попытаться их объяснить.

Вспомогательные материалы

Все программы, по которым выполняются расчеты и построения, должны быть приведены в папке рядом с файлом отчета.

Язык/среда программирования/расчетов

Для вычислений может быть использована любая подходящая среда или язык программирования. Но в целом предпочтение отдается Python, Matlab, R, т.к. они более приспособлены для научных вычислений и расчетов.

Описание заданий

1.1 Вытягивание шаров из одной корзины с определением числа шаров каждого цвета.

Корзина с шарами разных цветов. Количество шаров известно, нужно угадать распределение по цветам. Вытягивание с возвращением по несколько шаров.

Описание

В корзине N шаров разных цветов (всего m цветов). N и m известно.

Игрок последовательно тянет из корзины по d шаров, запоминает цвета, возвращает шары в корзину, шары перемешиваются и так далее. Всего выполняется K извлечений.

Необходимо определить количество шаров каждого цвета в корзине.

Исходные данные (task_1_balls.txt)

Соответствующий файл исходных данных имеет название task_1_balls.txt. Листинг с примером такого файла приведен ниже:

```
N: 280, m: 6, Nexр=10000
d_diap: [2, 3]
colors: ['Red', 'White', 'Black', 'Green', 'Blue', 'Yellow']
Experiments:
# 1, balls: Blue, Green
# 2, balls: Blue, Yellow, Blue
# 3, balls: Blue, Black, Black
# 4, balls: Blue, Yellow
...
```

В данном примере в корзине 280 шаров 6-ти цветов (красный, белый, черный, зеленый, синий, желтый). Всего проводится 10000 опытов. В первом эксперименте вытягиваются синий и зеленый шары, во втором – синий, желтый, синий и т.д.

Задание

1а. После каждого k извлечения необходимо вычислить ряд распределения апостериорных вероятностей определенных конфигураций сочетаний числа шаров $P_k(n_1, n_2, \dots, n_m)$, $\sum_{i=1}^m n_i = N$, где n_i – количество шаров i -го цвета. Представить соответствующие результаты визуально на графике в форме диаграмм распределений вероятностей конфигураций по ходу изменения числа опытов.

Зам. В самом начале все вероятности следует принять равными. Т.к. общее количество вариантов различных конфигураций сочетаний шаров равно:

$$N_{\text{общ}} = \overline{C_N^{m-1}} = C_{N+m-1}^{m-1},$$

то начальные вероятности равны:

$$P_0(n_1, n_2, \dots, n_m) = \frac{1}{C_{N+m-1}^{m-1}}, \forall n_1, n_2, \dots, n_m: \sum_{i=1}^m n_i = N$$

Зам. Общее количество конфигураций может быть довольно большим, поэтому следует предложить некоторую эвристическую схему, в соответствии с которой на каждом шаге следует рассматривать только определенный набор гипотез с вероятностью не ниже некоторого порога.

1б. Определять после каждого извлечения, какие конфигурации имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятных конфигураций.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2а. Вычислить вероятности нахождения в корзине определенного числа шаров для каждого цвета отдельно после каждого извлечения. Построить соответствующие графики зависимостей изменения этих вероятностей от числа опытов.

Зам. В данном варианте при рассмотрении отдельно каждого шага всего будет $(N + 1)$ гипотеза о количестве этих шаров.

2b. Определять для каждого цвета на каждом шаге наиболее вероятные значения числа шаров и визуализировать их.

2c. Сравнить полученные отдельные наиболее вероятные значения числа шаров заданного цвета с наиболее вероятной конфигурацией цветов в совокупности.

2d. Построить зависимость числа превалирующих гипотез от количества проведенных опытов.

3a. Определить приближенно количество шаров каждого цвета исходя из средних частот вынимания шаров определенного цвета. Построить соответствующие графики оценки среднего числа шаров каждого цвета как функцию от количества извлечений.

3b. Сравнить результаты с п. 2 и 3.

Зам. *Эволюцию динамики изменения* характеристик следует выполнять с грамотно подобранным шагом дискретности по времени. Не требуется визуализировать каждый шаг, но при этом важно не пропустить резких изменений. Наиболее рациональным вариантом является контроль изменения характеристик и адаптация шага дискретизации по времени в зависимости от этого. Т.е. если динамики по сути нет, то и строить одни и те же графики не имеет никакого смысла и наоборот, если происходит быстрое изменение какой-либо характеристики, следует уменьшить шаг дискретности и выполнять визуализацию более часто.

1.2. Угадывание комбинации карт

Описание

Ведущий вначале достает из полной игровой колоды в 52 карты какие-то N карт ($N < 52$). Затем он по очереди извлекает по d карт и сообщает второму игроку масть и номинал карт (от 2 до туза), после чего возвращает их обратно. Особенность игры заключается в том, что ведущий может соврать, исказив масть карты, а также ее номинал – соответствующие вероятности заданы и равны p_1 и p_2 .

Задача игрока – угадать, какие карты находятся у ведущего.

Исходные данные (task_1_cards.txt)

$N=8, d=3, p_1=0.2653, p_2=0.2406, N_{exp}=10000$

1, cards:♠:7, ♦:7, ♣:3

2, cards:♥:3, ♦:3, ♠:6

3, cards:♥:K, ♥:2, ♣:7

4, cards:♣:3, ♠:6, ♦:3

...

В данном примере ведущий загадал 8 карт, каждый раз он извлекает по 3 карты. С вероятностью 0.2653 он каждый раз фальсифицирует масть вытянутой карты, с вероятностью 0.2406 – ее номинал. Всего проводится 10000 экспериментов. В первый опыт ведущий сообщает карты :♠:7, ♦:7, ♣:3, во второй опыт - ♥:3, ♦:3, ♠:6 и т.д.

Задание

1а. После каждого k опыта необходимо вычислить ряд распределения апостериорных вероятностей определенных сочетаний карт. Представить соответствующие результаты визуально на графике в форме диаграмм распределений вероятностей конфигураций по ходу изменения числа опытов.

