

1. Дан набор значений 2,4,10,12,3,20,30,11,25. Предположим количество кластеров $k = 3$, и выбраны начальные средние значения $m_1 = 2$, $m_2 = 4$, $m_3 = 6$. Покажите, какие кластеры будут после первой итерации алгоритма k -средних, и рассчитайте новые значения центров кластеров для следующей итерации.

D	2	4	10	12	3	20	30	11	25	M
X — m_1	0	2	8	10	1	18	28	9	23	2
X — m_2	2	0	6	8	1	16	26	7	21	4
X — m_3	4	2	4	6	3	14	24	5	19	6
C1	2				3					2,5
C2		4								4
C3			10	12		20	30	11	25	18 Новые

2. Дан набор точек x и вероятности из принадлежности к кластерам C1 и C2.

x	$P(C_1 x)$	$P(C_2 x)$
2	0.9	0.1
3	0.8	0.1
7	0.3	0.7
9	0.1	0.9
2	0.9	0.1
1	0.8	0.2

Выполните следующие задание:

1. Найдите оценку максимального правдоподобия для средних μ_1 и μ_2
2. Предположим, что $\mu_1 = 2$, $\mu_2 = 7$ и $\sigma_1 = \sigma_2 = 1$. Найдите вероятности принадлежности точки $x = 5$ к кластерам C1 и C2. Априорные вероятности каждого кластера $P(C_1) = P(C_2) = 0.5$ и $P(x = 5) = 0.029$

Оценка макс. Правдоподобия:

$$\mu_i = (\sum n_{ij} w_{ij} x_j) / (\sum n_{ij} w_{ij})$$

n	1	2
$\sum x \cdot w$	9,8	13,9
$\sum w$	3,8	2,1
μ	2,57894736842105	6,61904761904762

Вероятность принадлежности:

$$f(x_1 | m_1, o_1) = 0,029$$

$$f(x_2 | m_2, o_2) = 0,054$$

$$P(C_1 | x) \approx 0.0758$$

$$P(C_2 | x) \approx 0.9241$$