

Контрольная работа по эконометрике

ЧАСТЬ I

Используя фактические значения независимого фактора (x) и результирующей переменной (y), провести эконометрическое исследование зависимости y от x .

1. Построить поле корреляции переменных y и x .
2. Выбрать и обосновать спецификацию уравнения регрессии (использовать уравнение вида $y = a_0 + a_1x + e$).
3. Рассчитать коэффициенты $\hat{a}_0$ и $\hat{a}_1$ выбранного уравнения. Для этого составить систему нормальных уравнений и найти ее решения методом определителей.
4. Построить уравнение прогноза и провести содержательный анализ его коэффициентов.
5. Рассчитать коэффициент парной линейной корреляции и сделать выводы о тесноте связи между переменными построенного уравнения.
6. Провести оценку значимости коэффициентов регрессии и корреляции с помощью t -критерия Стьюдента (при уровне значимости $\alpha=0,05$).
7. Провести оценку качества построенного уравнения регрессии с помощью F -критерия Фишера (при уровне значимости $\alpha=0,05$).
8. Вычислить коэффициент детерминации и проанализировать его значение.
9. Оценить построенное уравнение через среднюю ошибку аппроксимации.
10. Используя уравнение прогноза, построенное в п. 4, выполнить точечный и интервальный прогноз значений результирующей переменной y по значениям объясняющей переменной x , указанным в условии задачи.
11. Сделать все необходимые выводы по результатам выполнения каждого из пунктов задания.

Задача 14

Имеются данные (в у.е.) о среднемесячной заработной плате (y) и прожиточном минимуме (x):

Прожиточный минимум	100	105	110	98	110	131	90	98	111	140
Заработная плата	180	205	178	200	205	230	185	207	185	220

Определить размер среднемесячной заработной платы при размере прожиточного минимума, равном 150 у.е.

Решение.

В парной регрессии выбор вида математической функции $\hat{y}_x = f(x)$ может быть осуществлен тремя методами:

- 1) графическим;
- 2) аналитическим, т.е. исходя из теории изучаемой взаимосвязи;
- 3) экспериментальным.

При изучении зависимости между двумя признаками графический метод подбора вида уравнения регрессии достаточно нагляден. Он основан на поле корреляции.

Поле корреляции строится как график зависимости между двумя параметрами. Оно строится в таком порядке: по горизонтальной линии откладываются измерения величин измерения величин одной переменной, а по вертикальной оси - другой переменной.

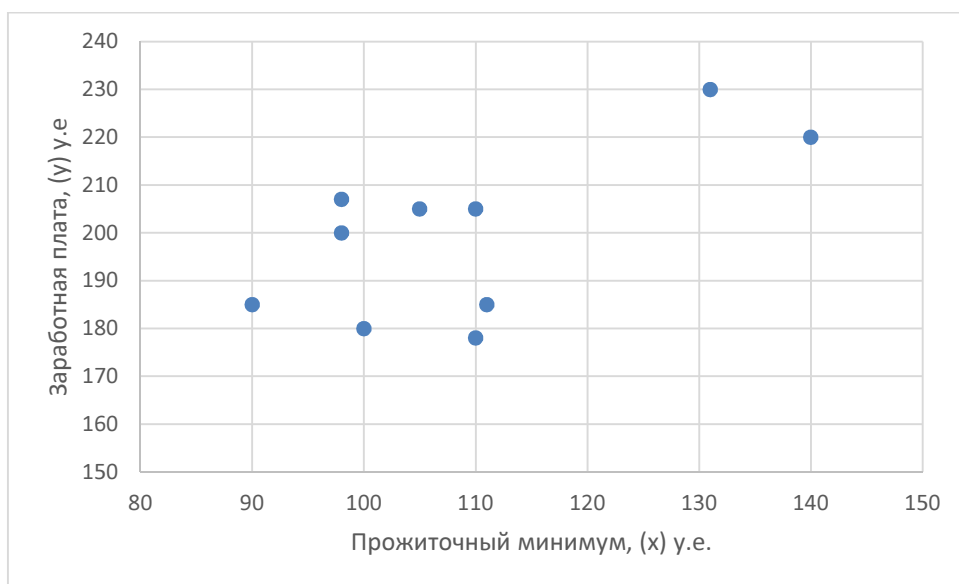


Рис 1. Поле корреляции

Анализ рис. 1 позволяет сделать вывод о возможной линейной статистической связи между исследуемыми переменными. При этом связь имеет положительную тенденцию, т.е. с ростом переменной X наблюдается увеличение Y .