Рассмотреть 3 случая – карты по мастям, карты по номиналам и карты по мастям и номиналам.

Зам. *Общее количество гипотез* зависит от того, что использовать в качестве гипотезы.

Если в качестве гипотез рассматривать набор карт, учитывающий масть и номинал, то общее количество комбинаций (гипотез) равно C_{52}^N .

Если в качестве гипотез рассматривать количество карт каждой масти, то число гипотез равно $N_{общ} = \overline{C_N^{4-1}} = C_{N+3}^3$

Если в качестве гипотез рассматривать количество карт каждого номинала, то число гипотез равно $N_{общ} = \overline{C_N^{13-1}} = C_{N+12}^{12}$

1б. Определять после каждого извлечения, какие комбинации карт имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятных комбинаций карт.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2а. Определить приближенно карты у ведущего исходя из средних частот их появления. Построить соответствующие графики изменения частот появления карт в зависимости от опыта.

2б. Сравнить результаты с п. 1.

1.3. Вытягивание шаров из случайно выбранной корзины. Угадывание корзины.

Набор корзин с шарами. Требуется угадать, где какая корзина путем последовательного извлечения шаров

Описание

Имеется N корзин, в каждой из которой находится известная комбинация шаров различного цвета. Игрок вначале случайно выбирает одну из корзин и затем последовательно K раз вынимает из нее по несколько $d = 2 \div 3$ шаров с возвращением. Перед каждым подходом игрока к корзине ведущий с заданной вероятностью может случайно заменить ее на какую-то другую.

По истечении опытов игроку необходимо определить, к какой из N корзин он скорее всего подходил.

Исходные данные (task_1_ball_boxes.txt)

```
Source data:
n_boxes: 6, m: 4, d: 4, p_change_box: 0.100, nExp = 10000
Box 1. Total: 250. Red: 39, White: 80, Black: 46, Green: 85
Box 2. Total: 240. Red: 41, White: 86, Black: 54, Green: 59
Box 3. Total: 300. Red: 63, White: 11, Black: 118, Green: 108
Box 4. Total: 230. Red: 96, White: 77, Black: 48, Green: 9
Box 5. Total: 220. Red: 22, White: 69, Black: 42, Green: 87
Box 6. Total: 280. Red: 76, White: 93, Black: 48, Green: 63
Experiments:
# 1, Balls: White, Black, Black, Green
# 2, Balls: Black, Black, Red, White
# 3, Balls: White, Black, Green, Red
# 4, Balls: White, White, Red, Green
...
```

В примере есть 6 корзин с шарами 4-х цветов. В первой унре 250 шаров, 39 красных, 80 белых, 46 черных и 85 зеленых и т.д. Извлекается в каждый опыт по $d=4$ шара. Вероятность замены искомой корзины равна 0.1. Всего 10000 опытов. В первом опыте вытягиваются белый, черный, черный, зеленый шары и т.д.

Задание

1а. После каждого k опыта необходимо вычислить ряд распределения апостериорных вероятностей гипотез – к какой из корзин подошел игрок. Представить соответствующие результаты визуальнo на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез.

Зам. В данной задаче всего N гипотез.

1b. Определять после каждого извлечения, какие корзины имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятной корзины.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2а. Определить приближенно частоту вынимания шаров каждого цвета из корзины (экспериментальный профиль). Рассчитать теоретические вероятности вынимания шаров каждого цвета из каждой корзины – получится N теоретических профилей для каждой корзины.

2b. Сравнить теоретический профиль с каждым из полученных экспериментальным и найти наиболее похожий. Сравнить с полученным результатов в п. 1

2с. Построить графики изменения экспериментальных профилей для различного количества опытов.

1.4. Вытягивание шаров из случайно перемешанных вначале корзин. Угадывание порядка расположения корзин.

Случайно перемешанные корзины с известным содержимым. Извлечение шаров с циклическим переключением.

Описание

N пронумерованных корзин с известным распределением шаров (различным) случайным образом переставляют. Затем игрок подходит к первой корзине, последовательно вынимает d шаров, запоминает их и кладет обратно, потом по аналогии вытаскивает d шаров из 2-ой корзины, смотрит их и возвращает и т.д. в цикле. Такая процедура выполняется несколько раз. Требуется, владея исходной информацией о распределении шаров по корзинам и о том, какие шары вынимались и перекладывались из корзин, вычислить на каждом шаге:

1. для каждой корзины – вероятность того, что она имеет номер $i, i \in \{1, 2, \dots, N\}$.
2. для набора их всех корзин – вероятность того, что этот набор корзин имеет последовательные номера $i_1, i_2, \dots, i_N, i_j \in \{1, 2, \dots, N\}, j = 1, 2, \dots, N$.

3. наиболее правдоподобную комбинацию номеров корзин.

task_1_ball_boxes_arrange.txt

Описание исходных данных

Исходные данные данного эксперимента совпадают с исходными данными в п. 1.3. Отличием является только то, что игрок в цикле вытягивает шары по очереди из каждой из корзин и никто не переставляет корзины. Пример исходных данных приведен ниже.

Source data:

n_boxes: 5, m: 5, d: 4, nExp: 10000

Box 1. Total: 230. Red: 11, White: 38, Black: 52, Green: 66, Blue: 63

Box 2. Total: 210. Red: 28, White: 7, Black: 69, Green: 39, Blue: 67

Box 3. Total: 290. Red: 84, White: 10, Black: 5, Green: 97, Blue: 94

Box 4. Total: 270. Red: 82, White: 10, Black: 50, Green: 77, Blue: 51

Box 5. Total: 270. Red: 63, White: 76, Black: 75, Green: 11, Blue: 45

Experiments:

1, Balls: Red, Red, Black, Green

2, Balls: Green, Black, Green, Blue

3, Balls: Blue, Green, Green, White

Задание

1а. После каждого k опыта необходимо вычислить ряд распределения апостериорных вероятностей гипотез – о том, какая корзина имеет какой номер. Представить соответствующие результаты визуально на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез.

