

Корреляционный анализ

Анализ некоторых показателей рынка жилой недвижимости

Вариант 9: Взяты данные с 81 по 90.

В предыдущем задании я использовала все данные кроме 81-90, поэтому для этого д/з я сделала перерасчет корреляционных матриц.

Таблица 1. Данные

№ п/п	S общая	Этаж квартиры	Цена за м ²
	x15	x18	x19
81	92,2	1	115 174
82	105,3	11	132 953
83	37,7	15	68 000
84	40,6	14	68 000
85	53,7	7	76 432
86	78,6	6	68 000
87	78,6	14	76 509
88	56,6	1	85 000
89	58,9	1	85 000
90	59,0	1	85 000

Интерпретация показателей:

- x_{15} – общая площадь квартиры, м²;
- x_{18} – этаж, на котором располагается квартира;
- x_{19} – цена за квадратный метр, руб.

1. оценить параметры:

- вектор средних;

Таблица 2. Параметры анализируемых показателей

	x15	x18	x19
Сумма	661,21	71,00	860068,00

Среднее	66,12	7,1	86006,80
----------------	-------	-----	----------

Вектор средних: $X_{\text{ср.}} = (66,12 \ 7,1 \ 86006,80)^T$

Выводы:

По результатам выборки средний размер общей площади квартиры составляет 66,12 м², средний номер этажа квартиры – 7,1, средняя цена за квадратный метр – 86006,80 руб.

– вектор среднеквадратических отклонений;

Таблица 3. Параметры анализируемых показателей

	x15	x18	x19
S	20,82	5,68	20523,97

Вектор среднеквадратических отклонений: $S = (20,82 \ 5,86 \ 20523,97)^T$

– корреляционную матрицу парных коэффициентов;

Таблица 4. Параметры анализируемых показателей

	x15x18	x18x19	x15x19
Сумма	4506,81	5818807,00	6023954950
Среднее	450,68	581880,70	6023954,95
r	-0,16	-0,25	0,79

Корреляционная матрица парных коэффициентов: $R =$

$$\begin{pmatrix} 1 & -0,16 & 0,79 \\ -0,16 & 1 & -0,25 \\ 0,79 & -0,25 & 1 \end{pmatrix}$$

Выводы:

Между всеми показателями выборки наблюдается линейная взаимосвязь. Между

общей площадью квартиры (x_{15}) и ценой за квадратный метр (x_{19}) наблюдается сильная положительная линейная взаимосвязь, близкая к функциональной (чем выше один показатель, тем выше другой).

Между этажом квартиры (x_{18}) и ценой за квадратный метр (x_{19}) наблюдается отрицательная взаимосвязь (чем ниже один показатель, тем выше другой).

Между этажом квартиры (x_{18}) и общей площадью (x_{15}) наблюдается низкая отрицательная (они практически независимы друг от друга).

2. оценить корреляционную матрицу частных коэффициентов;

Таблица 5. Параметры анализируемых показателей

R1515	0,94
R1518	-0,035
R1519	-0,75
R1818	0,38
R1819	0,121
R1919	0,974

Таблица 6. Параметры анализируемых показателей

	x15x18	x18x19	x15x19
r_{частн}	0,06	0,20	0,78

Корреляционная матрица частных коэффициентов: $R_{\text{частн.}} =$

1	0,06	0,78
0,06	1	0,20
0,78	0,20	1

Таблица 7. Выводы о корреляции показателей по частным коэффициентам

Коэффициенты	Парные	Частные	Выводы
$r_{1518/19} =$	-0,16	0,06	Показатель цены за квадратный метр станции метро (x_{19}) ослабляет взаимосвязь площади квартиры и этажа. «Чистая» взаимосвязь между показателями изначально была низка, однако при фиксированном влиянии цены за квадратный метр (x_{19}) связь практически исчезает
$r_{1519/18} =$	-0,25	0,20	Номер этажа (x_{18}) изменяет тип взаимосвязи цены за квадратный метр квартиры и ее общей площади. Взаимосвязь между общей площадью квартиры (x_{15}) и цены за квадратный метр (x_{19}) меняется с низкой отрицательной до низкой положительной.
$r_{1819/15} =$	0,79	0,78	Общая площадь (x_{15}) никак не влияет на взаимосвязь номера этажа (x_{18}) и цены за квадратный метр (x_{19}). «Чистая» линейная взаимосвязь между номером этажа квартиры (x_{18}) и ценой за квадратный метр (x_{19}) остается практически такой же..