Линейная регрессия описывается уравнением $y = a_0 + a_1x + e$

Расчёт неизвестных параметров уравнения выполним методом наименьших квадратов (МНК), построив систему нормальных уравнений и решая её, относительно неизвестных a_0 и a_1 .

$$\begin{cases} na_0 + a_1 \sum x = \sum y \\ a_0 \sum x + a_1 \sum x^2 = \sum xy \end{cases}$$

Для расчёта используем значения определителей второго порядка Δ , Δa_0 и Δa_1 .

Расчёт определителя системы выполним по формуле:

$$\Delta = n * \sum (X^2) - \sum X * \sum X$$

Расчёт определителя свободного члена уравнения выполним по формуле:

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

$$\Delta a_0 = \sum Y * \sum (X^2) - \sum (Y * X) * \sum X$$

Расчёт определителя коэффициента регрессии выполним по формуле:

$$\Delta a_1 = n * \sum (Y * X) - \sum Y * \sum X$$

Расчётные процедуры представим в разработочной таблице, в которую, кроме значений Y и X, войдут X^2 , $X*Y$, а также их итоговые значения, средние, сигмы и дисперсии для Y и X. См. табл.2

Таблица №2

№	X	$Y_{факт.}$	X^2	$Y * X$	$\hat{Y}_{расч.}$	$\left \frac{Y_{факт.} - \hat{Y}}{\bar{Y}} \right $	$\varepsilon', \%$	$(y - \hat{y}_x)^2$
1	100	180	10000	18000	192,568 6	0,0630	6,30	157,9692
2	105	205	11025	21525	196,295 1	0,0436	4,36	75,7744
3	110	178	12100	19580	200,021 7	0,1104	11,0 4	484,9561
4	98	200	9604	19600	191,077 9	0,0447	4,47	79,6030
5	110	205	12100	22550	200,021 7	0,0250	2,50	24,7833
6	131	230	17161	30130	215,673 3	0,0718	7,18	205,2538
7	90	185	8100	16650	185,115 4	0,0006	0,06	0,0133
8	98	207	9604	20286	191,077	0,0798	7,98	253,5117

					9			
9	111	185	12321	20535	200,767 0	0,0790	7,90	248,5994
10	140	220	19600	30800	222,381 1	0,0119	1,19	5,6699
Итого	1093	1995	121615	21965 6	1995	0,5299	52,9 9	1536,134 0
Средняя	109,3	199,5	—	—	—	—	5,30	—
σ	14,6632	16,524 2	—	—	—	—	—	—
D	215,01	273,05	—	—	—	—	—	—
Δ=	21501	—	—	—	—	—	—	—
Δa ₀ =	253791 7	a ₀ =	118,037 2	—	—	—	—	—
Δa ₁ =	16025	a ₁ =	0,7453	—	—	—	—	—

$$\Delta = n * \sum (X^2) - \sum X * \sum X = 10 * 121615 - 1093 * 1995 = 21501$$

$$\Delta a_0 = \sum Y * \sum (X^2) - \sum (Y * X) * \sum X = 1995 * 121615 - 219656 * 1093 = 2537917$$

$$\Delta a_1 = n * \sum (Y * X) - \sum Y * \sum X = 10 * 2196556 - 1093 * 1995 = 16025$$

Расчёт параметров уравнения регрессии даёт следующие результаты:

$$a_0 = \frac{\Delta a_0}{\Delta} = \frac{2537917}{21501} = 118,0372; \quad a_1 = \frac{\Delta a_1}{\Delta} = \frac{16025}{21501} = 0,7453.$$

В итоге, получаем теоретическое уравнение регрессии следующего вида:

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

$$\hat{y}_x = 118,0372 + 0,7453x$$

В уравнении коэффициент регрессии $a_1 = 0,7453$ означает, что при увеличении прожиточного минимума на 1 у.е. (от своей средней) заработная плата увеличивается на 0,7453 у.е. (от своей средней).