Зам. В данной задаче количество гипотез равно количеству вариантов переставить корзины, т.е. $N!$

1б. Определять после каждого извлечения, какая гипотеза о порядке корзин имеет наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятной корзины.

1с. Построить зависимость числа преобладающих гипотез от числа проведенных опытов.

2а. Рассмотреть каждую корзину по отдельности и в качестве гипотез выдвигать то, какой номер имеет соответствующая корзина. Всего таким образом для одной корзины получится N гипотез. Вычислить для каждой из корзин распределения вероятностей гипотез (о том, какой у нее номер) после каждого опыта. Представить результаты визуально по аналогии с п. 1а.

2б. Определить для каждой корзины наиболее вероятную гипотезу на каждом шаге и визуализировать эволюцию этой гипотезы.

2с. Объединить результаты для всех корзин, получить наиболее вероятную перестановку корзин и сравнить ее с полученной перестановкой в п.1. Провести анализ сравнения.

3а. Определить приближенно частоту вынимания шаров каждого цвета из каждой корзины (получится N экспериментальных профилей). Рассчитать теоретические вероятности вынимания шаров каждого цвета из каждой корзины – получится N теоретических профилей для каждой корзины.

3б. Сопоставить теоретические профили с каждым из полученных экспериментальных и найти их наиболее правдоподобное соответствие. Сравнить с полученным результатов в п. 1 и 2. Провести анализ сравнения.

3с. Привести графики изменения экспериментальных профилей для различного количества опытов.

1.5. Угадывание слов по информации о буквах

Описание

Ведущий загадывает слово. Затем в начале каждого хода ведущий сообщает некоторую информацию о слове:

- он выбирает наугад одну из букв в слове и сообщает ее без уточнения ее позиции в слове;

- он выбирает наугад одну из букв в слове и сообщает ее позицию и определенную информацию о ней (гласная/согласная/соединительная, звонкая/глухая, заглавная/строчная);

Игрок на основании информации, полученной от ведущего, строит гипотезы о том, какое это слово. Количество букв в слове в общем случае неизвестно.

1. Попытаться вначале угадать слово самостоятельно, используя как можно меньше подсказок.

2. Используя предоставленную базу слов, вычислить для каждого из слов последовательно вероятности того, что загадано именно это слово. Определить количество подсказок, начиная с которого можно с определенной уверенностью утверждать о том, какое слово загадано.

Построить распределения числа слов-кандидатов как функцию от номера подсказок.

По логике количество слов-кандидатов с каждым шагом должно уменьшаться.

Загаданное слово принадлежит одному из классов:

Город России (cities_rus.txt)

Страна мира (countries.txt)

Столица (capitals.txt)

Русское имя (names_all.txt).

Река (rivers.txt)

Русское существительное (имя нарицательное) (russian_nouns.txt)

В скобках выше приведены названия текстовых файлов, в которых приведены все возможные варианты существительных для каждого из классов.

Исходные данные (task_1_words.txt)

Ниже приведен пример файла с исходными данными.

Word type: capitals, nExp = 1000

Experiments:

1: "б"

2: "е"

3: "т"

4: "е"

5: "т"

6: "е"

7: "т"

8: 3: согласная звонкая

9: "р"

10: 6: строчная согласная глухая

11: "е"

Используется набор слов – столицы мира., количество опытов равно 1000. Далее приводятся результаты каждого из опытов. С 1 по 7 опыт сообщаются буквы б, е, т. В 8-м опыте сообщается, что 3-я буква – звонкая согласная и т.д.

Задание

1а. После каждого k опыта необходимо вычислить ряд распределения апостериорных вероятностей гипотез – какое слово загадано. Представить соответствующие результаты визуально на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез. Сузить количество гипотез можно по аналогии с п. 3а.

Зам. В данной задаче количество гипотез равно количеству слов в соответствующем файле со словами.

1b. Определять после каждого опыта, какое слово имеет наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятного слова.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2а. Рассмотреть каждую букву в слове по отдельности и в качестве гипотез выдвигать то, что это за буква (сам символ, заглавная/строчная, гласная/согласная и т.д.). Для гипотезы о том, что за символ общее количество гипотез равно общему количеству возможных символов, для гипотез о классах буквы количество гипотез равно количеству классов.

Вычислить для каждой из букв распределения вероятностей гипотез после каждого опыта. Представить результаты визуально по аналогии с п. 1а.

2b. Определить для каждой буквы наиболее вероятную гипотезу на каждом шаге и визуализировать эволюцию этой гипотезы.

2с. Объединить результаты для всех букв, получить наиболее вероятное слово и сравнить его с полученным п.1. Провести анализ сравнения.

Зам. Возможной проблемой в данном пункте будет то, что общее количество букв не задано. Возможным способом ее решения является отбрасывание всех букв, у которых номер в слове больше, чем номер последней буквы, сообщенной ведущим.

3а. Определить приближенно частоту тех или иных сообщаемых символов. Исходя из этого можно оценить примерно, сколько каждого из символов в слове. Затем сопоставьте эту информацию с тем, что ведущий говорит о позициях букв и их типе. В результате можно сузить количество возможных слов, используемых в качестве гипотез.

3b. Составьте для каждого слова из базы теоретические профили, содержащие в себе некоторую статистику – количество символов каждого типа, позиция с учетом типа символа. Далее из п. 3а определите аналогичный экспериментальный профиль.

3с. Сопоставить теоретические профили с экспериментальным и найдите их наиболее правдоподобное соответствие в форме слова (слов). Сравнить с полученными результатами в п. 1 и 2. Провести анализ сравнения.

3d. Привести графики изменения экспериментальных профилей для различного количества опытов.

