

Задача 2

а) Рассчитаем корр. корр. корр.

Артём Александр
3 группа
3 вариант
2 стр.

y	x ₁	x ₂	y ²	x ₁ ²	x ₂ ²
20	43	7,5	400	1849	56,25
12	32,1	6,4	144	1030,41	40,96
15,6	35,1	7	243,36	1232,01	49
12,5	32	6,2	156,25	1024	38,44
13,2	33	6	174,24	1089	36
14,6	32,5	5,8	213,16	1056,25	33,64

y x ₁	y x ₂	x ₁ x ₂
860	150	322,5
385,2	76,8	205,44
547,86	109,2	245,7
400	77,5	198,4
435,6	79,2	198
474,5	84,68	188,5

Пусть ξ и η_1, η_2 соответствуют y, x_1 и x_2 для каждого из объектов ξ^2, η_1^2 и η_2^2 , а также $\xi\eta_1, \xi\eta_2$ и $\eta_1\eta_2$. Следующим шагом посчитаем моменты $E\xi^2 = \frac{1331,01}{6}$

$$E\xi = \frac{87,9}{6} = 14,65$$

$$E\eta_1 = \frac{207,7}{6} = 34,616$$

$$E\eta_2 = \frac{38,9}{6} = 6,483$$

$$E\eta_1^2 = \frac{254,29}{6} = 42,38$$

$$E\eta_2^2 = \frac{254,29}{6} = 42,38$$

$$E\xi\eta_1 = \frac{577,38}{6} = 96,23$$

$$E\xi\eta_2 = \frac{577,38}{6} = 96,23$$

$$E\eta_1\eta_2 = \frac{1358,54}{6} = 226,42$$

$$S\xi = \sqrt{E\xi^2 - (E\xi)^2} = 2,686$$

$$S\eta_1 = \sqrt{E\eta_1^2 - (E\eta_1)^2} = 3,9$$

$$S\eta_2 = \sqrt{E\eta_2^2 - (E\eta_2)^2} = 0,59$$

$$r_{\xi\eta_1} = \frac{E\xi\eta_1 - E\xi \cdot E\eta_1}{S\xi \cdot S\eta_1} = \frac{517,143 - 14,65 \cdot 34,616}{2,686 \cdot 3,9} = 0,96$$

$$r_{\xi\eta_2} = \frac{E\xi\eta_2 - E\xi \cdot E\eta_2}{S\xi \cdot S\eta_2} = \frac{96,23 - 14,65 \cdot 6,483}{2,686 \cdot 0,59} = 0,79$$

$$r_{\eta_1\eta_2} = \frac{E\eta_1\eta_2 - E\eta_1 \cdot E\eta_2}{S\eta_1 \cdot S\eta_2} = \frac{226,42 - 34,616 \cdot 6,483}{0,59 \cdot 3,9} = 0,87$$

$$r_{\xi\eta_1} = \frac{517,143 - 14,65 \cdot 34,616}{2,686 \cdot 3,9} = 0,96$$

$$r_{\xi\eta_2} = \frac{96,23 - 14,65 \cdot 6,483}{2,686 \cdot 0,59} = 0,79$$

$$r_{\eta_1\eta_2} = \frac{226,42 - 34,616 \cdot 6,483}{0,59 \cdot 3,9} = 0,87$$

$$r_{\xi\eta_1} = \frac{517,143 - 14,65 \cdot 34,616}{2,686 \cdot 3,9} = 0,96$$

$$r_{\xi\eta_2} = \frac{96,23 - 14,65 \cdot 6,483}{2,686 \cdot 0,59} = 0,79$$

$$r_{\eta_1\eta_2} = \frac{226,42 - 34,616 \cdot 6,483}{0,59 \cdot 3,9} = 0,87$$

т.е. все 3 связи положительные, особенно $r_{\xi\eta_1}$ и все 3 связи положительные.