Свободный член уравнения $a_0 = 118,0372$ оценивает влияние прочих факторов, оказывающих воздействие на заработную плату

Определим теоретические значения результата $Y_{\text{расч.}}$. Для этого в полученное уравнение $\hat{y}_x = 118,0372 + 0,7453x$ последовательно подставим фактические значения фактора X и выполним расчёт.

Для оценки тесноты связи рассчитаем линейный коэффициент парной корреляции:

$$r_{YX} = a_1 * \frac{\sigma_X}{\sigma_Y} = 0,7453 * \frac{14,6632}{16,5242} = 0,6614$$

Коэффициент корреляции, равный 0,6614, показывает, что выявлена тесная заметная зависимость между прожиточным минимумом и заработной платой

Оценку статистической значимости параметров регрессии проведем с помощью t -статистики Стьюдента.

Выдвигаем гипотезу H_0 о статистически незначимом отличии показателей от нуля: $a_0 = a_1 = r_{xy} = 0$.

$t_{\text{табл}}$ для числа степеней свободы $df = m - 2 = 10 - 2 = 8$ и $\alpha = 0,05$ составит 2,31.

Определим случайные ошибки m_{a0} , m_{a1} , $m_{r_{xy}}$ (остаточная дисперсия на одну степень

свободы $S_{\text{ост}}^2 = \frac{\sum (y - \hat{y}_x)^2}{n - 2} = \frac{1536,1340}{8} = 192,0168$)

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=ekr

$$m_{a0} = \sqrt{\frac{S_{ocm}^2 \sum_{i=1}^n x_i^2}{n^2 \sigma_x^2}} = \sqrt{\frac{192,0168 * 121615}{10^2 * 14,6632}} = 32,9559;$$

$$m_{a1} = \sqrt{\frac{S_{ocm}^2}{n \sigma_x^2}} = \sqrt{\frac{192,0168}{10 * 14,6632}} = 0,2988$$

$$m_r = \sqrt{\frac{(1 - r_{yx}^2)}{(n - 2)}} = \sqrt{\frac{1 - (0,6614)^2}{8}} = 0,2652.$$

Тогда

$$t_{a0} = \frac{a_0}{m_{a0}} = \frac{178,0372}{32,9559} = 3,5817$$

$$t_{a1} = \frac{a_1}{m_{a1}} = \frac{0,7453}{0,2988} = 2,494$$

$$t_r = \frac{r_{xy}}{m_r} = \frac{0,6614}{0,2652} = 2,494$$

Фактические значения t – статистики превосходят табличные показания:

$$t_{(r=0)} = 2,494 > t_{табл} = 2,31, \quad t_{(a1=0)} = 2,494 > t_{табл} = 2,31,$$

$$t_{(a0=0)} = 3,5817 > t_{табл} = 2,31$$

поэтому параметры a_0 , a_1 и r_{xy} не случайно отличаются от нуля, а статистически значимы.

Для оценки статистической надёжности выявленной зависимости рассчитаем фактическое значение F-критерия Фишера – $F_{факт.}$ и сравним его с табличным значением – $F_{табл.}$. По результатам сравнения примем решения по нулевой гипотезе $H_0 : a_0 = a_1 = r_{yx} = 0$, то есть,

©**МатБюро** – Консультации по математике, экономике, праву, другим предметам

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

либо примем, либо отклоним её с вероятностью допустить ошибку, которая не превысит 5% (или с уровнем значимости $\alpha=0,05$).

В нашем случае,

$$F_{\text{факт.}} = \frac{r_{YX}^2}{1 - r_{YX}^2} : \frac{k}{n - k - 1} = \frac{0.6614^2}{1 - 0.6614^2} : \frac{1}{10 - 1 - 1} = 6.22$$

где k - число факторов в уравнении; n - число изучаемых объектов

Для обоснованного вывода сравним полученный результат с табличным значением критерия: $F_{\text{табл.}} = 5.32$ при степенях свободы d.f.1= $k-1=1$ и d.f.2= $n-k=10-2=8$ и уровне значимости $\alpha=0,05$.