1.6 Передача сообщения в двоичном коде по каналу связи с помехами

По каналу связи передаются буквы $[x_1; x_2; \dots; x_n]$ в двоичном коде. Последовательность переданных букв образует сообщение. Канал симметричный, вероятность искажения каждого отдельного символа (бита) равна q . В результате однократной передачи сообщения $X = [x^{(1)}, x^{(2)}, \dots, x^{(k)}]$ на приемной стороне принято сообщение $Y_1 = [y_1^{(1)}, y_1^{(2)}, \dots, y_1^{(k)}]$. В результате повторной передачи того же слова на приемной стороне принято слово $Y_2 = [y_2^{(1)}, y_2^{(2)}, \dots, y_2^{(k)}]$. В результате последней (m -й) передачи того же слова на приемной стороне принято слово $Y_m = [y_m^{(1)}, y_m^{(2)}, \dots, y_m^{(k)}]$.

Передаваемые буквы (алфавит) и их код приведен в табл.1 приложения. Для каждого студента есть свой вариант в виде текстового файла, названного по фамилии и имени студента. В файле задается число букв в сообщении, разрядность кода (количество бит, используемых при передаче одной буквы), шум (вероятности искажения символов 0-1 и 1-0, а также вероятность стирания символов), число посылок m и набор из m посылок (принятых сообщений $Y^{(i)}$).

Исходные данные (task_1_info_messages.txt)

Ниже приведен пример файла с исходными данными.

Число букв: 230	Разрядность кода: 7 бит	P_error_01: 0.166	P_error_10: 0.166	P_erase: 0.000	Число посылок: 10
Зашумленная посылка № 01: 0101000 1101101 1011110 1010010 0001011 0101010 0010000 0110110 1011111 1101011 0010100 0111001 0111010 1111101 0101011 1110010 0011100 0101111 1101100 0110110 1001011 1001101 1000110 0100100 0110011 1010010 0011110 0110000 0100111 0011011 0101111 1000101 1010111 0000000 0000111 0000011 0100000 1001001 0000000 0000001 1001110 0000001 1001001 1001100 1000100 ...					
....					

Возможно 4 типа канала связи:

- двоичный симметричный канал без стирания – $P_{10}=P_{01}$, $P_{\text{erase}}=0$
- двоичный симметричный канал со стиранием – $P_{10}=P_{01}$, $P_{\text{erase}} \neq 0$
- двоичный несимметричный канал без стирания – $P_{10} \neq P_{01}$, $P_{\text{erase}}=0$
- двоичный несимметричный канал со стиранием – $P_{10} \neq P_{01}$, $P_{\text{erase}} \neq 0$

В случае каналов со стиранием помимо 0 и 1 в сообщениях фигурирует символ '-', обозначающий, что произошло стирание символа из-за соответствующей ошибки при передаче.

Задание ч.1. Последовательная передача одинаковых сообщений

1.1. Определение переданного сообщения

- вычислите априорное распределение вероятностей исходных букв алфавита $p(x_i)$, рассмотрите два случая (все дальнейшие расчеты в п. 1.1 и 1.2 необходимо будет проделать для этих двух вариантов):
 - все символы равновероятны;
 - вероятности букв задаются исходя из известной информации о частоте букв в русском алфавите (таблица 2);
- вычислите апостериорное распределение вероятностей после 1-й, 2-й и m -й передач для каждой s буквы сообщения - $P(x_i / y_1^{(s)})$, $P(x_i / y_1^{(s)} y_2^{(s)})$, $P(x_i / y_1^{(s)} y_2^{(s)} y_3^{(s)})$, ... , $P(x_i / y_1^{(s)} y_2^{(s)} \dots y_m^{(s)})$; при расчете используются формулы (3), (4) и (6); следует учитывать, что для повторных посылок априорные вероятности будут совпадать с апостериорными для предыдущей посылки (см. формулу (6)).

- постройте график изменения апостериорного распределения вероятностей на примере любой 1-ой передаваемой буквы сообщения (n передач => n графиков друг под другом, на графике по оси X – номер символа, по оси Y – вероятность)
- по максимуму апостериорной вероятности определите наиболее вероятные буквы и составьте вариант исходного переданного сообщения для 1-й, 2-й и m-й посылки;
- проанализируйте, как повторные передачи сказались на принятии решения.

1.2. Расчет энтропии и количества информации

- Выберите в посылаемом сообщении произвольную букву (под номером s), далее все вычисления будут относиться к этой букве;
- Определите апостериорные вероятности, рассматривая каждую передачу независимо от другой; схема вычислений следующая $P(x_i) \rightarrow P(y_j / x_i) \rightarrow P(y_j) \rightarrow P(x_i / y_j)$; при расчете используйте формулы (3), (4) и (5).
- Определите условные энтропии $H(X / y_j)$ на сообщения y_j по формуле (2), среднее количество информации $I(X, y_j)$ об X, содержащееся в y_j по формуле (8).
- Определите среднюю условную энтропию $H(X/Y)$ по формуле (7) и среднюю взаимную информацию $I(X,Y)$ по формуле (9).
- Постройте графики изменения условной энтропии $H(X / y_j)$ и количества информации $I(X, y_j)$ от номера посылки.

1.3. Сравните результаты п. 1.1 и 1.2 при различных заданиях изначальных априорных вероятностей.

Задание ч. 2 Передача сообщения путем многократного дублирования

Рассмотрите m передач сообщений как передачу одного большого сообщения, в котором каждый символ многократно (m-кратно) дублируется

На входе $X_{new} = [x_{new}^{(1)}; x_{new}^{(2)}; \dots; x_{new}^{(k)}] = [x^{(1)} x^{(1)} \dots x^{(1)}, x^{(2)} x^{(2)} \dots x^{(2)}, \dots, x^{(k)} x^{(k)} \dots x^{(k)}]$,

На выходе $Y_{new} = [y_{new}^{(1)}; y_{new}^{(2)}; \dots; y_{new}^{(k)}] = [y_1^{(1)} y_2^{(1)} \dots y_m^{(1)}, y_1^{(2)} y_2^{(2)} \dots y_m^{(2)}, \dots, y_1^{(k)} y_2^{(k)} \dots y_m^{(k)}]$.

