

---

# Наивный байесовский классификатор

Имя входного файла:	стандартный ввод
Имя выходного файла:	стандартный вывод
Ограничение по времени:	1 секунда
Ограничение по памяти:	256 мегабайт

Реализуйте наивный байесовский классификатор.

Априорные вероятности классов оцениваются обыкновенным частотным методом.

Для оценки вероятности встречи слов в каждом классе используется модель Бернулли с аддитивным сглаживанием (сглаживание Лапласа)  $p(x) = \frac{\text{count}(x) + \alpha}{\sum_{y \in Q} \text{count}(y) + \alpha \cdot |Q|}$ , где  $x$  — рассматриваемое событие, а  $Q$  — множество всех событий.

Каждое слово это отдельный признак с двумя возможными событиями встретилось / не встретилось.

## Формат входных данных

В первой строке содержится целое положительное число  $K$  ( $1 \leq K \leq 10$ ) — число классов.

Во второй строке содержится  $K$  целых положительных чисел  $\lambda_C$  ( $1 \leq \lambda_C \leq 10$ ) — штрафы за ошибки классификации сообщений соответствующих классов.

В третьей строке содержится целое положительное число  $\alpha$  ( $1 \leq \alpha \leq 10$ ) — интенсивность аддитивного сглаживания.

Следующая строка содержит целое положительное число  $N$  ( $1 \leq N \leq 200$ ) — число сообщений в обучающей выборке.

Следующие  $N$  строк содержат описания соответствующих сообщений из обучающей выборки. Каждое сообщение в ней начинается с целого положительного числа  $C_i$  ( $1 \leq C_i \leq K$ ) — класса к которому относится  $i$ -е сообщение. Далее следует целое положительное число  $L_i$  ( $1 \leq L_i \leq 10^4$ ) — число слов в  $i$ -м сообщении. Затем следует содержание сообщения —  $L_i$  слов состоящих из маленьких латинских букв.

Далее в отдельной строке содержится целое положительное число  $M$  ( $1 \leq M \leq 200$ ) — число сообщений в проверочной выборке.

Следующие  $M$  строк содержат описания соответствующих сообщений из проверочной выборки. Каждое сообщение в ней начинается с целого положительного числа  $L_j$  ( $1 \leq L_j \leq 10^4$ ) — число слов в  $j$ -м сообщении. Затем следует содержание сообщения —  $L_j$  слов состоящих из маленьких латинских букв.

Гарантируется, что сумма длин всех сообщений в обучающей и проверочной выборках меньше чем  $2 \cdot 10^6$ .

## Формат выходных данных

Выведите  $M$  строк — результаты мягкой классификации оптимального наивного байесовского классификатора соответствующих сообщений из проверочной выборки.

Каждый  $j$ -й результат мягкой классификации должен содержать  $K$  чисел  $p_C$  — вероятности того, что  $j$ -е сообщение относится к классу  $C$ .

## Пример

стандартный ввод	стандартный вывод
3	0.4869739479 0.1710086840 0.3420173681
1 1 1	0.1741935484 0.7340501792 0.0917562724
1	0.4869739479 0.1710086840 0.3420173681
4	0.4869739479 0.1710086840 0.3420173681
1 2 ant emu	0.4869739479 0.3420173681 0.1710086840
2 3 dog fish dog	
3 3 bird emu ant	
1 3 ant dog bird	
5	
2 emu emu	
5 emu dog fish dog fish	
5 fish emu ant cat cat	
2 emu cat	
1 cat	

## Замечание

В примере условные вероятности выглядят следующим образом:

$p(w_x c_y)$	ant	bird	dog	emu	fish
$c_1$	3/4	1/2	1/2	1/2	1/4
$c_2$	1/3	1/3	2/3	1/3	2/3
$c_3$	2/3	2/3	1/3	2/3	1/3

Слово cat не рассматривается, так как оно ни разу не встретилось в обучающей выборке.

Для первого запроса  $p(c_1|M) \cdot p(M) = \frac{2}{4} \cdot \left(1 - \frac{3}{4}\right) \cdot \left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) \cdot \left(1 - \frac{1}{4}\right)$  и  $p(c_1|M) = \frac{3/256}{3/256+1/243+2/243}$