В силу того, что $F_{\text{факт.}} = 6.22 > F_{\text{табл.}} = 5.32$, нулевую гипотезу о статистической незначимости выявленной зависимости потребления от дохода можно отклонить с фактической вероятностью допустить ошибку значительно меньшей, чем традиционные 5%.

Коэффициент детерминации

$$R^2 = r_{xy}^2$$

$$R^2 = (0.6614)^2 = 0.4374$$

Коэффициент детерминации, равный 0.4374, устанавливает, что вариация заработной платы на 43,74% из 100% предопределена вариацией прожиточного минимума, роль прочих факторов, влияющих на заработную плату, определяется в 56,26%, что является заметной величиной.

Оценку качества модели дадим с помощью скорректированной средней ошибки аппроксимации:

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

$$\overline{\varepsilon'} = \frac{1}{n} * \sum \left| \frac{Y_{\text{факт}} - \hat{Y}}{\bar{Y}} \right| * 100 \% .$$

Вычисления приведены в графах 6,7 таблицы №2.

$$\overline{\varepsilon'} = 5.3 \%$$

В нашем случае скорректированная ошибка аппроксимации составляет 5,3%, что не превышает 10-12%. Она указывает на высокое качество построенной линейной модели и позволяет её использование для выполнения точных прогнозных расчётов даже при условии сравнительно небольшого изменения фактора X (относительно его среднего значения $\bar{X}$).

Найдем прогнозное значение результата $\hat{y}$, если прогнозное значение фактора $\hat{x}$ составит 150 у.е. Прогнозное значение результата сформируется на уровне $\hat{y} = 118,0372 + 0,7453 \cdot 150 = 229,8343$. То есть если прожиточный минимум увеличится до 150 у.е., то заработная плата составит 229,8343 у.е.

Ошибка прогноза составит:

$$m_{\hat{y}_p} = \sqrt{S_{\text{ост}}^2 \left(1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{n \cdot \sigma_x^2} \right)} = \sqrt{192,0168 \left(1 + \frac{1}{10} + \frac{(109,3 - 150)^2}{10 \cdot 14,6632} \right)} = 22,9186$$

Предельная ошибка прогноза, которая в 95% случаев не будет превышена, составит:

$$\Delta_{\hat{y}_p} = t\mu_{\hat{y}_p} = 2.31 \cdot 22,9186 = 52,9420$$

Доверительный интервал прогноза:

$$\gamma_{\hat{y}_p} = 229,8343 \pm 52,9420 \text{ и } 176,8923 \leq y_0^* \leq 282,77663$$

Контрольная работа по эконометрике выполнена в www.MatBuro.ru

©**МатБюро** – Консультации по математике, экономике, праву, другим предметам

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

Выполненный прогноз потребления является надежным ($p = 1 - \alpha = 1 - 0,05 = 0,95$) и точным.

ЧАСТЬ II

Используя фактические значения независимых переменных (x_1 и x_2) и результирующего показателя (y), провести эконометрическое исследование зависимости y от x_1 и x_2 :

1. Выбрать в качестве уравнения взаимосвязи переменных x_1 , x_2 и y линейное регрессионное уравнение вида $y = a_0 + a_1x_1 + a_2x_2 + e$.

2. Найти коэффициенты парной корреляции факторов и построить матрицу парных коэффициентов корреляции. Сделать выводы о связи переменных уравнения регрессии

3. Рассчитать коэффициенты $\hat{a}_0$, $\hat{a}_1$ и $\hat{a}_2$ выбранного уравнения. Для этого составить систему нормальных уравнений и найти ее решение методом определителей. Построить модель прогноза.

4. Вычислить индекс множественной корреляции.