При этом новый алфавит по сути – m-кратное дублирование старого алфавита:

$[x_{new1}; x_{new2}; \dots; x_{newn}] = [x_1 x_1 \dots x_1; x_2 x_2 \dots x_2; \dots; x_n x_n \dots x_n]$

2.1. Определение переданного сообщения

- вычислите априорное распределение вероятностей исходных букв алфавита $p(x_i)$ – рассмотрите два случая (по аналогии с п.1. все дальнейшие расчеты в п. 2.1 и 2.2 необходимо выполнить для этих двух вариантов):
 - все символы равновероятны;
 - вероятности букв задаются исходя из известной информации о частоте букв в русском алфавите (таблица 2);
- вычислите апостериорное распределение вероятностей для каждой 1 буквы сообщения - $P(x_{newi} / y_{new}^{(i)})$; при расчете используются формулы (3), (4);
- постройте график апостериорного распределения вероятностей на примере 1-ой передаваемой буквы сообщения
- по максимуму апостериорной вероятности определите наиболее вероятные буквы и составьте вариант исходного переданного сообщения – сравните его со случаем передачи сообщений последовательно

2.2. Расчет энтропии и количества информации

- Выберите в посылаемом сообщении ту же букву, что и использовалась в п. 1.2, далее все вычисления будут относиться к этой букве;
- Определите апостериорные вероятности; схема вычислений следующая $P(x_{new_i}) \rightarrow P(y_{new} / x_{new_i}) \rightarrow P(y_{new}) \rightarrow P(x_{new_i} / y_{new})$; при расчете используйте формулы (3), (4) и (5).
- Определите условную энтропию $H(X_{new} / y_{new})$ на сообщения y_{new} по формуле (2), среднее количество информации $I(X, y_{new})$ об X , содержащееся в y_{new} по формуле (8).
- Определите среднюю условную энтропию $H(X_{new} / Y_{new})$ по формуле (7) и среднюю взаимную информацию $I(X, Y_{new})$ по формуле (9).

Сравните результаты (энтропия, количество информации) с п.1.2 и объясните их.

П.6.1 Теоретические основы

Пусть X - ансамбль возможных сообщений, которые могут быть переданы по каналам связи. Априорные вероятности, с которыми на передающей стороне может появиться i -ое сообщение, обозначим через $p_i = p(x_i)$.

В качестве характеристики априорной неопределенности генерации того или иного сообщения на передающей стороне Шеннон предложил применить следующий функционал энтропии:

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (1)$$

В этом функционале используется логарифм по основанию 2 в связи с тем, что в цифровых каналах связи информация обычно представляется в двоичных кодах, и энтропия измеряется в *битах*.

Энтропия является характеристикой неопределенности состояния ансамбля X в отличие от характеристики возможности наступления того или иного события, то есть вероятности. Энтропия не зависит от значений, которые может принимать тот или иной элемент ансамбля и от способа его представления. В частности, элементами ансамбля могут быть словесные описания или фотографические изображения. Энтропия характеризует степень хаотичности ансамбля и принимает максимальное значение, когда вероятности $p(x_i)$ одинаковы.

Инструментарий теории информации, выдвинутой К.Шенноном, эффективно используется в теории и практике передачи информации и кодирования. В этих применениях ансамбль X - ансамбль возможных сообщений. Основные результаты теории информации относятся к передаче информации в двоичном коде и к каналам, ориентированным на передачу двоичных кодов. Передача такой информации осуществляется двоичными кодовыми словами, и при каждой такой передаче искажение отдельного символа в передаваемом слове заключается в том, что вместо 1 получатель принимает 0, или вместо нуля получатель принимает 1. При таких условиях канал передачи двоичной информации характеризуется вероятностью искажения символа. Схематическое представление простейшего канала связи, а именно, двоичного симметричного канала представлено на рис 5.

Симметричным этот канал называется потому, что вероятности искажения символов '0' и '1' одинаковы и равны q , как это показано на рисунке. Вероятность p - вероятность неискаженной передачи символа.

В случае, если канал не симметричный, вероятности искажения для разных символов различаются. В случае, если канал связи со стиранием, то с ненулевой вероятностью передаваемый символ может быть стерт – в этом случае фиксируется, что произошла ошибка при передаче (символ '-').

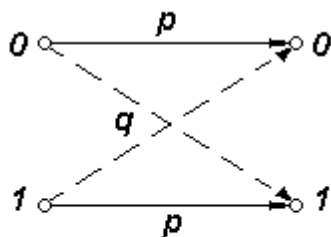


Рис.5. Двоичный симметричный канал

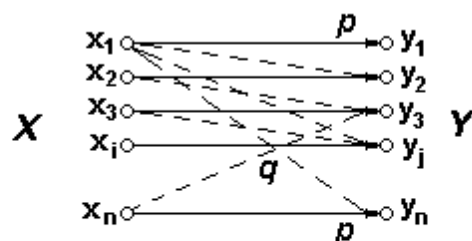


Рис.6. Передача сообщений по каналу с искажениями

Поскольку каждое сообщение $x_i \in X$ передается кодовым словом, которое может состоять из нескольких символов '0' или '1', возможностей искажения слова при передаче

по такому каналу гораздо больше (см. рис. 6, где принимаемые сообщения обозначены через $Y_1, Y_2, \dots, Y_j, \dots \in Y$). Вероятность искажения слова в общем случае отличается от вероятности искажения отдельного символа.