б) Построим матрицу парных коэф. корреляции

Артемов
Александр
3 группа
3 вариант
3 стр.

	y	x ₁	x ₂
y	1	0,96	0,79
x ₁	0,96	1	0,87
x ₂	0,79	0,87	1

в) Теперь рассчитаем множественный коэф. корреляции:

$$R_{y/x_1x_2} = \sqrt{\frac{R_{yx_1}^2 + R_{yx_2}^2 - 2R_{yx_1}R_{yx_2}R_{x_1x_2}}{1 - R_{x_1x_2}^2}} = \sqrt{\frac{0,96^2 + 0,79^2 - 2 \cdot 0,96 \cdot 0,79 \cdot 0,87}{1 - 0,87^2}}$$

$$= 0,964$$

г) Проверим значимость парных коэф. корреляции $\alpha = 0,05$
 Для этого выберем статистику $T = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} = \frac{2r}{\sqrt{1-r^2}}$
 где $n = 6$. Тогда $t_{кр} = t(0,05, 4)$
 т.е. $t_{кр} = 2,776$. Статистики имеют значения:

1) $T_{xy} = 6,86 > t_{кр} \Rightarrow T_{xy}$ попадает в критическую область $\Rightarrow R_{yx_1}$ значим

2) $T_{yx_2} = 2,577 < t_{кр} \Rightarrow T_{yx_2} \in \text{Доверит. области}$ $\Rightarrow R_{yx_2}$ не значим. $H_0^{yx_2}$ - принимается

3) $T_{x_1x_2} = 3,53 > t_{кр} \Rightarrow T_{x_1x_2} \in KO \Rightarrow R_{x_1x_2}$ значим на $\alpha = 0,05$
 $H_0^{x_1x_2}$ отвергается

г) $H_0: R_{y/x_1x_2}$ не значим: выберем статистику $F = \frac{R_{y/x_1x_2}^2}{(1 - R_{y/x_1x_2}^2)/(n-3)}$

$$F_{кр} = F(0,05; 2; 3) = 9,55 \text{ и } F_{табл} = 19,72 > F_{факт} \Rightarrow$$

H_0 отвергается и R_{y/x_1x_2} значим.

$$е) z_r = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right) \Rightarrow z_{yx_1} = 1,95 \quad z_{yx_2} = 1,07 \quad z_{x_1x_2} = 1,33$$

$t_{0,95} = 1,96 \Rightarrow$ Доверительные интервалы:

$$R_{yx_1}: (1,95 - 1,96 \cdot \sqrt{\frac{1}{3}}; 1,95 + 1,96 \cdot \sqrt{\frac{1}{3}}) \text{ т.е. } (0,818; 3,08)$$

$$R_{yx_2}: (1,07 - 1,96 \cdot \sqrt{\frac{1}{3}}; 1,07 + 1,96 \cdot \sqrt{\frac{1}{3}}) \text{ т.е. } (-0,06; 2,2)$$

$$R_{x_1x_2}: (1,33 - 1,96 \cdot \sqrt{\frac{1}{3}}; 1,33 + 1,96 \cdot \sqrt{\frac{1}{3}}) \text{ т.е. } (0,198; 2,46)$$

Задание 3) Для каждого построим матрицу расстояний

Артемиев
Александр
3 группа
3 вариант
4 стр.

расстояние =
евклидово
расстояние

	1 Мед	2 Мог	3 Дж	4 Вен	5 Хлоэ	6 Итан	7 Крис
1	0	40,01	110,07	440,09	65,03	18,87	47,12
2	40,01	0	150,03	400,08	25,02	52,81	13,93
3	110,07	150,03	0	550,02	175,01	102	155,32
4	440,09	400,08	550,02	0	375,07	450,7	395,03
5	65,03	25,02	175,01	375,07	0	77,13	23,32
6	18,87	52,81	102	450,7	77,13	0	62,65
7	47,12	13,93	155,32	395,03	23,32	62,65	0

1ый
этап

Мин расстояние $R(2,7) = 13,93$
 Следующее по величине $R(1,6) = 18,87$
 2 и 7 кластер и 3 кластер новые расстояния по формуле
 $R(1,6) = \frac{1}{2}R_{1m} + \frac{1}{2}R_{6n} - \frac{1}{2}|R_{1n} - R_{6m}|$ 2ой этап

	1	2,7	3	4	5	6
1	0	40,01	110,07	440,09	65,03	18,87
2,7	40,01	0	150,03	385,03	25,02	52,81
3	110,07	150,03	0	550,02	175,01	102
4	440,09	385,03	550,02	0	375,07	450,7
5	65,03	23,32	175,01	375,07	0	77,15
6	18,87	52,81	102	450,7	77,15	0