5. Оценить качество построенного уравнения:

а) определить значимость коэффициентов регрессии с помощью t -критерия Стьюдента (при уровне значимости $\alpha=0,05$);

б) с помощью F -критерия Фишера оценить статистическую надежность построенного уравнения (при уровне значимости $\alpha=0,05$);

в) рассчитать частные критерии Фишера и оценить целесообразность включения в построенное уравнение фактора x_1 после фактора x_2 и фактора x_2 после фактора x_1 ;

г) оценить значимость коэффициентов при переменных x_1 и x_2 уравнения через значения частных критериев Фишера. Сравнить полученные результаты с результатами оценки значимости коэффициентов по критерию Стьюдента (п. 5а).

6. Рассчитать средние коэффициенты эластичности и с их помощью оценить степень влияния независимых переменных x_1 и x_2 на зависимую переменную y .

7. Построить частные уравнения регрессии.

8. Сделать все необходимые выводы по результатам выполнения каждого из пунктов задания.

Задача 2

Имеются данные (в у.е.) о среднегодовой стоимости основных фондов (x_1), оборотных средств (x_2) и объеме произведённой продукции (y) ряда предприятий:

Стоимость основных фондов	15,4	2,6	1,6	12,0	3,7	11,5	14,1	5,6	2,4	16,4
Стоимость оборотных средств	10,7	4,3	3,5	5,4	3,7	3,6	4,3	2,7	3,7	5,1
Объем продукции	165	37,5	24,1	115	40,3	105	127	50,8	33,3	145,3

Решение.

В тех случаях, когда необходимо оценить влияние нескольких факторов на исследуемую величину, строится уравнение множественной регрессии.

Если связь является линейной, то уравнение линейной множественной регрессии запишется в виде:

$$Y = a_0 + a_1 x_1 + a_2 x_2 + e$$

Для удобства проведения расчетов поместим результаты промежуточных расчетов в таблицу:

Таблица 4.

№	y	x_1	x_2	yx_1	yx_2	$x_1 x_2$	x_1^2	x_2^2	y^2
1	2	3	4	5	6	7	8	9	10
1	165	15	11	2541	1765,5	164,8	237,16	114,5	27225
2	37,5	2,6	4,3	97,5	161,25	11,18	6,76	18,49	1406,25
3	24,1	1,6	3,5	38,56	84,35	5,6	2,56	12,25	580,81
4	115	12	5,4	1380	621	64,8	144	29,16	13225

5	40,3	3,7	3,7	149,11	149,11	13,69	13,69	13,69	1624,09
6	105	12	3,6	1207,5	378	41,4	132,25	12,96	11025
7	127	14	4,3	1790,7	546,1	60,63	198,81	18,49	16129
8	50,8	5,6	2,7	284,48	137,16	15,12	31,36	7,29	2580,64
9	33,3	2,4	3,7	79,92	123,21	8,88	5,76	13,69	1108,89
10	145	16	5,1	2382,9	741,03	83,64	268,96	26,01	21112,1
Сумма	843,3	85,3	47	9951,7	4706,7	469,7	1041,3	266,5	96016,8
Ср. знач.	84,33	8,53	4,7	995,17	470,67	46,97	104,13	26,65	9601,68

Найдем средние квадратические отклонения признаков:

$$\sigma_y = \sqrt{y^2 - \bar{y}^2} = \sqrt{9601,68 - 84,33^2} = 49,901$$

$$\sigma_{x1} = \sqrt{x_1^2 - \bar{x}_1^2} = \sqrt{104,13 - 8,53^2} = 5,601$$

$$\sigma_{x2} = \sqrt{x_2^2 - \bar{x}_2^2} = \sqrt{26,65 - 4,7^2} = 2,136$$

Рассчитаем парные коэффициенты корреляции

$$r_{yx1} = \frac{\text{cov}(y, x_1)}{\sigma_y \cdot \sigma_{x1}} = \frac{\overline{yx_1} - \bar{y} \cdot \bar{x}_1}{\sigma_y \cdot \sigma_{x1}} = \frac{995,17 - 84,33 \cdot 8,53}{49,901 \cdot 5,601} = 0,9869$$

$$r_{yx2} = \frac{\text{cov}(y, x_2)}{\sigma_y \cdot \sigma_{x2}} = \frac{\overline{yx_2} - \bar{y} \cdot \bar{x}_2}{\sigma_y \cdot \sigma_{x2}} = \frac{470,67 - 84,33 \cdot 4,7}{49,901 \cdot 2,136} = 0,6973$$

$$r_{x1x2} = \frac{\text{cov}(x_1, x_2)}{\sigma_{x1} \cdot \sigma_{x2}} = \frac{\overline{x_1x_2} - \bar{x}_1 \cdot \bar{x}_2}{\sigma_{x1} \cdot \sigma_{x2}} = \frac{46,97 - 8,53 \cdot 4,7}{5,601 \cdot 2,136} = 0,5752$$