Понятно, что при передаче по каналам связи сообщения (слова) $x_i \in X$ и при получении сообщения $y_j \in Y$, несмотря на искажения в каналах энтропия ансамбля X , несомненно, уменьшится и будет исчисляться условной энтропией на сообщение

$$H(X/Y) = - \sum_{i=1}^n p(x_i/y_j) \cdot \log_2 p(x_i/y_j). \quad (2)$$

Прим. Вероятности $p(x_i/y_j)$ называются апостериорными. Для их расчета необходимо использовать формулу Байеса

$$p(x_i/y_j) = \frac{p(y_j/x_i)p(x_i)}{p(y_j)} = \frac{p(y_j/x_i)p(x_i)}{\sum_k p(y_j/x_k)p(x_k)} \quad (3)$$

$p(x_i)$ - априорные вероятности, вначале они могут быть выбраны произвольно и как правило они принимаются одинаковыми и равными $1/N$, где N – общее количество слов x_i .

$p(y_j/x_i)$ - условная вероятность приема y_j при условии, что было послано x_i . Данную вероятность легко найти следующим образом. Если общее количество разрядов в слове k , в t разрядах произошла ошибка (они проинвертировались), то вероятность очевидно равна

$$p(y_j/x_i) = p^{k-t} q^t \quad (4)$$

Здесь по-прежнему p – вероятность правильной передачи одного бита, $q=1-p$ – вероятность ошибки.

$$p(y_j) = \sum_k p(y_j/x_k)p(x_k) \quad (5)$$

- вероятность приема y_j в данной задаче несет вспомогательный характер – она потребуется в том числе в дальнейших расчетах.

В случае, если при приеме одной и той же буквы несколько раз информация накапливается, для расчета апостериорных вероятностей последующих приемов в качестве априорных вероятностей необходимо использовать апостериорные вероятности после предыдущего принятого сообщения, т.е.:

$$p(x_i/y_1 y_2 \dots y_j) = \frac{p(y_j/x_i)p(x_i/y_1 y_2 \dots y_{j-1})}{p(y_j)} = \frac{p(y_j/x_i)p(x_i/y_1 y_2 \dots y_{j-1})}{\sum_k p(y_j/x_k)p(x_k/y_1 y_2 \dots y_{j-1})} \quad (6)$$

Для того, чтобы характеризовать всю систему приема - передачи, используют среднюю условную энтропию:

$$H(X/Y) = \sum_{j=1}^n p(y_j) \cdot H(X/y_j) = H(X,Y) - H(Y) \quad (7)$$

$$\text{где } H(Y) = \sum_{j=1}^n p(y_j) \cdot \log_2 p(y_j), \quad H(X,Y) = - \sum_{j=1}^n \sum_{i=1}^n p(x_i, y_j) \cdot \log_2 p(x_i, y_j)$$

$H(X,Y)$ - совместная энтропия двух ансамблей, равная

$$H(X,Y) = H(X/Y) + H(Y) = H(Y/X) + H(X)$$

Количество информации об x_i , полученное в одном сообщении y_j :

$$I(x_i : y_j) = \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)}.$$

Среднее количество информации об X , полученное в сообщении y_j

$$I(X : y_j) = - \sum_{i=1}^n p(x_i / y_j) \log p(x_i) - H(X / y_j) \quad (8)$$

Количество информации измеряется в тех же единицах, что и энтропия - в битах.
Средняя взаимная информация, содержащаяся в Y об X или в X об Y :

$$I(X : Y) = \sum_{j=1}^n p(y_j) I(X : y_j) = H(X) - H(X / Y) \quad (9)$$

то есть это количество информации численно равно количеству неопределенности, устраненной при получении одного сообщения.

1.7 Задача о лгунах, вариант 1

На входе информация в виде N букв (мы их не знаем). Каждый лгун может определенным образом исказить букву. На выходе получаем различные буквы.

Требуется угадать, какие буквы передавали на входе, используя априорную информацию о лгунах.

Исходные данные (task_1_liars1.txt)

Ниже приведен пример файла с исходными данными.

```
N: 6, m: 6, Nexр=10000
lettersAll: ['a', 'б', 'в', 'г', 'д', 'е', 'ж', 'з', 'и', 'й', 'к', 'л', 'м', 'н', 'о', 'п', 'р', 'с', 'т', 'у', 'ф', 'х', 'ц',
'ч', 'ш', 'щ', 'ъ', 'ы', 'ь', 'э', 'ю', 'я']
liar order: [0, 1, 2, 3, 4, 5]
liar subs:
liar #0: with prob=0.43000000000000005 change from ['м', 'г', 'ь', 'ю', 'ш', 'щ', 'о', 'э', 'б', 'ф',
'у', 'з', 'ч', 'к', 'л'] to ['в', 'с', 'п', 'ц', 'а']
liar #1: with prob=0.46000000000000001 change from ['у', 'с', 'ю', 'а', 'э', 'ж', 'р', 'н', 'ф', 'в',
'м', 'е', 'х'] to ['л', 'к', 'б', 'й', 'о', 'и']
liar #2: with prob=0.49000000000000001 change from ['ъ', 'и', 'к', 'у', 'ж', 'е', 'п', 'д', 'л', 'ц',
'х', 'н', 'ш'] to ['ф', 'б', 'я', 'ч', 'м']
liar #3: with prob=0.45000000000000007 change from ['ц', 'о', 'д', 'к', 'н', 'э', 'ь', 'ю', 'ы', 'ж']
to ['р', 'п', 'г', 'б', 'т']
liar #4: with prob=0.41000000000000003 change from ['а', 'с', 'е', 'г', 'э', 'л', 'р', 'и', 'ж', 'ю']
to ['о', 'в', 'н', 'ь', 'д', 'з', 'у']
liar #5: with prob=0.43000000000000005 change from ['с', 'п', 'ю', 'ж', 'й', 'щ', 'х', 'г', 'л', 'н',
'ь', 'м'] to ['б', 'е', 'в', 'и', 'д', 'т']
Experiments:
# 1, letters: ['б', 'и', 'ч', 'ц', 'б', 'т']
# 2, letters: ['т', 'б', 'ч', 'б', 'у', 'ц']
# 3, letters: ['ф', 'ж', 'д', 'ч', 'б', 'з']
# 4, letters: ['д', 'ф', 'б', 'п', 'я', 'в']
```

В данном случае количество букв при каждой передаче равно $N=5$, количество лгунов $m=6$. Все варианты букв – lettersAll. Порядок лгунов – по умолчанию, т.е. они не перемешаны.