3. при доверительной вероятности $\gamma=0,95$ построить доверительные интервалы для:
– парных коэффициентов;

$$Z' - t_{\gamma} \sqrt{\frac{1}{n-l-3}} \leq Z \leq Z' + t_{\gamma} \sqrt{\frac{1}{n-l-3}}$$

Корреляционная матрица парных коэффициентов: R =

$$\begin{matrix} 1 & -0,16 & 0,79 \\ -0,16 & 1 & -0,25 \\ 0,79 & -0,25 & 1 \end{matrix}$$

По таблице интегральной функции Лапласа из условия $\gamma=0,95$ следует, что $t_{\gamma}=1,96$.
По таблице Z-преобразования Фишера при условии $r_{1518} = -0,16$ следует, что $Z_{r_{1518}} = -0,1614$.

$$-0,1614 - 1,96 * (1/(10-0-3))^{(1/2)} \leq Z \leq -0,1614 + 1,96 * (1/(10-0-3))^{(1/2)}$$

$$-0,9022 \leq Z \leq 0,5794$$

По таблице Z-преобразования Фишера осуществляем переход к значениям парного коэффициента.

$$-0,72 \leq r_{1518} \leq 0,52$$

Выводы:

С вероятностью 0,95 можно утверждать, что парный коэффициент корреляции между общей площадью квартиры (x_{15}) и ее этажом (x_{18}) будет находиться в интервале от -0,72 до 0,52. Здесь произошел переход через 0 отметку, что свидетельствует о незначимости коэффициента.

Аналогично ищем интервалы для r_{1519} и r_{1819} .

По таблице Z-преобразования Фишера при условии $r_{1519} = 0,79$ следует, что $Z_{r_{1519}} = 1,0714$.

$$1,0714 - 1,96 * (1/(10-0-3))^{(1/2)} \leq Z \leq 1,0714 + 1,96 * (1/(10-0-3))^{(1/2)}$$

$$0,3305 \leq Z \leq 1,8122$$

По таблице Z-преобразования Фишера осуществляем переход к значениям парного коэффициента.

$$0,32 \leq r_{1519} \leq 0,95$$

Выводы:

С вероятностью 0,95 можно утверждать, что парный коэффициент корреляции между общей площадью квартиры (x_{15}) и ценой за квадратный метр (x_{19}) будет находиться в интервале от 0,32 до 0,95.

По таблице Z-преобразования Фишера при условии $r_{1819} = -0,25$ следует, что $Z_{r_{1819}} = -0,2554$.

$$0,2554 - 1,96 * (1/(10-0-3))^{(1/2)} \leq Z \leq 0,2554 + 1,96 * (1/(10-0-3))^{(1/2)}$$

$$-0,9962 \leq Z \leq 0,4854$$

По таблице Z-преобразования Фишера осуществляем переход к значениям парного коэффициента.

$$-0,76 \leq r_{1819} \leq 0,45$$

Выводы:

С вероятностью 0,95 можно утверждать, что парный коэффициент корреляции между этажом квартиры (x_{18}) и ценой за квадратный метр (x_{19}) будет находиться в интервале от (-0,76) до 0,45. Здесь произошел переход через 0 отметку, что свидетельствует о незначимости коэффициента.

— частных коэффициентов;

Аналогично ищем интервалы для частных коэффициентов.