Составим матрицу парных коэффициентов корреляции

	y	x ₁	x ₂
y	1	0,9869	0,6973
x ₁	0,9869	1	0,5752

x_2	0,6973	0,5752	1
-------	--------	--------	---

Полученные коэффициенты парной корреляции указывают на очень тесную связь факторов x_1 и x_2 с результатом y .

Т.к. $r_{x_1x_2} = 0.5752 < 0,7$ следовательно, межфакторная зависимость несильная, что хорошо отражается на качестве линейной модели регрессии.

Для оценки параметров уравнения множественной регрессии применяют метод наименьших квадратов (МНК). Для линейных уравнений строится следующая система нормальных уравнений, решение которой позволяет получить оценки параметров регрессии:

$$\begin{cases} na_0 + a_1 \sum x_{1i} + a_2 \sum x_{2i} = \sum y_i \\ a_0 \sum x_{1i} + a_1 \sum x_{1i}^2 + a_2 \sum x_{1i}x_{2i} = \sum y_i x_{1i} \\ a_0 \sum x_{2i} + a_1 \sum x_{1i}x_{2i} + a_2 \sum x_{2i}^2 = \sum y_i x_{2i} \end{cases}$$

Для ее решения может быть применен метод определителей:

$$a_0 = \frac{\Delta a_0}{\Delta}, \quad a_1 = \frac{\Delta a_1}{\Delta}, \quad a_2 = \frac{\Delta a_2}{\Delta}$$

$$\text{где } \Delta = \begin{vmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{vmatrix} - \text{определитель системы;}$$

$\Delta a_0, \Delta a_1, \Delta a_2$ - частные определители; которые получаются путем замены соответствующего столбца матрицы определителя системы данными левой части системы.

Таким образом получаем систему:

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

$$\begin{cases} 10a_0 + 85.3a_1 + 47a_2 = 843.3 \\ 85.3a_0 + 1041.3a_1 + 469.7a_2 = 9951.7 \\ 47a_0 + 469.7a_1 + 266.5a_2 = 4706.7 \end{cases}$$

Находим определители методом разложения по первой строке.

$$\Delta = 10 \cdot (1041,3 \cdot 266,5 - 469,7 \cdot 469,7) - 85,3 \cdot (85,3 \cdot 266,5 - 469,7 \cdot 47) + 47 \cdot (85,3 \cdot 469,7 - 1041,3 \cdot 47) = 95762,2352$$

Заменяем 1-ый столбец матрицы на вектор результата.

$$\Delta_{a_0} = \begin{vmatrix} 843.3 & 85.3 & 47 \\ 9951.69 & 1041.31 & 469.72 \\ 4706.71 & 469.72 & 266.52 \end{vmatrix} = -332966.5911$$

Заменяем 2-ый столбец матрицы на вектор результата.

$$\Delta_{a_1} = \begin{vmatrix} 10 & 843.3 & 47 \\ 85.3 & 9951.69 & 469.72 \\ 47 & 4706.71 & 266.52 \end{vmatrix} = 746959.2442$$

Заменяем 3-ый столбец матрицы на вектор результата.

$$\Delta_{a_2} = \begin{vmatrix} 10 & 85.3 & 843.3 \\ 85.3 & 1041.31 & 9951.69 \\ 47 & 469.72 & 4706.71 \end{vmatrix} = 433411.3899$$

Следовательно

$$a_0 = \frac{-332966.5911}{95762.2352} = -3.477$$

$$a_1 = \frac{746959.2442}{95762.2352} = 7.8001$$

$$a_2 = \frac{433411.3899}{95762.2352} = 4.5259$$

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

Таким образом, уравнение регрессии имеет вид:

$$y = -3.477 + 7.8001x_1 + 4.5259x_2$$

Уравнение показывает, что с увеличением стоимости основных фондов на 1 у.е. объем продукции в среднем увеличивается на 7,8001 у.е., с увеличением стоимости оборотных средств на 1 у.е. – увеличивается в среднем на 4,5259 у.е.