Далее идет информация по лгунам. К примеру лгун 0 меняет символы ['м', 'г', 'ь', 'ю', 'ш', 'щ', 'о', 'э', 'б', 'ф', 'у', 'з', 'ч', 'к', 'л'] на ['в', 'с', 'п', 'ц', 'а'] с вероятностью 0.43.

Далее идет информация по каждой посылке. В первый раз на выходе получены буквы letters_out: ['б', 'и', 'ч', 'ц', 'б', 'т'].

Задание

1а. После каждого k опыта необходимо вычислить ряд распределения апостериорных вероятностей гипотез по каждой из N передаваемых букв. Представить соответствующие результаты визуально на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез.

Зам. В данной задаче всего N гипотез.

1б. Определять после каждой передачи, какие последовательности букв имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятных гипотез.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2а. Определить приближенно частоты появления каждой из передаваемых букв для каждой из N передаваемых букв.

2b. Определите наиболее подходящие буквы по принципу максимума относительной частоты.

2c. Построить графики изменения динамики относительных частот для различного количества опытов.

1.8 Задача о лгунах, вариант 2

На входе информация в виде букв (мы их знаем, они меняются при каждой посылке), далее каждый лгун ее получает, может исказить и передает следующему лгуну, в результате на выходе получаем также буквы³³. Порядок лгунов не известен. Требуется угадать, где какой лгун.

Исходные данные (task_1_liars2.txt)

Ниже приведен пример файла с исходными данными.

```
N: 5, m: 6, Nexр=10000
lettersAll: ['a', 'б', 'в', 'г', 'д', 'е', 'ж', 'з', 'и', 'й', 'к', 'л', 'м', 'н', 'о', 'п', 'р', 'с', 'т', 'у', 'ф', 'х', 'ц',
'ч', 'ш', 'щ', 'ъ', 'ы', 'ь', 'э', 'ю', 'я']
liar subs:
liar #0: with prob=0.47000000000000001 change from ['х', 'с', 'в', 'ъ', 'ы', 'ж', 'и', 'к', 'а', 'ш',
'щ', 'й', 'о'] to ['р', 'э', 'п', 'ч', 'б', 'д', 'ф']
liar #1: with prob=0.45000000000000007 change from ['р', 'ф', 'г', 'э', 'с', 'ю', 'м', 'в', 'к', 'ц',
'ъ', 'н'] to ['и', 'ы', 'ш', 'б', 'я', 'з', 'щ']
liar #2: with prob=0.41000000000000003 change from ['и', 'ы', 'е', 'ч', 'в', 'х', 'э', 'м', 'у', 'б',
'ф', 'ъ', 'ж', 'щ'] to ['р', 'п', 'н', 'л', 'ц', 'а', 'о']
liar #3: with prob=0.41000000000000003 change from ['й', 'в', 'з', 'л', 'я', 'ы', 'ч', 'д', 'г', 'а',
'ъ'] to ['б', 'х', 'р', 'е', 'н', 'о', 'и']
liar #4: with prob=0.45000000000000007 change from ['ъ', 'щ', 'у', 'ю', 'е', 'р', 'ж', 'и', 'д', 'ц',
'п', 'ф', 'а'] to ['н', 'л', 'в', 'м', 'т', 'й', 'ы']
liar #5: with prob=0.41000000000000003 change from ['д', 'в', 'с', 'з', 'л', 'ч', 'у', 'б', 'о', 'ф',
'ш'] to ['е', 'н', 'п', 'п', 'г']
Experiments:
# 1, letters_in: ['н', 'щ', 'о', 'э', 'ъ'], letters_out: ['н', 'т', 'м', 'л', 'о']
# 2, letters_in: ['ы', 'ш', 'ю', 'п', 'о'], letters_out: ['е', 'ш', 'ю', 'п', 'о']
# 3, letters_in: ['ч', 'э', 'п', 'т', 'к'], letters_out: ['б', 'й', 'й', 'т', 'ш']
```

В данном случае количество букв при каждой передаче равно $N=5$, количество лгунов $m=6$. Все варианты букв – lettersAll.

Далее идет информация по лгунам. К примеру лгун 0 меняет символы ['х', 'с', 'в', 'ъ', 'ы', 'ж', 'и', 'к', 'а', 'ш', 'щ', 'й', 'о'] на ['р', 'э', 'п', 'ч', 'б', 'д', 'ф'] с вероятностью 0.47.

Далее идет информация по каждой посылке. В первый раз отправляются буквы letters_in: ['н', 'щ', 'о', 'э', 'ъ'], а на выходе буквы letters_out: ['н', 'т', 'м', 'л', 'о'].

Задание

1а. После каждого k опыта необходимо вычислить ряд распределения апостериорных вероятностей гипотез по каждому варианту перестановки лгунов (всего $m!$ гипотез). Представить соответствующие результаты визуально на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез.

Зам. В данной задаче всего $m!$ гипотез.

1б. Определять после каждой передачи, какие гипотезы имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятных гипотез.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2. Попытаться придумать частотную статистику, на основе которой можно предположить, где какой лгун находится.

1.9 Случайное число из таблицы целых чисел

Ведущий загадывает число в диапазоне от 1 до N. На каждом шаге он выбирает случайную операцию (сложение, умножение, деление без остатка) и число, далее он выполняет арифметическую операцию и сообщает результат в виде одной из цифр получившегося числа, т.е. цифру и ее порядок. Также ведущий периодически может сбрасывать все, т.е. как бы начинать игру с самого начала.

Нужно угадать число.