Корреляционная матрица частных коэффициентов: $R_{\text{частн.}} =$

$$\begin{matrix} 1 & 0,06 & 0,78 \\ 0,06 & 1 & 0,20 \\ 0,78 & 0,20 & 1 \end{matrix}$$

По таблице Z-преобразования Фишера при условии $r_{1518/19} = 0,06$ следует, что $Z_{r_{1518/19}} = 0,0601$.

$$0,0601 - 1,96 * (1/(10-1-3))^{(1/2)} \leq Z \leq 0,0601 + 1,96 * (1/(10-1-3))^{(1/2)} \\ -0,7400 \leq Z \leq 0,8602$$

По таблице Z-преобразования Фишера осуществляем переход к значениям частного коэффициента.

$$-0,63 \leq r_{23} \leq 0,43$$

Выводы:

С вероятностью 0,95 можно утверждать, что парный коэффициент корреляции между общей площадью квартиры (x_{15}) и ее этажом (x_{18}) (при фиксированном влиянии цены за квадратный метр (x_{19})) будет находиться в интервале от (-0,63) до 0,43. Здесь произошел переход через 0 отметку, что свидетельствует о незначимости коэффициента.

По таблице Z-преобразования Фишера при условии $r_{1519/18} = 0,78$ следует, что $Z_{r_{1519/18}} = 1,0454$.

$$1,0454 - 1,96 * (1/(10-1-3))^{(1/2)} \leq Z \leq 1,0454 + 1,96 * (1/(10-1-3))^{(1/2)} \\ 0,2452 \leq Z \leq 1,8455$$

По таблице Z-преобразования Фишера осуществляем переход к значениям частного коэффициента.

$$0,24 \leq r_{23} \leq 0,95$$

Выводы:

С вероятностью 0,95 можно утверждать, что парный коэффициент корреляции между общей площадью квартиры (x_{15}) и ценой за квадратный метр (x_{19}) (при фиксированном влиянии номера этажа (x_{18})) будет находиться в интервале от 0,24 до 0,95.

По таблице Z-преобразования Фишера при условии $r_{1819/15} = 0,20$ следует, что $Z_{r_{1819/15}} = 0,2027$.

$$0,2027 - 1,96 * (1/(10-1-3))^{(1/2)} \leq Z \leq 0,2027 + 1,96 * (1/(10-1-3))^{(1/2)} \\ -0,5974 \leq Z \leq 1,0028$$

По таблице Z-преобразования Фишера осуществляем переход к значениям частного коэффициента.

$$-0,53 \leq r_{23} \leq 0,76$$

Выводы:

С вероятностью 0,95 можно утверждать, что парный коэффициент корреляции между этажом (x_{18}) и ценой за квадратный метр (x_{19}) (при фиксированном влиянии общей площади квартиры (x_{15})) будет находится в интервале от (-0,53) до 0,76. Здесь произошел переход через 0 отметку, что свидетельствует о незначимости коэффициента.

4. найти точечную оценку множественного коэффициента корреляции для x_{15} :

$$r_{1/2,3,4,\dots,k} = \sqrt{1 - \frac{|R|}{R_{11}}}$$

$$r_{15/18,19} = (1 - 0,3532/0,94)^{(1/2)} = 0,7898$$

Выводы:

В качестве результативного признака был выбран x_{15} – общая площадь квартиры. Совокупное влияние этажа квартиры (x_{18}) и цены за квадратный метр (x_{19}) на общую площадь квартиры (x_{15}) положительно и стремится к линейной.

$$(r_{15/18,19})^2 = 0,6239$$

Выводы:

Вариация (дисперсия) общей площади квартиры (x_{15}) объяснена вариацией (дисперсией) этажа квартиры (x_{18}) и ценой за квадратный метр (x_{19}) на 62,39%.

6. проверить значимость множественного коэффициента корреляции при уровне значимости $\alpha=0,05$

$$F_{\text{набл}} = \frac{\frac{1}{l} r_{1/2, \dots, k}^2}{\frac{1}{n-l-1} (1 - r_{1/2, \dots, k}^2)}$$

$$F_{\text{набл}} = (1/2 * 0,6239) / (1 / (10 - 2 - 1) * (1 - 0,6239)) = 5,8061$$

По таблице Фишера $F_{\text{кр}}(0,05; l=2; n-l-1=7) = 4,74$

$F_{\text{набл}} > F_{\text{кр}}$, следовательно коэффициент значим.