мин расстояние
(1,6)
 Теперь объединим
 1 и 6
 Дальше я буду
 писать только
 как много расстояний
 матрицы = к
 она симметрична
 и $m_{ij} = m_{ji}$

3ый этап

	1,6	2,7	3	4	5
1,6	0				
2,7	40,01	0			
3	102	150,03	0		
4	440,09	385,03	550,02	0	
5	18,87	23,32	175,01	375,07	0

следующий шаг
 объединим
 1,6 и 5

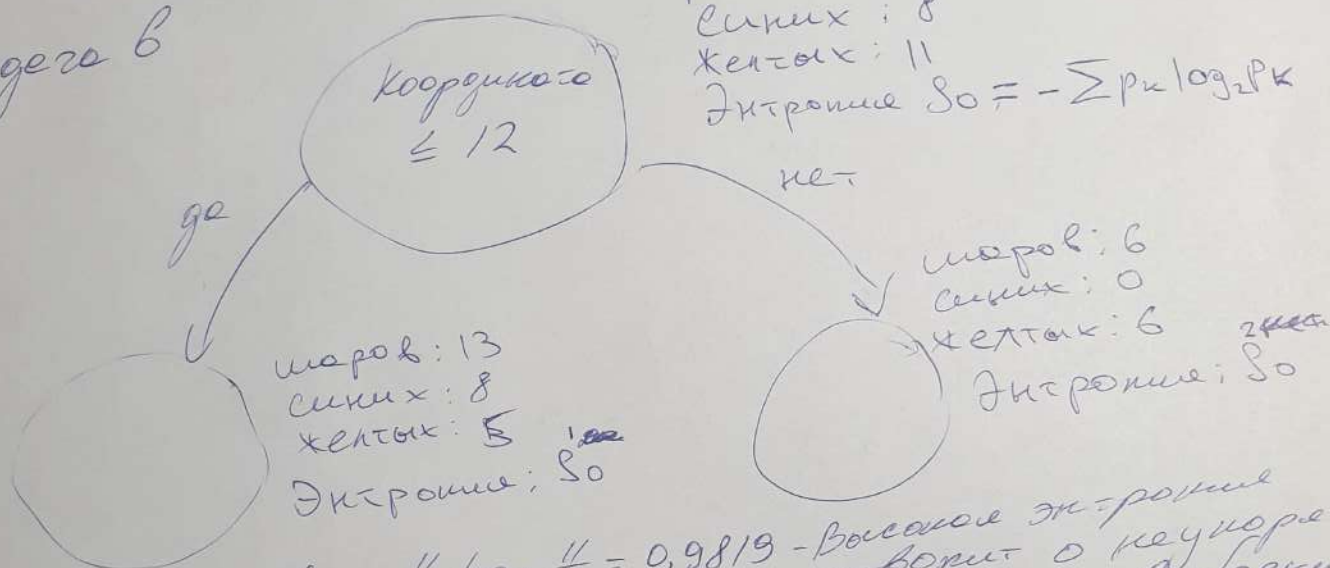
Задача 5. С одной стороны нам явно виден только один
 как это был столбец. Но относительно него ~~два~~ столбца
 не симметричны и это не норм. распределение. Я
 предположу, что здесь 2 смеси с μ 5 и 6
 столбцом

смесь	μ_i	σ_i	φ_i
1	7,13	7,93	0,857
2	4,80	0,37	0,43

Частоты я считаю
 приблизительно
 одинаковые, так
 как смесь вылив в
 моде и $\mu_1 > \mu_2$

σ_2 я взял просто как середину интервала

Задача 6



$$S_0 = -\frac{8}{19} \log_2 \frac{8}{19} - \frac{11}{19} \log_2 \frac{11}{19} = 0,9819$$

Высшая энтропия
 говорит о неупоре-
 дотченности выборки

$$S_0^2 = 0$$

т.к. синих шаров 0 и это говорит, что выборка
 с координатой > 12 полностью упорядочена

$$S_0' = -\frac{8}{13} \log_2 \frac{8}{13} - \frac{5}{13} \log_2 \frac{5}{13} = 0,9612$$