Исходные данные (task_1_numbers.txt)

Ниже приведен пример файла с исходными данными.

```
Nmin: 100000, Nmax: 1000000, Nexp=10000
Experiments:
# 1, result: {'op': '-', 'num': 855436, 'outType': 'digit', 'extraParams': 4, 'outVal': 5}
# 2, result: {'op': '+', 'num': 802675, 'outType': 'digit', 'extraParams': 1, 'outVal': 7}
# 3, result: {'op': '-', 'num': 557405, 'outType': 'digit', 'extraParams': 4, 'outVal': 0}
# 4, result: {'op': '*', 'num': 85, 'outType': 'digit', 'extraParams': 6, 'outVal': 5}
# 5, result: {'op': 'r', 'num': '-', 'outType': '-', 'extraParams': '-', 'outVal': '-', '-': '-'}
# 6, result: {'op': '//', 'num': 22, 'outType': 'digit', 'extraParams': 3, 'outVal': 0}
# 7, result: {'op': '+', 'num': 730946, 'outType': 'digit', 'extraParams': 1, 'outVal': 3}
```

В данном случае загаданное число находится в диапазоне от Nmin=100000 до Nmax=1000000. Количество изменений числа ведущим равно Nexp=10000. Далее идет информация по каждой попытке.

Поле op означает математическую операцию:

+ сложение, - вычитание, * умножение, // деление без остатка, r сброс к изначальному значению.

Далее параметр num в случае операций +,-,*,// содержит значение второго аргумента

Параметр outType означает тип ответа ведущего:

digit – цифра числа, sum – сумма цифр числа, - означает отсутствие ответа при сбросе.

В случае ответа типа digit параметр extraParams содержит номер цифры числа, которая будет дана в выходном параметре outVal.

Наконец параметр outVal содержит выходное значение, которое сообщает ведущий, которое зависит от типа outType и является числом, либо прочерком.

Задание

1а. После каждого k опыта необходимо вычислить ряд распределения апостериорных вероятностей гипотез по каждому варианту возможного значения числа. Представить соответствующие результаты визуально на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез.

Зам. В данной задаче всего $N_{\max}-N_{\min}+1$ гипотез.

1б. Определять после каждой попытки, какие гипотезы имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятных гипотез.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2. Попытаться придумать частотную статистику, на основе которой можно выделить полезную информацию о значении числа.

1.10 Стрельба и угадывание мишеней

Игрок стреляет по n мишеням. В каждой попытке делается несколько выстрелов по $k < n$ мишеням. За каждую мишень дается определенное число очков при попадании в нее. Также известны вероятности попасть в мишень. Игрок стреляет по выбранным мишеням и узнает сумму, которую он набрал. Требуется угадать, сколько очков дается за каждую мишень.

Исходные данные (task_1_numbers.txt)

Ниже приведен пример файла с исходными данными.

```
N: 10, Nexp=10000
prob_values: [0.25000000000000001, 0.19000000000000003, 0.27000000000000013, 0.15,
0.35000000000000002, 0.32000000000000002, 0.20000000000000004, 0.32000000000000002,
0.25000000000000001, 0.28000000000000014]
prizeValues: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Experiments:
# 1, result: {'shoots': {2: 1, 9: 1, 4: 2}, 'sum': 0}
# 2, result: {'shoots': {4: 2, 0: 2, 6: 2, 7: 1}, 'sum': 2}
# 3, result: {'shoots': {9: 1, 0: 2, 5: 2, 8: 2}, 'sum': 10}
# 4, result: {'shoots': {8: 2, 7: 1, 4: 2, 3: 1}, 'sum': 9}
```

В данном случае у нас имеется 10 мишеней. Вероятности попасть в мишени указаны в `prob_values`. Призовые значения мишеней содержатся в `prizeValues`. Но неизвестно, какой мишени какое из значений сопоставляется.

Далее идет информация по выстрелам.

В первый раз игрок стреляет 1 раз по 2 мишени, 1 раз по 9 мишени и 2 раза по 4 мишени и выбивает сумму 0.

Во второй раз игрок стреляет 2 раза по 4,0,6 мишени, 1 раз по 7 мишени и выбивает сумму 2 и т.д.

Задание

1а. После каждой стрельбы игрока необходимо вычислить ряд распределения апостериорных вероятностей гипотез по каждому варианту гипотез о том, какая мишень какому призу сопоставляется. Представить соответствующие результаты визуально на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез.

Зам. В данной задаче всего $N!$ гипотез.

1б. Определять после каждой попытки, какие гипотезы имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятных гипотез.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2. Попытаться придумать частотную статистику, на основе которой можно выделить полезную информацию о том, какой приз какой мишени сопоставляется.

1.11 Проверка лампочек

Среди N лампочек часть r неисправна. Случайно отбирается каждый раз по m лампочек и проверяется, сколько из них исправно. Требуется угадать количество неисправных лампочек.

Исходные данные (task_1_lamps_testing.txt)

Ниже приведен пример файла с исходными данными.

```
N: 131, Nexр=10000
Experiments:
# 1, result: {'m': 24, 's': 0}
# 2, result: {'m': 37, 's': 4}
# 3, result: {'m': 37, 's': 0}
# 4, result: {'m': 29, 's': 1}
```

Всего проверяется $N=131$ лампочка, количество проверок $N_{exp}=10000$. В первом эксперименте отбираются 24 лампочки, из них 0 неисправных, во втором 37, из них 4 неисправных и т.д.

Задание

1а. После каждой проверки необходимо вычислить ряд распределения апостериорных вероятностей гипотез о количестве неисправных лампочек. Представить соответствующие результаты визуально на графике в форме изменения с течением опытов диаграмм распределений вероятностей гипотез.

Зам. В данной задаче всего $N+1$ гипотеза.

1б. Определять после каждой попытки, какие гипотезы имеют наибольшую вероятность. Визуализировать эволюцию изменения наиболее вероятных гипотез.

1с. Построить зависимость числа превалирующих гипотез от числа проведенных опытов.

2. Попытаться придумать частотную статистику, на основе которой можно выделить полезную информацию о том, сколько лампочек неисправно.