Индекс множественной корреляции можно определить через матрицу парных коэффициентов корреляции:

$$R_{yx_1x_2} = \sqrt{1 - \frac{\Delta_r}{\Delta_{r_{11}}}},$$

где

$$\Delta_r = \begin{vmatrix} 1 & r_{yx_1} & r_{yx_2} \\ r_{yx_1} & 1 & r_{x_1x_2} \\ r_{yx_2} & r_{x_2x_1} & 1 \end{vmatrix} - \text{определитель матрицы парных коэффициентов корреляции;}$$

$$\Delta_{r_{11}} = \begin{vmatrix} 1 & r_{x_1x_2} \\ r_{x_2x_1} & 1 \end{vmatrix} - \text{определитель матрицы межфакторной корреляции.}$$

$$\Delta_r = \begin{vmatrix} 1 & 0.9869 & 0.6973 \\ 0.9869 & 1 & 0.5752 \\ 0.6973 & 0.5752 & 1 \end{vmatrix} = 0.0006$$

$$\Delta_{r_{11}} = \begin{vmatrix} 1 & 0.5752 \\ 0.5752 & 1 \end{vmatrix} = 0.6691$$

Коэффициент множественной корреляции

$$R_{yx_1x_2} = \sqrt{1 - \frac{0.0006}{0.6691}} = \sqrt{0.9991} = 0.9996$$

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

Зависимость y от x_1 и x_2 характеризуется как тесная, в которой 99,96 % вариации выпуска продукции определяются вариацией учтенных в модели факторов: стоимости основных фондов и стоимости оборотных средств. Прочие факторы, не включенные в модель, составляют соответственно 0,4% от общей вариации y .

Оценим статистическую значимость параметров регрессии с помощью t-критерия Стьюдента. Рассчитаем стандартные ошибки коэффициентов регрессии по формулам:

$$m_{a1} = \frac{\sigma_y \sqrt{1 - R_{yx1x2}^2}}{\sigma_{x1} \sqrt{1 - r_{x1x2}^2}} \cdot \frac{1}{\sqrt{n-3}} = \frac{49.9012 \sqrt{1 - 0.9991}}{5.6009 \sqrt{1 - 0.5752^2}} \cdot \frac{1}{\sqrt{10-3}} = 0.1228$$

$$m_{a2} = \frac{\sigma_y \sqrt{1 - R_{yx1x2}^2}}{\sigma_{x2} \sqrt{1 - r_{x1x2}^2}} \cdot \frac{1}{\sqrt{n-3}} = \frac{49.9012 \sqrt{1 - 0.9991}}{2.1359 \sqrt{1 - 0.5752^2}} \cdot \frac{1}{\sqrt{10-3}} = 0.3221$$

Фактические значения t-критерия Стьюдента:

$$t_{a1} = \frac{a_1}{m_{a1}} = \frac{7.8001}{0.1228} = 63.5077, \quad t_{a2} = \frac{a_2}{m_{a2}} = \frac{4.5259}{0.3221} = 14.0524$$

Табличное значение критерия при уровне значимости $\alpha = 0,05$ и числе степеней свободы $k = 7$ составит $t_{\text{табл}} (\alpha = 0,05; k = 7) = 2,365$.

Таким образом, признается статистическая значимость параметра a_1 , т.к. $t_{a1} > t_{\text{табл}}$, и параметра a_2 , т.к. $t_{a2} > t_{\text{табл}}$.