Энтропия после
 упорядочивания
 остатков не упорядочен-
 ных шаров, но все равно выборка остается не упорядочен-
 ной

$$I_6 = S_0 - \frac{N_1}{N} S_0' - \frac{N_2}{N} S_0^2 = S_0 - \frac{N_1}{N} S_0' = 0,9819 - \frac{13}{19} \cdot 0,9612$$

$= 0,3242$ - критерий информативности

Артёмов
 Александр
 Вруня
 Зверев

Задача 7 (какого посмотрим как
разбиваются данные при делении по
температуре в волюте °C

Артёмов
Александр
3 группа
3 вариант

1) $t \leq 20$ нет
 se
 idx комп
 1 70
 2 90
 3 80
 4 100
 5 75

2) $t \leq 21$ нет
 se
 idx комп
 1 70
 3 80
 2 90
 4 100
 5 75

3) $t \leq 23$ нет
 se
 idx комп
 1 70
 2 90
 3 80
 4 100
 5 75

4) $t \leq 25$ нет
 se
 idx комп
 1 70
 2 90
 3 80
 4 100
 5 75

Получаем значения в каждой группе:

$$P_1^+ = 0 + 122,9167 = 122,9167 \quad P_2^+ = 100 + 312,5 = 412,5$$

$$P_2^+ = 50 + 158,333 = 208,333 \quad P_3^+ = 166,667 + 0 = 166,667$$

Теперь разобьем по ширине полки

1) $l \leq 75$ нет
 se
 idx комп
 1 70
 2 90
 3 80
 4 100
 5 75

2) $l \leq 80$ нет
 se
 idx комп
 1 70
 3 80
 2 90
 4 100
 5 75

3) $l \leq 85$ нет
 se
 idx комп
 1 70
 2 90
 3 80
 4 100
 5 75

$$P_1^l = P_1^+ ; P_2^l = P_2^+ ; P_3^l = P_4^+ \Rightarrow \min D = P_1^l = P_4^l$$

Корень дерева будет температурой ≤ 20 и шириной полки ≤ 75

Задача 1) 1) Для начала посчитаем суммарную стоимость всех элементов для каждого местоположения.

Артемьев Александр
3 группа
Вариант 3
1 стр

id	Название	Стоимость
1	Ice castle	2400
2	Love story	1150
3	Dance Round in Round	3850
4	Fun Starts Here	1500
5	Newer Too Old	2300
6	Honey, I'm at Home's	2050

stem-and-leaf

stem	leaf
11	50
15	00
20	50
23	00
24	00
38	50

Анализируя график можно сделать вывод: данные распределены в диапазоне от 1150 до 3850 и данные распределены не равномерно, все значения сконцентрированы в диапазоне 1150-2400 и одно значение около 3850, которое значительно больше остальных. График построен взвешивая тысячи за stem и сотни за leaf. Построив график мы увидели форму и размах выборки.

2) Для оценки правого 3σ можно считать среднее: $\bar{x} = (300 + 100 + 750 + 200 + 600 + 500) \cdot \frac{1}{6} = 408,333$
 Теперь определим $E\xi^2$ (ξ - цена для сайта) $E\xi^2 = (300^2 + 100^2 + 750^2 + 200^2 + 600^2 + 500^2) \cdot \frac{1}{6} = \frac{100^2}{6} (9 + 1 + 56,25 + 4 + 36 + 25) = \frac{100^2 \cdot 131,25}{6} = \frac{1312500}{6} = 218750$ $(E\xi)^2 = 166733,33$

Вывод: $D_{\text{смет}} = \frac{D_{\text{смет}} \cdot n}{n-1}$, где $D_{\text{смет}} = E\xi^2 - (E\xi)^2 \Rightarrow$
 $\Rightarrow D_{\text{смет}} = 52016,611$; $n = 6 \Rightarrow D_{\text{смет}} = 62419,93$
 $\Rightarrow \sigma = \sqrt{D_{\text{смет}}} = 249,84$. Правило 3σ устанавливает

интервал: $(n - 3\sigma; n + 3\sigma) : (-341,187; 1157,853)$

Поскольку все элементы выборки в интервале \Rightarrow выбросов нет