Оценку надежности уравнения регрессии в целом и показателя тесноты связи R_{yx1x2} дает

F -критерий Фишера:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m}.$$

В нашем случае фактическое значение F -критерия Фишера:

$$F_{\text{факт}} = \frac{0.9991}{1 - 0.9991} \cdot \frac{10 - 2 - 1}{2} = 3928.38$$

Табличное значение критерия со степенями свободы $k_1 = 2$ и $k_2 = 7$, $F_{\text{табл}} = 4.74$

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

Получили, что $F_{\text{факт}} = 3928.38 > F_{\text{табл}} = 4.74$, т.е. вероятность случайно получить такое значение F -критерия не превышает допустимый уровень значимости 5%. Следовательно, полученное значение не случайно, оно сформировалось под влиянием существенных факторов, т.е. подтверждается статистическая значимость всего уравнения и показателя тесноты связи R_{yx1x2} .

С помощью частных F -критериев Фишера оценим целесообразность включения в уравнение множественной регрессии фактора X_1 после X_2 и фактора X_2 после X_1 при помощи формул:

$$F_{x1} = \frac{R^2 - r_{yx2}^2}{1 - R^2} (n - m - 1)$$

$$F_{x2} = \frac{R^2 - r_{yx1}^2}{1 - R^2} (n - m - 1)$$

Найдем R_{yx1}^2 и R_{yx2}^2 .

$$R_{yx1}^2 = r_{yx1}^2 = 0.9869^2 = 0.9740;$$

$$R_{yx2}^2 = r_{yx2}^2 = 0.6973^2 = 0.4862.$$

Имеем

$$F_{\text{част},x1} = \frac{0.9991 - 0.4862}{1 - 0.9991} \cdot (10 - 2 - 1) = 4033.232$$

$$F_{\text{част},x2} = \frac{0.9991 - 0.9740}{1 - 0.9991} \cdot (10 - 2 - 1) = 197.469.$$

Получили, что $F_{\text{част},x2} > F_{\text{табл}} = 4,74$. Следовательно, включение в модель фактора X_2 после того, как в модель включен фактор X_1 статистически целесообразно: прирост факторной дисперсии за счет дополнительного признака X_2 оказывается значительным, существенным; фактор X_2 включать в уравнение после фактора X_1 следует.

Поможем с заданиями по эконометрике: https://www.matburo.ru/sub_subject.php?p=eckr

Если рассмотреть вариант включения x_1 после x_2 , то результат расчета частного F -критерия $F_{\text{част},x1} > F_{\text{табл}} = 4,74$, т.е. вероятность его случайного формирования меньше принятого стандарта $\alpha = 0,05$ (5%). Следовательно, значение частного F -критерия для дополнительно включенного фактора x_1 не случайно, является статистически значимым, надежным, достоверным: прирост факторной дисперсии за счет дополнительного фактора x_1 является существенным. Фактор x_1 должен присутствовать в уравнении, в том числе в варианте, когда он дополнительно включается после фактора x_2 .

Значения t -критерия Стьюдента можно найти как корень из соответствующего частного F -критерия Фишера.

$$t_{a1} = \sqrt{F_{\text{част},x1}} = \sqrt{4033.232} = 63.5077$$

$$t_{a2} = \sqrt{F_{\text{част},x2}} = \sqrt{197.4692} = 14.0524.$$

Полученные значения совпадают с результатами оценки значимости коэффициентов по критерию Стьюдента.

На основе линейного уравнения множественной регрессии

$y = -3.477 + 7.8001x_1 + 4.5259x_2$ могут быть найдены частные уравнения регрессии, т.е. уравнения регрессии, которые связывают результативный признак с соответствующим фактором x_i при закреплении остальных факторов на среднем уровне.

При подстановке в эти уравнения средних значений соответствующих факторов они принимают вид парных уравнений линейной регрессии, т.е. имеем

$$y_{x1} = -3.477 + 7.8001x_1 + 4.5259\bar{x}_2 = -3.477 + 7.8001x_1 + 4.5259 \cdot 4.7 = 17.7947 + 7.8001x_1$$

$$y_{x2} = -3.477 + 7.8001\bar{x}_1 + 4.8259x_2 = -3.477 + 7.8001 \cdot 8.53 + 4.8259x_2 = 63.0579 + 4.8259x_2$$

.